Measure Theory

Volume I

V.I. Bogachev

Measure Theory

Volume I

Springer

Vladimir I. Bogachev
Department of Mechanics and Mathematics
Moscow State University
119992 Moscow
Russia

Mathematics Subject Classification (2000): 28-00, 46GXX, 60-XX

Library of Congress Control Number: 2006933997

ISBN-10 3-540-34513-2 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-34513-8 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media
springer.com

Typesetting by the author using a Springer LaTeX macro package
Cover design: deblik, Berlin

Printed on acid-free paper SPIN: 11693215 41/SPi 5 4 3 2 1 0

Preface

This book gives an exposition of the foundations of modern measure theory and offers three levels of presentation: a standard university graduate course, an advanced study containing some complements to the basic course (the material of this level corresponds to a variety of special courses), and, finally, more specialized topics partly covered by more than 850 exercises. The target readership includes graduate students interested in deeper knowledge of measure theory, instructors of courses in measure and integration theory, and researchers in all fields of mathematics. The book may serve as a source for many advanced courses or as a reference.

Volume 1 (Chapters 1–5) is devoted to the classical theory of measure and integral, created chiefly by H. Lebesgue and developed by many other mathematicians, in particular, by E. Borel, G. Vitali, W. Young, F. Riesz, D. Egoroff, N. Lusin, J. Radon, M. Fréchet, H. Hahn, C. Carathéodory, and O. Nikodym, whose results are presented in these chapters. Almost all the results in Chapters 1–5 were already known in the first third of the 20th century, but the methods of presentation, certainly, take into account later developments. The basic material designed for graduate students and oriented towards beginners covers approximately 100 pages in the first five chapters (i.e., less than 1/4 of those chapters) and includes the following sections: §1.1–1.7, §2.1–2.11, §3.2–3.4, §3.9, §4.1, §4.3, and some fragments of §5.1–5.4. It corresponds to a one-semester university course of real analysis (measure and integration theory) taught by the author at the Department of Mechanics and Mathematics at the Lomonosov Moscow University. The curriculum of this course is found at the end of the Bibliographical and Historical Comments. The required background includes only the basics of calculus (convergence of sequences and series, continuity of functions, open and closed sets in the real line, the Riemann integral) and linear algebra. Although knowledge of the Riemann integral is not formally assumed, I am convinced that the Riemann approach should be a starting point of the study of integration; acquaintance with the basics of the Riemann theory enables one to appreciate the depth and beauty of Lebesgue's creation. Some additional notions needed in particular sections are explained in the appropriate places. Naturally, the classical basic material of the first five chapters (without supplements) does not differ much from what is contained in many well-known textbooks on measure and integration or probability theory, e.g., Bauer [**70**], Halmos [**404**], Kolmogorov, Fomin [**536**], Loève [**617**], Natanson [**707**], Neveu [**713**], Parthasarathy [**739**], Royden [**829**], Shiryaev [**868**], and other books. An important feature of our exposition is that the listed sections contain only minimal material covered in real lectures. In particular, less attention than usual is given to measures on semirings etc. In general, the technical set-theoretic ingredients are considerably shortened. However, the corresponding material is not completely excluded: it is just transferred to supplements and exercises. In this way, one can substantially ease the first acquaintance with the subject when the abundance of definitions and set-theoretical constructions often make obstacles for understanding the principal ideas. Other sections of the main body of the book, supplements and exercises contain many things that are very useful in applications but seldom included in textbooks. There are two reasons why the standard course is included in full detail (rather than just mentioned in prerequisites): it makes the book completely self-contained and available to a much broader audience, in addition, many topics in the advanced material continue our discussion started in the basic course; it would be unnatural to give a continuation of a discussion without its beginning and origins. It should be noted that brevity of exposition has not been my priority; moreover, due to the described structure of the book, certain results are first presented in more special cases and only later are given in more general form. For example, our discussion of measures and integrals starts from finite measures, since the consideration of infinite values does not require new ideas, but for the beginner may overshadow the essence by rather artificial troubles with infinities. The organization of the book does not suggest reading from cover to cover; in particular, almost all sections in the supplements are independent of each other and are directly linked only to specific sections of the main part. A detailed table of contents is given. Here are brief comments on the structure of chapters.

In Chapter 1, the principal objects are countably additive measures on algebras and σ-algebras, and the main theorems are concerned with constructions and extensions of measures.

Chapter 2 is devoted to the construction of the Lebesgue integral, for which measurable functions are introduced first. The main theorems in this chapter are concerned with passage to the limit under the integral sign. The Lebesgue integral — one of the basic objects in this book — is not the most general type of integral. Apparently, its role in modern mathematics is explained by two factors: it possesses a sufficient and reasonable generality combined with aesthetic attractiveness.

In Chapter 3, we consider the most important operations on measures and functions: the Hahn–Jordan decomposition of signed measures, product measures, multiplication of measures by functions, convolutions of functions and measures, transformations of measures and change of variables. We discuss in detail finite and infinite products of measures. Fundamental theorems due to Radon&Nikodym and Fubini are presented.

Chapter 4 is devoted to spaces of integrable functions and spaces of measures. We discuss the geometric properties of the space L^p, study the uniform integrability, and prove several important theorems on convergence and boundedness of sequences of measures. Considerable attention is given to weak convergence and the weak topology in L^1. Finally, the structure properties of spaces of functions and measures are discussed.

In Chapter 5, we investigate connections between integration and differentiation and prove the classical theorems on the differentiability of functions of bounded variation and absolutely continuous functions and integration by parts. Covering theorems and the maximal function are discussed. The Henstock–Kurzweil integral is introduced and briefly studied.

Whereas the first volume presents the ideas that go back mainly to Lebesgue, the second volume (Chapters 6–10) is to a large extent the result of the development of ideas generated in 1930–1960 by a number of mathematicians, among which primarily one should mention A.N. Kolmogorov, J. von Neumann, and A.D. Alexandroff; other chief contributors are mentioned in the comments. The central subjects in Volume 2 are: transformations of measures, conditional measures, and weak convergence of measures. These three themes are closely interwoven and form the heart of modern measure theory. Typical measure spaces here are infinite dimensional: e.g., it is often convenient to consider a measure on the interval as a measure on the space $\{0,1\}^\infty$ of all sequences of zeros and ones. The point is that in spite of the fact that any reasonable measure space is isomorphic to an interval, a significant role is played by diverse additional structures on measure spaces: algebraic, topological, and differential. This is partly explained by the fact that many problems of modern measure theory grew under the influence of probability theory, the theory of dynamical systems, information theory, the theory of representations of groups, nonlinear analysis, and mathematical physics. All these fields brought into measure theory not only problems, methods, and terminology, but also inherent ways of thinking. Note also that the most fruitful directions in measure theory now border with other branches of mathematics.

Unlike the first volume, a considerable portion of material in Chapters 6–10 has not been presented in such detail in textbooks. Chapters 6–10 require also a deeper background. In addition to knowledge of the basic course, it is necessary to be familiar with the standard university course of functional analysis including elements of general topology (e.g., the textbook by Kolmogorov and Fomin covers the prerequisites). In some sections it is desirable to be familiar with fundamentals of probability theory (for this purpose, a concise book, Lamperti [**566**], can be recommended). In the second volume many themes touched on in the first volume find their natural development (for example, transformations of measures, convergence of measures, Souslin sets, connections between measure and topology).

Chapter 6 plays an important technical role: here we study various properties of Borel and Souslin sets in topological spaces and Borel mappings of

Souslin sets, in particular, several measurable selection and implicit function theorems are proved here. The birth of this direction is due to a great extent to the works of N. Lusin and M. Souslin. The exposition in this chapter has a clear set-theoretic and topological character with almost no measures. The principal results are very elegant, but are difficult in parts in the technical sense, and I decided not to hide these difficulties in exercises. However, this chapter can be viewed as a compendium of results to which one should resort in case of need in the subsequent chapters.

In Chapter 7, we discuss measures on topological spaces, their regularity properties, and extensions of measures, and examine the connections between measures and the associated functionals on function spaces. The branch of measure theory discussed here grew from the classical works of J. Radon and A.D. Alexandroff, and was strongly influenced (and still is) by general topology and descriptive set theory. The central object of the chapter is Radon measures. We also study in detail perfect and τ-additive measures. A separate section is devoted to the Daniell–Stone method. This method could have been explained already in Chapter 2, but it is more natural to place it close to the Riesz representation theorem in the topological framework. There is also a brief discussion of measures on locally convex spaces and their characteristic functionals (Fourier transforms).

In Chapter 8, directly linked only to Chapter 7, the theory of weak convergence of measures is presented. We prove several fundamental results due to A.D. Alexandroff, Yu.V. Prohorov and A.V. Skorohod, study the weak topology on spaces of measures and consider weak compactness. The topological properties of spaces of measures on topological spaces equipped with the weak topology are discussed. The concept of weak convergence of measures plays an important role in many applications, including stochastic analysis, mathematical statistics, and mathematical physics. Among many complementary results in this chapter one can mention a thorough discussion of convergence of measures on open sets and a proof of the Fichtenholz–Dieudonné–Grothendieck theorem.

Chapter 9 is devoted to transformations of measures. We discuss the properties of images of measures under mappings, the existence of preimages, various types of isomorphisms of measure spaces (for example, point, metric, topological), the absolute continuity of transformed measures, in particular, Lusin's (N)-property, transformations of measures by flows generated by vector fields, Haar measures on locally compact groups, the existence of invariant measures of transformations, and many other questions important for applications. The "nonlinear measure theory" discussed here originated in the 1930s in the works of G.D. Birkhoff, J. von Neumann, N.N. Bogolubov, N.M. Krylov, E. Hopf and other researchers in the theory of dynamical systems, and was also considerably influenced by other fields such as the integration on topological groups developed by A. Haar, A. Weil, and others. A separate section is devoted to the theory of Lebesgue spaces elaborated by V. Rohlin (such spaces are called here Lebesgue–Rohlin spaces).

Chapter 10 is close to Chapter 9 in its spirit. The principal ideas of this chapter go back to the works of A.N. Kolmogorov, J. von Neumann, J. Doob, and P. Lévy. It is concerned with conditional measures — the object that plays an exceptional role in measure theory as well as in numerous applications. We describe in detail connections between conditional measures and conditional expectations, prove the main theorems on convergence of conditional expectations, establish the existence of conditional measures under broad assumptions and clarify their relation to liftings. In addition, a concise introduction to the theory of martingales is given with views towards applications in measure theory. A separate section is devoted to ergodic theory — a fruitful field at the border of measure theory, probability theory, and mathematical physics. Finally, in this chapter we continue our study of Lebesgue–Rohlin spaces, and in particular, discuss measurable partitions.

Extensive complementary material is presented in the final sections of all chapters, where there are also a lot of exercises supplied with complete solutions or hints and references. Some exercises are merely theorems from the cited sources printed in a smaller font and are placed there to save space (so that the absence of hints means that I have no solutions different from the ones found in the cited works). The symbol $^\circ$ marks exercises recommendable for graduate courses or self-study. Note also that many solutions have been borrowed from the cited works, but sometimes solutions simpler than the original ones are presented (this fact, however, is not indicated). It should be emphasized that many exercises given without references are either taken from the textbooks listed in the bibliographical comments or belong to the mathematical folklore. In such exercises, I omitted the sources (which appear in hints, though), since they are mostly secondary. It is possible that some exercises are new, but this is never claimed for the obvious reason that a seemingly new assertion could have been read in one of hundreds papers from the list of references or even heard from colleagues and later recalled.

The book contains an extensive bibliography and the bibliographical and historical comments. The comments are made separately on each volume, the bibliography in Volume 1 contains the works cited only in that volume, and Volume 2 contains the cumulative bibliography, where the works cited only in Volume 1 are marked with an asterisk. For each item in the list of references we indicate all pages where it is cited. The comments, in addition to remarks of a historical or bibliographical character, give references to works on many special aspects of measure theory, which could not be covered in a book of this size, but the information about which may be useful for the reader. A detailed subject index completes the book (Volume 1 contains only the index for that volume, and Volume 2 contains the cumulative index).

For all assertions and formulas we use the triple enumeration: the chapter number, section number, and assertion number (all assertions are numbered independently of their type within each section); numbers of formulas are given in brackets.

This book is intended as a complement to the existing large literature of advanced graduate-text type and provides the reader with a lot of material from many parts of measure theory which does not belong to the standard course but is necessary in order to read research literature in many areas. Modern measure theory is so vast that it cannot be adequately presented in one book. Moreover, even if one attempts to cover all the directions in a universal treatise, possibly in many volumes, due depth of presentation will not be achieved because of the excessive amount of required information from other fields. It appears that for an in-depth study not so voluminous expositions of specialized directions are more suitable. Such expositions already exist in a several directions (for example, the geometric measure theory, Hausdorff measures, probability distributions on Banach spaces, measures on groups, ergodic theory, Gaussian measures). Here a discussion of such directions is reduced to a minimum, in many cases just to mentioning their existence.

This book grew from my lectures at the Lomonosov Moscow University, and many related problems have been discussed in lectures, seminar talks and conversations with colleagues at many other universities and mathematical institutes in Moscow, St.-Petersburg, Kiev, Berlin, Bielefeld, Bonn, Oberwolfach, Paris, Strasburg, Cambridge, Warwick, Rome, Pisa, Vienna, Stockholm, Copenhagen, Zürich, Barcelona, Lisbon, Athens, Edmonton, Berkeley, Boston, Minneapolis, Santiago, Haifa, Kyoto, Beijing, Sydney, and many other places. Opportunities to work in the libraries of these institutions have been especially valuable. Through the years of work on this book I received from many individuals the considerable help in the form of remarks, corrections, additional references, historical comments etc. Not being able to mention here all those to whom I owe gratitude, I particularly thank H. Airault, E.A. Alekhno, E. Behrends, P.A. Borodin, G. Da Prato, D. Elworthy, V.V. Fedorchuk, M.I. Gordin, M.M. Gordina, V.P. Havin, N.V. Krylov, P. Lescot, G. Letta, A.A. Lodkin, E. Mayer-Wolf, P. Malliavin, P.-A. Meyer, L. Mejlbro, E. Priola, V.I. Ponomarev, Yu.V. Prohorov, M. Röckner, V.V. Sazonov, B. Schmuland, A.N. Shiryaev, A.V. Skorohod, O.G. Smolyanov, A.M. Stepin, V.N. Sudakov, V.I. Tarieladze, S.A. Telyakovskii, A.N. Tikhomirov, F. Topsøe, V.V. Ulyanov, H. von Weizsäcker, and M. Zakai. The character of presentation was considerably influenced by discussions with my colleagues at the chair of theory of functions and functional analysis at the Department of Mechanics and Mathematics of the Lomonosov Moscow University headed by the member of the Russian Academy of Science P.L. Ulyanov. For checking several preliminary versions of the book, numerous corrections, improvements and other related help I am very grateful to A.V. Kolesnikov, E.P. Krugova, K.V. Medvedev, O.V. Pugachev, T.S. Rybnikova, N.A. Tolmachev, R.A. Troupianskii, Yu.A. Zhereb'ev, and V.S. Zhuravlev. The book took its final form after Z. Lipecki read the manuscript and sent his corrections, comments, and certain materials that were not available to me. I thank J. Boys for careful copyediting and the editorial staff at Springer-Verlag for cooperation.

Moscow, August 2006 Vladimir Bogachev

Contents

Contents of Volume 2

CHAPTER 1

Constructions and extensions of measures

I compiled these lectures not assuming from the reader any knowledge other than is found in the under-graduate programme of all departments; I can even say that not assuming anything except for acquaintance with the definition and the most elementary properties of integrals of continuous functions. But even if there is no necessity to know much before reading these lectures, it is yet necessary to have some practice of thinking in such matters.

H. Lebesgue. Intégration et la récherche des fonctions primitives.

1.1. Measurement of length: introductory remarks

Many problems discussed in this book grew from the following question: which sets have length? This question clear at the first glance leads to two other questions: what is a "set" and what is a "number" (since one speaks of a qualitative measure of length)? We suppose throughout that some answers to these questions have been given and do not raise them further, although even the first constructions of measure theory lead to situations requiring greater certainty. We assume that the reader is familiar with the standard facts about real numbers, which are given in textbooks of calculus, and for "set theory" we take the basic assumptions of the "naive set theory" also presented in textbooks of calculus; sometimes the axiom of choice is employed. In the last section the reader will find a brief discussion of major set-theoretic problems related to measure theory. We use throughout the following set-theoretic relations and operations (in their usual sense): $A \subset B$ (the inclusion of a set A to a set B), $a \in A$ (the inclusion of an element a in a set A), $A \cup B$ (the union of sets A and B), $A \cap B$ (the intersection of sets A and B), $A\backslash B$ (the complement of B in A, i.e., the set of all points from A not belonging to B). Finally, let $A \triangle B$ denote the symmetric difference of two sets A and B, i.e., $A \triangle B = (A \cup B)\backslash(A \cap B)$. We write $A_n \uparrow A$ if $A_n \subset A_{n+1}$ and $A = \bigcup_{n=1}^\infty A_n$; we write $A_n \downarrow A$ if $A_{n+1} \subset A_n$ and $A = \bigcap_{n=1}^\infty A_n$.

The restriction of a function f to a set A is denoted by $f|_A$.

The standard symbols $\mathbb{N} = \{1, 2, \ldots\}$, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}^n$ denote, respectively, the sets of all natural, integer, rational numbers, and the n-dimensional Euclidean space. The term "positive" means "strictly positive" with the exception of some special situations with the established terminology (e.g., the positive part of a function may be zero); similarly with "negative".

The following facts about the set $\mathbb{R}^1$ of real numbers are assumed to be known.

1) The sets $U \subset \mathbb{R}^1$ such that every point x from U belongs to U with some interval of the form $(x-\varepsilon, x+\varepsilon)$, where $\varepsilon > 0$, are called open; every open set is the union of a finite or countable collection of pairwise disjoint intervals or rays. The empty set is open by definition.

2) The closed sets are the complements to open sets; a set A is closed precisely when it contains all its limit points. We recall that a is called a limit point for A if every interval centered at a contains a point $b \neq a$ from A. It is clear that any unions and finite intersections of open sets are open. Thus, the real line is a topological space (more detailed information about topological spaces is given in Chapter 6).

It is clear that any intersections and finite unions of closed sets are closed. An important property of $\mathbb{R}^1$ is that the intersection of any decreasing sequence of nonempty bounded closed sets is nonempty. Depending on the way in which the real numbers have been introduced, this claim is either an axiom or is derived from other axioms. The principal concepts related to convergence of sequences and series are assumed to be known.

Let us now consider the problem of measurement of length. Let us aim at defining the length λ of subsets of the interval $I = [0, 1]$. For an interval J of the form (a, b), $[a, b)$, $[a, b]$ or $(a, b]$, we set $\lambda(J) = |b-a|$. For a finite union of disjoint intervals $J_1,\ldots,J_n$, we set $\lambda\bigl(\bigcup_{i=1}^n J_i\bigr) = \sum_{i=1}^n \lambda(J_i)$. The sets of the indicated form are called *elementary*. We now have to make a non-trivial step and extend measure to non-elementary sets. A natural way of doing this, which goes back to antiquity, consists of approximating non-elementary sets by elementary ones. But how to approximate? The construction that leads to the so-called *Jordan measure* (which should be more precisely called the *Peano–Jordan measure* following the works Peano [**741**], Jordan [**472**]), is this: a set $A \subset I$ is Jordan measurable if for any $\varepsilon > 0$, there exist elementary sets A_ε and B_ε such that $A_\varepsilon \subset A \subset B_\varepsilon$ and $\lambda(B_\varepsilon \backslash A_\varepsilon) < \varepsilon$. It is clear that when $\varepsilon \to 0$, the lengths of A_ε and B_ε have a common limit, which one takes for $\lambda(A)$. Are all the sets assigned lengths after this procedure? No, not at all. For example, the set $\mathbb{Q} \cap I$ of rational numbers in the interval is not Jordan measurable. Indeed, it contains no elementary set of positive measure. On the other hand, any elementary set containing $\mathbb{Q} \cap I$ has measure 1. The question arises naturally about extensions of λ to larger domains. It is desirable to preserve the nice properties of length, which it possesses on the class of Jordan measurable sets. The most important of these properties are the additivity (i.e., $\lambda(A \cup B) = \lambda(A) + \lambda(B)$ for any disjoint sets A and B in the domain) and the invariance with respect to translations. The first property is even fulfilled in the following stronger form of countable additivity: if disjoint sets A_n together with their union $A = \bigcup_{n=1}^\infty A_n$ are Jordan measurable, then $\lambda(A) = \sum_{n=1}^\infty \lambda(A_n)$. As we shall see later, this problem admits solutions. The most important of them suggested by Lebesgue a century ago and leading to Lebesgue measurability consists of changing the way of approximating by elementary sets. Namely, by analogy with the ancient construction one introduces the outer measure λ^* for *every* set $A \subset I$ as the infimum of sums of measures of elementary sets forming countable covers of A. Then a set A is called Lebesgue measurable if the equality $\lambda^*(A) + \lambda^*(I\backslash A) = \lambda(I)$ holds, which can also be expressed in the form of the equality $\lambda^*(A) = \lambda_*(A)$, where the inner measure λ_* is defined *not* by means of inscribed sets as in the case of the Jordan measure, but by the equality $\lambda_*(A) = \lambda(I) - \lambda^*(I\backslash A)$. An equivalent description of the Lebesgue measurability in terms of approximations by elementary sets is this: for any $\varepsilon > 0$ there exists an elementary set A_ε such that $\lambda^*(A \triangle A_\varepsilon) < \varepsilon$. Now, unlike the Jordan measure, no inclusion of sets is required, i.e., "skew approximations" are admissible. This minor nuance leads to a substantial enlargement of the class of measurable sets. The enlargement is so great that the question of the existence of sets to which no measure is assigned becomes dependent on accepting or not accepting certain special set-theoretic axioms. We shall soon verify that the collection of Lebesgue measurable sets is closed with respect to countable unions, countable intersections, and complements. In addition, if we define the measure of a set A as the limit of measures of elementary sets approximating it in the above sense, then the extended measure turns out to be countably additive. All these claims will be derived from more general results. The role of the countable additivity is obvious from the very beginning: if one approximates a disc by unions of rectangles or triangles, then countable unions arise with necessity.

It follows from what has been said above that in the discussion of measures the key role is played by issues related to domains of definition and extensions. So the next section is devoted to principal classes of sets connected with domains of measures. It turns out in this discussion that the specifics of length on subsets of the real line play no role and it is reasonable from the very beginning to speak of measures of an arbitrary nature. Moreover, this point of view becomes necessary for considering measures on general spaces, e.g., manifolds or functional spaces, which is very important for many branches of mathematics and theoretical physics.

1.2. Algebras and σ-algebras

One of the principal concepts of measure theory is an algebra of sets.

1.2.1. Definition. *An algebra of sets $\mathcal{A}$ is a class of subsets of some fixed set X (called the space) such that*

(i) *X and the empty set belong to $\mathcal{A}$*;

(ii) *if A, $B \in \mathcal{A}$, then $A \cap B \in \mathcal{A}$, $A \cup B \in \mathcal{A}$, $A\backslash B \in \mathcal{A}$.*

In place of the condition $A\backslash B \in \mathcal{A}$ one could only require that $X\backslash B \in \mathcal{A}$ whenever $B \in \mathcal{A}$, since $A\backslash B = A \cap (X\backslash B)$ and $A \cup B = X\backslash\bigl((X\backslash A) \cap (X\backslash B)\bigr)$. It is sufficient as well to require in (ii) only that $A\backslash B \in \mathcal{A}$ for all $A, B \in \mathcal{A}$, since $A \cap B = A\backslash(A\backslash B)$.

Sometimes in the definition of an algebra the inclusion $X \in \mathcal{A}$ is replaced by the following wider assumption: there exists a set $E \in \mathcal{A}$ called the unit of the algebra such that $A \cap E = A$ for all $A \in \mathcal{A}$. It is clear that replacing X by E we arrive at our definition on a smaller space. It should be noted that not all of the results below extend to this wider concept.

1.2.2. Definition. *An algebra of sets $\mathcal{A}$ is called a σ-algebra if for any sequence of sets A_n in $\mathcal{A}$ one has $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.*

1.2.3. Definition. *A pair $(X, \mathcal{A})$ consisting of a set X and a σ-algebra $\mathcal{A}$ of its subsets is called a measurable space.*

The basic set (space) on which a σ-algebra or measure are given is most often denoted in this book by X; other frequent symbols are E, M, S (from "ensemble", "Menge", "set"), and Ω, a generally accepted symbol in probability theory. For denoting a σ-algebra it is traditional to use script Latin capitals (e.g., $\mathcal{A}$, $\mathcal{B}$, $\mathcal{E}$, $\mathcal{F}$, $\mathcal{L}$, $\mathcal{M}$, $\mathcal{S}$), Gothic capitals $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{F}$, $\mathfrak{L}$, $\mathfrak{M}$, $\mathfrak{S}$ (i.e., A, B, F, L, M and S) and Greek letters (e.g., $\varSigma$, $\varLambda$, $\varGamma$, $\varXi$), although when necessary other symbols are used as well.

In the subsequent remarks and exercises some other classes of sets are mentioned such as semialgebras, rings, semirings, σ-rings, etc. These classes slightly differ in the operations they admit. It is clear that in the definition of a σ-algebra in place of stability with respect to countable unions one could require stability with respect to countable intersections. Indeed, by the formula $\bigcup_{n=1}^{\infty} A_n = X \backslash \bigcap_{n=1}^{\infty} (X \backslash A_n)$ and the stability of any algebra with respect to complementation it is seen that both properties are equivalent.

1.2.4. Example. The collection of finite unions of all intervals of the form $[a, b]$, $[a, b)$, $(a, b]$, (a, b) in the interval $[0, 1]$ is an algebra, but not a σ-algebra.

Clearly, the collection 2^X of all subsets of a fixed set X is a σ-algebra. The smallest σ-algebra is $(X, \varnothing)$. Any other σ-algebra of subsets of X is contained between these two trivial examples.

1.2.5. Definition. *Let $\mathcal{F}$ be a family of subsets of a space X. The smallest σ-algebra of subsets of X containing $\mathcal{F}$ is called the σ-algebra generated by $\mathcal{F}$ and is denoted by the symbol $\sigma(\mathcal{F})$. The algebra generated by $\mathcal{F}$ is defined as the smallest algebra containing $\mathcal{F}$.*

The smallest σ-algebra and algebra mentioned in the definition exist indeed.

1.2.6. Proposition. *Let X be a set. For any family $\mathcal{F}$ of subsets of X there exists a unique σ-algebra generated by $\mathcal{F}$. In addition, there exists a unique algebra generated by $\mathcal{F}$.*

PROOF. Set $\sigma(\mathcal{F}) = \bigcap_{\mathcal{F} \subset \mathcal{A}} \mathcal{A}$, where the intersection is taken over all σ-algebras of subsets of the space X containing all sets from $\mathcal{F}$. Such σ-algebras exist: for example, 2^X; their intersection by definition is the collection of all sets that belong to each of such σ-algebras. By construction, $\mathcal{F} \subset \sigma(\mathcal{F})$. If we are given a sequence of sets $A_n \in \sigma(\mathcal{F})$, then their intersection, union and complements belong to any σ-algebra $\mathcal{A}$ containing $\mathcal{F}$, hence belong to $\sigma(\mathcal{F})$, i.e., $\sigma(\mathcal{F})$ is a σ-algebra. The uniqueness is obvious from the fact that the existence of a σ-algebra $\mathcal{B}$ containing $\mathcal{F}$ but not containing $\sigma(\mathcal{F})$ contradicts the definition of $\sigma(\mathcal{F})$, since $\mathcal{B}\cap\sigma(\mathcal{F})$ contains $\mathcal{F}$ and is a σ-algebra. The case of an algebra is similar. □

Note that it follows from the definition that the class of sets formed by the complements of sets in $\mathcal{F}$ generates the same σ-algebra as $\mathcal{F}$. It is also clear that a countable class may generate an uncountable σ-algebra. For example, the intervals with rational endpoints generate the σ-algebra containing all single-point sets.

The algebra generated by a family of sets $\mathcal{F}$ can be easily described explicitly. To this end, let us add to $\mathcal{F}$ the empty set and denote by $\mathcal{F}_1$ the collection of all sets of this enlarged collection together with their complements. Then we denote by $\mathcal{F}_2$ the class of all finite intersections of sets in $\mathcal{F}_1$. The class $\mathcal{F}_3$ of all finite unions of sets in $\mathcal{F}_2$ is the algebra generated by $\mathcal{F}$. Indeed, it is clear that $\mathcal{F}\subset\mathcal{F}_1\subset\mathcal{F}_2\subset\mathcal{F}_3$ and that $\varnothing\in\mathcal{F}_3$. The class $\mathcal{F}_3$ admits any finite intersections, since if $A=\bigcup_{i=1}^n A_i$, $B=\bigcup_{j=1}^k B_j$, where $A_i, B_j\in\mathcal{F}_2$, then we have $A\cap B=\bigcup_{i\le n, j\le k} A_i\cap B_j$ and $A_i\cap B_j\in\mathcal{F}_2$. In addition, $\mathcal{F}_3$ is stable under complements. Indeed, if $E=E_1\cup\cdots\cup E_n$, where $E_i\in\mathcal{F}_2$, then $X\backslash E=\bigcap_{i=1}^n(X\backslash E_i)$. Since $E_i=E_{i,1}\cap\cdots\cap E_{i,k_i}$, where $E_{i,j}\in\mathcal{F}_1$, one has $X\backslash E_i=\bigcup_{j=1}^{k_i}(X\backslash E_{i,j})$, where $D_{i,j}:=X\backslash E_{i,j}\in\mathcal{F}_1$. Hence $X\backslash E=\bigcap_{i=1}^n\bigcup_{j=1}^{k_i}D_{i,j}$, which belongs to $\mathcal{F}_3$ by the stability of $\mathcal{F}_3$ with respect to finite unions and intersections. On the other hand, it is clear that $\mathcal{F}_3$ belongs to the algebra generated by $\mathcal{F}$.

One should not attempt to imagine the elements of the σ-algebra generated by the class $\mathcal{F}$ in a constructive form by means of countable unions, intersections or complements of the elements in $\mathcal{F}$. The point is that the above-mentioned operations can be repeated in an unlimited number of steps in any order. For example, one can form the class $\mathcal{F}_\sigma$ of countable unions of closed sets in the interval, then the class $\mathcal{F}_{\sigma\delta}$ of countable intersections of sets in $\mathcal{F}_\sigma$, and continue this process inductively. One will be obtaining new classes all the time, but even their union does not exhaust the σ-algebra generated by the closed sets (the proof of this fact is not trivial; see Exercises 6.10.30, 6.10.31, 6.10.32 in Chapter 6). In §1.10 we study the so-called A-operation, which gives all sets in the σ-algebra generated by intervals, but produces also other sets. Let us give an example where one can explicitly describe the σ-algebra generated by a class of sets.

1.2.7. Example. Let $\mathcal{A}_0$ be a σ-algebra of subsets in a space X. Suppose that a set $S\subset X$ does not belong to $\mathcal{A}_0$. Then the σ-algebra $\sigma\bigl(\mathcal{A}_0\cup\{S\}\bigr)$, generated by $\mathcal{A}_0$ and the set S coincides with the collection of all sets of the form

$$E=(A\cap S)\cup\bigl(B\cap(X\backslash S)\bigr),\quad \text{where } A,B\in\mathcal{A}_0. \tag{1.2.1}$$

PROOF. All sets of the form (1.2.1) belong to the σ-algebra $\sigma\bigl(\mathcal{A}_0\cup\{S\}\bigr)$. On the other hand, the sets of the indicated type form a σ-algebra. Indeed,

$$X\backslash E=\bigl((X\backslash A)\cap S\bigr)\cup\bigl((X\backslash B)\cap(X\backslash S)\bigr),$$

since x does not belong to E precisely when either x belongs to S but not to A, or x belongs neither to S, nor to B. In addition, if the sets E_n are represented in the form (1.2.1) with some $A_n, B_n\in\mathcal{A}_0$, then $\bigcap_{n=1}^\infty E_n$ and $\bigcup_{n=1}^\infty E_n$ also have the form (1.2.1). For example, $\bigcap_{n=1}^\infty E_n$ has the form (1.2.1) with $A=\bigcap_{n=1}^\infty A_n$ and $B=\bigcap_{n=1}^\infty B_n$. Finally, all sets in $\mathcal{A}_0$ are obtained in the form (1.2.1) with $A=B$, and for obtaining S we take $A=X$ and $B=\varnothing$. $\square$

In considerations involving σ-algebras the following simple properties of the set-theoretic operations are often useful.

1.2.8. Lemma. *Let $(A_\alpha)_{\alpha\in\Lambda}$ be a family of subsets of a set X and let $f\colon E\to X$ be an arbitrary mapping of a set E to X. Then*

$$X\backslash\bigcup_{\alpha\in\Lambda}A_\alpha=\bigcap_{\alpha\in\Lambda}(X\backslash A_\alpha),\quad X\backslash\bigcap_{\alpha\in\Lambda}A_\alpha=\bigcup_{\alpha\in\Lambda}(X\backslash A_\alpha),\tag{1.2.2}$$

$$f^{-1}\Bigl(\bigcup_{\alpha\in\Lambda}A_\alpha\Bigr)=\bigcup_{\alpha\in\Lambda}f^{-1}(A_\alpha),\quad f^{-1}\Bigl(\bigcap_{\alpha\in\Lambda}A_\alpha\Bigr)=\bigcap_{\alpha\in\Lambda}f^{-1}(A_\alpha).\tag{1.2.3}$$

PROOF. Let $x\in X\backslash\bigcup_{\alpha\in\Lambda}A_\alpha$, i.e., $x\notin A_\alpha$ for all $\alpha\in\Lambda$. The latter is equivalent to the inclusion $x\in\bigcap_{\alpha\in\Lambda}(X\backslash A_\alpha)$. Other relationships are proved in a similar manner. $\square$

1.2.9. Corollary. *Let $\mathcal{A}$ be a σ-algebra of subsets of a set X and f an arbitrary mapping from a set E to X. Then the class $f^{-1}(\mathcal{A})$ of all sets of the form $f^{-1}(A)$, where $A\in\mathcal{A}$, is a σ-algebra in E.*

In addition, for an arbitrary σ-algebra $\mathcal{B}$ of subsets of E, the class of sets $\{A\subset X\colon f^{-1}(A)\in\mathcal{B}\}$ is a σ-algebra. Furthermore, for any class of sets $\mathcal{F}$ in X, one has $\sigma\bigl(f^{-1}(\mathcal{F})\bigr)=f^{-1}\bigl(\sigma(\mathcal{F})\bigr)$.

PROOF. The first two assertions are clear from the lemma. Since the class $f^{-1}\bigl(\sigma(\mathcal{F})\bigr)$ is a σ-algebra by the first assertion, we obtain the inclusion $\sigma\bigl(f^{-1}(\mathcal{F})\bigr)\subset f^{-1}\bigl(\sigma(\mathcal{F})\bigr)$. Finally, by the second assertion, we have $f^{-1}\bigl(\sigma(\mathcal{F})\bigr)\subset\sigma\bigl(f^{-1}(\mathcal{F})\bigr)$ because $f^{-1}(\mathcal{F})\subset\sigma\bigl(f^{-1}(\mathcal{F})\bigr)$. $\square$

Simple examples show that the class $f(\mathcal{B})$ of all sets of the form $f(B)$, where $B\in\mathcal{B}$, is not always an algebra.

1.2.10. Definition. *The Borel σ-algebra of $\mathbb{R}^n$ is the σ-algebra $\mathcal{B}(\mathbb{R}^n)$ generated by all open sets. The sets in $\mathcal{B}(\mathbb{R}^n)$ are called Borel sets. For any set $E\subset\mathbb{R}^n$, let $\mathcal{B}(E)$ denote the class of all sets of the form $E\cap B$, where $B\in\mathcal{B}(\mathbb{R}^n)$.*

The class $\mathcal{B}(E)$ can also be defined as the σ-algebra generated by the intersections of E with open sets in $\mathbb{R}^n$. This is clear from the following: if the latter σ-algebra is denoted by $\mathcal{E}$, then the family of all sets $B \in \mathcal{B}(\mathbb{R}^n)$ such that $B \cap E \in \mathcal{E}$ is a σ-algebra containing all open sets, i.e., it coincides with $\mathcal{B}(\mathbb{R}^n)$. The sets in $\mathcal{B}(E)$ are called Borel sets of the space E and $\mathcal{B}(E)$ is called the Borel σ-algebra of the space E. One should keep in mind that such sets may not be Borel in $\mathbb{R}^n$ unless, of course, E itself is Borel in $\mathbb{R}^n$. For example, one always has $E \in \mathcal{B}(E)$, since $E \cap \mathbb{R}^n = E$.

It is clear that $\mathcal{B}(\mathbb{R}^n)$ is also generated by the class of all closed sets.

1.2.11. Lemma. *The Borel σ-algebra of the real line is generated by any of the following classes of sets*:

(i) *the collection of all intervals*;

(ii) *the collection of all intervals with rational endpoints*;

(iii) *the collection of all rays of the form $(-\infty, c)$, where c is rational*;

(iv) *the collection of all rays of the form $(-\infty, c]$, where c is rational*;

(v) *the collection of rays of the form $(c, +\infty)$, where c rational*;

(vi) *the collection of all rays of the form $[c, +\infty)$, where c is rational. Finally, the same is true if in place of rational numbers one takes points of any everywhere dense set.*

Proof. It is clear that all the sets indicated above are Borel, since they are either open or closed. Therefore, the σ-algebras generated by the corresponding families are contained in $\mathcal{B}(\mathbb{R}^1)$. Since every open set on the real line is the union of an at most countable collection of intervals, it suffices to show that any interval (a, b) is contained in the σ-algebras corresponding to the classes (i)–(vi). This follows from the fact that (a, b) is the union of intervals of the form (a_n, b_n), where a_n and b_n are rational, and also is the union of intervals of the form $[a_n, b_n)$ with rational endpoints, whereas such intervals belong to the σ-algebra generated by the rays $(-\infty, c)$, since they can be written as differences of rays. In a similar manner, the differences of the rays of the form (c, ∞) give the intervals $(a_n, b_n]$, from which by means of unions one constructs the intervals (a, b). $\square$

It is clear from the proof that the Borel σ-algebra is generated by the closed intervals with rational endpoints. It is seen from this, by the way, that disjoint classes of sets may generate one and the same σ-algebra.

1.2.12. Example. The collection of all single-point sets in a space X generates the σ-algebra consisting of all sets that are either at most countable or have at most countable complements. In addition, this σ-algebra is strictly smaller than the Borel one if $X = \mathbb{R}^1$.

Proof. Denote by $\mathcal{A}$ the family of all sets $A \subset X$ such that either A is at most countable or $X \backslash A$ is at most countable. Let us verify that $\mathcal{A}$ is a σ-algebra. Since X is contained in $\mathcal{A}$ and $\mathcal{A}$ is closed under complementation, it suffices to show that $A := \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ whenever $A_n \in \mathcal{A}$. If all A_n are at most countable, this is obvious. Suppose that among the sets A_n there is at least one set A_{n_1} whose complement is at most countable. The complement of A is contained in the complement of A_{n_1}, hence is at most countable as well, i.e., $A \in \mathcal{A}$. All one-point sets belong to $\mathcal{A}$, hence the σ-algebra $\mathcal{A}_0$ generated by them is contained in $\mathcal{A}$. On the other hand, it is clear that any set in $\mathcal{A}$ is an element of $\mathcal{A}_0$, whence it follows that $\mathcal{A}_0 = \mathcal{A}$. □

Let us give definitions of several other classes of sets employed in measure theory.

1.2.13. Definition. (i) *A family $\mathcal{R}$ of subsets of a set X is called a ring if it contains the empty set and the sets $A \cap B$, $A \cup B$ and $A \backslash B$ belong to $\mathcal{R}$ for all $A, B \in \mathcal{R}$;*

(ii) *A family $\mathcal{S}$ of subsets of a set X is called a semiring if it contains the empty set, $A \cap B \in \mathcal{S}$ for all $A, B \in \mathcal{S}$ and, for every pair of sets $A, B \in \mathcal{S}$ with $A \subset B$, the set $B \backslash A$ is the union of finitely many disjoint sets in $\mathcal{S}$. If $X \in \mathcal{S}$, then $\mathcal{S}$ is called a semialgebra*;

(iii) *A ring is called a σ-ring if it is closed with respect to countable unions. A ring is called a δ-ring if it is closed with respect to countable intersections.*

As an example of a ring that is not an algebra, let us mention the collection of all bounded sets on the real line. The family of all intervals in the interval $[a, b]$ gives an example of a semiring that is not a ring. According to the following lemma, the collection of all finite unions of elements of a semiring is a ring (called the ring generated by the given semiring). It is clear that this is the minimal ring containing the given semiring.

1.2.14. Lemma. *For any semiring $\mathcal{S}$, the collection of all finite unions of sets in $\mathcal{S}$ forms a ring $\mathcal{R}$. Every set in $\mathcal{R}$ is a finite union of pairwise disjoint sets in $\mathcal{S}$. If $\mathcal{S}$ is a semialgebra, then $\mathcal{R}$ is an algebra.*

Proof. It is clear that the class $\mathcal{R}$ admits finite unions. Suppose that $A = A_1 \cup \cdots \cup A_n$, $B = B_1 \cup \cdots \cup B_k$, where $A_i, B_j \in \mathcal{S}$. Then we have $A \cap B = \bigcup_{i \le n, j \le k} A_i \cap B_j \in \mathcal{R}$. Hence $\mathcal{R}$ admits finite intersections. In addition,

$$A \backslash B = \bigcup_{i=1}^{n} \Bigl(A_i \backslash \bigcup_{j=1}^{k} B_j \Bigr) = \bigcup_{i=1}^{n} \bigcap_{j=1}^{k} (A_i \backslash B_j).$$

Since the set $A_i \backslash B_j = A_i \backslash (A_i \cap B_j)$ is a finite union of sets in $\mathcal{S}$, one has $A \backslash B \in \mathcal{R}$. Clearly, A can be written as a union of a finitely many disjoint sets in $\mathcal{S}$ because $\mathcal{S}$ is closed with respect to intersections. The last claim of the lemma is obvious. □

Note that for any σ-algebra $\mathcal{B}$ in a space X and any set $A \subset X$, the class $\mathcal{B}_A := \{B \cap A \colon\ B \in \mathcal{B}\}$ is a σ-algebra in the space A. This σ-algebra is called the trace σ-algebra.

1.3. Additivity and countable additivity of measures

Functions with values in $(-\infty, +\infty)$ will be called real or real-valued. In the cases where we discuss functions with values in the extended real line $[-\infty, +\infty]$, this will always be specified.

1.3.1. Definition. *A real-valued set function μ defined on a class of sets $\mathcal{A}$ is called additive* (*or finitely additive*) *if*

$$\mu\Bigl(\bigcup_{i=1}^{n} A_i\Bigr) = \sum_{i=1}^{n} \mu(A_i) \tag{1.3.1}$$

for all n and all disjoint sets $A_1, \ldots, A_n \in \mathcal{A}$ such that $\bigcup_{i=1}^{n} A_i \in \mathcal{A}$.

In the case where $\mathcal{A}$ is closed with respect to finite unions, the finite additivity is equivalent to the equality

$$\mu(A \cup B) = \mu(A) + \mu(B) \tag{1.3.2}$$

for all disjoint sets $A, B \in \mathcal{A}$.

If the domain of definition of an additive real-valued set function μ contains the empty set $\varnothing$, then $\mu(\varnothing) = 0$. In particular, this is true for any additive set function on a ring or an algebra.

It is also useful to consider the property of *subadditivity* (also called the *semiadditivity*):

$$\mu\Bigl(\bigcup_{i=1}^{n} A_i\Bigr) \le \sum_{i=1}^{n} \mu(A_i) \tag{1.3.3}$$

for all $A_i \in \mathcal{A}$ with $\bigcup_{i=1}^{n} A_i \in \mathcal{A}$. Any additive nonnegative set function on an algebra is subadditive (see below).

1.3.2. Definition. *A real-valued set function μ on a class of sets $\mathcal{A}$ is called countably additive if*

$$\mu\Bigl(\bigcup_{n=1}^{\infty} A_n\Bigr) = \sum_{n=1}^{\infty} \mu(A_n) \tag{1.3.4}$$

for all pairwise disjoint sets A_n in $\mathcal{A}$ such that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$. A countably additive set function defined on an algebra is called a measure.

It is readily seen from the definition that the series in (1.3.4) converges absolutely because its sum is independent of rearrangements of its terms.

1.3.3. Proposition. *Let μ be an additive real set function on an algebra* (*or a ring*) *of sets $\mathcal{A}$. Then the following conditions are equivalent*:

(i) *the function μ is countably additive,*

(ii) *the function μ is continuous at zero in the following sense: if $A_n \in \mathcal{A}$, $A_{n+1} \subset A_n$ for all $n \in \mathbb{N}$ and $\bigcap_{n=1}^{\infty} A_n = \varnothing$, then*

$$\lim_{n\to\infty} \mu(A_n) = 0, \tag{1.3.5}$$

(iii) *the function μ is continuous from below, i.e., if $A_n \in \mathcal{A}$ are such that $A_n \subset A_{n+1}$ for all $n \in \mathbb{N}$ and $\bigcup_{n=1}^\infty A_n \in \mathcal{A}$, then*

$$\mu\Bigl(\bigcup_{n=1}^\infty A_n\Bigr) = \lim_{n\to\infty} \mu(A_n). \tag{1.3.6}$$

PROOF. (i) Let μ be countably additive and let the sets $A_n \in \mathcal{A}$ decrease monotonically to the empty set. Set $B_n = A_n \backslash A_{n+1}$. The sets B_n belong to $\mathcal{A}$ and are disjoint and their union is A_1. Hence the series $\sum_{n=1}^\infty \mu(B_n)$ converges. Then $\sum_{n=N}^\infty \mu(B_n)$ tends to zero as $N \to \infty$, but the sum of this series is $\mu(A_N)$, since $\bigcup_{n=N}^\infty B_n = A_N$. Hence we arrive at condition (ii).

Suppose now that condition (ii) is fulfilled. Let $\{B_n\}$ be a sequence of pairwise disjoint sets in $\mathcal{A}$ whose union B is an element of $\mathcal{A}$ as well. Set $A_n = B\backslash \bigcup_{k=1}^n B_k$. It is clear that $\{A_n\}$ is a sequence of monotonically decreasing sets in $\mathcal{A}$ with the empty intersection. By hypothesis, $\mu(A_n) \to 0$. By the finite additivity this means that $\sum_{k=1}^n \mu(B_k) \to \mu(B)$ as $n \to \infty$. Hence μ is countably additive. Clearly, (iii) follows from (ii), for if the sets $A_n \in \mathcal{A}$ increase monotonically and their union is the set $A \in \mathcal{A}$, then the sets $A\backslash A_n \in \mathcal{A}$ decrease monotonically to the empty set. Finally, by the finite additivity (iii) yields the countable additivity of μ. $\square$

The reader is warned that there is no such equivalence for semialgebras (see Exercise 1.12.75).

1.3.4. Definition. *A countably additive measure μ on a σ-algebra of subsets of a space X is called a probability measure if $\mu \geq 0$ and $\mu(X) = 1$.*

1.3.5. Definition. *A triple $(X, \mathcal{A}, \mu)$ is called a measure space if μ is a nonnegative measure on a σ-algebra $\mathcal{A}$ of subset of a set X. If μ is a probability measure, then $(X, \mathcal{A}, \mu)$ is called a probability space.*

Nonnegative not identically zero measures are called *positive measures*.

Additive set functions are also called additive measures, but to simplify the terminology we use the term measure only for *countably additive measures on algebras or rings*. Countably additive measures are also called σ-additive measures.

1.3.6. Definition. *A measure defined on the Borel σ-algebra of the whole space $\mathbb{R}^n$ or its subset is called a Borel measure.*

It is clear that if $\mathcal{A}$ is an algebra, then the additivity is just equality (1.3.2) for arbitrary disjoint sets in $\mathcal{A}$. Similarly, if $\mathcal{A}$ is a σ-algebra, then the countable additivity is equality (1.3.4) for arbitrary sequences of disjoint sets in $\mathcal{A}$. The above given formulations are convenient for two reasons. First, the validity of the corresponding equalities is required only for those collections of sets for which both parts make sense. Second, as we shall see later, under natural hypotheses, additive (or countably additive) set functions admit additive (respectively, countably additive) extensions to larger classes of sets that admit unions of the corresponding type.

1.3.7. Example. Let $\mathcal{A}$ be the algebra of sets $A \subset \mathbb{N}$ such that either A or $\mathbb{N}\backslash A$ is finite. For finite A, let $\mu(A) = 0$, and for A with a finite complement let $\mu(A) = 1$. Then μ is an additive, but not countably additive set function.

PROOF. It is clear that $\mathcal{A}$ is indeed an algebra. Relation (1.3.2) is obvious for disjoint sets A and B if A is finite. Finally, A and B in $\mathcal{A}$ cannot be infinite simultaneously being disjoint. If μ were countably additive, we would have had $\mu(\mathbb{N}) = \sum_{n=1}^{\infty} \mu(\{n\}) = 0$. □

There exist additive, but not countably additive set functions on σ-algebras (see Example 1.12.28). The simplest countably additive set function is identically zero. Another example: let X be a nonempty set and let $a \in X$; Dirac's measure δ_a at the point a is defined as follows: for every $A \subset X$, $\delta_a(A) = 1$ if $a \in A$ and $\delta_a(A) = 0$ otherwise. Let us give a slightly less trivial example.

1.3.8. Example. Let $\mathcal{A}$ be the σ-algebra of all subsets of $\mathbb{N}$. For every set $A = \{n_k\}$, let $\mu(A) = \sum_k 2^{-n_k}$. Then μ is a measure on $\mathcal{A}$.

In order to construct less trivial examples (say, Lebesgue measure), we need auxiliary technical tools discussed in the next section.

Note several simple properties of additive and countably additive set functions.

1.3.9. Proposition. *Let μ be a nonnegative additive set function on an algebra or a ring $\mathcal{A}$.*

(i) *If $A, B \in \mathcal{A}$ and $A \subset B$, then $\mu(A) \le \mu(B)$.*

(ii) *For any collection $A_1, \ldots, A_n \in \mathcal{A}$ one has*

$$\mu\Bigl(\bigcup_{i=1}^{n} A_i\Bigr) \le \sum_{i=1}^{n} \mu(A_i).$$

(iii) *The function μ is countably additive precisely when in addition to the additivity it is countably subadditive in the following sense: for any sequence $\{A_n\} \subset \mathcal{A}$ with $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ one has*

$$\mu\Bigl(\bigcup_{n=1}^{\infty} A_n\Bigr) \le \sum_{n=1}^{\infty} \mu(A_n).$$

PROOF. Assertion (i) follows, since $\mu(B\backslash A) \ge 0$. Assertion (ii) is easily verified by induction taking into account the nonnegativity of μ and the relation $\mu(A \cup B) = \mu(A\backslash B) + \mu(B\backslash A) + \mu(A \cap B)$.

If μ is countably additive and the union of sets $A_n \in \mathcal{A}$ belongs to $\mathcal{A}$, then according to Proposition 1.3.3 one has

$$\mu\Bigl(\bigcup_{i=1}^{n} A_i\Bigr) \to \mu\Bigl(\bigcup_{i=1}^{\infty} A_i\Bigr),$$

which by (ii) gives the estimate indicated in (iii). Finally, such an estimate combined with the additivity yields the countable additivity. Indeed, let B_n be pairwise disjoint sets in $\mathcal{A}$ whose union B belongs to $\mathcal{A}$ as well. Then for any $n \in \mathbb{N}$ we have

$$\sum_{k=1}^{n} \mu(B_k) = \mu\Bigl(\bigcup_{k=1}^{n} B_k\Bigr) \le \mu(B) \le \sum_{k=1}^{\infty} \mu(B_k),$$

whence it follows that $\sum_{k=1}^{\infty} \mu(B_k) = \mu(B)$. □

1.3.10. Proposition. *Let $\mathcal{A}_0$ be a semialgebra (see Definition 1.2.13). Then every additive set function μ on $\mathcal{A}_0$ uniquely extends to an additive set function on the algebra $\mathcal{A}$ generated by $\mathcal{A}_0$ (i.e., the family of all finite unions of sets in $\mathcal{A}_0$). This extension is countably additive provided that μ is countably additive on $\mathcal{A}_0$. The same is true in the case of a semiring $\mathcal{A}$ and the ring generated by it.*

PROOF. By Lemma 1.2.14 the collection of all finite unions of elements of $\mathcal{A}_0$ is an algebra (or a ring when $\mathcal{A}_0$ is a semiring). It is clear that any set in $\mathcal{A}$ can be represented as a union of disjoint elements of $\mathcal{A}_0$. Set

$$\mu(A) = \sum_{i=1}^{n} \mu(A_i)$$

if $A_i \in \mathcal{A}_0$ are pairwise disjoint and their union is A. The indicated extension is obviously additive, but we have to verify that it is well-defined, i.e., is independent of partitioning A into parts in $\mathcal{A}_0$. Indeed, if $B_1, \ldots, B_m$ are pairwise disjoint sets in $\mathcal{A}_0$ whose union is A, then by the additivity of μ on the algebra $\mathcal{A}_0$ one has the equality $\mu(A_i) = \sum_{j=1}^{m} \mu(A_i \cap B_j)$, $\mu(B_j) = \sum_{i=1}^{n} \mu(A_i \cap B_j)$, whence the desired conclusion follows. Let us verify the countable additivity of the indicated extension in the case of the countable additivity on $\mathcal{A}_0$. Let $A, A_n \in \mathcal{A}$, $A = \bigcup_{n=1}^{\infty} A_n$ be such that $A_n \cap A_k = \varnothing$ if $n \neq k$. Then

$$A = \bigcup_{j=1}^{N} B_j, \quad A_n = \bigcup_{i=1}^{N_n} B_{n,i},$$

where $B_j, B_{n,i} \in \mathcal{A}_0$. Set $C_{n,i,j} := B_{n,i} \cap B_j$. The sets $C_{n,i,j}$ are pairwise disjoint and

$$B_j = \bigcup_{n=1}^{\infty} \bigcup_{i=1}^{N_n} C_{n,i,j}, \quad B_{n,i} = \bigcup_{j=1}^{N} C_{n,i,j}.$$

By the countable additivity of μ on $\mathcal{A}_0$ we have

$$\mu(B_j) = \sum_{n=1}^{\infty} \sum_{i=1}^{N_n} \mu(C_{n,i,j}), \quad \mu(B_{n,i}) = \sum_{j=1}^{N} \mu(C_{n,i,j}),$$

and by the definition of μ on $\mathcal{A}$ one has the following equality:

$$\mu(A) = \sum_{j=1}^{N} \mu(B_j), \quad \mu(A_n) = \sum_{i=1}^{N_n} \mu(B_{n,i}).$$

We obtain from these equalities that $\mu(A) = \sum_{n=1}^{\infty} \mu(A_n)$, since both quantities equal the sum of all $\mu(C_{n,i,j})$. That it is possible to interchange the summations in n and j is obvious from the fact that the series in n converge and the sums in j and i are finite. □

1.4. Compact classes and countable additivity

In this section, we give a sufficient condition for the countable additivity, which is satisfied for most of the measures encountered in real applications.

1.4.1. Definition. *A family $\mathcal{K}$ of subsets of a set X is called a compact class if, for any sequence K_n of its elements with $\bigcap_{n=1}^{\infty} K_n = \varnothing$, there exists N such that $\bigcap_{n=1}^{N} K_n = \varnothing$.*

The terminology is explained by the following basic example.

1.4.2. Example. An arbitrary family of compact sets in $\mathbb{R}^n$ (more generally, in a topological space) is a compact class.

PROOF. Indeed, let K_n be compact sets whose intersection is empty. Suppose that for every n the set $E_n = \bigcap_{i=1}^{n} K_i$ contains some element x_n. We may assume that no element of the sequence $\{x_n\}$ is repeated infinitely often, since otherwise it is a common element of all E_n. By the compactness of K_1 there exists a point x each neighborhood of which contains infinitely many elements of the sequence $\{x_n\}$. All sets E_n are compact and $x_i \in E_n$ whenever $i \geq n$, hence the point x belongs to all E_n, which is a contradiction. □

Note that some authors call the above-defined compact classes countably compact or semicompact and in the definition of compact classes require the following stronger property: if the intersection of a (possibly uncountable) collection of sets in $\mathcal{K}$ is empty, then the intersection of some its finite subcollection is empty as well. See Exercise 1.12.105 for an example distinguishing the two properties. Although such a terminology is more consistent from the point of view of topology (see Exercise 6.10.66 in Chapter 6), we shall not follow it.

1.4.3. Theorem. *Let μ be a nonnegative additive set function on an algebra $\mathcal{A}$. Suppose that there exists a compact class $\mathcal{K}$ approximating μ in the following sense: for every $A \in \mathcal{A}$ and every $\varepsilon > 0$, there exist $K_\varepsilon \in \mathcal{K}$ and $A_\varepsilon \in \mathcal{A}$ such that $A_\varepsilon \subset K_\varepsilon \subset A$ and $\mu(A \backslash A_\varepsilon) < \varepsilon$. Then μ is countably additive. In particular, this is true if the compact class $\mathcal{K}$ is contained in $\mathcal{A}$ and for any $A \in \mathcal{A}$ one has the equality*

$$\mu(A) = \sup_{K \subset A,\ K \in \mathcal{K}} \mu(K).$$

PROOF. Suppose that the sets $A_n \in \mathcal{A}$ are decreasing and their intersection is empty. Let us show that $\mu(A_n) \to 0$. Let us fix $\varepsilon > 0$. By hypothesis, there exist $K_n \in \mathcal{K}$ and $B_n \in \mathcal{A}$ such that $B_n \subset K_n \subset A_n$ and $\mu(A_n \backslash B_n) < \varepsilon 2^{-n}$. It is clear that $\bigcap_{n=1}^{\infty} K_n \subset \bigcap_{n=1}^{\infty} A_n = \varnothing$. By the definition of a compact class, there exists N such that $\bigcap_{n=1}^{N} K_n = \varnothing$. Then $\bigcap_{n=1}^{N} B_n = \varnothing$. Note that one has

$$A_N = \bigcap_{n=1}^{N} A_n \subset \bigcup_{n=1}^{N} (A_n \backslash B_n).$$

Indeed, let $x \in A_N$, i.e., $x \in A_n$ for all $n \leq N$. If x does not belong to $\bigcup_{n=1}^{N} (A_n \backslash B_n)$, then $x \notin A_n \backslash B_n$ for all $n \leq N$. Then $x \in B_n$ for every $n \leq N$, whence we obtain $x \in \bigcap_{n=1}^{N} B_n$, which is a contradiction. The above proved equality yields the estimate

$$\mu(A_N) \leq \sum_{n=1}^{N} \mu(A_n \backslash B_n) \leq \sum_{n=1}^{N} \varepsilon 2^{-n} \leq \varepsilon.$$

Hence $\mu(A_n) \to 0$, which implies the countable additivity of μ. □

1.4.4. Example. Let I be an interval in $\mathbb{R}^1$, $\mathcal{A}$ the algebra of finite unions of intervals in I (closed, open and half-open). Then the usual length λ_1, which assigns the value $b - a$ to the interval with the endpoints a and b and extends by additivity to their finite disjoint unions, is countably additive on the algebra $\mathcal{A}$.

PROOF. Finite unions of closed intervals form a compact class and approximate from within finite unions of arbitrary intervals. □

1.4.5. Example. Let I be a cube in $\mathbb{R}^n$ of the form $[a, b]^n$ and let $\mathcal{A}$ be the algebra of finite unions of the parallelepipeds in I that are products of intervals in $[a, b]$. Then the usual volume λ_n is countably additive on $\mathcal{A}$. We call λ_n *Lebesgue measure.*

PROOF. As in the previous example, finite unions of closed parallelepipeds form a compact approximating class. □

It is shown in Theorem 1.12.5 below that the compactness property can be slightly relaxed.

The previous results justify the introduction of the following concept.

1.4.6. Definition. *Let m be a nonnegative function on a class $\mathcal{E}$ of subsets of a set X and let $\mathcal{P}$ be a class of subsets of X, too. We say that $\mathcal{P}$ is an approximating class for m if, for every $E \in \mathcal{E}$ and every $\varepsilon > 0$, there exist $P_\varepsilon \in \mathcal{P}$ and $E_\varepsilon \in \mathcal{E}$ such that $E_\varepsilon \subset P_\varepsilon \subset E$ and $|m(E) - m(E_\varepsilon)| < \varepsilon$.*

1.4.7. Remark. (i) The reasoning in Theorem 1.4.3 actually proves the following assertion. Let μ be a nonnegative additive set function on an algebra $\mathcal{A}$ and let $\mathcal{A}_0$ be a subalgebra in $\mathcal{A}$. Suppose that there exists a compact class $\mathcal{K}$ approximating μ on $\mathcal{A}_0$ with respect to $\mathcal{A}$ in the following sense: for any $A \in \mathcal{A}_0$ and any $\varepsilon > 0$, there exist $K_\varepsilon \in \mathcal{K}$ and $A_\varepsilon \in \mathcal{A}$ such that $A_\varepsilon \subset K_\varepsilon \subset A$ and $\mu(A \backslash A_\varepsilon) < \varepsilon$. Then μ is countably additive on $\mathcal{A}_0$.

(ii) The compact class $\mathcal{K}$ in Theorem 1.4.3 need not be contained in $\mathcal{A}$. For example, if $\mathcal{A}$ is the algebra generated by all intervals in $[0,1]$ with rational endpoints and μ is Lebesgue measure, then the class $\mathcal{K}$ of all finite unions of closed intervals with irrational endpoints is approximating for μ and has no intersection with $\mathcal{A}$. However, it will be shown in §1.12(ii) that one can always replace $\mathcal{K}$ by a compact class $\mathcal{K}'$ that is contained in $\sigma(\mathcal{A})$ and approximates the countably additive extension of μ on $\sigma(\mathcal{A})$. It is worth noting that there exists a countably additive extension of μ to the σ-algebra generated by $\mathcal{A}_0$ and $\mathcal{K}$ (see Theorem 1.12.34).

Note that so far in the considered examples we have been concerned with the countable additivity on algebras. However, as we shall see below, any countably additive measure on an algebra automatically extends (in a unique way) to a countably additive measure on the σ-algebra generated by this algebra.

We shall see in Chapter 7 that the class of measures possessing a compact approximating class is very large (so that it is not easy even to construct an example of a countably additive measure without compact approximating classes). Thus, the described sufficient condition of countable additivity has a very universal character. Here we only give the following result.

1.4.8. Theorem. *Let μ be a nonnegative countably additive measure on the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ in the space $\mathbb{R}^n$. Then, for any Borel set $B \subset \mathbb{R}^n$ and any $\varepsilon > 0$, there exist an open set U_ε and a compact set K_ε such that $K_\varepsilon \subset B \subset U_\varepsilon$ and $\mu(U_\varepsilon \backslash K_\varepsilon) < \varepsilon$.*

Proof. Let us show that for any $\varepsilon > 0$ there exists a closed set $F_\varepsilon \subset B$ such that

$$\mu(B \backslash F_\varepsilon) < \varepsilon/2.$$

Then, by the countable additivity of μ, the set F_ε itself can be approximated from within up to $\varepsilon/2$ by $F_\varepsilon \cap U$, where U is a closed ball of a sufficiently large radius. Denote by $\mathcal{A}$ the class of all sets $A \in \mathcal{B}(\mathbb{R}^n)$ such that, for any $\varepsilon > 0$, there exist a closed set F_ε and an open set U_ε with $F_\varepsilon \subset A \subset U_\varepsilon$ and $\mu(U_\varepsilon \backslash F_\varepsilon) < \varepsilon$. Every closed set A belongs to $\mathcal{A}$, since one can take for F_ε the set A itself, and for U_ε one can take some open δ-neighborhood A^δ of the set A, i.e., the union of all open balls of radius δ with centers at the points in A. When δ is decreasing to zero, the open sets A^δ are decreasing to A, hence their measures approach the measure of A. Let us show that $\mathcal{A}$ is a σ-algebra. If this is done, then the theorem is proven, for the closed sets generate the Borel σ-algebra. By construction, the class $\mathcal{A}$ is closed with respect to the operation of complementation. Hence it remains to verify the stability of $\mathcal{A}$ with respect to countable unions. Let $A_j \in \mathcal{A}$ and let $\varepsilon > 0$. Then there exist a closed set F_j and an open set U_j such that $F_j \subset A_j \subset U_j$ and $\mu(U_j \backslash F_j) < \varepsilon 2^{-j}$, $j \in \mathbb{N}$.

The set $U = \bigcup_{j=1}^{\infty} U_j$ is open and the set $Z_k = \bigcup_{j=1}^{k} F_j$ is closed for any $k \in \mathbb{N}$. It remains to observe that $Z_k \subset \bigcup_{j=1}^{\infty} A_j \subset U$ and for k large enough one has the estimate $\mu(U \backslash Z_k) < \varepsilon$. Indeed, $\mu\bigl(\bigcup_{j=1}^{\infty}(U_j \backslash F_j)\bigr) < \sum_{j=1}^{\infty} \varepsilon 2^{-j} = \varepsilon$ and by the countable additivity $\mu(Z_k) \to \mu\bigl(\bigcup_{j=1}^{\infty} F_j\bigr)$ as $k \to \infty$. $\square$

This result shows that the measurability can be defined (as it is actually done in some textbooks) in the spirit of the Jordan–Peano construction via inner approximations by compact sets and outer approximations by open sets. Certainly, it is necessary for this to define first the measure of open sets, which determines the measures of compacts. In the case of an interval this creates no problem, since open sets are built from disjoint intervals, which by virtue of the countable additivity uniquely determines its measure from the measures of intervals. However, already in the case of a square there is no such disjoint representation of open sets, and the aforementioned construction is not as effective here.

Finally, it is worth mentioning that Lebesgue measure considered above on the algebra generated by cubes could be defined at once on the Borel σ-algebra by the equality $\lambda_n(B) := \inf \sum_{j=1}^{\infty} \lambda_n(I_j)$, where inf is taken over all at most countable covers of B by cubes I_j. In fact, exactly this will be done below, however, a justification of the fact that the indicated equality gives a countably additive measure is not trivial and will be given by some detour, where the principal role will be played by the idea of compact approximations and the construction of outer measure, with which the next section is concerned.

1.5. Outer measure and the Lebesgue extension of measures

It is shown in this section how to extend countably additive measures from algebras to σ-algebras. Extensions from rings are considered in §1.11.

For any nonnegative set function μ that is defined on a certain class $\mathcal{A}$ of subsets in a space X and contains X itself, the formula

$$\mu^*(A) = \inf\biggl\{\sum_{n=1}^{\infty} \mu(A_n) \Big| A_n \in \mathcal{A},\ A \subset \bigcup_{n=1}^{\infty} A_n\biggr\}$$

defines a new set function defined already for every $A \subset X$. The same construction is applicable to set functions with values in $[0, +\infty]$. If X does not belong to $\mathcal{A}$, then μ^* is defined by the above formula on all sets A that can be covered by a countable sequence of elements of $\mathcal{A}$, and all other sets are assigned the infinite value. An alternative definition of μ^* on a set A that cannot be covered by a sequence from $\mathcal{A}$ is to take the supremum of the values of μ^* on the sets contained in A and covered by sequences from $\mathcal{A}$ (see Example 1.12.130). The function μ^* is called the outer measure, although it need not be additive. In Section 1.11 below we discuss in more detail Carathéodory outer measures, not necessarily originated from additive set functions.

1.5.1. Definition. *Suppose that μ is a nonnegative set function on domain $\mathcal{A} \subset 2^X$. A set A is called μ-measurable (or Lebesgue measurable with respect to μ) if, for any $\varepsilon > 0$, there exists $A_\varepsilon \in \mathcal{A}$ such that*

$$\mu^*(A \triangle A_\varepsilon) < \varepsilon.$$

The class of all μ-measurable sets is denoted by $\mathcal{A}_\mu$.

We shall be interested in the case where μ is a countably additive measure on an algebra $\mathcal{A}$.

Note that the definition of measurability given by Lebesgue (for an interval X) was the equality $\mu^*(A) + \mu^*(X\backslash A) = \mu(X)$. It is shown below that for additive functions on algebras this definition (possibly not so intuitively transparent) is equivalent to the one given above (see Theorem 1.11.8 and also Proposition 1.5.11 for countably additive measures). In addition, we discuss below the definition of the Carathéodory measurability, which is also equivalent to the above definition in the case of nonnegative additive set functions on algebras, but is much more fruitful in the general case.

1.5.2. Example. (i) Let $\varnothing \in \mathcal{A}$ and $\mu(\varnothing) = 0$. Then $\mathcal{A} \subset \mathcal{A}_\mu$ (if $A \in \mathcal{A}$, one can take $A_\varepsilon = A$). In addition, any set A with $\mu^*(A) = 0$ is μ-measurable, for one can take $A_\varepsilon = \varnothing$.

(ii) Let $\mathcal{A}$ be the algebra of finite unions of intervals from Example 1.4.4 with the usual length λ. Then, the λ-measurability of A is equivalent to the following: for each $\varepsilon > 0$, one can find a set E that is a finite union of intervals and two sets A'_ε and A''_ε with

$$A = (E \cup A'_\varepsilon)\backslash A''_\varepsilon,\ \lambda^*(A'_\varepsilon) \le \varepsilon,\ \lambda^*(A''_\varepsilon) \le \varepsilon.$$

(iii) Let $X = [0,1]$, $\mathcal{A} = \{\varnothing, X\}$, $\mu(X) = 1$, $\mu(\varnothing) = 0$. Then μ is a countably additive measure on $\mathcal{A}$ and $\mathcal{A}_\mu = \mathcal{A}$. Indeed, $\mu^*(E) = 1$ for any $E \neq \varnothing$. Hence the whole interval is the only nonempty set that can be approximated up to $\varepsilon < 1$ by a set from $\mathcal{A}$.

Note that μ^* is *monotone*, i.e., $\mu^*(A) \le \mu^*(B)$ if $A \subset B$. However, even if μ is a countably additive measure on a σ-algebra $\mathcal{A}$, the corresponding outer measure μ^* may not be countably additive on the class of all sets.

1.5.3. Example. Let X be a two-point set $\{0,1\}$ and let $\mathcal{A} = \{\varnothing, X\}$. Set $\mu(\varnothing) = 0$, $\mu(X) = 1$. Then $\mathcal{A}$ is a σ-algebra and μ is countably additive on $\mathcal{A}$, but μ^* is not additive on the σ-algebra of all sets, since $\mu^*(\{0\}) = 1$, $\mu^*(\{1\}) = 1$, and $\mu^*(\{0\} \cup \{1\}) = 1$.

1.5.4. Lemma. *Let μ be a nonnegative set function on a class $\mathcal{A}$. Then the function μ^* is countably subadditive, i.e.,*

$$\mu^*\Bigl(\bigcup_{n=1}^\infty A_n\Bigr) \le \sum_{n=1}^\infty \mu^*(A_n) \tag{1.5.1}$$

for any sets A_n.

PROOF. Let $\varepsilon > 0$ and $\mu^*(A_n) < \infty$. For any n, there exists a collection $\{B_{n,k}\}_{k=1}^\infty \subset \mathcal{A}$ such that $A_n \subset \bigcup_{k=1}^\infty B_{n,k}$ and

$$\sum_{k=1}^\infty \mu(B_{n,k}) \le \mu^*(A_n) + \frac{\varepsilon}{2^n}.$$

Then $\bigcup_{n=1}^\infty A_n \subset \bigcup_{n=1}^\infty \bigcup_{k=1}^\infty B_{n,k}$ and hence

$$\mu^*\Bigl(\bigcup_{n=1}^\infty A_n\Bigr) \le \sum_{n=1}^\infty \sum_{k=1}^\infty \mu(B_{n,k}) \le \sum_{n=1}^\infty \mu^*(A_n) + \varepsilon.$$

Since ε is arbitrary, we arrive at (1.5.1). □

1.5.5. Lemma. *In the situation of the previous lemma, for any sets A and B such that $\mu^*(B) < \infty$ one has the inequality*

$$|\mu^*(A) - \mu^*(B)| \le \mu^*(A \bigtriangleup B). \tag{1.5.2}$$

PROOF. We observe that $A \subset B \cup (A \bigtriangleup B)$, whence by the subadditivity of μ^* we obtain the estimate

$$\mu^*(A) \le \mu^*(B) + \mu^*(A \bigtriangleup B),$$

i.e., $\mu^*(A) - \mu^*(B) \le \mu^*(A \bigtriangleup B)$. The estimate $\mu^*(B) - \mu^*(A) \le \mu^*(A \bigtriangleup B)$ is obtained in a similar manner. □

1.5.6. Theorem. *Let μ be a nonnegative countably additive set function on an algebra $\mathcal{A}$. Then*:

(i) *one has $\mathcal{A} \subset \mathcal{A}_\mu$, and the outer measure μ^* coincides with μ on $\mathcal{A}$*;

(ii) *the collection $\mathcal{A}_\mu$ of all μ-measurable sets is a σ-algebra and the restriction of μ^* to $\mathcal{A}_\mu$ is countably additive*;

(iii) *the function μ^* is a unique nonnegative countably additive extension of μ to the σ-algebra $\sigma(\mathcal{A})$ generated by $\mathcal{A}$ and a unique nonnegative countably additive extension of μ to $\mathcal{A}_\mu$.*

PROOF. (i) It has already been noted that $\mathcal{A} \subset \mathcal{A}_\mu$. Let $A \in \mathcal{A}$ and $A \subset \bigcup_{n=1}^\infty A_n$, where $A_n \in \mathcal{A}$. Then $A = \bigcup_{n=1}^\infty (A \cap A_n)$. Hence by Proposition 1.3.9(iii) we have

$$\mu(A) \le \sum_{n=1}^\infty \mu(A \cap A_n) \le \sum_{n=1}^\infty \mu(A_n),$$

whence we obtain $\mu(A) \le \mu^*(A)$. By definition, $\mu^*(A) \le \mu(A)$. Therefore, $\mu(A) = \mu^*(A)$.

(ii) First we observe that the complement of a measurable set A is measurable. This is seen from the formula $(X\backslash A) \bigtriangleup (X\backslash A_\varepsilon) = A \bigtriangleup A_\varepsilon$. Next, the union of two measurable sets A and B is measurable. Indeed, let $\varepsilon > 0$ and let A_ε, $B_\varepsilon \in \mathcal{A}$ be such that $\mu^*(A \bigtriangleup A_\varepsilon) < \varepsilon/2$ and $\mu^*(B \bigtriangleup B_\varepsilon) < \varepsilon/2$. Since

$$(A \cup B) \bigtriangleup (A_\varepsilon \cup B_\varepsilon) \subset (A \bigtriangleup A_\varepsilon) \cup (B \bigtriangleup B_\varepsilon),$$

one has

$$\mu^*\Big((A \cup B) \bigtriangleup (A_\varepsilon \cup B_\varepsilon)\Big) \le \mu^*\Big((A \bigtriangleup A_\varepsilon) \cup (B \bigtriangleup B_\varepsilon)\Big) < \varepsilon.$$

Therefore, $A \cup B \in \mathcal{A}_\mu$. In addition, by what has already been proven, we have $A \cap B = X\backslash\big((X\backslash A) \cup (X\backslash B)\big) \in \mathcal{A}_\mu$. Hence $\mathcal{A}_\mu$ is an algebra.

Let us now establish two less obvious properties of the outer measure. First we verify its additivity on $\mathcal{A}_\mu$. Let $A, B \in \mathcal{A}_\mu$, where $A \cap B = \varnothing$. Let us fix $\varepsilon > 0$ and find $A_\varepsilon, B_\varepsilon \in \mathcal{A}$ such that

$$\mu^*(A \bigtriangleup A_\varepsilon) < \varepsilon/2 \quad \text{and} \quad \mu^*(B \bigtriangleup B_\varepsilon) < \varepsilon/2.$$

By Lemma 1.5.5, taking into account that μ^* and μ coincide on $\mathcal{A}$, we obtain

$$\mu^*(A \cup B) \ge \mu(A_\varepsilon \cup B_\varepsilon) - \mu^*\Big((A \cup B) \bigtriangleup (A_\varepsilon \cup B_\varepsilon)\Big). \tag{1.5.3}$$

By the inclusion $(A\cup B)\bigtriangleup(A_\varepsilon \cup B_\varepsilon) \subset (A\bigtriangleup A_\varepsilon)\cup(B\bigtriangleup B_\varepsilon)$ and the subadditivity of μ^* one has the inequality

$$\mu^*\Big((A \cup B) \bigtriangleup (A_\varepsilon \cup B_\varepsilon)\Big) \le \mu^*(A \bigtriangleup A_\varepsilon) + \mu^*(B \bigtriangleup B_\varepsilon) \le \varepsilon. \tag{1.5.4}$$

By the inclusion $A_\varepsilon \cap B_\varepsilon \subset (A \bigtriangleup A_\varepsilon) \cup (B \bigtriangleup B_\varepsilon)$ we have

$$\mu(A_\varepsilon \cap B_\varepsilon) = \mu^*(A_\varepsilon \cap B_\varepsilon) \le \mu^*(A \bigtriangleup A_\varepsilon) + \mu^*(B \bigtriangleup B_\varepsilon) \le \varepsilon.$$

Hence the estimates $\mu(A_\varepsilon) \ge \mu^*(A) - \varepsilon/2$ and $\mu(B_\varepsilon) \ge \mu^*(B) - \varepsilon/2$ yield

$$\mu(A_\varepsilon \cup B_\varepsilon) = \mu(A_\varepsilon) + \mu(B_\varepsilon) - \mu(A_\varepsilon \cap B_\varepsilon) \ge \mu^*(A) + \mu^*(B) - 2\varepsilon.$$

Taking into account relationships (1.5.3) and (1.5.4) we obtain

$$\mu^*(A \cup B) \ge \mu^*(A) + \mu^*(B) - 3\varepsilon.$$

Since ε is arbitrary, one has $\mu^*(A \cup B) \ge \mu^*(A) + \mu^*(B)$. By the reverse inequality $\mu^*(A \cup B) \le \mu^*(A) + \mu^*(B)$, we conclude that

$$\mu^*(A \cup B) = \mu^*(A) + \mu^*(B).$$

The next important step is a verification of the fact that countable unions of measurable sets are measurable. It suffices to prove this for disjoint sets $A_n \in \mathcal{A}_\mu$. Indeed, in the general case one can write $B_n = A_n \backslash \bigcup_{k=1}^{n-1} A_k$. Then the sets B_n are pairwise disjoint and measurable according to what we have already proved; they have the same union as the sets A_n. Dealing now with disjoint sets, we observe that by the finite additivity of μ^* on $\mathcal{A}_\mu$ the following relations are valid:

$$\sum_{k=1}^n \mu^*(A_k) = \mu^*\Big(\bigcup_{k=1}^n A_k\Big) \le \mu^*\Big(\bigcup_{k=1}^\infty A_k\Big) \le \mu(X).$$

Hence $\sum\limits_{k=1}^\infty \mu^*(A_k) < \infty$. Let $\varepsilon > 0$. We can find n such that

$$\sum_{k=n+1}^\infty \mu^*(A_k) < \frac{\varepsilon}{2}.$$

By using the measurability of finite unions one can find a set $B \in \mathcal{A}$ such that $\mu^*\Bigl(\Bigl(\bigcup_{k=1}^n A_k\Bigr) \triangle B\Bigr) < \varepsilon/2$. Since

$$\Bigl(\bigcup_{k=1}^\infty A_k\Bigr) \triangle B \subset \Bigl(\Bigl(\bigcup_{k=1}^n A_k\Bigr) \triangle B\Bigr) \cup \Bigl(\bigcup_{k=n+1}^\infty A_k\Bigr),$$

we obtain

$$\begin{aligned}\mu^*\Bigl(\Bigl(\bigcup_{k=1}^\infty A_k\Bigr) \triangle B\Bigr) &\le \mu^*\Bigl(\Bigl(\bigcup_{k=1}^n A_k\Bigr) \triangle B\Bigr) + \mu^*\Bigl(\bigcup_{k=n+1}^\infty A_k\Bigr)\\ &\le \frac{\varepsilon}{2} + \sum_{k=n+1}^\infty \mu^*(A_k) < \varepsilon.\end{aligned}$$

Thus, $\bigcup_{k=1}^\infty A_k$ is measurable. Therefore, $\mathcal{A}_\mu$ is a σ-algebra. It remains to note that the additivity and countable subadditivity of μ^* on $\mathcal{A}_\mu$ yield the countable additivity (see Proposition 1.3.9).

(iii) We observe that $\sigma(\mathcal{A}) \subset \mathcal{A}_\mu$, since $\mathcal{A}_\mu$ is a σ-algebra containing $\mathcal{A}$. Let ν be some nonnegative countably additive extension of μ to $\sigma(\mathcal{A})$. Let $A \in \sigma(\mathcal{A})$ and $\varepsilon > 0$. It has been proven that $A \in \mathcal{A}_\mu$, hence there exists $B \in \mathcal{A}$ with $\mu^*(A \triangle B) < \varepsilon$. Therefore, there exist sets $C_n \in \mathcal{A}$ such that $A \triangle B \subset \bigcup_{n=1}^\infty C_n$ and $\sum_{n=1}^\infty \mu(C_n) < \varepsilon$. Then we obtain

$$|\nu(A) - \nu(B)| \le \nu(A \triangle B) \le \sum_{n=1}^\infty \nu(C_n) = \sum_{n=1}^\infty \mu(C_n) < \varepsilon.$$

Since $\nu(B) = \mu(B) = \mu^*(B)$, we finally obtain

$$\begin{aligned}|\nu(A) - \mu^*(A)| &= |\nu(A) - \nu(B) + \mu^*(B) - \mu^*(A)|\\ &\le |\nu(A) - \nu(B)| + |\mu^*(B) - \mu^*(A)| \le 2\varepsilon.\end{aligned}$$

We arrive at the equality $\nu(A) = \mu^*(A)$ because ε is arbitrary. This reasoning also shows the uniqueness of a nonnegative countably additive extension of μ to $\mathcal{A}_\mu$, since we have only used that $A \in \mathcal{A}_\mu$ (however, as noted below, it is important that we deal with nonnegative extensions). □

A control question: where does the above proof employ the countable additivity of μ?

1.5.7. Example. Let $\mathcal{A}$ be the algebra of all finite subsets of $\mathbb{N}$ and their complements and let μ equal 0 on finite sets and 1 on their complements. Then μ is additive and the single-point sets $\{n\}$ cover $\mathbb{N}$, hence $\mu^*(\mathbb{N}) = 0 < \mu(\mathbb{N})$.

It is worth noting that in the above theorem μ has no signed countably additive extensions from $\mathcal{A}$ to $\sigma(\mathcal{A})$, which follows by (iii) and the Jordan decomposition constructed in Chapter 3 (see §3.1), but it may have signed extensions to $\mathcal{A}_\mu$. For example, this happens if we take $X = \{0,1\}$ and let $\mathcal{A} = \sigma(\mathcal{A}) = \{\varnothing, X\}$, $\mu \equiv 0$, $\nu(\{0\}) = 1$, $\nu(\{1\}) = -1$, $\nu(X) = 0$.

An important special case, to which the extension theorem applies, is the situation of Example 1.4.5. Since the σ-algebra generated by the cubes with edges parallel to the coordinate axes is the Borel σ-algebra, we obtain a countably additive Lebesgue measure λ_n on the Borel σ-algebra of the cube (and even on a larger σ-algebra), which extends the elementary volume. This measure is considered in greater detail in §1.7. By Theorem 1.5.6, the Lebesgue measure of any Borel (as well as any measurable) set B in the cube is $\lambda_n^*(B)$. Now the question arises why we do not define at once the measure on the Borel σ-algebra of the cube by this formula. The point is that there is a difficulty in the verification of the additivity of the obtained set function. This difficulty is circumvented by considering the algebra generated by the parallelepipeds, where the additivity is obvious.

With the aid of the proven theorem one can give a new description of measurable sets.

1.5.8. Corollary. *Let μ be a nonnegative countably additive set function on an algebra $\mathcal{A}$. A set A is μ-measurable precisely when there exist two sets $A', A'' \in \sigma(\mathcal{A})$ such that*

$$A' \subset A \subset A'' \quad \textit{and} \quad \mu^*(A''\backslash A') = 0.$$

Moreover, one can take for A' a set of the form $\bigcup_{n=1}^\infty \bigcap_{k=1}^\infty A_{n,k}$, $A_{n,k} \in \mathcal{A}$, and for A'' a set of the form $\bigcap_{n=1}^\infty \bigcup_{k=1}^\infty B_{n,k}$, $B_{n,k} \in \mathcal{A}$.

Proof. Let $A \in \mathcal{A}_\mu$. Then, for any $\varepsilon > 0$, there exists a set $A_\varepsilon \in \sigma(\mathcal{A})$ such that $A \subset A_\varepsilon$ and $\mu^*(A) \geq \mu^*(A_\varepsilon) - \varepsilon$. Indeed, by definition there exist sets $A_n \in \mathcal{A}$ with $A \subset \bigcup_{n=1}^\infty A_n$ and $\mu^*(A) \geq \sum_{n=1}^\infty \mu(A_n) - \varepsilon$. Let $A_\varepsilon = \bigcup_{n=1}^\infty A_n$. It is clear that $A \subset A_\varepsilon$, $A_\varepsilon \in \sigma(\mathcal{A}) \subset \mathcal{A}_\mu$ and by the countable additivity of μ^* on $\mathcal{A}_\mu$ we have $\mu^*(A_\varepsilon) \leq \sum_{n=1}^\infty \mu(A_n)$. Set

$$A'' = \bigcap_{n=1}^\infty A_{1/n}.$$

Then $A \subset A'' \in \sigma(\mathcal{A}) \subset \mathcal{A}_\mu$ and $\mu^*(A) = \mu^*(A'')$, since

$$\mu^*(A) \geq \mu^*(A_{1/n}) - 1/n \geq \mu^*(A'') - 1/n$$

for all n. Note that for constructing A'' the measurability of A is not needed. Let us apply this to the complement of A and find a set $B \in \sigma(\mathcal{A}) \subset \mathcal{A}_\mu$ such that $X\backslash A \subset B$ and $\mu(B) = \mu^*(X\backslash A)$. Set $A' = X\backslash B$. Then we obtain $A' \subset A$, and by the additivity of μ^* on the σ-algebra $\mathcal{A}_\mu$ and the inclusion $A, B \in \mathcal{A}_\mu$ we have

$$\mu^*(A') = \mu(X) - \mu^*(B) = \mu(X) - \mu^*(X\backslash A) = \mu^*(A),$$

which is the required relation. Conversely, suppose that such sets A' and A'' exist. Since A is the union of A' and a subset of $A''\backslash A'$, it suffices to verify that every subset C in $A''\backslash A'$ belongs to $\mathcal{A}_\mu$. This is indeed true because $\mu^*(C) \leq \mu^*(A''\backslash A') = \mu^*(A'') - \mu^*(A') = 0$ by the additivity of μ^* on $\mathcal{A}_\mu$ and the inclusion $A'', A' \in \sigma(\mathcal{A}) \subset \mathcal{A}_\mu$. □

The uniqueness of extension yields the following useful result.

1.5.9. Corollary. *For the equality of two nonnegative Borel measures μ and ν on the real line it is necessary and sufficient that they coincide on all open intervals* (*or all closed intervals*).

Proof. Any closed interval is the intersection of a decreasing sequence of open intervals and any open interval is the union of an increasing sequence of closed intervals. By the countable additivity the equality of μ and ν on open intervals is equivalent to their equality on closed intervals and implies the equality of both measures on the algebra generated by intervals in $\mathbb{R}^1$. Since this algebra generates $\mathcal{B}(\mathbb{R}^1)$, our assertion follows by the uniqueness of a countably additive extension from an algebra to the generated σ-algebra. $\square$

The countably additive extension described in Theorem 1.5.6 is called the *Lebesgue extension* or the *Lebesgue completion* of the measure μ, and the measure space $(X, \mathcal{A}_\mu, \mu)$ is called the Lebesgue completion of $(X, \mathcal{A}, \mu)$. In addition, $\mathcal{A}_\mu$ is called the Lebesgue completion of the σ-algebra $\mathcal{A}$ with respect to μ. This terminology is related to the fact that the measure μ on $\mathcal{A}_\mu$ is complete in the sense of the following definition.

1.5.10. Definition. *A nonnegative countably additive measure μ on a σ-algebra $\mathcal{A}$ is called complete if $\mathcal{A}$ contains all subsets of every set in $\mathcal{A}$ with μ-measure zero. In this case we say that the σ-algebra $\mathcal{A}$ is complete with respect to the measure μ.*

It is clear from the definition of outer measure that if $A \subset B \in \mathcal{A}_\mu$ and $\mu(B) = 0$, then $A \in \mathcal{A}_\mu$ and $\mu(A) = 0$. It is easy to construct an example of a countably additive measure on a σ-algebra that is not complete: it suffices to take the identically zero measure on the σ-algebra consisting of the empty set and the interval $[0, 1]$. As a less trivial example let us mention Lebesgue measure on the σ-algebra of all Borel subsets of the interval constructed according to Example 1.4.4. This measure is considered below in greater detail; we shall see that there exist compact sets of zero Lebesgue measure containing non-Borel subsets.

Let us note the following simple but useful criterion of measurability of a set in terms of outer measure (which is, as already remarked, the original Lebesgue definition).

1.5.11. Proposition. *Let μ be a nonnegative countably additive measure on an algebra $\mathcal{A}$. Then, a set A belongs to $\mathcal{A}_\mu$ if and only if one has*

$$\mu^*(A) + \mu^*(X \backslash A) = \mu(X).$$

This is also equivalent to the equality $\mu^*(E \cap A) + \mu^*(E \backslash A) = \mu^*(E)$ for all sets $E \subset X$.

Proof. Let us verify the sufficiency of the first condition (then the stronger second one is sufficient too). Let us find μ-measurable sets B and C such that $A \subset B$, $X \backslash A \subset C$, $\mu(B) = \mu^*(A)$, $\mu(C) = \mu^*(X \backslash A)$. The existence of such sets has been established in the proof of Corollary 1.5.8. Clearly, $D = X\backslash C \subset A$ and

$$\mu(B) - \mu(D) = \mu(B) + \mu(C) - \mu(X) = 0.$$

Hence $\mu^*(A \triangle B) = 0$, whence the measurability of A follows.

Let us now prove that the second condition above is necessary. By the subadditivity of the outer measure it suffices to verify that $\mu^*(E\cap A)+\mu^*(E\backslash A) \le \mu^*(E)$ for any $E \subset X$ and any measurable A. It follows from (1.5.2) that it suffices to establish this inequality for all $A \in \mathcal{A}$. Let $\varepsilon > 0$ and let sets $A_n \in \mathcal{A}$ be such that $E \subset \bigcup_{n=1}^\infty A_n$ and $\mu^*(E) \ge \sum_{n=1}^\infty \mu(A_n) - \varepsilon$. Then $E \cap A \subset \bigcup_{n=1}^\infty (A_n \cap A)$ and $E\backslash A \subset \bigcup_{n=1}^\infty (A_n\backslash A)$, whence we obtain

$$\begin{aligned}\mu^*(E \cap A) + \mu^*(E\backslash A) &\le \sum_{n=1}^\infty \mu(A_n \cap A) + \sum_{n=1}^\infty \mu(A_n\backslash A)\\ &= \sum_{n=1}^\infty \mu(A_n) \le \mu^*(E) + \varepsilon.\end{aligned}$$

Since ε is arbitrary, our claim is proven. □

Note that this criterion of measurability can be formulated as the equality $\mu^*(A) = \mu_*(A)$ if we define the inner measure by the equality

$$\mu_*(A) := \mu(X) - \mu^*(X\backslash A),$$

as Lebesgue actually did. It is important that in this case one must not use the definition of inner measure in the spirit of the Jordan measure as the supremum of measures of the sets from $\mathcal{A}$ inscribed in A. Below we shall return to the discussion of outer measures and see that the last property in Proposition 1.5.11 can be taken for a definition of measurability, which leads to very interesting results. In turn, this proposition will be extended to finitely additive set functions.

Let us observe that any set $A \in \mathcal{A}_\mu$ can be made a measure space by restricting μ to the class of μ-measurable subsets of A, which is a σ-algebra in A. The obtained measure μ_A (or $\mu|_A$) is called the restriction of μ to A. Restrictions to arbitrary sets are considered in §1.12(iv).

We close this section by proving the following property of continuity from below for outer measure.

1.5.12. Proposition. *Let μ be a nonnegative measure on a σ-algebra $\mathcal{A}$. Suppose that sets A_n are such that $A_n \subset A_{n+1}$ for all $n \in \mathbb{N}$. Then, one has*

$$\mu^*\Bigl(\bigcup_{n=1}^\infty A_n\Bigr) = \lim_{n\to\infty} \mu^*(A_n). \tag{1.5.5}$$

PROOF. According to Corollary 1.5.8, there exist μ-measurable sets B_n such that $A_n \subset B_n$ and $\mu(B_n) = \mu^*(A_n)$. Set

$$B = \bigcup_{n=1}^\infty \bigcap_{k=n}^\infty B_k.$$

One has $A_n \subset B_k$ if $k \geq n$, hence $A_n \subset B$ and $\bigcup_{n=1}^{\infty} A_n \subset B$. Therefore,

$$\mu^*\Bigl(\bigcup_{n=1}^{\infty} A_n\Bigr) \leq \mu(B) = \lim_{n\to\infty} \mu\Bigl(\bigcap_{k=n}^{\infty} B_k\Bigr) \leq \limsup_{n\to\infty} \mu(B_n) = \lim_{n\to\infty} \mu^*(A_n).$$

Since the reverse inequality is also true, the claim is proven. $\square$

1.6. Infinite and σ-finite measures

We have so far been discussing finite measures, but one has to deal with infinite measures as well. The simplest (and most important) example is Lebesgue measure on $\mathbb{R}^n$. There are several ways of introducing set functions with infinite values. The first one is to admit set functions with values in the extended real line. For simplicity let us confine ourselves to nonnegative set functions. Let $c + \infty = \infty$ for any $c \in [0, +\infty]$. Now we can define the finite or countable additivity of set functions on algebras and σ-algebras (or rings, semirings, semialgebras) in the same way as above. In particular, we keep the definitions of outer measure and measurability. In this situation we use the term "a countably additive measure with values in $[0, +\infty]$". Similarly, one can consider measures with values in $(-\infty, +\infty]$ or $[-\infty, +\infty)$. A certain drawback of this approach is that rather pathological measures arise such as the countably additive measure that assigns $+\infty$ to all nonempty sets.

1.6.1. Definition. *Let $\mathcal{A}$ be a σ-algebra in a space X and let μ be a set function on $\mathcal{A}$ with values in $[0, +\infty]$ that satisfies the condition $\mu(\varnothing) = 0$ and is countably additive in the sense that $\mu\Bigl(\bigcup_{j=1}^{\infty} A_j\Bigr) = \sum_{j=1}^{\infty} \mu(A_j)$ for all pairwise disjoint sets $A_j \in \mathcal{A}$, where infinite values are admissible as well. Then μ is called a measure with values in $[0, +\infty]$. We call μ a σ-finite measure if $X = \bigcup_{n=1}^{\infty} X_n$, where $X_n \in \mathcal{A}$, $\mu(X_n) < \infty$.*

A desire to consider only measures with real but possibly unbounded values leads to modification of requirements on domains of definitions of measures; this is the second option. Here the concepts of a ring and δ-ring of sets introduced in Definition 1.2.13 become useful. For example, a natural domain of definition of Lebesgue measure on $\mathbb{R}^n$ could be the collection $\mathcal{L}_n^0$ of all sets of finite Lebesgue measure, i.e., all sets $E \subset \mathbb{R}^n$ such that measures of the sets $E_k := E \cap \{x\colon |x_i| \leq k,\ i = 1, \ldots, n\}$ in cubes (where we have already defined Lebesgue measure) are uniformly bounded in k. Lebesgue measure on $\mathcal{L}_n^0$ is given by the formula $\lambda_n(E) = \lim_{k\to\infty} \lambda_n(E_k)$. It is clear that the class $\mathcal{L}_n^0$ is a δ-ring. Lebesgue measure is countably additive on $\mathcal{L}_n^0$ (see below). In the next section we discuss the properties of Lebesgue measure on $\mathbb{R}^n$ in greater detail.

In what follows when considering infinite measures we always specify which definition we have in mind. Some additional information about measures with values in the extended real line (including their extensions and measurability with respect to such measures) is given in the final section and exercises.

1.6.2. Lemma. *Let $\mathcal{R}$ be a ring of subsets of a space X (i.e., $\mathcal{R}$ is closed with respect to finite intersections and unions, $\varnothing \in \mathcal{R}$ and $A\backslash B \in \mathcal{R}$ for all $A, B \in \mathcal{R}$). Let μ be a countably additive set function on $\mathcal{R}$ with values in $[0, +\infty]$ such that there exist sets $X_n \in \mathcal{R}$ with $X = \bigcup_{n=1}^{\infty} X_n$ and $\mu(X_n) < \infty$. Denote by μ_n the Lebesgue extension of the measure μ regarded on the set $S_n := \bigcup_{j=1}^{n} X_j$ equipped with the algebra of sets consisting of the intersections of elements in $\mathcal{R}$ with S_n. Let $\mathcal{L}_{\mu_n}$ denote the class of all μ_n-measurable sets. Let*

$$\mathcal{A} = \bigl\{A \subset X\colon\ A \cap S_n \in \mathcal{L}_{\mu_n}\ \forall\, n \in \mathbb{N},\ \overline{\mu}(A) := \lim_{n\to\infty} \mu_n(A \cap S_n) < \infty\bigr\}.$$

Then $\mathcal{A}$ is a ring closed with respect to countable intersections (i.e., a δ-ring) and $\overline{\mu}$ is a σ-additive measure whose restriction to every set S_n coincides with μ.

Proof. Let $A_i \in \mathcal{A}$ be pairwise disjoint sets with union in $\mathcal{A}$. We denote this union by A. For every n, the sets $A_i \cap S_n$ are disjoint too, hence

$$\mu_n(A \cap S_n) = \sum_{i=1}^{\infty} \mu_n(A_i \cap S_n).$$

Since $A \in \mathcal{A}$, the left-hand side of this equality is increasing to $\overline{\mu}(A)$. Therefore, $\sum_{i=1}^{\infty} \mu_n(A_i \cap S_n) \le \overline{\mu}(A)$ for all n, whence it follows by the equality $\lim\limits_{n\to\infty} \mu_n(A_i \cap S_n) = \overline{\mu}(A_i)$ for every i that $\sum_{i=1}^{\infty} \overline{\mu}(A_i) \le \overline{\mu}(A)$. This yields that $\overline{\mu}$ is a countably additive measure. Let $E \in \mathcal{R}$. Then the sets $E \cap \bigcup_{i=1}^{n} X_i$ belong to $\mathcal{R}$ and increase to E, which gives $\mu(E) = \overline{\mu}(E)$. Other claims are obvious. □

1.6.3. Remark. Suppose that in the situation of Lemma 1.6.2 the space X is represented as the union of another sequence of sets X_n' in $\mathcal{R}$ with finite measures. Then, as is clear from the lemma, this sequence yields the same extension of μ and the same class $\mathcal{A}$.

1.6.4. Example. Let $\mathcal{L}_n$ be the class of all sets $E \subset \mathbb{R}^n$ such that all the sets $E_k := E \cap \{x\colon\ |x_i| \le k,\ i = 1, \ldots, n\}$ are Lebesgue measurable. Then $\mathcal{L}_n$ is a σ-algebra, on which the function $\lambda_n(E) = \lim\limits_{k\to\infty} \lambda_n(E_k)$ is a σ-finite measure (called Lebesgue measure on $\mathbb{R}^n$). The σ-algebra $\mathcal{L}_n$ contains the above-considered δ-ring $\mathcal{L}_n^0$. If we apply the previous lemma to the ring of all bounded Lebesgue measurable sets, then we arrive at the δ-ring $\mathcal{L}_n^0$.

In addition to Lebesgue measure, σ-finite measures arise as Haar measures on locally compact groups and Riemannian volumes on manifolds. Sometimes in diverse problems of analysis, algebra, geometry and probability theory one has to deal with products of finite and σ-finite measures. Although the list of infinite measures encountered in real problems is not very large, it is useful to have a terminology which enables one to treat various concrete examples in a unified way. Many of our earlier-obtained assertions remain valid for infinite measures. We only give the following result extending Theorem 1.5.6, which is directly seen from the reasoning there (the details of proof are left as Exercise 1.12.78); this result also follows from Theorem 1.11.8 below.

1.6.5. Proposition. *Let μ be a countably additive measure on an algebra $\mathcal{A}$ with values in $[0,+\infty]$. Then $\mathcal{A}_\mu$ is a σ-algebra, $\mathcal{A}\subset\mathcal{A}_\mu$, and the function μ^* is a countably additive measure on $\mathcal{A}_\mu$ with values in $[0,+\infty]$ and coincides with μ on $\mathcal{A}$.*

However, there are exceptions. For example, for infinite measures, the countable additivity does not imply that the measures of sets A_n monotonically decreasing to the empty set approach zero. The point is that all the sets A_n may have infinite measures. In many books measures are defined from the very beginning as functions with values in $[0,+\infty]$. Then, in theorems, one has often to impose various additional conditions (moreover, different in different theorems; the reader will find a lot of examples in the exercises on infinite measures in Chapters 1–4). It appears that at least in a graduate course it is better to first establish all theorems for bounded measures, then observe that most of them remain valid for σ-finite measures, and finally point out that further generalizations are possible, but they require additional hypotheses. Our exposition will be developed according to this principle.

1.7. Lebesgue measure

Let us return to the situation considered in Example 1.4.5 and briefly discussed after Theorem 1.5.6. Let I be a cube in $\mathbb{R}^n$ of the form $[a,b]^n$, $\mathcal{A}_0$ the algebra of finite unions of parallelepipeds in I with edges parallel to the coordinate axes. As we know, the usual volume λ_n is countably additive on $\mathcal{A}_0$. Therefore, one can extend λ_n to a countably additive measure, also denoted by λ_n, on the σ-algebra $\mathcal{L}_n(I)$ of all λ_n-measurable sets in I, which contains the Borel σ-algebra. We write $\mathbb{R}^n$ as the union of the increasing sequence of cubes $I_k=\{|x_i|\le k, i=1,\ldots,n\}$ and denote by λ_n the σ-finite measure generated by Lebesgue measures on the cubes I_k according to the construction of the previous section (see Example 1.6.4). Let

$$\mathcal{L}_n=\{E\subset\mathbb{R}^n\colon\ E\cap I_k\in\mathcal{L}_n(I_k),\ \forall\, k\in\mathbb{N}\}.$$

1.7.1. Definition. *The above-defined measure λ_n on $\mathcal{L}_n$ is called Lebesgue measure on $\mathbb{R}^n$. The sets in $\mathcal{L}_n$ are called Lebesgue measurable.*

In the case where a subset of $\mathbb{R}^n$ is regarded with Lebesgue measure, it is customary to use the terms "measure zero set", "measurable set" etc. without explicitly mentioning Lebesgue measure. We also follow this tradition.

For defining Lebesgue measure of a set $E\in\mathcal{L}_n$ one can use the formula

$$\lambda_n(E)=\lim_{k\to\infty}\lambda_n(E\cap I_k)$$

as well as the formula

$$\lambda_n(E)=\sum_{j=1}^{\infty}\lambda_n(E\cap Q_j),$$

where Q_j are pairwise disjoint cubes that are translations of $[-1,1)^n$ and whose union is all of $\mathbb{R}^n$. Since the σ-algebra generated by the parallelepipeds of the above-mentioned form is the Borel σ-algebra $\mathcal{B}(I)$ of the cube I, we see that all Borel sets in the cube I, hence in $\mathbb{R}^n$ as well, are Lebesgue measurable.

Lebesgue measure can also be regarded on the δ-ring $\mathcal{L}_n^0$ of all sets of finite Lebesgue measure.

In the case of $\mathbb{R}^1$ Lebesgue measure of a set E is the sum of the series of $\lambda_1\bigl(E\cap(n,n+1]\bigr)$ over all integer numbers n.

The translation of a set A by a vector h, i.e., the set of all points of the form $a+h$, where $a\in A$, is denoted by $A+h$.

1.7.2. Lemma. *Let W be an open set in the cube $I=(-1,1)^n$. Then there exists an at most countable family of open pairwise disjoint cubes Q_j in W of the form $Q_j=c_jI+h_j$, $c_j>0$, $h_j\in W$, such that the set $W\backslash\bigcup_{j=1}^\infty Q_j$ has Lebesgue measure zero.*

PROOF. Let us employ Exercise 1.12.48 and write W as $W=\bigcup_{j=1}^\infty W_j$, where W_j are open cubes whose edges are parallel to the coordinate axes and have lengths $q2^{-p}$ with positive integer p,q, and whose centers have the coordinates of the form $l2^{-m}$ with integer l and positive integer m. Next we restructure the cubes W_j as follows. We delete all cubes W_j that are contained in W_1 and set $Q_1=W_1$. Let us take the first cube W_{n_2} in the remaining sequence and represent the interior of the body $W_{n_2}\backslash Q_1$ as the finite union of open pairwise disjoint cubes $Q_2,\ldots,Q_{m_2}$ of the same type as the cubes W_j and some pieces of the boundaries of these new cubes. This is possible by our choice of the initial cubes. Next we delete all the cubes W_j that are contained in $\bigcup_{i=1}^{m_2}Q_i$, take the first cube in the remaining sequence and construct a partition of its part that is not contained in the previously constructed cubes in the same way as explained above. Continuing the described process, we obtain pairwise disjoint cubes that cover W up to a measure zero set, namely, up to a countable union of boundaries of these cubes. $\square$

In Exercise 1.12.72, it is suggested that the reader modify this reasoning to make it work for any Borel measure. We have only used above that the boundaries of our cubes have measure zero. Note that the lengths of the edges of the constructed cubes are rational.

1.7.3. Theorem. *Let A be a Lebesgue measurable set of finite measure. Then*:

(i) $\lambda_n(A+h)=\lambda_n(A)$ *for any vector* $h\in\mathbb{R}^n$;
(ii) $\lambda_n\bigl(U(A)\bigr)=\lambda_n(A)$ *for any orthogonal linear operator* U *on* $\mathbb{R}^n$;
(iii) $\lambda_n(\alpha A)=|\alpha|^n\lambda_n(A)$ *for any real number* α.

PROOF. It follows from the definition of Lebesgue measure that it suffices to prove the listed properties for bounded measurable sets.

(i) Let us take a cube I centered at the origin such that the sets A and $A+h$ are contained in some cube inside I. Let $\mathcal{A}_0$ be the algebra generated by all cubes in I with edges parallel to the coordinate axes. When evaluating the outer measure of A it suffices to consider only sets $B \in \mathcal{A}_0$ with $B+h \subset I$. Since the volumes of sets in $\mathcal{A}_0$ are invariant under translations, the sets $A+h$ and A have equal outer measures. For every $\varepsilon > 0$, there exists a set $A_\varepsilon \in \mathcal{A}_0$ with $\lambda_n^*(A \triangle A_\varepsilon) < \varepsilon$. Then

$$\lambda_n^*\bigl((A+h) \triangle (A_\varepsilon + h)\bigr) = \lambda_n^*\bigl((A \triangle A_\varepsilon) + h\bigr) = \lambda_n^*(A \triangle A_\varepsilon) < \varepsilon,$$

whence we obtain the measurability of $A+h$ and the desired equality.

(ii) As in (i), it suffices to prove our claim for sets in $\mathcal{A}_0$. Hence it remains to show that, for any closed cube K with edges parallel to the coordinate axes, one has the equality

$$\lambda_n\bigl(U(K)\bigr) = \lambda_n(K). \tag{1.7.1}$$

Suppose that this is not true for some cube K, i.e.,

$$\lambda_n\bigl(U(K)\bigr) = r\lambda_n(K),$$

where $r \neq 1$. Let us show that for every ball $Q \subset I$ centered at the origin one has

$$\lambda_n\bigl(U(Q)\bigr) = r\lambda_n(Q) \quad \text{if} \quad U(Q) \subset I. \tag{1.7.2}$$

Let d be the length of the edge of K. Let us take an arbitrary natural number p and partition the cube K into p^n equal smaller closed cubes K_j that have equal edges of length d/p and disjoint interiors (i.e., may have in common only parts of faces). The cubes $U(K_j)$ are translations of each other and have equal measures as proved above. It is readily seen that faces of any cube have measure zero. Hence $\lambda_n\bigl(U(K)\bigr) = p^n\lambda_n\bigl(U(K_1)\bigr)$. Therefore, $\lambda_n\bigl(U(K_1)\bigr) = r\lambda_n(K_1)$. Then (1.7.2) is true for any cube of the form $qK+h$, where q is a rational number. This yields equality (1.7.2) for the ball Q. Indeed, by additivity this equality extends to finite unions of cubes with edges parallel to the coordinate axes. Next, for any $\varepsilon > 0$, one can find two such unions E_1 and E_2 with $E_1 \subset Q \subset E_2$ and $\lambda_n(E_2 \backslash E_1) < \varepsilon$. To this end, it suffices to take balls Q' and Q'' centered at the origin such that $Q' \subset Q \subset Q''$ with strict inclusions and a small measure of $Q'' \backslash Q'$. Then one can find a finite union E_1 of cubes of the indicated form with $Q' \subset E_1 \subset Q$ and an analogous union E_2 with $Q \subset E_2 \subset Q''$. It remains to observe that $U(Q) = Q$, and (1.7.2) leads to contradiction.

(iii) The last claim is obvious for sets in $\mathcal{A}_0$, hence as claims (i) and (ii), it extends to arbitrary measurable sets. □

It is worth noting that property (iii) of Lebesgue measure is a corollary of property (i), since by (i) it is valid for all cubes and $\alpha = 1/m$, where m is any natural number, then it extends to all rational α, which yields the general case by continuity. It is seen from the proof that property (ii) also follows from property (i). Property (i) characterizes Lebesgue measure up to a constant factor (see Exercise 1.12.74). There is an alternative derivation of property (ii) from properties (i) and (iii), employing the invariance of the ball with respect to rotations and the following theorem, which is very interesting in its own right.

1.7.4. Theorem. *Let W be a nonempty open set in $\mathbb{R}^n$. Then, there exists a countable collection of pairwise disjoint open balls $U_j \subset W$ such that the set $W\backslash\bigcup_{j=1}^{\infty} U_j$ has measure zero.*

PROOF. It suffices to prove the theorem in the case where $\lambda_n(W) < \infty$ (we may even assume that W is contained in a cube). Let $K = (-1,1)^n$ and let V be the open ball inscribed in K. It is clear that $\lambda_n(V) = \alpha\lambda_n(K)$, where $0 < \alpha < 1$. Set $q = 1 - \alpha$. Let us take a number $\beta > 1$ such that $q\beta < 1$. By Lemma 1.7.2, the set W can be written as the union of a measure zero set and a sequence of open pairwise disjoint cubes K_j of the form $K_j = c_j K + h_j$, where $c_j > 0$ and $h_j \in \mathbb{R}^n$. In every cube K_j we inscribe the open ball $V_j = c_j V + h_j$. Since $\lambda_n(V_j)/\lambda_n(K_j) = \alpha$, we obtain

$$\lambda_n(K_j\backslash V_j) = \lambda_n(K_j) - \lambda_n(V_j) = q\lambda_n(K_j).$$

Hence

$$\lambda_n\Bigl(W\backslash\bigcup_{j=1}^{\infty} V_j\Bigr) = \sum_{j=1}^{\infty}\lambda_n(K_j\backslash V_j) = q\sum_{j=1}^{\infty}\lambda_n(K_j) = q\lambda_n(W).$$

Let us take a finite number of these cubes such that

$$\lambda_n\Bigl(W\backslash\bigcup_{j=1}^{N_1} V_j\Bigr) \le \beta q\lambda_n(W).$$

Set $V_j^{(1)} = V_j$, $j \le N_1$. Let us repeat the described construction for the open set W_1 obtained from W by deleting the closures of the balls $V_1,\ldots,V_{N_1}$ (we observe that a finite union of closed sets is closed). We obtain pairwise disjoint open balls $V_j^{(2)} \subset W_1$, $j \le N_2$, such that

$$\lambda_n\Bigl(W_1\backslash\bigcup_{j=1}^{N_2} V_j^{(2)}\Bigr) \le \beta q\lambda_n(W_1) \le (\beta q)^2\lambda_n(W).$$

By induction, we obtain a countable family of pairwise disjoint open balls $V_j^{(k)}$, $j \le N_k$, with the following property: if Z_k is the union of the closures of the balls $V_1^{(k)},\ldots,V_{N_k}^{(k)}$ and $W_k = W_{k-1}\backslash Z_k$, where $W_0 = W$, then

$$\lambda_n\Bigl(W_k\backslash\bigcup_{j=1}^{N_{k+1}} V_j^{(k+1)}\Bigr) \le (\beta q)^{k+1}\lambda_n(W).$$

Since $(\beta q)^k \to 0$, the set $W\backslash\bigcup_{k=1}^{\infty}\bigcup_{j=1}^{N_k} V_j^{(k)}$ has measure zero. □

It is clear that in the formulation of this theorem the balls U_j can be replaced by any sets of the form $c_j S + h_j$, where S is a fixed bounded set of positive measure. Indeed, the proof only employed the translation invariance of Lebesgue measure and the relation $\lambda_n(rA) = r^n\lambda_n(A)$ for $r > 0$. In

Chapter 5 (Corollary 5.8.3) this theorem will be extended to arbitrary Borel measures.

Note that it follows by Theorem 1.7.3 that Lebesgue measure of any rectangular parallelepiped $P \subset I$ (not necessarily with edges parallel to the coordinate axes) equals the product of lengths of its edges. Clearly, any countable set has Lebesgue measure zero. As the following example of the Cantor set (named after the outstanding German mathematician Georg Cantor) shows, there exist uncountable sets of Lebesgue measure zero as well.

1.7.5. Example. Let $I = [0,1]$. Denote by $J_{1,1}$ the interval $(1/3, 2/3)$. Let $J_{2,1}$ and $J_{2,2}$ denote the intervals $(1/9, 2/9)$ and $(7/9, 8/9)$, which are the middle thirds of the intervals obtained after deleting $J_{1,1}$. Continue this process inductively by deleting the open middle intervals. After the nth step we obtain 2^n closed intervals; at the next step we delete their open middle thirds $J_{n+1,1}, \ldots, J_{n+1,2^n}$, after which there remains 2^{n+1} closed intervals, and the process continues. The set $C = I \backslash \bigcup_{n,j} J_{n,j}$ is called the Cantor set. It is compact, has cardinality of the continuum, but its Lebesgue measure is zero.

PROOF. The set C is compact, since its complement is open. In order to see that C has cardinality of the continuum, we write the points in $[0,1]$ in the ternary expansion, i.e., $x = \sum_{j=1}^{\infty} x_j 3^{-j}$, where x_j takes values $0, 1, 2$. As in the decimal expansion, this representation is not unique, since, for example, the sequence $(1, 1, 2, 2, \ldots)$ corresponds to the same number as the sequence $(1, 2, 0, 0, \ldots)$. However, this non-uniqueness is only possible for points of some countable set, which we denote by M. It is verified by induction that after the nth step of deleting there remain the points x such that $x_j = 0$ or $x_j = 2$ if $j \leq n$. Thus, $C \backslash M$ consists of all points whose ternary expansion involves only 0 and 2, whence it follows that C has cardinality of the set of all reals. Finally, in order to show that C has zero measure, it remains to verify that the complement of C in $[0,1]$ has measure 1. By induction one verifies that the measure of the set $J_{n,1} \cup \cdots \cup J_{n,2^{n-1}}$ equals $2^{n-1}3^{-n}$. Since $\sum_{n=1}^{\infty} 2^{n-1}3^{-n} = 1$, our claim is proven. □

1.7.6. Example. Let $\varepsilon > 0$ and let $\{r_n\}$ be the set of all rational numbers in $[0,1]$. Set $K = [0,1] \backslash \bigcup_{n=1}^{\infty} (r_n - \varepsilon 4^{-n}, r_n + \varepsilon 4^{-n})$. Then K is a compact set without inner points and its Lebesgue measure is not less than $1 - \varepsilon$ because the measure of the complement does not exceed $2\varepsilon \sum_{n=1}^{\infty} 4^{-n}$.

Thus, a compact set of positive measure may have the empty interior. A similar example (but with some additional interesting properties) can be constructed by a modification of the construction of the Cantor set. Namely, at every step one deletes a bit less than the middle third so that the sum of the deleted intervals becomes $1 - \varepsilon$.

Note that any subset of the Cantor set has measure zero, too. Therefore, the family of all measurable sets has cardinality equal to that of the class of all subsets of the real line. As we shall see below, the Borel σ-algebra has cardinality of the continuum. Hence among subsets of the Cantor set there are non-Borel Lebesgue measurable sets. The existence of non-Borel Lebesgue measurable sets will be established below in a more constructive way by means of the Souslin operation.

Now the question naturally arises how large the class of all Lebesgue measurable sets is and whether it includes all the sets. It turns out that an answer to this question depends on additional set-theoretic axioms and cannot be given in the framework of the "naive set theory" without the axiom of choice. In any case, as the following example due to Vitali shows, by means of the axiom of choice it is easy to find an example of a nonmeasurable (in the Lebesgue sense) set.

1.7.7. Example. Let us declare two points x and y in $[0,1]$ equivalent if the number $x-y$ is rational. It is clear that the obtained relation is indeed an equivalence relation, i.e., 1) $x \sim x$, 2) $y \sim x$ if $x \sim y$, 3) $x \sim z$ if $x \sim y$ and $y \sim z$. Hence we obtain the equivalence classes each of which contains points with rational mutual differences, and the differences between any representatives of different classes are irrational. Let us now choose in every class exactly one representative and denote the constructed set by E. It is the axiom of choice that enables one to construct such a set. The set E cannot be Lebesgue measurable. Indeed, if its measure equals zero, then the measure of $[0,1]$ equals zero as well, since $[0,1]$ is covered by countably many translations of E by rational numbers. The measure of E cannot be positive, since for different rational p and q, the sets $E+p$ and $E+q$ are disjoint and have equal positive measures. One has $E+p \subset [0,2]$ if $p \in [0,1]$, hence the interval $[0,2]$ would have infinite measure.

However, one should have in mind that the axiom of choice may be replaced by a proposition (added to the standard set-theoretic axioms) that makes all subsets of the real line measurable. Some remarks about this are made in §1.12(x).

Note also that even if we use the axiom of choice, there still remains the question: does there exist *some* extension of Lebesgue measure to a countably additive measure on the class of all subsets of the interval? The above example only says that such an extension cannot be obtained by means of the Lebesgue completion. An answer to this question also depends on additional set-theoretic axioms (see §1.12(x)). In any case, the Lebesgue extension is not maximal: by Theorem 1.12.14, for every set $E \subset [0,1]$ that is not Lebesgue measurable, one can extend Lebesgue measure to a countably additive measure on the σ-algebra generated by all Lebesgue measurable sets in $[0,1]$ and the set E.

Closing our discussion of the properties of Lebesgue measure let us mention the Jordan (Peano–Jordan) measure.

1.7.8. Definition. *A bounded set E in $\mathbb{R}^n$ is called Jordan measurable if, for each $\varepsilon > 0$, there exist sets U_ε and V_ε that are finite unions of cubes such that $U_\varepsilon \subset E \subset V_\varepsilon$ and $\lambda_n(V_\varepsilon \backslash U_\varepsilon) < \varepsilon$.*

It is clear that when $\varepsilon \to 0$, there exists a common limit of the measures of U_ε and V_ε, called the Jordan measure of the set E. It is seen from the definition that every Jordan measurable set E is Lebesgue measurable and its Lebesgue measure coincides with its Jordan measure. However, the converse is false: for example, the set of rational numbers in the interval is not Jordan measurable. The collection of all Jordan measurable sets is a ring (see Exercise 1.12.77), on which the Jordan measure coincides with Lebesgue measure. Certainly, the Jordan measure is countably additive on its domain and its Lebesgue extension is Lebesgue measure. In Exercise 3.10.75 one can find a useful sufficient condition of the Jordan measurability.

1.8. Lebesgue–Stieltjes measures

Let μ be a nonnegative Borel measure on $\mathbb{R}^1$. Then the function

$$t \mapsto F(t) = \mu\bigl((-\infty, t)\bigr)$$

is bounded, nondecreasing (i.e., $F(t) \le F(s)$ whenever $t \le s$; such functions are also called increasing), left continuous, i.e., $F(t_n) \to F(t)$ as $t_n \uparrow t$, which follows by the countable additivity μ, and one has $\lim\limits_{t\to-\infty} F(t) = 0$. These conditions turn out also to be sufficient in order that the function F be generated by some measure according to the above formula. The function F is called the *distribution function* of the measure μ. Note that the distribution function is often defined by the formula $F(t) = \mu\bigl((-\infty, t]\bigr)$, which leads to different values at the points of positive μ-measure (the jumps of the function F are exactly the points of positive μ-measure).

1.8.1. Theorem. *Let F be a bounded, nondecreasing, left continuous function with $\lim\limits_{t\to-\infty} F(t) = 0$. Then, there exists a unique nonnegative Borel measure on $\mathbb{R}^1$ such that*

$$F(t) = \mu\bigl((-\infty, t)\bigr) \quad \textit{for all } t \in \mathbb{R}^1.$$

PROOF. It is known from the elementary calculus that the function F has an at most countable set D of points of discontinuity. Clearly, there is a countable set S in $\mathbb{R}^1 \backslash D$ that is everywhere dense in $\mathbb{R}^1$. Let us consider the class $\mathcal{A}$ of all sets of the form $A = \bigcup_{i=1}^n J_i$, where J_i is an interval of one of the following four types: (a, b), $[a, b]$, $(a, b]$ or $[a, b)$, where a and b either belong to S or coincide with $-\infty$ or $+\infty$. It is readily seen that $\mathcal{A}$ is an algebra. Let us define the set function μ on $\mathcal{A}$ as follows: if A is an interval with endpoints a and b, where $a \le b$, then $\mu(A) = F(b) - F(a)$, and if A is a finite union of disjoint intervals J_i, then $\mu(A) = \sum_i \mu(J_i)$. It is clear that the function μ is well-defined and additive. For the proof of countable additivity μ on $\mathcal{A}$, it suffices to observe that the class of finite unions of compact intervals is compact and is approximating. Indeed, if J is an open or semiopen interval, e.g., $J = (a, b)$, where a and b belong to S (or coincide with the points $+\infty$, $-\infty$), then, by the continuity of F at the points of S, we have

$F(b)-F(a) = \lim\limits_{i\to\infty}[F(b_i)-F(a_i)]$, where $a_i \downarrow a$, $b_i \uparrow b$, $a_i,\ b_i \in S$. If $a=-\infty$, then the same follows by the condition $\lim\limits_{t\to-\infty} F(t) = 0$. Let us extend μ to a countably additive measure on the Borel σ-algebra $\mathcal{B}(\mathbb{R}^1)$ (note that $\mathcal{B}(\mathbb{R}^1)$ is generated by the algebra $\mathcal{A}$, since S is dense). We have $F(t) = \mu\bigl((-\infty,t)\bigr)$ for all t (and not only for $t\in S$). This follows by the left continuity of both functions and their coincidence on a countable everywhere dense set. The uniqueness of μ is clear from the fact that the function F uniquely determines the values of μ on intervals.

We observe that due to Proposition 1.3.10, we could also use the semi-algebra of semiclosed intervals of the form $(-\infty,b)$, $[a,b)$, $[a,+\infty)$, where $a,b\in S$. $\square$

The measure μ constructed from the function F as described above is called the Lebesgue–Stieltjes measure with distribution function F. Similarly, by means of the distribution functions of n variables (representing measures of sets $(-\infty,x_1)\times\cdots\times(-\infty,x_n)$) one defines Lebesgue–Stieltjes measures on $\mathbb{R}^n$ (see Exercise 1.12.156).

1.9. Monotone and σ-additive classes of sets

In this section, we consider two more classes of sets that are frequently used in measure theory.

1.9.1. Definition. *A family $\mathcal{E}$ of subsets of a set X is called a monotone class if $\bigcup_{n=1}^\infty E_n \in \mathcal{E}$ for every increasing sequence of sets $E_n \in \mathcal{E}$ and $\bigcap_{n=1}^\infty E_n \in \mathcal{E}$ for every decreasing sequence of sets $E_n \in \mathcal{E}$.*

1.9.2. Definition. *A family $\mathcal{E}$ of subsets of a set X is called a σ-additive class if the following conditions are fulfilled*:

(i) *$X\in\mathcal{E}$,*

(ii) *$E_2\backslash E_1 \in \mathcal{E}$ provided that $E_1,\ E_2\in\mathcal{E}$ and $E_1\subset E_2$,*

(iii) *$\bigcup_{n=1}^\infty E_n \in \mathcal{E}$ provided that $E_n\in\mathcal{E}$ are pairwise disjoint.*

Note that in the presence of conditions (i) and (ii), condition (iii) can be restated as follows: $E_1\cup E_2\in\mathcal{E}$ for every disjoint pair $E_1,\ E_2\in\mathcal{E}$ and $\bigcup_{n=1}^\infty E_n\in\mathcal{E}$ whenever $E_n\in\mathcal{E}$ and $E_n\subset E_{n+1}$ for all $n\in\mathbb{N}$.

Given a class $\mathcal{E}$ of subsets of X, we have the smallest monotone class containing $\mathcal{E}$ (called the monotone class generated by $\mathcal{E}$), and the smallest σ-additive class containing $\mathcal{E}$ (called the σ-additive class generated by $\mathcal{E}$). These minimal classes are, respectively, the intersections of all monotone and all σ-additive classes containing $\mathcal{E}$.

The next result called the monotone class theorem is frequently used in measure theory.

1.9.3. Theorem. (i) *Let $\mathcal{A}$ be an algebra of sets. Then the σ-algebra generated by $\mathcal{A}$ coincides with the monotone class generated by $\mathcal{A}$.*

(ii) *If the class $\mathcal{E}$ is closed under finite intersections, then the σ-additive class generated by $\mathcal{E}$ coincides with the σ-algebra generated by $\mathcal{E}$.*

PROOF. (i) Denote by $\mathcal{M}(\mathcal{A})$ the monotone class generated by $\mathcal{A}$. Since $\sigma(\mathcal{A})$ is a monotone class, one has $\mathcal{M}(\mathcal{A}) \subset \sigma(\mathcal{A})$. Let us prove the inverse inclusion. To this end, let us show that $\mathcal{M}(\mathcal{A})$ is a σ-algebra. It suffices to prove that $\mathcal{M}(\mathcal{A})$ is an algebra. We show first that the class $\mathcal{M}(\mathcal{A})$ is closed with respect to complementation. Let

$$\mathcal{M}_0 = \{B\colon\ B, X\backslash B \in \mathcal{M}(\mathcal{A})\}.$$

The class $\mathcal{M}_0$ is monotone, which is obvious, since $\mathcal{M}(\mathcal{A})$ is a monotone class and one has the equalities

$$X\backslash \bigcap_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty}(X\backslash B_n), \quad X\backslash \bigcup_{n=1}^{\infty} B_n = \bigcap_{n=1}^{\infty}(X\backslash B_n).$$

Since $\mathcal{A} \subset \mathcal{M}_0 \subset \mathcal{M}(\mathcal{A})$, one has $\mathcal{M}_0 = \mathcal{M}(\mathcal{A})$.

Let us verify that $\mathcal{M}(\mathcal{A})$ is closed with respect to finite intersections. Let $A \in \mathcal{M}(\mathcal{A})$. Set

$$\mathcal{M}_A = \bigl\{B \in \mathcal{M}(\mathcal{A})\colon\ A \cap B \in \mathcal{M}(\mathcal{A})\bigr\}.$$

If $B_n \in \mathcal{M}_A$ are monotonically increasing sets, then

$$A \cap \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty}(A \cap B_n) \in \mathcal{M}(\mathcal{A}).$$

The case where the sets B_n are decreasing is similar. Hence $\mathcal{M}_A$ is a monotone class. If $A \in \mathcal{A}$, then we have $\mathcal{A} \subset \mathcal{M}_A \subset \mathcal{M}(\mathcal{A})$, whence we obtain that $\mathcal{M}_A = \mathcal{M}(\mathcal{A})$. Now let $A \in \mathcal{A}$ and $B \in \mathcal{M}(\mathcal{A})$. Then, according to the equality $\mathcal{M}(\mathcal{A}) = \mathcal{M}_A$, we have $A \cap B \in \mathcal{M}(\mathcal{A})$, which gives $A \in \mathcal{M}_B$. Thus, $\mathcal{A} \subset \mathcal{M}_B \subset \mathcal{M}(\mathcal{A})$. Therefore, $\mathcal{M}_B = \mathcal{M}(\mathcal{A})$ for all $B \in \mathcal{M}(\mathcal{A})$, which means that $\mathcal{M}(\mathcal{A})$ is closed with respect to finite intersections. It follows that $\mathcal{M}(\mathcal{A})$ is an algebra as required.

(ii) Denote by $\mathcal{S}$ the σ-additive class generated by $\mathcal{E}$. It is clear that $\mathcal{S} \subset \sigma(\mathcal{E})$, since $\sigma(\mathcal{E})$ is a σ-additive class. Let us show the inverse inclusion. To this end, we show that $\mathcal{S}$ is a σ-algebra. It suffices to verify that the class $\mathcal{S}$ is closed with respect to finite intersections. Set

$$\mathcal{S}_0 = \{A \in \mathcal{S}\colon\ A \cap E \in \mathcal{S} \ \text{ for all } E \in \mathcal{E}\}.$$

Note that $\mathcal{S}_0$ is a σ-additive class. Indeed, $X \in \mathcal{S}_0$. Let $A, B \in \mathcal{S}_0$ and $A \subset B$. Then, for any $E \in \mathcal{E}$, we have $(B\backslash A) \cap E = (B \cap E)\backslash(A \cap E) \in \mathcal{S}$, since the intersections $A \cap E, B \cap E$ belong to $\mathcal{S}$ and $\mathcal{S}$ is a σ-additive class. Similarly, it is verified that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{S}_0$ for any pairwise disjoint sets $A_n \in \mathcal{S}_0$. Since $\mathcal{E} \subset \mathcal{S}_0$, one has $\mathcal{S}_0 = \mathcal{S}$. Thus, $A \cap E \in \mathcal{S}$ for all $A \in \mathcal{S}$ and $E \in \mathcal{E}$. Now set

$$\mathcal{S}_1 = \{A \in \mathcal{S}\colon\ A \cap B \in \mathcal{S} \ \text{ for all } B \in \mathcal{S}\}.$$

Let us show that $\mathcal{S}_1$ is a σ-additive class. Indeed, $X \in \mathcal{S}_1$. If $A_1, A_2 \in \mathcal{S}_1$, $A_1 \subset A_2$, then $A_2\backslash A_1 \in \mathcal{S}_1$, since for all $B \in \mathcal{S}$, by the definition of $\mathcal{S}_1$, we obtain $(A_2\backslash A_1) \cap B = (A_2 \cap B)\backslash(A_1 \cap B) \in \mathcal{S}$. Similarly, it is verified that $\bigcup_{n=1}^{\infty} B_n \in \mathcal{S}_1$ for any sequence of disjoint sets in $\mathcal{S}_1$. Since $\mathcal{E} \subset \mathcal{S}_1$ as proved above, one has $\mathcal{S}_1 = \mathcal{S}$. Therefore, $A \cap B \in \mathcal{S}$ for all $A, B \in \mathcal{S}$. Thus, $\mathcal{S}$ is a σ-algebra. □

As an application of Theorem 1.9.3 we prove the following useful result.

1.9.4. Lemma. *If two probability measures μ and ν on a measurable space $(X, \mathcal{A})$ coincide on some class of sets $\mathcal{E} \subset \mathcal{A}$ that is closed with respect to finite intersections, then they coincide on the σ-algebra generated by $\mathcal{E}$.*

PROOF. Let $\mathcal{B} = \bigl\{A \in \mathcal{A}\colon\ \mu(A) = \nu(A)\bigr\}$. By hypothesis, $X \in \mathcal{B}$. If $A, B \in \mathcal{B}$ and $A \subset B$, then $B\backslash A \in \mathcal{B}$. In addition, if sets A_i in $\mathcal{B}$ are pairwise disjoint, then their union also belongs to $\mathcal{B}$. Hence $\mathcal{B}$ is a σ-additive class. Therefore, the σ-additive class $\mathcal{S}$ generated by $\mathcal{E}$ is contained in $\mathcal{B}$. By Theorem 1.9.3(ii) one has $\mathcal{S} = \sigma(\mathcal{E})$. Therefore, $\sigma(\mathcal{E}) \subset \mathcal{B}$. □

1.10. Souslin sets and the A-operation

Let B be a Borel set in the plane and let A be its projection to one of the axes. Is A a Borel set? One can hardly imagine that the correct answer to this question is negative. This answer was found due to efforts of several eminent mathematicians investigating the structure of Borel sets. A result of those investigations was the creation of descriptive set theory, in particular, the invention of the A-operation. It was discovered that the continuous images of the Borel sets coincide with the result of application of the A-operation to the closed sets. This section is an introduction to the theory of Souslin sets discussed in greater detail in Chapter 6. In spite of an introductory and relatively elementary character of this section, it contains complete proofs of two deep facts of measure theory: the measurability of Souslin sets and, as a consequence, the measurability of sets that are images of Borel sets under continuous mappings.

Denote by $\mathbb{N}^{\infty}$ the set of all infinite sequences (n_i) with natural components.

1.10.1. Definition. *Let X be a nonempty set and let $\mathcal{E}$ be some class of its subsets. We say that we are given a Souslin scheme* (*or a table of sets*) *$\{A_{n_1,\ldots,n_k}\}$ with values in $\mathcal{E}$ if, to every finite sequence $(n_1, \ldots, n_k)$ of natural numbers, there corresponds a set $A_{n_1,\ldots,n_k} \in \mathcal{E}$. The A-operation* (*or the Souslin operation*) *over the class $\mathcal{E}$ is a mapping that to every Souslin scheme $\{A_{n_1,\ldots,n_k}\}$ with values in $\mathcal{E}$ associates the set*

$$A = \bigcup_{(n_i)\in\mathbb{N}^{\infty}} \bigcap_{k=1}^{\infty} A_{n_1,\ldots,n_k}. \tag{1.10.1}$$

The sets of this form are called $\mathcal{E}$-Souslin or $\mathcal{E}$-analytic. The collection of all such sets along with the empty set is denoted by $S(\mathcal{E})$.

Certainly, if $\varnothing \in \mathcal{E}$ (or if $\mathcal{E}$ contains disjoint sets), then $\varnothing \in S(\mathcal{E})$ automatically.

1.10.2. Example. By means of the A-operation one can obtain any countable unions and countable intersections of elements in the class $\mathcal{E}$.

PROOF. In the first case, it suffices to take $A_{n_1,\ldots,n_k} = A_{n_1}$, and in the second, $A_{n_1,\ldots,n_k} = A_k$. □

A Souslin scheme is called *monotone* (or regular) if

$$A_{n_1,\ldots,n_k,n_{k+1}} \subset A_{n_1,\ldots,n_k}.$$

If the class $\mathcal{E}$ is closed under finite intersections, then any Souslin scheme with values in $\mathcal{E}$ can be replaced by a monotone one giving the same result of the A-operation. Indeed, set

$$A^*_{n_1,\ldots,n_k} = A_{n_1} \cap A_{n_1,n_2} \cap \cdots \cap A_{n_1,\ldots,n_k}.$$

We need the following technical assertion. Let $\bigl(\mathbb{N}^\infty\bigr)^\infty$ denote the space of all sequences $\eta = (\eta^1, \eta^2, \ldots)$ with $\eta^i \in \mathbb{N}^\infty$.

1.10.3. Lemma. *There exist bijections*

$$\beta\colon\ \mathbb{N}\times\mathbb{N} \to \mathbb{N} \quad \textit{and} \quad \Psi\colon\ \mathbb{N}^\infty \times \bigl(\mathbb{N}^\infty\bigr)^\infty \to \mathbb{N}^\infty$$

with the property: for all $m, n \in \mathbb{N}$, $\sigma = (\sigma_i) \in \mathbb{N}^\infty$ and $(\tau^i) \in \bigl(\mathbb{N}^\infty\bigr)^\infty$, where $\tau^i = (\tau^i_j) \in \mathbb{N}^\infty$, the collections $\sigma_1, \ldots, \sigma_m$ and $\tau^m_1, \ldots, \tau^m_n$ are uniquely determined by the first $\beta(m,n)$ components of the element $\Psi\bigl(\sigma, (\tau^i)\bigr)$.

PROOF. Set $\beta(m,n) = 2^{m-1}(2n-1)$. It is clear that β is a bijection of $\mathbb{N}\times\mathbb{N}$ onto $\mathbb{N}$, since, for any $l \in \mathbb{N}$, there exists a unique pair of natural numbers (m,n) with $l = 2^{m-1}(2n-1)$. Set also $\varphi(l) := m$, $\psi(l) := n$, where $\beta(m,n) = l$. Let $\sigma = (\sigma_i) \in \mathbb{N}^\infty$ and $(\tau^i) \in \bigl(\mathbb{N}^\infty\bigr)^\infty$, where $\tau^i = (\tau^i_j) \in \mathbb{N}^\infty$. Finally, set

$$\Psi\bigl(\sigma, (\tau^i)\bigr) = \Bigl(\beta\bigl(\sigma_1, \tau^{\varphi(1)}_{\psi(1)}\bigr), \ldots, \beta\bigl(\sigma_l, \tau^{\varphi(l)}_{\psi(l)}\bigr), \ldots\Bigr).$$

For every $\eta = (\eta_i) \in \mathbb{N}^\infty$, the equation $\Psi\bigl(\sigma, (\tau^i)\bigr) = \eta$ has a unique solution $\sigma_i = \varphi(\eta_i)$, $\tau^i_j = \psi(\eta_{\beta(i,j)})$. Hence Ψ is bijective. Since $m \le \beta(m,n)$ and $\beta(m,k) \le \beta(m,n)$ whenever $k \le n$, it follows from the form of the solution that the first $\beta(m,n)$ components of $\Psi\bigl(\sigma, (\tau^i)\bigr)$ uniquely determine the first m components of σ and the first n components of τ^m. □

The next theorem describes a number of important properties of Souslin sets.

1.10.4. Theorem. (i) *One has $S\bigl(S(\mathcal{E})\bigr) = S(\mathcal{E})$. In particular, the class $S(\mathcal{E})$ is closed under countable unions and countable intersections.*

(ii) *If the complement of every set in $\mathcal{E}$ belongs to $S(\mathcal{E})$ (for example, is an at most countable union of elements of $\mathcal{E}$) and $\varnothing \in \mathcal{E}$, then the σ-algebra $\sigma(\mathcal{E})$ generated by $\mathcal{E}$ is contained in the class $S(\mathcal{E})$.*

PROOF. (i) Let $A_{n_1,\ldots,n_k}^{\nu_1,\ldots,\nu_m} \in \mathcal{E}$ and let

$$A = \bigcup_{(n_i)\in\mathbb{N}^\infty} \bigcap_{k=1}^{\infty} A_{n_1,\ldots,n_k}, \quad A_{n_1,\ldots,n_k} = \bigcup_{\nu\in\mathbb{N}^\infty} \bigcap_{m=1}^{\infty} A_{n_1,\ldots,n_k}^{\nu_1,\ldots,\nu_m}.$$

Keeping the notation of the above lemma, for any natural numbers $\eta_1, \ldots, \eta_l$ we find $\sigma \in \mathbb{N}^\infty$ and $\tau = (\tau^m) \in (\mathbb{N}^\infty)^\infty$ such that $\eta_1 = \Psi(\sigma,\tau)_1, \ldots, \eta_l = \Psi(\sigma,\tau)_l$. Certainly, σ and τ are not uniquely determined, but according to the lemma, the collections $\sigma_1,\ldots,\sigma_{\varphi(l)}$ and $\tau_1^{\varphi(l)},\ldots,\tau_{\psi(l)}^{\varphi(l)}$ are uniquely determined by the numbers $\eta_1, \ldots, \eta_l$. Hence we may set

$$B(\eta_1, \ldots, \eta_l) = A_{\sigma_1,\ldots,\sigma_{\varphi(l)}}^{\tau_1^{\varphi(l)},\ldots,\tau_{\psi(l)}^{\varphi(l)}} \in \mathcal{E}.$$

Then, denoting by $\eta = (\eta_l)$ and $\sigma = (\sigma_m)$ elements of $\mathbb{N}^\infty$ and by (τ^m) with $\tau^m = (\tau_n^m)$ elements of $(\mathbb{N}^\infty)^\infty$, we have

$$\begin{aligned}
\bigcup_{\eta} \bigcap_{l=1}^{\infty} B(\eta_1, \ldots, \eta_l) &= \bigcup_{\sigma,(\tau^m)} \bigcap_{l=1}^{\infty} B\Big(\Psi\big(\sigma,(\tau^m)\big)_1, \ldots, \Psi\big(\sigma,(\tau^m)\big)_l\Big) \\
&= \bigcup_{\sigma,(\tau^m)} \bigcap_{l=1}^{\infty} A_{\sigma_1,\ldots,\sigma_{\varphi(l)}}^{\tau_1^{\varphi(l)},\ldots,\tau_{\psi(l)}^{\varphi(l)}} = \bigcup_{\sigma,(\tau^m)} \bigcap_{m,n=1}^{\infty} A_{\sigma_1,\ldots,\sigma_m}^{\tau_1^m,\ldots,\tau_n^m} \\
&= \bigcup_{\sigma} \bigcup_{(\tau^m)} \bigcap_{m=1}^{\infty} \bigcap_{n=1}^{\infty} A_{\sigma_1,\ldots,\sigma_m}^{\tau_1^m,\ldots,\tau_n^m} = \bigcup_{\sigma} \bigcap_{m=1}^{\infty} \bigcup_{\tau^m} \bigcap_{n=1}^{\infty} A_{\sigma_1,\ldots,\sigma_m}^{\tau_1^m,\ldots,\tau_n^m} \\
&= \bigcup_{\sigma} \bigcap_{m=1}^{\infty} A_{\sigma_1,\ldots,\sigma_m} = A.
\end{aligned}$$

Thus, $S\big(S(\mathcal{E})\big) \subset S(\mathcal{E})$. The inverse inclusion is obvious.

(ii) Set

$$\mathcal{F} = \big\{B \in S(\mathcal{E}) \colon\ X\backslash B \in S(\mathcal{E})\big\}.$$

Let us show that $\mathcal{F}$ is a σ-algebra. By construction, $\mathcal{F}$ is closed under complementation. Let $B_n \in \mathcal{F}$. Then $\bigcap_{n=1}^\infty B_n \in S(\mathcal{E})$ according to assertion (i). Similarly, $X\backslash\bigcap_{n=1}^\infty B_n = \bigcup_{n=1}^\infty (X\backslash B_n) \in S(\mathcal{E})$. By hypothesis, $\varnothing \in \mathcal{F}$. Therefore, $\mathcal{F}$ is a σ-algebra. Since by hypothesis $\mathcal{E} \subset \mathcal{F}$, we obtain $\sigma(\mathcal{E}) \subset \mathcal{F} \subset S(\mathcal{E})$. □

It is clear that the condition $X\backslash E \in S(\mathcal{E})$ for $E \in \mathcal{E}$ is also necessary in order that $\sigma(\mathcal{E}) \subset S(\mathcal{E})$. The class $S(\mathcal{E})$ may not be closed with respect to complementation even in the case where $\mathcal{E}$ is a σ-algebra. As we shall see later, this happens, for example, with $\mathcal{E} = \mathcal{B}(\mathbb{R}^1)$. If we apply the A-operation to the class of all compact (or closed) sets in $\mathbb{R}^n$, then the hypothesis in assertion (ii) of the above theorem is satisfied, since every nonempty open set in $\mathbb{R}^n$ is a countable union of closed cubes. Below we consider this example more carefully.

The next fundamental result shows that the A-operation preserves measurability. This assertion is not at all obvious and, moreover, it is very surprising, since the A-operation involves uncountable unions.

1.10.5. Theorem. *Suppose that μ is a finite nonnegative measure on a σ-algebra $\mathcal{M}$. Then, the class $\mathcal{M}_\mu$ of all μ-measurable sets is closed with respect to the A-operation. Moreover, given a family of sets $\mathcal{E} \subset \mathcal{M}$ that is closed with respect to finite unions and countable intersections, one has*

$$\mu^*(A) = \sup\bigl\{\mu(E)\colon\ E \subset A,\ E \in \mathcal{E}\bigr\}$$

for every $\mathcal{E}$-Souslin set A. In particular, every $\mathcal{E}$-Souslin set is μ-measurable.

PROOF. The first claim is a simple corollary of the second one applied to the family $\mathcal{E} = \mathcal{M}_\mu$. So we prove the second claim. Let a set A be constructed by means of a monotone table of sets $E_{n_1,\ldots,n_k} \in \mathcal{E}$. Let $\varepsilon > 0$. For every collection $m_1, \ldots, m_k$ of natural numbers, denote by $D_{m_1,\ldots,m_k}$ the union of the sets $E_{n_1,\ldots,n_k}$ over all $n_1 \le m_1$,…,$n_k \le m_k$. Let

$$M_{m_1,\ldots,m_k} := \bigcup_{(n_i)\in\mathbb{N}^\infty,\ n_1\le m_1,\ldots,n_k\le m_k} \ \bigcap_{j=1}^{\infty} E_{n_1,\ldots,n_j}.$$

It is clear that as $m \to \infty$, the sets M_m monotonically increase to A, and the sets $M_{m_1,\ldots,m_k,m}$ with fixed $m_1, \ldots, m_k$ monotonically increase to $M_{m_1,\ldots,m_k}$. By Proposition 1.5.12, there is a number m_1 with $\mu^*(M_{m_1}) > \mu^*(A) - \varepsilon 2^{-1}$. Then we can find a number m_2 with $\mu^*(M_{m_1,m_2}) > \mu^*(M_{m_1}) - \varepsilon 2^{-2}$. Continuing this construction by induction, we obtain a sequence of natural numbers m_k such that

$$\mu^*(M_{m_1,m_2,\ldots,m_k}) > \mu^*(M_{m_1,m_2,\ldots,m_{k-1}}) - \varepsilon 2^{-k}.$$

Therefore, for all k one has

$$\mu^*(M_{m_1,m_2,\ldots,m_k}) > \mu^*(A) - \varepsilon.$$

By the stability of $\mathcal{E}$ with respect to finite unions we have $D_{m_1,\ldots,m_k} \in \mathcal{E}$, and the stability of $\mathcal{E}$ with respect to countable intersections yields the inclusion $E := \bigcap_{k=1}^\infty D_{m_1,\ldots,m_k} \in \mathcal{E}$. Since $M_{m_1,\ldots,m_k} \subset D_{m_1,\ldots,m_k}$, we obtain by the previous estimate $\mu^*(D_{m_1,m_2,\ldots,m_k}) > \mu^*(A) - \varepsilon$, whence it follows that $\mu(E) \ge \mu^*(A) - \varepsilon$, since the sets $D_{m_1,m_2,\ldots,m_k}$ decrease to E.

It remains to prove that $E \subset A$. Let $x \in E$. Then, for all k we have $x \in D_{m_1,\ldots,m_k}$. Hence $x \in E_{n_1,\ldots,n_k}$ for some collection $n_1, \ldots, n_k$ such that $n_1 \le m_1$,…,$n_k \le m_k$. Such collections will be called admissible. Our task is to construct an infinite sequence $n_1, n_2, \ldots$ such that all its initial intervals $n_1, \ldots, n_k$ are admissible. In this case $x \in \bigcap_{k=1}^\infty E_{n_1,\ldots,n_k} \subset A$. In order to construct such a sequence let us observe that, for any $k > 1$, we have admissible collections of k numbers. An admissible collection $n_1, \ldots, n_k$ is called extendible if, for every $l \ge k$, there exists an admissible collection $p_1, \ldots, p_l$ with $p_1 = n_1, \ldots, p_k = n_k$. Let us now observe that there exists at least one extendible collection n_1 of length 1. Indeed, suppose the contrary. Since every initial interval $n_1, \ldots, n_k$ in any admissible collection $n_1, \ldots, n_k, \ldots, n_l$ is admissible by the inclusion $E_{n_1,\ldots,n_l} \subset E_{n_1,\ldots,n_k}$, we obtain that for every $n \leq m_1$ there exists the maximal length $l(n)$ of admissible collections with the number n at the first position. Therefore, the lengths of all admissible collections are uniformly bounded and we arrive at a contradiction. Similarly, the extendible collection n_1 is contained in some extendible collection n_1, n_2 and so on. The obtained sequence possesses the desired property. □

1.10.6. Corollary. *If $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ are measurable spaces and a mapping $f\colon X \to Y$ be such that $f^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}$, then for every set $E \in S(\mathcal{B})$, the set $f^{-1}(E)$ belongs to $S(\mathcal{A})$ and hence is measurable with respect to every measure on $\mathcal{A}$.*

PROOF. It follows from (1.10.1) that $f^{-1}(E) \in S(\mathcal{A})$. □

Another method of proof of Theorem 1.10.5 is described in Exercise 6.10.60 in Chapter 6. A thorough study of Souslin sets and related problems in measure theory is accomplished in Chapters 6 and 7. However, even now we are able to derive from Theorem 1.10.5 very useful corollaries.

1.10.7. Definition. *The sets obtained by application of the A-operation to the class of closed sets in $\mathbb{R}^n$ are called the Souslin sets in the space $\mathbb{R}^n$.*

It is clear that the same result is obtained by applying the A-operation to the class of all compact sets in $\mathbb{R}^n$. Indeed, if A is contained in a cube K, then closed sets $A_{\nu_1,\ldots,\nu_k}$ that generate A can be replaced by the compacts $A_{\nu_1,\ldots,\nu_k} \cap K$. Any unbounded Souslin set A can be written as the union of its intersections $A \cap K_j$ with increasing cubes K_j. It remains to use that the class of sets constructed by the A-operation from compact sets admits countable unions.

As was mentioned above, it follows by Theorem 1.10.4 that Borel sets in $\mathbb{R}^n$ are Souslin. Note also that if L is a linear subspace in $\mathbb{R}^n$ of dimension $k < n$, then the intersection of L with any Souslin set A in $\mathbb{R}^n$ is Souslin in the space L. This follows by the fact that the intersection of any closed set with L is closed in L. Conversely, any Souslin set in L is Souslin in $\mathbb{R}^n$ as well.

1.10.8. Proposition. *The image of any Souslin set under a continuous mapping from $\mathbb{R}^n$ to $\mathbb{R}^d$ is Souslin.*

PROOF. Let a set A have the form (1.10.1), where the sets $A_{n_1,\ldots,n_k}$ are compact (as we know, such a representation is possible for every Souslin set). As noted above, we may assume that $A_{n_1,\ldots,n_k,n_{k+1}} \subset A_{n_1,\ldots,n_k}$ for all k. Let $f\colon \mathbb{R}^n \to \mathbb{R}^d$ be a continuous mapping. It is clear that

$$f(A) = \bigcup_{(n_i) \in \mathbb{N}^\infty} f\Bigl(\bigcap_{k=1}^{\infty} A_{n_1,\ldots,n_k}\Bigr).$$

It remains to observe that the sets $B_{n_1,\dots,n_k} = f(A_{n_1,\dots,n_k})$ are compact by the continuity of f and that

$$f\Bigl(\bigcap_{k=1}^{\infty} A_{n_1,\dots,n_k}\Bigr) = \bigcap_{k=1}^{\infty} f(A_{n_1,\dots,n_k}).$$

Indeed, the left-hand side of this equality is contained in the right-hand side for any sets and mappings. Let $y \in \bigcap_{k=1}^{\infty} f(A_{n_1,\dots,n_k})$. Then, for every k, there exists $x_k \in A_{n_1,\dots,n_k}$ with $f(x_k) = y$. If for infinitely many indices k the points x_k coincide with one and the same point x, then $x \in \bigcap_{k=1}^{\infty} A_{n_1,\dots,n_k}$ by the monotonicity of $A_{n_1,\dots,n_k}$. Clearly, $f(x) = y$. Hence it remains to consider the case where the sequence $\{x_k\}$ contains infinitely many distinct points. Since this sequence is contained in the compact set A_{n_1}, there exists a limit point x of $\{x_k\}$. Then $x \in A_{n_1,\dots,n_k}$ for all k, since $x_j \in A_{n_1,\dots,n_k}$ for all $j \geq k$ and $A_{n_1,\dots,n_k}$ is a closed set. Thus, $x \in \bigcap_{k=1}^{\infty} A_{n_1,\dots,n_k}$. By the continuity of f we obtain $f(x) = y$. □

1.10.9. Corollary. *The image of any Borel set $B \subset \mathbb{R}^n$ under a continuous mapping $f\colon \mathbb{R}^n \to \mathbb{R}^d$ is a Souslin set. In particular, the set $f(B)$ is Lebesgue measurable.*

In particular, the orthogonal projection of a Borel set is Souslin, hence measurable. We shall see in Chapter 6 that Souslin sets in $\mathbb{R}^n$ coincide with the orthogonal projections of Borel sets in $\mathbb{R}^{n+1}$ (thus, Souslin sets can be defined without the A-operation) and that there exist non-Borel Souslin sets. It is easily verified that the product of two Borel sets in $\mathbb{R}^n$ is Borel in $\mathbb{R}^{2n}$. Indeed, it suffices to check that $A\times\mathbb{R}^n \in \mathcal{B}(\mathbb{R}^{2n})$ if $A \in \mathcal{B}(\mathbb{R}^n)$. This is true for any open set A, hence for any Borel set A, since the class of all Borel sets A with such a property is obviously a σ-algebra.

1.10.10. Example. Let A and B be nonempty Borel sets in $\mathbb{R}^n$. Then the vector sum of the sets A and B defined by the equality

$$A + B := \{a + b\colon\ a \in A, b \in B\}$$

is a Souslin set. In addition, the convex hull $\operatorname{conv} A$ of the set A, i.e., the smallest convex set containing A, is Souslin as well. Indeed, $A+B$ is the image of the Borel set $A\times B$ in $\mathbb{R}^{2n}$ under the continuous mapping $(x, y) \mapsto x + y$. The convex hull of A consists of all sums of the form

$$\sum_{i=1}^{k} t_i a_i, \text{ where } t_i \geq 0,\ \textstyle\sum_{i=1}^{k} t_i = 1,\ a_i \in A,\ k \in \mathbb{N}.$$

For every fixed k, the set S of all points $(t_1, \dots, t_k) \in \mathbb{R}^k$ such that $\sum_{i=1}^{k} t_i = 1$ and $t_i \geq 0$ is Borel. Hence the set $A^k\times S$ in $(\mathbb{R}^n)^k\times\mathbb{R}^k$ is Borel as well and its image under the mapping $(a_1, \dots, a_k, t_1, \dots, t_k) \mapsto \sum_{i=1}^{k} t_i a_i$ is Souslin.

1.11. Carathéodory outer measures

In this section, we discuss in greater detail constructions of measures by means of the so-called Carathéodory outer measures. We have already encountered the principal idea in the consideration of extensions of countably additive measures from an algebra to a σ-algebra, but now we do not assume that an "outer measure" is generated by an additive measure.

1.11.1. Definition. *A set function* $\mathfrak{m}$ *defined on the class of all subsets of a set* X *and taking values in* $[0, +\infty]$ *is called an outer measure on* X (*or a Carathéodory outer measure*) *if*:
(i) $\mathfrak{m}(\varnothing) = 0$;
(ii) $\mathfrak{m}(A) \le \mathfrak{m}(B)$ *whenever* $A \subset B$, *i.e.*, $\mathfrak{m}$ *is monotone*;
(iii) $\mathfrak{m}\Bigl(\bigcup_{n=1}^\infty A_n\Bigr) \le \sum_{n=1}^\infty \mathfrak{m}(A_n)$ *for all* $A_n \subset X$.

An important example of a Carathéodory outer measure is the function μ^* discussed in §1.5.

1.11.2. Definition. *Let* $\mathfrak{m}$ *be a set function with values in* $[0, +\infty]$ *defined on the class of all subsets of a space* X *such that* $\mathfrak{m}(\varnothing) = 0$. *A set* $A \subset X$ *is called Carathéodory measurable with respect to* $\mathfrak{m}$ (*or Carathéodory* $\mathfrak{m}$*-measurable*) *if, for every set* $E \subset X$, *one has the equality*

$$\mathfrak{m}(E \cap A) + \mathfrak{m}(E \backslash A) = \mathfrak{m}(E). \tag{1.11.1}$$

The class of all Carathéodory $\mathfrak{m}$*-measurable sets is denoted by* $\mathfrak{M}_{\mathfrak{m}}$.

Thus, a measurable set splits every set according to the requirement of additivity of $\mathfrak{m}$ (see also Exercise 1.12.150 in this relation).

Let us note at once that in general the measurability does not follow from the equality

$$\mathfrak{m}(A) + \mathfrak{m}(X \backslash A) = \mathfrak{m}(X) \tag{1.11.2}$$

even in the case of an outer measure with $\mathfrak{m}(X) < \infty$. Let us consider the following example.

1.11.3. Example. Let $X = \{1, 2, 3\}$, $\mathfrak{m}(\varnothing) = 0$, $\mathfrak{m}(X) = 2$, and let $\mathfrak{m}(A) = 1$ for all other sets A. It is readily verified that $\mathfrak{m}$ is an outer measure. Here every subset $A \subset X$ satisfies (1.11.2), but for $A = \{1\}$ and $E = \{1, 2\}$ equality (1.11.1) does not hold (its left-hand side equals 2 and the right-hand side equals 1). It is easy to see that only two sets $\varnothing$ and X are $\mathfrak{m}$-measurable.

In this example the class $\mathfrak{M}_{\mathfrak{m}}$ of all Carathéodory $\mathfrak{m}$-measurable sets is smaller than the class $\mathcal{A}_{\mathfrak{m}}$ from Definition 1.5.1, since for the outer measure $\mathfrak{m}$ on the class of all sets the family $\mathcal{A}_{\mathfrak{m}}$ is the class of all sets. However, we shall see later that in the case where $\mathfrak{m} = \mu^*$ is the outer measure generated by a countably additive measure μ with values in $[0, +\infty]$ defined on a σ-algebra, the class $\mathfrak{M}_{\mathfrak{m}}$ may be larger than $\mathcal{A}_\mu$ (Exercise 1.12.129). On the other hand, under reasonable assumptions, the classes $\mathfrak{M}_{\mu^*}$ and $\mathcal{A}_\mu$ coincide.

Below a class of outer measures is singled out such that the corresponding measurability is equivalent to (1.11.2). This class embraces all outer measures generated by countably additive measures on algebras (see Proposition 1.11.7 and Theorem 1.11.8).

1.11.4. Theorem. *Let* $\mathfrak{m}$ *be a set function with values in* $[0,+\infty]$ *on the class of all sets in a space* X *such that* $\mathfrak{m}(\varnothing)=0$. *Then*:

(i) $\mathfrak{M}_{\mathfrak{m}}$ *is an algebra and the function* $\mathfrak{m}$ *is additive on* $\mathfrak{M}_{\mathfrak{m}}$.

(ii) *For every sequence of pairwise disjoint sets* $A_i \in \mathfrak{M}_{\mathfrak{m}}$ *one has*

$$\mathfrak{m}\Bigl(E\cap\bigcup_{i=1}^{n}A_i\Bigr)=\sum_{i=1}^{n}\mathfrak{m}(E\cap A_i),\quad \forall\, E\subset X,$$

$$\mathfrak{m}\Bigl(E\cap\bigcup_{i=1}^{\infty}A_i\Bigr)=\sum_{i=1}^{\infty}\mathfrak{m}(E\cap A_i)+\lim_{n\to\infty}\mathfrak{m}\Bigl(E\cap\bigcup_{i=n}^{\infty}A_i\Bigr),\quad \forall\, E\subset X.$$

(iii) *If the function* $\mathfrak{m}$ *is an outer measure on the set* X, *then the class* $\mathfrak{M}_{\mathfrak{m}}$ *is a* σ*-algebra and the function* $\mathfrak{m}$ *with values in* $[0,+\infty]$ *is countably additive on* $\mathfrak{M}_{\mathfrak{m}}$. *In addition, the measure* $\mathfrak{m}$ *is complete on* $\mathfrak{M}_{\mathfrak{m}}$.

Proof. (i) It is obvious from (1.11.1) that $\varnothing\in\mathfrak{M}_{\mathfrak{m}}$ and that the class $\mathfrak{M}_{\mathfrak{m}}$ is closed with respect to complementation. Suppose that sets A_1, A_2 belong to $\mathfrak{M}_{\mathfrak{m}}$ and let $E\subset X$. By the measurability of A_1 and A_2 we have

$$\begin{aligned}\mathfrak{m}(E)&=\mathfrak{m}(E\cap A_1)+\mathfrak{m}(E\backslash A_1)\\&=\mathfrak{m}(E\cap A_1)+\mathfrak{m}\bigl((E\backslash A_1)\cap A_2\bigr)+\mathfrak{m}\bigl((E\backslash A_1)\backslash A_2\bigr)\\&=\mathfrak{m}(E\cap A_1)+\mathfrak{m}\bigl((E\backslash A_1)\cap A_2\bigr)+\mathfrak{m}\bigl(E\backslash (A_1\cup A_2)\bigr).\end{aligned}$$

According to the equality $E\cap A_1=E\cap(A_1\cup A_2)\cap A_1$ and the measurability of A_1 one has

$$\mathfrak{m}\bigl(E\cap(A_1\cup A_2)\bigr)=\mathfrak{m}(E\cap A_1)+\mathfrak{m}\bigl((E\backslash A_1)\cap A_2\bigr).\tag{1.11.3}$$

Hence we obtain

$$\mathfrak{m}(E)=\mathfrak{m}\bigl(E\cap(A_1\cup A_2)\bigr)+\mathfrak{m}\bigl(E\backslash(A_1\cup A_2)\bigr).$$

Thus, $A_1\cup A_2\in\mathfrak{M}_{\mathfrak{m}}$, i.e., $\mathfrak{M}_{\mathfrak{m}}$ is an algebra. For disjoint sets A_1 and A_2 by taking $E=X$ in (1.11.3) we obtain the equality $\mathfrak{m}(A_1\cup A_2)=\mathfrak{m}(A_1)+\mathfrak{m}(A_2)$.

(ii) Let $A_i\in\mathfrak{M}_{\mathfrak{m}}$ be disjoint. Set

$$S_n=\bigcup_{i=1}^{n}A_i,\quad R_n=\bigcup_{i=n}^{\infty}A_i.$$

Then by equality (1.11.3) we have

$$\mathfrak{m}(E\cap S_n)=\mathfrak{m}(E\cap A_n)+\mathfrak{m}(E\cap S_{n-1}).$$

By induction this yields the first equality in assertion (ii). Next, by the equalities $R_1\cap S_{n-1}=S_{n-1}$ and $R_1\backslash S_{n-1}=R_n$ one has

$$\mathfrak{m}(E\cap R_1)=\mathfrak{m}(E\cap S_{n-1})+\mathfrak{m}(E\cap R_n)=\sum_{i=1}^{n-1}\mathfrak{m}(E\cap A_i)+\mathfrak{m}(E\cap R_n).$$

This gives the second equality in assertion (ii), since the sequence $\mathfrak{m}(E \cap R_n)$ is decreasing by the equality

$$\mathfrak{m}(E \cap R_n) = \mathfrak{m}(E \cap R_{n+1}) + \mathfrak{m}(E \cap A_n),$$

which follows from the measurability of A_n and the relations $R_n \backslash A_n = R_{n+1}$ and $R_n \cap A_n = A_n$.

(iii) Suppose now that $\mathfrak{m}$ is countably subadditive and that sets $A_i \in \mathfrak{M}_\mathfrak{m}$ are disjoint. Let $A = \bigcup_{i=1}^\infty A_i$. The second equality in (ii) yields that for any $E \subset X$ one has $\mathfrak{m}(E \cap A) \geq \sum_{i=1}^\infty \mathfrak{m}(E \cap A_i)$, which by the countable subadditivity gives

$$\mathfrak{m}(E \cap A) = \sum_{i=1}^\infty \mathfrak{m}(E \cap A_i). \tag{1.11.4}$$

We already know that $S_n = A_1 \cup \cdots \cup A_n \in \mathfrak{M}_\mathfrak{m}$. It follows by the first equality in assertion (ii) that

$$\mathfrak{m}(E) = \mathfrak{m}(E \cap S_n) + \mathfrak{m}(E \backslash S_n) \geq \sum_{i=1}^n \mathfrak{m}(E \cap A_i) + \mathfrak{m}(E \backslash A).$$

By (1.11.4) we obtain $\mathfrak{m}(E) \geq \mathfrak{m}(E \cap A) + \mathfrak{m}(E \backslash A)$. By subadditivity the reverse inequality is true as well, i.e., $A \in \mathfrak{M}_\mathfrak{m}$. Hence $\mathfrak{M}_\mathfrak{m}$ is an algebra closed with respect to countable unions of disjoint sets. This means that $\mathfrak{M}_\mathfrak{m}$ is a σ-algebra. By taking $E = X$ in (1.11.4) we obtain the countable additivity of $\mathfrak{m}$ on $\mathfrak{M}_\mathfrak{m}$. We verify that $\mathfrak{m}$ is complete on $\mathfrak{M}_\mathfrak{m}$. Let $\mathfrak{m}(A) = 0$. Then, for any set E, we have $\mathfrak{m}(E \cap A) + \mathfrak{m}(E \backslash A) = \mathfrak{m}(E)$, as $0 \leq \mathfrak{m}(E \cap A) \leq \mathfrak{m}(A) = 0$, and $\mathfrak{m}(E \backslash A) = \mathfrak{m}(E)$, as $\mathfrak{m}(E \backslash A) \leq \mathfrak{m}(E) \leq \mathfrak{m}(E \backslash A) + \mathfrak{m}(A) = \mathfrak{m}(E \backslash A)$. □

Note that the countably additive measure $\mu := \mathfrak{m}|_{\mathfrak{M}_\mathfrak{m}}$ on $\mathfrak{M}_\mathfrak{m}$, where $\mathfrak{m}$ is an outer measure, gives rise to a usual outer measure μ^* as we did before. However, this outer measure may differ from the original function $\mathfrak{m}$ (certainly, on the sets in $\mathfrak{M}_\mathfrak{m}$ both outer measures coincide). Say, in Example 1.11.3 we obtain $\mu^*(A) = 2$ for any nonempty set A different from X. Some additional information is given in Exercises 1.12.125 and 1.12.126.

In applications, outer measures are often constructed by the so-called Method I described in the following example and already employed in §1.5, where in Lemma 1.5.4 the countable subadditivity has been established.

1.11.5. Example. Let $\mathfrak{X}$ be a family of subsets of a X such that $\varnothing \in \mathfrak{X}$. Suppose that we are given a function $\tau\colon\ \mathfrak{X} \to [0, +\infty]$ with $\tau(\varnothing) = 0$. Set

$$\mathfrak{m}(A) = \inf\Bigl\{\sum_{n=1}^\infty \tau(X_n)\colon\ X_n \in \mathfrak{X}, A \subset \bigcup_{n=1}^\infty X_n\Bigr\}, \tag{1.11.5}$$

where in the case of absence of such sets X_n we set $\mathfrak{m}(A) := \infty$. Then $\mathfrak{m}$ is an outer measure. It is denoted by τ^*.

This construction will be used in §3.10(iii) for defining the so-called Hausdorff measures. Exercise 1.12.130 describes a modification of the construction of $\mathfrak{m}$ that differs as follows: if there are no sequences of sets in $\mathfrak{X}$ covering A, then the value $\mathfrak{m}(A)$ is defined as $\sup \mathfrak{m}(A')$ over those $A' \subset A$ for which such sequences exist.

It should be emphasized that it is not claimed in the above example that the constructed outer measure extends τ. In general, this may be false. In addition, sets in the original family $\mathfrak{X}$ may be nonmeasurable with respect to $\mathfrak{m}$. Let us consider the corresponding counter-examples. Let us take for X the set $\mathbb{N}$ and for $\mathfrak{X}$ the family of all singletons and the whole set X. Let $\tau(n) = 2^{-n}$, $\tau(X) = 2$. Then $\mathfrak{m}(X) = 1$ and X is measurable with respect to $\mathfrak{m}$. If we take for X the interval $[0,1]$ and for τ the outer Lebesgue measure defined on the class $\mathfrak{X}$ of all sets, then the obtained function $\mathfrak{m}$ coincides with the initial function τ and the collection of $\mathfrak{m}$-measurable sets coincides with the class of the usual Lebesgue measurable sets, which is smaller than $\mathfrak{X}$. In Exercise 1.12.121 it is suggested to construct a similar example with an additive function τ on a σ-algebra of all sets in the interval.

Let us now specify one important class of outer measures.

1.11.6. Definition. *An outer measure $\mathfrak{m}$ on X is called regular if, for every set $A \subset X$, there exists an $\mathfrak{m}$-measurable set B such that $A \subset B$ and $\mathfrak{m}(A) = \mathfrak{m}(B)$.*

For example, the outer measure λ^* constructed from Lebesgue measure on the interval is regular, since one can take for B the set $\bigcap_{n=1}^{\infty} A_n$, where the sets A_n are measurable, $A \subset A_n$ and $\lambda(A_n) < \lambda^*(A) + 1/n$ (such a set is called a measurable envelope of A, see §1.12(iv)). More general examples are given below.

1.11.7. Proposition. *Let $\mathfrak{m}$ be a regular outer measure on X with $\mathfrak{m}(X) < \infty$. Then, the $\mathfrak{m}$-measurability of a set A is equivalent to the equality*

$$\mathfrak{m}(A) + \mathfrak{m}(X \backslash A) = \mathfrak{m}(X). \tag{1.11.6}$$

Proof. The necessity of (1.11.6) is obvious. Let us verify its sufficiency. Let E be an arbitrary set in X, $C \in \mathfrak{M}_{\mathfrak{m}}$, $E \subset C$, $\mathfrak{m}(C) = \mathfrak{m}(E)$. It suffices to show that

$$\mathfrak{m}(E) \geq \mathfrak{m}(E \cap A) + \mathfrak{m}(E \backslash A), \tag{1.11.7}$$

since the reverse inequality follows by the subadditivity. Note that

$$\mathfrak{m}(A \backslash C) + \mathfrak{m}\big((X \backslash A) \backslash C\big) \geq \mathfrak{m}(X \backslash C). \tag{1.11.8}$$

By the measurability of C one has

$$\mathfrak{m}(A) = \mathfrak{m}(A \cap C) + \mathfrak{m}(A \backslash C), \tag{1.11.9}$$

$$\mathfrak{m}(X \backslash A) = \mathfrak{m}\big(C \cap (X \backslash A)\big) + \mathfrak{m}\big((X \backslash A) \backslash C\big). \tag{1.11.10}$$

It follows by (1.11.6), (1.11.9) and (1.11.10) combined with the subadditivity of $\mathfrak{m}$ that

$$\begin{aligned} \mathfrak{m}(X) &= \mathfrak{m}(A \cap C) + \mathfrak{m}(A \backslash C) + \mathfrak{m}\big(C \cap (X \backslash A)\big) + \mathfrak{m}\big((X \backslash A) \backslash C\big) \\ &\geq \mathfrak{m}(C) + \mathfrak{m}(X \backslash C) = \mathfrak{m}(X). \end{aligned}$$

Therefore, the inequality in the last chain is in fact an equality. Subtracting from it (1.11.8), which is possible, since $\mathfrak{m}$ is finite, we arrive at the estimate

$$\mathfrak{m}(C \cap A) + \mathfrak{m}(C \backslash A) \le \mathfrak{m}(C).$$

Finally, the last estimate along with the inclusion $E \subset C$ and monotonicity of $\mathfrak{m}$ yields

$$\mathfrak{m}(E \cap A) + \mathfrak{m}(E \backslash A) \le \mathfrak{m}(C) = \mathfrak{m}(E).$$

Hence we have proved (1.11.7). □

Example 1.11.3 shows that Method I from Example 1.11.5 does not always yield regular outer measures. According to Exercise 1.12.122, if $\mathfrak{X} \subset \mathfrak{M}_{\mathfrak{m}}$, then Method I gives a regular outer measure. Yet another useful result in this direction is contained in the following theorem.

1.11.8. Theorem. *Let X, $\mathfrak{X}$, τ, and $\mathfrak{m}$ be the same as in Example* 1.11.5*. Suppose, in addition, that $\mathfrak{X}$ is an algebra* (*or a ring*) *and the function τ is additive. Then, the outer measure $\mathfrak{m}$ is regular and all sets in the class $\mathfrak{X}$ are measurable with respect to $\mathfrak{m}$. If τ is countably additive, then $\mathfrak{m}$ coincides with τ on $\mathfrak{X}$.*

Finally, if $\tau(X) < \infty$, then $\mathfrak{M}_{\mathfrak{m}} = \mathfrak{X}_\tau$, i.e., in this case the definition of the Carathéodory measurability is equivalent to Definition 1.5.1*.*

Proof. It suffices to verify that all sets in $\mathfrak{X}$ are measurable with respect to $\mathfrak{m}$; then the regularity will follow by Exercise 1.12.122. Let $A \in \mathfrak{X}$. In order to prove the inclusion $A \in \mathfrak{M}_{\mathfrak{m}}$, it suffices to show that, for every set E with $\mathfrak{m}(E) < \infty$, one has the estimate

$$\mathfrak{m}(E) \ge \mathfrak{m}(E \cap A) + \mathfrak{m}\bigl(E \cap (X \backslash A)\bigr).$$

Let $\varepsilon > 0$. There exist sets $X_n \in \mathfrak{X}$ with $E \subset \bigcup_{n=1}^\infty X_n$ and

$$\sum_{n=1}^\infty \tau(X_n) < \mathfrak{m}(E) + \varepsilon.$$

The condition that $\mathfrak{X}$ is a ring yields $X_n \cap A \in \mathfrak{X}$ and $X_n \cap (X \backslash A) = X_n \backslash A \in \mathfrak{X}$. Hence by the additivity of τ on $\mathfrak{X}$ we have for all n

$$\tau(X_n) = \tau(X_n \cap A) + \tau\bigl(X_n \cap (X \backslash A)\bigr).$$

Since

$$E \cap A \subset \bigcup_{n=1}^\infty (X_n \cap A), \quad E \cap (X \backslash A) \subset \bigcup_{n=1}^\infty \bigl(X_n \cap (X \backslash A)\bigr),$$

we obtain

$$\begin{aligned}\mathfrak{m}(E)+\varepsilon &> \sum_{n=1}^{\infty}\tau(X_n)=\sum_{n=1}^{\infty}\tau(X_n\cap A)+\sum_{n=1}^{\infty}\tau\bigl(X_n\cap(X\backslash A)\bigr)\\ &\geq \sum_{n=1}^{\infty}\mathfrak{m}(X_n\cap A)+\sum_{n=1}^{\infty}\mathfrak{m}\bigl(X_n\cap(X\backslash A)\bigr)\\ &\geq \mathfrak{m}(E\cap A)+\mathfrak{m}\bigl(E\cap(X\backslash A)\bigr).\end{aligned}$$

The required inequality is established, since ε is arbitrary. In the general case, one has $\mathfrak{m}\leq\tau$ on $\mathfrak{X}$, but for a countably additive function τ it is easy to obtain the reverse inequality.

Let us now verify that in the case $\tau(X)<\infty$, Definition 1.5.1 gives the same class of τ-measurable sets as Definition 1.11.2 applied to the outer measure $\mathfrak{m}=\tau^*$. Let $A\in\mathfrak{M}_{\mathfrak{m}}$ and $\varepsilon>0$. There exist sets $A_n\in\mathfrak{X}$ with $A\subset\bigcup_{n=1}^{\infty}A_n$ and $\mathfrak{m}(A)\geq\sum_{n=1}^{\infty}\tau(A_n)-\varepsilon$. Since $\mathfrak{m}(A_n)\leq\tau(A_n)$, taking into account the countable additivity of $\mathfrak{m}$ on the σ-algebra $\mathfrak{M}_{\mathfrak{m}}$, which contains $\mathfrak{X}$, we obtain

$$\mathfrak{m}(A)\geq\sum_{n=1}^{\infty}\mathfrak{m}(A_n)-\varepsilon\geq\mathfrak{m}\Bigl(\bigcup_{n=1}^{\infty}A_n\Bigr)-\varepsilon.$$

Therefore, $\mathfrak{m}\bigl(\bigcup_{n=1}^{\infty}A_n\backslash A\bigr)\leq\varepsilon$. By using the countable additivity of $\mathfrak{m}$ once again, we obtain $\mathfrak{m}\bigl(A\,\triangle\,\bigcup_{n=1}^{k}A_n\bigr)\leq 2\varepsilon$ for k sufficiently large. Since ε is arbitrary it follows that $A\in\mathfrak{X}_\tau$. Conversely, if $A\in\mathfrak{X}_\tau$, then, for every $\varepsilon>0$, there exists a set $A_\varepsilon\in\mathfrak{X}$ with $\mathfrak{m}(A\,\triangle\,A_\varepsilon)\leq\varepsilon$. One has $\mathfrak{X}\subset\mathfrak{M}_{\mathfrak{m}}$. By the countable additivity of $\mathfrak{m}$ on $\mathfrak{M}_{\mathfrak{m}}$, we obtain that A belongs to the Lebesgue completion of $\mathfrak{M}_{\mathfrak{m}}$. The completeness of $\mathfrak{M}_{\mathfrak{m}}$ yields the inclusion $A\in\mathfrak{M}_{\mathfrak{m}}$. □

1.11.9. Corollary. *If a countably additive set function with values in* $[0,+\infty]$ *is defined on a ring, then it has a countably additive extension to the* σ*-algebra generated by the given ring.*

Unlike the case of an algebra, the aforementioned extension is not always unique (as an example, consider the space $X=\{0\}$ with the zero measure on the ring $\mathfrak{X}=\{\varnothing\}$). It is easy to prove the uniqueness of a countably additive extension of a σ-finite measure τ from a ring $\mathfrak{X}$ to the generated σ-ring (see Exercise 1.12.159); if a measure τ on a ring $\mathfrak{X}$ is such that the corresponding outer measure $\mathfrak{m}$ on $\mathfrak{M}_{\mathfrak{m}}$ is σ-finite, then $\mathfrak{m}$ is a unique countably extension of τ also to $\sigma(\mathfrak{X})$ (see Exercise 1.12.159). In the above example the measure $\mathfrak{m}$ is not σ-finite because $\mathfrak{m}(\{0\})=\infty$.

Let us stress again that in general the outer measure $\mathfrak{m}$ may differ from τ on $\mathfrak{X}$ (see Exercise 1.12.121). Finally, we recall that if a function τ on an algebra $\mathfrak{X}$ is countably additive, then the associated outer measure $\mathfrak{m}$ coincides with τ on $\mathfrak{X}$. For infinite measures, it may happen that the class $\mathfrak{X}_\tau$ is strictly contained in $\mathfrak{M}_{\tau^*}$ (see Exercise 1.12.129).

Closing our discussion of Carathéodory outer measures let us prove a criterion of $\mathfrak{m}$-measurability of all Borel sets for an outer measure on $\mathbb{R}^n$. We recall that the distance from a point a to a set B is the number

$$\mathrm{dist}\,(a,B) := \inf_{b\in B} |b-a|.$$

1.11.10. Theorem. *Let $\mathfrak{m}$ be a Carathéodory outer measure on $\mathbb{R}^n$. In order that all Borel sets be $\mathfrak{m}$-measurable, it is necessary and sufficient that the following condition be fulfilled:*

$$\mathfrak{m}(A\cup B) = \mathfrak{m}(A) + \mathfrak{m}(B) \quad \textit{whenever} \;\; d(A,B) > 0, \tag{1.11.11}$$

where $d(A,B) := \inf_{a\in A, b\in B} |a-b|$, and $d(A,\varnothing) := +\infty$.

PROOF. Let $\mathfrak{M}_{\mathfrak{m}}$ contain all closed sets and $d(A,B) = d > 0$. We take disjoint closed sets

$$C_1 = \{x\colon\ \mathrm{dist}\,(x,A) \le d/4\} \supset A \quad \text{and} \quad C_2 = \{x\colon\ \mathrm{dist}\,(x,B) \le d/4\} \supset B$$

and observe that by Theorem 1.11.4(ii) one has

$$\mathfrak{m}\bigl((A\cup B)\cap(C_1\cup C_2)\bigr) = \mathfrak{m}\bigl((A\cup B)\cap C_1\bigr) + \mathfrak{m}\bigl((A\cup B)\cap C_2\bigr),$$

which yields (1.11.11), since

$$(A\cup B)\cap C_1 = A,\ (A\cup B)\cap C_2 = B,\ (A\cup B)\cap(C_1\cup C_2) = A\cup B.$$

Let (1.11.11) be fulfilled. It suffices to verify that every closed set C is $\mathfrak{m}$-measurable. Due to the subadditivity of $\mathfrak{m}$, the verification reduces to proving the estimate

$$\mathfrak{m}(A) \ge \mathfrak{m}(A\cap C) + \mathfrak{m}(A\backslash C), \quad \forall\, A\subset \mathbb{R}^n. \tag{1.11.12}$$

If $\mathfrak{m}(A) = \infty$, then (1.11.12) is true. So we assume that $\mathfrak{m}(A) < \infty$. The sets $C_n := \{x\colon\ \mathrm{dist}\,(x,C) \le n^{-1}\}$ monotonically decrease to C. Obviously, one has $d(A\backslash C_n, A\cap C) \ge n^{-1}$. Therefore,

$$\mathfrak{m}(A\backslash C_n) + \mathfrak{m}(A\cap C) = \mathfrak{m}\bigl((A\backslash C_n)\cup(A\cap C)\bigr) \le \mathfrak{m}(A). \tag{1.11.13}$$

Let us show that

$$\lim_{n\to\infty} \mathfrak{m}(A\backslash C_n) = \mathfrak{m}(A\backslash C). \tag{1.11.14}$$

Let us consider the sets $D_k := \bigl\{x\in A\colon\ (k+1)^{-1} < \mathrm{dist}\,(x,C) \le k^{-1}\bigr\}$. Then $A\backslash C = \bigcup_{k=n}^{\infty} D_k \bigcup (A\backslash C_n)$. Hence

$$\mathfrak{m}(A\backslash C_n) \le \mathfrak{m}(A\backslash C) \le \mathfrak{m}(A\backslash C_n) + \sum_{k=n}^{\infty} \mathfrak{m}(D_k).$$

Now, for proving (1.11.14), it suffices to observe that the series of $\mathfrak{m}(D_k)$ converges. Indeed, one has $d(D_k, D_j) > 0$ if $j \ge k+2$. By (1.11.11) and induction this gives the relation $\sum_{k=1}^{N} \mathfrak{m}(D_{2k}) = \mathfrak{m}\Bigl(\bigcup_{k=1}^{N} D_{2k}\Bigr) \le \mathfrak{m}(A)$ and a similar relation for odd numbers. According to (1.11.13) and (1.11.14) we obtain

$$\mathfrak{m}(A\backslash C) + \mathfrak{m}(A\cap C) = \lim_{n\to\infty} \mathfrak{m}(A\backslash C_n) + \mathfrak{m}(A\cap C) \le \mathfrak{m}(A).$$

The proof of (1.11.12) is complete. So the theorem is proven. □

It is seen from our reasoning that it applies to any metric space in place of $\mathbb{R}^n$. We shall return to this subject in §7.14(x).

1.12. Supplements and exercises

1.12(i). Set operations

The following result of Sierpiński contains several useful modifications of Theorem 1.9.3 on monotone classes.

Let us consider the following list of operations on sets in a given set X and indicate the corresponding notation:

a finite union $\cup f$, a countable union $\cup c$, the union of an increasing sequence of sets $\lim\uparrow$, a disjoint union $\sqcup f$, a countable disjoint union $\sqcup c$, a finite intersection $\cap f$, a countable intersection $\cap c$, the intersection of a decreasing sequence of sets $\lim\downarrow$, the difference of sets $\backslash$, the difference of a set and its subset $-$.

Note that the symbols f and c indicate the finite and countable character of the corresponding operations and that in the operation $A\backslash B$ the set B may not belong to A, unlike the operation $-$. Every operation O in this list has the dual operation denoted by the symbol O^d and defined as follows:

$$(\cup f)^d := \cap f,\ (\cup c)^d := \cap c,\ (\lim\uparrow)^d := \lim\downarrow,\ (\sqcup f)^d := -,\ (\sqcup c)^d := -, \tag{1.12.1}$$
$$(\cap f)^d := \cup f,\ (\cap c)^d := \cup c, (\lim\downarrow)^d := \lim\uparrow,\ (\backslash)^d := \cup f,\ (-)^d := \sqcup f.$$

The property of a family $\mathcal{F}$ of subsets of X to be closed with respect to some of the above operations is understood in the natural way; for example, "$\mathcal{F}$ is closed with respect to $\lim\uparrow$" means that if sets $F_n \in \mathcal{F}$ increase, then their union belongs to $\mathcal{F}$ as well. It is readily verified that if we are given a class $\mathcal{F}$ of subsets of X and a collection of operations from the above list, then there is the smallest class of sets that contains $\mathcal{F}$ and is closed with respect to the given operations.

1.12.1. Theorem. *Let $\mathcal{F}$ and $\mathcal{G}$ be two classes of subsets of X such that $\mathcal{G} \subset \mathcal{F}$ and the class $\mathcal{F}$ is closed with respect to some collection of operations $\mathcal{O} = (O_1, O_2, \ldots)$ from* (1.12.1). *Denote by $\mathcal{F}_0$ the smallest class of sets that contains $\mathcal{G}$ and is closed with respect to the operations from the same collection $\mathcal{O}$. Then the following assertions are true*:

(i) *if $G\cap G' \in \mathcal{F}_0$ for all $G, G' \in \mathcal{G}$, then the class $\mathcal{F}_0$ is closed with respect to finite intersections*;

(ii) *if* $O^d \in \mathcal{O}$ *for every operation* $O \in \mathcal{O}$ *and* $X\backslash G \in \mathcal{F}_0$ *for all* $G \in \mathcal{G}$, *then the class* $\mathcal{F}_0$ *is closed with respect to complementation*; *in particular, if* $\mathcal{O} = (\cup c, \cap c)$, *then* $\mathcal{F}_0 = \sigma(\mathcal{G})$;

(iii) *if all the conditions in* (i) *and* (ii) *are fulfilled, then the algebra generated by* $\mathcal{G}$ *is contained in* $\mathcal{F}$, *and if* $\mathcal{O} = (\lim\uparrow, \lim\downarrow)$, *then* $\mathcal{F}_0 = \sigma(\mathcal{G})$.

A proof analogous to that of the monotone class theorem is left as Exercise 1.12.100. Another result due to Sierpiński gives a modification of the theorem on σ-additive classes.

1.12.2. Theorem. *Let* $\mathcal{E}$ *be a class of subsets in a space* X *containing the empty set. Denote by* $\mathcal{E}_{\sqcup,\delta}$ *the smallest class of sets in* X *that contains* $\mathcal{E}$ *and is closed with respect to countable unions of pairwise disjoint sets and any countable intersections. If* $X\backslash E \in \mathcal{E}_{\sqcup,\delta}$ *for all* $E \in \mathcal{E}$, *then* $\mathcal{E}_{\sqcup,\delta} = \sigma(\mathcal{E})$.

Proof. Let $\mathcal{A} := \{A \in \mathcal{E}_{\sqcup,\delta} \colon\ X\backslash A \in \mathcal{E}_{\sqcup,\delta}\}$. It suffices to show that the class $\mathcal{A}$ is closed with respect to countable unions of pairwise disjoint sets and any countable intersections, since it will coincide then with the class $\mathcal{E}_{\sqcup,\delta}$, hence the latter will be closed under complementation, i.e., will be a σ-algebra. If sets $A_n \in \mathcal{A}$ are disjoint, then their union belongs to $\mathcal{E}_{\sqcup,\delta}$ by the definition of $\mathcal{E}_{\sqcup,\delta}$, and the complement of their union is $\bigcap_{n=1}^\infty (X\backslash A_n)$, which also belongs to $\mathcal{E}_{\sqcup,\delta}$, since $X\backslash A_n \in \mathcal{E}_{\sqcup,\delta}$. Hence $\mathcal{A}$ admits countable unions of disjoint sets. If $B_n \in \mathcal{A}$, then $\bigcap_{n=1}^\infty B_n \in \mathcal{E}_{\sqcup,\delta}$. Finally, observe that $X\backslash\bigcap_{n=1}^\infty B_n$ can be written in the form

$$\bigcup_{n=1}^\infty (X\backslash B_n) = \bigcup_{n=1}^\infty \Bigl[(X\backslash B_n) \cap \Bigl(\bigcap_{k=1}^{n-1} B_k\Bigr)\Bigr]. \tag{1.12.2}$$

Indeed, the right-hand side obviously belongs to the left one. If x belongs to the left-hand side, then, for some n, we have $x \notin B_n$. If x does not belong to the right-hand side, then $x \notin \bigcap_{k=1}^{n-1} B_k$ and $x \in B_1$. Hence there exists a number m between 1 and $n-2$ such that $x \in \bigcap_{k=1}^m B_k$ and $x \notin \bigcap_{k=1}^{m+1} B_k$. Then $x \in (X\backslash B_{m+1}) \cap \bigl(\bigcap_{k=1}^m B_k\bigr)$, which belongs to the right-hand side of (1.12.2), contrary to our assumption. It is clear that the sets whose union is taken in the right-hand side of (1.12.2) are pairwise disjoint and belong to $\mathcal{E}_{\sqcup,\delta}$ because we have $X\backslash B_n, B_k \in \mathcal{E}_{\sqcup,\delta}$. Thus, $\mathcal{E}_{\sqcup,\delta}$ admits countable intersections. □

1.12.3. Example. The smallest class of subsets of the real line that contains all open sets and is closed under countable unions of pairwise disjoint sets and any countable intersections is the Borel σ-algebra. The same is true if in place of all open sets one takes all closed sets.

Proof. If $\mathcal{E}$ is the class of all open sets, then the theorem applies directly, since the complement of any open set is closed and hence is the countable intersection of a sequence of open sets.

Now let $\mathcal{E}$ be the class of all closed sets. Let us verify that the complements of sets in $\mathcal{E}$ belong to the class $\mathcal{E}_{\sqcup,\delta}$. These complements are open, hence are disjoint unions of intervals or rays. Hence it remains to show that every open interval (a,b) belongs to $\mathcal{E}_{\sqcup,\delta}$. This is not completely obvious, since the open interval cannot be represented in the form of a disjoint union of a sequence of closed intervals. However, one can find a sequence of pairwise disjoint nondegenerate closed intervals $I_n \subset (a,b)$ such that their union S is everywhere dense in (a,b). Let us now verify that $B := (a,b)\backslash S \in \mathcal{E}_{\sqcup,\delta}$. We observe that the closure $\overline{B}$ of the set B consists of B and the countable set $M = \{x_k\}$ formed by the points a and b and the endpoints of the intervals I_n. Hence $B = \bigcap_{m=1}^{\infty} \overline{B}\backslash\{x_1,\ldots,x_m\}$. The set $\overline{B}$ is nowhere dense compact. This enables us to represent each of the sets $\overline{B}\backslash\{x_1,\ldots,x_m\}$ in the form of the union of disjoint compact sets. Let us do this for $\overline{B}\backslash\{x_1\}$, the reasoning for other sets is similar. Since $\overline{B}$ has no interior, the open complement of $\overline{B}$ contains a sequence of points l_j increasing to x_1 and a sequence of points r_j decreasing to x_1. We may assume that $l_1 < a$, $r_1 > b$. The sets $(l_j, l_{j+1}) \cap \overline{B}$ and $(r_{j+1}, r_j) \cap \overline{B}$ are compact, since the points $l_j, l_{j+1}, r_{j+1}, r_j$ belong to the complement of $\overline{B}$ with some neighborhoods. These sets give the desired decomposition of $\overline{B}\backslash\{x_1\}$. □

In Chapter 6 one can find some additional information related to the results in this subsection.

1.12(ii). Compact classes

A compact class approximating a measure may not consist of measurable sets. For example, if $\mathcal{A}$ is the σ-algebra on $[0,1]^2$ consisting of the sets $B\times[0,1]$, where $B \in \mathcal{B}([0,1])$, μ is the restriction of Lebesgue measure to $\mathcal{A}$, and $\mathcal{K}$ is the class of all compact sets in $[0,1]^2$, then $\mathcal{K}$ is approximating for μ, but the interval $I := [0,1]\times\{0\}$ does not belong to $\mathcal{A}_\mu$, since $\mu^*(I) = 1$ and I does not contain nonempty sets from $\mathcal{A}$. In addition, a compact approximating class may not be closed with respect to unions and intersections. The next result shows that one can always "improve" the original approximating compact class by replacing it with a compact class that consists of measurable sets, approximates the measure, and is stable under finite unions and countable intersections.

1.12.4. Proposition. (i) *Let $\mathcal{K}$ be a compact class of subsets of a set X. Then, the minimal class $\mathcal{K}_{s\delta}$ which contains $\mathcal{K}$ and is closed with respect to finite unions and countable intersections, is compact as well* (*more precisely, $\mathcal{K}_{s\delta}$ coincides with the class of at most countable intersections of finite unions of elements of $\mathcal{K}$*).

(ii) *In addition, if $\mathcal{E}$ is a compact class of subsets of a set Y, then the class of products $K\times E$, $K \in \mathcal{K}$, $E \in \mathcal{E}$, is compact as well.*

(iii) *If a nonnegative measure μ on an algebra* (*or semialgebra*) *$\mathcal{A}_0$ has an approximating compact class $\mathcal{K}$, then there exists a compact class $\mathcal{K}'$ that is contained in $\sigma(\mathcal{A}_0)$, approximates μ on $\sigma(\mathcal{A}_0)$, and is stable under finite unions and countable intersections.*

PROOF. (i) We show first that the class $\mathcal{K}_s$ of finite unions of sets in $\mathcal{K}$ is compact. Let $A_i = \bigcup_{n=1}^{m_i} K_i^n$, where $K_i^n \in \mathcal{K}$, be such that $\bigcap_{i=1}^k A_i \neq \varnothing$ for all $k \in \mathbb{N}$. Denote by M the set of all sequences $\nu = (\nu_i)$ such that $\nu_i \leq m_i$ for all $i \geq 1$. Let M_k be the collection of all sequences ν in M such that $\bigcap_{i=1}^k K_i^{\nu_i} \neq \varnothing$. Note that the sets M_k are nonempty for all k. This follows from the relation

$$\bigcup_{\nu \in M} \bigcap_{i=1}^{k} K_i^{\nu_i} = \bigcap_{i=1}^{k} A_i \neq \varnothing,$$

which is easily seen from the fact that $x \in \bigcap_{i=1}^k A_i$ precisely when there exist $\nu_i \leq m_i$, $i = 1, \ldots, k$, with $x \in K_i^{\nu_i}$. In addition, the sets M_k are decreasing. We prove that there is a sequence ν in their intersection. This means that the intersection $\bigcap_{n=1}^\infty A_n$ is nonempty, since it contains the set $\bigcap_{n=1}^\infty K_n^{\nu_n}$, which is nonempty by the compactness of the class $\mathcal{K}$ and the fact that the sets $\bigcap_{n=1}^k K_n^{\nu_n}$ are nonempty.

In order to prove the relation $\bigcap_{k=1}^\infty M_k \neq \varnothing$ let us choose an element $\nu^{(k)} = (\nu_n^{(k)})_{n=1}^\infty$ in every set M_k. Since $\nu_n^{(k)} \leq m_n$ for all n and k, there exist infinitely many indices k such that the numbers $\nu_1^{(k)}$ coincide with one and the same number ν_1. By induction, we construct a sequence of natural numbers $\nu = (\nu_i)$ such that, for every n, there exist infinitely many indices k with the property that $\nu_i^{(k)} = \nu_i$ for all $i = 1, \ldots, n$. This means that $\nu \in M_n$, since the membership in M_n is determined by the first n coordinates of a sequence, and for all $k > n$ we have $\nu^{(k)} \in M_n$ by the inclusion $\nu^{(k)} \in M_k \subset M_n$. Thus, ν belongs to all M_n.

The compactness of the class $\mathcal{K}_s$ obviously yields the compactness of the class $\mathcal{K}_{s\delta}$ of all at most countable intersections of sets in $\mathcal{K}_s$. It is clear that this is the smallest class that contains $\mathcal{K}$ and is closed with respect to finite unions and at most countable intersections (observe that a finite union of several countable intersections of finite unions of sets in $\mathcal{K}$ can be written as a countable intersection of finite unions).

(ii) If the intersections $\bigcap_{n=1}^N (K_n \times E_n)$, where $K_n \in \mathcal{K}$, $E_n \in \mathcal{E}$, are nonempty, then $\bigcap_{n=1}^N K_n$ and $\bigcap_{n=1}^N E_n$ are nonempty as well, which by the compactness of $\mathcal{K}$ and $\mathcal{E}$ gives points $x \in \bigcap_{n=1}^\infty K_n$ and $y \in \bigcap_{n=1}^\infty E_n$. Then $(x, y) \in \bigcap_{n=1}^\infty (K_n \times E_n)$.

(iii) According to (i) we can assume that $\mathcal{K}$ is stable under finite unions and countable intersections. Let $\mathcal{K}' = \mathcal{K} \cap \sigma(\mathcal{A}_0)$. Clearly, $\mathcal{K}'$ is a compact class. Let us show that $\mathcal{K}'$ approximates μ on $\mathcal{A}_0$. Given $A \in \mathcal{A}_0$ and $\varepsilon > 0$, we can construct inductively sets $A_n \in \mathcal{A}_0$ and $K_n \in \mathcal{K}$ such that

$$A \supset K_1 \supset A_1 \supset K_2 \supset A_2 \supset \cdots \quad \text{and} \quad \mu(A_n \backslash A_{n+1}) < \varepsilon 2^{-n-1}, \ A_0 := A.$$

We observe that $\bigcap_{n=1}^\infty A_n = \bigcap_{n=1}^\infty K_n$. Denoting this set by K we have $K \in \mathcal{K}'$, since $\sigma(\mathcal{A}_0)$ and $\mathcal{K}$ admit countable intersections. In addition, $K \subset A$ and $\mu(A \backslash K) < \varepsilon$. Finally, $\mathcal{K}'$ approximates μ on $\sigma(\mathcal{A}_0)$. Indeed, for every $A \in \sigma(\mathcal{A}_0)$ and every $\varepsilon > 0$, one can find sets $A_n \in \mathcal{A}$ such that $A_0 :=$

$\bigcap_{n=1}^{\infty} A_n \subset A$ and $\mu(A\backslash A_0) < \varepsilon$. To this end, it suffices to find sets $B_n \in \mathcal{A}$ covering $X\backslash A$ such that the measure of their union is less than $\mu(X\backslash A) + \varepsilon$ and take $A_n = X\backslash B_n$. There exist sets $K_n \in \mathcal{K}'$ such that $K_n \subset A_n$ and $\mu(A_n\backslash K_n) < \varepsilon 2^{-n}$. Let $K := \bigcap_{n=1}^{\infty} K_n$. Then $K \subset A_0$, $\mu(A_0\backslash K) < \mu(K) + \varepsilon$ and $K \in \mathcal{K}'$ because $\mathcal{K}'$ is stable under countable intersections. □

Assertion (ii) will be reinforced in Lemma 3.5.3. The class of sets of the form $K{\times}E$, where $K \in \mathcal{K}$, $E \in \mathcal{E}$, is denoted by $\mathcal{K}{\times}\mathcal{E}$ (the usual understanding of the product of sets $\mathcal{K}$ and $\mathcal{E}$ as the collection of pairs (K, E) does not lead to confusion here).

It is worth noting that if μ is a finite nonnegative measure on a σ-algebra $\mathcal{A}$, then, by assertion (iii) above, the existence of a compact approximating class for μ does not depend on whether we consider μ on $\mathcal{A}$ or on its completion $\mathcal{A}_\mu$. We know that an approximating compact class $\mathcal{K}$ need not be contained in $\mathcal{A}_\mu$. However, according to Theorem 1.12.34 stated below, there is a countably additive extension of μ to the σ-algebra generated by $\mathcal{A}$ and $\mathcal{K}$.

A property somewhat broader than compactness is monocompactness, considered in the following result of Mallory [**647**], which strengthens Theorem 1.4.3.

1.12.5. Theorem. *Let $\mathcal{R}$ be a semiring and let μ be an additive nonnegative function on $\mathcal{R}$ such that there exists a class of sets $\mathcal{M} \subset \mathcal{R}$ with the following property: if sets $M_n \in \mathcal{M}$ are nonempty and decreasing, then $\bigcap_{n=1}^{\infty} M_n$ is nonempty (such a class is called monocompact). Suppose that*

$$\mu(R) = \sup\{\mu(M)\colon\ M \in \mathcal{M}, M \subset R\} \quad \textit{for all } R \in \mathcal{R}.$$

Then μ is countably additive on $\mathcal{R}$.

Proof. Let $R = \bigcup_{n=1}^{\infty} R_n$, where $R_n \in \mathcal{R}$. It suffices to show that

$$\mu(R) \le \sum_{n=1}^{\infty} \mu(R_n).$$

Suppose the opposite. Then there exists a number c such that

$$\sum_{n=1}^{\infty} \mu(R_n) < c < \mu(R).$$

Let us take $M \in \mathcal{M}$ with $M \subset R$ and $\mu(M) > c$. We can write $M\backslash R_1$ as a disjoint union

$$M\backslash R_1 = \bigcup_{j=1}^{m_1} R^j, \quad R^j \in \mathcal{R}.$$

Let us find $M_1, \ldots, M_{m_1} \in \mathcal{M}$ with $M_j \subset R^j$ and $\sum_{j=1}^{m_1} \mu(M_j) + \mu(R_1) > c$. By induction, we construct sets $M_{j_1,\ldots,j_n} \in \mathcal{M}$ as follows. If $M_{j_1,\ldots,j_n}$ are already constructed, then we find finitely many disjoint sets $R^{j_1,\ldots,j_n,j} \in \mathcal{R}$

whose union is $M_{j_1,\ldots,j_n}\backslash R_{n+1}$, and also a set $M_{j_1,\ldots,j_n,j} \in \mathcal{M}$ such that one has $M_{j_1,\ldots,j_n,j} \subset R^{j_1,\ldots,j_n,j}$ and

$$\sum_{j_1,\ldots,j_n,j} \mu(M_{j_1,\ldots,j_n,j}) + \sum_{i=1}^{n} \mu(R_i) > c.$$

Note that $\sum\limits_{j_1,\ldots,j_n,j} \mu(M_{j_1,\ldots,j_n,j}) > 0$ due to our choice of c. Hence there exists a sequence of indices j_i such that $M_{j_1,\ldots,j_k} \neq \varnothing$ for all k (such a sequence is found by induction by choosing $j_1,\ldots,j_{k-1}$ with $\mu(M_{j_1,\ldots,j_{k-1}}) > 0$). Thus, $\bigcap_{k=1}^{\infty} M_{j_1,\ldots,j_k}$ is nonempty, whence it follows that $R \neq \bigcup_{n=1}^{\infty} R_n$, which is a contradiction. □

Fremlin [**326**] constructed an example that distinguishes compact and monocompact measures, i.e., there is a probability measure possessing a monocompact approximating class, but having no compact (countably compact by the terminology of the cited work) approximating classes.

1.12(iii). Metric Boolean algebra

Let $(X, \mathcal{A}, \mu)$ be a measure space with a finite nonnegative measure μ. In this subsection we discuss a natural metric structure on the set of all μ-measurable sets.

Suppose first that μ is a bounded nonnegative additive set function on an algebra $\mathcal{A}$. Set

$$d(A, B) = \mu(A \triangle B), \quad A, B \in \mathcal{A}.$$

The function d is called the Fréchet–Nikodym metric. Let us introduce the following relation on $\mathcal{A}$: $A \sim B$ if $d(A, B) = 0$. Clearly, $A \sim B$ if and only if A and B differ in a measure zero set. This is an equivalence relation:

1) $A \sim A$, 2) if $A \sim B$, then $B \sim A$, 3) if $A \sim B$ and $B \sim C$, then $A \sim C$. Denote by $\mathcal{A}/\mu$ the set of all equivalence classes for this relation. The function d has a natural extension to $\mathcal{A}/\mu \times \mathcal{A}/\mu$:

$$d(\widetilde{A}, \widetilde{B}) = d(A, B)$$

if A and B represent the classes $\widetilde{A}$ and $\widetilde{B}$, respectively. By the additivity of μ, this definition does not depend on our choice of representatives in the equivalence classes. The function d makes the set $\mathcal{A}/\mu$ a metric space. The triangle inequality follows, since for all $A, B, C \in \mathcal{A}$ one has the inclusion $A\triangle C \subset (A\triangle B)\cup(B\triangle C)$, whence we obtain $\mu(A\triangle C) \leq \mu(A\triangle B)+\mu(B\triangle C)$. By means of representatives of classes, one introduces the operations of intersection, union, and complementation on $\mathcal{A}/\mu$. The metric space $(\mathcal{A}/\mu, d)$ is called the metric Boolean algebra generated by $(\mathcal{A}, \mu)$. Note that the function μ is naturally defined on $\mathcal{A}/\mu$ and is Lipschitzian on $(\mathcal{A}/\mu, d)$. This follows by the inequality $|\mu(A) - \mu(B)| \leq \mu(A \triangle B) = d(A, B)$.

A measure μ is called separable if the metric space $(\mathcal{A}/\mu, d)$ is separable, i.e., contains a countable everywhere dense subset. The separability of μ is equivalent to the existence of an at most countable collection of sets $A_n \in \mathcal{A}$

such that, for every $A \in \mathcal{A}$ and $\varepsilon > 0$, there exists n with $\mu(A \triangle A_n) < \varepsilon$. The last property can be taken as a definition of separability for infinite measures. Lebesgue measure and many other measures encountered in applications are separable, but nonseparable measures exist as well. Concerning separable measures, see Exercises 1.12.102 and 4.7.63 and §7.14(iv).

1.12.6. Theorem. *Let μ be a bounded nonnegative additive set function on an algebra $\mathcal{A}$.*

(i) *The function μ is countably additive if and only if $d(A_n, \varnothing) \to 0$ as $A_n \downarrow \varnothing$.*

(ii) *If $\mathcal{A}$ is a σ-algebra and μ is countably additive, then the metric space $(\mathcal{A}/\mu, d)$ is complete.*

Proof. (i) It suffices to note that $A_n \triangle \varnothing = A_n$ and $d(A_n, \varnothing) = \mu(A_n)$.
(ii) Let $\{\widetilde{A}_n\}$ be a Cauchy sequence in $(\mathcal{A}/\mu, d)$ and A_n a representative of the class $\widetilde{A}_n$. Let us show that there exists a set $A \in \mathcal{A}$ such that $d(A_n, A) \to 0$. It suffices to show that there is a convergent subsequence in $\{A_n\}$. Hence, passing to a subsequence, we may assume that $\mu(A_k \triangle A_n) < 2^{-n}$ for all n and $k \geq n$. Set

$$A = \limsup_{n\to\infty} A_n := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k.$$

We show that $d(A_n, A) \to 0$. Let $\varepsilon > 0$. The sets $\bigcap_{n=1}^{N} \bigcup_{k=n}^{\infty} A_k$ increase to A. By the countable additivity of μ there exists a number N such that

$$\mu\Bigl(\bigcup_{k=N}^{\infty} A_k \backslash A\Bigr) = \mu\Bigl(\bigcap_{n=1}^{N} \bigcup_{k=n}^{\infty} A_k \backslash A\Bigr) < \varepsilon.$$

Then, for all $m \geq N$, we have

$$\mu\Bigl(\bigcup_{k=m}^{\infty} A_k \backslash A\Bigr) < \varepsilon.$$

Since $\mu(A_m \triangle A_k) \geq \mu(A_k \backslash A_m)$, we obtain for all m sufficiently large that

$$\mu\Bigl(\bigcup_{k=m}^{\infty} A_k \backslash A_m\Bigr) \leq \sum_{k=m+1}^{\infty} \mu(A_k \backslash A_m) \leq \sum_{k=m+1}^{\infty} 2^{-k} < \varepsilon,$$

whence we have $\mu(A_m \triangle A) < 2\varepsilon$, since $A, A_m \subset \bigcup_{k=m}^{\infty} A_k$. □

We remark that in assertion (ii) the space $(\mathcal{A}/\mu, d)$ is complete even if $\mathcal{A}$ is not complete with respect to μ, which is natural, since every set in the completed σ-algebra $\mathcal{A}_\mu$ coincides up to a measure zero set with an element of $\mathcal{A}$, hence belongs to the same equivalence class. Note also that the consideration of $(\mathcal{A}/\mu, d)$ is simplified if we employ the concepts of the theory of integration developed in Chapters 2 and 4 and deal with the indicator functions of sets rather than with sets themselves.

Now let $\mathcal{A}$ be a σ-algebra and let μ be countably additive.

1.12.7. Definition. The set $A \in \mathcal{A}$ is called an atom of the measure μ if $\mu(A) > 0$ and every set $B \subset A$ from $\mathcal{A}$ has measure either 0 or $\mu(A)$.

If two atoms A_1 and A_2 are distinct in the sense that $d(A, B) > 0$ (i.e., A and B are not equivalent), then $\mu(A_1 \cap A_2) = 0$. Hence there exists an at most countable set $\{A_n\}$ of pairwise non-equivalent atoms. The measure μ is called purely atomic if $\mu\bigl(X \backslash \bigcup_{n=1}^{\infty} A_n\bigr) = 0$. If there are no atoms, then the measure μ is called atomless.

1.12.8. Example. Lebesgue measure λ is atomless on every measurable set A in $[a, b]$. Moreover, for any $\alpha \in [0, \lambda(A)]$, there exists a set $B \subset A$ such that $\lambda(B) = \alpha$.

PROOF. The function $F(x) = \lambda\bigl(A \cap [a, x)\bigr)$ is continuous on $[a, b]$ by the countable additivity of Lebesgue measure. It remains to apply the mean value theorem. □

1.12.9. Theorem. *Let $(X, \mathcal{A}, \mu)$ be a measure space with a finite nonnegative measure μ. Then, for every $\varepsilon > 0$, there exists a finite partition of X into pairwise disjoint sets $X_1, \ldots, X_n \in \mathcal{A}$ with the following property: either $\mu(X_i) \le \varepsilon$, or X_i is an atom of measure greater than ε.*

PROOF. There exist only finitely many non-equivalent atoms $A_1, \ldots, A_p$ of measure greater than ε. Then the space $Y = X \backslash \bigcup_{i=1}^{p} A_i$ has no atoms of measure greater than ε. Let us show that every set $B \in \mathcal{A}$, contained in Y and having positive measure, contains a set C such that $0 < \mu(C) \le \varepsilon$. Indeed, suppose that there exists a set B for which this is false. Then $\mu(B) > \varepsilon$ (otherwise we may take $C = B$) and hence B is not an atom. Therefore, there exists a set $B_1 \in \mathcal{A}$ with $\varepsilon < \mu(B_1) < \mu(B)$. Then $\mu(B \backslash B_1) > \varepsilon$ (otherwise we arrive at a contradiction with our choice of B) and for the same reason the set $C_1 = B \backslash B_1$ contains a subset $B_2 \in \mathcal{A}$ with $\varepsilon < \mu(B_2) < \mu(C_1)$. Note that $\mu(C_1 \backslash B_2) > \varepsilon$. Let $C_2 = C_1 \backslash B_2$ and in C_2 we find a set $B_3 \in \mathcal{A}$ with $\varepsilon < \mu(B_3) < \mu(C_2)$. Continuing by induction, we obtain an infinite sequence of pairwise disjoint sets B_n of measure greater than ε, which is impossible, since $\mu(Y) < \infty$.

Now for every $A \in \mathcal{A}$ we set

$$\eta(A) = \sup\bigl\{\mu(B)\colon\ B \subset A, B \in \mathcal{A}, \mu(B) \le \varepsilon\bigr\}.$$

According to what has been proven above, one has that $0 < \eta(A) \le \varepsilon$ if $A \subset Y$ and $\mu(A) > 0$. We may find a set $B_1 \in \mathcal{A}$ in Y such that $0 < \mu(B_1) \le \eta(Y)$, provided that $\mu(Y) > \varepsilon$; if $\mu(Y) \le \varepsilon$, then the proof is complete. By using the above established property of subsets of Y, we construct by induction a sequence of pairwise disjoint sets $B_n \in \mathcal{A}$ such that $B_n \subset Y$ and

$$\frac{1}{2}\eta\Bigl(Y \backslash \bigcup_{i=1}^{n} B_i\Bigr) \le \mu(B_{n+1}) \le \varepsilon.$$

If at some step it is impossible to continue this construction, then this completes the proof. Let $B_0 = Y \backslash \bigcup_{i=1}^\infty B_i$. Then

$$\eta(B_0) \le \eta\Bigl(Y \backslash \bigcup_{i=1}^n B_i\Bigr) \le 2\mu(B_{n+1})$$

for all n. The series of measures of B_n converges, hence $\mu(B_n) \to 0$, whence we have $\eta(B_0) = 0$. Therefore, $\mu(B_0) = 0$. It remains to take a number k such that $\sum_{i=k}^\infty \mu(B_i) < \varepsilon$. The sets $A_1, \ldots, A_p$, $B_1, \ldots, B_k$, $\bigcup_{i=k+1}^\infty B_i \bigcup B_0$ form a desired partition. $\square$

1.12.10. Corollary. *Let μ be an atomless measure. Then, for every $\alpha \in [0, \mu(X)]$, there exists a set $A \in \mathcal{A}$ such that $\mu(A) = \alpha$.*

Proof. By using the previous theorem one can construct an increasing sequence of sets $A_n \in \mathcal{A}$ such that $\mu(A_n) \to \alpha$. Indeed, let $\alpha > 0$. We can partition X into finitely many parts X_j with $\mu(X_j) < 1/2$. Let us take the biggest number m with $\mu\bigl(\bigcup_{j=1}^m X_j\bigr) \le \alpha$. Letting $A_1 := \bigcup_{j=1}^m X_j$ we have $\mu(A_1) \ge \alpha - 1/2$. In the same manner we find a set $B_1 \subset X \backslash A_1$ with $\mu(B_1) \ge \alpha - \mu(A_1) - 1/3$ and take $A_2 := A_1 \cup B_1$. We proceed by induction and obtain sets A_{n+1} of the form $A_n \cup B_n$, where $B_n \subset X \backslash A_n$ and $\mu(B_n) \ge \alpha - \mu(A_n) - (n+1)^{-1}$. Now we can take $A = \bigcup_{n=1}^\infty A_n$. $\square$

We remark that in the case of infinite measures the Fréchet–Nikodym metric can be considered on the class of sets of finite measure. Another related metric is considered in Exercise 1.12.152.

1.12(iv). Measurable envelope, measurable kernel and inner measure

Let $(X, \mathcal{B}, \mu)$ be a measure space with a finite nonnegative measure μ. We observe that the restriction of μ to a measurable subset A is again a measure defined on the trace σ-algebra $\mathcal{B}_A$ of the space A that consists of the sets $A \cap B$, where $B \in \mathcal{B}$. The following construction enables one to restrict μ to arbitrary sets A, possibly nonmeasurable, if we define $\mathcal{B}_A$ as above. The trace σ-algebra $\mathcal{B}_A$ is also called the restriction of the σ-algebra $\mathcal{B}$ to A and denoted by the symbol $\mathcal{B} \cap A$.

For any set $A \subset X$, there exists a set $\widetilde{A} \in \mathcal{B}$ (called a *measurable envelope* of A) with

$$A \subset \widetilde{A} \text{ and } \mu(\widetilde{A}) = \mu^*(A). \tag{1.12.3}$$

For such a set (which is not unique) we can take

$$\widetilde{A} = \bigcap_{n=1}^\infty A_n, \text{ where } A_n \in \mathcal{B},\ A_n \supset A \text{ and } \mu(A_n) \le \mu^*(A) + 1/n. \tag{1.12.4}$$

Informally speaking, $\widetilde{A}$ is a minimal measurable set containing A.

By (1.12.3) and the definition of outer measure it follows that if we have $A \subset B \subset \widetilde{A}$ and $B \in \mathcal{B}$, then $\mu(\widetilde{A} \triangle B) = 0$.

1.12.11. Definition. *The restriction μ_A (denoted also by $\mu|_A$) of the measure μ to $\mathcal{B}_A$ is defined by the formula*

$$\mu_A(B \cap A) := \mu|_A(B \cap A) := \mu(B \cap \widetilde{A}), \quad B \in \mathcal{B},$$

where $\widetilde{A}$ is an arbitrary measurable envelope of A.

It is easily seen that this definition does not depend on our choice of $\widetilde{A}$ and that the function μ_A is countably additive. If $A \in \mathcal{B}$, then we obtain the usual restriction.

1.12.12. Proposition. *The measure μ_A coincides with the restriction of the outer measure μ^* to $\mathcal{B}_A$.*

PROOF. Let $B \in \mathcal{B}$. Then

$$\mu^*(B \cap A) \le \mu^*(B \cap \widetilde{A}) = \mu(B \cap \widetilde{A}) = \mu_A(B \cap A).$$

On the other hand, if $B \cap A \subset C$, where $C \in \mathcal{B}$, then

$$A \subset \widetilde{A} \backslash \bigl(B \cap (\widetilde{A} \backslash C)\bigr).$$

By the definition of a measurable envelope we obtain $\mu\bigl(B \cap (\widetilde{A} \backslash C)\bigr) = 0$. Hence

$$\mu(B \cap \widetilde{A}) \le \mu(B \cap C) + \mu\bigl(B \cap (\widetilde{A} \backslash C)\bigr) = \mu(B \cap C) \le \mu(C),$$

which yields by taking inf over C that $\mu(B \cap \widetilde{A}) \le \mu^*(B \cap A)$. □

By analogy with a measurable envelope one can define a measurable kernel $\underline{A}$ of an arbitrary set A. Namely, let us first define the *inner measure* of a set A by the formula

$$\mu_*(A) = \sup\bigl\{\mu(B) \colon\ B \subset A, B \in \mathcal{B}\bigr\}.$$

A measurable kernel of a set A is a set $\underline{A} \in \mathcal{B}$ such that

$$\underline{A} \subset A \quad \text{and} \quad \mu(\underline{A}) = \mu_*(A).$$

For $\underline{A}$ one can take the union of a sequence of sets $B_n \in \mathcal{B}$ such that $B_n \subset A$ and $\mu(B_n) \ge \mu_*(A) - 1/n$. Obviously, a measurable kernel is not unique, but if a set C from $\mathcal{B}$ is contained in A, then $\mu(C \backslash \underline{A}) = 0$. Informally speaking, $\underline{A}$ is a maximal measurable subset of A.

Outer and inner measures are also denoted by the symbols μ_e and μ_i, respectively (from “mesure extérieure” and “mesure intérieure”).

Note that the nonmeasurable set in Example 1.7.7 has inner measure 0 (otherwise E would contain a measurable set E_0 of positive measure, which gives disjoint sets $E_0 + r_n$ with equal positive measures). The following modification of this example produces an even more exotic set.

1.12.13. Example. The real line with Lebesgue measure λ contains a set E such that

$$\lambda_*(E) = 0 \quad \text{and} \quad \lambda^*(E \cap A) = \lambda(A) = \lambda^*(A \backslash E)$$

for any Lebesgue measurable set A. The same is true for the interval $[0,1]$.

PROOF. Similarly to Example 1.7.7, we find a set E_0 containing exactly one representative from every equivalence class for the following equivalence relation: $x \sim y$ if $x - y = n + m\sqrt{2}$, where $m, n \in \mathbb{Z}$. Set

$$E = \Bigl\{e + 2n + m\sqrt{2}\colon\ e \in E_0, m, n \in \mathbb{Z}\Bigr\}.$$

In the case of the interval we consider the intersection of E with $[0,1]$. Let $A \subset E$ be a measurable set. Note that the set $A - A = \bigl\{a_1 - a_2\colon\ a_1, a_2 \in A\bigr\}$ contains no points of the form $2n+1+m\sqrt{2}$ with integer n and m. Therefore, $A - A$ contains no intervals, hence $\lambda(A) = 0$ (see Exercise 1.12.62). Thus, $\lambda_*(E) = 0$. We observe that the complement of E coincides with $E + 1$ (in the case of $[0,1]$ one has $[0,1]\backslash E \subset (E+1) \cup (E-1)$). Indeed, the difference between any point x and its representative in E_0 is a number of the form $n + m\sqrt{2}$. Hence $x = e + n + m\sqrt{2}$ is either in E (if n is even) or in $E+1$. On the other hand, $E \cap (E+1) = \varnothing$, since E_0 contains only one representative from every class. Therefore, the complement of E has inner measure 0. This means that $\lambda^*(A \cap E) = \lambda(A)$ for any Lebesgue measurable set A, since

$$\lambda^*(A \cap E) = \lambda(A) - \lambda_*\bigl(A\backslash(A \cap E)\bigr) = \lambda(A) - \lambda_*(A\backslash E),$$

where the number $\lambda_*(A\backslash E)$ does not exceed the inner measure of the complement of E, i.e., equals zero. Similarly, $\lambda^*(A\backslash E) = \lambda(A)$. $\square$

1.12(v). Extensions of measures

The next result shows that one can always extend a measure whose domain does not coincide with the class of all subsets of the given space. It follows that a measure has no maximal countably additive extension unless it can be extended to all subsets.

1.12.14. Theorem. *Let μ be a finite nonnegative measure on a σ-algebra $\mathcal{B}$ in a space X and let S be a set such that $\mu_*(S) = \alpha < \mu^*(S) = \beta$, where $\mu_*(S) = \sup\bigl\{\mu(B)\colon\ B \subset S, B \in \mathcal{B}\bigr\}$. Then, for any $\gamma \in [\alpha, \beta]$, there exists a countably additive measure ν on the σ-algebra $\sigma(\mathcal{B} \cup S)$ generated by $\mathcal{B}$ and S such that $\nu = \mu$ on $\mathcal{B}$ and $\nu(S) = \gamma$.*

PROOF. Suppose first that $\mu_*(S) = 0$ and $\mu^*(S) = \mu(X)$. We may assume that $\mu(X) = 1$. Set

$$\mathcal{E}_S = \Bigl\{E = (S \cap A) \cup \bigl((X\backslash S) \cap B\bigr)\colon\ A, B \in \mathcal{B}\Bigr\}. \qquad (1.12.5)$$

As we have seen in Example 1.2.7, $\mathcal{E}_S$ is the σ-algebra generated by S and $\mathcal{B}$. Now we set

$$\nu\Bigl((S \cap A) \cup \bigl((X\backslash S) \cap B\bigr)\Bigr) = \gamma\mu(A) + (1-\gamma)\mu(B).$$

Let us show that the set function ν is well-defined, i.e., if

$$E = (S \cap A) \cup \bigl((X\backslash S) \cap B\bigr) = (S \cap A_0) \cup \bigl((X\backslash S) \cap B_0\bigr),$$

where $A_0, B_0 \in \mathcal{B}$, then $\nu(E)$ does not depend on which of the two representations of E we use. To this end, it suffices to note that $\mu(A_0) = \mu(A)$ and $\mu(B_0) = \mu(B)$. Indeed, $A \cap S = A_0 \cap S$. Then the measurable sets $A\backslash A_0$ and $A_0\backslash A$ are contained in $X\backslash S$ and have measure zero, since $\mu^*(S) = \mu(X)$. Therefore, one has $\mu(A \bigtriangleup A_0) = 0$. Similarly we obtain $\mu(B \bigtriangleup B_0) = 0$, since $\mu^*(X\backslash S) = \mu(X)$ by the equality $\mu_*(S) = 0$. By construction we have $\nu(S) = \gamma\mu(X) = \gamma$. If $A = B \in \mathcal{B}$, then $\nu(B) = \gamma\mu(B) + (1-\gamma)\mu(B) = \mu(B)$.

Let us show that ν is a countably additive measure. Let E_n be pairwise disjoint sets in $\mathcal{E}_S$, generated by pairs of sets $(A_n, B_n) \in \mathcal{B}$ according to (1.12.5). Then the sets $A_n \cap S$ are pairwise disjoint. Therefore, if $n \neq k$, the measurable sets $A_n \cap A_k$ are contained in $X\backslash S$ and hence have measure zero. Therefore, $\mu\bigl(\bigcup_{n=1}^\infty A_n\bigr) = \sum_{n=1}^\infty \mu(A_n)$. Similarly, $\mu\bigl(\bigcup_{n=1}^\infty B_n\bigr) = \sum_{n=1}^\infty \mu(B_n)$. This shows that $\nu\bigl(\bigcup_{n=1}^\infty E_n\bigr) = \sum_{n=1}^\infty \nu(E_n)$. Thus, in the considered case the theorem is proven.

In the general case, let us take a measurable envelope $\widetilde{S}$ of the set S (see (1.12.4). Let $\underline{S}$ be a measurable kernel of S. Then $\mu(\underline{S}) = \mu_*(S) = \alpha$. Set

$$X_0 = \widetilde{S}\backslash\underline{S}, \quad S_0 = S\backslash\underline{S}.$$

The restriction of the measure μ to X_0 is denoted by μ_0. Note that we have $\mu_0^*(S_0) = \mu_0(X_0) = \beta - \alpha$ and $(\mu_0)_*(S_0) = 0$. According to the previous step, there exists a measure ν_0 on the space X_0 with the σ-algebra $\mathcal{E}_{S_0}$ generated by S_0 and all sets $B \in \mathcal{B}$ with $B \subset X_0$ such that $\nu_0(S_0) = \gamma - \alpha$ and ν_0 coincides with μ_0 on all sets $B \subset X_0$ in $\mathcal{B}$. The collection of all sets of the form

$$E = A \cup E_0 \cup B, \quad \text{where } A, B \in \mathcal{B}, A \subset X\backslash\widetilde{S}, B \subset \underline{S}, E_0 \in \mathcal{E}_{S_0},$$

is the σ-algebra $\mathcal{E}$ generated by S and $\mathcal{B}$. Let us consider the measure

$$\nu(E) = \mu(A) + \nu_0(E_0) + \mu(B).$$

It is readily seen that ν is a countably additive measure on $\mathcal{E}$ equal to μ on $\mathcal{B}$, and that $\nu(S) = \mu(\varnothing) + \nu_0(S_0) + \mu(\underline{S}) = \gamma - \alpha + \alpha = \gamma$.

It is easily verified that the formula

$$\nu(E) := \mu^*(E \cap S) + \mu_*\bigl(E \cap (X\backslash S)\bigr), \quad E \in \mathcal{E}_S,$$

gives an extension of the measure μ with $\nu(S) = \mu^*(S)$. The closely related Nikodym's approach is described in Exercise 3.10.37. $\square$

The assertion on existence of extensions can be generalized to arbitrary families of pairwise disjoint sets. For countable families of additional sets this is due to Bierlein [**89**]; the general case was considered in Ascherl, Lehn [**40**].

1.12.15. Theorem. *Let $(X, \mathcal{B}, \mu)$ be a probability space and let $\{Z_\alpha\}$ be a family of pairwise disjoint subsets in X. Then, there exists a probability measure ν that extends μ to the σ-algebra generated by $\mathcal{B}$ and $\{Z_\alpha\}$.*

PROOF. First we consider a countable family of pairwise disjoint sets Z_n. Let us choose measurable envelopes $\widetilde{Z}_n$ of the sets Z_n. Let

$$B_1 = \widetilde{Z}_1, \quad B_n = \widetilde{Z}_n \backslash \bigcup_{i=1}^{n-1} \widetilde{Z}_i, \quad n > 1.$$

Then the sets B_n belong to $\mathcal{B}$ and are disjoint. We shall show that the set $S = \bigcup_{n=1}^{\infty}(B_n \backslash Z_n)$ has inner measure zero. Note first that

$$\mu_*(B_n \backslash Z_n) \leq \mu_*(\widetilde{Z}_n \backslash Z_n) = 0$$

for all $n \geq 1$, since $B_n \subset \widetilde{Z}_n$. Now let $C \in \mathcal{B}$, $C \subset \bigcup_{n=1}^{\infty}(B_n \backslash Z_n)$. Then $\mu(C) = \sum_{n=1}^{\infty} \mu(C \cap B_n) = 0$, since $C \cap B_n \subset B_n \backslash Z_n$. Thus, $\mu_*(S) = 0$. By Theorem 1.12.14, there exists an extension of the measure μ to a countably additive measure ν_0 on the σ-algebra $\mathcal{A}$ generated by $\mathcal{B}$ and S such that $\nu_0(S) = 0$. Denote by ν the Lebesgue completion of ν_0. All subsets of the set S belong to $\mathcal{A}_{\nu_0}$ and the measure ν vanishes on them. In particular, $\nu(B_n \backslash Z_n) = 0$. Note that

$$Z_n \backslash B_n \subset \bigcup_{i=1}^{n-1} (B_i \backslash Z_i). \tag{1.12.6}$$

Indeed, if $x \in Z_n \backslash B_n$, then $x \in Z_n \bigcap \bigcup_{i=1}^{n-1} \widetilde{Z}_i \subset \widetilde{Z}_n \bigcap \bigcup_{i=1}^{n-1} B_i$. Then $x \in B_i$ for some $i < n$. Clearly, $x \notin Z_i$, since $Z_i \cap Z_n = \varnothing$. Hence $x \in B_i \backslash Z_i$. By (1.12.6) we obtain $\nu(Z_n \backslash B_n) = 0$. Thus, we have $\nu(B_n \triangle Z_n) = 0$, which means the ν-measurability of all sets Z_n.

In the case of an uncountable family we set

$$c = \sup\Bigl\{\mu_*(S)\colon\ S = \bigcup_{n=1}^{\infty} Z_{\alpha_n}\Bigr\},$$

where sup is taken over all countable subfamilies $\{Z_{\alpha_n}\}$ of the initial family of sets. By using the countable additivity of μ, it is readily verified that there exists a countable family $N = \{\alpha_n\}$ such that $\mu_*(S) = c$, where $S = \bigcup_{n=1}^{\infty} Z_{\alpha_n}$. According to the previous step, the measure μ extends to a countably additive measure ν_0 on the σ-algebra $\mathcal{A}$ generated by $\mathcal{B}$ and the sets Z_{α_n}. Denote by $\mathcal{E}$ the class of all sets of the form

$$E = A \triangle C, \quad \text{where } A \in \mathcal{A},\ C \subset \bigcup_{j=1}^{\infty} Z_{\beta_j},\ \beta_j \notin N.$$

It is readily verified that $\mathcal{E}$ is a σ-algebra. It is clear that $\mathcal{A} \subset \mathcal{E}$ (since one can take $C = \varnothing$) and that $Z_\alpha \in \mathcal{E}$ for all α (since for $\alpha \notin N$ one can take $A = \varnothing$). Finally, let $\nu(A \triangle C) := \nu_0(A)$. This definition is non-ambiguous, which follows from the above-established non-ambiguity of Definition 1.12.11. To this end, however, it is necessary to verify that if $E = A_1 \triangle C_1$ is another representation of the above form, then the set $A \triangle A_1$ has ν_0-measure zero. Since this set is contained in a countable union of the sets Z_{β_j}, $\beta_j \notin N$, we have to show that the set $Z = \bigcup_{j=1}^{\infty} Z_{\beta_j}$ has inner measure zero with respect to ν_0. This is not completely obvious: although Z has zero inner measure with respect to μ, in the process of extending a measure the inner measure may increase. In our case, however, this does not happen. Indeed, suppose that Z contains a set E of positive ν_0-measure. By the construction of ν_0 (the Lebesgue completion of the extension explicitly described above) it follows that for E one can take a set of the form $E = (A_1 \cap S) \cup \bigl(A_2 \cap (X\backslash S)\bigr)$, where $A_1, A_2 \in \mathcal{B}$, $S = \bigcup_{n=1}^{\infty}(B_n\backslash Z_{\alpha_n})$ with some sets $B_n \in \mathcal{B}$ constructed at the first step of our proof. We have $\nu_0(E) = \mu(A_2)$. Then, the set E and its subset $E_0 = A_2 \cap (X\backslash S)$ have equal ν_0-measures. Since the sets B_n are pairwise disjoint, the set $X\backslash S$ is the union of the sets $\bigcup_{n=1}^{\infty}(B_n \cap Z_{\alpha_n})$ and $X\backslash\bigcup_{n=1}^{\infty} B_n$. But A_2 does not meet the sets Z_{α_n}, for it is contained in Z. Therefore, we obtain $E_0 = A_2 \cap \bigl(X\backslash\bigcup_{n=1}^{\infty} B_n\bigr) \in \mathcal{B}$ and hence $\mu(E_0) = \nu_0(E_0) > 0$. This contradicts the equality $\mu_*(Z) = 0$. By the above reasoning we also obtain that ν is a countably additive measure that extends the measure ν_0, hence extends the measure μ as well. □

The question arises whether the assumption that the additional sets in the above theorem are disjoint is essential. Under the continuum hypothesis, there exists a countable family of sets $E_j \subset [0,1]$ such that Lebesgue measure has no extensions to a countably additive measure on a σ-algebra containing all E_j. This assertion goes back to Banach and Kuratowski [**57**], and its proof is found in Corollary 3.10.3. The same is true under Martin's axiom defined below in §1.12(x); see a short reasoning in Mauldin [**659**]. On the other hand, it is proved in Carlson [**168**] that if the system of axioms ZFC (the Zermelo–Fraenkel system with the axiom of choice) is consistent, then it remains consistent with the statement that Lebesgue measure is extendible to any σ-algebra obtained by adding any countable sequence of sets. For yet another extension result, see Exercise 1.12.149.

Generalizations of Theorem 1.12.15 are obtained in Weber [**1007**] and Lipecki [**616**], where disjoint collections are replaced by well-ordered collections.

In Chapter 7 we discuss extensions to σ-algebras not necessarily obtained by adding disjoint families.

1.12(vi). Some interesting sets

In this subsection, we consider several interesting examples of measurable and nonmeasurable sets on the real line.

1.12.16. Example. There exists a Borel set B on the real line such that, for every nonempty interval J, the sets $B \cap J$ and $(\mathbb{R}^1\backslash B) \cap J$ have positive measures.

PROOF. Let $\{I_n\}$ be all nondegenerate intervals in $[0,1]$ with rational endpoints. Let us find a nowhere dense compact set $A_1 \subset I_1$ of positive measure. The set $I_1\backslash A_1$ contains an interval, hence there is a nowhere dense compact set $B_1 \subset I_1\backslash A_1$ of positive measure. Similarly, there exist nowhere dense compact sets $A_2 \subset I_2\backslash(A_1 \cup B_1)$ and $B_2 \subset I_2\backslash(A_1 \cup B_1 \cup A_2)$ with $\lambda(A_2) > 0$ and $\lambda(B_2) > 0$. By induction, we construct in $[0,1]$ a sequence of pairwise disjoint nowhere dense compact sets A_n and B_n of positive measure such that $B_n \subset I_n\backslash A_n$. If A_i and B_i are already constructed for $i \le n$, the set $I_{n+1}\backslash\bigcup_{i=1}^n (A_i \cup B_i)$ contains some interval, since the union of finitely many nowhere dense compact sets is a nowhere dense compact set. In this interval one can find disjoint nowhere dense compact sets A_{n+1} and B_{n+1} of positive measure and continue our construction. Let $E = \bigcup_{n=1}^\infty B_n$. If we are given an interval in $[0,1]$, then it contains the interval I_m for some m. According to our construction, I_m contains sets A_{m+1} and B_{m+1}, i.e., the intersections of I_m with E and $[0,1]\backslash E$ have positive measures. Finally, let us set $B = \bigcup_{z=-\infty}^{+\infty}(E+z)$. □

Let us introduce several concepts and facts related to ordered sets and ordinal numbers. A detailed exposition of these issues (including the transfinite induction) is given in the following books: Dudley [**251**], Jech [**459**], Kolmogorov, Fomin [**536**], Natanson [**707**]. A set T is called partially ordered if it is equipped with a partial order, i.e., some pairs $(t,s) \in T\times T$ are linked by a relation $t \le s$ satisfying the conditions: 1) $t \le t$, 2) if $t \le s$ and $s \le u$, then $t \le u$ for all $s,t,u \in T$. Sometimes such a relation is called a partial pre-order, and the definition of a partial order includes the requirement of antisymmetry: if $t \le s$ and $s \le t$, then $t = s$. But we do not require this. We write $t < s$ if $t \le s$ and $t \ne s$. The set T is called linearly ordered if all its elements are pairwise comparable and, in addition, if $t \le s$ and $s \le t$, then $t = s$. An element m of a partially ordered set is called maximal if there is no element x with $x > m$. A minimal element is defined by analogy.

A set is called well-ordered if it is linearly ordered and every nonempty subset of it has a minimal element. For example, the sets $\mathbb{N}$ and $\mathbb{R}^1$ with their natural orderings are linearly ordered, $\mathbb{N}$ is well-ordered, but $\mathbb{R}^1$ is not.

The interval (α,β) in a well-ordered set M is defined as the set of all points x such that $\alpha < x < \beta$. A set of the form $\{x \in M\colon\ x < \alpha\}$ is called an initial interval in M (the point α is not included). The closed interval $[\alpha,\beta]$ is the interval (α,β) with the added endpoints. Two well-ordered sets are called order-isomorphic if there is a one-to-one order-preserving correspondence between them. A class of order-isomorphic well-ordered sets is called an ordinal number or an ordinal. Ordinal numbers corresponding to infinite sets are called transfinite numbers or transfinites. If we are given two well-ordered sets A and B that represent distinct ordinal numbers α and β, then either A is order-isomorphic to some initial interval in B, or B is order-isomorphic to some initial interval in A. In the first case, we write $\alpha < \beta$, and in the second $\beta < \alpha$. Thus, given any two distinct ordinals, one is less than the other. Any set consisting of ordinal numbers is also well-ordered (unlike subsets of $\mathbb{R}^1$ with their usual ordering). The set $W(\alpha)$ of all ordinal numbers less than α is a well-ordered set of the type α. If we are given a set X of cardinality κ, then by means of the axiom of choice it can be well-ordered (Zermelo's theorem), i.e., there exist ordinals corresponding to sets of cardinality κ. Therefore, among such ordinals there is the smallest one $\omega(\kappa)$. Similarly, one defines the smallest uncountable ordinal number ω_1 (the smallest ordinal number corresponding to an uncountable set), which is sometimes used in measure theory for constructing various exotic examples. The least uncountable cardinality is denoted by $\aleph_1$. The continuum hypothesis is the equality $\aleph_1 = \mathfrak{c}$. The first (i.e., the smallest) infinite ordinal is denoted by ω_0.

The next example is a typical application of well-ordered sets.

1.12.17. Example. There exists a set $B \subset \mathbb{R}$ (called the *Bernstein set*) such that this set and its complement have nonempty intersections with all uncountable closed subsets of the real line. The intersection of B with every set of positive Lebesgue measure is nonmeasurable.

PROOF. It is clear that there exist the continuum of closed sets on the real line (since the complement of any closed set is a countable union of intervals) and that the collection of all uncountable closed sets has cardinality of the continuum $\mathfrak{c}$. Let us employ the following fact: the set of all ordinal numbers smaller than $\omega(\mathfrak{c})$ (the first ordinal number corresponding to sets of cardinality of the continuum) has cardinality of the continuum $\mathfrak{c}$. Hence the set of all uncountable closed sets on the real line can be parameterized by infinite ordinal numbers less than $\omega(\mathfrak{c})$, and represented in the form $\{F_\alpha, \alpha < \omega(\mathfrak{c})\}$. By means of transfinite induction, in every F_α we can choose two points x_α and y_α such that all selected points are distinct. Indeed, the sets F_α can be well-ordered. By using that the set of indices α is well-ordered, we pick the first (in the sense of the established order) elements $x_1, y_1 \in F_1$ for the first element in the index set. If $1 < \alpha < \mathfrak{c}$ and pairwise distinct elements x_β, y_β are already found for all $\beta < \alpha$, we take for x_α, y_α the first elements in the set $F_\alpha \backslash \bigcup_{\beta<\alpha} \{x_\beta, y_\beta\}$, which is infinite, since F_α has cardinality of the continuum according to Exercise 1.12.111, and the cardinality of the set of indices not exceeding α has cardinality less than $\mathfrak{c}$. By the transfinite induction principle, elements x_α, y_α are defined for all $\alpha < \omega(\mathfrak{c})$. It remains to take $B = \{x_\alpha, \alpha < \omega(\mathfrak{c})\}$. It is clear that $y_\alpha \in \mathbb{R}\backslash B$ and $x_\alpha \in F_\alpha \cap B$, $y_\alpha \in F_\alpha \cap (\mathbb{R}\backslash B)$. The last claim is obvious from the fact that any set of positive measure contains a compact set of positive measure. $\square$

It will be shown in Chapter 6 (Corollary 6.7.13) that every uncountable Souslin set contains an uncountable compact subset. Hence the Bernstein set contains no uncountable Souslin subsets. This is employed in the following lemma.

1.12.18. Lemma. *Let T be a set of cardinality of the continuum and let $E \subset \mathbb{R} \times T$. Suppose that, for any $x \in \mathbb{R}$, the section $E_x = \{t\colon (x,t) \in E\}$ is finite and that, for any $T' \subset T$, the set $\{x\colon E_x \cap T' \neq \varnothing\}$ is Lebesgue measurable. Then, there exist a set Z of Lebesgue measure zero and an at most countable set $S \subset T$ such that $E_x \subset S$ for all $x \in \mathbb{R}\backslash Z$.*

Proof. Without loss of generality we may take for T a set of cardinality of the continuum such that it contains no uncountable Souslin subsets (for example, the Bernstein set). Note that there exists a Borel set N of measure zero such that the set $D := E \cap ((\mathbb{R}\backslash N) \times \mathbb{R})$ has the following property: for any open set U, the set $\{x\colon D_x \cap U \neq \varnothing\}$ is Borel. Indeed, let $\{U_n\}$ be the sequence of all intervals with rational endpoints. By hypothesis, we have $\{x\colon U_n \cap E_x \neq \varnothing\} = B_n \cup N_n$, where $B_n \in \mathcal{B}(\mathbb{R})$ and $\lambda(N_n) = 0$. We find measure zero Borel sets N_n' with $N_n \subset N_n'$ and put $N = \bigcup_{n=1}^{\infty} N_n'$. An arbitrary nonempty open set U is the union of finitely or countably many sets U_n. Hence in order to establish the indicated property of the set N, it suffices to verify that the sets $\{x\colon D_x \cap U_n \neq \varnothing\}$ are Borel. To this end, we observe that $\{x\colon D_x \cap U_n \neq \varnothing\} = B_n \cup N_n \backslash N = B_n \backslash N$. Let us now show that D is Borel. It follows from our assumption that the sets D_x are finite. Hence

$$D = \bigcap_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \Big\{(x,r)\colon\ |r - r_m| < 1/n,\ D_x \cap (r_m - 1/n, r_m + 1/n) \neq \varnothing\Big\},$$

where $\{r_m\}$ are all rational numbers. Indeed, the left-hand side of this relation always belongs to the right-hand side, and if (x,r) does not belong to D, then, for some n, we have $|r - t| > (2n)^{-1}$ for all t from the finite set D_x, hence (x,r) does not belong to the right-hand side of this relation. Thus, D is the countable intersection of countable unions of the sets

$$(r_m - 1/n, r_m + 1/n) \times \big\{x\colon\ D_x \cap (r_m - 1/n, r_m + 1/n) \neq \varnothing\big\},$$

which are Borel as shown above. Thus, D is a Borel set. Let S be the projection of D to the second factor. Then S is a Souslin set. According to our choice of T, the set S is at most countable. It is clear that N and S are as required. □

Now we can prove the following interesting result.

1.12.19. Theorem. *Let $\{A_t\}_{t\in T}$ be some family of measure zero sets covering the real line such that every point belongs only to finitely many of them. Then, there exists a subfamily $T' \subset T$ such that the set $\bigcup_{t\in T'} A_t$ is nonmeasurable.*

Proof. Let $E = \{(x,t)\colon\ t \in T, x \in A_t\}$. If, for each $T' \subset T$, the set $\bigcup_{t\in T'} A_t$ is measurable, then E satisfies the hypotheses of the above lemma. Hence there exist a measure zero set Z and an at most countable set $S \subset T$ such that $E_x \subset S$ for all $x \in \mathbb{R}^1 \backslash Z$. Then $\mathbb{R}^1 \backslash Z \subset \bigcup_{s\in S} A_s$, which is a contradiction. □

Let us recall that a Hamel basis (or an algebraic basis) in a linear space L is a collection of linearly independent vectors v_α such that every vector in L is a finite linear combination of v_α. If $\mathbb{R}$ is regarded as a linear space over the real field, then any nonzero vector serves as a basis. However, the situation changes if we regard $\mathbb{R}$ over the field $\mathbb{Q}$ of rational numbers: now there is no finite basis. But it is known (see Kolmogorov, Fomin [**536**]) that in this case there exists a Hamel basis as well and any basis has cardinality of the continuum. It is interesting that the metric properties of Hamel bases of the space $\mathbb{R}$ over $\mathbb{Q}$ may be very different.

1.12.20. Lemma. *Each Hamel basis of* $\mathbb{R}$ *over* $\mathbb{Q}$ *has inner Lebesgue measure zero, and there exist Lebesgue measurable Hamel bases.*

PROOF. Let H be a Hamel basis and $h \in H$. In the case $\lambda_*(H) > 0$, where λ is Lebesgue measure, the set H contains a compact set of positive measure. According to Exercise 1.12.62, the set $\{h_1 - h_2, h_1, h_2 \in H\}$ contains a nonempty interval. Hence there exist $h_1, h_2 \in H$ and nonzero $q \in \mathbb{Q}$ such that $h_1 - h_2 = qh$, which contradicts the linear independence of vectors of our basis over $\mathbb{Q}$.

In order to construct a measurable Hamel basis, we apply Exercise 1.12.61 and take two sets A and B of measure zero such that $\{a+b, a \in A, b \in B\} = \mathbb{R}$. Let $M = A \cup B$. Then M has measure zero. It remains to observe that there exists a Hamel basis consisting of elements of M. As in the proof of the existence of a Hamel basis, it suffices to take a set $H \subset M$ that is a maximal (in the sense of inclusion) linearly independent set over $\mathbb{Q}$. Then H is a Hamel basis, since the linear span of H over $\mathbb{Q}$ contains M, hence it equals $\mathbb{R}$. $\square$

1.12.21. Example. There exists a Lebesgue nonmeasurable Hamel basis of $\mathbb{R}$ over $\mathbb{Q}$.

PROOF. We give a proof under the assumption of the continuum hypothesis, although this hypothesis is not necessary (Exercise 1.12.66). Let us take any Hamel basis H. By using that it has cardinality of the continuum we can establish a one-to-one correspondence $\alpha \mapsto h_\alpha$ between ordinal numbers $\alpha < \mathfrak{c}$ and elements of H. For any $\alpha < \mathfrak{c}$ and any nonzero $q \in \mathbb{Q}$, we denote by $V_{\alpha,q}$ the collection of all numbers of the form $q_1 h_{\alpha_1} + \cdots + q_n h_{\alpha_n} + q h_\alpha$, where $q_i \in \mathbb{Q}$ and $\alpha_i < \alpha$. According to the continuum hypothesis, every set $V_{\alpha,q}$ is countable (since its cardinality is less than $\mathfrak{c}$), and their union gives $\mathbb{R}\backslash\{0\}$. Let us write $V_{\alpha,q}$ as a countable sequence $\{h^n_{\alpha,q}\}$ and, for every $k \in \mathbb{N}$, consider $M_{k,q} = \bigcup_{\alpha<\mathfrak{c}} h^k_{\alpha,q}$. If we prove that the sets $M_{k,q}$ are linearly independent, then they can be complemented to Hamel bases $H_{k,q}$. The union of the latter sets contains the union of the sets $M_{k,q}$ and hence equals $\mathbb{R}\backslash\{0\}$, whence it follows that a countable collection of bases $H_{k,q}$ contains nonmeasurable sets because they all have inner measure zero. For the proof of linear independence of $M_{k,q}$ we consider a collection of distinct elements $h^k_{\alpha_1,q}, \ldots, h^k_{\alpha_n,q} \in M_{k,q}$, where $\alpha_1 < \cdots < \alpha_n < \mathfrak{c}$. Let $q_1, \ldots, q_n \in \mathbb{Q}$ and let $j \geq 1$ be the maximum of the indices of nonzero q_i. The expansion of $q_j h^k_{\alpha_j,q}$ with respect to the basis H contains the element $q_j q h_{\alpha_j}$, whereas the expansions of all other $q_i h^k_{\alpha_i,q}$ do not involve h_{α_j}, whence it follows that $q_1 h^k_{\alpha_1,q} + \cdots + q_n h^k_{\alpha_n,q} \neq 0$. $\square$

The next example is a deep theorem due to Besicovitch; its compact proof can be found in Stein [**906**, Chapter X]. Let R be a rectangle in the plane with the longer side length 1. Denote by $\widetilde{R}$ its translation to 2 in the positive direction parallel to the longer side, i.e., if e is the unit vector in the right half-plane giving the direction of the longer side, then $\widetilde{R} = R+2e$. The known methods of constructing the Besicovitch set (see Stein [**906**]) are based on the following assertions.

1.12.22. Lemma. *For any $\varepsilon > 0$, there exist a number $N = N_\varepsilon \in \mathbb{N}$ and 2^N rectangles $R_1, \ldots, R_{2^N} \subset \mathbb{R}^2$ with the side lengths 1 and 2^{-N} such that $\lambda_2\bigl(\bigcup_{j=1}^{2^N} R_j\bigr) < \varepsilon$, and the above-defined rectangles $\widetilde{R}_j$ are pairwise disjoint, so that $\lambda_2\bigl(\bigcup_{j=1}^{2^N} \widetilde{R}_j\bigr) = 1$, where λ_2 is Lebesgue measure on $\mathbb{R}^2$.*

1.12.23. Lemma. *Let P be a parallelogram in the plane with two sides in the lines $y = 0$ and $y = 1$. Then, for any $\varepsilon > 0$, one can find a number $N = N_\varepsilon \in \mathbb{N}$ and N parallelograms $P_1, \ldots, P_N$ in P such that each of them has two sides in the lines $y = 0$ and $y = 1$, $\lambda_2\bigl(\bigcup_{i=1}^{N} P_i\bigr) < \varepsilon$, and every interval in P with the endpoints in the lines $y = 0$ and $y = 1$ can be parallely translated to one of P_i.*

1.12.24. Example. There exists a compact set $K \subset \mathbb{R}^2$ (the *Besicovitch set*) of measure zero such that, for any straight line l in $\mathbb{R}^2$, the set K contains a unit interval parallel to l.

PROOF. Consequently applying the previous lemma, we obtain a sequence of compact sets $K_1 \supset K_2 \supset \cdots \supset K_j \supset \cdots$, where K_1 is the square $0 \leq x, y \leq 1$, with the following properties: $\lambda_2(K_j) \leq 1/j$ and, for any closed interval I joining the horizontal sides of K_1, the set K_j contains a closed interval obtained by a parallel transport of I. The set $\bigcap_{j=1}^{\infty} K_j$ has measure zero and contains a parallel transport of every interval of length 1 whose angle with the axis of ordinates lies between $-\pi/4$ and $\pi/4$. The union of two sets of such a type is a desired compact set. $\square$

Sets of the indicated type give a solution to the so-called Kakeya problem: what is a minimal measure of a set that contains unit intervals in all directions? Concerning this problem, see Wolff [**1024**].

Kahane [**479**] considered the set F of all line segments joining the points of the compact set E in the interval $[0,1]$ of the axis of abscissas described in Exercise 1.12.155 and the points of the form $(-2x, 1)$, $x \in E$. This set has zero measure, but contains translations of line segments of unit length whose angles with the axis of ordinates fill in some interval, so that a suitable union of finitely many sets of this type is a Besicovitch set. It is possible to prove the existence of a Besicovitch type set without any explicit construction. A class of random Besicovitch sets is described in Alexander [**11**]. Körner [**542**] considered the set $\mathcal{P}$ of all compact subsets $P \subset [-1,1]\times[0,1]$ with the following two properties: (i) P is a union of line segments joining points of the interval $[-1,1]$ in the axis of abscissas and points of the interval $[0,1]$ in the axis of ordinates, (ii) P contains a translation of each line segment of unit length. It is shown that $\mathcal{P}$ is closed in the space $\mathcal{K}$ of all compact sets in the plane equipped with the Hausdorff metric, and the collection of all compact sets in $\mathcal{P}$ of measure zero is a second category set in $\mathcal{P}$, hence is not empty.

Finally, let us mention the following surprising example due to Nikodym. Its construction is quite involved and may be read in the books by Guzmán [**386**] and Falconer [**277**].

1.12.25. Example. There exists a Borel set $A \subset [0,1]\times[0,1]$ (the *Nikodym set*) of Lebesgue measure 1 such that, for every point $x \in A$, there exists a straight line l_x whose intersection with A is exactly the point x.

The Nikodym set is especially surprising in connection with Fubini's theorem discussed in Chapter 3; see also Exercise 3.10.59, where the discussion concerns interesting Davies sets that are related to the Nikodym set.

1.12(vii). Additive, but not countably additive measures

In this subsection, it is explained how to construct additive measures on σ-algebras that are not countably additive. Unlike our constructive example on an algebra, here one has to employ non-constructive methods based on the axiom of choice. More precisely, we need the following Hahn–Banach theorem, which is proven in courses on functional analysis by means of the axiom of choice (see Kolmogorov, Fomin [**536**]).

1.12.26. Theorem. *Let L be a real linear space and let p be a real function with the following properties*:

(a) $p(\alpha x) = \alpha p(x)$ *for all* $\alpha \geq 0$ *and* $x \in L$;

(b) $p(x+y) \leq p(x) + p(y)$ *for all* $x, y \in L$.

Suppose that L_0 is a linear subspace in L and that l is a linear function on L_0 such that $l(x) \leq p(x)$ for all $x \in L_0$. Then l extends to a linear function $\widehat{l}$ on all of L such that $\widehat{l}(x) \leq p(x)$ for all $x \in L$.

Functions p with properties (a) and (b) are called sublinear. If, in addition, $p(-x) = p(x)$, then p is called a seminorm. For example, the norm of a normed space (see Chapter 4) is sublinear. Let us give less trivial examples that are employed for constructing some interesting linear functions.

1.12.27. Example. The following functions p are sublinear:

(i) let L be the space of all bounded real sequences $x = (x_n)$ with its natural linear structure (the operations are defined coordinate-wise) and let

$$p(x) = \inf S(x, a_1, \ldots, a_n), \quad S(x, a_1, \ldots, a_n) := \sup_{k\geq 1} \frac{1}{n}\sum_{i=1}^n x_{k+a_i},$$

where inf is taken over all natural n and all finite sequences $a_1, \ldots, a_n \in \mathbb{N}$;

(ii) let L be the space of all bounded real functions on the real line with its natural linear structure and let

$$p(f) = \inf S(f, a_1, \ldots, a_n), \quad S(f, a_1, \ldots, a_n) := \sup_{t\in\mathbb{R}} \frac{1}{n}\sum_{i=1}^n f(t+a_i),$$

where inf is taken over all natural n and all finite sequences $a_1, \ldots, a_n \in \mathbb{R}$;

(iii) let L be the space of all bounded real functions on the real line and let

$$p(f) = \inf\left\{ \limsup_{t\to+\infty} \frac{1}{n}\sum_{i=1}^n f(t+a_i)\right\},$$

where inf is taken over all natural n and all finite sequences $a_1, \ldots, a_n \in \mathbb{R}$;

(iv) let L be the space of all bounded real sequences $x = (x_n)$ and let

$$p(x) = \inf S(x, a_1, \ldots, a_n), \quad S(x, a_1, \ldots, a_n) := \limsup_{k\to\infty} \frac{1}{n}\sum_{i=1}^n x_{k+a_i},$$

where inf is taken over all natural n and all finite sequences $a_1, \ldots, a_n \in \mathbb{N}$.

PROOF. Claim (i) follows from (ii). Let us show (ii). It is clear that $|p(f)| < \infty$ and $p(\alpha f) = \alpha p(f)$ if $\alpha \geq 0$. Let $f, g \in L$. Take $\varepsilon > 0$ and find $a_1, \ldots, a_n, b_1, \ldots, b_m$ such that

$$\sup_{t\in\mathbb{R}} \frac{1}{n}\sum_{i=1}^n f(t+a_i) < p(f) + \varepsilon, \quad \sup_{t\in\mathbb{R}} \frac{1}{m}\sum_{i=1}^m g(t+b_i) < p(g) + \varepsilon.$$

We observe that

$$\begin{aligned}&\sup_{t\in\mathbb{R}} \frac{1}{nm}\sum_{i=1}^n\sum_{j=1}^m (f+g)(t+a_i+b_j)\\ &\leq \sup_{t\in\mathbb{R}} \frac{1}{n}\sum_{i=1}^n \frac{1}{m}\sum_{j=1}^m f(t+a_i+b_j) + \sup_{t\in\mathbb{R}} \frac{1}{m}\sum_{j=1}^m \frac{1}{n}\sum_{i=1}^n g(t+a_i+b_j).\end{aligned}$$

For fixed t and b_j we have $n^{-1}\sum\limits_{i=1}^n f(t+a_i+b_j) \leq S(f, a_1, \ldots, a_n)$, whence it follows that

$$\sup_{t\in\mathbb{R}} \frac{1}{n}\sum_{i=1}^n \frac{1}{m}\sum_{j=1}^m f(t+a_i+b_j) \leq S(f, a_1, \ldots, a_n).$$

A similar estimate for g yields

$$p(f+g) \leq S(f, a_1, \ldots, a_n) + S(g, b_1, \ldots, b_m) < p(f) + p(g) + 2\varepsilon,$$

which shows that $p(f+g) \leq p(f) + p(g)$, since ε is arbitrary. The proof of (iii) is similar, and (iv) follows from (iii). $\square$

Let us now consider applications to constructing some interesting set functions.

1.12.28. Example. On the σ-algebra of all subsets in $\mathbb{N}$, there exists a nonnegative additive function ν that vanishes on all finite sets and equals 1 on $\mathbb{N}$; in particular, ν is not countably additive.

PROOF. Let us consider the space L of all bounded sequences with the function p from assertion (iv) in the previous example and take the subspace L_0 of all convergent sequences. Set $l(x) = \lim\limits_{n\to\infty} x_n$ if $x \in L_0$. Note that $l(x) = p(x)$, since for fixed a_i and n we have $\limsup\limits_{k\to\infty} n^{-1}\sum\limits_{i=1}^{n} x_{k+a_i} = \lim\limits_{k\to\infty} x_k$. Let us extend l to a linear function $\widehat{l}$ on L with $\widehat{l} \le p$. If $x \in L$ and $x_n \le 0$ for all n, then $p(x) \le 0$ and hence $\widehat{l}(x) \le 0$. Therefore, $\widehat{l}(x) \ge 0$ if $x_n \ge 0$. If $x = (x_1, \ldots, x_n, 0, 0, \ldots)$, then $\widehat{l}(x) = l(x) = 0$. Finally, $\widehat{l}(1, 1, \ldots) = 1$. For every set $E \subset \mathbb{N}$, let $\nu(E) := \widehat{l}(I_E)$, where I_E is the indicator of the set E, i.e., the sequence having in the nth position either 1 or 0 depending on whether n is in E or not. Finite sets are associated with finite sequences, hence ν vanishes on them. The value of ν on $\mathbb{N}$ is 1, and the additivity of ν follows by the additivity of $\widehat{l}$ and the fact that $I_{E_1\cup E_2} = I_{E_1} + I_{E_2}$ for disjoint E_1 and E_2. It is obvious that ν is not countably additive. □

The following assertion is justified in a similar manner (its proof is delegated to Exercise 2.12.102 in the next chapter because it is naturally related to the concept of the integral, although can be given without it).

1.12.29. Example. On the σ-algebra of all subsets in $[0, 1)$, there exists a nonnegative additive set function ζ that coincides with Lebesgue measure on all Lebesgue measurable sets and $\zeta(E + h) = \zeta(E)$ for all $E \subset [0, 1)$ and $h \in [0, 1)$, where in the formation of $E + h$ the sum $e + h \ge 1$ is replaced by $e + h - 1$.

If we do not require that the additive function ζ should extend Lebesgue measure, then there is a simpler example.

1.12.30. Example. There exists an additive nonnegative set function ζ defined on all bounded sets on the real line and invariant with respect to translations such that $\zeta\bigl([0, 1)\bigr) = 1$.

PROOF. Let L be the space of bounded functions on the real line with the sublinear function p from Example 1.12.27(ii). By the Hahn–Banach theorem, there exists a linear function l on L with $l(f) \le p(f)$ for all $f \in L$. Indeed, on $L_0 = 0$ we set $l_0(0) = 0$. Note that $l(-f) = -l(f) \le p(-f)$, whence

$$-p(-f) \le l(f) \le p(f), \quad \forall\, f \in L.$$

If $f \ge 0$, then $p(-f) \le 0$ by the definition of p, hence $l(f) \ge 0$. Next, $p(1) = 1$, $p(-1) = -1$, which gives $l(1) = 1$. It is clear that $|l(f)| \le \sup\limits_t |f(t)|$, since $p(f) \le \sup\limits_t |f(t)|$. Finally, for all $h \in \mathbb{R}^1$ we have $l(f) = l(f(\cdot + h))$ for each $f \in L$. Indeed, let $g(t) = f(t + h) - f(t)$. We verify that $l(g) = 0$. Let $h_k = (k - 1)h$ if $k = 1, \ldots, n + 1$. Then

$$p(g) \le S(g, h_1, \ldots, h_{n+1}) = \sup_t \frac{1}{n+1}[f(t + (n+1)h) - f(t)] \le \frac{2\sup\limits_s |f(s)|}{n+1},$$

which tends to zero as $n \to \infty$. Thus, $p(g) \le 0$. Similarly, we obtain the estimate $p(-g) \le 0$. Therefore, $l(g) = 0$. Now it remains to set $\zeta(A) = l(\overline{I}_A)$ for all $A \subset [0,1)$, where $\overline{I}_A$ is the 1-periodic extension of I_A to the real line. By the above-established properties of l we obtain a nonnegative additive set function on $[0,1)$ that is invariant with respect to translations within the set $[0,1)$. In addition, $\zeta\bigl([0,1)\bigr) = 1$, since $\overline{I}_{[0,1)} = 1$. For any bounded set A, we find n with $A \subset [-n,n)$ and set

$$\zeta(A) = \sum_{j=-n}^{n-1} \zeta\Bigl(\bigl(A \cap [j,j+1)\bigr) - j\Bigr).$$

It is readily verified that we obtain a desired function. □

We observe that ζ coincides with Lebesgue measure on all intervals.

1.12(viii). Abstract inner measures

Having considered Carathéodory outer measures, it is natural to turn to superadditive functions. In this subsection, we present some results in this direction.

A set function η defined on the family of all subsets in a space X and taking values in $[0,+\infty]$ is called an abstract inner measure if $\eta(\varnothing) = 0$ and:

(a) $\eta(A \cup B) \ge \eta(A) + \eta(B)$ for all disjoint A and B,

(b) $\eta\bigl(\bigcap_{n=1}^\infty A_n\bigr) = \lim\limits_{n\to\infty} \eta(A_n)$ for every decreasing sequence of sets such that $\eta(A_1) < \infty$,

(c) if $\eta(A) = \infty$, then, for every number c, there exists $B \subset A$ such that $c \le \eta(B) < \infty$.

It follows from (a) that $\eta\bigl(\bigcup_{n=1}^\infty E_n\bigr) \ge \sum_{n=1}^\infty \eta(E_n)$ for all pairwise disjoint sets E_n. In addition, $\eta(B) \le \eta(A)$ whenever $B \subset A$ because we have $\eta(A\backslash B) \ge 0$, i.e., η is monotone.

If μ is a nonnegative countably additive measure on a σ-algebra $\mathcal{A}$, then the function μ_* has properties (a) and (b), which is readily verified (one can either directly verify property (b) by using measurable kernels of the sets E_n or refer to the properties of μ^* and the equality $\mu_*(A) = \mu(X) - \mu^*(X\backslash A)$ for finite measures). For finite (or semifinite) measures μ property (c) is fulfilled, too. In fact, this property will be fulfilled for any measure if we define μ_* by

$$\mu_*(A) := \sup\bigl\{\mu(B)\colon\ B \subset A,\ B \in \mathcal{A},\ \mu(B) < \infty\bigr\}. \qquad (1.12.7)$$

Suppose that $\mathcal{F}$ is a family of subsets of a set X with $\varnothing \in \mathcal{F}$. Let $\tau\colon \mathcal{F} \to [0,+\infty]$ be a set function with $\tau(\varnothing) = 0$. We define the function τ_* on all sets $A \subset X$ by the formula

$$\tau_*(A) = \sup\Bigl\{\sum_{j=1}^\infty \tau(F_j)\colon\ F_j \in \mathcal{F},\ F_j \subset A \text{ are disjoint}\Bigr\}. \qquad (1.12.8)$$

Note that τ_* can also be defined by the formula

$$\tau_*(A) = \sup\biggl\{\sum_{j=1}^{n} \tau(F_j)\colon\ n \in \mathbb{N},\ F_j \in \mathcal{F},\ F_j \subset A \text{ are disjoint}\biggr\}. \quad (1.12.9)$$

This follows by the equality $\tau(\varnothing) = 0$. Note the following obvious estimate:

$$\tau_*(F) \geq \tau(F), \quad \forall\, F \in \mathcal{F}.$$

It is seen from the definition that τ_* is superadditive. Certainly, this function (as any other one) generates the class $\mathfrak{M}_{\tau_*}$ (see Definition 1.11.2) that is an algebra, on which τ_* is additive by Theorem 1.11.4. The question arises of the countable additivity of the function τ_* on this algebra and its relation to τ. Obviously, if $\tau\colon 2^X \to [0, +\infty]$ with $\tau(\varnothing) = 0$ is superadditive on the family of all sets, then $\tau_* = \tau$ because $\sum_{j=1}^{\infty} \tau(F_j) \leq \tau\bigl(\bigcup_{j=1}^{\infty} F_j\bigr) \leq \tau(A)$ for all pairwise disjoint sets $F_j \subset A$.

1.12.31. Proposition. (i) *Let τ be an abstract inner measure on a space X. Then $\mathfrak{M}_\tau$ is a σ-algebra and τ is countably additive on $\mathfrak{M}_\tau$.*

(ii) *Suppose that on a σ-algebra $\mathcal{A}$ we are given a measure μ with values in* $[0, +\infty]$. *Then, the function $\tau = \mu_*$ defined by* (1.12.7) *is an abstract inner measure and if the measure μ is finite, then the measure τ on the domain $\mathfrak{M}_\tau$ extends μ.*

Proof. (i) Under condition (b) the function τ is countably additive on the algebra $\mathfrak{M}_\tau$ by Theorem 1.11.4(ii) and this does not employ condition (a). Let us show that $\mathfrak{M}_\tau$ is a σ-algebra. For simplification of our reasoning we assume that τ has only finite values (the general case is similar and uses condition (c)). As noted above, condition (a) yields that $\tau(B) \leq \tau(A)$ if $B \subset A$, i.e., τ is monotone. Let $A_n \in \mathfrak{M}_\tau$ increase to A. For any $E \subset X$, by the monotonicity of τ and (b) we have

$$\tau(E \cap A) + \tau(E \backslash A) \geq \lim_{n\to\infty} \tau(E \cap A_n) + \lim_{n\to\infty} \tau(E \backslash A_n) = \tau(A).$$

Since (a) yields the converse, we obtain $A \in \mathfrak{M}_\tau$. Assertion (ii) has already been explained. Here one has $\mathcal{A} \subset \mathfrak{M}_{\mu_*}$ and if $\mu(X) < \infty$, then $\mu_*|_{\mathcal{A}} = \mu$. □

It should be noted that for a measure μ on an algebra $\mathcal{A}$ that is not a σ-algebra, the function μ_* may fail to have property (b). For example, this is the case for the usual length on the algebra $\mathcal{A}$ generated by intervals in $[0, 1]$: the set $\mathcal{R}$ of irrational numbers has inner measure 0 (evaluated, of course, by means of $\mathcal{A}$!) and is the intersection of a sequence of decreasing sets with finite complements and inner measures 1. However, inner measures are a very efficient tool for constructing and extending measures. Here and in the next subsection, we consider rather abstract examples whose real content is seen when dealing with inner compact regular set functions on topological spaces (see Chapter 7).

1.12.32. Proposition. *Let $\mathcal{F}$ be a family of subsets of a space X and let $\mu\colon \mathcal{F} \to [0,+\infty]$ be such that $\varnothing \in \mathcal{F}$ and $\mu(\varnothing) = 0$. Suppose that we have the identity*

$$\mu(A) = \mu_*(A \cap B) + \mu_*(A \backslash B), \quad \forall\, A, B \in \mathcal{F},$$

and that there exists a compact class $\mathcal{K}$ such that

$$\mu(A) \le \sup\{\mu_*(K)\colon\ K \in \mathcal{K},\, K \subset A\}, \quad \forall\, A \in \mathcal{F}.$$

Then:

(i) *the class $\mathfrak{M}_{\mu_*}$ is an algebra, $\mathcal{F} \subset \mathfrak{M}_{\mu_*}$, the function μ_* is countably additive on $\mathfrak{M}_{\mu_*}$ and coincides with μ on $\mathcal{F}$*;

(ii) $\lim\limits_{n\to\infty} \mu_*(A_n) = 0$ *if $A_n \subset X$, $A_n \downarrow \varnothing$ and $\mu_*(A_1) < \infty$.*

Proof. (i) It is clear that μ_* extends μ, since we can take $A = B$ in the above equality. According to Exercise 1.12.127, we have $\mathcal{F} \subset \mathfrak{M}_{\mu_*}$. By Theorem 1.11.4, the class $\mathfrak{M}_{\mu_*}$ is an algebra and μ_* is additive on $\mathfrak{M}_{\mu_*}$. The countable additivity will be established below.

(ii) Let $A_n \downarrow \varnothing$, $\mu_*(A_1) < \infty$ and $\varepsilon > 0$. We may assume that the class $\mathcal{K}$ is closed with respect to finite unions and countable intersections, passing to the smallest compact class $\widetilde{\mathcal{K}} \supset \mathcal{K}$ with such a property. Let us find $C_n \in \mathcal{K}$ with

$$C_n \subset A_n, \quad \mu_*(A_n) \le \mu_*(C_n) + \varepsilon 2^{-n-1}.$$

For this purpose we take a number $c \in \bigl(\mu_*(A_n) - \varepsilon 2^{-n-1}, \mu_*(A_n)\bigr)$ and find disjoint sets $F_1, \ldots, F_m \in \mathcal{F}$ such that $F_1 \cup \cdots \cup F_m \subset A_n$ and $c < \mu(F_1) + \cdots + \mu(F_m)$. Then we find $K_j \subset F_j$ such that $c < \mu(K_1) + \cdots + \mu(K_m)$ and take $C_n = K_1 \cup \cdots \cup K_m$. Similarly one verifies that there exist sets $M_n \in \mathfrak{M}_{\mu_*}$ with

$$M_n \subset C_n \quad \text{and} \quad \mu_*(C_n) \le \mu_*(M_n) + \varepsilon 2^{-n-1}.$$

It is easy to see that $\mu_*(A_n \backslash M_n) \le \varepsilon 2^{-n}$. One has $\bigcap_{n=1}^\infty C_n = \varnothing$, as $C_n \subset A_n$. Hence $\bigcap_{n=1}^k C_n = \varnothing$ for some k. By using the additivity of μ_* and the relation $\bigcap_{n=1}^k M_n \subset \bigcap_{n=1}^k C_n = \varnothing$, we obtain

$$\begin{aligned} \mu_*(A_n) &\le \mu_*(C_n) + \varepsilon 2^{-n-1} \le \mu_*(M_n) + \varepsilon 2^{-n} \\ &= \mu_*\Bigl(M_n \backslash \bigcap_{i=1}^k M_i\Bigr) + \varepsilon 2^{-n} \le \sum_{i=1}^k \mu_*(M_n \backslash M_i) + \varepsilon 2^{-n}. \end{aligned}$$

For $n > k \ge i$ we have

$$\mu_*(M_n \backslash M_i) \le \mu_*(A_n \backslash M_i) \le \mu_*(A_i \backslash M_i) \le \varepsilon 2^{-i},$$

whence we obtain $\mu_*(A_n) \le \varepsilon$.

It remains to show the countable additivity of μ_* on $\mathfrak{M}_{\mu_*}$. To this end, it suffices to verify that if $M, M_n \in \mathfrak{M}_{\mu_*}$ and $M \subset \bigcup_{n=1}^\infty M_n$, then $\mu_*(M) \le \sum_{n=1}^\infty \mu_*(M_n)$. Let $B_1 = M_1$ and $B_n = M_n \backslash (M_1 \cup \cdots \cup M_{n-1})$, $n > 1$. Then the sets $B_n \in \mathfrak{M}_{\mu_*}$ are disjoint and $M \subset \bigcup_{n=1}^\infty B_n$. Let $R_n = \bigcup_{j=n}^\infty B_j$.

Suppose that the series of $\mu_*(M_n)$ converges to $c < \infty$. If $\mu_*(M) > c$, then, for any $C \subset M$ with $\mu_*(C) > c$, we have $\mu_*(C \cap R_n) = \infty$. This follows from what has already been proven, since by Theorem 1.11.4 we have

$$\mu_*(C) = \sum_{n=1}^{\infty} \mu_*(C \cap B_n) + \lim_{n\to\infty} \mu_*(C \cap R_n),$$

and $C \cap R_n \downarrow \varnothing$. As shown above, one can find $C_0 \in \mathcal{K}$ with $C_0 \subset M$ and $\mu_*(C_0) > c$. Then $\mu_*(C_0 \cap R_1) = \infty$. By induction we construct $C_n \in \mathcal{K}$ such that $C_{n+1} \subset C_n \cap R_{n+1}$ and $\mu_*(C_n) > c$. This leads to a contradiction, since $C_n \downarrow \varnothing$ and hence for some p we have $C_p = C_1 \cap \cdots \cap C_p = \varnothing$, whereas one has $\mu_*(\varnothing) = 0$. $\square$

1.12.33. Theorem. *Let $\mathcal{K}$ be a compact class of sets in X that contains the empty set and is closed with respect to formation of finite unions and countable intersections, and let $\mu\colon \mathcal{K} \to [0, +\infty)$ be a set function satisfying the condition*

$$\mu(A) = \mu_*(A \cap B) + \mu_*(A \backslash B), \quad \forall\, A, B \in \mathcal{K},$$

or, which is equivalent, the condition

$$\mu(A) = \mu(A \cap B) + \sup\{\mu(K)\colon\ K \in \mathcal{K},\, K \subset A \backslash B\}, \quad \forall\, A, B \in \mathcal{K}.$$

Then:

(i) $\mathfrak{M}_{\mu_*}$ *is a σ-algebra and μ_* is countably additive on $\mathfrak{M}_{\mu_*}$ as a function with values in* $[0, +\infty]$;

(ii) $\mathcal{K} \subset \mathfrak{M}_{\mu_*}$ *and μ_* extends μ;*

(iii) $\mu_*(A) = \sup\{\mu(K)\colon\ K \subset A,\, K \in \mathcal{K}\}$ *for all $A \subset X$;*

(iv) $M \in \mathfrak{M}_{\mu_*}$ *precisely when $M \cap K \in \mathfrak{M}_{\mu_*}$ for all $K \in \mathcal{K}$;*

(v) $\lim\limits_{n\to\infty} \mu_*(A_n) = \mu_*(A)$ *if $A_n \downarrow A$ and $\mu_*(A_1) < \infty$.*

Proof. Since $\mu(\varnothing) = 2\mu_*(\varnothing)$, one has $\mu(\varnothing) = \mu_*(\varnothing) = 0$. By the above proposition with $\mathcal{F} = \mathcal{K}$ we obtain that $\mathfrak{M}_{\mu_*}$ is an algebra, on which μ_* is countably additive and (ii) is true. In particular, μ is additive on $\mathcal{K}$, which gives (iii) (this also follows by Exercise 1.12.124). Let us verify (v). Let $\varepsilon > 0$. By (iii) we can find $K_1 \subset A_1$ with $K_1 \in \mathcal{K}$ and $\mu_*(A_1) \le \mu(K_1) + \varepsilon/2$. By induction we construct sets $K_n \in \mathcal{K}$ with

$$K_n \subset A_n \cap K_{n-1}, \quad \mu_*(A_n \cap K_{n-1}) \le \mu(K_n) + \varepsilon 2^{-n}.$$

By using the decrease of A_j and the inclusion $\mathcal{K} \subset \mathfrak{M}_{\mu_*}$, we obtain

$$\begin{aligned}
\mu_*(A_{j+1}) + \mu(K_j) &\le \mu(K_{j+1}) + \mu_*(A_j \backslash K_j) + \mu(K_j) + \varepsilon 2^{-j-1} \\
&\le \mu(K_{j+1}) + \mu_*(A_j \backslash K_j) + \mu_*(A_j \cap K_j) + \varepsilon 2^{-j-1} \\
&\le \mu(K_{j+1}) + \mu_*(A_j) + \varepsilon 2^{-j-1}.
\end{aligned}$$

Set $K = \bigcap_{n=1}^{\infty} K_n$. Then $K \subset A$ and $K \in \mathcal{K} \subset \mathfrak{M}_{\mu_*}$. Since $K_n \backslash K \downarrow \varnothing$, by the above proposition we have $\mu_*(K_n \backslash K) \to 0$. Therefore,

$$\begin{aligned}
\mu_*(A_n) &= \mu_*(A_1) + \sum_{j=1}^{n-1} [\mu_*(A_{j+1}) - \mu_*(A_j)] \\
&\le \mu(K_1) + \frac{\varepsilon}{2} + \sum_{j=1}^{n-1} [\mu_*(K_{j+1}) - \mu_*(K_j) + \varepsilon 2^{-j-1}] \\
&\le \mu(K_n) + \varepsilon \le \mu_*(A) + \mu_*(K_n \backslash K) + \varepsilon.
\end{aligned}$$

Hence $\mu_*(A) \le \lim\limits_{n\to\infty} \mu_*(A_n) \le \mu_*(A)$.

Let us verify that $\mathfrak{M}_{\mu_*}$ is a σ-algebra. It suffices to show that if $M_n \in \mathfrak{M}_{\mu_*}$ and $M_n \downarrow M$, then $M \in \mathfrak{M}_{\mu_*}$. Let $A \subset X$. If $K \in \mathcal{K}$ and $K \subset A$, then

$$\mu(K) = \mu_*(K \cap M_n) + \mu_*(K \backslash M_n) \le \mu_*(K \cap M_n) + \mu_*(A \backslash M).$$

By using (v) and taking into account that μ is finite on $\mathcal{K}$, we obtain passing to the limit as $n \to \infty$ that

$$\mu(K) \le \mu_*(K \cap M) + \mu_*(A \backslash M) \le \mu_*(A \cap M) + \mu_*(A \backslash M).$$

According to (iii) we have $\mu_*(A) \le \mu_*(A \cap M) + \mu_*(A \backslash M)$. Since the reverse inequality is true as well, one has $M \in \mathfrak{M}_{\mu_*}$. Thus, (i) is established.

It remains to show (iv). Clearly, if $M \in \mathfrak{M}_{\mu_*}$ and $K \in \mathcal{K}$, then we have $K \cap M \in \mathfrak{M}_{\mu_*}$, since $\mathcal{K}$ belongs to the algebra $\mathfrak{M}_{\mu_*}$. Conversely, let $K \cap M \in \mathfrak{M}_{\mu_*}$ for all $K \in \mathcal{K}$. For every $A \subset X$, we have whenever $K \subset A$ and $K \in \mathcal{K}$

$$\begin{aligned}
\mu(K) &= \mu_*\bigl(K \cap (M \cap K)\bigr) + \mu_*\bigl(K \backslash (M \cap K)\bigr) \\
&\le \mu_*(A \cap M) + \mu_*(A \backslash M) \le \mu_*(A).
\end{aligned}$$

Taking sup over K we obtain by (iii) that $M \in \mathfrak{M}_{\mu_*}$.

If we have the second condition of the theorem, then $\mu(\varnothing) = 0$, whence $\mu(A) = \sup\{\mu(K)\colon\ K \in \mathcal{K}, K \subset A\}$ if $A \in \mathcal{K}$. Hence $\mu(B \cup C) = \mu(B) + \mu(C)$ if $B, C \in \mathcal{K}$, $B \cap C = \varnothing$. Hence μ_* coincides with μ on $\mathcal{K}$. So we have (iii) and the first condition of the theorem. The converse is true as well. □

The proof of the next theorem, which can be read in Fremlin [**327**, §413], combines the functions ν_* and ν^*.

1.12.34. Theorem. *Let $\mathcal{R}$ be a ring of subsets of a space X, let $\mathcal{K}$ be some class of subsets of X closed with respect to formation of finite intersections and finite disjoint unions, and let ν be a finite nonnegative additive function on $\mathcal{R}$ such that $\mathcal{K}$ is an approximating class for ν. Then the following assertions are true.*

(i) *If every element of $\mathcal{K}$ is contained in an element of $\mathcal{R}$, then ν extends to a finite nonnegative additive function $\widetilde{\nu}$ defined on a ring $\widetilde{\mathcal{R}}$ that contains $\mathcal{R}$ and $\mathcal{K}$, such that $\mathcal{K}$ is an approximating class for $\widetilde{\nu}$ and, for each $R \in \widetilde{\mathcal{R}}$ and $\varepsilon > 0$, there exists $R_\varepsilon \in \mathcal{R}$ with $\widetilde{\nu}(R \triangle R_\varepsilon) < \varepsilon$.*

(ii) *If $\mathcal{R}$ a σ-algebra, ν is countably additive, and $\mathcal{K}$ admits countable intersections, then ν extends to a measure $\widetilde{\nu}$ defined on a σ-algebra $\mathcal{A}$ containing $\mathcal{R}$ and $\mathcal{K}$, such that $\mathcal{K}$ remains an approximating class for $\widetilde{\nu}$ and, for each $R \in \mathcal{R}$, there exists $A \in \mathcal{A}$ with $\widetilde{\nu}(R \triangle A) = 0$.*

It is readily seen that unlike superadditive functions, a subadditive function $\mathfrak{m}$ may not be monotone, i.e., may not satisfy the condition $\mathfrak{m}(A) \leq \mathfrak{m}(B)$ whenever $A \subset B$. A *submeasure* is a finite nonnegative monotone subadditive function $\mathfrak{m}$ on an algebra $\mathfrak{A}$ such that $\mathfrak{m}(\varnothing) = 0$. A submeasure $\mathfrak{m}$ is called exhaustive if, for each sequence of disjoint sets $A_n \in \mathfrak{A}$, one has the equality $\lim\limits_{n\to\infty} \mathfrak{m}(A_n) = 0$. A submeasure $\mathfrak{m}$ is called uniformly exhaustive if, for each $\varepsilon > 0$, there exists n such that, in every collection of disjoint sets $A_1, \ldots, A_n \in \mathfrak{A}$, there exists A_i with $\mathfrak{m}(A_i) < \varepsilon$. Clearly, a uniformly exhaustive submeasure is exhaustive. A submeasure $\mathfrak{m}$ is called Maharam if $\lim\limits_{n\to\infty} \mathfrak{m}(A_n) = 0$ as $A_n \downarrow \varnothing$, $A_n \in \mathfrak{A}$. Recently, Talagrand [**932**] has constructed a counter-example to a long-standing open problem (the so-called control measure problem) that asked whether for every Maharam submeasure $\mathfrak{m}$ on a σ-algebra $\mathfrak{A}$, there exists a finite nonnegative measure μ with the same class of zero sets as $\mathfrak{m}$. It is known that this problem is equivalent to the following one: is every exhaustive submeasure uniformly exhaustive? Thus, both questions are answered negatively.

1.12(ix). Measures on lattices of sets

In applications one often encounters set functions defined not on algebras or semirings, but on lattices of sets. The results in this subsection are employed in Chapter 10 in our study of disintegrations.

1.12.35. Definition. *A class $\mathfrak{R}$ of subsets in a space X is called a lattice of sets if it contains the empty set and is closed with respect to finite intersections and unions.*

Unlike an algebra, a lattice may not be closed under complementation. Typical examples are: (a) the collection of all compact sets in a topological space X, (b) the collection of all open sets in a given space X. Sometimes in the definition of a lattice it is required that $X \in \mathfrak{R}$. Certainly, this can be always achieved by simply adding X to $\mathfrak{R}$, which does not affect the stability with respect to formation of unions and intersections.

A finite nonnegative set function β on a lattice $\mathfrak{R}$ is called modular if one has $\beta(\varnothing) = 0$ and

$$\beta(R_1 \cup R_2) + \beta(R_1 \cap R_2) = \beta(R_1) + \beta(R_2), \quad \forall\, R_1,\, R_2 \in \mathfrak{R}. \tag{1.12.10}$$

If in (1.12.10) we replace the equality sign by "$\leq$", then we obtain the definition of a submodular function, and the change of "$=$" to "$\geq$" gives the definition of a supermodular function. If $\mathfrak{R}$ is an algebra, then the modular functions are precisely the additive ones. We recall that a set function β is called *monotone* if $\beta(R_1) \leq \beta(R_2)$ whenever $R_1 \subset R_2$.

1.12.36. Proposition. *Let β be a monotone submodular function on a lattice $\mathfrak{R}$ and $X \in \mathfrak{R}$. Then, there exists a monotone modular function α on $\mathfrak{R}$ such that $\alpha \le \beta$ and $\alpha(X) = \beta(X)$.*

The proof is delegated to Exercise 1.12.148.

1.12.37. Corollary. *Suppose that β is a monotone supermodular function on a lattice $\mathfrak{R}$ and $X \in \mathfrak{R}$. Then, there exists a monotone modular function γ on $\mathfrak{R}$ such that $\gamma \ge \beta$ and $\gamma(X) = \beta(X)$.*

Proof. Let us consider the set function

$$\beta_0(C) = \beta(X) - \beta(X\backslash C)$$

on the lattice $\mathfrak{R}_0 = \{C\colon\ X\backslash C \in \mathfrak{R}\}$. It is readily verified that β_0 is monotone and submodular. According to the above proposition, there exists a monotone modular function α_0 on $\mathfrak{R}_0$ with $\alpha_0 \le \beta_0$ and $\alpha_0(X) = \beta_0(X)$. Now set $\gamma(R) = \alpha_0(X) - \alpha_0(X\backslash R)$, $R \in \mathfrak{R}$. Then $\gamma(X) = \beta(X)$ and $\gamma(R) \ge \beta(R)$, since $\alpha_0(X\backslash R) \le \beta_0(X\backslash R)$. □

1.12.38. Lemma. *Let β be a monotone modular set function on a lattice $\mathfrak{R}$, $X \in \mathfrak{R}$, and $\beta(X) = 1$. Then, there exists a monotone modular set function ζ on $\mathfrak{R}$ such that $\beta \le \zeta$, $\zeta(X) = 1$, and*

$$\zeta(R) + \zeta_*(X\backslash R) = 1, \quad \forall\, R \in \mathfrak{R}. \tag{1.12.11}$$

Proof. The set Ψ of all monotone modular set functions ψ on $\mathfrak{R}$ satisfying the conditions $\psi(X) = 1$ and $\psi \ge \beta$, is partially ordered by the relation $\le$. Each linearly ordered part of Ψ has an upper bound in Ψ given as the supremum of that part (this upper bound is modular, since the considered part is linearly ordered). By Zorn's lemma Ψ has a maximal element ζ. Corollary 1.12.37 yields (1.12.11), since otherwise the function ζ is not maximal. To see this, it suffices to show that for any fixed $R_0 \in \mathfrak{R}$, there is a function $\psi \in \Psi$ such that $\psi(R_0) + \psi_*(X\backslash R_0) = 1$. Let

$$\tau_1(R) := \sup\{\beta(R \cap S)\colon\ S \in \mathfrak{R},\, S \cap R_0 = \varnothing\}, \quad R \in \mathfrak{R}.$$

The function τ_1 is modular. Indeed, given $R_1, R_2 \in \mathfrak{R}$, for every $\varepsilon > 0$, one can find $S_i \in \mathfrak{R}$, $i = 1, \ldots, 4$, such that $S_i \cap R_0 = \varnothing$ and the sum of the quantities $\tau_1(R_1) - \beta(R_1 \cap S_1)$, $\tau_1(R_2) - \beta(R_2 \cap S_2)$, $\tau_1(R_1 \cap R_2) - \beta(R_1 \cap R_2 \cap S_3)$, $\tau_1(R_1 \cup R_2) - \beta\bigl((R_1 \cup R_2) \cap S_4\bigr)$ is less than ε. The same estimate holds if we replace all S_i by $S := S_1 \cup \cdots \cup S_4$. Then $\beta(R_1 \cap S) + \beta(R_2 \cap S)$ equals $\beta(R_1 \cap R_2 \cap S) + \beta\bigl((R_1 \cup R_2) \cap S\bigr)$, since β is modular and $(R_1 \cup R_2) \cap S = (R_1 \cap S) \cup (R_2 \cap S)$. The function $\beta - \tau_1$ is modular and monotone as well, which is seen from the fact that if $R_1 \subset R_2$, $R_i \in \mathfrak{R}$ and $S \in \mathfrak{R}$, then

$$\beta(R_1) + \beta(R_2 \cap S) = \beta(R_1 \cap S) + \beta\bigl(R_1 \cup (R_2 \cap S)\bigr) \le \beta(R_1 \cap S) + \beta(R_2).$$

Let

$$\tau_2(R) := \sup\{\beta(S) - \tau_1(S)\colon\ S \in \mathfrak{R},\, S \cap R_0 \subset R\}, \quad R \in \mathfrak{R}.$$

It is readily verified that the function τ_2 is monotone and supermodular. By the above corollary there exists a monotone modular function τ_3 on $\mathfrak{R}$ with

$\tau_3 \geq \tau_2$ and $\tau_3(X) = \tau_2(X) = 1 - \tau_1(X)$. Let $\psi = \tau_1 + \tau_3$. The function ψ is monotone and modular. For all $R \in \mathfrak{R}$, we have $\psi(R) \geq \tau_1(R) + \tau_2(R) \geq \beta(R)$, since $\tau_2(R) \geq \beta(R) - \tau_1(R)$. Finally, by the monotonicity of $\beta - \tau_1$ one has

$$\psi(R_0) \geq \tau_2(R_0) = \beta(X) - \tau_1(X) \geq 1 - \psi_*(X\backslash R_0).$$

Since $\psi(R_0) + \psi_*(X\backslash R_0) \leq 1$, we obtain the required equality. □

1.12.39. Corollary. *Suppose that in the proven lemma $\mathfrak{R}$ is a compact class closed with respect to formation of countable intersections. Set*

$$\mathcal{E} = \{E \subset X\colon\ \zeta_*(E) + \zeta_*(X\backslash E) = 1\}.$$

Then $\mathcal{E}$ is a σ-algebra and the restriction of ζ_* to $\mathcal{E}$ is countably additive.

Proof. Let us show that $\mathcal{E} = \mathfrak{M}_{\zeta_*}$. Let $E \in \mathcal{E}$ and $A \subset X$. Then $\zeta_*(A) \geq \zeta_*(A \cap E) + \zeta_*(A\backslash E)$. Let us verify the reverse inequality. Let $\varepsilon > 0$. We can find $R_1, R_2, R_3 \in \mathfrak{R}$ such that $R_1 \subset A$, $R_2 \subset E$, $R_3 \subset X\backslash E$ and $\zeta_*(A) \leq \zeta(R_1) + \varepsilon$, $\zeta_*(E) \leq \zeta(R_2) + \varepsilon$, $\zeta_*(X\backslash E) \leq \zeta(R_3) + \varepsilon$. Then $\zeta_*(A \cap E) \geq \zeta(R_1 \cap R_2)$, $\zeta_*(A\backslash E) \geq \zeta(R_1 \cap R_3)$. Since $\zeta(R_2) + \zeta(R_3) \geq 1 - 2\varepsilon$, by the modularity of ζ we obtain

$$\begin{aligned}\zeta_*(A \cap E) + \zeta_*(A\backslash E) &\geq \zeta(R_1 \cap R_2) + \zeta(R_1 \cap R_3) = \zeta\bigl(R_1 \cap (R_2 \cup R_3)\bigr)\\ &= \zeta(R_1) + \zeta(R_2 \cup R_3) - \zeta(R_1 \cup R_2 \cup R_3) \geq \zeta(R_1) - 2\varepsilon.\end{aligned}$$

Hence $E \in \mathfrak{M}_{\zeta_*}$. By Theorem 1.11.4 we obtain our assertion. □

1.12(x). Set-theoretic problems in measure theory

We have already seen that constructions of nonmeasurable sets involve certain set-theoretic axioms such as the axiom of choice. The question arises whether this is indispensable and what the situation is in the framework of the naive set theory without the axiom of choice. In addition, one might also ask the following question: even if there exist sets that are nonmeasurable in the Lebesgue sense, is it possible to extend Lebesgue measure to a countably additive measure on all sets (i.e., not necessarily by means of the Lebesgue completion and not necessarily with the property of the translation invariance)? Here we present a number of results in this direction. First, by admitting the axiom of choice, we consider the problem of the existence of nontrivial measures defined on all subsets of a given set, and then several remarks are made on the role of the axiom of choice.

Let X be a set of cardinality $\aleph_1$, i.e., X is equipotent to the set of all ordinal numbers that are smaller than the first uncountable ordinal number. Note that X is uncountable and can be well-ordered in such a way that every element is preceded by an at most countable set of elements. The following theorem is due to Ulam [**967**].

1.12.40. Theorem. *If a finite countably additive measure μ is defined on all subsets of the set X of cardinality $\aleph_1$ and vanishes on all singletons, then it is identically zero.*

PROOF. It suffices to consider only nonnegative measures (see §3.1 in Chapter 3). By hypothesis, X can be well-ordered in such a way that, for every y, the set $\{x\colon\, x < y\}$ is at most countable. There is an injective mapping $x \mapsto f(x,y)$ of this set into $\mathbb{N}$. Thus, for every pair (x,y) with $x<y$ one has a natural number $f(x,y)$. For every $x \in X$ and every natural n, we have the set

$$A_x^n = \{y \in X\colon\; x<y, f(x,y)=n\}.$$

For fixed n, the sets A_x^n, $x\in X$, are pairwise disjoint. Indeed, let $y \in A_x^n \cap A_z^n$, where $x \neq z$. We may assume that $x<z$. This is, however, impossible, since $x<y$, $z<y$ and hence $f(x,y)\neq f(z,y)$ by the injectivity of the function $f(\,\cdot\,,y)$. Therefore, by the countable additivity of the measure, for every n, there can be an at most countable set of points x such that $\mu(A_x^n)>0$. Since X is uncountable, there exists a point $x\in X$ such that $\mu(A_x^n)=0$ for all n. Hence $A=\bigcup_{n=1}^\infty A_x^n$ has measure zero. It remains to observe that the set $X\backslash A$ is at most countable, since it is contained in the set $\{y\colon\, y\le x\}$, which is at most countable by hypothesis. Indeed, if $y>x$, then $y\in A_x^n$, where $n=f(x,y)$. Therefore, $\mu(X\backslash A)=0$, which completes the proof. $\square$

Another proof will be given in Corollary 3.10.3 in Chapter 3.

We recall that one of the forms of the continuum hypothesis is the assertion that the cardinality of the continuum $\mathfrak{c}$ equals $\aleph_1$.

1.12.41. Corollary. *Assume the continuum hypothesis. Then, any finite countably additive measure that is defined on all subsets of a set of cardinality of the continuum and vanishes on all singletons is identically zero.*

One more set-theoretic axiom employed in this circle of problems is called Martin's axiom. A topological space X is said to satisfy the countable chain condition if every disjoint family of its open subsets is at most countable. Martin's axiom (MA) can be introduced as the assertion that, in every nonempty compact space satisfying the countable chain condition, the intersection of less than $\mathfrak{c}$ open dense sets is not empty. The continuum hypothesis (CH) is equivalent to the same assertion valid for all compacts (not necessarily satisfying the countable chain condition). Thus, CH implies MA. It is known that each of the axioms CH, MA and MA–CH (Martin's axiom with the negation of the continuum hypothesis) is consistent with the system of axioms ZFC (this is the notation for the Zermelo–Fraenkel system with the axiom of choice), i.e., if ZFC is consistent, then it remains consistent after adding any of these three axioms. In this book, none of these axioms is employed in main theorems, but sometimes they turn out to be useful for constructing certain exotic counter-examples or play some role in the situations where one is concerned with the validity of certain results in their maximal generality. Concerning the continuum hypothesis and Martin's axiom, see Jech [**458**], Kuratowski, Mostowski [**555**], Fremlin [**323**], Sierpiński [**879**].

Ulam's theorem leads to the notion of a measurable cardinal. For brevity, cardinal numbers are called cardinals. A cardinal κ is called *real measurable* if there exist a space of cardinality κ and a probability measure ν defined on the family of all its subsets and vanishing on all singletons. If ν assumes the values 0 and 1 only, then κ is called *two-valued measurable.* Real nonmeasurable cardinals (i.e., the ones that are not real measurable) are called Ulam numbers. The terminology here is opposite to the one related to the measurability of sets or functions: nonmeasurable cardinals are "nice". It is clear that the countable cardinality is nonmeasurable. Since every cardinal less than a nonmeasurable one is nonmeasurable as well, the nonmeasurable cardinals form some initial interval in the "collection of all cardinal numbers" (possibly embracing all cardinals as seen from what is said below). Anyway, this "interval" is very large, which is clear from the following Ulam–Tarski theorem (for a proof, see Federer [**282**, §2.1], Kharazishvili [**507**]).

1.12.42. Theorem. (i) *If a cardinal β is the immediate successor of a nonmeasurable cardinal α, then β is nonmeasurable.* (ii) *If the cardinality of a set M of nonmeasurable cardinals is nonmeasurable, then the supremum of M is nonmeasurable as well.*

A cardinal κ is called inaccessible if the class of all smaller cardinal numbers has no maximal element and there is no subset of cardinality less than κ whose supremum equals κ. The previous theorem means that if there exist measurable cardinals, then the smallest one is inaccessible. The cardinal $\aleph_1$ in Theorem 1.12.40 is the successor of the countable cardinal $\aleph_0$, which makes it nonmeasurable. The two-valued nonmeasurability of cardinality $\mathfrak{c}$ of the continuum is proved without use of the continuum hypothesis, which follows from Exercise 1.12.108 or from the following result (see Jech [**459**], Kuratowski, Mostowski [**555**, Ch. IX, §3], Kharazishvili [**507**]).

1.12.43. Proposition. *If a cardinal κ is two-valued nonmeasurable, then so is the cardinal 2^κ.*

This proposition yields that the cardinal $\mathfrak{c}$ is not two-valued measurable. Martin's axiom implies that the cardinal $\mathfrak{c}$ is not real measurable. If $\mathfrak{c}$ is not real measurable, then real measurable and two-valued measurable cardinals coincide. The following theorem (see Jech [**459**]) summarizes the basic facts related to measurable cardinals.

1.12.44. Theorem. *The supposition that measurable cardinals do not exist is consistent with the ZFC. In addition, if either of the following assertions is consistent with the ZFC, then so are all of them*:

(i) *two-valued measurable cardinals exist*;

(ii) *real measurable cardinals exist*;

(iii) *the cardinal $\mathfrak{c}$ is real measurable*;

(iv) *Lebesgue measure can be extended to a measure on the σ-algebra of all subsets in* $[0, 1]$.

Nonmeasurable cardinals will be encountered in Chapter 7 in our discussion of supports of measures in metric spaces. Some additional information about measurable and nonmeasurable cardinals can be found in Buldygin, Kharazishvili [**142**], Kharazishvili [**506**], [**507**], [**508**], [**511**], Fremlin [**323**], [**325**], Jech [**459**], Solovay [**898**].

We recall that the axiom of choice does not exclude countably additive extensions of Lebesgue measure to all sets, but only makes impossible the existence of such extensions with the property of translation invariance (in the next subsection there are remarks on invariant extensions), in particular, it does not enable one to exhaust all sets by means of the Lebesgue completion.

It is now natural to discuss what happens if we restrict the use of the axiom of choice. It is reasonable to admit the countable form of the axiom of choice, i.e., the possibility of choosing representatives from any countable collection of nonempty sets. At least, without it, there is no measure theory, nor even the theory of infinite series (see Kanovei [**490**]). It turns out that if we permit the use of the countable form of the axiom of choice, then, as shown by Solovay [**897**], there exists a model of set theory such that all sets on the real line are Lebesgue measurable (see also Jech [**458**, §20]). Certainly, the full axiom of choice is excluded here. Another interesting related result deals with the so-called axiom of determinacy. For the formulation, we have to define the following game G_A of two players I and II, associated with every set A consisting of infinite sequences $a = (a_0, a_1, , \ldots)$ of natural numbers a_n. The game proceeds as follows. Player I writes a number $b_0 \in \mathbb{N}$, then player II writes a number $b_1 \in \mathbb{N}$ and so on; the players know all the previous moves. If the obtained sequence $b = (b_0, b_1, \ldots)$ belongs to A, then I wins, otherwise II wins. The set A and game G_A are called determined if one of the players I or II has a winning strategy (i.e., a rule to make steps corresponding to the steps of the opposite side leading to victory). For example, if A consists of a single sequence $a = (a_i)$, then II has a winning strategy: it suffices to write $b_1 \neq a_1$ at the very first move. The axiom of determinacy (AD) is the statement that every set $A \subset \mathbb{N}^\infty$ is determined. In Kanovei [**490**] one can find interesting consequences of the axiom of determinacy, of which the most interesting for us are the measurability of all sets of reals (see also Martin [**657**]) and the real measurability of the cardinal $\aleph_1$. Thus, on the one hand, the axiom of determinacy excludes some paradoxical sets, but, on the other hand, it gives some objects impossible under the full axiom of choice.

1.12(xi). Invariant extensions of Lebesgue measure

We already know that Lebesgue measure can be extended to a countably additive measure on the σ-algebra obtained by adding a given nonmeasurable set to the class of Lebesgue measurable sets. However, such an extension may not be invariant with respect to translations. Szpilrajn-Marczewski [**928**] proved that there exists an extension of Lebesgue measure λ on the real line to a countably additive measure l that is defined on some σ-algebra $\mathfrak{L}$ strictly containing the σ-algebra of Lebesgue measurable sets, and is complete and invariant with respect to translations (i.e., if $A \in \mathfrak{L}$, then $A + t \in \mathfrak{L}$ and

$l(A+t)=l(A)$ for all t). It was proved in Kodaira, Kakutani [**525**] that there exists a countably additive extension of Lebesgue measure that is invariant with respect to translations and is nonseparable, i.e., there exists no countable collection of sets approximating all measurable sets in the sense of measure. It was shown in Kakutani, Oxtoby [**483**] that there also exist nonseparable extensions of Lebesgue measure that are invariant with respect to all isometries.

Besides countably additive, finitely additive extensions invariant with respect to translations or isometries have been considered, too. In this direction Banach [**49**] proved that on the class of all bounded sets in $\mathbb{R}^1$ and $\mathbb{R}^2$ there exist nontrivial additive set functions m invariant with respect to all isometries, i.e., translations and linear isometries (moreover, one can ensure the coincidence of m with Lebesgue measure on all measurable sets, but one can also obtain the equality $m(E)=1$ for some set E of Lebesgue measure zero). There are no such functions on $\mathbb{R}^3$, which was first proved by F. Hausdorff. This negative result was investigated by Banach and Tarski [**60**], who proved the following theorem; a proof is found in Stromberg [**915**], Wise, Hall [**1022**, Example 6.1], and also in Wagon [**1001**].

1.12.45. Theorem. *Let A and B be bounded sets in $\mathbb{R}^3$ with nonempty interiors. Then, for some $n\in\mathbb{N}$, one can partition A into pieces $A_1,\ldots,A_n$ and B into pieces $B_1,\ldots,B_n$ such that, for every i, the set A_i is congruent to the set B_i.*

If A is a ball and B consists of two disjoint balls of the same radius, then $n=5$ suffices in this theorem, but $n=4$ is not enough.

Let $\mathcal{R}_n$ be the ring of bounded Lebesgue measurable sets in $\mathbb{R}^n$. Banach [**49**] investigated the following question (posed by Ruziewicz): is it true that every finitely additive measure on $\mathcal{R}_n$ that is invariant with respect to isometries is proportional to Lebesgue measure? Banach gave negative answers for $n=1,2$. G.A. Margulis [**655**] proved that for $n\geq 3$ the answer is positive. W. Sierpiński raised the following question (see Szpilrajn [**928**]): does there exist a maximal countably additive extension of Lebesgue measure on $\mathbb{R}^n$, invariant with respect to isometries? A negative answer to this question was given only half a century later in Ciesielski, Pelc [**182**] (see also Ciesielski [**180**]), where it was proved that, for any group G of isometries of the space $\mathbb{R}^n$ containing all parallel translations, one can write $\mathbb{R}^n$ as the union of a sequence of sets Z_n, each of which is absolutely G-null (earlier under the continuum hypothesis, a solution was given by S.S. Pkhakadze and A. Hulanicki, see references in [**182**]). Here an absolutely G-null set is a set Z such that, for each σ-finite G-invariant measure m, there exists a G-invariant extension defined on Z, and all such extensions vanish on Z (a countably additive σ-finite measure m is called G-invariant if it is defined on some σ-algebra $\mathcal{M}$ such that $g(A)\in\mathcal{M}$ and $m\bigl(g(A)\bigr)=m(A)$ for all $g\in G$, $A\in\mathcal{M}$). For the group of parallel translations, this result was obtained earlier by A.B. Kharazishvili, who proved under the continuum hypothesis a more general assertion (see [**507**]). On this subject and related problems, see Hadwiger [**392**], Kharazishvili [**507**], [**510**], [**512**], Lubotzky [**625**], von Neumann [**712**], Sierpiński [**880**], and Wagon [**1001**].

1.12(xii). Whitney's decomposition

In Lemma 1.7.2, we have represented any open set as a union of closed cubes with disjoint interiors. However, the behavior of diameters of such cubes could be quite irregular. It was observed by Whitney that one can achieve that these diameters be comparable with the distance to the boundary of the set. As above, for nonempty sets A and B we denote by $d(A,B)$ the infimum of the distances between the points in A and B.

1.12.46. Theorem. *Let Ω be an open set in $\mathbb{R}^n$ and let $Z := \mathbb{R}^n\backslash\Omega$ be nonempty. Then, there exists an at most countable family of closed cubes Q_k with edges parallel to the coordinate axes such that*:
(i) *the interiors of Q_k are disjoint and $\Omega = \bigcup_{k=1}^{\infty} Q_k$,*
(ii) $\operatorname{diam} Q_k \le d(Q_k, Z) \le 4\operatorname{diam} Q_k$.

PROOF. In the reasoning that follows we mean by cubes only closed cubes with edges parallel to the coordinate axes. Let $\mathcal{S}_k$ be a net of cubes obtained by translating the cube $[0,2^{-k}]^n$ by all vectors whose coordinates are multiples of 2^{-k}. The cubes in $\mathcal{S}_k$ have edges 2^{-k} and diameters $\sqrt{n}2^{-k}$. Set

$$\Omega_k := \Bigl\{x\in\Omega\colon\ 2\sqrt{n}2^{-k} < \operatorname{dist}(x,Z) \le 2\sqrt{n}2^{-k+1}\Bigr\},\quad k\in\mathbb{Z}.$$

It is clear that $\Omega = \bigcup_{k\in\mathbb{Z}}\Omega_k$. Now we can choose a preliminary collection $\mathcal{F}$ of cubes in the above nets. To this end, let us consider the cubes in $\mathcal{S}_k$. If a cube $Q\in\mathcal{S}_k$ meets Ω_k, then we include it in $\mathcal{F}$. Thus,

$$\mathcal{F} = \bigcup_{k=-\infty}^{\infty}\bigl\{Q\in\mathcal{S}_k\colon\ Q\cap\Omega_k\ne\varnothing\bigr\}.$$

It is clear that the union of all cubes in $\mathcal{F}$ covers Ω. Let us show that

$$\operatorname{diam} Q \le d(Q,Z) \le 4\operatorname{diam} Q,\quad \forall\, Q\in\mathcal{F}. \tag{1.12.12}$$

A cube Q from $\mathcal{F}$ belongs to $\mathcal{S}_k$ for some k. Hence it has the diameter $\sqrt{n}2^{-k}$ and there exists $x\in Q\cap\Omega_k$. Therefore,

$$d(Q,Z)\le \operatorname{dist}(x,Z)\le 2\sqrt{n}2^{-k+1}.$$

On the other hand,

$$d(Q,Z)\ge \operatorname{dist}(x,Z)-\operatorname{diam} Q > 2\sqrt{n}2^{-k}-\sqrt{n}2^{-k}.$$

It follows by (1.12.12) that all cubes Q are contained in Ω. However, cubes in $\mathcal{F}$ may not be disjoint. For this reason some further work on $\mathcal{F}$ is needed. Let us show that for every cube $Q\in\mathcal{F}$, there exists a unique cube from $\mathcal{F}$ that contains Q and is maximal in the sense that it is not contained in a larger cube from $\mathcal{F}$, and that such maximal cubes have disjoint interiors. Then the collection of such maximal cubes is a desired one: they have all the necessary properties, in particular, their union equals the union of cubes in $\mathcal{F}$, i.e., equals Ω. For the proof of the existence of maximal cubes, let us observe that two cubes $Q' \in \mathcal{S}_k$ and $Q'' \in \mathcal{S}_m$ may have common inner points only if one of them is entirely contained in the other (i.e., if there are common inner points and $k < m$, then we have $Q'' \subset Q'$). This is clear from the construction of $\mathcal{S}_k$. Now let $Q \in \mathcal{F}$. If $Q \subset Q' \in \mathcal{F}$, then we obtain by (1.12.12) that $\operatorname{diam} Q' \leq 4 \operatorname{diam} Q$. By the above observation we see that, for any two cubes $Q', Q'' \in \mathcal{F}$ containing Q, either $Q' \subset Q''$ or $Q'' \subset Q'$. Together with the previous estimate of diameter this proves the existence and uniqueness of a maximal cube $K(Q) \in \mathcal{F}$ containing Q. For the same reasons, maximal cubes $K(Q_1)$ and $K(Q_2)$, corresponding to distinct $Q_1, Q_2 \in \mathcal{F}$, either coincide or have disjoint interiors. Indeed, otherwise one of them would strictly belong to the other, say, $K(Q_1) \subset K(Q_2)$. Then $Q_1 \subset K(Q_2)$, contrary to the uniqueness of a maximal cube for Q_1. Deleting from the collection of cubes $K(Q)$ the repeating ones (if different Q' and Q'' give one and the same maximal cube), we obtain the required sequence. □

Exercises

1.12.47.° Suppose we are given a family of open sets in $\mathbb{R}^n$. Show that this family contains an at most countable subfamily with the same union.

HINT: consider a countable everywhere dense set of points x_k in the union W of the given sets W_α; for every point x_k, take all open balls $K(x_k, r_j)$ centered at x_k, having rational radii r_j and contained in at least one of the sets W_α; for every $U(x_k, r_j)$, pick a set $W_{\alpha_{k,j}} \supset U(x_k, r_j)$ and consider the obtained family.

1.12.48.° Let W be a nonempty open set in $\mathbb{R}^n$. Prove that W is the union of an at most countable collection of open cubes whose edges are parallel to the coordinate axes and have lengths of the form $p2^{-q}$, where $p, q \in \mathbb{N}$, and whose centers have coordinates of the form $m2^{-k}$, where $m \in \mathbb{Z}$, $k \in \mathbb{N}$.

HINT: observe that the union of all cubes in W of the indicated type is W.

1.12.49.° Let μ be a nonnegative measure on a ring $\mathcal{R}$. Prove that the class of all sets $Z \in \mathcal{R}$ of measure zero is a ring.

1.12.50.° Let μ be an arbitrary finite Borel measure on a closed interval I. Show that there exists a first category set E (i.e., a countable union of nowhere dense sets) such that $\mu(I \backslash E) = 0$.

HINT: it suffices to find, for each n, a compact set K_n without inner points such that $\mu(K_n) > \mu(I) - 2^{-n}$. By using that μ has an at most countable set of points a_j of nonzero measure, one can find a countable everywhere dense set of points s_j of μ-measure zero. Around every point s_j there is an interval $U_{n,j}$ with $\mu(U_{n,j}) < 2^{-j-n}$. Now we take the compact set $K_n = I \backslash \bigcup_{j=1}^\infty U_{n,j}$.

1.12.51.° Let $\mathcal{S}$ be some collection of subsets of a set X such that it is closed with respect to finite unions and finite intersections and contains the empty set (for example, the class of all closed sets or the class of all open sets in $[0,1]$). Show that the class of all sets of the form $A \backslash B$, $A, B \in \mathcal{S}$, $B \subset A$, is a semiring, and the class of all sets of the form $(A_1\backslash B_1)\cup\cdots\cup(A_n\backslash B_n)$, $A_i, B_i\in\mathcal{S}$, $B_i\subset A_i$, $n\in\mathbb{N}$, is the ring generated by $\mathcal{S}$.

HINT: verify that $(A\backslash B)\backslash(C\backslash D)=\bigl(A\backslash\bigl(B\cup(A\cap C)\bigr)\bigr)\cup\bigl((A\cap D)\backslash(B\cap D)\bigr)$ if $B\subset A$, $D\subset C$; next verify that the class of the indicated unions is closed with respect to intersections.

1.12.52.° Let m be an additive set function on a ring of sets $\mathcal{R}$. Prove the following Poincaré formula for all $A_1,\ldots,A_n\in\mathcal{R}$:

$$m\Bigl(\bigcup_{i=1}^n A_i\Bigr)=\sum_{i=1}^n m(A_i)-\sum_{1\le i<j\le n} m(A_i\cap A_j)$$
$$+\sum_{1\le i<j<k\le n} m(A_i\cap A_j\cap A_k)-\cdots+(-1)^{n+1}m\Bigl(\bigcap_{i=1}^n A_i\Bigr).$$

1.12.53.° Let $\mathcal{R}_1$ and $\mathcal{R}_2$ be two semirings of sets. Prove that

$$\mathcal{R}_1\times\mathcal{R}_2=\{R_1\times R_2\colon\ R_1\in\mathcal{R}_1, R_2\in\mathcal{R}_2\}$$

is a semiring. Show that $\mathcal{R}_1\times\mathcal{R}_2$ may not be a ring even if $\mathcal{R}_1$ and $\mathcal{R}_2$ are algebras.

1.12.54.° Let $\mathcal{F}$ be some collection of sets in a space X. Prove that every set A in the σ-algebra $\sigma(\mathcal{F})$ generated by $\mathcal{F}$ is contained in the σ-algebra generated by an at most countable subcollection $\{F_n\}\subset\mathcal{F}$.

HINT: verify that the union of all σ-algebras $\sigma(\{F_n\})$ generated by at most countable subcollections $\{F_n\}\subset\mathcal{F}$ is a σ-algebra.

1.12.55.° (Brown, Freilich [**134**]) The aim of this exercise is to show that Proposition 1.2.6 may be false if a σ-algebra is defined in the broader sense mentioned in §1.2. Suppose we are given a set X and a collection $\mathcal{S}$ of its subsets such that the union of all sets in $\mathcal{S}$ is $Y\subset X$. Prove that the following conditions are equivalent: (i) Y is an at most countable union of sets in $\mathcal{S}$; (ii) there exists a smallest family of sets $\mathcal{A}$ with the following properties: $\mathcal{A}$ is a σ-algebra on some subset $Z\subset X$ (i.e., Z is the unit of this σ-algebra) and $\mathcal{S}\subset\mathcal{A}$, where a smallest family is a family that is contained in every other family with the stated properties. Consider the example where $X=[0,1]$, $Y=[0,1/2]$, $\mathcal{S}$ is the class of all at most countable subsets of Y.

HINT: if Y is not the countable union of elements in $\mathcal{S}$, then Y does not belong to the class $\mathcal{P}$ of all sets $A\subset Y$ such that $A\subset\bigcup_{n=1}^\infty S_n$, where $S_n\in\mathcal{S}$. Let us fix $z\in X\backslash Y$ and consider the class $\mathcal{E}$ of all sets $E\subset Y\cup\{z\}$ such that either $E\in\mathcal{P}$ or $(Y\cup\{z\})\backslash E\in\mathcal{P}$. It is readily verified that $\mathcal{E}$ is a σ-algebra. One has $Y\notin\mathcal{E}$. If there exists a smallest family of sets $\mathcal{A}$ with the properties indicated in (ii), then the corresponding set Z cannot be smaller than Y, i.e., $Z=Y$ and hence $Y\in\mathcal{A}$. Therefore, $\mathcal{A}$ does not belong to $\mathcal{E}$, which gives a contradiction.

1.12.56. (Broughton, Huff [**132**]) Suppose we are given a sequence of σ-algebras $\mathcal{A}_n$ in a space X such that $\mathcal{A}_n$ is strictly contained in $\mathcal{A}_{n+1}$ for each n. Prove that $\bigcup_{n=1}^\infty\mathcal{A}_n$ is not a σ-algebra.

HINT: we may assume that there is a nonempty set $B\in\mathcal{A}_1$ not equal to X. If, for some n, we have $B\cap\mathcal{A}_{n+1}=B\cap\mathcal{A}_n$ and the same is true for $X\backslash B$, then $\mathcal{A}_{n+1}=\mathcal{A}_n$, which is a contradiction. Hence one can find $E\in\mathcal{A}_1$ and infinitely many p_k with $p_{k+1}>p_k$ such that $(E\cap\mathcal{A}_{p_k+1})\backslash(E\cap\mathcal{A}_{p_k})\neq\varnothing$. Then the classes $E\cap\mathcal{A}_{p_k}$ are strictly increasing σ-algebras on E. By induction, we construct a subsequence $\mathcal{A}_{j_1}, \mathcal{A}_{j_2}, \ldots$, where $j_{k+1} > j_k$, and sets $E_1 \supset E_2 \supset \ldots$ with $E_k \in \mathcal{A}_{j_k}$ and $E_{k+1} \in (E_k \cap \mathcal{A}_{j_{k+1}}) \backslash (E_k \cap \mathcal{A}_{j_k})$. We obtain disjoint sets $F_k := E_k \backslash E_{k+1}$, $F_k \in \mathcal{A}_{j_{k+1}} \backslash \mathcal{A}_{j_k}$. We may assume that $X = \bigcup_{k=1}^\infty F_k$. Let $\pi\colon X \to \mathbb{N}$, $\pi(F_k) = k$ and let $\mathcal{A}'_n := \{A\colon \pi^{-1}(A) \in \mathcal{A}_n\}$. It is easily verified that, for every n, there is the smallest set $B_n \in \mathcal{A}'_n$ with $n \in B_n$. Then $B_n \subset \{k \geq n\}$, $B_n \neq \{n\}$. If $m \in B_n$, then $B_m \subset B_n$, since $B_m \cap B_n \in \mathcal{A}'_m$. Let $n_1 := 1$. We find by induction $n_{k+1} \in B_{n_k}$, $n_{k+1} > n_k$. Then $B_{n_1} \supset B_{n_2} \supset \ldots$ Let $E := \{n_2, n_4, n_6, \ldots\}$. If $\pi^{-1}(E) \in \mathcal{A}_n$, i.e., $E \in \mathcal{A}'_n$, then $E \in \mathcal{A}'_{n_{2k}}$ for some k, whence one has $\{n_{2k}, n_{2k+2}, \ldots\} \in \mathcal{A}'_{n_{2k}}$ and $B_{n_{2k}} \subset \{n_{2k}, n_{2k+2}, \ldots\}$, contrary to the inclusion $n_{2k+1} \in B_{n_{2k}}$.

1.12.57.° Show that every set of positive Lebesgue measure contains a nonmeasurable subset.

1.12.58. Prove that there exists a sequence of sets $A_n \subset [0,1]$ such that for all n one has $A_{n+1} \subset A_n$, $\bigcap_{n=1}^\infty A_n = \varnothing$ and $\lambda^*(A_n) = 1$, where λ is Lebesgue measure.

Hint: let $\{r_n\}$ be some enumeration of the rational numbers and let $E \subset [0,1]$ be the nonmeasurable set from Vitali's example. Show that the sets

$$E_n := \bigl(E \cup (E + r_1) \cup \cdots \cup (E + r_n)\bigr) \cap [0,1]$$

have inner measure zero and take $A_n := [0,1] \backslash E_n$.

1.12.59. Show that every nonempty perfect set contains a nonempty perfect subset of Lebesgue measure zero. In particular, every set of positive Lebesgue measure contains a measure zero compact set of cardinality of the continuum.

Hint: it suffices to consider a compact set K of positive measure without isolated points; then, similarly to the construction of the classical Cantor set, delete from K the countable union of sets $J_n \cap K$, where J_n are disjoint intervals, in such a way that the remaining set is perfect, nonempty and has measure zero.

1.12.60.° Let C be the Cantor set in $[0,1]$. Show that

$$C + C := \{c_1 + c_2\colon\ c_1, c_2 \in C\} = [0,2], \quad C - C := \{c_1 - c_2\colon\ c_1, c_2 \in C\} = [-1,1].$$

Hint: the sets $C + C$ and $C - C$ are compact, hence it suffices to verify that they contain certain everywhere dense subsets in the indicated intervals, which can be done by using the description of C in terms of the ternary expansion.

1.12.61.° Give an example of two closed sets $A, B \subset \mathbb{R}$ of Lebesgue measure zero such that the set $A + B := \{a + b\colon\ a \in A, b \in B\}$ is $\mathbb{R}$.

Hint: take for A the Cantor set and for B the union of translations of A to all integer numbers.

1.12.62.° (Steinhaus [**910**]) Let A be a set of positive Lebesgue measure on the real line. Show that the set $A - A := \{a_1 - a_2\colon\ a_1, a_2 \in A\}$ contains some interval. Prove an analogous assertion for $\mathbb{R}^n$ (obtained in Rademacher [**775**]).

Hint: there is a compact set $K \subset A$ with $\lambda(K) > 0$; take an open set U with $K \subset U$ and $\lambda(U) < 2\lambda(K) = \lambda(K) + \lambda(K + h)$ and observe that there exists $\varepsilon > 0$ such that $K + h \subset U$ whenever $|h| < \varepsilon$; then $\lambda\bigl(K \cup (K + h)\bigr) \leq \lambda(U)$ for such h, whence $K \cap (K + h) \neq \varnothing$.

1.12.63. (P.L. Ulyanov, see Bary [**66**, Appendix, §23]) Let $E \subset [0,1]$ be a measurable set of positive measure. (i) Prove that for every sequence $\{h_n\}$ converging to zero and every $\varepsilon > 0$, there exist a measurable set $E_\varepsilon \subset E$ and a subsequence $\{h_{n_k}\}$

such that $\lambda(E_\varepsilon) > \lambda(E) - \varepsilon$ and for all $x \in E_\varepsilon$ we have $x + h_{n_k} \in E$, $x - h_{n_k} \in E$ for all k.

(ii) Prove that there exist a measurable set $E_0 \subset E$ and a sequence of numbers $h_n > 0$ converging to zero such that $\lambda(E_0) = \lambda(E)$ and for every $x \in E_0$, we have $x + h_n \in E$ for all $n \geq n(x)$.

HINT: (i) choose numbers n_k such that

$$\lambda\bigl(E \triangle (E + h_{n_k})\bigr) \leq \varepsilon 8^{-k}, \ \lambda\bigl(E \triangle (E - h_{n_k})\bigr) \leq \varepsilon 8^{-k},$$

and take $E_\varepsilon = \bigcap_{k=1}^\infty \bigl((E + h_{n_k}) \cap (E - h_{n_k})\bigr)$. (ii) For $\{2^{-n}\}$ and $\varepsilon_1 = 1/2$, take the set $E_{1/2}$ according to (i) and proceed by induction: if for some n we have chosen a set $E_{2^{-n}}$ according to (i) and a subsequence $\{h_k^{(n)}\}$ in $\{2^{-n}\}$, then when choosing $E_{2^{-n-1}}$ for the number $n+1$, we take a subsequence in $\{h_k^{(n)}\}$. Let $E_0 = \bigcup_{n=1}^\infty E_{2^{-n}}$ and $h_n := h_n^{(n)}$.

1.12.64. Let A be a set of positive Lebesgue measure in $\mathbb{R}^n$ and let $k \in \mathbb{N}$. Prove that there exist a set B of positive Lebesgue measure and a number $\delta > 0$ such that the sets $B_{i_1,\ldots,i_n} := B + \delta(i_1, \ldots, i_n)$, where $i_j \in \{1, \ldots, k\}$, are disjoint and are contained in A.

1.12.65. (Jones [**469**]) In this exercise, by a Hamel basis we mean a Hamel basis of the space $\mathbb{R}^1$ over the field of rational numbers.

(i) Let M be a set in $[0, 1]$ and let $\lambda_*(M - M) > 0$. Prove that M contains a Hamel basis. Deduce that the Cantor set contains a Hamel basis and that every set of positive measure contains a Hamel basis.

(ii) Prove that there exists a Hamel basis containing a nonempty perfect set.

(iii) Let H be a Hamel basis and $DE := \{e_1 - e_2, e_1, e_2 \in E, e_1 \geq e_2\}$ for any set E. Prove that $\lambda^*(D^n H) > 0$ for some n and $\lambda_*(D^n H) = 0$ for all n, where D^n is defined inductively.

(iv) Let H be a Hamel basis and $TE := \{e_1 + e_2 - e_3, e_1, e_2, e_3 \in E\}$ for any set E. Prove that $\lambda^*(T^n H) > 0$ for some n and $\lambda_*(T^n H) = 0$ for all n.

1.12.66. Prove the existence of a nonmeasurable (in the sense of Lebesgue) Hamel basis of $\mathbb{R}^1$ over $\mathbb{Q}$ without using the continuum hypothesis (see Example 1.12.21).

HINT: let $\omega_{\mathfrak{c}}$ be the smallest ordinal number corresponding to the cardinality of the continuum. The family of all compacts of positive measure has cardinality $\mathfrak{c}$ and hence can be put in some one-to-one correspondence $\alpha \mapsto K_\alpha$ with ordinal numbers $\alpha < \omega_{\mathfrak{c}}$. By means of transfinite induction we find a family of elements $h_\alpha \in K_\alpha$ linearly independent over $\mathbb{Q}$. Namely, if such elements h_β are already found for all $\beta < \alpha$, where $\alpha < \mathfrak{c}$, then the collection of all linear combinations of these elements with rational coefficients has cardinality less than that of the continuum. Hence K_α contains an element h_α that is not such a linear combination. Let us complement the constructed family $\{h_\alpha, \alpha < \mathfrak{c}\}$ to a Hamel basis. We obtain a nonmeasurable set, since if it were measurable, then, according to what we proved earlier, it would have measure zero, which is impossible because the constructed family meets every compact set in $[0, 1]$ of positive measure.

1.12.67. Prove that there exists a bounded set E of measure zero such that $E + E$ is nonmeasurable.

HINT: let $H = \{h_\alpha\}$ be a Hamel basis over $\mathbb{Q}$ of zero measure with $h_\alpha \in [0, 1]$, $A = \{rh\colon\ r \in \mathbb{Q} \cap [0, 1], h \in H\}$. Set $E_1 := A + A$; it is readily seen that E_1 has inner measure zero because otherwise $E_1 - E_1$ would contain an interval, which is impossible, since any point in $E_1 - E_1$ is a linear combination of four vectors in H. If E_1 is nonmeasurable, then we take $E = A$; otherwise we set $E_2 := E_1 + E_1$ and construct inductively $E_{n+1} := E_n + E_n$. In finitely many steps we obtain a desired set, since $E_n - E_n$ cannot contain an interval and the union of all E_n covers $[0,1]$.

1.12.68. (Ciesielski, Fejzić, Freiling [**181**]) Show that every set $E \subset \mathbb{R}$ contains a subset A with $\lambda_*(A+A) = 0$ and $\lambda^*(A+A) = \lambda^*(E+E)$, where λ is Lebesgue measure.

1.12.69. (Sodnomov [**895**]) Let $E \subset \mathbb{R}^1$ be a set of positive Lebesgue measure. Then, there exists a perfect set P with $P + P \subset E$.

1.12.70. Let $\beta \in (0,1)$. The operation $T(\beta)$ over a finite family of disjoint intervals $I_1, \ldots, I_n$ of nonzero length consists of deleting from every I_j the open interval with the same center as I_j and length $\beta\lambda(I_j)$. Given a sequence of numbers $\beta_n \in (0,1)$, let us define inductively compacts K_n obtained by consequent application of the operations $T(\beta_1),\ldots,T(\beta_n)$, starting with the interval $I = [0,1]$.

(i) Show that $\lambda\bigl(\bigcap_{n=1}^\infty K_n\bigr) = \lim\limits_{n\to\infty} \prod_{i=1}^n (1-\beta_i)$. In particular, letting $\beta_n = 1 - \alpha^{\frac{1}{n(n+1)}}$, where $\alpha \in (0,1)$, we have $\lambda\bigl(\bigcap_{n=1}^\infty K_n\bigr) = \alpha$.

(ii) Show that there exists a sequence of pairwise disjoint nowhere dense compact sets A_n with the following properties: $\lambda(A_n) = 2^{-n}$ and the intersection of A_{n+1} with each interval contiguous to the set $\bigcup_{j=1}^n A_j$ has a positive measure.

(iii) Show that the intersections of the set $A := \bigcup_{n=1}^\infty A_{2n-1}$ and its complement with every interval $I \subset [0,1]$ have positive measures.

Hint: see George [**351**, p. 62, 63].

1.12.71.° Prove that Lebesgue measure of every measurable set $E \subset \mathbb{R}^n$ equals the infimum of the sums $\sum_{k=1}^\infty \lambda_n(U_k)$ over all sequences of open balls U_k covering E.

Hint: observe that it suffices to prove the claim for open E and in this case use the fact that one can inscribe in E a disjoint collection of open balls V_j such that the set $E \backslash \bigcup_{j=1}^\infty V_j$ has measure zero, and then cover this set with a sequence of balls W_i with the sum of measures majorized by a given $\varepsilon > 0$.

1.12.72. Suppose that μ is a countably additive measure with values in $[0,+\infty]$ on the σ-algebra of Borel sets in $\mathbb{R}^n$ and is finite on balls, and let W be a nonempty open set in $\mathbb{R}^n$. Prove that there exists an at most countable collection of disjoint open cubes Q_j in W with edges parallel to the coordinate axes such that $\mu\bigl(W \backslash \bigcup_{j=1}^\infty W_j\bigr) = 0$.

Hint: we may assume that W is contained in a cube I; in the proof of Lemma 1.7.2 one can choose all cubes in such a way that their boundaries have μ-measure zero; to this end, we observe that at most countably many affine hyperplanes parallel to the coordinate hyperplanes have positive μ-measure. In addition, given a countable set of points t_i on the real line, the set of points of the form $r + t_i$, where r is binary-rational (i.e., $r = m2^{-k}$ with integer m, k), is countable as well; therefore, one can find $\alpha \neq 0$ such that the required cubes have edges of length $m2^{-k}$, where $m \in \mathbb{Z}$, $k \in \mathbb{N}$, and centers with coordinates of the form $\alpha + m2^{-k}$.

1.12.73.° Show that a set $E \subset \mathbb{R}$ is Lebesgue measurable precisely when for every $\varepsilon > 0$, there exist open sets U and V such that $E \subset U$, $U \backslash E \subset V$ and $\lambda(V) < \varepsilon$.

1.12.74.° Let μ be a Borel probability measure on the cube $I = [0,1]^n$ such that $\mu(A) = \mu(B)$ for any Borel sets $A, B \subset I$ that are translations of one another. Show that μ coincides with Lebesgue measure λ_n.

HINT: observe that μ coincides with λ_n on all cubes in I with edges parallel to the axes and having binary-rational lengths (the boundaries of such cubes have measure zero with respect to μ by the countable additivity and the hypothesis). It follows that μ coincides with λ_n on the algebra generated by the indicated cubes.

1.12.75.° (i) Show that for any countably additive function $\mu\colon \mathfrak{R} \to [0, +\infty)$ on a semiring $\mathfrak{R}$ and any $A, A_n \in \mathfrak{R}$ such that A_n either increase or decrease to A, one has the equality $\mu(A) = \lim\limits_{n\to\infty} \mu(A_n)$.

(ii) Give an example showing that the properties indicated in (i) do not imply the countable additivity of a nonnegative additive set function on a semiring.

HINT: (ii) consider the semiring of sets of the form $\mathbb{Q}\cap(a,b)$, $\mathbb{Q}\cap(a,b]$, $\mathbb{Q}\cap[a,b)$, $\mathbb{Q}\cap[a,b]$, where $\mathbb{Q}$ is the set of rational numbers in $[0,1]$; on such sets let μ equal $b-a$.

1.12.76.° Give an example of a nonnegative additive set function μ on a semiring $\mathfrak{R}$ such that $\mu(A) = \lim\limits_{n\to\infty} \mu(A_n)$ whenever $A, A_n \in \mathfrak{R}$ and A_n either increase or decrease to A, but the additive extension of μ to the ring generated by $\mathfrak{R}$ does not possess this property.

HINT: see Exercise 1.12.75.

1.12.77.° (i) Show that a bounded set $E \subset \mathbb{R}^n$ is Jordan measurable (see Definition in §1.1) precisely when the boundary of E (the set of points each neighborhood of which contains points from the set E and from its complement) has measure zero. (ii) Show that the collection of all Jordan measurable sets in an interval or in a cube is a ring.

1.12.78.° Prove Proposition 1.6.5.

1.12.79.° Show that a bounded nonnegative measure μ on a σ-algebra $\mathcal{A}$ is complete precisely when $\mathcal{A} = \mathcal{A}_\mu$; In particular, the Lebesgue extension of any complete measure coincides with the initial measure.

1.12.80.° Give an example of a σ-finite measure on a σ-algebra that is not σ-finite on some sub-σ-algebra.

HINT: consider Lebesgue measure on $\mathbb{R}^1$ and the sub-σ-algebra of all sets that are either at most countable or have at most countable complements.

1.12.81.° Let A_n be subsets of a space X. Show that

$$\{x\colon\, x \in A_n \text{ for infinitely many } n\} = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k.$$

1.12.82.° Let μ be a probability measure and let $A_1, \ldots, A_n$ be measurable sets with $\sum_{i=1}^n \mu(A_i) > n - 1$. Prove that $\mu\bigl(\bigcap_{i=1}^n A_i\bigr) > 0$.

HINT: observe that $\sum_{i=1}^n \mu(C_i) = \sum_{i=1}^n (1 - \mu(A_i)) < 1$, where C_i is the complement of A_i.

1.12.83.° **(Baire category theorem)** Let M_j, $j \in \mathbb{N}$, be closed sets in $\mathbb{R}^d$ such that their union is a closed cube. Prove that at least one of the sets M_j has inner points. Generalize to the case where M_j are closed sets in a complete metric space X with $\bigcup_{j=1}^\infty M_j = X$. A set in a metric space is called nowhere dense if its closure has no interior; a countable union of nowhere dense sets is said to be a first category set. The above result can be formulated as follows: a complete nonempty metric space is not a first category set.

HINT: assuming the opposite, construct a sequence of decreasing closed balls U_j with radii $r_j \to 0$ such that $U_j \cap M_j = \varnothing$.

1.12.84. Prove that $\mathbb{R}^1$ cannot be written as the union of a family of pairwise disjoint nondegenerate closed intervals.

HINT: verify that such a family must be countable and that the family of all endpoints of the given intervals is closed and has no isolated points; apply the Baire theorem. One can also use that a closed set without isolated points is uncountable (see Proposition 6.1.17 in Chapter 6).

1.12.85. Show that $\mathbb{R}^n$ with $n > 1$ cannot be written as the union of a family of closed balls with pairwise disjoint interiors.

HINT: apply Exercise 1.12.84 to a straight line which passes through the origin, contains no points of tangency of the given balls and is not tangent to any of them.

1.12.86.° Show that the σ-algebra $\mathcal{B}(\mathbb{R}^1)$ of all Borel subsets of the real line is the smallest class of sets that contains all closed sets and admits countable intersections and countable unions.

HINT: use that the indicated smallest class is monotone and contains the algebra of finite unions of rays and intervals; another approach is to verify that the collection of all sets belonging to the above class along with their complements is a σ-algebra and contains all closed sets. A stronger assertion is found in Example 1.12.3.

1.12.87. (i) Prove that the union of an arbitrary family of nondegenerate closed intervals on the real line is measurable.

(ii) Prove that the union of an arbitrary family of nondegenerate rectangles in the plane is measurable.

(iii) Prove that the union of an arbitrary family of nondegenerate triangles in the plane is measurable.

HINT: (i) it suffices to verify that the union of the family of all intervals I_α of length not smaller than $1/k$ is measurable for each k; there exists an at most countable subfamily I_{α_n} such that the union of their interiors equals the union of the interiors of all I_α; the set $\bigcup_\alpha I_\alpha \backslash \bigcup_{n=1}^\infty I_{\alpha_n}$ is at most countable, since every point is isolated (such a point may be only an endpoint of some interval I_α, and an interval of length $1/k$ cannot contain three such points). (ii) Consider all rectangles E_α with the shorter side length at least $1/k$; take a countable subfamily E_{α_n} with the union of interiors equal to the union of the interiors of all E_α and observe that any circle of a sufficiently small radius can meet at most finitely many sides of those rectangles E_α that are not covered by the rectangles E_{α_n}. (iii) Modify the proof of (ii) for triangles, considering subfamilies of triangles with sides at least $1/k$ and angles belonging to $[1/k, \pi - 1/k]$. We note that these assertions follow by the Vitali covering theorem proven in Chapter 5 (Theorem 5.5.2).

1.12.88. (Nikodym [**716**]) For any sequence of sets E_n let

$$\limsup_{n\to\infty} E_n := \bigcap_{n=1}^\infty \bigcup_{k=n}^\infty E_k, \quad \liminf_{n\to\infty} E_n := \bigcup_{n=1}^\infty \bigcap_{k=n}^\infty E_k.$$

Let $(X, \mathcal{A}, \mu)$ be a probability space. Prove that a sequence of sets $A_n \in \mathcal{A}$ converges to a set $A \in \mathcal{A}$ in the Fréchet–Nikodym metric $d(B_1, B_2) = \mu(B_1 \triangle B_2)$ precisely when every subsequence in $\{A_n\}$ contains a further subsequence $\{E_n\}$ such that

$$A = \limsup_{n\to\infty} E_n = \liminf_{n\to\infty} E_n$$

up to a measure zero set.

HINT: see Theorem 1.12.6; this also follows by Theorem 2.2.5 in Chapter 2.

1.12.89.° Let $(X, \mathcal{A}, \mu)$ be a space with a probability measure, let $A_n \in \mathcal{A}_\mu$, and let

$$B := \{x\colon\ x \in A_n \text{ for infinitely many } n\},$$

i.e., $B = \bigcap_{k=1}^\infty \bigcup_{n=k}^\infty A_n$ according to Exercise 1.12.81.

(i) **(Borel–Cantelli lemma)** Show that if $\sum_{n=1}^\infty \mu(A_n) < \infty$, then $\mu(B) = 0$.

(ii) Prove that if $\mu(A_n) \geq \varepsilon > 0$ for all n, then $\mu(B) \geq \varepsilon$.

(iii) (Pták [**772**]) Show that if $\mu(B) > 0$, then one can find a subsequence $\{n_k\}$ such that $\mu(\bigcap_{k=1}^m A_{n_k}) > 0$ for all m.

HINT: the sets $B_k := \bigcup_{n=k}^\infty A_n$ decrease and one has $\mu(B_k) \leq \sum_{n=k}^\infty \mu(A_n)$, $\mu(B_k) \geq \mu(A_k)$. If $\mu(B) > 0$, we find the first number n_1 with $\mu(B \cap A_{n_1}) > 0$, then we find $n_2 > n_1$ with $\mu(B \cap A_{n_1} \cap A_{n_2}) > 0$ and so on. See also Exercise 2.12.35.

1.12.90. (i) Construct a sequence of sets $E_n \subset [0,1]$ of measure $\sigma > 0$ such that the intersection of each subsequence in this sequence has measure zero.

(ii) Let μ be a probability measure and let A_n be μ-measurable sets such that $\mu(A_n) \geq \varepsilon > 0$ for all $n \in \mathbb{N}$. Show that there exists a subsequence n_k such that $\bigcap_{k=1}^\infty A_{n_k}$ is nonempty.

(iii) (Erdős, Kestelman, Rogers [**270**]) Let A_n be Lebesgue measurable sets in $[0,1]$ with $\lambda(A_n) \geq \varepsilon > 0$ for all $n \in \mathbb{N}$. Show that there exists a subsequence n_k such that $\bigcap_{k=1}^\infty A_{n_k}$ is uncountable (see a stronger assertion in Exercise 3.10.107).

HINT: (i) define E_n inductively: $E_1 = (0, 1/2)$, $E_2 = (0, 1/4) \cup (3/4, 1)$ and so on; the set E_{n+1} consists of 2^n intervals $J_{n,k}$ that are the left halves of the intervals $J_{n-1,k}$ and the left halves of the contiguous intervals to the intervals $J_{n-1,k}$. (ii) Follows by the previous exercise.

1.12.91. Let a function $\alpha\colon \mathbb{N} \to [0, +\infty)$ be such that $\sum_{k=1}^\infty \alpha(k) < \infty$. Prove that the set E of all $x \in (0,1)$ such that, for infinitely many natural numbers q, there exists a natural number p such that p and q are relatively prime and $|x - p/q| < \alpha(q)/q$, has measure zero. In Exercise 10.10.57 in Chapter 10 see a converse assertion.

HINT: for fixed q, let E_q be the set of all $x \in (0,1)$ such that, for some $p \in \mathbb{N}$, one has $|x - p/q| < \alpha(q)/q$. This set consists of the intervals of length $2\alpha(q)/q$ centered at the points p/q, $p = 1, \ldots, q$, whence $\lambda(E_q) \leq 2\alpha(q)$. By the Borel–Cantelli lemma, $\lambda(E) = 0$.

1.12.92. (Gillis [**354**], [**355**]) Let $E_k \subset [0,1]$ be measurable sets and let $\lambda(E_k) \geq \alpha$ for all k, where $\alpha \in (0,1)$. Prove that for all $p \in \mathbb{N}$ and $\varepsilon > 0$, there exist $k_1 < \cdots < k_p$ such that $\lambda(E_{k_1} \cap \cdots \cap E_{k_p}) > \alpha^p - \varepsilon$.

1.12.93. (i) Let $E \subset [0,1]$ be a set of Lebesgue measure zero. Prove that there exists a convergent series with positive terms a_n such that, for any $\varepsilon > 0$, the set E can be covered by a sequence of intervals I_n of length at most εa_n. (ii) Show that there is no such series that would suit every measure zero set.

1.12.94. (Wesler [**1010**]; Mergelyan [**682**] for $n = 2$) Let U_k be disjoint open balls of radii r_k in the unit ball U in $\mathbb{R}^n$ such that $U \backslash \bigcup_{k=1}^{\infty} U_k$ has measure zero. Show that $\sum_{k=1}^{\infty} r_k^{n-1} = \infty$.

HINT: see Crittenden, Swanson [**192**], Larman [**569**], and Wesler [**1010**].

1.12.95. (i) Let $\alpha = n^{-1}$, where $n \in \mathbb{N}$. Prove that for any sets A and B in $[0, 1]$ of positive Lebesgue measure, there exist points $x, y \in [0, 1]$ such that $\lambda(A \cap [x, y]) = \alpha\lambda(A)$ and $\lambda(B \cap [x, y]) = \alpha\lambda(B)$. (ii) Show that if $\alpha \in (0, 1)$ does not have the form n^{-1} with $n \in \mathbb{N}$, then assertion (i) is false.

HINT: see George [**351**, p. 59].

1.12.96. A set $S \subset \mathbb{R}^1$ is called a Sierpiński set if $S \cap Z$ is at most countable for every set Z of Lebesgue measure zero.

(i) Under the continuum hypothesis show the existence of a Sierpiński set.

(ii) Prove that no Sierpiński set is measurable.

HINT: see Kharazishvili [**511**].

1.12.97. Let A be a set in $\mathbb{R}^d$ of Lebesgue measure greater than 1. Prove that there exist two distinct points $x, y \in A$ such that the vector $x - y$ has integer coordinates.

1.12.98.° Prove that each convex set in $\mathbb{R}^d$ is Lebesgue measurable.

HINT: show that the boundary of a bounded convex set has measure zero.

1.12.99. Let A be a bounded convex set in $\mathbb{R}^d$ and let A^ε be the set of all points with the distance from A at most ε. Prove that $\lambda_d(A^\varepsilon)$, where λ_d is Lebesgue measure, is a polynomial of degree d in ε.

HINT: verify the claim for convex polyhedra.

1.12.100.° Prove Theorem 1.12.1.

1.12.101.° Let $(X, \mathcal{A}, \mu)$ be a probability space, $\mathcal{B}$ a sub-σ-algebra in $\mathcal{A}$, and let $\mathcal{B}^\mu$ be the σ-algebra generated by $\mathcal{B}$ and all sets of measure zero in $\mathcal{A}_\mu$.

(i) Show that $E \in \mathcal{B}^\mu$ precisely when there exists a set $B \in \mathcal{B}$ such that $E \triangle B \in \mathcal{A}_\mu$ and $\mu(E \triangle B) = 0$.

(ii) Give an example demonstrating that $\mathcal{B}^\mu$ may be strictly larger than the σ-algebra $\mathcal{B}_\mu$ that is the completion of $\mathcal{B}$ with respect to the measure $\mu|_{\mathcal{B}}$.

HINT: (i) the sets of the indicated form belong to $\mathcal{B}^\mu$ and form a σ-algebra. (ii) Take Lebesgue measure λ on the σ-algebra of all measurable sets in $[0, 1]$ and $\mathcal{B} = \{\varnothing, [0, 1]\}$. Then $\mathcal{B}_\lambda = \mathcal{B}$.

1.12.102.° Let μ be a probability measure on a σ-algebra $\mathcal{A}$. Suppose that $\mathcal{A}$ is countably generated, i.e., is generated by an at most countable family of sets. Show that the measure μ is separable. Give an example showing that the converse is false.

HINT: if $\mathcal{A}$ is generated by sets A_n, then the algebra $\mathcal{A}_0$ generated by those sets is at most countable. It remains to use that, for any $A \in \mathcal{A}$ and $\varepsilon > 0$, there exists $A_0 \in \mathcal{A}_0$ such that $\mu(A \triangle A_0) < \varepsilon$. As an example of a separable measure on a σ-algebra that is not countably generated, one can take Lebesgue measure on the σ-algebra of Lebesgue measurable sets in an interval (see §6.5). Another example: Lebesgue measure on the σ-algebra of all sets in $[0, 1]$ that are either at most countable or have at most countable complements.

1.12.103. Let $(X, \mathcal{A}, \mu)$ be a measure space with a finite nonnegative measure μ and let $\mathcal{A}/\mu$ be the corresponding metric Boolean algebra with the metric d

introduced in §1.12(iii). Prove that the mapping $A \mapsto X\backslash A$ from $\mathcal{A}/\mu$ to $\mathcal{A}/\mu$ and the mappings $(A,B) \mapsto A \cup B$, $(A,B) \mapsto A \cap B$ from $(\mathcal{A}/\mu)^2$ to $\mathcal{A}/\mu$ are continuous.

1.12.104. Let μ be a separable probability measure on a σ-algebra $\mathcal{A}$ and let $\{X_t\}_{t\in T}$ be an uncountable family of sets of positive measure. Show that there exists a countable subfamily $\{t_n\} \subset T$ such that $\mu\bigl(\bigcap_{n=1}^\infty X_{t_n}\bigr) > 0$.

HINT: in the separable measure algebra $\mathcal{A}/\mu$ the given family has a point of accumulation X' with $\mu(X') > 0$, since an uncountable set cannot have the only accumulation point corresponding to the equivalence class of measure zero sets; there exist indices t_n with $\mu(X' \triangle X_{t_n}) < \mu(X')2^{-n}$.

1.12.105.° Let $\mathcal{A}$ be the class of all subsets on the real line that are either at most countable or have at most countable complements. If the complement of a set $A \in \mathcal{A}$ is at most countable, then we set $\mu(A) = 1$, otherwise we set $\mu(A) = 0$. Then $\mathcal{A}$ is a σ-algebra and μ is a probability measure on $\mathcal{A}$, the collection $\mathcal{K}$ of all sets with at most countable complements is a compact class, approximating μ, but there is no class $\mathcal{K}' \subset \mathcal{A}$ approximating μ and having the property that every (not necessarily countable) collection in $\mathcal{K}'$ with empty intersection has a finite subcollection with empty intersection.

HINT: if such a class $\mathcal{K}'$ exists, then, for every $x \in \mathbb{R}^1$, there is a set $K_x \in \mathcal{K}'$ such that $K_x \subset \mathbb{R}^1\backslash\{x\}$ and $\mu(K_x) > 0$. Then $\mu(K_x) = 1$ and hence each finite intersection of such sets is nonempty, but the intersection of all K_x is empty.

1.12.106.° Let μ be an atomless probability measure on a measurable space $(X,\mathcal{A})$ and let $\mathcal{F} \subset \mathcal{A}$ be a countable family of sets of positive measure. Show that there exists a set $A \in \mathcal{A}$ such that $0 < \mu(A \cap F) < \mu(F)$ for all $F \in \mathcal{F}$.

HINT: let $\mathcal{F} = \{F_n\}$ and $\mathcal{F}_n = \{A \in \mathcal{A}\colon\ \mu(A\cap F_n) = 0 \text{ or } \mu(A\cap F_n) = \mu(F_n)\}$. Then $\mathcal{F}_n$ is closed in $\mathcal{A}/\mu$. Since μ is atomless, the sets $\mathcal{F}_n$ are nowhere dense in $\mathcal{A}/\mu$. By Baire's theorem the intersection of their complements is not empty.

1.12.107. Let $\mathbb{Q}$ be the set of all rational numbers equipped with the σ-algebra $2^{\mathbb{Q}}$ of all subsets and let the measure μ on $2^{\mathbb{Q}}$ with values in $[0,+\infty]$ be defined as the cardinality of a set. Let $\nu = 2\mu$. Show that the distinct measures μ and ν coincide on all open sets in $\mathbb{Q}$ (with the induced topology), and on all sets from the algebra that consists of finite disjoint unions of sets of the form $\mathbb{Q}\cap(a,b]$ and $\mathbb{Q}\cap(c,+\infty)$, where $a,b,c \in \mathbb{Q}$ or $c = -\infty$ (this algebra generates $2^{\mathbb{Q}}$).

HINT: nonempty sets of the above types are infinite.

1.12.108. Prove that there exists no countably additive measure defined on all subsets of the space $X = \{0,1\}^\infty$ that assumes only two values 0 and 1 and vanishes on all singletons.

HINT: let $X_n = \{(x_i) \in X\colon\ x_n = 0\}$; if such a measure μ exists, then, for any n, either $\mu(X_n) = 1$ or $\mu(X_n) = 0$; denote by Y_n that of the two sets X_n and $X\backslash X_n$ which has measure 1; then $\bigcap_{n=1}^\infty Y_n$ has measure 1 as well and is a singleton.

1.12.109. Prove that for every Borel set $E \subset \mathbb{R}^n$, there exists a Borel set $\widehat{E}$ that differs from E in a measure zero set and has the following property: for every point x at the boundary $\partial\widehat{E}$ of the set $\widehat{E}$ and every $r > 0$, one has

$$0 < \lambda_n\bigl(\widehat{E}\cap B(x,r)\bigr) < \omega_n r^n,$$

where $B(x,r)$ is the ball centered at x with the radius r and ω_n is the measure of the unit ball.

HINT: let E_0 be the set of all x such that $\lambda_n\bigl(E\cap B(x,r)\bigr)=0$ for some $r>0$, and let E_1 be the set of all x such that $\lambda_n\bigl(E\cap B(x,r)\bigr)=\omega_n r^n$ for some $r>0$. Consider $\widehat{E}=(E\cup E_1)\backslash E_0$ and use the fact that E_0 and E_1 are open.

1.12.110. Prove that every uncountable set $G\subset \mathbb{R}$ that is the intersection of a sequence of open sets contains a nowhere dense closed set Z of Lebesgue measure zero that can be continuously mapped onto $[0,1]$.

HINT: see Oxtoby [**733**, Lemma 5.1] or Chapter 6.

1.12.111. Prove that every uncountable set $G\subset \mathbb{R}$ that is the intersection of a sequence of open sets has cardinality of the continuum.

HINT: apply the previous exercise (see also Chapter 6, §6.1).

1.12.112. (i) Prove that the class of all Souslin subsets of the real line is obtained by applying the A-operation to the collection of all open sets. (ii) Show that in (i) it suffices to take the collection of all intervals with rational endpoints.

HINT: (i) use that every closed set is the intersection of a countable sequence of open sets and that $S(\mathcal{E})$ is closed with respect to the A-operation.

1.12.113. Prove that the classes of all Souslin and all Borel sets on the real line (or in the space $\mathbb{R}^n$) have cardinality of the continuum.

1.12.114. Let $(X,\mathcal{A},\mu)$ be a space with a finite nonnegative measure μ such that there exists a set E that is not μ-measurable. Prove that there exists $\varepsilon>0$ with the following property: if A and B are measurable, $E\subset A$, $X\backslash E\subset B$, then $\mu(A\cap B)\geq\varepsilon$.

HINT: assuming the converse one can find measurable sets A_n and B_n with $E\subset A_n$, $X\backslash E\subset B_n$, $\mu(A_n\cap B_n)<n^{-1}$; let $A=\bigcap_{n=1}^\infty A_n$, $B=\bigcap_{n=1}^\infty B_n$; then $E\subset A$, $X\backslash E\subset B$, $\mu(A\cap B)=0$, whence one has $\mu^*(E)+\mu^*(X\backslash E)\leq\mu(X)$ and hence we obtain the equality $\mu^*(E)+\mu^*(X\backslash E)=\mu(X)$.

1.12.115. Construct an example of a separable probability measure μ on a σ-algebra $\mathcal{A}$ such that, for every countably generated σ-algebra $\mathcal{E}\subset\mathcal{A}$, the completion of $\mathcal{E}$ with respect to μ is strictly smaller than $\mathcal{A}$.

HINT: see Example 9.8.1 in Chapter 9.

1.12.116. (Zink [**1052**]) Let (X,S,μ) be a measure space with a complete atomless separable probability measure μ and let $\mu^*(E)>0$. Then, there exist nonmeasurable sets E_1 and E_2 such that $E_1\cap E_2=\varnothing$, $E_1\cup E_2=E$ and one has $\mu^*(E_1)=\mu^*(E_2)=\mu^*(E)$.

1.12.117.° Let $\mathfrak{m}$ be a Carathéodory outer measure on a space X. Prove that a set A is Carathéodory measurable precisely when for all $B\subset A$ and $C\subset X\backslash A$ one has $\mathfrak{m}(B\cup C)=\mathfrak{m}(B)+\mathfrak{m}(C)$.

HINT: if A is Carathéodory measurable, then in the definition of measurability one can take $E=B\cup C$; if one has the indicated property, then an arbitrary set E can be written in the form $E=B\cup C$, $B=E\cap A$, $C=E\backslash A$.

1.12.118.° Suppose that $\mathfrak{m}_1$ and $\mathfrak{m}_2$ are outer measures on a space X. Show that $\max(\mathfrak{m}_1,\mathfrak{m}_2)$ is an outer measure too.

1.12.119.° (Young [**1029**]) Let $(X,\mathcal{A},\mu)$ be a measure space with a finite nonnegative measure μ. Prove that a set $A\subset X$ belongs to $\mathcal{A}_\mu$ precisely when for each set B disjoint with A one has the equality $\mu^*(A\cup B)=\mu^*(A)+\mu^*(B)$.

HINT: for the proof of sufficiency take $B = X\backslash A$; the necessity follows by the previous exercise.

1.12.120.° Let $\mathfrak{m}$ be a Carathéodory outer measure on a space X. Prove that for any $E \subset X$ the function $\mathfrak{m}_E(B) = \mathfrak{m}(B \cap E)$ is a Carathéodory outer measure and all $\mathfrak{m}$-measurable sets are $\mathfrak{m}_E$-measurable.

1.12.121. Let τ be an additive, but not countably additive nonnegative set function that is defined on the class of all subsets of $[0,1]$ and coincides with Lebesgue measure on all Lebesgue measurable sets (see Example 1.12.29). Show that the corresponding outer measure $\mathfrak{m}$ from Example 1.11.5 is identically zero under the continuum hypothesis.

HINT: Theorem 1.11.8 yields the $\mathfrak{m}$-measurability of all sets, $\mathfrak{m}$ is countably additive on $\mathfrak{M}_{\mathfrak{m}}$ and $\mathfrak{m}(\{x\}) = 0$ for each x.

1.12.122. Prove that if $\mathfrak{X} \subset \mathfrak{M}_{\mathfrak{m}}$, then Method I from Example 1.11.5 gives a regular outer measure.

1.12.123. Let $\mathcal{S}$ be a collection of subsets of a set X, closed with respect to finite unions and finite intersections and containing the empty set, i.e., a lattice of sets (e.g., the class of all closed sets or the class of all open sets in $[0,1]$).

(i) Suppose that on $\mathcal{S}$ we have a modular set function m, i.e., $m(\varnothing) = 0$ and $m(A \cup B) + m(A \cap B) = m(A) + m(B)$ for all $A, B \in \mathcal{S}$. Show that by the equality $m(A\backslash B) = m(A) - m(B)$, $A, B \in \mathcal{S}$, $B \subset A$, the function m uniquely extends to an additive set function (which, in particular, is well-defined) on the semiring formed by the differences of elements in $\mathcal{S}$ (see Exercise 1.12.51), and then uniquely extends to an additive set function on the ring generated by $\mathcal{S}$.

(ii) Give an example showing that in (i) one cannot replace the modularity by the additivity even if m is nonnegative, monotone and subadditive on $\mathcal{S}$.

HINT: (i) use Exercise 1.12.51 and Proposition 1.3.10; in order to verify that m is well-defined we observe that if $A_1\backslash A_1' = A_2\backslash A_2'$, where $A_i, A_i' \in \mathcal{S}$, $A_i' \subset A_i$, then $m(A_1) + m(A_2') = m(A_2) + m(A_1')$ because $A_1 \cup A_2' = A_2 \cup A_1'$, $A_1 \cap A_2' = A_1' \cap A_2$, which is easily verified; see the details in Kelley, Srinivasan [**502**, Chapter 2, p. 23, Theorem 2]. (ii) Take $X = \{0,1,2\}$ and $\mathcal{S}$ consisting of X, $\varnothing$, $\{0,1\}$, $\{1,2\}$, $\{1\}$ with $m(X) = 2$, $m(\varnothing) = 0$ and $m = 1$ on all other sets in $\mathcal{S}$.

1.12.124. Suppose that $\mathcal{F}$ is a family of subsets of a set X, $\varnothing \in \mathcal{F}$. Let $\tau\colon \mathcal{F} \to [0,+\infty]$ be a set function with $\tau(\varnothing) = 0$. Let us define τ_* on all sets $A \subset X$ by formula (1.12.8).

(i) Prove that if $A_1, \ldots, A_n \subset X$ are disjoint sets and $A_1 \cup \cdots \cup A_n \subset A$, then one has $\tau_*(A) \geq \sum_{j=1}^n \tau_*(A_j)$.

(ii) Prove that τ_* coincides with τ on $\mathcal{F}$ if and only if, for all pairwise disjoint sets $F_1, \ldots, F_n \in \mathcal{F}$ and all $F \in \mathcal{F}$ with $\bigcup_{j=1}^n F_j \subset F$, one has $\tau(F) \geq \sum_{j=1}^n \tau(F_j)$.

(iii) Prove that if τ satisfies the condition in (ii) and the class $\mathcal{F}$ is closed with respect to finite unions of disjoint sets, then

$$\tau_*(A) = \sup\{\tau(F),\ F \in \mathcal{F},\ F \subset A\}, \quad \forall\, A \subset X.$$

HINT: (i) Let $\tau_*(A) < \infty$ and $\varepsilon > 0$. For every i, there exist disjoint sets $F_{ij} \in \mathcal{F}$, $j \leq n(i)$, such that $\bigcup_{j=1}^{n(i)} F_{ij} \subset A_i$ and $\tau_*(A_i) \leq \varepsilon 2^{-i} + \sum_{j=1}^{n(i)} \tau(F_{ij})$. All sets F_{ij} are pairwise disjoint and are contained in A. Therefore,

$$\sum_{i=1}^{n} \tau_*(A_i) \le \sum_{i=1}^{n} \varepsilon 2^{-i} + \sum_{i=1}^{n}\sum_{j=1}^{n(i)} \tau(F_{ij}) \le \varepsilon + \tau_*(A),$$

whence we obtain the claim, since ε is arbitrary.

(ii) Let $F_j, F \in \mathcal{F}$, $F_j \subset F$, where the sets F_j are pairwise disjoint. Then the inequality $\tau(F) \ge \sum_{j=1}^{n} \tau(F_j)$ yields the inequality $\tau(F) \ge \tau_*(F)$. Since the reverse inequality is obvious from the definition, we obtain the equality $\tau_* = \tau$ on $\mathcal{F}$. On the other hand, this equality obviously implies the indicated inequality.

(iii) Let $F_1, \ldots, F_n \in \mathcal{F}$ be disjoint sets and let $E := \bigcup_{j=1}^{n} F_j \subset A$. Then, by hypothesis, we have $E \in \mathcal{F}$ and $\sum_{j=1}^{n} \tau(F_j) \le \tau(E) \le \sup\bigl\{\tau(F)\colon\ F \in \mathcal{F}, F \subset A\bigr\}$, whence $\tau_*(A) \le \sup\bigl\{\tau(F)\colon\ F \in \mathcal{F}, F \subset A\bigr\}$; the reverse inequality is trivial.

1.12.125. Let $\mathcal{F}$ and τ be the same as in the previous exercise. (i) Prove that the outer measure τ^* coincides with τ on $\mathcal{F}$ precisely when $\tau(F) \le \sum_{n=1}^{\infty} \tau(F_n)$ whenever $F, F_n \in \mathcal{F}$ and $F \subset \bigcup_{n=1}^{\infty} F_n$.

(ii) Prove that if the condition in (i) is fulfilled and the class $\mathcal{F}$ is closed with respect to countable unions, then

$$\tau^*(A) = \inf\{\tau(F),\ F \in \mathcal{F},\ A \subset F\}, \quad \forall\, A \subset X.$$

Hint: the proof is similar to the reasoning in the previous exercise.

1.12.126. Suppose that $\mathcal{F}$ is a class of subsets of a space X, $\varnothing \in \mathcal{F}$. Let $\tau\colon \mathcal{F} \to [0, +\infty]$ be a set function with $\tau(\varnothing) = 0$. Prove that the following conditions are equivalent:

(i) τ^* coincides with τ on $\mathcal{F}$ and $\mathcal{F} \subset \mathfrak{M}_{\tau^*}$;

(ii) $\tau(A) = \tau^*(A \cap B) + \tau^*(A \backslash B)$ for all $A, B \in \mathcal{F}$.

Hint: (i) implies (ii) by the additivity of τ^* on $\mathfrak{M}_{\tau^*}$. Let (ii) be fulfilled. Letting $B = \varnothing$, we get $\tau(A) = \tau^*(A)$, $A \in \mathcal{F}$. Suppose that $F \in \mathcal{F}$ and $E \subset X$. Let $F_j \in \mathcal{F}$ and $E \subset \bigcup_{j=1}^{\infty} F_j$. Then

$$\sum_{j=1}^{\infty} \tau(F_j) = \sum_{j=1}^{\infty} \tau^*(F_j \cap F) + \sum_{j=1}^{\infty} \tau(F_j \backslash F) \ge \tau^*(E \cap F) + \tau^*(E \backslash F).$$

Taking the infimum over $\{F_j\}$, we obtain $\tau^*(E) \ge \tau^*(E \cap F) + \tau^*(E \backslash F)$, i.e., we have $F \in \mathfrak{M}_{\tau^*}$.

1.12.127. Suppose that $\mathcal{F}$ is a class of subsets of a space X, $\varnothing \in \mathcal{F}$. Let $\tau\colon \mathcal{F} \to [0, +\infty]$ be a set function with $\tau(\varnothing) = 0$. Denote by τ_* the corresponding inner measure (see formula (1.12.8)). Prove that the following conditions are equivalent:

(i) τ_* coincides with τ on $\mathcal{F}$ and $\mathcal{F} \subset \mathfrak{M}_{\tau_*}$;

(ii) $\tau(A) = \tau_*(A \cap B) + \tau_*(A \backslash B)$, $\forall\, A, B \in \mathcal{F}$.

Hint: the proof is completely analogous to the previous exercise, one has only take finitely many disjoint $F_j \subset A$; see also Glazkov [**360**], Hoffmann-Jørgensen [**440**, 1.26].

1.12.128. (i) Show that if in the situation of the previous exercise we have one of the equivalent conditions (i) and (ii), then on the algebra $\mathcal{A}_{\mathcal{F}}$ generated by $\mathcal{F}$, there exists an additive set function τ_0 that coincides with τ on $\mathcal{F}$.

(ii) Show that if, in addition to the hypotheses in (i), it is known that

$$\tau_*(F) \le \sum_{n=1}^{\infty} \tau_*(F_n) \quad \text{whenever } F, F_n \in \mathcal{A}_{\mathcal{F}} \text{ and } F \subset \bigcup_{n=1}^{\infty} F_n,$$

then there exists a countably additive measure μ on $\sigma(\mathcal{F})$ that coincides with τ on $\mathcal{F}$.

HINT: according to Theorem 1.11.4, the function τ_* is additive on $\mathfrak{M}_{\tau_*}$ and $\mathfrak{M}_{\tau_*}$ is an algebra. Since the algebra $\mathfrak{M}_{\tau_*}$ contains $\mathcal{F}$ by hypothesis, it also contains the algebra generated by $\mathcal{F}$. The second claim follows by the cited theorem, too.

1.12.129. Let $(X, \mathcal{A}, \mu)$ be a measure space, where $\mathcal{A}$ is a σ-algebra and μ is a countably additive measure with values in $[0, +\infty]$. Denote by $\mathfrak{L}_\mu$ the class of all sets $E \subset X$ for each of which there exist two sets $A_1, A_2 \in \mathcal{A}$ with $A_1 \subset E \subset A_2$ and $\mu(A_2 \backslash A_1) = 0$.

(i) Show that $\mathfrak{L}_\mu$ is a σ-algebra, coincides with $\mathcal{A}_\mu$ and belongs to $\mathfrak{M}_{\mu^*}$.

(ii) Show that if the measure μ is σ-finite, then $\mathfrak{L}_\mu$ coincides with $\mathfrak{M}_{\mu^*}$.

(iii) Let $X = [0,1]$, let $\mathcal{A}$ be the σ-algebra generated by all singletons, and let the measure μ with values in $[0, +\infty]$ be defined as follows: $\mu(A)$ is the cardinality of A, $A \in \mathcal{A}$. Show that $\mathfrak{M}_{\mu^*}$ contains all sets, but $[0, 1/2] \notin \mathfrak{L}_\mu$.

HINT: (iii) show that $\mu^*(A)$ is the cardinality of A and that $\mathfrak{L}_\mu = \mathcal{A}$, by using that nonempty sets have measure at least 1.

1.12.130. Let us consider the following modification of Example 1.11.5. Let $\mathfrak{X}$ be a family of subsets of a set X such that $\varnothing \in \mathfrak{X}$. Suppose that we are given a function $\tau\colon \mathfrak{X} \to [0, +\infty]$ with $\tau(\varnothing) = 0$. Set

$$\widetilde{\mathfrak{m}}(A) = \inf\Bigl\{\sum_{n=1}^{\infty} \tau(X_n)\colon\ X_n \in \mathfrak{X}, A \subset \bigcup_{n=1}^{\infty} X_n\Bigr\}$$

if such sets X_n exist and otherwise let $\widetilde{\mathfrak{m}}(A) = \sup \widetilde{\mathfrak{m}}(A')$, where sup is taken over all sets $A' \subset A$ that can be covered by a sequence of sets in $\mathfrak{X}$.

(i) Show that $\widetilde{\mathfrak{m}}$ is an outer measure.

(ii) Let $X = [0,1]{\times}[0,1]$, $\mathfrak{X} = \bigl\{[a,b){\times}t, a, b, t \in [0,1], a \le b\bigr\}$, $\tau([a,b){\times}t) = b - a$. Let $\mathfrak{m}$ be given by formula (1.11.5). Show that $\mathfrak{m}$ and $\widetilde{\mathfrak{m}}$ do not coincide and that there exists a set $E \in \mathfrak{M}_{\mathfrak{m}} \cap \mathfrak{M}_{\widetilde{\mathfrak{m}}}$ such that $\mathfrak{m}(E) \ne \widetilde{\mathfrak{m}}(E)$.

HINT: (i) is verified similarly to the case of $\mathfrak{m}$; (ii) for E take the diagonal in the square.

1.12.131. Let μ be a measure with values in $[0, +\infty]$ defined on a measurable space $(X, \mathcal{A})$. The measure μ is called decomposable if there exists a partition of X into pairwise disjoint sets $X_\alpha \in \mathcal{A}$ of finite measure (indexed by elements α of some set Λ) with the following properties: (a) if $E \cap X_\alpha \in \mathcal{A}$ for all α, then $E \in \mathcal{A}$, (b) $\mu(E) = \sum\limits_\alpha \mu(E \cap X_\alpha)$ for each set $E \in \mathcal{A}$, where convergence of the series $\sum\limits_\alpha c_\alpha$, $c_\alpha \ge 0$, to a finite number s means by definition that among the numbers c_α at most countably many are nonzero and the corresponding series converges to s, and the divergence of such a series to $+\infty$ means the divergence of some of its countable subseries.

(i) Give an example of a measure that is not decomposable.

(ii) Show that a measure μ is decomposable precisely when there exists a partition of X into disjoint sets X_α of positive measure having property (a) and property (b'): if $A \in \mathcal{A}$ and $\mu(A \cap X_\alpha) = 0$ for all α, then $\mu(A) = 0$.

1.12.132. Let μ be a measure with values in $[0, +\infty]$ defined on a measurable space $(X, \mathcal{A})$. The measure μ is called semifinite if every set of infinite measure has a subset of finite positive measure.

(i) Give an example of a measure with values in $[0, +\infty]$ that is not semifinite.

(ii) Give an example of a semifinite measure that is not σ-finite.

(iii) Prove that for any measure μ with values in $[0, +\infty]$, defined on a σ-algebra $\mathcal{A}$, the formula $\mu_0(A) := \sup\{\mu(B)\colon\ B \subset A, B \in \mathcal{A}, \mu(B) < \infty\}$ defines a semifinite measure with values in $[0, +\infty]$ and μ is semifinite precisely when $\mu = \mu_0$.

(iv) Show that every decomposable measure is semifinite.

(v) Give an example of a semifinite measure μ with values in $[0, +\infty]$ that is defined on an algebra $\mathcal{A}$ and has infinitely many semifinite extensions to $\sigma(\mathcal{A})$.

Hint: (v) let $X = \mathbb{R}^1$, let $\mathcal{A}$ be the class of all finite sets and their complements, and let $\mu(A)$ be the cardinality (denoted Card) of $A \cap \mathbb{Q}$. For any $s \geq 0$ and $A \in \sigma(\mathcal{A})$, let $\mu_s(A) = \mathrm{Card}(A \cap \mathbb{Q})$ if $A \cap (\mathbb{R}^1 \backslash \mathbb{Q})$ is at most countable, $\mu_s(A) = s + \mathrm{Card}(A \cap \mathbb{Q})$ if $(\mathbb{R}^1 \backslash A) \cap (\mathbb{R}^1 \backslash \mathbb{Q})$ is at most countable.

1.12.133. Let μ be a measure μ with values in $[0, +\infty]$ defined on a measurable space $(X, \mathcal{A})$. A set E is called locally measurable if $E \cap A \in \mathcal{A}$ for every $A \in \mathcal{A}$ with $\mu(A) < \infty$. The measure μ is called saturated if every locally measurable set belongs to $\mathcal{A}$.

(i) Let $X = \mathbb{R}$, $\mathcal{A} = \{\mathbb{R}, \varnothing\}$, $\mu(\mathbb{R}) = +\infty$, $\mu(\varnothing) = 0$. Show that μ is a complete measure with values in $[0, +\infty]$ that is not saturated.

(ii) Show that every σ-finite measure is saturated.

(iii) Show that locally measurable sets form a σ-algebra.

(iv) Show that every measure with values in $[0, +\infty]$ can be extended to a saturated measure on the σ-algebra $\mathcal{L}$ of all locally measurable sets by the formula $\overline{\mu}(E) = \mu(E)$ if $E \in \mathcal{A}$, $\overline{\mu}(E) = +\infty$ if $E \notin \mathcal{A}$.

(v) Construct an example showing that $\overline{\mu}$ may not be a unique saturated extension of μ to the σ-algebra $\mathcal{L}$.

Hint: (i) observe that every set in X is locally measurable with respect to μ; (iii) use that $(X \backslash E) \cap A = A \backslash (A \cap E)$; (v) let $\mu_0(A) = 0$ if A is countable and $\mu_0(A) = \infty$ if A is uncountable; observe that μ_0 is saturated.

1.12.134. Let $(X, \mathcal{A}, \mu)$ be a measure space, where μ takes values in $[0, +\infty]$. The measure μ is called Maharam (or localizable) if μ is semifinite and each collection $\mathcal{M} \subset \mathcal{A}$ has the essential supremum in the following sense: there exists a set $E \in \mathcal{A}$ such that all sets $M \backslash E$, where $M \in \mathcal{M}$, have measure zero and if $E' \in \mathcal{A}$ is another set with such a property, then $E \backslash E'$ is a measure zero set.

(i) Prove that every decomposable measure is Maharam.

(ii) Give an example of a complete Maharam measure that is not decomposable.

Hint: (i) let the sets X_α, $\alpha \in \Lambda$, give a decomposition of the measure space $(X, \mathcal{A}, \mu)$ and $\mathcal{M} \subset \mathcal{A}$. Denote by $\mathcal{F}$ the family of all sets $F \in \mathcal{A}$ with $\mu(F \cap M) = 0$ for all $M \in \mathcal{M}$. It is clear that $\mathcal{F}$ contains the empty set and admits countable unions. For every α, let $c_\alpha := \sup\{\mu(F \cap X_\alpha), F \in \mathcal{F}\}$ and choose $F_{\alpha,n} \in \mathcal{F}$ such that $\lim\limits_{n\to\infty} \mu(F_{\alpha,n} \cap X_\alpha) = c_\alpha$. Let $F_\alpha := \bigcup_{n=1}^\infty F_{\alpha,n}$ and $\Psi := \bigcup_{\alpha\in\Lambda}(F_\alpha \cap X_\alpha)$. Then $\Psi \cap X_\alpha = F_\alpha$ and hence $\Psi \in \mathcal{A}$. Therefore, $E := X \backslash \Psi \in \mathcal{A}$. For any $M \in \mathcal{M}$ we have

$$\mu(M \backslash E) = \mu(M \cap \Psi) = \sum_\alpha \mu(M \cap \Psi \cap X_\alpha) = \sum_\alpha \mu(M \cap F_\alpha \cap X_\alpha) = 0$$

by the definition of $\mathcal{F}$. If E' is another set with such a property, then $X\backslash E' \in \mathcal{F}$ and $\Psi' := \Psi \cup (X\backslash E') \in \mathcal{F}$. Now it is readily shown that $\mu(\Psi \cap X_\alpha) = \mu(\Psi' \cap X_\alpha)$ for all α, whence $\mu\bigl((\Psi'\backslash\Psi) \cap X_\alpha\bigr) = 0$, i.e., $\mu(\Psi'\backslash\Psi) = 0$ and $\mu(E\backslash E') = 0$. (ii) Examples with various additional properties can be found in Fremlin [**327**, §216].

1.12.135. A measure with values in $[0, +\infty]$ is called locally determined if it is semifinite and saturated. Let μ be a measure with values in $[0, +\infty]$ defined on a measurable space $(X, \mathcal{A})$. Let $\mathcal{L}_\mu$ be the σ-algebra of locally $\mathcal{A}_\mu$-measurable sets, i.e., all sets L such that $L \cap A \in \mathcal{A}_\mu$ for all $A \in \mathcal{A}_\mu$ with $\mu(A) < \infty$. Let

$$\widetilde{\mu}(L) = \sup\{\mu(L \cap A)\colon\ A \in \mathcal{A}_\mu, \mu(A) < \infty\}, \quad L \in \mathcal{L}_\mu.$$

(i) Show that the measure $\widetilde{\mu}$ is locally determined and complete and that one has $\widetilde{\mu}(A) = \mu(A)$ whenever $A \in \mathcal{A}_\mu$ and $\mu(A) < \infty$.

(ii) Show that if μ is decomposable, then so is $\widetilde{\mu}$ and in this case $\widetilde{\mu}$ coincides with the completion of μ.

(iii) Show that if μ is Maharam, then so is $\widetilde{\mu}$.

(iv) Show that the measure μ is complete and locally determined precisely when one has $\mu = \widetilde{\mu}$.

Hint: the detailed verification of these simple assertions can be found, e.g., in Fremlin [**327**].

1.12.136. Let $(X, \mathcal{A})$ be a measurable space and let a measure μ on $\mathcal{A}$ with values in $[0, +\infty]$ be complete and locally determined. Suppose that there exists a family $\mathcal{D}$ of pairwise disjoint sets of finite measure in $\mathcal{A}$ such that if $E \in \mathcal{A}$ and $\mu(E \cap D) = 0$ for all $D \in \mathcal{D}$, then $\mu(E) = 0$. Prove that the measure μ is decomposable.

Hint: see Fremlin [**327**, §213O].

1.12.137. Let X be a set of cardinality of the continuum and let Y be a set of cardinality greater than that of the continuum. For every $E \subset X \times Y$, the sets $\{(a, y) \in E\}$ with fixed $a \in X$ will be called vertical sections of E, and the sets $\{(x, b) \in E\}$ with fixed $b \in Y$ will be called horizontal sections of E. Denote by $\mathcal{A}$ the class of all sets $A \subset X \times Y$ such that all their horizontal and vertical sections are either at most countable or have at most countable complements in the corresponding sections of $X \times Y$. Let $\gamma(A)$ be the number of those horizontal sections of the complement of A that are at most countable. Similarly, by means of vertical sections we define the function $v(A)$. Let $\mu(A) = \gamma(A) + v(A)$.

(i) Prove that $\mathcal{A}$ is a σ-algebra and that γ, v, and μ are countably additive measures with values in $[0, +\infty]$.

(ii) Prove that μ is semifinite in the sense of Exercise 1.12.132.

(iii) Prove that μ is not decomposable in the sense of Exercise 1.12.131.

Hint: (ii) if $(X \times Y)\backslash A$ has infinite number of finite or countable horizontal sections, then, given $N \in \mathbb{N}$, one can take points $y_1, \ldots, y_N \in Y$, giving such sections; let us take the set B such that the horizontal sections of its complement at the points y_i coincide with the corresponding sections of the complement of A, and all other sections of the complement of B coincide with $X \times y$; then $B \subset A$ and $\gamma(B) = N$, $v(B) = 0$. (iii) If sets E_α give a partition of $X \times Y$ and $\mu(E_\alpha) < \infty$, then the cardinality of this family of sets cannot be smaller than that of Y. Indeed, otherwise, since E_α is contained in a finite union of sets of the form $a \times Y$ and $X \times b$, one would find a set $X \times y$ whose intersection with every E_α is a set with the uncountable complement in $X \times y$, whence $\mu\bigl((X \times y) \cap E_\alpha\bigr) = 0$ for all α, but we have $\mu(X\times y)=1$. On the other hand, for every $x\in X$, there is a unique set E_{α_x} with $\mu\bigl((x\times Y)\cap E_{\alpha_x}\bigr)=1$, and since the complement of $(x\times Y)\cap E_{\alpha_x}$ in $x\times Y$ is at most countable, the set $x\times Y$ meets at most countably many sets E_α. Hence the cardinality of the family $\{E_\alpha\}$ is that of the continuum, which is a contradiction.

1.12.138. Let $X=[0,1]\times\{0,1\}$ and let $\mathcal{A}$ be the class of all sets $E\subset X$ such that the sections $E_x:=\{y\colon (x,y)\in E\}$ are either empty or coincide with $\{0,1\}$ for all x, excepting possibly the points of an at most countable set. Show that $\mathcal{A}$ is a σ-algebra and the function μ that to every set E assigns the cardinality of the intersection of E with the first coordinate axis, is a complete and semifinite countably additive measure with values in $[0,+\infty]$, but the measure generated by the outer measure μ^* is not semifinite.

1.12.139. (Luther [**639**]) Let μ be a measure with values in $[0,+\infty]$ defined on a ring $\mathcal{R}$, let $\overline{\mu}$ be the restriction of μ^* to the σ-ring $\mathcal{S}$ generated by $\mathcal{R}$, and let $\mathcal{R}_0$ and $\mathcal{S}_0$ be the subclasses in $\mathcal{R}$ and $\mathcal{S}$ consisting of all sets of finite measure. Set

$$\widetilde{\mu}(E)=\limsup\{\overline{\mu}(P\cap E),P\in\mathcal{R}_0\},\ \ E\in\mathcal{S}.$$

(i) Prove that the following conditions are equivalent: (a) μ is semifinite, (b) $\widetilde{\mu}$ is an extension of μ to $\mathcal{S}$, (c) any measure ν on $\mathcal{S}$ with values in $[0,+\infty]$ that agrees with μ on $\mathcal{R}_0$ coincides with μ on $\mathcal{R}$.

(ii) Show that any measure ν on $\mathcal{S}$ with values in $[0,+\infty]$ that agrees with μ on $\mathcal{R}_0$, coincides with $\widetilde{\mu}$ and $\overline{\mu}$ on $\mathcal{S}_0$, and that $\widetilde{\mu}\le\nu\le\overline{\mu}$ on $\mathcal{S}$.

(iii) Prove that the following conditions are equivalent: (a) $\overline{\mu}$ is semifinite, (b) μ is semifinite and has a unique extension to $\mathcal{S}$, (c) $\widetilde{\mu}=\overline{\mu}$, (d) for all $E\in\mathcal{S}$ one has $\overline{\mu}(E)=\limsup\{\overline{\mu}(P\cap E),P\in\mathcal{R}_0\}$.

(iv) Prove that if the measure $\overline{\mu}$ is σ-finite, then μ has a unique extension to $\mathcal{S}$.

(v) Give an example showing that in (iv) it is not sufficient to require the existence of some σ-finite extension of μ.

1.12.140. (Luther [**640**]) Let μ be a measure with values in $[0,+\infty]$ defined on a σ-ring $\mathcal{R}$. Prove that $\mu=\mu_1+\mu_2$, where μ_1 is a semifinite measure on $\mathcal{R}$, the measure μ_2 can assume only the values 0 and ∞, and in every set $R\in\mathcal{R}$ there exists a subset $R'\in\mathcal{R}$ such that $\mu_1(R')=\mu_1(R)$ and $\mu_2(R')=0$.

1.12.141. Let $\mathcal{E}_1$ and $\mathcal{E}_2$ be two algebras of subsets of Ω and let μ_1,μ_2 be two additive real functions on $\mathcal{E}_1$ and $\mathcal{E}_2$, respectively (or μ_1,μ_2 take values in the extended real line and vanish at $\varnothing$). (a) Show that the equality $\mu_1(E)=\mu_2(E)$ for all $E\in\mathcal{E}_1\cap\mathcal{E}_2$ is necessary and sufficient for the existence of an additive function μ that extends μ_1 and μ_2 to some algebra $\mathcal{F}$ containing $\mathcal{E}_1$ and $\mathcal{E}_2$. (b) Show that if $\mu_1,\mu_2\ge 0$, then the existence of a common nonnegative extension μ is equivalent to the following relations: $\mu_1(C)\ge\mu_2(D)$ for all $C\in\mathcal{E}_1$, $D\in\mathcal{E}_2$ with $D\subset C$ and $\mu_1(E)\le\mu_2(F)$ for all $E\in\mathcal{E}_1$, $F\in\mathcal{E}_2$ with $E\subset F$.

HINT: see Rao, Rao [**786**, §3.6, p. 82].

1.12.142. Let $(X,\mathcal{A},\mu)$ be a probability space and let μ^* be the corresponding outer measure. For a set $E\subset X$, we denote by $\mathfrak{m}_E$ the restriction of μ^* to the class of all subsets of E. Show that $\mathfrak{m}_E$ coincides with the outer measure on the space E generated by the restriction μ_E of μ to E in the sense of Definition 1.12.11. In particular, $\mathfrak{m}_E$ is a regular Carathéodory outer measure.

HINT: let $\widetilde{E}$ be a measurable envelope of E; for any set $B\subset E$ one has

$$\mathfrak{m}_E(B)=\inf\{\mu(A)\colon\ A\in\mathcal{A},B\subset A\}.$$

By the definition of μ_E we have

$$\mu_E^*(B) = \inf\{\mu_E(C)\colon\ C \in \mathcal{A}_E, B \subset C\} = \inf\{\mu(A \cap \widetilde{E})\colon\ A \in \mathcal{A}, B \subset A \cap E\}.$$

Clearly, one has $\mathfrak{m}_E(B) \geq \mu_E^*(B)$. On the other hand, given $\varepsilon > 0$, we find a set $A_\varepsilon \in \mathcal{A}$ such that $\mu(A_\varepsilon \cap \widetilde{E}) < \mu_E^*(B) + \varepsilon$. Hence $\mu(A_\varepsilon) < \mu_E^*(B) + \varepsilon$ and $B \subset A_\varepsilon$, which yields the estimate $\mathfrak{m}_E(B) \leq \mu_E^*(B) + \varepsilon$. Hence $\mathfrak{m}_E(B) \leq \mu_E^*(B)$.

1.12.143. Suppose that μ is a measure with values in $[0, +\infty]$ on a measurable space $(X, \mathcal{A})$. Let μ^* and μ_* be the corresponding outer and inner measures and let $\mathfrak{m} := (\mu^* + \mu_*)/2$.

(i) (Carathéodory [**164**, p. 693]) Show that $\mathfrak{m}$ is a Carathéodory outer measure. Denote by ν the measure generated by $\mathfrak{m}$.

(ii) Let $X = \{0, 1\}$, $\mathcal{A} = \{X, \varnothing\}$, $\mu(X) = 1$. Show that $\mu \neq \nu$.

(iii) (Fremlin [**324**]) Prove that if μ is Lebesgue measure on $[0, 1]$, then $\mu = \nu$.

1.12.144. Let $\mathfrak{m}$ be a Carathéodory outer measure on a space X and let $\varphi\colon\ [0, +\infty] \to [0, +\infty)$ be a bounded concave function such that $\varphi(0) = 0$ and $\varphi(t) > 0$ if $t \neq 0$. Let $d(A, B) = \varphi\bigl(\mathfrak{m}(A \triangle B)\bigr)$, $A, B \in \mathfrak{M}_{\mathfrak{m}}$. Denote by $\widetilde{\mathfrak{M}}_\mu$ the factor-space of the space $\mathfrak{M}_{\mathfrak{m}}$ by the ring of $\mathfrak{m}$-zero sets. Show that $(\widetilde{\mathfrak{M}}_\mu, d)$ is a complete metric space.

1.12.145. (Steinhaus [**910**]) Let E be a set of positive measure on the real line. Prove that, for every finite set F, the set E contains a subset similar to F, i.e., having the form $c + tF$, where $t \neq 0$.

1.12.146. (i) Let μ be an atomless probability measure on a measurable space $(X, \mathcal{A})$. Show that every point $x \in X$ belongs to $\mathcal{A}_\mu$ and has μ-measure zero.

(ii) (Marczewski [**651**]) Prove that if a probability measure μ on a measurable space $(X, \mathcal{A})$ is atomless, then there exist nonempty sets of μ-measure zero.

Hint: (i) let us fix a point $x \in X$ and take its measurable envelope E. Then $\mu(E) = 0$. Indeed, if $c = \mu(E) > 0$, we find a set $A \in \mathcal{A}$ such that $A \subset E$ and $\mu(A) = c/2$, which is possible since μ is atomless. Then either $x \in A$ or $x \in E\backslash A$ and $\mu(A) = \mu(E\backslash A) = c/2$, which contradicts the fact that E is a measurable envelope of x. Alternatively, one can use the following fact that will be established in §9.1 of Chapter 9: there exists a function f from X to $[0, 1]$ such that for every $t \in [0, 1]$ one has $\mu(x\colon\ f(x) < t) = t$. It follows that for every $t \in [0, 1]$ the set $f^{-1}(t)$ has μ-measure zero. Assertion (ii) easily follows. Moreover, by the second proof, there exists an uncountable set of μ-measure zero.

1.12.147. (Kindler [**517**]) Let $\mathcal{S}$ be a family of subsets of a set Ω with $\varnothing \in \mathcal{S}$ and let $\alpha, \beta\colon\ \mathcal{S} \to (-\infty, +\infty]$ be two set functions vanishing at $\varnothing$. Prove that the following conditions are equivalent:

(i) there exists an additive set function μ on the set of all subsets of Ω taking values in $(-\infty, +\infty]$ and satisfying the condition $\alpha \leq \mu|_{\mathcal{S}} \leq \beta$;

(ii) if $A_i, B_j \in \mathcal{S}$ and $\sum_{i=1}^n I_{A_i} = \sum_{j=1}^m I_{B_j}$, then $\sum_{i=1}^n \alpha(A_i) \leq \sum_{j=1}^m \beta(B_j)$.

1.12.148. Prove Proposition 1.12.36. Moreover, show that there is a nonnegative additive function α on the set of all subsets of X with $\alpha|_{\mathfrak{R}} \leq \beta$ and $\alpha(X) = \beta(X)$.

Hint: (a) by induction on n we prove the following fact: if $R_1, \ldots, R_n \in \mathfrak{R}$, then there are $R_1', \ldots, R_n' \in \mathfrak{R}$ such that $R_1' \subset R_2' \subset \ldots \subset R_n'$, $\sum_{i=1}^n I_{R_i} = \sum_{i=1}^n I_{R_i'}$ and $\sum_{i=1}^n \beta(R_i) \geq \sum_{i=1}^n \beta(R_i')$. For the inductive step to $n+1$, given $R_1, \ldots, R_{n+1} \in \mathfrak{R}$, set $S_{n+1} = R_{n+1}$ and use the inductive hypothesis to find $S_1, \ldots, S_n \in \mathfrak{R}$ such that $S_1 \subset \ldots \subset S_n$, $\sum_{i=1}^n I_{R_i} = \sum_{i=1}^n I_{S_i}$ and $\sum_{i=1}^n \beta(R_i) \geq \sum_{i=1}^n \beta(S_i)$. Now set $S_n' = S_{n+1} \cap S_n$, $S_i' = S_i$ for $i < n$. There are $R_1', \ldots, R_n' \in \mathfrak{R}$ such that $R_1' \subset R_2' \subset \ldots \subset R_n'$, $\sum_{i=1}^n I_{S_i'} = \sum_{i=1}^n I_{R_i'}$ and $\sum_{i=1}^n \beta(S_i') \geq \sum_{i=1}^n \beta(R_i')$. Let $R_{n+1}' = S_{n+1}' = S_n \cup S_{n+1}$. Then $S_i, S_i', R_i' \in \mathfrak{R}$. As $I_{R_n'} \leq \sum_{i=1}^n I_{S_i'}$, one has $R_n' \subset \bigcup_{i=1}^n S_i' \subset S_n \subset R_{n+1}'$. In addition,

$$\sum_{i=1}^{n+1} I_{R_i'} = \sum_{i=1}^{n} I_{S_i'} + I_{S_{n+1}'} = \sum_{i=1}^{n-1} I_{S_i} + I_{S_n \cap S_{n+1}} + I_{S_n \cup S_{n+1}} = \sum_{i=1}^{n+1} I_{S_i} = \sum_{i=1}^{n+1} I_{R_i}.$$

Finally,

$$\begin{aligned}\sum_{i=1}^{n+1} \beta(R_i') &\leq \sum_{i=1}^{n} \beta(S_i') + \beta(S_{n+1}') = \sum_{i=1}^{n-1} \beta(S_i) + \beta(S_n \cap S_{n+1}) + \beta(S_n \cup S_{n+1}) \\ &\leq \sum_{i=1}^{n-1} \beta(S_i) + \beta(S_n) + \beta(S_{n+1}) = \sum_{i=1}^{n} \beta(S_i) + \beta(S_{n+1}) \leq \sum_{i=1}^{n+1} \beta(R_i).\end{aligned}$$

(b) We may assume that $\beta(X) = 1$. Let us show that if $R_1, \ldots, R_n \in \mathfrak{R}$ are such that $\sum_{i=1}^n I_{R_i}(x) \geq m$ for all x, where $m \in \mathbb{N}$, then $\sum_{i=1}^n \beta(R_i) \geq m$. Let R_i' be as in (a). It suffices to verify our claim for the sets R_i'. As $R_i' \subset R_{i+1}'$, one has $R_n' = \cdots = R_{n-m+1}' = X$. Hence $\beta(R_j') = 1$ for $j \geq n+m-1$.

(c) On the linear space L of finitely valued functions on X we set

$$p(f) = \inf\Bigl\{\sum_{i=1}^n \alpha_i \beta(R_i)\colon\ R_i \in \mathfrak{R}, \alpha_i \geq 0, f \leq \sum_{i=1}^n \alpha_i I_{R_i}\Bigr\}.$$

It is readily verified that $p(f+g) \leq p(f) + p(g)$ and $p(\alpha f) = \alpha p(f)$ for all $f, g \in L$, $\alpha \geq 0$. In addition, $p(1) \geq 1$. Indeed, otherwise we can find $R_i \in \mathfrak{R}$ and $\alpha_i \geq 0$, $i = 1, \ldots, n$, of the form $\alpha_i = n_i/m$, where $n_i, m \in \mathbb{N}$, such that $\sum_{i=1}^n \alpha_i \beta(R_i) < 1$. Set $M := \{(i,j)\colon\ 1 \leq i \leq n, 1 \leq j \leq n_i\}$ and $R_{ij} = R_i$ if $(i,j) \in M$. Then

$$\sum_{(i,j)\in M} I_{R_{ij}} = \sum_{i=1}^n n_i I_{R_i} = m \sum_{i=1}^n \alpha_i I_{R_i} \geq m,$$

but

$$\sum_{(i,j)\in M} \beta(R_{ij}) = \sum_{i=1}^n n_i \beta(R_i) = m \sum_{i=1}^n \alpha_i \beta(R_i) < m,$$

which contradicts (b). By the Hahn–Banach theorem, there is a linear functional λ on L such that $\lambda(1) = p(1) \geq 1$ and $\lambda \leq p$. Let $\nu(E) := \lambda(I_E)$, $E \subset X$. Then $\nu(E) \leq \beta(R)$ if $E \subset R \in \mathfrak{R}$. Let $\alpha(E) := \nu^+(E) := \sup_{A\subset E} \nu(E)$. Then α is nonnegative and additive (see Proposition 3.10.16 in Ch. 3) and $\alpha(R) \leq \beta(R)$ if $R \in \mathfrak{R}$. Finally, $1 \leq \nu(X) \leq \alpha(X) \leq \beta(X) = 1$.

1.12.149. Let $(X, \mathcal{A}, \mu)$ be a probability space and let $\mathcal{S}$ be a family of subsets in X such that $\mu_*\bigl(\bigcup_{n=1}^\infty S_n\bigr) = 0$ for every countable collection $\{S_n\} \subset \mathcal{S}$. Prove that there exists a probability measure $\widetilde{\mu}$ defined on some σ-algebra $\widetilde{\mathcal{A}}$ such that $\mathcal{A}, \mathcal{S} \subset \widetilde{\mathcal{A}}$, $\widetilde{\mu}$ extends μ and vanishes on $\mathcal{S}$, and for each $A \in \widetilde{\mathcal{A}}$ there exists $A' \in \mathcal{A}$ with $\widetilde{\mu}(A \bigtriangleup A') = 0$.

HINT: let $\mathcal{Z}$ be the class of all subsets in X that can be covered by an at most countable subfamily in $\mathcal{S}$. It is clear that $\mu_*(Z) = 0$ if $Z \in \mathcal{Z}$. Let

$$\widetilde{\mathcal{A}} := \{A \bigtriangleup Z, A \in \mathcal{A}, Z \in \mathcal{Z}\}.$$

It is easily seen that $\widetilde{\mathcal{A}}$ is a σ-algebra and contains $\mathcal{A}$ and $\mathcal{S}$. Set $\widetilde{\mu}(A \bigtriangleup Z) := \mu(A)$ for $A \in \mathcal{A}$ and $Z \in \mathcal{Z}$. The definition is unambiguous because if $A \bigtriangleup Z = A' \bigtriangleup Z'$, $A, A' \in \mathcal{A}$, $Z, Z' \in \mathcal{Z}$, then $A \bigtriangleup A' = Z \bigtriangleup Z'$, whence $\mu(A \bigtriangleup A') = \mu_*(Z \bigtriangleup Z') = 0$, since $Z \bigtriangleup Z' \in \mathcal{Z}$. Note that $\widetilde{\mu}(Z) = 0$ for $Z \in \mathcal{Z}$, since one can take $A = \varnothing$. The countable additivity of $\widetilde{\mu}$ is easily verified.

1.12.150.° Let μ be a bounded nonnegative measure on a σ-algebra $\mathcal{A}$ in a space X. Denote by $\mathcal{E}$ the class of all sets $E \subset X$ such that

$$\mu^*(E) = \mu^*(E \backslash A) + \mu^*(E \cap A) \quad \text{for all } A \in \mathcal{A}.$$

Is it true that the function μ^* is additive on $\mathcal{E}$?

HINT: no. Let us consider the following example due to O.V. Pugachev. Let $X = \{1, -1, i, -i\}$. We define a measure μ on a σ-algebra $\mathcal{A}$ consisting of eight sets as follows:

$$\mu(\varnothing) = 0, \quad \mu(X) = 3,$$

$$\mu(1) = \mu(-1) = \mu(\{i, -i\}) = 1, \quad \mu(\{1, -1\}) = \mu(\{1, i, -i\}) = \mu(\{-1, i, -i\}) = 2.$$

Clearly, the domain of definition of μ is indeed a σ-algebra. It is easily seen that μ is additive, hence countably additive. For every $E \subset X$, we have

$$\mu^*(E) = \mu^*(E \backslash A) + \mu^*(E \cap A)$$

for all $A \in \mathcal{A}$, but μ^* is not additive on the algebra of all subsets in X.

1.12.151. (Radó, Reichelderfer [**777**, p. 260]) Let Φ be a finite nonnegative set function defined on the family $\mathcal{U}$ of all open sets in $(0, 1)$ such that:

(i) $\Phi\bigl(\bigcup_{n=1}^\infty U_n\bigr) = \sum_{n=1}^\infty \Phi(U_n)$ for every countable family of pairwise disjoint sets $U_n \in \mathcal{U}$,

(ii) $\Phi(U_1) \le \Phi(U_2)$ whenever $U_1, U_2 \in \mathcal{U}$ and $U_1 \subset U_2$,

(iii) $\Phi(U) = \lim\limits_{\varepsilon \to 0} \Phi(U_\varepsilon)$ for every $U \in \mathcal{U}$, where U_ε is the set of all points in U with distance more than ε from the boundary of U.

Is it true that Φ has a countably additive extension to the Borel σ-algebra of $(0, 1)$?

HINT: no; let $\Phi(U) = 1$ if $[1/4, 1/2] \subset U$ and $\Phi(U) = 0$ otherwise.

1.12.152. Let μ be a nonnegative σ-finite measure on a measurable space $(X, \mathcal{A})$ and let M_0 be the class of all sets of finite μ-measure. Let

$\sigma_\mu(A, B) = \mu(A \bigtriangleup B)/\mu(A \cup B)$ if $\mu(A \cup B) > 0$, $\sigma_\mu(A, B) = 0$ if $\mu(A \cup B) = 0$.

(i) (Marczewski, Steinhaus [**653**]) (a) Show that σ_μ is a metric on the space of equivalence classes in M_0, where $A \sim B$ whenever $\mu(A \bigtriangleup B) = 0$.

(b) Show that if $A_n, A \in M_0$ and $\sigma_\mu(A_n, A) \to 0$, then $\mu(A_n \bigtriangleup A) \to 0$.

(c) Show that if $\mu(A_n \bigtriangleup A) \to 0$ and $\mu(A) > 0$, then $\sigma_\mu(A_n, A) \to 0$.

(d) Observe that $\sigma_\mu(\varnothing, B) = 1$ if $\mu(B) > 0$ and deduce that in the case of Lebesgue measure on $[0, 1]$, the identity mapping $(M_0, d) \to (M_0, \sigma_0)$, where d is the Fréchet–Nikodym metric, is discontinuous at the point corresponding to $\varnothing$.

(ii) (Gładysz, Marczewski, Ryll-Nardzewski [**359**]) For all $A_1, \ldots, A_n \in M_0$ let

$$\sigma_\mu(A_1, \ldots, A_n) = \frac{\mu\bigl((A_1 \cup \cdots \cup A_n) \backslash (A_1 \cap \cdots \cap A_n)\bigr)}{\mu(A_1 \cup \cdots \cup A_n)}$$

if $\mu(A_1\cup\cdots\cup A_n)>0$ and $\sigma_\mu(A_1,\ldots,A_n)=0$ if $\mu(A_1\cup\cdots\cup A_n)=0$. Prove the inequality

$$\sigma_\mu(A_1,\ldots,A_n)\le\frac{1}{n-1}\sum_{i<j}\sigma_\mu(A_i,A_j).$$

Deduce that if $\sigma_\mu(A_i,A_j)<2/n$ for all $1\le i<j\le n$, then $\mu(A_1\cap\cdots\cap A_n)>0$.

1.12.153.° Let $A_1,\ldots,A_n$ be measurable sets in a probability space $(\Omega,\mathcal{A},P)$. Prove that

$$0\le\sum_{i=1}^n P(A_i)-P\Bigl(\bigcup_{i=1}^n A_i\Bigr)\le\sum_{1\le i<j\le n}P(A_i\cap A_j).$$

HINT: by using induction on n and the easily verified fact that A_n is the union of the disjoint sets $B_1:=\bigl(\bigcup_{i=1}^n A_i\bigr)\backslash\bigl(\bigcup_{i=1}^{n-1}A_i\bigr)$ and $B_2:=\bigcup_{i=1}^{n-1}(A_i\cap A_n)$ we obtain

$$\begin{aligned}\sum_{i=1}^n P(A_i)-P\Bigl(\bigcup_{i=1}^n A_i\Bigr)&=\sum_{i=1}^{n-1}P(A_i)-P\Bigl(\bigcup_{i=1}^{n-1}A_i\Bigr)+P(A_n)-P(B_1)\\&\le\sum_{1\le i<j\le n-1}P(A_i\cap A_j)+P(B_2).\end{aligned}$$

It remains to observe that $P(B_2)\le\sum_{i=1}^{n-1}P(A_i\cap A_n)$. More general inequalities of this type are considered in Galambos, Simonelli [**336**].

1.12.154. (Darji, Evans [**203**]) Let A be a measurable set in the unit cube I of $\mathbb{R}^n$, let $F\subset I\backslash A$ be a finite set, and let $\varepsilon>0$. Show that there exists a finite set $S\subset A$ with the following property: for every partition $\mathcal{P}$ of the cube I into finitely many parallelepipeds of the form $[a_i,b_i]\times\cdots\times[a_n,b_n]$ with pairwise disjoint interiors, letting $B:=\bigcup\{P\in\mathcal{P}\colon\ P\cap F\ne\varnothing, P\cap S=\varnothing\}$ we have $\lambda_n(A\cap B)<\varepsilon$.

1.12.155. (Kahane [**479**]) Let E be the set of all points in $[0,1]$ of the form $x=3\sum_{n=1}^\infty\varepsilon_n4^{-n}$, $\varepsilon_n\in\{0,1\}$. Show that $E+\dfrac12E=[0,3/2]$, but for almost all real λ, the set $E+\lambda E$ has measure zero.

1.12.156. Multivariate distribution functions admit the following characterization. For any vectors $x=(x_1,\ldots,x_n)$, $y=(y_1,\ldots,y_n)$ let

$$[x,y):=[x_1,y_1)\times\cdots\times[x_n,y_n).$$

Given a function F on $\mathbb{R}^n$ let $F[x,y):=\sum\limits_u s(u)F(u)$, where the summation is taken over all corner points u of the set $[x,y)$ and $s(u)$ equals $+1$ or -1 depending on whether the number of indices k with $u_k=y_k$ is even or odd. Prove that the function F on $\mathbb{R}^n$ is the distribution function of some probability measure precisely when the following conditions are fulfilled: 1) $F[x,y)\ge0$ whenever $x<y$ coordinatewise, 2) $F(x^j)\to F(x)$ whenever the vectors x^j increase to x, 3) $F(x)\to0$ as $\max_k x_k\to-\infty$ and $F(x)\to1$ as $\min_k x_k\to+\infty$.

HINT: see Vestrup [**976**, §2.3, 2.4].

1.12.157. Let $\mathcal{A}$ be a σ-algebra of subsets of $\mathbb{N}$. Show that $\mathcal{A}$ is generated by some finite or countable partition of $\mathbb{N}$ into disjoint sets, so that every element of $\mathcal{A}$ is an at most countable union of elements of this partition.

HINT: let $n\sim m$ if n and m cannot be separated by a set from $\mathcal{A}$. It is readily verified that we obtain an equivalence relation. Every equivalence class K is an element of $\mathcal{A}$. Indeed, let us fix some $k\in K$. For every $n\in\mathbb{N}\backslash K$, there is a set

$A_n \in \mathcal{A}$ such that $k \in A_n$, $n \notin A_n$. Then $K = \bigcap_{n=1}^{\infty} A_n$. Indeed, $\bigcap_{n=1}^{\infty} A_n \subset K$ by construction. On the other hand, if $l \in K$ and $l \notin \bigcap_{n=1}^{\infty} A_n$, then k is separated from l by the set $\bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$. Hence we obtain an at most countable family of disjoint sets $M_n \in \mathcal{A}$ with union $\mathbb{N}$ such that every element of $\mathcal{A}$ is a finite or countable union of some of these sets.

1.12.158. (i) Let $\mathcal{A}$ be a σ-algebra of subsets of $\mathbb{N}$ and let μ be a probability measure on $\mathcal{A}$. Show that μ extends to a probability measure on the class of all subsets of $\mathbb{N}$.

(ii) Let $\mathcal{A}$ be the σ-algebra generated by singletons of a set X and let $\mathcal{A}_0$ be its sub-σ-algebra. Show that any measure μ on $\mathcal{A}_0$ extends to a measure on $\mathcal{A}$.

Hint: (i) apply Exercise 1.12.157 (cf. Hanisch, Hirsch, Renyi [**406**]; the result also follows as a special case of extension of measures on Souslin spaces, which is considered in Volume 2). (ii) Observe that μ is concentrated at countably many atoms, and any atom is either countable or has a countable complement.

1.12.159. Let μ be a countably additive measure with values in $[0, +\infty]$ on a ring $\mathfrak{X}$ of subsets of a space X.

(i) Suppose that μ is σ-finite, i.e., $X = \bigcup_{n=1}^{\infty} X_n$, where one has $X_n \in \mathfrak{X}$ and $\mu(X_n) < \infty$. Show that μ has a unique countably additive extension to the σ-ring $\Sigma(\mathfrak{X})$ generated by $\mathfrak{X}$.

(ii) Suppose that the measure $\mathfrak{m} := \mu^*$ is σ-finite on $\mathfrak{X}_{\mathfrak{m}}$. Show that it is a unique extension of μ to $\sigma(\mathfrak{X})$.

Hint: (i) according to Corollary 1.11.9, μ^* is a countably additive extension of μ to $\Sigma(\mathfrak{X})$ (even to $\sigma(\mathfrak{X})$). Let ν be another countably additive extension of μ to $\Sigma(\mathfrak{X})$. We show that $\mu^* = \nu$ on $\Sigma(\mathfrak{X})$. Let $E \in \Sigma(\mathfrak{X})$. We may assume that $X_n \subset X_{n+1}$. It suffices to show that $\mu^*(E \cap X_n) = \nu(E \cap X_n)$ for every n. This follows by the uniqueness result in the case of algebras because it is readily seen that the set $E \cap X_n$ belongs to the σ-algebra generated by the intersections of sets in $\mathfrak{X}$ with X_n. (ii) See Vulikh [**1000**, Ch. IV, §5].

1.12.160. Two sets A and B on the real line are called metrically separated if, for every $\varepsilon > 0$, there exist open sets A_ε and B_ε such that $A \subset A_\varepsilon$ and $B \subset B_\varepsilon$ with $\lambda(A_\varepsilon \cap B_\varepsilon) < \varepsilon$, where λ is Lebesgue measure.

(i) Show that if sets A and B are metrically separated, then there exist Borel sets A_0 and B_0 such that $A \subset A_0$ and $B \subset B_0$ with $\lambda(A_0 \cap B_0) = 0$.

(ii) Let A be a Lebesgue measurable set on the real line and let $A = A_1 \cup A_2$, where the sets A_1 and A_2 are metrically separated. Show that A_1 and A_2 are Lebesgue measurable.

Hint: (i) let A_n and B_n be open sets such that $A \subset A_n$, $B \subset B_n$, and $\lambda(A_n \cap B_n) < n^{-1}$. Take the sets $A_0 := \bigcap_{n=1}^{\infty} A_n$ and $B_0 := \bigcap_{n=1}^{\infty} B_n$. (ii) According to (i) there exist Borel sets B_1 and B_2 with $A_1 \subset B_1$, $A_2 \subset B_2$, and $\lambda(B_1 \cap B_2) = 0$. Let $E := A \cap (B_1 \backslash A_1)$. It is readily verified that $E \subset B_1 \cap B_2$. Hence $\lambda(E) = 0$, which shows that A_1 is Lebesgue measurable.

CHAPTER 2

The Lebesgue integral

Any measurement is subject to unavoidable errors, and the general total consists of a given number of the smallest capricious particulars, but in the large, the average of all these minor caprices vanishes, and then God's fundamental law appears, the law which alone turns slaves into the true masters of everything undertaken and forthcoming.

D.I. Mendeleev. Intimate thoughts.

2.1. Measurable functions

In this section, we study measurable functions. In spite of its name, the concept of measurability of functions is defined in terms of σ-algebras and *is not connected with measures*. Connections with measures arise when the given σ-algebra is the σ-algebra of all sets measurable with respect to a fixed measure. This important special case is considered at the end of the section.

2.1.1. Definition. *Let $(X, \mathcal{A})$ be a measurable space, i.e., a space with a σ-algebra. A function $f\colon X \to \mathrm{I\!R}^1$ is called measurable with respect to $\mathcal{A}$ (or $\mathcal{A}$-measurable) if $\{x\colon f(x) < c\} \in \mathcal{A}$ for every $c \in \mathrm{I\!R}^1$.*

The simplest example of an $\mathcal{A}$-measurable function is the indicator I_A of a set $A \in \mathcal{A}$ defined as follows: $I_A(x) = 1$ if $x \in A$ and $I_A(x) = 0$ if $x \notin A$. The indicator of a set A is also called the characteristic function of A or the indicator function of A. The set $\{x\colon I_A(x) < c\}$ is empty if $c \leq 0$, equals the complement of A if $c \in (0, 1]$ and coincides with X if $c > 1$. It is clear that the inclusion $A \in \mathcal{A}$ is also necessary for the $\mathcal{A}$-measurability of I_A.

2.1.2. Theorem. *A function f is measurable with respect to a σ-algebra $\mathcal{A}$ if and only if $f^{-1}(B) \in \mathcal{A}$ for all sets $B \in \mathcal{B}(\mathrm{I\!R}^1)$.*

Proof. Let f be $\mathcal{A}$-measurable. Denote by $\mathcal{E}$ the collection of all sets $B \in \mathcal{B}(\mathrm{I\!R}^1)$ such that $f^{-1}(B) \in \mathcal{A}$. We show that $\mathcal{E}$ is a σ-algebra. Indeed, if $B_n \in \mathcal{E}$, then (see Lemma 1.2.8)

$$f^{-1}\Bigl(\bigcup_{n=1}^{\infty} B_n\Bigr) = \bigcup_{n=1}^{\infty} f^{-1}(B_n) \in \mathcal{A}, \quad f^{-1}(\mathrm{I\!R}^1 \backslash B_n) = X \backslash f^{-1}(B_n) \in \mathcal{A}.$$

Since $\mathcal{E}$ contains the rays $(-\infty, c)$, we obtain that $\mathcal{B}(\mathrm{I\!R}^1) \subset \mathcal{E}$, i.e., $\mathcal{B}(\mathrm{I\!R}^1) = \mathcal{E}$. The converse assertion is obvious, since the rays are Borel sets. $\square$

Let us write f in the form $f = f^+ - f^-$, where

$$f^+(x) := \max\bigl(f(x), 0\bigr), \quad f^-(x) := \max\bigl(-f(x), 0\bigr).$$

It is clear that the $\mathcal{A}$-measurability of f is equivalent to the $\mathcal{A}$-measurability of both functions f^+ and f^-. For example, if $c > 0$, we have the equality $\{x\colon f(x) < c\} = \{x\colon f^+(x) < c\}$.

It is clear from the definition that the restriction $f|_E$ of any $\mathcal{A}$-measurable function f to an arbitrary set $E \subset X$ is measurable with respect to the σ-algebra $\mathcal{A}_E = \{A \cap E\colon A \in \mathcal{A}\}$.

The following more general definition is frequently useful.

2.1.3. Definition. *Let $(X_1, \mathcal{A}_1)$ and $(X_2, \mathcal{A}_2)$ be two spaces with σ-algebras. A mapping $f\colon X_1 \to X_2$ is called measurable with respect to the pair $(\mathcal{A}_1, \mathcal{A}_2)$ (or $(\mathcal{A}_1, \mathcal{A}_2)$-measurable) if $f^{-1}(B) \in \mathcal{A}_1$ for all $B \in \mathcal{A}_2$.*

In the case where $(X_2, \mathcal{A}_2) = \bigl(\mathbb{R}^1, \mathcal{B}(\mathbb{R}^1)\bigr)$, we arrive at the definition of a measurable function. In another special case where X_1 and X_2 are metric (or topological) spaces with their Borel σ-algebras $\mathcal{A}_1 = \mathcal{B}(X_1)$ and $\mathcal{A}_2 = \mathcal{B}(X_2)$, i.e., the σ-algebras generated by open sets, we obtain the notion of a *Borel* (or *Borel measurable*) mapping. In particular, a real function on a set $E \subset \mathbb{R}^n$ is called Borel if it is $\mathcal{B}(E)$-measurable.

2.1.4. Example. Every continuous function f on a set $E \subset \mathbb{R}^n$ is Borel measurable, since the set $\{x\colon f(x) < c\}$ is open for any c, hence Borel.

An important class of $\mathcal{A}$-measurable functions is the collection of all *simple functions*, i.e., $\mathcal{A}$-measurable functions f with finitely many values. Thus, any simple function f has the form $f = \sum_{i=1}^n c_i I_{A_i}$, where $c_i \in \mathbb{R}^1$, $A_i \in \mathcal{A}$, in other words, f is a finite linear combination of indicators of sets in $\mathcal{A}$. Obviously, the converse is also true.

The following theorem describes the basic properties of measurable functions.

2.1.5. Theorem. *Suppose that functions f, g, f_n, where $n \in \mathbb{N}$, are measurable with respect to a σ-algebra $\mathcal{A}$. Then*:

(i) *the function $\varphi \circ f$ is measurable with respect to $\mathcal{A}$ for any Borel function $\varphi\colon \mathbb{R}^1 \to \mathbb{R}^1$; in particular, this is true if φ is continuous*;

(ii) *the function $\alpha f + \beta g$ is measurable with respect to $\mathcal{A}$ for all α, $\beta \in \mathbb{R}^1$*;

(iii) *the function fg is measurable with respect to $\mathcal{A}$*;

(iv) *if $g(x) \neq 0$, then the function f/g is measurable with respect to $\mathcal{A}$*;

(v) *if there exists a finite limit $f_0(x) = \lim\limits_{n\to\infty} f_n(x)$ for all x, then the function f_0 is measurable with respect to $\mathcal{A}$*;

(vi) *if the functions $\sup_n f_n(x)$ and $\inf_n f_n(x)$ are finite for all x, then they are measurable with respect to $\mathcal{A}$.*

PROOF. Claim (i) follows by the equality

$$(\varphi \circ f)^{-1}(B) = f^{-1}\bigl(\varphi^{-1}(B)\bigr).$$

By (i), for the proof of (ii) it suffices to consider the case $\alpha = \beta = 1$ and observe that

$$\begin{aligned}\bigl\{x\colon\ f(x)+g(x)<c\bigr\} &= \bigl\{x\colon\ f(x)<c-g(x)\bigr\}\\ &= \bigcup_{r_n}\Bigl(\bigl\{x\colon\ f(x)<r_n\bigr\}\cap\bigl\{x\colon\ r_n<c-g(x)\bigr\}\Bigr),\end{aligned}$$

where the union is taken over all rational numbers r_n. The right-hand side of this relation belongs to $\mathcal{A}$, since the functions f and g are measurable with respect to $\mathcal{A}$. Claim (iii) follows by the equality $2fg = \bigl[(f+g)^2 - f^2 - g^2\bigr]$ and the already-proven assertions; in particular, the square of a measurable function is measurable by (i). Noting that the function φ given by the equality $\varphi(x) = 1/x$ if $x \neq 0$ and $\varphi(0) = 0$, is Borel (a simple verification of this is left as an exercise for the reader), we obtain (iv). The least obvious in all the assertions in the theorem is (v), which, however, is clear from the following easily verified relations:

$$\bigl\{x\colon\ f_0(x)<c\bigr\} = \bigcup_{k=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{m=n+1}^{\infty}\Bigl\{x\colon\ f_m(x)<c-\frac{1}{k}\Bigr\}.$$

For the proof of (vi) we observe that

$$\sup_n f_n(x) = \lim_{n\to\infty}\max\bigl(f_1(x),\ldots,f_n(x)\bigr).$$

By (v), it suffices to show the measurability of $\max(f_1,\ldots,f_n)$. By induction, this reduces to $n=2$. It remains to observe that

$$\Bigl\{x\colon\ \max\bigl(f_1(x),f_2(x)\bigr)<c\Bigr\} = \bigl\{x\colon\ f_1(x)<c\bigr\}\cap\bigl\{x\colon\ f_2(x)<c\bigr\}.$$

The assertion for inf is verified similarly (certainly, one can also use the equality $\inf_n f_n = -\sup_n(-f_n)$). The theorem is proven. $\square$

2.1.6. Remark. For functions f with values on the extended real line $\overline{\mathbb{R}} = [-\infty,+\infty]$ we define the $\mathcal{A}$-measurability by requiring the inclusions

$$f^{-1}(-\infty),\ f^{-1}(+\infty)\in\mathcal{A}$$

and the $\mathcal{A}$-measurability of f on $f^{-1}(\mathbb{R})$. This is equivalent to the measurability in the sense of Definition 2.1.3 if $\overline{\mathbb{R}}$ is equipped with the σ-algebra $\mathcal{B}(\overline{\mathbb{R}})$ consisting of Borel sets of the usual line with possible addition of the points $-\infty$, $+\infty$. Then, for functions with values in $\overline{\mathbb{R}}$, assertions (i), (v), (vi) of the above theorem remain valid, and for the validity of assertions (ii), (iii), (iv) one has to consider functions f and g with values either in $[-\infty,+\infty)$ or in $(-\infty,+\infty]$. The algebraic operations for such values are defined in the following natural way: $+\infty + c = +\infty$ if $c\in(-\infty,+\infty]$, $+\infty\cdot 0 = 0$, $+\infty\cdot c = +\infty$ if $c>0$, $+\infty\cdot c = -\infty$ if $c<0$.

2.1.7. Lemma. *Let functions f_n be measurable with respect to a σ-algebra $\mathcal{A}$ in a space X. Then, the set L of all points $x\in X$ such that $\lim\limits_{n\to\infty} f_n(x)$*

exists and is finite belongs to $\mathcal{A}$. The same is true for the sets L^- and L^+ of all those points where the limit equals $-\infty$ and $+\infty$.

PROOF. The set L coincides with the set of all points x where the sequence $\{f_n(x)\}$ is fundamental, hence

$$L = \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n,j\geq m} \Bigl\{x\colon\ |f_n(x) - f_j(x)| \leq \frac{1}{k}\Bigr\} \in \mathcal{A}.$$

This equality is verified as follows: x belongs to the right-hand side precisely when, for each k, there exists m such that $|f_n(x) - f_j(x)| \leq 1/k$ whenever $n, j \geq m$. This is exactly the fundamentality of $\{f_n(x)\}$. For L^- and L^+ proofs are similar. $\square$

2.1.8. Lemma. *Suppose that $\mathcal{A}$ is a σ-algebra of subsets of a space X. Then, for any bounded $\mathcal{A}$-measurable function f, there exists a sequence of simple functions f_n convergent to f uniformly on X.*

PROOF. Let $c = \sup_{x\in X} |f(x)| + 1$. For every $n \in \mathbb{N}$ we partition $[-c, c)$ into n disjoint intervals $I_j = \bigl[-c + 2c(j-1)n^{-1}, -c + 2cjn^{-1}\bigr)$ of length $2cn^{-1}$. Let $A_j = f^{-1}(I_j)$. It is clear that $A_j \in \mathcal{A}$ and $\bigcup_{j=1}^n A_j = X$. Let c_j be the middle point of I_j. Let us define the function f_n by the equality $f_n(x) = c_j$ for $x \in A_j$. Then f_n is a simple function and

$$\sup_{x\in X} |f(x) - f_n(x)| \leq cn^{-1},$$

since the function f maps A_j to I_j, and f_n takes A_j to the middle point of I_j, which is at the distance at most cn^{-1} from any point in I_j. $\square$

2.1.9. Corollary. *Suppose that $\mathcal{A}$ is a σ-algebra of subsets of a space X. Then, for every $\mathcal{A}$-measurable function f, there exists a sequence of simple functions f_n convergent to f at every point.*

PROOF. Let us consider the functions g_n defined by $g_n(x) = f(x)$ if $f(x) \in [-n, n]$ and $g_n(x) = 0$ otherwise. We can find simple functions f_n such that $|f_n(x) - g_n(x)| \leq n^{-1}$. It is clear that $\lim_{n\to\infty} f_n(x) = f(x)$ for all $x \in X$. $\square$

Once again we draw the reader's attention to the fact that so far no measures have been involved in our discussion of measurable functions. Suppose now that we have a nonnegative countably additive measure μ on a σ-algebra $\mathcal{A}$ of subsets of a space X.

2.1.10. Definition. *Let $(X, \mathcal{A}, \mu)$ be a measure space. A real function f on X is called μ-measurable if it is measurable with respect to the σ-algebra $\mathcal{A}_\mu$ of all μ-measurable sets. In addition, we agree to call μ-measurable also any function f that is defined and $\mathcal{A}_\mu$-measurable on $X\backslash Z$, where Z is a set of μ-measure zero (that is, f may be undefined or infinite on Z). The set of all μ-measurable functions is denoted by $\mathcal{L}^0(\mu)$.*

Thus, the μ-measurability of a function f means that, for any $c \in \mathbb{R}^1$, the set $\{x\colon\ f(x) < c\}$ belongs to the Lebesgue completion of $\mathcal{A}$ with respect to μ (and that f is defined on a full measure set, i.e., outside a measure zero set). It is clear that the class of μ-measurable functions (even everywhere defined) may be wider than the class of $\mathcal{A}$-measurable functions, since no completeness of $\mathcal{A}$ with respect to μ is assumed. If a μ-measurable function f is not defined on a set Z of measure zero, then, defining it on Z in an arbitrary way (say, letting $f|_Z = 0$), we make it μ-measurable in the sense of the first part of the given definition. It will be clear from the sequel that a somewhat broader concept of measurability of functions allowed by the second part of our definition is technically convenient. Normally, in concrete situations, when one speaks of a measurable function, it is clear whether it is supposed to be defined everywhere or only almost everywhere and this circumstance is never specified. However, one can easily imagine situations where such a specification is necessary. For example, suppose one has to consider a family of functions f_α on $[0,1]$, where $\alpha \in [0,1]$, such that the function f_α is not defined at the point α. Then from the formal point of view, these functions have no common points of domain of definition at all.

For functions with values in $[-\infty, +\infty]$ (possibly infinite on a set of positive measure), the μ-measurability is understood as follows: $f^{-1}(-\infty)$ and $f^{-1}(+\infty)$ belong to $\mathcal{A}_\mu$, and on the set $\{|f| < \infty\}$ the function f is μ-measurable. Such functions *are not included* in $\mathcal{L}^0(\mu)$ (we do not consider such functions at all); in order to avoid confusion, it is preferable to call them mappings rather than functions.

2.1.11. Proposition. *Let μ be a nonnegative measure on a σ-algebra $\mathcal{A}$. Then, for every μ-measurable function f, one can find a set $Y \in \mathcal{A}$ and a function g measurable with respect to $\mathcal{A}$ such that $f(x) = g(x)$ for all $x \in Y$ and $\mu(X \backslash Y) = 0$.*

PROOF. We may assume that f is defined and finite everywhere. By Corollary 2.1.9, there exists a sequence of $\mathcal{A}_\mu$-measurable simple functions f_n pointwise convergent to f. The function f_n assumes finitely many distinct values on sets $A_1, \ldots, A_k \in \mathcal{A}_\mu$. Every set A_i contains a set B_i from $\mathcal{A}$ such that $\mu(A_i \backslash B_i) = 0$. Let us consider the function g_n that coincides with f_n on the union of the sets B_i and equals 0 outside this union. Clearly, g_n is an $\mathcal{A}$-measurable simple function, and there is a measure zero set $Z_n \in \mathcal{A}$ such that $f_n(x) = g_n(x)$ if $x \notin Z_n$. Let $Y = X \backslash \bigcup_{n=1}^{\infty} Z_n$. Then $Y \in \mathcal{A}$ and $\mu(X \backslash Y) = 0$. Let $g(x) = f(x)$ if $x \in Y$ and $g(x) = 0$ otherwise. For every $x \in Y$ one has $f(x) = \lim\limits_{n\to\infty} f_n(x) = \lim\limits_{n\to\infty} g_n(x)$. Hence f is $\mathcal{A}$-measurable on Y. Therefore, g is $\mathcal{A}$-measurable on X. $\square$

It follows by this proposition that for a *bounded* μ-measurable function f, there exist two $\mathcal{A}$-measurable functions f_1 and f_2 such that

$$f_1(x) \le f(x) \le f_2(x) \text{ for all } x \text{ and } \mu\bigl(x\colon\ f_1(x) \ne f_2(x)\bigr) = 0.$$

Indeed, let $f_1 = f_2 = g$ on Y. Outside Y we set $f_1(x) = \inf f$, $f_2 = \sup f$.

2.2. Convergence in measure and almost everywhere

Let $(X, \mathcal{A}, \mu)$ be a measure space with a nonnegative measure μ. We say that some property for points in X is fulfilled almost everywhere (or μ-almost everywhere) on X if the set Z of all points in X that do not have this property belongs to $\mathcal{A}_\mu$ and has measure zero with respect to μ. We use the following abbreviations for "almost everywhere": a.e., μ-a.e. If a function g equals a function f a.e., then it is called a modification or version of f. It is clear from the definition of $\mathcal{A}_\mu$ that there exists a set $Z_0 \in \mathcal{A}$ such that $Z \subset Z_0$ and $\mu(Z_0) = 0$, i.e., the corresponding property is fulfilled outside some measure zero set from $\mathcal{A}$. This circumstance should be kept in mind when dealing with incomplete measures. The complement of a measure zero set is called a set of full measure.

For example, one can speak of a.e. convergence of a sequence of functions f_n, fundamentality a.e. of $\{f_n\}$, nonnegativity a.e. of a function etc. It is clear that a.e. convergence of $\{f_n\}$ follows from convergence of $\{f_n(x)\}$ for each x (called pointwise convergence), and the latter follows from uniform convergence of $\{f_n\}$. A deeper connection between almost everywhere convergence and uniform convergence is described by the following theorem due to the eminent Russian mathematician D. Egoroff.

2.2.1. Theorem. *Let $(X, \mathcal{A}, \mu)$ be a space with a finite nonnegative measure μ and let μ-measurable functions f_n be such that μ-almost everywhere there is a finite limit $f(x) := \lim\limits_{n\to\infty} f_n(x)$. Then, for every $\varepsilon > 0$, there exists a set $X_\varepsilon \in \mathcal{A}$ such that $\mu(X \backslash X_\varepsilon) < \varepsilon$ and the functions f_n converge to f uniformly on X_ε.*

PROOF. The assertion reduces to the case where the sequence $\{f_n(x)\}$ converges at every point because we can redefine the functions f_n on the measure zero set on which at least one of them is not defined or there is no convergence. Then

$$X_n^m := \bigcap_{i\geq n} \Big\{x\colon\ |f_i(x) - f(x)| < \frac{1}{m}\Big\} \in \mathcal{A}_\mu.$$

We observe that $X_n^m \subset X_{n+1}^m$ for all $m,\ n \in \mathbb{N}$, and that $\bigcup_{n=1}^\infty X_n^m = X$, since for fixed m, for any x, there exists a number n such that $|f_i(x) - f(x)| < 1/m$ whenever $i \geq n$. Let $\varepsilon > 0$. By the countable additivity of μ, for each m, there exists a number $k(m)$ with $\mu(X \backslash X_{k(m)}^m) < \varepsilon 2^{-m}$. Set $X_\varepsilon = \bigcap_{m=1}^\infty X_{k(m)}^m$. Then $X_\varepsilon \in \mathcal{A}_\mu$ and

$$\mu(X \backslash X_\varepsilon) = \mu\Big(\bigcup_{m=1}^\infty (X \backslash X_{k(m)}^m)\Big) \leq \sum_{m=1}^\infty \mu(X \backslash X_{k(m)}^m) \leq \varepsilon \sum_{m=1}^\infty 2^{-m} = \varepsilon.$$

Finally, for fixed m, we have $|f_i(x) - f(x)| < 1/m$ for all $x \in X_\varepsilon$ and all $i \geq k(m)$, which means uniform convergence of the sequence $\{f_n\}$ to f on the set X_ε. It remains to take in X_ε a subset (denoted by the same symbol) from $\mathcal{A}$ of the same measure. $\square$

Simple examples show that Egoroff's theorem does not extend to the case $\varepsilon = 0$. For example, the sequence of functions $f_n\colon x \mapsto x^n$ on $(0,1)$ converges at every point to zero, but it cannot converge uniformly on a set $E \subset (0,1)$ with Lebesgue measure 1, since every neighborhood of the point 1 contains points from E and then $\sup_{x\in E} f_n(x) = 1$ for every n. The property of convergence established by Egoroff is called *almost uniform convergence.*

Let us consider yet another important type of convergence of measurable functions.

2.2.2. Definition. *Suppose we are given a measure space $(X, \mathcal{A}, \mu)$ with a finite measure μ and a sequence of μ-measurable functions f_n.*

(i) *The sequence $\{f_n\}$ is called fundamental* (*or Cauchy*) *in measure if, for every $c > 0$, one has*

$$\lim_{N\to\infty} \sup_{n,k\geq N} \mu\bigl(x\colon\ |f_n(x) - f_k(x)| \geq c\bigr) = 0.$$

(ii) *The sequence $\{f_n\}$ is said to converge in measure to a μ-measurable function f if, for every $c > 0$, one has*

$$\lim_{n\to\infty} \mu\bigl(x\colon\ |f(x) - f_n(x)| \geq c\bigr) = 0.$$

Note that if a sequence of functions f_n converges in measure, then it is fundamental in measure. Indeed, the set $\bigl\{x\colon |f_n(x) - f_k(x)| \geq c\bigr\}$ is contained in the set

$$\bigl\{x\colon\ |f(x) - f_n(x)| \geq c/2\bigr\} \cup \bigl\{x\colon\ |f(x) - f_k(x)| \geq c/2\bigr\}.$$

Note also that if a sequence $\{f_n\}$ converges in measure to functions f and g, then $f = g$ almost everywhere. Hence up to a redefinition of functions on measure zero sets, the limit in the sense of convergence in measure is unique. Indeed, for every $c > 0$ we have

$$\begin{aligned}\mu\bigl(x\colon\ |f(x) - g(x)| \geq c\bigr) &\leq \mu\bigl(x\colon\ |f(x) - f_n(x)| \geq c/2\bigr)\\ &\quad + \mu\bigl(x\colon\ |f_n(x) - g(x)| \geq c/2\bigr) \to 0,\end{aligned}$$

whence $\mu\bigl(x\colon\ |f(x) - g(x)| > 0\bigr) = 0$, since the set of points where the function $|f - g|$ is positive is the union of the sets of points where it is at least n^{-1}.

Let us clarify connections between convergence in measure and convergence almost everywhere.

2.2.3. Theorem. *Let $(X, \mathcal{A}, \mu)$ be a measure space with a finite measure. If a sequence of μ-measurable functions f_n converges almost everywhere to a function f, then it converges to f in measure.*

Proof. Let $c > 0$ and

$$A_n = \bigl\{x\colon\ |f(x) - f_i(x)| < c,\ \forall\, i \geq n\bigr\}.$$

The sets A_n are μ-measurable and $A_n \subset A_{n+1}$. It is clear that the set $\bigcup_{n=1}^\infty A_n$ contains all points at which $\{f_n\}$ converges to f. Hence $\mu(X) = \mu\bigl(\bigcup_{n=1}^\infty A_n\bigr)$. By the countable additivity of μ we have $\mu(A_n) \to \mu(X)$, i.e., $\mu(X\backslash A_n) \to 0$. It remains to observe that $\bigl(x\colon\ |f(x) - f_n(x)| \geq c\bigr) \subset X\backslash A_n$. $\square$

The converse assertion is false: there exists a sequence of measurable functions on $[0,1]$ that converges to zero in Lebesgue measure but does not converge at any point at all.

2.2.4. Example. For every $n \in \mathbb{N}$ we partition $[0,1]$ into 2^n intervals $I_{n,k} = [(k-1)2^{-n}, k2^{-n})$, $k = 1, \ldots, 2^n$, of length 2^{-n}. Let $f_{n,k}(x) = 1$ if $x \in I_{n,k}$ and $f_{n,k}(x) = 0$ if $x \notin I_{n,k}$. We write the functions $f_{n,k}$ in a single sequence

$$f_n = (f_{1,1}, f_{1,2}, f_{2,1}, f_{2,2}, \ldots)$$

such that the function $f_{n+1,k}$ follows the functions $f_{n,j}$. The sequence $\{f_n\}$ converges to zero in Lebesgue measure, since the length of the interval on which the function f_n is nonzero tends to zero as n increases. However, there is no convergence at any point x, since the sequence $\{f_n(x)\}$ contains infinitely many elements 0 and 1.

The next theorem due to F. Riesz gives a partial converse to Theorem 2.2.3.

2.2.5. Theorem. *Let $(X, \mathcal{A}, \mu)$ be a space with a finite measure.*

(i) *If a sequence of μ-measurable functions f_n converges to f in measure μ, then there exists its subsequence $\{f_{n_k}\}$ that converges to f almost everywhere.*

(ii) *If a sequence of μ-measurable functions f_n is fundamental in measure μ, then it converges in measure μ to some measurable function f.*

PROOF. Let $\{f_n\}$ be fundamental in measure. Let us show that there exists a sequence of natural numbers $n_k \to \infty$ such that

$$\mu\Bigl(x\colon\ |f_n(x) - f_j(x)| \geq 2^{-k}\Bigr) \leq 2^{-k}, \quad \forall\, n,\, j \geq n_k.$$

Indeed, we find a number n_1 with

$$\mu\Bigl(x\colon\ |f_n(x) - f_j(x)| \geq 2^{-1}\Bigr) \leq 2^{-1}, \quad \forall\, n,\, j \geq n_1.$$

Next we find a number $n_2 > n_1$ with

$$\mu\Bigl(x\colon\ |f_n(x) - f_j(x)| \geq 2^{-2}\Bigr) \leq 2^{-2}, \quad \forall\, n,\, j \geq n_2.$$

Continuing this process, we obtain a desired sequence $\{n_k\}$. Let us show that the sequence $\{f_{n_k}\}$ converges a.e. To this end, it suffices to show that it is a.e. fundamental. Set

$$E_j = \Bigl\{x\colon\ |f_{n_{j+1}}(x) - f_{n_j}(x)| \geq 2^{-j}\Bigr\}.$$

Since

$$\mu\Bigl(\bigcup_{j=k}^{\infty} E_j\Bigr) \leq \sum_{j=k}^{\infty} 2^{-j} = 2^{-k+1} \to 0 \quad \text{as } k \to \infty,$$

the set $Z = \bigcap_{k=1}^{\infty} \bigcup_{j=k}^{\infty} E_j$ has μ-measure zero. If $x \in X \backslash Z$, then the sequence $\{f_{n_k}(x)\}$ is fundamental. Indeed, there exists a number k such that x does not belong to $\bigcup_{j=k}^{\infty} E_j$, i.e., $x \notin E_j$ for all $j \geq k$. By definition this means that $|f_{n_{j+1}}(x) - f_{n_j}(x)| < 2^{-j}$ for all $j \geq k$. Hence, for every fixed $m \geq k$, for all $i > j > m$ one has the estimate

$$|f_{n_i}(x) - f_{n_j}(x)| \leq |f_{n_i}(x) - f_{n_{i-1}}(x)| + |f_{n_{i-1}}(x) - f_{n_{i-2}}(x)| + \ldots$$
$$+ |f_{n_{j+1}}(x) - f_{n_j}(x)| \leq \sum_{l=j}^{\infty} 2^{-l} \leq 2^{-j+1} \leq 2^{-m},$$

which means that $\{f_{n_k}(x)\}$ is fundamental. Thus, the selected subsequence $\{f_{n_k}\}$ converges almost everywhere to some function f. Then one has convergence in measure as well, which yields assertion (ii). Finally, assertion (i) follows from the above-noted fact that any sequence convergent in measure is fundamental in measure. In addition, the limit of the selected subsequence coincides almost everywhere with the limit of $\{f_n\}$ in measure due to the uniqueness of the limit in measure up to a redefinition of a function on a set of measure zero. □

2.2.6. Corollary. *Let μ be a finite measure and let two sequences of μ-measurable functions f_n and g_n converge in measure μ to functions f and g, respectively. Suppose that Ψ is a continuous function on some set $Y \subset \mathbb{R}^2$ such that $\big(f(x), g(x)\big) \in Y$ and $\big(f_n(x), g_n(x)\big) \in Y$ for all x and all n. Then, the functions $\Psi(f_n, g_n)$ converge in measure μ to the function $\Psi(f, g)$. In particular, $f_n g_n \to fg$ and $\alpha f_n + \beta g_n \to \alpha f + \beta g$ in measure μ for all real numbers α and β.*

PROOF. According to Exercise 2.12.29 the functions $\Psi(f, g)$ and $\Psi(f_n, g_n)$ are measurable. If our claim is false, then there exist $c > 0$ and a subsequence j_n such that

$$\mu\Big(x\colon\ \big|\Psi\big(f(x), g(x)\big) - \Psi\big(f_{j_n}(x), g_{j_n}(x)\big)\big| > c\Big) > c \tag{2.2.1}$$

for all n. By the Riesz theorem, $\{j_n\}$ contains a subsequence $\{i_n\}$ such that $f_{i_n}(x) \to f(x)$ and $g_{i_n}(x) \to g(x)$ a.e. Due to the continuity of Ψ we obtain

$$\Psi\big(f_{i_n}(x), g_{i_n}(x)\big) \to \Psi\big(f(x), g(x)\big) \quad \text{a.e.},$$

whence $\Psi(f_{i_n}, g_{i_n}) \to \Psi(f, g)$ in measure, which contradicts (2.2.1). The remaining claims follow by the proven claim applied to the functions $\Psi(x, y) = xy$ and $\Psi(x, y) = \alpha x + \beta y$. □

2.2.7. Remark. We shall see later that convergence in measure can be described by a metric (Exercise 4.7.60). It can be seen directly from the definition that convergence in measure possesses the following property: if functions f_n converge in measure μ to a function f, and, for every fixed n, the functions $f_{n,k}$ converge in measure μ to the function f_n, then there exist numbers $k_n \geq n$ such that the sequence f_{n,k_n} converges in measure μ to f. The choice of k_n is made inductively. First we find a number k_1 with

$$\mu\Big(x\colon\ |f_{1,k_1}(x) - f_1(x)| \geq 2^{-1}\Big) \leq 2^{-1}.$$

If we have already found increasing numbers $k_1, \ldots, k_{n-1}$ such that $k_j \geq j$ and

$$\mu\Bigl(x\colon\ |f_{j,k_j}(x) - f_j(x)| \geq 2^{-j}\Bigr) \leq 2^{-j} \quad \text{for } j = 1, \ldots, n-1,$$

then we can find $k_n > \max(k_{n-1}, n)$ such that

$$\mu\Bigl(x\colon\ |f_{n,k_n}(x) - f_n(x)| \geq 2^{-n}\Bigr) \leq 2^{-n}.$$

For the proof of convergence of $\{f_{n,k_n}\}$ to f in measure μ it suffices to observe that, for every fixed $c > 0$, for all n with $2^{-n} < c/2$ one has the inclusion

$$\Bigl\{x\colon\ |f_{n,k_n}(x) - f(x)| \geq c\Bigr\}$$
$$\subset \Bigl\{x\colon\ |f_{n,k_n}(x) - f_n(x)| \geq 2^{-n}\Bigr\} \bigcup \Bigl\{x\colon\ |f_n(x) - f(x)| \geq c/2\Bigr\},$$

where the measure of the set on the right tends to zero. It is interesting to note that a.e. convergence cannot be described by a metric or by a topology (Exercise 2.12.70).

This remark enables one to construct approximations in measure by functions from given classes.

2.2.8. Lemma. *Let K be a compact set on the real line, U an open set containing K, and f a continuous function on K. Then, there exists a continuous function g on the real line such that $g = f$ on K, $g = 0$ outside U and*

$$\sup_{x\in\mathbb{R}^1} |g(x)| = \sup_{x\in K} |f(x)|.$$

PROOF. It suffices to consider the case where U is bounded. The set $U\backslash K$ is a finite or countable union of pairwise disjoint open intervals. Set $g = 0$ outside U, $g = f$ on K, and on every interval (a, b) constituting U we define g with the aid of linear interpolation of the values at the endpoints of this interval: $g\bigl(ta + (1-t)b\bigr) = tg(a) + (1-t)g(b)$. The obtained function has the required properties. $\square$

2.2.9. Proposition. *For every measurable function f on an interval I with Lebesgue measure, there exists a sequence of continuous functions f_n convergent to f in measure.*

PROOF. The functions g_n defined by the equality

$$g_n(x) = f(x) \text{ if } |f(x)| \leq n,\ g_n(x) = n \operatorname{sign} f(x) \text{ if } |f(x)| > n,$$

are measurable and converge to f pointwise, hence in measure. Each of the functions g_n is the uniform limit of simple functions. According to Remark 2.2.7, it suffices to prove our claim for all functions of the form $f = \sum_{i=1}^n c_i I_{A_i}$, where A_i are disjoint measurable sets in I. Moreover, we may assume that the sets A_i are compact, since every A_i is approximated from inside by compact sets in the sense of measure. Then, for any $m \in \mathbb{N}$, there exist disjoint open sets U_i that are finite unions of intervals such that

$A_i \subset U_i$ and $\lambda\bigl(\bigcup_{i=1}^n (U_i\backslash A_i)\bigr) < m^{-1}$. Let $c = \max_{i\le n} |c_i|$. According to Lemma 2.2.8, there exists a continuous function $f_m\colon I \to [-c, c]$ such that $f_m = f$ on $\bigcup_{i=1}^n A_i$ and $f_m = 0$ outside $\bigcup_{i=1}^n U_i$. Thus, the measure of the set $\{f_m \neq f\}$ does not exceed m^{-1}, whence we obtain convergence of $\{f_m\}$ to f in measure. □

The nature of measurable functions on an interval with Lebesgue measure is clarified in the following classical Lusin theorem.

2.2.10. Theorem. *A function f on an interval I with Lebesgue measure is measurable precisely when for each $\varepsilon > 0$, there exist a continuous function f_ε and a compact set K_ε such that $\lambda(I\backslash K_\varepsilon) < \varepsilon$ and $f = f_\varepsilon$ on K_ε.*

PROOF. The sufficiency of the above condition is seen from the fact that if it is satisfied, then the set $\{x\colon f(x) < c\}$ coincides up to a measure zero set with the Borel set $\bigcup_{n=1}^\infty \bigl\{x \in K_{1/n}\colon f_{1/n}(x) < c\bigr\}$. Let us verify its necessity. By using the previous proposition, we choose a sequence of continuous functions f_n convergent in measure to f. Applying the Riesz theorem and passing to a subsequence, we may assume that $f_n \to f$ a.e. By Egoroff's theorem, there exists a measurable set F_ε such that $\lambda(I\backslash F_\varepsilon) < \varepsilon/2$ and $f_n \to f$ uniformly on F_ε. Next we find a compact set $K_\varepsilon \subset F_\varepsilon$ with $\lambda(F_\varepsilon\backslash K_\varepsilon) < \varepsilon/2$ and observe that $f|_{K_\varepsilon}$ is continuous being the uniform limit of continuous functions. It remains to note that, by Lemma 2.2.8, the function $f|_{K_\varepsilon}$ can be extended to a continuous function f_ε on I. □

2.2.11. Remark. It is worth noting that Proposition 2.2.9 and Theorem 2.2.10 with the same proofs remain valid for arbitrary bounded Borel measures on an interval. In Chapter 7 (see §7.1, §7.14(ix)) we return to Lusin's theorem in the case of measures on topological spaces.

2.3. The integral for simple functions

Let $(X, \mathcal{A}, \mu)$ be a space with a finite nonnegative measure. For any simple function f on X that assumes finitely many values c_i on disjoint sets A_i, $i = 1, \ldots, n$, the Lebesgue integral of f with respect to μ is defined by the equality

$$\int_X f(x)\,\mu(dx) := \sum_{i=1}^n c_i\mu(A_i).$$

That the integral is well-defined is obvious from the additivity of measure, which enables one to deal with the case where all the values c_i are distinct.

If $A \in \mathcal{A}$, then the integral of f over the set A is defined as the integral of the simple function $I_A f$, i.e.,

$$\int_A f(x)\,\mu(dx) = \sum_{i=1}^n c_i\mu(A_i \cap A).$$

The following brief notation for the integral of a function f over a set A with respect to a measure μ is used:

$$\int_A f\,d\mu.$$

2.3.1. Definition. *A sequence $\{f_n\}$ of simple functions is called fundamental in the mean or mean fundamental (or fundamental in $L^1(\mu)$, which is explained below) if, for every $\varepsilon > 0$, there exists a number n such that*

$$\int_X |f_i(x) - f_j(x)|\,\mu(dx) < \varepsilon \quad \textit{for all } i,\, j \geq n.$$

Note that a sequence is fundamental in the mean exactly when it is fundamental with respect to the metric

$$\varrho(f,g) := \|f-g\|_{L^1(\mu)} := \int_X |f(x) - g(x)|\,\mu(dx)$$

on the space of equivalence classes of simple functions, where two functions are equivalent if they coincide almost everywhere. We discuss this in greater detail in Chapter 4.

2.3.2. Lemma. *The Lebesgue integral on simple functions enjoys the following properties*:

(i) *if $f \geq 0$, then*

$$\int_X f(x)\,\mu(dx) \geq 0;$$

(ii) *the inequality*

$$\Bigl|\int_X f(x)\,\mu(dx)\Bigr| \leq \int_X |f(x)|\,\mu(dx) \leq \sup_{x\in X} |f(x)|\,\mu(X)$$

holds;

(iii) *if α, $\beta \in \mathbb{R}^1$, then*

$$\int_X \bigl[\alpha f(x) + \beta g(x)\bigr]\,\mu(dx) = \alpha \int_X f(x)\,\mu(dx) + \beta \int_X g(x)\,\mu(dx).$$

In particular, if A and B are disjoint sets in $\mathcal{A}$, then

$$\int_{A\cup B} f(x)\,\mu(dx) = \int_A f(x)\,\mu(dx) + \int_B f(x)\,\mu(dx). \tag{2.3.1}$$

PROOF. Assertions (i) and (ii) are obvious from the definition. In addition, the definition yields the equality

$$\int_X \alpha f(x)\,\mu(dx) = \alpha \int_X f(x)\,\mu(dx).$$

Hence it suffices to verify claim (iii) for $\alpha = \beta = 1$. Let f assume distinct values c_i on sets A_i, $i = 1, \ldots, n$, and let g assume distinct values b_j on sets B_j,

$j = 1, \dots, m$. Then the sets $A_i \cap B_j \in \mathcal{A}$ are disjoint and $f + g = a_i + b_j$ on the set $A_i \cap B_j$. Hence

$$\begin{aligned}\int_X \bigl[f(x) + g(x)\bigr]\,\mu(dx) &= \sum_{i\le n,\, j\le m} (a_i + b_j)\mu(A_i \cap B_j)\\ &= \sum_{i\le n} a_i\mu(A_i) + \sum_{j\le m} b_j\mu(B_j)\\ &= \int_X f(x)\,\mu(dx) + \int_X g(x)\,\mu(dx),\end{aligned}$$

since $\sum_{j\le m}\mu(A_i \cap B_j) = \mu(A_i)$ and $\sum_{i\le n}\mu(A_i \cap B_j) = \mu(B_j)$. The last claim in (iii) follows by the equality $I_{A\cup B} = I_A + I_B$. □

2.3.3. Corollary. *If f and g are simple functions and $f \le g$ almost everywhere, then*

$$\int_X f(x)\,\mu(dx) \le \int_X g(x)\,\mu(dx).$$

Proof. Let $A = \bigl\{x\colon\ f(x) \le g(x)\bigr\}$. Then $A \in \mathcal{A}$ and $\mu(X\backslash A) = 0$. Let $c = \sup_{x\in X}[|f(x)| + |g(x)|]$. We have $g - f + cI_{X\backslash A} \ge 0$. By definition, the integral of the function $cI_{X\backslash A}$ equals zero. Hence the inequality we prove follows by assertions (i) and (iii) in Lemma 2.3.2. □

The second assertion in the next lemma expresses a very important property of the uniform absolute continuity of any sequence fundamental in the mean.

2.3.4. Lemma. *Suppose that a sequence of simple functions f_n is fundamental in the mean. Then*:

(i) *the sequence*

$$\int_X f_n(x)\,\mu(dx)$$

converges to a finite limit;

(ii) *for every $\varepsilon > 0$, there exists $\delta > 0$ such that, for each set D with $\mu(D) < \delta$ and all n, one has the estimate*

$$\int_D |f_n(x)|\,\mu(dx) \le \varepsilon.$$

Proof. (i) It suffices to observe that according to what has been proven earlier, one has

$$\Bigl|\int_X f_n(x)\,\mu(dx) - \int_X f_k(x)\,\mu(dx)\Bigr| \le \int_X |f_n(x) - f_k(x)|\,\mu(dx).$$

(ii) We find N such that

$$\int_X |f_n(x) - f_j(x)|\,\mu(dx) \le \frac{\varepsilon}{2}, \quad \forall\, n,\, j \ge N.$$

Let $C = \max\limits_{x\in X,\, i\le N} |f_i(x)| + 1$ and $\delta = \varepsilon(2C)^{-1}$. If $\mu(D) < \delta$ and $n \geq N$, then

$$\begin{aligned}\int_D |f_n(x)|\,\mu(dx) &= \int_D |f_n(x) - f_N(x) + f_N(x)|\,\mu(dx)\\ &\le \int_D |f_n(x) - f_N(x)|\,\mu(dx) + \int_D |f_N(x)|\,\mu(dx)\\ &\le \frac{\varepsilon}{2} + C\delta \le \varepsilon.\end{aligned}$$

If $n < N$, then we have

$$\int_D |f_n(x)|\,\mu(dx) \le C\mu(D) \le \varepsilon.$$

The lemma is proven. □

2.4. The general definition of the Lebesgue integral

In this section, a triple $(X, \mathcal{A}, \mu)$ denotes a space X with a σ-algebra $\mathcal{A}$ and a finite nonnegative measure μ on $\mathcal{A}$.

In the definition of the integral it is convenient to employ the extended concept of a measurable function given in Definition 2.1.10 and admit functions that are defined almost everywhere (i.e., may be undefined or infinite on sets of measure zero). The idea of the following definition is to obtain the integral by means of completion, which is much in the spirit of defining measurable sets by means of approximations by elementary ones.

2.4.1. Definition. *Let a function f be defined and finite μ-a.e.* (*i.e., f may be undefined or infinite on a set of measure zero*). *The function f is called Lebesgue integrable with respect to the measure μ* (*or μ-integrable*) *if there exists a sequence of simple functions f_n such that $f_n(x) \to f(x)$ almost everywhere and the sequence $\{f_n\}$ is fundamental in the mean. The finite value*

$$\lim_{n\to\infty} \int_X f_n(x)\,\mu(dx),$$

which exists by Lemma 2.3.4, *is called the Lebesgue integral of the function f and is denoted by*

$$\int_X f(x)\,\mu(dx) \quad \textit{or by} \quad \int_X f\,d\mu.$$

Let $\mathcal{L}^1(\mu)$ be the collection of all μ-integrable functions.

Obviously, any μ-integrable function is μ-measurable. Let us show that the value of the integral is independent of our choice of a sequence $\{f_n\}$ involved in its definition. It is to be noted that in the next section we give an equivalent definition of the integral that does not require the justification of its correctness. Exercises 2.12.56, 2.12.57, and 2.12.58 contain other frequently used definitions of the Lebesgue integral equivalent to the one given above (see also Exercises 2.12.59, 2.12.60, and 2.12.61). The most constructive is the definition from Exercise 2.12.57: the integral is the limit of the so-called Lebesgue sums

$$\sum_{k=-\infty}^{+\infty} \varepsilon k\, \mu\bigl(x\colon\ \varepsilon k \le f(x) < \varepsilon(k+1)\bigr)$$

as $\varepsilon \to 0$, where the absolute convergence of the series for some $\varepsilon > 0$ is required (i.e., convergence separately for positive and negative k); then it follows automatically that the sum is finite for every $\varepsilon > 0$ and the above limit exists. In particular, it suffices to consider ε of the form $\varepsilon = 1/n$, $n \in \mathbb{N}$. The corresponding Lebesgue sums become

$$\sum_{k=-\infty}^{+\infty} \frac{k}{n}\, \mu\Bigl(x\colon\ \frac{k}{n} \le f(x) < \frac{k+1}{n}\Bigr).$$

These facts will be obvious from the subsequent discussion.

2.4.2. Lemma. *Let $\{f_n\}$ and $\{g_n\}$ be two sequences of simple functions that are mean fundamental and converge almost everywhere to one and the same function f. Then the integrals of f_n and g_n converge to the same value.*

PROOF. Let $\varepsilon > 0$. By Lemma 2.3.4, there exists $\delta > 0$ such that for any set D with $\mu(D) < \delta$, one has the estimate

$$\Bigl|\int_D f_n(x)\, \mu(dx)\Bigr| + \Bigl|\int_D g_n(x)\, \mu(dx)\Bigr| \le \varepsilon, \quad \forall\, n \in \mathbb{N}. \tag{2.4.1}$$

By Egoroff's theorem, there exists a set $X_\delta \in \mathcal{A}$ such that $\mu(X\backslash X_\delta) < \delta$ and on the set X_δ the sequences $\{f_n\}$ and $\{g_n\}$ converge to f uniformly. Hence there exists a number N such that

$$\sup_{x \in X_\delta} |f_n(x) - g_n(x)| \le \varepsilon, \quad \forall\, n \ge N. \tag{2.4.2}$$

Then, by (2.4.1) and (2.4.2), we obtain for $n \ge N$

$$\begin{aligned}
&\Bigl|\int_X f_n(x)\, \mu(dx) - \int_X g_n(x)\, \mu(dx)\Bigr| \\
&\le \Bigl|\int_{X_\delta} \bigl[f_n(x) - g_n(x)\bigr]\, \mu(dx) + \int_{X\backslash X_\delta} f_n(x)\, \mu(dx) - \int_{X\backslash X_\delta} g_n(x)\, \mu(dx)\Bigr| \\
&\quad \le \varepsilon\mu(X) + \Bigl|\int_{X\backslash X_\delta} f_n(x)\, \mu(dx)\Bigr| + \Bigl|\int_{X\backslash X_\delta} g_n(x)\, \mu(dx)\Bigr| \le \varepsilon\bigl(\mu(X) + 1\bigr),
\end{aligned}$$

which proves our claim. □

The reader is warned that in order that a function f be integrable it is not sufficient to represent it as the pointwise limit of simple functions f_n with the convergent sequence of integrals. For example, as we shall see below, the function $f(x) = x^{-1}$ on the interval $[-1, 1]$ with Lebesgue measure is not Lebesgue integrable, although it can be easily represented as the limit of odd simple functions f_n whose integrals over $[-1, 1]$ vanish. The fundamentality of $\{f_n\}$ in the mean is a key condition. Almost everywhere convergence is needed to identify the limit of $\{f_n\}$ with a point function, not just with an abstract element of the completion of the metric space corresponding to simple functions. Let us recall that the completion of a metric space M can be defined by means of a metric space of fundamental sequences from the elements of M. The above definition employs this idea, but does not entirely reduce to it.

2.4.3. Lemma. *Suppose that f is a μ-integrable function and $A \in \mathcal{A}_\mu$. Then, the function fI_A is μ-integrable as well.*

PROOF. We may assume that $A \in \mathcal{A}$ because there is a set $B \in \mathcal{A}$ such that $B \subset A$ and $\mu(A\backslash B) = 0$, i.e., $I_A = I_B$ a.e. Let $\{f_n\}$ be a sequence of simple functions that is fundamental in the mean and converges to f almost everywhere. Then the functions $g_n = f_n I_A$ are simple as well, converge to fI_A almost everywhere, and the sequence $\{g_n\}$ is fundamental in the mean, which follows by the estimate $|g_n - g_m| \le |f_n - f_m|$ and Corollary 2.3.3. $\square$

This lemma implies the following definition.

2.4.4. Definition. *The Lebesgue integral of a function f over a set $A \in \mathcal{A}_\mu$ is defined as the integral of the function fI_A over the whole space if the latter is integrable.*

It is clear that any integrable function is integrable over every set in $\mathcal{A}_\mu$. The integral of the function f over the set A is denoted by the symbols

$$\int_A f(x)\,\mu(dx) \quad \text{and} \quad \int_A f\,d\mu.$$

In the case where we integrate over the whole space X, the indication of the domain of integration may be omitted and then we use the notation

$$\int f\,d\mu.$$

In the case of Lebesgue measure on $\mathbb{R}^n$, we also write

$$\int_A f(x)\,dx.$$

We observe that by definition any two functions f and g that are equal almost everywhere, either both are integrable or both are not integrable, and in the case of integrability their integrals are equal. In particular, an arbitrary function (possibly infinite) on every set of measure zero is integrable and has zero integral. It is often useful not to distinguish functions that are equal almost everywhere. Such functions are called equivalent. To this end, in place of the space $\mathcal{L}^1(\mu)$ one considers the space $L^1(\mu)$ (an alternate notation: $L^1(X,\mu)$) whose elements are equivalence classes in $\mathcal{L}^1(\mu)$ consisting of almost everywhere equal functions. We return to this in §2.11 and Chapter 4.

No completeness of the measure μ is assumed above, but it is clear that one can also take $\mathcal{A}_\mu$ for $\mathcal{A}$. Moreover, according to our definition, we obtain the same class of integrable functions if we replace $\mathcal{A}$ by $\mathcal{A}_\mu$ in the case where $\mathcal{A}_\mu$ is larger than $\mathcal{A}$. Indeed, although in the latter case we increase the class of simple functions, this does not affect the class of integrable functions, since every $\mathcal{A}_\mu$-simple function coincides almost everywhere with some $\mathcal{A}$-measurable function.

2.5. Basic properties of the integral

As in the previous section, $(X, \mathcal{A}, \mu)$ stands for a measure space with a finite nonnegative measure μ.

2.5.1. Theorem. *The Lebesgue integral defined in the previous section possesses the following properties*:

(i) *if f is an integrable function and $f \geq 0$ a.e., then*

$$\int_X f(x)\,\mu(dx) \geq 0;$$

(ii) *if a function f is integrable, then the function $|f|$ is integrable as well and*

$$\Bigl|\int_X f(x)\,\mu(dx)\Bigr| \leq \int_X |f(x)|\,\mu(dx);$$

(iii) *every $\mathcal{A}_\mu$-measurable bounded function f is integrable and*

$$\Bigl|\int_X f(x)\,\mu(dx)\Bigr| \leq \sup_{x\in X} |f(x)|\,\mu(X);$$

(iv) *if two functions f and g are integrable, then, for all α, $\beta \in \mathbb{R}^1$, the function $\alpha f + \beta g$ is integrable and*

$$\int_X \bigl[\alpha f(x) + \beta g(x)\bigr]\,\mu(dx) = \alpha \int_X f(x)\,\mu(dx) + \beta \int_X g(x)\,\mu(dx).$$

In particular, if A and B are disjoint sets in $\mathcal{A}_\mu$, then, for every integrable function f, one has

$$\int_{A\cup B} f(x)\,\mu(dx) = \int_A f(x)\,\mu(dx) + \int_B f(x)\,\mu(dx);$$

(v) *if integrable functions f and g are such that $f(x) \leq g(x)$ a.e., then*

$$\int_X f\,d\mu \leq \int_X g\,d\mu.$$

Proof. (i) There is a sequence of simple functions f_n that is fundamental in the mean and converges to f almost everywhere. Then the functions $|f_n|$ are simple, $|f_n| \to |f|$ a.e., which due to the nonnegativity of f a.e. implies that $|f_n| \to f$ a.e. In addition, one has

$$\int_X \bigl||f_n(x)| - |f_m(x)|\bigr|\,\mu(dx) \leq \int_X |f_n(x) - f_m(x)|\,\mu(dx),$$

since $\Bigl||t| - |s|\Bigr| \leq |t - s|$ for all t, $s \in \mathbb{R}^1$. It remains to use that the integrals of the functions $|f_n|$ are nonnegative.

Claim (ii) is clear from the reasoning in (i).

(iii) If a measurable function f takes values in $[-c, c]$, then by Lemma 2.1.8 one can find a sequence of simple functions f_n with values in $[-c, c]$ uniformly convergent to f. It remains to apply assertion (ii) of Lemma 2.3.2.

(iv) If two mean fundamental sequences of simple functions f_n and g_n are such that $f_n \to f$ and $g_n \to g$ a.e., then $h_n = \alpha f_n + \beta g_n \to \alpha f + \beta g$ a.e. and

$$\int_X |h_n - h_m|\, d\mu \le |\alpha| \int_X |f_n - f_m|\, d\mu + |\beta| \int_X |g_n - g_m|\, d\mu,$$

which means that $\{h_n\}$ is fundamental in the mean. It remains to use the linearity of the integral on simple functions.

Claim (v) follows by the linearity of the integral and claim (i), since one has $g(x) - f(x) \ge 0$ almost everywhere. $\square$

Let us now give an equivalent definition of the Lebesgue integral used in many books. An advantage of this definition is its somewhat greater constructibility, and its drawback is the necessity to consider first nonnegative functions. If this characterization of integrability is taken as a definition, then one can also prove the linearity of the integral. Let us set

$$f^+ = \max(f, 0), \quad f^- = \max(-f, 0).$$

2.5.2. Theorem. *A nonnegative μ-measurable function f is integrable precisely when the following value is finite*:

$$I(f) := \sup\left\{\int_X \varphi\, d\mu \colon\ \varphi \le f \ \textit{a.e.},\ \varphi \ \textit{is simple}\right\}.$$

In this case $I(f)$ coincides with the integral of f. The integrability of an arbitrary measurable function f is equivalent to the finiteness of $I(f^+)$ and $I(f^-)$, and then $I(f^+) - I(f^-)$ coincides with the integral of f.

Proof. We may deal with a version of f that is $\mathcal{A}$-measurable and nonnegative. Let $f_n(x) = k4^{-n}$ if $f(x) \in \bigl[k4^{-n}, (k+1)4^{-n}\bigr)$, $k = 0, \ldots, 8^n - 1$, $f_n(x) = 2^n$ if $f(x) \ge 2^n$. Then the functions f_n are simple, $f_n \le f$, $f_{n+1} \ge f_n$ and $f_n \to f$. The integrals of f_n are increasing. If f is integrable, then these integrals are majorized by the integral of f and hence converge to some number $I \le I(f)$. It is clear that

$$I(f) \le \int f\, d\mu.$$

Taking into account the estimate $f_n \le f_m$ for $n \le m$, we conclude that $\{f_n\}$ is fundamental in the mean. Hence I coincides with the integral of f, which yields the equality $I = I(f)$. Conversely, if $I(f)$ is finite, then again we obtain that $\{f_n\}$ is fundamental in the mean, which gives the integrability of f. The case of a sign-alternating function reduces to the considered one due to the linearity of the integral. $\square$

A simple corollary of property (v) in Theorem 2.5.1 is the following frequently used Chebyshev inequality.

2.5.3. Theorem. *For any μ-integrable function f and any $R>0$ one has*

$$\mu\bigl(x\colon\ |f(x)|\ge R\bigr)\le \frac{1}{R}\int_X |f(x)|\,\mu(dx). \tag{2.5.1}$$

PROOF. Set $A_R=\bigl\{x\colon\ |f(x)|\ge R\bigr\}$. It is clear that $R\cdot I_{A_R}(x)\le |f(x)|$ for all x. Hence the integral of the function $R\cdot I_{A_R}$ is majorized by the integral of $|f|$, which yields (2.5.1). □

2.5.4. Corollary. *If*

$$\int_X |f|\,d\mu=0,$$

then $f=0$ a.e.

2.5.5. Proposition. *A nonnegative μ-measurable function f is integrable with respect to μ precisely when*

$$\sup_{n\ge 1}\int_X \min(f,n)\,d\mu<\infty.$$

PROOF. We may deal with an $\mathcal{A}$-measurable version of f. The functions $f_n:=\min(f,n)$ are bounded and $\mathcal{A}$-measurable. Suppose that their integrals are uniformly bounded. There exist simple functions g_n such that we have $|f_n(x)-g_n(x)|\le n^{-1}$ for all x. Since $f_n(x)\to f(x)$, one has $g_n(x)\to f(x)$. Whenever $n\ge k$, we have $|f_n-f_k|=f_n-f_k$, hence

$$\begin{aligned}\int |g_n-g_k|\,d\mu &= \int |g_n-f_n+f_n-f_k+f_k-g_k|\,d\mu\\ &\le \int |g_n-f_n|\,d\mu+\int |f_n-f_k|\,d\mu+\int |f_k-g_k|\,d\mu\\ &\le \frac{1}{n}\mu(X)+\int f_n\,d\mu-\int f_k\,d\mu+\frac{1}{k}\mu(X).\end{aligned}$$

It remains to observe that the sequence

$$\int f_n\,d\mu$$

is fundamental, since it is increasing and bounded. Thus, the sequence $\{g_n\}$ is fundamental in the mean. The converse is obvious. □

2.5.6. Corollary. *Suppose that f is a μ-measurable function such that $|f(x)|\le g(x)$ a.e., where g is a μ-integrable function. Then the function f is μ-integrable as well.*

PROOF. The functions f^+ and f^- are μ-measurable and

$$\min(f^+,n)\le \min(g,n)\quad\text{and}\quad \min(f^-,n)\le\min(g,n).$$

Hence the functions f^+ and f^- are integrable and so is their difference, i.e., the function f. □

This corollary yields the integrability of a measurable function f such that the function $|f|$ is integrable. Certainly, the hypothesis of measurability of f cannot be omitted, since there exist nonmeasurable functions f with $|f(x)| \equiv 1$.

The next theorem establishes a very important property of the *absolute continuity of the Lebesgue integral*.

2.5.7. Theorem. *Let f be a μ-integrable function. Then, for every $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$\int_D |f(x)|\,\mu(dx) < \varepsilon \quad \textit{if } \mu(D) < \delta.$$

PROOF. There is a mean fundamental sequence of simple functions f_n convergent to $|f|$ almost everywhere. By Lemma 2.3.4, there exists $\delta > 0$ such that

$$\Bigl|\int_D f_n(x)\,\mu(dx)\Bigr| < \frac{\varepsilon}{2}, \quad \forall\, n \in \mathbb{N},$$

for any set D with $\mu(D) < \delta$. It remains to observe that

$$\int_D |f(x)|\,\mu(dx) = \lim_{n\to\infty} \int_D f_n(x)\,\mu(dx),$$

since $f_n I_D \to |f| I_D$ a.e. and the sequence $\{f_n I_D\}$ is fundamental in the mean. □

Let us consider functions with countably many values.

2.5.8. Example. Suppose that a function f assumes countably many values c_n on disjoint μ-measurable sets A_n. Then, the integrability of f with respect to μ is equivalent to convergence of the series $\sum_{n=1}^{\infty} |c_n|\mu(A_n)$. In addition,

$$\int_X f\,d\mu = \sum_{n=1}^{\infty} c_n \mu(A_n).$$

PROOF. It is clear that the function f is measurable. Let us consider the simple functions $f_n = \sum_{i=1}^{n} c_i I_{A_i}$. Then $|f_n| \le |f|$. If the function f is integrable, then the integrals of the functions $|f_n|$ are majorized by the integral of $|f|$, whence $\sup_n \sum_{i=1}^{n} |c_i|\mu(A_i) < \infty$, which means convergence of the above series. If this series converges, then the sequence $\{f_n\}$, as is readily seen, is fundamental in the mean, which implies the integrability of f because $f_n(x) \to f(x)$ for each x. We also obtain the announced expression for the integral of f. □

2.6. Integration with respect to infinite measures

In this section, we discuss integration over spaces with infinite measures. Let μ be a countably additive measure defined on a σ-algebra $\mathcal{A}$ in a space X and taking values in $[0, +\infty]$.

2.6.1. Definition. *If μ is an infinite measure, then a function f is called simple if it is $\mathcal{A}$-measurable, assumes only finitely many values and satisfies the condition $\mu\bigl(x\colon f(x) \neq 0\bigr) < \infty$. The integrability and integral with respect to an infinite measure are defined in the same manner as in the case of a space with finite measure, i.e., with the aid of Definition* 2.4.1*, where we set $0 \cdot \mu\bigl(x\colon f(x) = 0\bigr) = 0$ for any simple function f.*

With this definition many basic properties of the integral remain valid (although there are exceptions, for example, bounded functions may not be integrable). The integral for infinite measures can also be defined in the spirit of Theorem 2.5.2.

The next result shows that the integral with respect to an arbitrary infinite measure reduces to the integral with respect to a σ-finite measure (obtained by restricting the initial measure), and the latter can be reduced, if we like, to the integral with respect to some finite measure. In particular, it follows that the integral with respect to an infinite measure is well-defined and possesses the principal properties of the integral established in the previous section.

2.6.2. Proposition. (i) *If a function f is integrable with respect to a countably additive measure μ with values in $[0, +\infty]$, then the measure μ is σ-finite on the set $\bigl\{x\colon f(x) \neq 0\bigr\}$.*

(ii) *Let μ be a σ-finite measure on a space X that is the union of an increasing sequence of μ-measurable subsets X_n of finite measure. Then, the function f is integrable with respect to μ precisely when the restrictions of f to the sets X_n are integrable and*

$$\sup_n \int_{X_n} |f|\, d\mu < \infty.$$

In this case, one has

$$\int_X f\, d\mu = \lim_{n\to\infty} \int_{X_n} f\, d\mu = \sum_{n=1}^{\infty} \int_{X_n \backslash X_{n-1}} f\, d\mu, \quad X_0 = \varnothing. \tag{2.6.1}$$

(iii) *For any σ-finite measure μ, there exists a strictly positive μ-integrable function ϱ with countably many values. The function f is integrable with respect to μ precisely when the function f/ϱ is integrable with respect to the bounded measure $\nu = \varrho \cdot \mu$ defined by the equality*

$$\nu(A) := \int_A \varrho(x)\, \mu(dx), \quad A \in \mathcal{A}.$$

In addition,

$$\int_X f\, d\mu = \int_X \frac{f}{\varrho}\, d\nu. \tag{2.6.2}$$

Proof. (i) Let us take a mean fundamental sequence of simple functions f_n convergent almost everywhere to f. The set $X_0 = \bigcup_{n=1}^{\infty}\bigl\{x\colon f_n(x) \neq 0\bigr\}$

is a countable union of sets of finite measure. Since $f = \lim_{n\to\infty} f_n$ a.e., one has $f = 0$ a.e. on the set $X\backslash X_0$.

(ii) Let a function f be integrable. As in the case of a finite measure, this yields the integrability of $|f|$. Then, as is readily seen, the restrictions of $|f|$ to X_n are integrable. Hence the integrals of $|f|$ over X_n (which are well-defined according to what has been proven for finite measures) are majorized by the integral of $|f|$ over X. Moreover, if $\{f_j\}$ and $\{g_j\}$ are mean fundamental sequences of simple functions almost everywhere convergent to f, then the restrictions of f_j and g_j to each set X_n converge in the mean to the restriction of f to X_n. Given $\varepsilon > 0$, one can find a number N such that

$$\int_X |f_j - f_k|\,d\mu + \int_X |g_j - g_k|\,d\mu \le \varepsilon, \quad \forall\, j,k \ge N.$$

Next we find n such that

$$\int_{X\backslash X_n} \bigl[|f_N| + |g_N|\bigr]\,d\mu \le \varepsilon.$$

Then, for $j \ge N$, we have

$$\begin{aligned}\int_X |f_j - g_j|\,d\mu &= \int_{X_n} |f_j - g_j|\,d\mu + \int_{X\backslash X_n} |f_j - g_j|\,d\mu \\ &\le \int_{X_n} |f_j - g_j|\,d\mu + \int_{X\backslash X_n} \bigl[|f_j - f_N| + |f_N - g_N| + |g_N - g_j|\bigr]\,d\mu \\ &\le \int_{X_n} |f_j - g_j|\,d\mu + 2\varepsilon.\end{aligned}$$

It follows that the integrals of f_j and g_j converge to a common limit, which means that the integral is well-defined for infinite measures, too.

Conversely, if the integrals of $|f|$ over the sets X_n are uniformly bounded, then, since the sets X_n are increasing, there exists a finite limit

$$\lim_{n\to\infty} \int_{X_n} |f|\,d\mu.$$

Let us choose numbers $C_{n,j} > 0$ such that

$$\sum_{n=1}^{\infty} \int_{|f|\ge C_{n,j}} |f|\,d\mu < 2^{-j}.$$

It is easy to find a sequence of simple functions f_j with the following properties: for $n = 1,\ldots,j$ on every set $X_{n,j} = \bigl\{x \in X_n\backslash X_{n-1} \colon\ |f(x)| \le C_{n,j}\bigr\}$ one has the inequality $|f_j - f| \le 2^{-j}2^{-n}\bigl(1 + \mu\bigl(X_n\bigr)\bigr)^{-1}$, and outside the union of these sets one has $f_j = 0$. It is clear that the sequence $\{f_j\}$ is fundamental in the mean and converges almost everywhere to f. This reasoning yields relation (2.6.1) as well.

(iii) We observe that if A_n are pairwise disjoint sets of finite μ-measure with union X, then the function ϱ equal to $2^{-n}\bigl(\mu(A_n)+1\bigr)^{-1}$ on A_n is integrable with respect to μ. Set

$$\nu(A) = \int_A \varrho(x)\,\mu(dx), \quad A \in \mathcal{A}.$$

By using that, for every fixed n, the function $A \mapsto \mu(A \cap A_n)$ is a countably additive measure, it is readily verified that ν is a bounded countably additive measure. Equality (2.6.2) holds for indicators of all sets in $\mathcal{A}$ that are contained in one of the sets A_n. Hence it remains valid for all μ-simple functions and consequently for all μ-integrable functions. Then it is clear that the integrability of f with respect to μ is equivalent to the integrability of f/ϱ with respect to ν. Indeed, if a sequence of simple functions f_j converges to f μ-a.e. and is fundamental in $L^1(\mu)$, then $\{f_j/\varrho\}$ converges to f/ϱ ν-a.e. and is fundamental in $L^1(\nu)$. Conversely, if $f/\varrho \in \mathcal{L}^1(\nu)$, then there is a sequence of simple functions g_j fundamental in $L^1(\nu)$ that is ν-a.e. convergent to f/ϱ. Let $X_n = \bigcup_{i=1}^n A_i$. Then $g_j\varrho I_{X_j}$ are simple functions convergent μ-a.e. to f, and the sequence $\{g_j\varrho I_{X_j}\}$ is fundamental in $L^1(\mu)$. □

2.6.3. Remark. Given a sequence of μ-integrable functions f_j, the set X_0 mentioned in the proof of (i) can be chosen in such way that $f_j = 0$ almost everywhere outside X_0 for each j.

For the reader's convenience we summarize the basic properties of the integral with respect to infinite measures that are immediate corollaries of the results in the previous section and the above proposition.

2.6.4. Proposition. *Let $(X, \mathcal{A})$ be a measurable space and let μ be a measure on $\mathcal{A}$ with values in $[0, +\infty]$. Then, all the assertions of the previous section, excepting assertion* (iii) *of Theorem* 2.5.1*, are true for μ.*

2.6.5. Remark. The measurability and integrability of complex functions f with respect to a measure μ are defined as the measurability and integrability of the real and imaginary parts of f, denoted by $\mathrm{Re}\, f$ and $\mathrm{Im}\, f$, respectively. Set

$$\int f\,d\mu := \int \mathrm{Re}\, f\,d\mu + i\int \mathrm{Im}\, f\,d\mu.$$

For mappings with values in $\mathbb{R}^n$, the measurability and integrability are defined analogously, i.e., coordinate-wise. Thus, the integral of a mapping $f = (f_1, \ldots, f_n)$ with integrable components f_i is the vector whose coordinates are the integrals of f_i. We draw attention to the fact that the coordinate-wise measurability of the mapping $f = (f_1, \ldots, f_n)$ with respect to a σ-algebra $\mathcal{A}$ is equivalent to the inclusion $f^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}(\mathbb{R}^n)$ (see Lemma 2.12.5).

2.7. The completeness of the space L^1

In this section, we show that the space of Lebesgue integrable functions possesses the important property of completeness, i.e., every mean fundamental sequence converges in the mean (the Riemann integral does not have this property). As in the case of simple functions, we introduce the corresponding notion.

2.7.1. Definition. (i) *A sequence of functions f_n that are integrable with respect to a measure μ (possibly with values in $[0,+\infty]$) is called fundamental in the mean or mean fundamental if, for every $\varepsilon > 0$, there exists a number N such that*

$$\int_X |f_n(x) - f_k(x)|\,\mu(dx) < \varepsilon, \quad \forall\, n,\, k \geq N.$$

(ii) *We say that a sequence of μ-integrable functions f_n converges to a μ-integrable function f in the mean if*

$$\lim_{n\to\infty} \int_X |f(x) - f_n(x)|\,\mu(dx) = 0.$$

Mean fundamental or mean convergent sequences are also called fundamental or convergent in $L^1(\mu)$.

Such a convergence is just convergence with respect to the natural norm of the space $L^1(\mu)$, which is discussed in greater detail in Chapter 4.

First we consider the case where μ is a bounded measure and then extend the results to measures with values in $[0,+\infty]$.

2.7.2. Lemma. *Suppose that a sequence of simple functions φ_j is fundamental in the mean and converges a.e. to φ. Then*

$$\lim_{j\to\infty} \int_X |\varphi(x) - \varphi_j(x)|\,\mu(dx) = 0. \tag{2.7.1}$$

PROOF. Let $\varepsilon > 0$. By Lemma 2.3.4 applied to the sequence $\{\varphi_j\}$ and the absolute continuity of the Lebesgue integral, there exists $\delta > 0$ such that for all n one has

$$\int_D \bigl[|\varphi(x)| + |\varphi_n(x)|\bigr]\,\mu(dx) < \varepsilon$$

for any set D with measure less than δ. By Egoroff's theorem, there exists a set X_δ such that $\mu(X\backslash X_\delta) < \delta$ and on X_δ the sequence $\{\varphi_n\}$ converges to φ uniformly. Hence there exists a number N such that for all $j \geq N$ one has $\sup_{x\in X_\delta} |\varphi_j(x) - \varphi(x)| < \varepsilon$. Then, for all $n \geq N$, we have

$$\begin{aligned}
&\int_X |\varphi(x) - \varphi_n(x)|\,\mu(dx) \\
&\leq \int_{X_\delta} |\varphi(x) - \varphi_n(x)|\,\mu(dx) + \int_{X\backslash X_\delta} |\varphi(x) - \varphi_n(x)|\,\mu(dx) \leq \varepsilon\mu(X) + \varepsilon,
\end{aligned}$$

which proves (2.7.1). □

2.7.3. Theorem. *If a sequence of μ-integrable functions f_n is fundamental in the mean, then it converges in the mean to some μ-integrable function f.*

PROOF. By the definition of integrability of f_n and Lemma 2.7.2 we obtain that, for every n, one can find a simple function g_n such that

$$\int_X |f_n(x) - g_n(x)|\,\mu(dx) \le \frac{1}{n}. \tag{2.7.2}$$

Then, the sequence $\{g_n\}$ is fundamental in the mean, since

$$\begin{aligned}\int_X |g_n(x) - g_k(x)|\,\mu(dx) \\ \le \int_X \Bigl[|g_n(x) - f_n(x)| + |f_n(x) - f_k(x)| + |f_k(x) - g_k(x)|\Bigr]\,\mu(dx) \\ \le \frac{1}{n} + \frac{1}{k} + \int_X |f_n(x) - f_k(x)|\,\mu(dx).\end{aligned}$$

In addition, by the Chebyshev inequality, one has

$$\mu\Bigl(x\colon\ |g_n(x) - g_k(x)| \ge c\Bigr) \le c^{-1}\int_X |g_n(x) - g_k(x)|\,\mu(dx),$$

hence the sequence $\{g_k\}$ is fundamental in measure and converges in measure to some function f. By the Riesz theorem, there exists a subsequence $\{g_{n_k}\}$ convergent to f almost everywhere. By definition, the function f is integrable. Relations (2.7.1) and (2.7.2) yield mean convergence of $\{f_n\}$ to f, since

$$\int_X |f(x) - f_n(x)|\,\mu(dx) \le \int_X \Bigl[|f(x) - g_n(x)| + |g_n(x) - f_n(x)|\Bigr]\,\mu(dx).$$

The theorem is proven. □

According to the terminology introduced in Chapter 4, the proven fact means the completeness of the normed space $L^1(\mu)$.

2.7.4. Corollary. *If a mean fundamental sequence of μ-integrable functions f_n converges almost everywhere to a function f, then the function f is integrable and the sequence $\{f_n\}$ converges to f in the mean.*

It is clear from Proposition 2.6.2 and Remark 2.6.3 that the results of this section remain valid for infinite countably additive measures.

2.7.5. Corollary. *The assertions of Theorem* 2.7.3 *and Corollary* 2.7.4 *are true in the case where μ is a countably additive measure with values in* $[0, +\infty]$.

The result of this section gives a new proof of the completeness of the measure algebra $(\mathcal{A}/\mu, d)$ verified in §1.12(iii). To this end, we identify any measurable set A with its indicator function and observe that the indicator functions form a closed set in $L^1(\mu)$ and that $\mu(A \triangle B)$ coincides with the integral of $|I_A - I_B|$.

2.8. Convergence theorems

In this section, we prove the three principal theorems on convergence of integrable functions; these theorems bear the names of Lebesgue, Beppo Levi, and Fatou. As usual, we suppose first that μ is a bounded nonnegative measure on a space X with a σ-algebra $\mathcal{A}$. The most important in the theory of integral is the following *Lebesgue dominated convergence theorem.*

2.8.1. Theorem. *Suppose that μ-integrable functions f_n converge almost everywhere to a function f. If there exists a μ-integrable function Φ such that*

$$|f_n(x)| \le \Phi(x) \quad \textit{a.e. for every } n,$$

then the function f is integrable and

$$\int_X f(x)\,\mu(dx) = \lim_{n\to\infty} \int_X f_n(x)\,\mu(dx). \tag{2.8.1}$$

In addition,

$$\lim_{n\to\infty} \int_X |f(x) - f_n(x)|\,\mu(dx) = 0.$$

PROOF. The function f is measurable, since it is the limit of an almost everywhere convergent sequence of measurable functions. The integrability of f follows by the estimate $|f| \le \Phi$ a.e. Let $\varepsilon > 0$. By the absolute continuity of the Lebesgue integral, there exists $\delta > 0$ such that

$$\int_D \Phi(x)\,\mu(dx) < \frac{\varepsilon}{4} \quad \text{if } \mu(D) < \delta.$$

By Egoroff's theorem, there is a set X_δ such that $\mu(X\backslash X_\delta) < \delta$ and the functions f_n converge to f uniformly on X_δ. Hence there exists a number N such that

$$|f_n(x) - f(x)| \le \frac{\varepsilon}{2\mu(X) + 1}$$

for all $n \ge N$. Therefore, for $n \ge N$ we have

$$\begin{aligned}
&\int_X |f(x) - f_n(x)|\,\mu(dx) \\
&\qquad \le \int_{X\backslash X_\delta} |f(x) - f_n(x)|\,\mu(dx) + \int_{X_\delta} |f(x) - f_n(x)|\,\mu(dx) \\
&\qquad\qquad \le 2\int_{X\backslash X_\delta} \Phi(x)\,\mu(dx) + \frac{\varepsilon}{2\mu(X)+1}\mu(X_\delta) \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

The theorem is proven. $\square$

The next very important result is the *monotone convergence theorem* due to Lebesgue and Beppo Levi.

2.8.2. Theorem. *Let $\{f_n\}$ be a sequence of μ-integrable functions such that $f_n(x) \le f_{n+1}(x)$ a.e. for each $n \in \mathbb{N}$. Suppose that*

$$\sup_n \int_X f_n(x)\,\mu(dx) < \infty.$$

Then, the function $f(x) = \lim\limits_{n\to\infty} f_n(x)$ *is almost everywhere finite and integrable. In addition, equality* (2.8.1) *holds true.*

Proof. For $n \le m$ we have

$$\int_X |f_m - f_n|\,d\mu = \int_X (f_m - f_n)\,d\mu = \int_X f_m\,d\mu - \int_X f_n\,d\mu.$$

Since the sequence of integrals of the functions f_n is increasing and bounded, it is convergent. Therefore, the above equality implies that the sequence $\{f_n\}$ is fundamental in the mean, hence converges in the mean to some integrable function g. Mean convergence yields convergence in measure (due to the Chebyshev inequality). By the Riesz theorem some subsequence $\{f_{n_k}\} \subset \{f_n\}$ converges to g almost everywhere. By the monotonicity, the whole sequence $f_n(x)$ converges to $g(x)$ for almost all x, whence we obtain the equality $f(x) = g(x)$ almost everywhere. In particular, $f(x) < \infty$ a.e. The last claim follows by the Lebesgue theorem, since $|f_n(x)| \le |f(x)| + |f_1(x)|$ a.e. for each n. □

The third frequently used result is *Fatou's theorem* (sometimes it is called Fatou's lemma).

2.8.3. Theorem. *Let* $\{f_n\}$ *be a sequence of nonnegative* μ*-integrable functions convergent to a function* f *almost everywhere and let*

$$\sup_n \int_X f_n(x)\,\mu(dx) \le K < \infty.$$

Then, the function f *is* μ*-integrable and*

$$\int_X f(x)\,\mu(dx) \le K.$$

Moreover,

$$\int_X f(x)\,\mu(dx) \le \liminf_{n\to\infty} \int_X f_n(x)\,\mu(dx).$$

Proof. Set $g_n(x) = \inf_{k\ge n} f_k(x)$. Then

$$0 \le g_n \le f_n, \quad g_n \le g_{n+1}.$$

Hence the functions g_n are integrable and form a monotone sequence, and their integrals are majorized by K. By the monotone convergence theorem, almost everywhere there exists a finite limit

$$g(x) = \lim_{n\to\infty} g_n(x),$$

the function g is integrable, its integral equals the limit of the integrals of the functions g_n and does not exceed K. It remains to observe that $f(x) = g(x)$ a.e. by convergence of $\{f_n(x)\}$ a.e. The last claim follows by applying what we have already proven to a suitably chosen subsequence. □

2.8.4. Corollary. *Let $\{f_n\}$ be a sequence of nonnegative μ-integrable functions such that*

$$\sup_n \int_X f_n(x)\,\mu(dx) \le K < \infty.$$

Then, the function $\liminf\limits_{n\to\infty} f_n$ is μ-integrable and one has

$$\int_X \liminf_{n\to\infty} f_n(x)\,\mu(dx) \le \liminf_{n\to\infty} \int_X f_n(x)\,\mu(dx) \le K.$$

PROOF. Note that

$$\liminf_{n\to\infty} f_n(x) = \lim_{k\to\infty} \inf_{n\ge k} f_n(x)$$

and apply Fatou's theorem. □

2.8.5. Theorem. *The dominated convergence theorem and Fatou's theorem remain valid if in place of almost everywhere convergence in their hypotheses we require convergence of $\{f_n\}$ to f in measure μ.*

PROOF. Since $\{f_n\}$ has a subsequence convergent to f almost everywhere, we obtain at once the analog of Fatou's theorem for convergence in measure, as well as the conclusion of the Lebesgue theorem for the chosen subsequence. It remains to observe that then our claim is true for the whole sequence $\{f_n\}$. Indeed, otherwise we could find a subsequence f_{n_k} such that

$$\int_X |f_{n_k} - f|\,d\mu \ge c > 0$$

for all k, but this is impossible because we would choose in $\{f_{n_k}\}$ a further subsequence convergent a.e., thus arriving at a contradiction. □

We now extend our results to measures with values in $[0, +\infty]$.

2.8.6. Corollary. *The dominated convergence theorem, monotone convergence theorem, Fatou's theorem, Corollary* 2.8.4 *and Theorem* 2.8.5 *remain valid in the case when μ is an unbounded countably additive measure with values in* $[0, +\infty]$.

PROOF. In order to extend these theorems to unbounded measures, one can apply Proposition 2.6.2 and Remark 2.6.3. Indeed, let μ be an unbounded measure and let $f_n(x) \to f(x)$ a.e., where the functions f_n are integrable. According to Remark 2.6.3, there exists a measurable set X_0 such that the measure μ on X_0 is σ-finite, i.e., X_0 is the countable union of pairwise disjoint sets $X_n \in \mathcal{A}$ of finite measure, and all functions f_n and f vanish on the complement of X_0. Let us take a function ϱ with countably many values that is strictly positive on X_0 and integrable with respect to μ (such a function has been constructed in Proposition 2.6.2). Let us consider the bounded measure $\nu = \varrho \cdot \mu$. The functions $F_n = f_n/\varrho$ and $F = f/\varrho$ are integrable with respect to ν and $F_n \to F$ ν-a.e. If the functions f_n are majorized by a μ-integrable function Φ, then the function $\Psi = \Phi/\varrho$ turns out to be ν-integrable and majorizes the sequence $\{F_n\}$. According to the dominated convergence theorem for the measure ν and the functions F_n, we obtain the corresponding assertion for μ and f_n. In a similar manner one extends to infinite measures all other results of this section. □

If one introduces the integral according to Theorem 2.5.2, then one can prove first the Beppo Levi theorem and derive from it the Lebesgue and Fatou theorems.

By using the Lebesgue dominated convergence theorem one proves the following assertion about continuity and differentiability of integrals with respect to a parameter.

2.8.7. Corollary. *Let μ be a nonnegative measure (possibly with values in $[0, +\infty]$) on a space X and let a function $f\colon\ X\times(a,b)\to \mathbb{R}^1$ be such that for every $\alpha\in(a,b)$ the function $x\mapsto f(x,\alpha)$ is integrable.*

(i) *Suppose that for μ-a.e. x the function $\alpha\mapsto f(x,\alpha)$ is continuous and there exists an integrable function Φ such that for every fixed α we have $|f(x,\alpha)|\le\Phi(x)$ μ-a.e. Then, the function*

$$J\colon\ \alpha\mapsto\int_X f(x,\alpha)\,\mu(dx)$$

is continuous.

(ii) *Suppose that, for μ-a.e. x, the function $\alpha\mapsto f(x,\alpha)$ is differentiable and there exists a μ-integrable function Φ such that for μ-a.e. x we have $|\partial f(x,\alpha)/\partial\alpha|\le\Phi(x)$ for all α simultaneously. Then, the function J is differentiable and*

$$J'(\alpha)=\int_X\frac{\partial f(x,\alpha)}{\partial\alpha}\,\mu(dx).$$

PROOF. Assertion (i) is obvious from the Lebesgue theorem. (ii) Let α be fixed and let $t_n\to 0$. Then, by the mean value theorem, for μ-a.e. x, there exists $\xi=\xi(x,\alpha,n)$ such that

$$\bigl|t_n^{-1}\bigl(f(x,\alpha+t_n)-f(x,\alpha)\bigr)\bigr|=|\partial f(x,\xi)/\partial\alpha|\le\Phi(x).$$

The above ratio converges to $\partial f(x,\alpha)/\partial\alpha$. By the Lebesgue theorem, the limit $\lim\limits_{n\to\infty}t_n^{-1}\bigl(J(\alpha+t_n)-J(\alpha)\bigr)$ equals the integral of $\partial f(x,\alpha)/\partial\alpha$. □

Exercise 2.12.68 contains a modification of assertion (ii), ensuring the differentiability at a single point.

Considering the functions $f_n(x) = nI_{(0,1/n]}(x)$ that converge to zero pointwise on $(0,1]$, we see that in the dominated convergence theorem one cannot omit the integrable majorant condition, and that in Fatou's theorem one cannot always interchange the limit and integral. An interesting consequence of the absence of integrable majorants is found in Exercise 10.10.43 in Chapter 10. However, it may happen that the functions f_n converge to f in the mean without having a common integrable majorant (Exercise 2.12.41).

In addition, there is no need to require the existence of integrable majorants in the following interesting theorem due to Young (see Young [**1034**, p. 315]).

2.8.8. Theorem. *Suppose we are given three sequences of μ-integrable functions $\{f_n\}$, $\{g_n\}$, and $\{h_n\}$ (where μ may take values in $[0,+\infty]$) such that*

$$g_n(x) \le f_n(x) \le h_n(x) \quad a.e.$$

and

$$\lim_{n\to\infty} f_n(x) = f(x), \ \lim_{n\to\infty} g_n(x) = g(x), \ \lim_{n\to\infty} h_n(x) = h(x).$$

Let g and h be integrable and let

$$\lim_{n\to\infty} \int_X h_n\, d\mu = \int_X h\, d\mu, \quad \lim_{n\to\infty} \int_X g_n\, d\mu = \int_X g\, d\mu.$$

Then f is integrable and

$$\lim_{n\to\infty} \int_X f_n\, d\mu = \int_X f\, d\mu.$$

PROOF. It is clear that f is integrable, since $g(x) \le f(x) \le h(x)$ a.e., whence we obtain $|f(x)| \le |g(x)| + |h(x)|$ a.e. By Fatou's theorem we obtain the relation

$$\begin{aligned}\int_X f\, d\mu - \int_X g\, d\mu &= \int_X \lim_{n\to\infty} (f_n - g_n)\, d\mu \\ &\le \liminf_{n\to\infty} \int_X (f_n - g_n)\, d\mu = \liminf_{n\to\infty} \int_X f_n\, d\mu - \int_X g\, d\mu,\end{aligned}$$

whence one has

$$\int_X f\, d\mu \le \liminf_{n\to\infty} \int_X f_n\, d\mu.$$

Similarly, by using h_n we obtain

$$\int_X f\, d\mu \ge \limsup_{n\to\infty} \int_X f_n\, d\mu.$$

Note that we could also apply the concept of the uniform absolute continuity (see §4.5 and Exercise 4.7.71). □

In Young's theorem, the functions f_n may not converge to f in the mean, but if $g_n \le 0 \le h_n$, then we also have mean convergence, which follows at once from this theorem and the estimate $0 \le |f_n - f| \le h_n - g_n + |f|$. A simple corollary of Young's theorem is the following useful fact obtained in the works of Vitali (and also Young and Fichtenholz) for Lebesgue measure and later rediscovered by Scheffé in the general case (it is called in the literature the "Scheffé theorem"; it appears that the name "Vitali–Scheffé theorem" is more appropriate).

2.8.9. Theorem. *If nonnegative μ-integrable functions f_n converge a.e. to a μ-integrable function f (where μ is a measure with values in $[0,+\infty]$) and*

$$\lim_{n\to\infty}\int_X f_n\,d\mu = \int_X f\,d\mu,$$

then

$$\lim_{n\to\infty}\int_X |f-f_n|\,d\mu = 0.$$

For functions f_n of arbitrary sign convergent a.e. to f, the mean convergence of f_n to f is equivalent to convergence of the integrals of $|f_n|$ to the integral of $|f|$.

PROOF. Since $0 \le |f_n - f| \le |f_n| + |f|$, Young's theorem applies. □

An interesting generalization of this result is contained in Proposition 4.7.30 in Chapter 4.

All the results in this section have exceptional significance in the theory of measure and integration, which we shall see below. So, as an application of these results we consider just one, but rather typical example of how Fatou's theorem works.

2.8.10. Example. Suppose we are given a sequence of integrable functions f_n on a space X with a probability measure μ and that there exists $M > 0$ such that, for all $n \in \mathbb{N}$, we have

$$\int_X \left| f_n(x) - \int f_n\,d\mu \right|^2 \mu(dx) \le M \int_X |f_n|\,d\mu.$$

Then either

$$\limsup_{n\to\infty}\int_X |f_n|\,d\mu < \infty \quad \text{and} \quad \liminf_{n\to\infty}|f_n(x)| < \infty \quad \text{a.e.}$$

or

$$\limsup_{n\to\infty}\int_X |f_n|\,d\mu = \infty \quad \text{and} \quad \limsup_{n\to\infty}|f_n(x)| = \infty \quad \text{a.e.}$$

In particular, if for a.e. x the sequence of numbers $f_n(x)$ is bounded, then the integrals of $|f_n|$ are uniformly bounded.

PROOF. We observe that

$$\int_X \left| |f_n(x)| - \int_X |f_n|\,d\mu \right|^2 \mu(dx) = \int_X |f_n|^2\,d\mu - \left|\int_X |f_n|\,d\mu\right|^2$$
$$\le \int_X |f_n|^2\,d\mu - \left|\int_X f_n\,d\mu\right|^2 = \int_X \left| f_n(x) - \int_X f_n\,d\mu\right|^2 \mu(dx) \le M\int_X |f_n|\,d\mu,$$

since the absolute value of the integral of f_n does not exceed the integral of $|f_n|$. This inequality, weaker than in the theorem, will actually be used. Let J_n be the integral of $|f_n|$. If the numbers J_n are bounded, then, by Fatou's theorem, one has $\liminf\limits_{n\to\infty}|f_n(x)| < \infty$ a.e. Otherwise, passing to a subsequence, we may assume that $J_n \to \infty$. The above-mentioned inequality yields

$$\int_X \Bigl| |f_n(x)|/\sqrt{J_n} - \sqrt{J_n} \Bigr|^2 \mu(dx) \le M.$$

By Fatou's theorem, one has $\liminf\limits_{n\to\infty} \Bigl| |f_n(x)|/\sqrt{J_n} - \sqrt{J_n} \Bigr| < \infty$ a.e., whence it follows that $\limsup\limits_{n\to\infty} |f_n(x)| = \infty$ a.e. □

Exercise 2.12.95 contains a generalization of this example. In Chapter 4 and other exercises in this chapter, other useful results related to limits under the integral sign are given.

2.9. Criteria of integrability

The definition of the integral is almost never used for verification of the integrability of concrete functions. Very efficient and frequently practically used sufficient conditions of integrability are given by the Beppo Levi and Fatou theorems. In real problems, one of the most obvious criteria of integrability of measurable functions is employed: majorization in the absolute value by an integrable function. In this section, we derive from this trivial criterion several less obvious ones and obtain the integrability criteria in terms of convergence of series or Riemannian integrals over the real line.

2.9.1. Theorem. *Let $(X, \mathcal{A}, \mu)$ be a space with a finite nonnegative measure and let f be a μ-measurable function. Then, the integrability of f with respect to μ is equivalent to convergence of the series*

$$\sum_{n=1}^{\infty} n\,\mu\bigl(x\colon\ n \le |f(x)| < n+1\bigr), \tag{2.9.1}$$

and is also equivalent to convergence of the series

$$\sum_{n=1}^{\infty} \mu\bigl(x\colon\ |f(x)| \ge n\bigr). \tag{2.9.2}$$

PROOF. Let $A_0 = \bigl\{x\colon\ |f(x)| < 1\bigr\}$. Set $A_n = \bigl\{x\colon\ n \le |f(x)| < n+1\bigr\}$ for $n \in \mathbb{N}$. Then the sets A_n are μ-measurable disjoint sets whose union is the whole space up to a measure zero set. The function g defined by the equality $g|_{A_n} = n$, $n = 0, 1, \ldots$, is obviously μ-measurable and one has $g(x) \le |f(x)| \le g(x)+1$. Therefore, the function g is integrable precisely when f is integrable. According to Example 2.5.8, the integrability of g is equivalent to convergence of the series (2.9.1). It remains to observe that the series (2.9.1) and (2.9.2) converge or diverge simultaneously. Indeed, $\bigl\{x\colon\ |f(x)| \ge n\bigr\} = \bigcup_{k=n}^{\infty} A_k$, whence we obtain

$$\mu\bigl(x\colon\ |f(x)| \ge n\bigr) = \sum_{k=n}^{\infty} \mu(A_k).$$

Thus, taking the sum in n, we count the number $\mu(A_n)$ on the right-hand side n times. □

2.9.2. Example. (i) A function f measurable with respect to a bounded nonnegative measure μ is integrable in every degree $p \in (0, \infty)$ precisely when the function $\mu(x\colon\ |f(x)| > t)$ decreases faster than any power of t as $t \to +\infty$.

(ii) The function $|\ln x|^p$ on $(0, 1)$ is integrable with respect to Lebesgue measure for all $p > -1$, and the function x^α is integrable if $\alpha > -1$.

For infinite measures the indicated criteria do not work, since they do not take into account sets of small values of $|f|$. They can be modified for infinite measures, but we give instead a universal criterion. One of its advantages is a reduction of the problem to a certain Riemannian integral.

2.9.3. Theorem. *Let μ be a countably additive measure with values in $[0, +\infty]$ and let f be a μ-measurable function. Then, the μ-integrability of f is equivalent to the integrability of the function $t \mapsto \mu\bigl(x\colon\ |f(x)| > t\bigr)$ on $(0, +\infty)$ with respect to Lebesgue measure. In addition,*

$$\int_X |f(x)|\,\mu(dx) = \int_0^\infty \mu\bigl(x\colon\ |f(x)| > t\bigr)\,dt. \tag{2.9.3}$$

Proof. There are three different proofs of (2.9.3) in this book: see Theorem 3.4.7 in Chapter 3, where a simple geometric reasoning involving double integrals is given, and Exercise 5.8.112 in Chapter 5, where an even shorter proof is based on integration by parts. Here no additional facts are needed. Let f be integrable. Then, for any n, the function f_n equal $|f(x)|$ if $n^{-1} \le |f(x)| \le n$ and 0 otherwise is integrable as well. If we prove (2.9.3) for f_n in place of f, then, as $n \to \infty$, we obtain this equality for f, since the integrals of f_n converge to the integral of $|f|$, and the sets $\{x\colon\ f_n(x) > t\}$ increase for every t to $\{x\colon\ |f(x)| > t\}$ so that the monotone convergence theorem applies. The function f_n is nonzero on a set of finite measure. Thus, the general case is reduced to the case of a finite measure and bounded function. The next obvious step is a reduction to simple functions; it is accomplished by choosing a sequence of simple functions g_n uniformly convergent to f. Clearly, $\mu(x\colon\ |f(x)| > t) = \lim\limits_{n\to\infty} \mu(x\colon\ |g_n(x)| > t)$ for each t, with the exception of an at most countable set of points t, where $\mu(x\colon\ |f(x)| = t) > 0$ (this is readily verified). Hence it remains to obtain (2.9.3) for simple functions. This case is verified directly: if $|f|$ assumes values $c_1 < \cdots < c_n$ on sets $A_1, \ldots, A_n$, then on $[c_{j-1}, c_j)$ the function $\mu(x\colon\ |f(x)| > t)$ equals $\mu(B_{n+1-j})$, where $B_j := A_{n+1-j} \cup \cdots \cup A_n$ for $j = 1, \ldots, n$. The reader can easily provide the details.

If the function $\mu\bigl(x\colon\ |f(x)| > t\bigr)$ on the half-line is integrable, then the functions $\mu\bigl(x\colon\ |f_n(x)| > t\bigr)$, where the functions f_n are defined above, are integrable as well. It is clear that the set $\{|f| \ge 1/n\}$ has finite measure. Hence the bounded functions f_n are integrable. According to (2.9.3) the integrals of f_n are majorized by the integral of $\mu\bigl(x\colon\ |f(x)| > t\bigr)$ over the half-line, which yields the integrability of f by Fatou's theorem. $\square$

2.10. Connections with the Riemann integral

We assume that the reader is familiar with the definition of the Riemann integral (see, e.g., Rudin [**834**]). In particular, the Riemann integral of the indicator function of an interval is the interval length, hence for piecewise constant functions on an interval the Riemann integral coincides with the Lebesgue one.

2.10.1. Theorem. *If a function f is Riemann integrable in the proper sense on the interval $I=[a,b]$, then it is Lebesgue integrable on I and its Riemann and Lebesgue integrals are equal.*

PROOF. We may assume that $b-a=1$. For every $n\in\mathbb{N}$ we partition the interval $I=[a,b]$ into disjoint intervals $[a,a+2^{-n}),\ldots,[b-2^{-n},b]$ of length 2^{-n}. These intervals are denoted by $I_1,\ldots,I_{2^n}$. Let $m_k=\inf_{x\in I_k}f(x)$, $M_k=\sup_{x\in I_k}f(x)$. Let us consider step functions f_n and g_n defined as follows: $f_n=m_k$ on I_k, $g_n=M_k$ on I_k, $k=1,\ldots,2^n$. It is clear that $f_n(x)\le f(x)\le g_n(x)$. In addition, $f_n(x)\le f_{n+1}(x)$, $g_{n+1}(x)\le g_n(x)$. Hence the limits $\varphi(x):=\lim\limits_{n\to\infty}f_n(x)$ and $\psi(x):=\lim\limits_{n\to\infty}g_n(x)$ exist, and one has $\varphi(x)\le f(x)\le\psi(x)$. It is known from the elementary calculus that the Riemann integrability of f implies the equality

$$\lim_{n\to\infty}\int_a^b f_n(x)\,dx=\lim_{n\to\infty}\int_a^b g_n(x)\,dx=R(f), \tag{2.10.1}$$

where $R(f)$ denotes the Riemann integral of f (we also use the aforementioned coincidence of the Riemann and Lebesgue integrals for piecewise constant functions). The functions φ and ψ are bounded and Lebesgue measurable (being the limits of step functions), hence they are Lebesgue integrable. It is clear that

$$\int_a^b f_n(x)\,dx\le\int_a^b\varphi(x)\,dx\le\int_a^b\psi(x)\,dx\le\int_a^b g_n(x)\,dx$$

for all n. By (2.10.1) the Lebesgue integrals of the functions φ and ψ equal $R(f)$, hence $\varphi(x)=\psi(x)$ a.e., since $\varphi(x)\le\psi(x)$. Therefore, $\varphi=f=\psi$ a.e., which yields our claim. $\square$

There exist functions on an interval that have improper Riemann integrals but are not Lebesgue integrable (see Exercise 2.12.37). However, the existence of the absolute improper Riemann integral implies the Lebesgue integrability.

2.10.2. Theorem. *Suppose that a function f is integrable on an interval I (bounded or unbounded) in the improper Riemann sense along with the function $|f|$. Then f is Lebesgue integrable on I and its improper Riemann integral equals its Lebesgue integral.*

PROOF. We consider the case where the interval $I=(a,b]$ is bounded and f is integrable in the proper Riemann sense on every interval $[a+\varepsilon,b]$, $\varepsilon>0$. The case where $a=-\infty$, is similar, and the general case reduces to finitely many considered ones. Let $f_n = f$ on $[a+n^{-1}, b]$, $f_n = 0$ on $(a, a+n^{-1})$. By the Riemann integrability, the function f is Lebesgue measurable on the interval $[a+n^{-1}, b]$, hence the function f_n is measurable. It is clear that $f_n \to f$ pointwise, hence f is measurable on $(a, b]$. By the improper integrability of $|f|$, the functions $|f_n| \le |f|$ have the uniformly bounded Lebesgue integrals (equal to their Riemann integrals by the previous theorem). By the Beppo Levi theorem (or by the Fatou theorem), the function $|f|$ is Lebesgue integrable. By the dominated convergence theorem, the Lebesgue integrals of the functions f_n over $(a, b]$ approach the Lebesgue integral of f. Hence the Lebesgue integral of f equals the improper Riemann integral. $\square$

It is worth noting that even the absolute improper Riemann integral has no completeness property from §2.7: let us take step functions on $[0, 1]$ convergent in the mean to the indicator of the compact set from Example 1.7.6 (or the set from Exercise 2.12.28).

Closing our discussion of the links between the Riemann and Lebesgue integrals we observe that the Lebesgue integral of a function of a real variable can be expressed by means of certain generalized Riemann sums, although not as constructively as the Riemann integral. For example, if the function f has a period 1 and is integrable over its period, then its integral over $[0, 1]$ equals the limit of the sums

$$2^{-n}\sum_{k=1}^{2^n} f(x_0 + k2^{-n})$$

for a.e. x_0. Concerning this question, see Exercise 2.12.63, Exercise 4.7.101 in Chapter 4, the section on the Henstock–Kurzweil integral in Chapter 5, and Example 10.3.18 in Chapter 10.

2.11. The Hölder and Minkowski inequalities

Let $(X, \mathcal{A}, \mu)$ be a space with a nonnegative measure μ (finite or with values in $[0, +\infty]$) and let $p \in (0, +\infty)$. Let $\mathcal{L}^p(\mu)$ denote the set of all μ-measurable functions f such that $|f|^p$ is μ-integrable. In particular, $\mathcal{L}^1(\mu)$ is the set of all μ-integrable functions. Let $\mathcal{L}^0(\mu)$ denote the class of all μ-a.e. finite μ-measurable functions. Two μ-measurable functions f and g are called equivalent if $f = g$ μ-a.e. The corresponding notation is $f \sim g$. It is clear that if $f \sim g$ and $g \sim h$, then $f \sim h$ and $g \sim f$. In addition, $f \sim f$. Thus, we obtain an equivalence relation and the collection $\mathcal{L}^0(\mu)$ of all measurable functions is partitioned into disjoint classes of pairwise equivalent functions. We denote by $L^0(\mu)$ and $L^p(\mu)$ the corresponding factor-spaces of the spaces $\mathcal{L}^0(\mu)$ and $\mathcal{L}^p(\mu)$ with respect to this equivalence relation. Thus, $L^p(\mu)$ is the set of all equivalence classes of μ-measurable functions f such that $|f|^p$ is integrable. The same notation is used for complex-valued functions. In the case of Lebesgue measure on $\mathbb{R}^n$ or on a set $E \subset \mathbb{R}^n$ we use the symbols $\mathcal{L}^p(\mathbb{R}^n)$, $L^p(\mathbb{R}^n)$, $\mathcal{L}^p(E)$, and $L^p(E)$ without explicit indication of measure. In place of $L^p([a, b])$ and $L^p\big([a, +\infty)\big)$ we write $L^p[a, b]$ and $L^p[a, +\infty)$.

Sometimes it is necessary to explicitly indicate the space X in the above notation; then the symbols $\mathcal{L}^p(X,\mu)$, $L^p(X,\mu)$ are employed. It is customary in books and articles on measure theory to allow the deliberate ambiguity of notation in the expressions of the type "a function f in L^p", where one should say "a function f in $\mathcal{L}^p$" or the "equivalence class of a function f in L^p". Normally this does not lead to misunderstanding and may be even helpful in formulations as an implicit indication that the assertion is valid not only for an individual function, but for the whole equivalence class. We do not always strictly distinguish between functions and their classes, too. However, one should remember that from the formal point of view, an expression like "a continuous function f from L^p" is not perfectly correct, although one can hardly advise the precise expression "the equivalence class of $f \in L^p$ contains a continuous function". Certainly, one can simply say "a continuous function $f \in \mathcal{L}^p(\mu)$".

For $1 \le p < \infty$ we set

$$\|f\|_p := \|f\|_{L^p(\mu)} := \Bigl(\int_X |f|^p \, d\mu\Bigr)^{1/p}, \quad f \in \mathcal{L}^p(\mu).$$

The same notation is used for elements of $L^p(\mu)$.

Finally, let $\mathcal{L}^\infty(\mu)$ be the set of all *bounded everywhere defined* μ-measurable functions. For $f \in \mathcal{L}^\infty(\mu)$ we set

$$\|f\|_{L^\infty(\mu)} := \|f\|_\infty := \inf_{\widetilde{f}\sim f} \sup_{x\in X} |\widetilde{f}(x)|.$$

A function f is called *essentially bounded* if it coincides μ-a.e. with a bounded function. Then the number $\|f\|_\infty$ is defined as above. An alternative notation is $\operatorname{esssup}|f|$, $\operatorname{vrai\,sup}|f|$.

In the study of the spaces $\mathcal{L}^p(\mu)$ and the corresponding normed spaces $L^p(\mu)$ considered in Chapter 4, we need the following Hölder inequality, which is very important in its own right, being one of the most frequently used inequalities in the theory of integration.

2.11.1. Theorem. *Suppose that $1 < p < \infty$, $q = p(p-1)^{-1}$, $f \in \mathcal{L}^p(\mu)$, $g \in \mathcal{L}^q(\mu)$. Then $fg \in \mathcal{L}^1(\mu)$ and $\|fg\|_1 \le \|f\|_p \|g\|_q$, i.e., one has*

$$\int_X |fg| \, d\mu \le \Bigl(\int_X |f|^p \, d\mu\Bigr)^{1/p} \Bigl(\int_X |g|^q \, d\mu\Bigr)^{1/q}. \tag{2.11.1}$$

Proof. The function fg is defined a.e. and measurable. It is readily shown (see Exercise 2.12.87) that for all nonnegative a and b one has the inequality $ab \le \frac{a^p}{p} + \frac{b^q}{q}$. Therefore,

$$\frac{|f(x)|}{\|f\|_p} \frac{|g(x)|}{\|g\|_q} \le \frac{1}{p} \frac{|f(x)|^p}{\|f\|_p^p} + \frac{1}{q} \frac{|g(x)|^q}{\|g\|_q^q}.$$

The right-hand side of this inequality is integrable and its integral equals 1, hence the left-hand side is integrable as well and its integral does not exceed 1, which is equivalent to (2.11.1). □

2.11.2. Corollary. *Under the hypotheses of the above theorem one has*

$$\int_X fg\,d\mu \le \Bigl(\int_X |f|^p\,d\mu\Bigr)^{1/p}\Bigl(\int_X |g|^q\,d\mu\Bigr)^{1/q}. \tag{2.11.2}$$

In Exercise 2.12.89 the conditions for the equality in (2.11.2) are investigated.

An immediate corollary of the Hölder inequality is the following Cauchy–Bunyakowsky inequality (also called Cauchy–Bunyakowsky–Schwarz inequality), which, however, can be easily proved directly: see Chapter 4, §4.3.

2.11.3. Corollary. *Suppose that $f, g \in \mathcal{L}^2(\mu)$. Then $fg \in \mathcal{L}^1(\mu)$ and*

$$\int_X fg\,d\mu \le \Bigl(\int_X |f|^2\,d\mu\Bigr)^{1/2}\Bigl(\int_X |g|^2\,d\mu\Bigr)^{1/2}. \tag{2.11.3}$$

Letting $g = I_{\{f\neq 0\}}$ we arrive at the following estimate.

2.11.4. Corollary. *Suppose that $f \in \mathcal{L}^p(\mu)$ and $\mu\bigl(x\colon\ f(x)\neq 0\bigr) < \infty$. Then*

$$\int_X |f|\,d\mu \le \mu\bigl(x\colon\ f(x)\neq 0\bigr)^{1/q}\Bigl(\int_X |f|^p\,d\mu\Bigr)^{1/p}, \quad q = p(p-1)^{-1}.$$

Sometimes the following generalized Hölder inequality is useful; its partial case where $r = 1$, $p_1 = p$, $p_2 = q$ we have just proved.

2.11.5. Corollary. *Let $1 \le r, p_1, \ldots, p_n < \infty$, $1/p_1 + \cdots + 1/p_n = 1/r$, and let $f_1 \in \mathcal{L}^{p_1}(\mu), \ldots,\ f_n \in \mathcal{L}^{p_n}(\mu)$. Then $f_1 \cdots f_n \in \mathcal{L}^r(\mu)$ and one has*

$$\Bigl(\int_X |f_1\cdots f_n|^r\,d\mu\Bigr)^{1/r} \le \Bigl(\int_X |f_1|^{p_1}\,d\mu\Bigr)^{1/p_1}\cdots\Bigl(\int_X |f_n|^{p_n}\,d\mu\Bigr)^{1/p_n}. \tag{2.11.4}$$

Proof. We may assume that $r = 1$, passing to new exponents $p_i' = p_i/r$. For $n = 2$ inequality (2.11.4) is already known. We argue by induction on n and suppose that the desired inequality is known for $n-1$. Let us apply the usual Hölder inequality with the exponents p_1 and q given by the equality $1/q = 1/p_2 + \cdots + 1/p_n$ to the integral of the product $|f_1||f_2\cdots f_n|$ and estimate it by $\|f_1\|_{p_1}\|f_2\cdots f_n\|_q$. Now we apply the inductive assumption and obtain

$$\|f_2\cdots f_n\|_q \le \|f_2\|_{p_2}\cdots\|f_n\|_{p_n},$$

which completes the proof. □

Hölder's inequality may help to establish membership in L^p.

2.11.6. Example. Let μ be a finite nonnegative measure. Suppose that a μ-measurable function f satisfies the following condition: there exist $p \in (1, \infty)$ and $M \ge 0$ such that, for every function $\varphi \in \mathcal{L}^\infty(\mu)$, one has $f\varphi \in L^1(\mu)$ and

$$\int_X f\varphi\,d\mu \le M\|\varphi\|_{L^p(\mu)}.$$

Then $f \in L^q(\mu)$, where $q = p(p-1)^{-1}$, and $\|f\|_{L^q(\mu)} \le M$.

Indeed, taking for φ the functions $\varphi_n := \operatorname{sgn} f|f|^{p-1}I_{\{|f|\le n\}}$, we obtain

$$\int_{\{|f|\le n\}} |f|^p\,d\mu \le M\Bigl(\int_{\{|f|\le n\}} |f|^p\,d\mu\Bigr)^{1/q},$$

which gives the estimate $\|fI_{\{|f|\le n\}}\|_{L^p(\mu)} \le M$. By Fatou's theorem we arrive at the desired conclusion. The same is true for infinite measures if the hypothesis is fulfilled for all $\varphi \in \mathcal{L}^\infty(\mu)\cap\mathcal{L}^q(\mu)$.

We recall that Chebyshev's inequality estimates large deviations of a function from above. As observed in Salem, Zygmund [**842**], one can estimate moderate deviations of functions from below by using Hölder's inequality.

2.11.7. Proposition. *Let μ be a probability measure on a measurable space $(X,\mathcal{A})$, let $f\in\mathcal{L}^p(\mu)$, where $1<p<\infty$, and let $q=p(p-1)^{-1}$. Then one has*

$$\mu\Bigl(x\colon\ |f(x)|\ge\lambda\|f\|_{L^1(\mu)}\Bigr)\ge(1-\lambda)^q\frac{\|f\|^q_{L^1(\mu)}}{\|f\|^q_{L^p(\mu)}},\quad \lambda\in[0,1]. \qquad (2.11.5)$$

PROOF. Letting $A=\{x\colon\ |f(x)|\ge\lambda\|f\|_{L^1(\mu)}\}$ and $g=|f|I_A$ one has

$$\Bigl(\int_X g\,d\mu\Bigr)^p \le \mu(A)^{p/q}\int_X g^p\,d\mu \le \mu(A)^{p/q}\int_X |f|^p\,d\mu.$$

Since $\|f\|_{L^1(\mu)}\le\|g\|_{L^1(\mu)}+\lambda\|f\|_{L^1(\mu)}$, i.e., $(1-\lambda)\|f\|_{L^1(\mu)}\le\|g\|_{L^1(\mu)}$, we obtain $(1-\lambda)^p\|f\|^p_{L^1(\mu)}\le\mu(A)^{p/q}\|f\|^p_{L^p(\mu)}$, which yields the claim. □

2.11.8. Example. Suppose that μ is a probability measure. Let a sequence $\{f_n\}\subset\mathcal{L}^2(\mu)$ be such that $0<\alpha\le\|f_n\|_{L^2(\mu)}\le\beta\|f_n\|_{L^1(\mu)}$ with some constants α, β. Then, for every $\lambda\in(0,1)$, the set of all points x such that $|f_n(x)|\ge\lambda\alpha\beta^{-1}$ for infinitely many numbers n has measure at least $(1-\lambda)^2\beta^{-2}$.

PROOF. We have

$$\mu\Bigl(x\colon\ |f_n(x)|\ge\lambda\alpha\beta^{-1}\Bigr)\ge(1-\lambda)^2\beta^{-2}.$$

It remains to refer to Exercise 1.12.89. □

Let us now turn to the following Minkowskiinequality.

2.11.9. Theorem. *Suppose that $p\in[1,+\infty)$ and $f,g\in\mathcal{L}^p(\mu)$. Then $f+g\in\mathcal{L}^p(\mu)$ and one has*

$$\Bigl(\int_X |f+g|^p\,d\mu\Bigr)^{1/p}\le\Bigl(\int_X |f|^p\,d\mu\Bigr)^{1/p}+\Bigl(\int_X |g|^p\,d\mu\Bigr)^{1/p}. \qquad (2.11.6)$$

PROOF. The function $f+g$ is defined a.e. and measurable. For $p=1$ inequality (2.11.6) is obvious. For $p>1$ we have $|f+g|^p\le 2^p(|f|^p+|g|^p)$, hence $|f+g|^p\in\mathcal{L}^1(\mu)$. We observe that

$$|f(x)+g(x)|^p\le|f(x)+g(x)|^{p-1}|f(x)|+|f(x)+g(x)|^{p-1}|g(x)|. \qquad (2.11.7)$$

Since $|f+g|^{p-1} \in \mathcal{L}^{p/(p-1)}(\mu) = \mathcal{L}^q(\mu)$, by the Hölder inequality one has

$$\int_X |f+g|^{p-1}|f|\,d\mu \le \Bigl(\int_X |f+g|^p\,d\mu\Bigr)^{1/q}\Bigl(\int_X |f|^p\,d\mu\Bigr)^{1/p}.$$

Estimating in a similar manner the integral of the second summand on the right-hand side of (2.11.7), we arrive at the estimate

$$\int_X |f+g|^p\,d\mu \le \Bigl(\int_X |f+g|^p\,d\mu\Bigr)^{1/q}\biggl[\Bigl(\int_X |f|^p\,d\mu\Bigr)^{1/p} + \Bigl(\int_X |g|^p\,d\mu\Bigr)^{1/p}\biggr].$$

Noting that $1-1/q=1/p$, we obtain $\|f+g\|_p \le \|f\|_p + \|g\|_p$. $\square$

Although one can take sums of functions in the spaces $\mathcal{L}^p(\mu)$ and multiply them by numbers (on sets of full measure), these spaces *are not linear*, since the indicated operations are not associative: for example, if a function f is not defined at a point x, then neither is $f+(-f)$, but this function must be everywhere zero because in a linear space there is only one zero element. Certainly, one could take in $\mathcal{L}^p(\mu)$ a subset consisting of all everywhere defined finite functions, which is a linear space, but it is more reasonable to pass to the space $L^p(\mu)$.

2.12. Supplements and exercises

2.12(i). The σ-algebra generated by a class of functions

Let $\mathcal{F}$ be a class of real functions on a set X.

2.12.1. Definition. *The smallest σ-algebra with respect to which all functions in $\mathcal{F}$ are measurable is called the σ-algebra generated by the class $\mathcal{F}$ and is denoted by $\sigma(\mathcal{F})$.*

It is clear that $\sigma(\mathcal{F})$ is the σ-algebra generated by all sets of the form $\{f<c\}$, $f\in\mathcal{F}$, $c\in\mathbb{R}^1$. Indeed, the σ-algebra generated by these sets belongs to $\sigma(\mathcal{F})$ and all functions in $\mathcal{F}$ are measurable with respect to it.

The simplest example of the σ-algebra generated by a class of functions is the case when the class $\mathcal{F}$ consists of a single function f. In this case

$$\sigma(\{f\}) = \bigl\{f^{-1}(B)\colon\ B\in\mathcal{B}(\mathbb{R}^1)\bigr\}.$$

Let $\mathbb{R}^\infty$ be the countable product of real lines, i.e., the space of all real sequences $x=(x_i)$. We denote by $\mathcal{B}(\mathbb{R}^\infty)$ the σ-algebra generated by all sets of the form

$$C_{i,t} = \{x\in\mathbb{R}^\infty\colon\ x_i<t\},\quad i\in\mathbb{N},\ t\in\mathbb{R}^1.$$

The sets in $\mathcal{B}(\mathbb{R}^\infty)$ are called Borel sets in $\mathbb{R}^\infty$. Functions on $\mathbb{R}^\infty$ measurable with respect to $\mathcal{B}(\mathbb{R}^\infty)$ are called Borel or Borel measurable.

2.12.2. Lemma. *Let $\mathcal{F}$ be a class of functions on a nonempty set X. Then, the σ-algebra $\sigma(\mathcal{F})$ generated by them coincides with the class of all sets of the form*

$$E_{(f_i),B} = \bigl\{x\colon\ \bigl(f_1(x),\ldots,f_n(x),\ldots\bigr)\in B\bigr\},\quad f_i\in\mathcal{F},\ B\in\mathcal{B}(\mathbb{R}^\infty). \tag{2.12.1}$$

PROOF. It is clear that sets of the indicated type form a σ-algebra. We denote it by $\mathcal{E}$. This σ-algebra contains all sets $\{f<c\}$, where $f\in\mathcal{F}$, $c\in\mathbb{R}^1$. Indeed, if we take all f_n equal f and put $B=C_{1,t}$, then $E_{(f_i),B}=\{f<t\}$. Hence $\sigma(\mathcal{F})\subset\mathcal{E}$.

On the other hand, $E_{(f_i),B}\in\sigma(\mathcal{F})$ for $B\in\mathcal{B}(\mathbb{R}^\infty)$. Indeed, it is readily verified that for fixed $f_1,\ldots,f_n,\ldots$ the class

$$\mathcal{B}_0=\bigl\{B\in\mathcal{B}(\mathbb{R}^\infty)\colon\ E_{(f_i),B}\in\sigma(\mathcal{F})\bigr\}$$

is a σ-algebra. The sets $C_{i,t}$ belong to $\mathcal{B}_0$ by the definition of $\sigma(\mathcal{F})$. Hence, $\mathcal{B}(\mathbb{R}^\infty)\subset\mathcal{B}_0$ as claimed. It follows that $\mathcal{E}\subset\sigma(\mathcal{F})$, whence $\mathcal{E}=\sigma(\mathcal{F})$. $\square$

2.12.3. Theorem. *Let $\mathcal{F}$ be a class of functions on a nonempty set X. Then, a function g on X is measurable with respect to $\sigma(\mathcal{F})$ precisely when g has the form*

$$g(x)=\psi\bigl(f_1(x),\ldots,f_n(x),\ldots\bigr), \tag{2.12.2}$$

where $f_i\in\mathcal{F}$ and ψ is a Borel function on $\mathbb{R}^\infty$. If $\mathcal{F}$ is a finite family $\{f_1,\ldots,f_n\}$, then for ψ one can take a Borel function on $\mathbb{R}^n$.

PROOF. If g is the indicator of a set E, then our claim follows by the above lemma: writing E in the form (2.12.1) with some $f_i\in\mathcal{F}$ and $B\in\mathcal{B}(\mathbb{R}^\infty)$, we take $\psi=I_B$. If g is a finite linear combination of the indicators of sets $E_1,\ldots,E_k$ with coefficients $c_1,\ldots,c_k$, then the functions $f_i^{(j)}$ involved in the representation of E_j, can be arranged in a single sequence $\{f_i\}$ in such a way that to the functions $f_i^{(j)}$, $j=1,\ldots,k$, there will correspond the subsequences $J_i^{(j)}$. Set $\varphi_j(x_1,x_2,\ldots):=\psi_j(x_{J_1^{(j)}},x_{J_2^{(j)}},\ldots)$. It is clear that φ_j is a Borel function on $\mathbb{R}^\infty$. Then g can be written in the form

$$g=c_1\psi_1(f_{J_1^1},f_{J_2^1},\ldots)+\cdots+c_k\psi_k(f_{J_1^k},f_{J_2^k},\ldots)=\sum_{j=1}^k c_j\varphi_j(f_1,f_2,\ldots).$$

Finally, in the general case, there exists a sequence of simple functions g_k pointwise convergent to g. Let us represent every function g_k in the form (2.12.2) with some functions $f_i^{(k)}\in\mathcal{F}$ and Borel functions ψ_k on $\mathcal{B}(\mathbb{R}^\infty)$. We can arrange the functions $f_i^{(k)}$ in a single sequence $\{f_i\}$. As above, we can write $g_k=\varphi_k(f_1,f_2,\ldots)$, where φ_k is a Borel function on $\mathbb{R}^\infty$ (which is the composition of ψ_k with a projection to certain coordinates). Denote by Ω the set of all $(x_i)\in\mathbb{R}^\infty$ such that $\psi(x):=\lim\limits_{k\to\infty}\varphi_k(x)$ exists and is finite. Then $\Omega\in\mathcal{B}(\mathbb{R}^\infty)$. Letting $\psi=0$ outside Ω, we obtain a Borel function on $\mathbb{R}^\infty$. It remains to observe that $g(x) = \psi\bigl(f_1(x), f_2(x), \ldots\bigr)$. Indeed, for any $x \in X$, the sequence $\varphi_k\bigl(f_1(x), f_2(x), \ldots\bigr)$ converges to $g(x)$. Therefore,

$$\bigl(f_1(x), f_2(x), \ldots\bigr) \in \Omega \quad \text{and} \quad \psi\bigl(f_1(x), f_2(x), \ldots\bigr) = g(x).$$

In the case when the family $\mathcal{F}$ consists of n functions, it suffices to take functions ψ on $\mathbb{R}^n$. □

It is easily seen that the σ-algebra generated by a family of sets coincides with the σ-algebra generated by the indicator functions of those sets.

2.12.4. Example. Let $\{A_n\}$ be a countable collection of subsets of a space X. Then, the σ-algebra generated by $\{A_n\}$ coincides with the σ-algebra generated by the function

$$\psi(x) = \sum_{n=1}^{\infty} 3^{-n} I_{A_n}(x)$$

and is the class of all sets of the form $\psi^{-1}(B)$, $B \in \mathcal{B}(\mathbb{R}^1)$.

PROOF. It is clear that the function ψ is measurable with respect to the σ-algebra $\sigma(\{A_n\})$. Hence the σ-algebra $\sigma(\{\psi\})$ belongs to $\sigma(\{A_n\})$. The inverse inclusion follows from the fact that $I_{A_n} = \theta_n \circ \psi$, where θ_n are Borel functions on $[0,1)$ defined as follows: for any number z with the ternary expansion $z = \sum_{n=1}^{\infty} c_n 3^{-n}$, where $c_n = 0,1,2$, we set $\theta_n(z) := c_n$. For all points z whose ternary expansion is not unique (such points form a countable set) we take for representatives finite sums (for example, the sequence $(0,2,2,2,\ldots)$ is identified with $(1,0,0,0,\ldots)$). It is clear that the step functions θ_n are Borel. □

2.12(ii). Borel mappings on $\mathbb{R}^n$

As in the case of real functions, the mapping $f\colon\ \mathbb{R}^n \to \mathbb{R}^k$ is called Borel if it is $\bigl(\mathcal{B}(\mathbb{R}^n), \mathcal{B}(\mathbb{R}^k)\bigr)$-measurable, i.e., the preimage of any Borel set in $\mathbb{R}^k$ is Borel in $\mathbb{R}^n$. If we write f in the coordinate form $f = (f_1, \ldots, f_k)$, then f is Borel exactly when so are all coordinate functions f_i. This is clear from the following general assertion.

2.12.5. Lemma. *Let $(X, \mathcal{B}_X)$, $(Y_1, \mathcal{B}_1)$,…,$(Y_k, \mathcal{B}_k)$ be measurable spaces and let the space $Y = Y_1 \times \cdots \times Y_k$ be equipped with the σ-algebra $\mathcal{B}_Y$ generated by the sets $B_1 \times \cdots \times B_k$, $B_i \in \mathcal{B}_i$. Then, the mapping $f = (f_1, \ldots, f_k)\colon\ X \to Y$ is $(\mathcal{B}_X, \mathcal{B}_Y)$-measurable precisely when all functions f_i are $(\mathcal{B}_X, \mathcal{B}_i)$-measurable.*

PROOF. If the mapping f is measurable with respect to the indicated σ-algebras, then every component f_i is measurable with respect to $(\mathcal{B}_X, \mathcal{B}_i)$ by the measurability of the projection $(y_1, \ldots, y_k) \mapsto y_i$ with respect to $(\mathcal{B}_Y, \mathcal{B}_i)$, which follows directly from the definition of $\mathcal{B}_Y$. Now suppose that every function f_i is measurable with respect to $(\mathcal{B}_X, \mathcal{B}_i)$. Then $f^{-1}(B_1 \times \cdots \times B_k) = \bigcap_{i=1}^{k} f_i^{-1}(B_i) \in \mathcal{B}_X$ for all $B_i \in \mathcal{B}_i$. The class of all sets $E \subset Y$ with

$f^{-1}(E) \in \mathcal{B}_X$ is a σ-algebra. Since this class contains the products $B_1 \times \cdots \times B_k$, generating $\mathcal{B}_Y$, it contains the whole σ-algebra $\mathcal{B}_Y$. □

It is easily seen that the composition of two Borel mappings is a Borel mapping and that every continuous mapping $f\colon \mathbb{R}^n \to \mathbb{R}^k$ is Borel. Therefore, as already explained in §1.10, for any set $A \in \mathcal{B}(\mathbb{R}^n)$, the set $A \times \mathbb{R}^d$ is Borel in $\mathbb{R}^n \times \mathbb{R}^d$ (as the preimage of A under the natural projection), hence $A \times B \in \mathcal{B}(\mathbb{R}^n \times \mathbb{R}^d)$ whenever $A \in \mathcal{B}(\mathbb{R}^n)$, $B \in \mathcal{B}(\mathbb{R}^d)$.

2.12.6. Proposition. *Let $f\colon \mathbb{R}^n \to \mathbb{R}^k$ be a Borel mapping. Then its graph $\Gamma_f = \bigl\{(x, f(x))\colon\, x \in \mathbb{R}^n\bigr\}$ is a Borel subset of $\mathbb{R}^n \times \mathbb{R}^k$.*

PROOF. It follows by the previous lemma that $(x, y) \mapsto \bigl(f(x), y\bigr)$ from $\mathbb{R}^n \times \mathbb{R}^k$ to $\mathbb{R}^k \times \mathbb{R}^k$ is a Borel mapping. By the continuity of the function $(z, y) \mapsto \|y - z\|$ we conclude that the function $g\colon\ (x, y) \mapsto \|y - f(x)\|$ is Borel. It remains to observe that $\Gamma_f = g^{-1}(0)$. □

2.12.7. Corollary. *Let $f\colon \mathbb{R}^n \to \mathbb{R}^k$ be a Borel mapping and let $B \subset \mathbb{R}^n$ be a Borel set. Then $f(B)$ is a Souslin set. In particular, $f(B)$ is measurable with respect to any Borel measure.*

PROOF. As we proved, the graph of the mapping f is a Borel subset of the space $\mathbb{R}^n \times \mathbb{R}^k$. The projection of this graph to $\mathbb{R}^k$ is $f(B)$. □

We shall see in Chapter 6 that every Souslin set is the continuous image of a Borel set, whence it follows that Corollary 2.12.7 remains valid for any Souslin set B as well.

2.12.8. Corollary. *Let f be a bounded Borel function on $\mathbb{R}^n \times \mathbb{R}^k$. Then, the function $g(x) = \sup_{y \in \mathbb{R}^k} f(x, y)$ is measurable with respect to any Borel measure on $\mathbb{R}^n$.*

PROOF. For every $c \in \mathbb{R}^1$, the set $\bigl\{x \in \mathbb{R}^n\colon\ g(x) > c\bigr\}$ coincides with the projection to $\mathbb{R}^n$ of the Borel set $\bigl\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^k\colon\ f(x, y) > c\bigr\}$. □

We note that the considered function g may not be Borel (see Exercise 6.10.42 in Chapter 6).

2.12(iii). The functional monotone class theorem

The next theorem is a functional version of the monotone class theorem.

2.12.9. Theorem. *Let $\mathcal{H}$ be a class of real functions on a set Ω such that $1 \in \mathcal{H}$ and let $\mathcal{H}_0$ be a subset in $\mathcal{H}$. Then, either of the following conditions yields that $\mathcal{H}$ contains all bounded functions measurable with respect to the σ-algebra $\mathcal{E}$ generated by $\mathcal{H}_0$:*

(i) *$\mathcal{H}$ is a closed linear subspace in the space of all bounded functions on Ω with the norm $\|f\| := \sup_\Omega |f(\omega)|$ such that $\lim\limits_{n\to\infty} f_n \in \mathcal{H}$ for every increasing uniformly bounded sequence of nonnegative functions $f_n \in \mathcal{H}$, and,* *in addition, $\mathcal{H}_0$ is closed with respect to multiplication (i.e., $fg \in \mathcal{H}_0$ for all functions $f,\ g \in \mathcal{H}_0$).*

(ii) *$\mathcal{H}$ is closed with respect to the formation of uniform limits and monotone limits and $\mathcal{H}_0$ is an algebra of functions (i.e., $f+g, cf, fg \in \mathcal{H}_0$ for all $f, g \in \mathcal{H}_0$, $c \in \mathbb{R}^1$) and $1 \in \mathcal{H}_0$.*

(iii) *$\mathcal{H}$ is closed with respect to monotone limits and $\mathcal{H}_0$ is a linear space containing 1 such that $\min(f,g) \in \mathcal{H}_0$ for all $f, g \in \mathcal{H}_0$.*

PROOF. (i) Let us denote by $\mathcal{H}_1$ the linear space generated by 1 and $\mathcal{H}_0$. Condition (i) yields that the class $\mathcal{H}_1$ consists of all functions of the form $c_0 + c_1 h_1 + \cdots + c_n h_n$, $c_i \in \mathbb{R}^1$, $h_i \in \mathcal{H}_0$, and is an algebra of functions, i.e., a linear space closed with respect to multiplication. By Zorn's lemma, there exists a maximal algebra of functions $\mathcal{H}_2$ with $\mathcal{H}_1 \subset \mathcal{H}_2 \subset \mathcal{H}$. It is clear that by the maximality the algebra $\mathcal{H}_2$ is closed with respect to the uniform limits. Then $|f| \in \mathcal{H}_2$ for all $f \in \mathcal{H}_2$, since the function $|f|$ is the uniform limit of a sequence of functions of the form $P_n(f)$, where P_n is a polynomial. Hence $f^+ = \max(f,0) = (f+|f|)/2 \in \mathcal{H}_2$ for all $f \in \mathcal{H}_2$. Similarly, $\min(f,0) \in \mathcal{H}_2$. Therefore, $\mathcal{H}_2$ admits the operations max and min. Finally, we observe that if $\{g_n\}$ is a bounded increasing sequence of nonnegative functions in $\mathcal{H}_2$, then $g = \lim\limits_{n\to\infty} g_n \in \mathcal{H}_2$. Indeed, functions of the form $\sum_{k=0}^n \psi_k g^k$, where $\psi_k \in \mathcal{H}_2$, form an algebra, which we denote by $\mathcal{H}_3$. One has $\mathcal{H}_3 \subset \mathcal{H}$, since $\psi g^k \in \mathcal{H}$ for all $\psi \in \mathcal{H}_2$ and $k \in \mathbb{N}$. Indeed, $\psi^+ g^k$ and $\psi^- g^k$ are monotone limits of the sequences $\psi^+ g_n^k, \psi^- g_n^k \in \mathcal{H}_2$. By the maximality of $\mathcal{H}_2$ we have $\mathcal{H}_3 = \mathcal{H}_2$.

Suppose now that a function f is measurable with respect to $\mathcal{E}$. Since it is the uniform limit of a sequence of $\mathcal{E}$-measurable functions with finitely many values, for the proof of the inclusion $f \in \mathcal{H}$ it suffices to show that $I_A \in \mathcal{H}$ for all $A \in \mathcal{E}$. Let

$$\mathcal{B} = \{B \subset \Omega\colon I_B \in \mathcal{H}_2\}.$$

The class $\mathcal{B}$ is closed with respect to formation of finite intersections and complementation, since $I_{A\cap B} = I_A I_B$ and $1 \in \mathcal{H}_2$. Moreover, $\mathcal{B}$ is a σ-algebra, since $\mathcal{H}_2$ admits monotone limits. Since $\mathcal{E}$ is the σ-algebra generated by the sets $\{\psi > c\}$, where $\psi \in \mathcal{H}_0$ and $c \in \mathbb{R}^1$, it remains to verify that $A = \{\psi > c\} \in \mathcal{B}$. This follows from the fact that I_A is the pointwise limit of the increasing sequence of functions $\psi_n = \min\bigl(1, n(\psi - c)^+\bigr)$. As shown above, $\psi_n \in \mathcal{H}_2$, whence $I_A \in \mathcal{H}_2$.

Assertions (ii) and (iii) are proved similarly with the aid of minor modifications of the above reasoning. □

2.12.10. Example. Let μ and ν be two probability measures on a measurable space $(X, \mathcal{A})$ and let $\mathcal{F}$ be a family of $\mathcal{A}$-measurable functions such that $fg \in \mathcal{F}$ for all $f, g \in \mathcal{F}$ and every function $f \in \mathcal{F}$ has equal integrals with respect to μ and ν. Then, every bounded function measurable with respect to the σ-algebra $\sigma(\mathcal{F})$ generated by $\mathcal{F}$ also has equal integrals with respect to μ and ν. In particular, if $\sigma(\mathcal{F}) = \mathcal{A}$, then $\mu = \nu$.

PROOF. Let $\mathcal{H}$ be the class of all bounded $\mathcal{A}$-measurable functions with equal integrals with respect to μ and ν. Clearly, $\mathcal{H}$ is a linear space that is closed under uniform limits and monotone limits of uniformly bounded sequences (which follows by the standard convergence theorems). Let us set $\mathcal{H}_0 := \mathcal{F}$. Now assertion (i) of the above theorem applies. $\square$

2.12.11. Example. Two Borel probability measures on $\mathbb{R}^n$ coincide provided that they assign equal integrals to every bounded smooth function. Indeed, let $\mathcal{H}_0 = C_b^\infty(\mathbb{R}^n)$ and let $\mathcal{H}$ be the class of all bounded Borel functions with equal integrals with respect to both measures.

2.12(iv). Baire classes of functions

The pointwise limit of a sequence of continuous functions on an interval is a Borel function, but is not necessarily continuous. We know that any Borel function coincides almost everywhere with the pointwise limit of a sequence of continuous functions. Is it possible in this statement to say "everywhere" in place of "almost everywhere"? No, since the pointwise limit of continuous functions must have points of continuity (Exercise 2.12.73). R. Baire [**46**] introduced certain classes of functions that enable one to obtain all Borel functions by consecutive limit operations starting from continuous functions. The zero Baire class B_0 is the class of all continuous functions on $[0,1]$. The Baire classes B_n for $n = 1, 2, \ldots$, are defined inductively: B_n consists of all functions f that do not belong to B_{n-1}, but have the form

$$f(x) = \lim_{j\to\infty} f_j(x), \quad x \in [0,1], \tag{2.12.3}$$

where $f_j \in B_{n-1}$. However, as we shall later see, the classes B_n do not exhaust the collection of all Borel functions. If a function f belongs to no class B_n, but is representable in the form (2.12.3) with some $f_j \in B_{n_j}$, then we write $f \in B_\omega$.

In order to obtain all Borel functions, we have to introduce the Baire classes B_α with transfinite numbers corresponding to countable sets. Namely, by means of transfinite induction, for every ordinal number α (see §1.12(vi)) corresponding to a countable well-ordered set, we denote by B_α the class of all functions f that do not belong to the classes B_β with $\beta < \alpha$, but have the form (2.12.3), where $f_j \in B_{\beta_j}$ and $\beta_j < \alpha$.

In the same manner one defines the Baire classes of functions on an arbitrary metric (or topological) space. We shall need below the Baire classes of functions on the plane.

It is readily verified that if f is a function of some Baire class B_α and φ is a continuous function on the real line, then the function $\varphi \circ f$ is of Baire class α or less. In addition, the uniform limit of a sequence of functions of Baire class α or less also belongs to some Baire class B_β with $\beta \le \alpha$ (see Exercises 2.12.75 and 2.12.76).

2.12.12. Proposition. *The union of all Baire classes B_α coincides with the class of all Borel functions.*

Proof. Let B be the class of all Baire functions. It is clear that the class B is a linear space and is closed with respect to the pointwise limits. Since B contains all continuous functions, it follows by Theorem 2.12.9 that the class B contains all bounded functions that are measurable with respect to the σ-algebra generated by all continuous functions, i.e., B contains all bounded Borel functions. Hence B contains all Borel functions. On the other hand, all functions in all Baire classes are Borel, which follows by transfinite induction and the fact that the class of Borel functions is closed with respect to the pointwise limits. $\square$

For a proof of the following theorem due to Lebesgue, see Natanson [**707**, Ch. XV, §2].

2.12.13. Theorem. *For any ordinal number $\alpha \geq 1$ that is either finite or corresponds to a countable well-ordered set, there exists a function F_α on $[0,1]\times[0,1]$ such that F_α is a function of some Baire class (as a function on the plane) and, for any function f of the class less than α, there exists $t \in [0,1]$ with $f(x) = F_\alpha(x,t)$ for all $x \in [0,1]$.*

2.12.14. Corollary. *All Baire classes B_α are nonempty.*

Proof. If some Baire class B_α is empty, then so are all higher classes, hence any Baire function is of Baire class less than α. Let us take the function F_α from the previous theorem and set $F(x,t) = \max\bigl(F_\alpha(x,t), 0\bigr)$ and

$$\varphi(x,t) = \lim_{n\to\infty} \frac{nF(x,t)}{1+nF(x,t)}.$$

It is clear that the function φ assumes only the values 0 and 1. According to Exercise 2.12.77, the function $\varphi(x,x)$ belongs to some Baire class. Then the function $1-\varphi(x,x)$ also does. Therefore, for some t_0 we have $1-\varphi(x,x) = F_\alpha(x,t_0) = F(x,t_0)$ for all $x \in [0,1]$. This leads to a contradiction: if $\varphi(t_0,t_0) = 0$, then $F(t_0,t_0) = 1$, whence we obtain $\varphi(t_0,t_0) = 1$, and if $\varphi(t_0,t_0) = 1$, then $F(t_0,t_0) = 0$ and hence $\varphi(t_0,t_0) = 0$. $\square$

The Dirichlet function equal to 1 at all rational points and 0 at all irrational points belongs to the second Baire class, but not to the first class (see Exercise 2.12.78); however, it can be made continuous by redefining on a measure zero set. There exist Lebesgue measurable functions on $[0,1]$ that cannot be made functions in the first Baire class by redefining on a set of measure zero (Exercise 2.12.79). Vitali proved (see [**984**]) that the situation is different for the second class.

2.12.15. Example. Every Lebesgue measurable function f on the interval $[0,1]$ coincides almost everywhere with a function g that belongs to one of the Baire classes B_0, B_1, B_2.

PROOF. Passing to the function $\operatorname{arctg} f$ and applying Exercise 2.12.76, we may assume that the function f is bounded. There exists a sequence of continuous functions f_n such that $f(x) = \lim_{n\to\infty} f_n(x)$ for almost all x. It is clear that one can choose a uniformly bounded sequence with such a property. For any fixed n, the functions $f_{n,k} = \max(f_n, \ldots, f_{n+k})$ are continuous, uniformly bounded and $f_{n,k} \le f_{n,k+1}$. Hence the functions $g_n(x) = \lim_{k\to\infty} f_{n,k}(x)$ belong to the zero or first Baire class. These functions are uniformly bounded and $g_{n+1} \le g_n$. Therefore, the function $g(x) = \lim_{n\to\infty} g_n(x)$ is of Baire class 2 or less. It is clear that $g(x)$ coincides with the limit of $f_n(x)$ everywhere, where this limit exists, i.e., almost everywhere. Thus, $g = f$ a.e. $\square$

2.12(v). Mean value theorems

It is known from the elementary calculus that the integral of a continuous function over a compact interval equals the product of the interval length and some value of the function on that interval. Here we discuss analogous assertions for the Lebesgue integral. If a function f is Lebesgue integrable on $[a, b]$ and $m \le f \le M$, then the integral of f lies between $m(b - a)$ and $M(b-a)$ and hence equals $c(b-a)$ for some $c \in [m, M]$. But c may not belong to the range of f. For this reason, the following assertion is usually called the first mean value theorem for the Lebesgue integral.

2.12.16. Theorem. *If a function $f \ge 0$ is integrable on $[a, b]$ and a function g is continuous, then there exists $\xi \in [a, b]$ such that*

$$\int_a^b f(t)g(t)\,dt = g(\xi)\int_a^b f(t)\,dt.$$

PROOF. Let I be the integral of f over $[a, b]$. Then the integral of fg lies between $I \min g$ and $I \max g$. $\square$

The next useful result is often called *the second mean value theorem.*

2.12.17. Theorem. *Suppose that a function f is integrable on (a, b) and a function φ is bounded on (a, b) and increasing. Then, there exists a point $\xi \in [a, b]$ such that*

$$\int_a^b \varphi(x)f(x)\,dx = \varphi(a+0)\int_a^\xi f(x)\,dx + \varphi(b-0)\int_\xi^b f(x)\,dx, \qquad (2.12.4)$$

where $\varphi(a + 0)$ and $\varphi(b - 0)$ denote the right and left limits, respectively. If, in addition, φ is nonnegative, then there exists a point $\eta \in [a, b]$ such that

$$\int_a^b \varphi(x)f(x)\,dx = \varphi(b-0)\int_\eta^b f(x)\,dx. \qquad (2.12.5)$$

PROOF. Suppose first that φ and f are continuously differentiable functions on $[a, b]$. Set

$$F(x) = \int_a^x f(t)\,dt.$$

By the Newton–Leibniz formula we have

$$\int_a^b \varphi(x)f(x)\,dx = \varphi(b)F(b) - \varphi(a)F(a) - \int_a^b \varphi'(x)F(x)\,dx. \qquad (2.12.6)$$

Since $\varphi' \geq 0$, we obtain

$$\bigl[\min_x F(x)\bigr] \int_a^b \varphi'(x)\,dx \leq \int_a^b \varphi'(x)F(x)\,dx \leq \bigl[\max_x F(x)\bigr] \int_a^b \varphi'(x)\,dx.$$

By the mean value theorem there exists a point $\xi \in [a, b]$ such that

$$\int_a^b \varphi'(x)F(x)\,dx = F(\xi) \int_a^b \varphi'(x)\,dx = F(\xi)[\varphi(b) - \varphi(a)].$$

Substituting this equality in (2.12.6), we arrive at (2.12.4).

In the general case, we can find two sequences of continuously differentiable functions f_n and φ_n on $[a, b]$ such that the functions f_n converge to f in the mean, the functions φ_n are nondecreasing, $\sup_{n,x} |\varphi_n(x)| < \infty$ and $\varphi_n(x) \to \varphi(x)$ at all points of continuity of φ. For φ_n one can take

$$\varphi_n(x) := \int_0^1 \varphi(x - n^{-1}y)p(y)\,dy,$$

where p is a nonnegative smooth function vanishing outside $[0, 1]$ and having the integral 1, where we set $\varphi(x) = \varphi(a + 0)$ if $x \leq a$. It is clear that $\varphi_n(a) = \varphi(a + 0)$, $|\varphi_n(x)| \leq \sup_t |\varphi(t)|$, and the functions φ_n are increasing and continuously differentiable. The latter follows from the equality

$$\varphi_n(x) = n \int_{a-1}^b \varphi(z)p(nx - nz)\,dz,$$

which is obtained by changing variables, and the theorem on differentiation of the Lebesgue integral with respect to a parameter. By the dominated convergence theorem we obtain that $\varphi_n(x) \to \varphi(x)$ at all points x where φ is left continuous, in particular, $\varphi_n(b) \to \varphi(b - 0)$. Since the set of points of discontinuity of φ is at most countable, one has $\varphi_n(x) \to \varphi(x)$ almost everywhere. Hence the integrals of $\varphi_n f_n$ converge to the integral of φf. Let $\xi_n \in [a, b]$ be certain points corresponding to φ_n and f_n in (2.12.4). The sequence ξ_n has a limit point $\xi \in [a, b]$. Passing to a subsequence we may assume that $\xi_n \to \xi$. In order to see that ξ is a required point, it remains to observe that

$$\int_a^{\xi_n} f_n(x)\,dx - \int_a^{\xi} f(x)\,dx = \int_a^{\xi_n} [f_n(x) - f(x)]\,dx + \int_{\xi}^{\xi_n} f(x)\,dx \to 0,$$

since $f_n \to f$ in the mean on $[a, b]$ and the integrals of the function $|f|$ over intervals of length $|\xi_n - \xi|$ tend to zero as $n \to \infty$ by the absolute continuity of the Lebesgue integral.

In the case where $\varphi \geq 0$, it suffices to show that the right-hand side of (2.12.4) belongs to the closed interval formed by the values of the continuous function

$$\Psi(x) = \varphi(b-0)\int_x^b f(t)\,dt$$

on $[a,b]$. For example, if the integral of f over $[a,\xi]$ is nonnegative, then

$$\begin{aligned}\varphi(b-0)\int_\xi^b f(x)\,dx &\le \varphi(a+0)\int_a^\xi f(x)\,dx + \varphi(b-0)\int_\xi^b f(x)\,dx\\ &\le \varphi(b-0)\int_a^b f(x)\,dx,\end{aligned}$$

whence the claim follows. □

2.12(vi). The Lebesgue–Stieltjes integral

In Chapter 1, we considered Lebesgue–Stieltjes measures on the real line: to every left continuous increasing function F having the limit 0 at $-\infty$ and the limit 1 at $+\infty$, a Borel probability measure μ with $F(t) = \mu\bigl((-\infty,t)\bigr)$ was associated. Let g be a μ-integrable function.

2.12.18. Definition. *The quantity*

$$\int_{-\infty}^{\infty} g(t)\,dF(t) := \int_{\mathbb{R}} g(t)\,\mu(dt) \tag{2.12.7}$$

is called the Lebesgue–Stieltjes integral of the function f with respect to the function F.

This definition can be easily extended to all functions F of the form $F = c_1F_1 + c_2F_2$, where F_1, F_2 are the distribution functions of probability measures μ_1 and μ_2 and c_1, c_2 are constant numbers. Then, one takes for μ the measure $c_1\mu_1 + c_2\mu_2$ (signed measures are discussed in Chapter 3). One defines similarly the Lebesgue–Stieltjes integral over closed or open intervals. In certain applications, one is given the distribution function F, and not the measure μ directly, and for this reason the notation for the integral by means of the left-hand side of (2.12.7) is convenient and helpful in calculations. If g assumes finitely many values c_i on intervals $[a_i, b_i)$ and vanishes outside those intervals, then

$$\int g(t)\,dF(t) = \sum_{i=1}^n c_i[F(b_i) - F(a_i)].$$

For continuous functions g on $[a,b]$, the Lebesgue–Stieltjes integral can be expressed as a limit of sums of the Riemannian type. It should be noted that one can develop in this spirit the Riemann–Stieltjes integral, but we shall not do this. In Exercise 5.8.112 in Chapter 5 one can find the integration by parts formula for the Lebesgue–Stieltjes integral.

2.12(vii). Integral inequalities

In the theory of measure and integral and its applications, an important role is played by various integral inequalities. For example, we have already encountered the Chebyshev inequality and the Hölder and Minkowski inequalities. In this subsection we derive several other frequently used inequalities. The first of them is Jensen's inequality.

We recall that a real function Ψ defined on an interval $\mathrm{Dom}(\Psi) = (a, b)$ (possibly unbounded) is called convex if

$$\Psi\bigl(tx + (1-t)y\bigr) \le t\Psi(x) + (1-t)\Psi(y), \quad \forall\, x, y \in \mathrm{Dom}(\Psi), \forall\, t \in [0,1].$$

If Ψ is bounded in a one-sided neighborhood of a finite boundary point a or b, then such a point is included in $\mathrm{Dom}(\Psi)$ and the value at this point is defined by continuity. The following sufficient condition for convexity is frequently used in practice: Ψ is twice differentiable and $\Psi'' \ge 0$. The proof reduces to the case $x = 0$, $y = 1$. Passing to $\Psi(x) - x\Psi(1) - (1-x)\Psi(0)$ we reduce everything to the case $\Psi(0) = \Psi(1) = 0$. Now we have to verify that $\Psi \le 0$. If this is not so, there exists a point of maximum $\xi \in (0,1)$ with $\Psi(\xi) > 0$. Then $\Psi'(\xi) = 0$, whence $\Psi'(t) \ge 0$ for $t \ge \xi$ due to $\Psi'' \ge 0$. Hence $\Psi(1) \ge \Psi(\xi) > 0$, a contradiction.

Here are typical examples of convex functions: e^x, $|x|^\alpha$ with $\alpha \ge 1$.

We observe that for any point $x_0 \in \mathrm{Dom}(\Psi)$, there exists a number $\lambda(x_0)$ such that

$$\Psi(x) \ge \Psi(x_0) + \lambda(x_0)(x - x_0), \quad \forall\, x \in \mathrm{Dom}(\Psi). \tag{2.12.8}$$

For $\lambda(x_0)$ one can take any number between the lower derivative

$$\Psi'_-(x_0) = \liminf_{h\to 0} h^{-1}\bigl(\Psi(x_0 + h) - \Psi(x_0)\bigr)$$

and the upper derivative

$$\Psi'_+(x_0) = \limsup_{h\to 0} h^{-1}\bigl(\Psi(x_0 + h) - \Psi(x_0)\bigr)$$

(see Exercise 2.12.88).

By using this property of convex functions (which can be taken as a definition) one obtains the following Jensen inequality.

2.12.19. Theorem. *Suppose that μ is a probability measure on a space $(X, \mathcal{A})$. Let f be a μ-integrable function with values in the domain of definition of a convex function Ψ such that the function $\Psi(f)$ is integrable. Then one has*

$$\Psi\biggl(\int_X f(x)\,\mu(dx)\biggr) \le \int_X \Psi\bigl(f(x)\bigr)\,\mu(dx). \tag{2.12.9}$$

PROOF. Let x_0 be the integral of f. It is readily verified that x_0 belongs to $\mathrm{Dom}(\Psi)$. Substituting $f(x)$ in place of x in (2.12.8) we obtain

$$\Psi\bigl(f(x)\bigr) \ge \Psi(x_0) + \lambda(x_0)\bigl[f(x) - x_0\bigr].$$

Let us integrate this equality and observe that the integral of the second summand on the right is zero. Hence we arrive at (2.12.9). □

A number of useful inequalities can be obtained by choosing concrete functions Ψ in the general Jensen inequality.

2.12.20. Corollary. *Let μ be a probability measure on a measurable space $(X,\mathcal{A})$. Let f be a μ-integrable function such that the function $\exp f$ is integrable. Then*

$$\exp\Bigl(\int_X f(x)\,\mu(dx)\Bigr) \le \int_X \exp f(x)\,\mu(dx). \tag{2.12.10}$$

Letting $\Psi(t) = |t|^\alpha$ with $\alpha > 1$, we obtain the following Lyapunov inequality (which also follows by Hölder's inequality).

2.12.21. Corollary. *Let μ be a probability measure on a measurable space $(X,\mathcal{A})$. Let f be a function such that the function $|f|^p$ is integrable for some $p \ge 1$. Then, for any $r \in (0,p]$, the function $|f|^r$ is integrable and*

$$\Bigl(\int_X |f|^r\,d\mu\Bigr)^{1/r} \le \Bigl(\int_X |f|^p\,d\mu\Bigr)^{1/p}. \tag{2.12.11}$$

In the case of a general measure space, a similar estimate is available.

2.12.22. Corollary. *Let μ be a nonnegative measure (possibly with values in $[0,+\infty]$) on a measurable space $(X,\mathcal{A})$. Let f be a function such that the function $|f|^p$ is integrable for some $p \ge 1$ and $\mu\bigl(x\colon\ f(x) \ne 0\bigr) < \infty$. Then, for any $r \in (0,p]$, the function $|f|^r$ is integrable and*

$$\Bigl(\int_X |f|^r\,d\mu\Bigr)^{1/r} \le \mu\Bigl(x\colon\ f(x)\ne 0\Bigr)^{1/r-1/p}\Bigl(\int_X |f|^p\,d\mu\Bigr)^{1/p}. \tag{2.12.12}$$

For the proof we set $\Omega := \{f \ne 0\}$ and take the probability measure $\mu(\Omega)^{-1}\mu|_\Omega$. Note that (2.12.12) is better than (2.12.11) if $0 < \mu(\Omega) < 1$.

The next two integral inequalities are employed in information theory and probability theory (see Liese, Vajda [**613**]).

2.12.23. Theorem. *Let f and g be positive integrable functions on a space X with a nonnegative measure μ. Then*

$$\begin{aligned}\int_X f\ln f\,d\mu - \int_X f\,d\mu\Bigl(\ln\int_X f\,d\mu\Bigr) \qquad\qquad\qquad\qquad & \\ \ge \int_X f\ln g\,d\mu - \int_X f\,d\mu\Bigl(\ln\int_X g\,d\mu\Bigr),&\end{aligned} \tag{2.12.13}$$

provided that $f\ln f$ and $f\ln g$ are integrable. In addition, the equality is only possible when $f = cg$ a.e. for some number c.

Proof. Suppose first that f and g have equal integrals. The inequality $\ln x \le x-1$ on $(0,\infty)$ yields the estimate $f\ln g - f\ln f = f\ln(g/f) \le g-f$ (it suffices to take $x=g/f$). By integrating we obtain the inequality

$$\int_X f\ln f\,d\mu \ge \int_X f\ln g\,d\mu.$$

It is clear that the equality is only possible when one has $f\ln(g/f)=g-f$ a.e., which is equivalent to $f=g$ a.e. In the general case, writing the last inequality for the functions $f\|f\|^{-1}_{L^1(\mu)}$ and $g\|g\|^{-1}_{L^1(\mu)}$ with equal integrals and using that the integral of $f/\|f\|_{L^1(\mu)}$ is 1, we arrive at (2.12.13). □

The quantity

$$\int f\ln f\,d\mu$$

is called the entropy of f. The following estimate is named the Pinsker–Kullback–Csiszár inequality (Pinsker [**757**] obtained it with some constant and then Csiszár and Kullback justified it in the form presented below, see Csiszár [**194**]).

2.12.24. Theorem. *Let μ and ν be two probability measures on a measurable space $(X,\mathcal{A})$ and let $\nu = f\cdot\mu$, where $f>0$. Then*

$$\|\mu-\nu\|^2 := \biggl(\int_X |f-1|\,d\mu\biggr)^2 \le 2\int_X f\ln f\,d\mu,$$

where the infinite value is allowed on the right-hand side.

Proof. Let $E:=\{f\le 1\}$, $\nu(E)=a$, $t=\mu(E)$. It is clear that $a\le t$. One has

$$\int_X |f-1|\,d\mu = \int_E (1-f)\,d\mu + \int_{X\backslash E}(f-1)\,d\mu = 2(t-a).$$

If $a=1$ or $t=1$, then $f=1$ a.e. So we assume further that $a,t\in(0,1)$. In addition, we assume that the function $f\ln f$ is integrable because otherwise one has $+\infty$ on the right due to the boundedness of the function $f\ln f$ on the set E. Applying inequality (2.12.13) to the probability density g that equals a/t on E and $(1-a)/(1-t)$ on $X\backslash E$, we obtain

$$\int_X f\ln f\,d\mu \ge a\ln\frac{a}{t} + (1-a)\ln\frac{1-a}{1-t}.$$

Now it suffices to observe that for all $a\le t\le 1$ one has the inequality

$$\psi_a(t) := 2(t-a)^2 - a\ln\frac{a}{t} - (1-a)\ln\frac{1-a}{1-t} \le 0,$$

which follows by the relation $\psi_a'(t) = (a-t)\bigl(4-t^{-1}(1-t)^{-1}\bigr)\le 0$ for all $t\ge a$ and the equality $\psi_a(a)=0$. □

Several other important integral inequalities will be obtained in §3.10.

Exercises

2.12.25.° Let $(X, \mathcal{A})$, $(Y, \mathcal{B})$, and $(Z, \mathcal{E})$ be measurable spaces. Suppose that a mapping $f\colon X \to Y$ is $(\mathcal{A}, \mathcal{B})$-measurable and a mapping $g\colon Y \to Z$ is $(\mathcal{B}, \mathcal{E})$-measurable. Show that the composition $g \circ f\colon X \to Z$ is $(\mathcal{A}, \mathcal{E})$-measurable.

2.12.26.° Suppose that measurable functions f_n on $[0,1]$ converge almost everywhere to zero. Show that there exist numbers $C_n > 0$ such that $\lim\limits_{n\to\infty} C_n = \infty$, but the sequence $C_n f_n$ converges almost everywhere to zero.

2.12.27.° Suppose that measurable functions f_n on $[0,1]$ converge almost everywhere to zero. Prove that there exist numbers $\varepsilon_n > 0$ and a measurable finite function g such that $\lim\limits_{n\to\infty} \varepsilon_n = 0$ and $|f_n(x)| \le \varepsilon_n g(x)$ almost everywhere for every n.

HINT: in Exercise 2.12.26 take $\varepsilon_n = C_n^{-1}$.

2.12.28.° Construct a measurable set in $[0,1]$ such that every function on $[0,1]$ that almost everywhere equals its indicator function is discontinuous almost everywhere (and is not Riemann integrable, see Exercise 2.12.38).

HINT: take a set such that the intersections of this set and its complement with every interval have positive measures.

2.12.29.° Suppose that functions f and g are measurable with respect to a σ-algebra $\mathcal{A}$ and that a function Ψ on the plane is continuous on the set of values of the mapping (f,g). Show that the function $\Psi(f,g)$ is measurable with respect to $\mathcal{A}$.

HINT: letting Y be the range of (f,g), use that the sets $\{\Psi < c\}$ are open in Y, i.e., $\{\Psi < c\} = Y \cap U$, where U is open in the plane.

2.12.30.° Let $\mathcal{A}$ be the σ-algebra generated by all singletons in a space X. Prove that a function f is measurable with respect to $\mathcal{A}$ if and only if it is constant on the complement of some at most countable set.

HINT: the indicated condition is sufficient for the $\mathcal{A}$-measurability of f, since all at most countable sets belong to $\mathcal{A}$. The converse follows by the fact that the above condition is fulfilled for all simple functions.

2.12.31.° Let μ be a probability measure, let $\{c_\alpha\}$ be a family of real numbers, and let f be a μ-measurable function. Show that

$$\mu\bigl(x\colon f(x) \ge \sup_\alpha c_\alpha\bigr) \ge \inf_\alpha \mu\bigl(x\colon f(x) \ge c_\alpha\bigr).$$

HINT: let $r = \inf_\alpha \mu\bigl(x\colon f(x) \ge c_\alpha\bigr)$ and let α_n be such that the numbers c_{α_n} are increasing to $\sup_\alpha c_\alpha$. One has $\mu\bigl(x\colon f(x) \ge c_{\alpha_n}\bigr) \ge r$ for all n, whence the claim follows by the σ-additivity of μ.

2.12.32. (Davies [**207**]) Let μ be a finite nonnegative measure on a space X. Prove that a function $f\colon X \to \mathbb{R}^1$ is measurable with respect to μ precisely when for each μ-measurable set A with $\mu(A) > 0$ and each $\varepsilon > 0$, there exists a μ-measurable set $B \subset A$ such that $\mu(B) > 0$ and $\sup\limits_{x,y\in B} |f(x) - f(y)| \le \varepsilon$.

HINT: the necessity of this condition is clear from the fact that the set A is covered by the sets $\bigl\{x \in A\colon n\varepsilon \le f(x) < (n+1)\varepsilon\bigr\}$. For the proof of sufficiency, one can construct a sequence of μ-measurable functions f_n with countably many values, uniformly convergent to f on a set of full measure. To this end, for fixed $\varepsilon > 0$ and any set E of positive measure, we consider the class $\mathcal{B}(E,\varepsilon)$ of all measurable sets $B \subset E$ with $\mu(B) > 0$ and $\sup\limits_{x,y\in B} |f(x) - f(y)| \leq \varepsilon$ (this class is nonempty by our hypothesis) and put $\delta_1 = \sup\{\mu(B)\colon B \in \mathcal{B}(X,\varepsilon)\}$; we choose $B_1 \in \mathcal{B}(X,\varepsilon)$ with $\mu(B_1) > \delta_1/2$. Let us repeat the described construction for the set $X\backslash B_1$ and find $B_2 \subset X\backslash B_1$ with $\mu(B_2) > \delta_2/2$, where $\delta_2 = \sup\{\mu(B), B \in \mathcal{B}(X\backslash B_1,\varepsilon)\}$. By induction, we obtain μ-measurable sets B_n with $B_n \subset X\backslash(B_1 \cup \cdots \cup B_{n-1})$, $\mu(B_n) > \delta_n/2$, $\delta_n = \sup\{\mu(B), B \in \mathcal{B}(X\backslash(B_1 \cup \cdots \cup B_{n-1}),\varepsilon)\}$. This process will be finite only in the case if X is covered by finitely many sets B_n up to a measure zero set. In the general case we obtain a sequence of sets B_n covering X up to a set of measure zero. Indeed, otherwise there exists a set $E \subset X\backslash\bigcup_{n=1}^\infty B_n$ such that $\mu(E) = \delta > 0$ and $\sup\limits_{x,y\in E} |f(x) - f(y)| \leq \varepsilon$. It is clear that $\delta_n \to 0$ and hence there exists $\delta_k < \delta/2$. This leads to a contradiction, since $E \subset X\backslash\bigcup_{n=1}^k B_n$, whence $\mu(E) \leq \delta_k < \delta$. It remains to choose a point x_n in every set B_n and put $g|_{B_n} = f(x_n)$. Then $|g(x) - f(x)| \leq \varepsilon$ for all $x \in \bigcup_{n=1}^\infty B_n$.

2.12.33.° (M. Fréchet) Suppose that a sequence of measurable functions f_n on a probability space $(X,\mathcal{A},\mu)$ converges a.e. to a function f and, for every n, there is a sequence of measurable functions $f_{n,m}$ a.e. convergent to f_n. Prove that there exist subsequences $\{n_k\}$ and $\{m_k\}$ such that $f_{n_k,m_k} \to f$ a.e.

Hint: use Remark 2.2.7 (or the metrizability of convergence in measure) and the Riesz theorem.

2.12.34.° Investigate for which real α and β the function $x^\alpha \sin(x^\beta)$ is Lebesgue integrable on (a) $(0,1)$, (b) $(0,+\infty)$, (c) $(1,+\infty)$. Answer the same question for the proper and improper Riemann integrability.

2.12.35.° (Alekhno, Zabreĭko [**8**]) Let μ be a finite nonnegative measure on a measurable space $(X,\mathcal{A})$ and let $\{f_n\}$ be a sequence of μ-measurable functions. Suppose that it is not true that this sequence converges to zero μ-a.e. Prove that there exist a subsequence $\{f_{n_k}\}$ in $\{f_n\}$, measurable sets A_k with $\mu(A_k) > 0$ and $A_{k+1} \subset A_k$ for all k, and $\varepsilon > 0$ such that $|f_{n_k}(x)| \geq \varepsilon$ for all $x \in A_k$ and all k.

Hint: let $g_m(x) := \sup_{n\geq m} |f_n(x)|$. Since the sequence $\{g_m\}$ decreases and does not converge to zero on some positive measure set, it is readily seen that there exists $\varepsilon > 0$ such that the set $E := \bigcap_{m\geq 1}\{x\colon g_m(x) > \varepsilon\}$ has positive measure. Letting $E_n := \{x \in E\colon |f_n(x)| > \varepsilon\}$, we find n_1 such that $\mu(E_{n_1}) > 0$, then we find $n_2 > n_1$ such that $\mu(E_{n_1} \cap E_{n_2}) > 0$ and so on. Finally, let $A_k := E_{n_1} \cap \cdots \cap E_{n_k}$.

2.12.36.° Investigate for which real α and β the function $x^\alpha(\ln x)^\beta$ is Lebesgue integrable on (a) $(0,1)$, (b) $(0,+\infty)$.

2.12.37.° Let J_n be a sequence of disjoint intervals in $[0,1]$, convergent to the origin, $|J_n| = 4^{-n}$, and let $f = n^{-1}/|J_{2n}|$ on J_{2n}, $f = -n^{-1}/|J_{2n+1}|$ on J_{2n+1}, and let f be zero at all other points. Show that f is Riemann integrable in the improper sense, but is not Lebesgue integrable.

2.12.38. (i) (H. Lebesgue, G. Vitali) Show that a bounded function is Riemann integrable on an interval (or a cube) precisely when the set of its discontinuity points has measure zero.

(ii) Prove that a function f on $[a,b]$ is Riemann integrable precisely when, for each $\varepsilon > 0$, there exist step functions g and h such that $|f(x) - g(x)| \le h(x)$ and

$$\int_a^b h(x)\,dx \le \varepsilon.$$

HINT: (i) see Rudin [**834**, Theorem 10.33], Zorich [**1053**, Ch. XI, §1]; (ii) apply (i) and the Chebyshev inequality.

2.12.39.° Suppose that a sequence of μ-integrable functions f_n converges to f in $L^1(\mu)$ and a sequence of μ-measurable functions φ_n converges to φ μ-a.e. and is uniformly bounded. Show that the functions $\varphi_n f_n$ converge to φf in $L^1(\mu)$.

HINT: observe that the assertion reduces to the case of a bounded measure and use the uniform integrability of $\{f_n\}$.

2.12.40.° Let a function $f \ge 0$ be integrable with respect to a measure μ. Prove the equality

$$\int f\,d\mu = \lim_{r\downarrow 1} \sum_{n=-\infty}^{\infty} r^n \mu\bigl(x\colon\ r^n \le f(x) < r^{n+1}\bigr).$$

HINT: let $f_r = \sum_{n=-\infty}^{\infty} r^n I_{f^{-1}[r^n, r^{n+1})}$, then one has $f_r \le f \le r f_r$.

2.12.41.° (i) Construct a sequence of nonnegative functions f_n on $[0,1]$ convergent to zero pointwise such that their integrals tend to zero, but the function $\Phi(x) = \sup_n f_n(x)$ is not integrable. In particular, the functions f_n have no common integrable majorant.

(ii) Construct a sequence of functions $f_n \ge 0$ on $[0,1]$ such that their integrals tend to zero, but $\sup_n f_n(x) = +\infty$ for every x.

HINT: (i) take $f_n(x) = nI_{[(n+1)^{-1}, n^{-1}]}$, $x \in [0,1]$; (ii) take the functions $f_{n,k}$ from Example 2.2.4 and consider $nf_{n,k}$.

2.12.42. Let μ be a probability measure on a space X and let $\{f_n\}$ be a sequence of μ-integrable functions that converges μ-a.e. to a μ-integrable function f such that the integrals of f_n converge to the integral of f. Prove that for any $\varepsilon > 0$ there exist a measurable set E and a number $N \in \mathbb{N}$ such that for all $n \ge N$ one has

$$\Bigl|\int_{X\backslash E} f_n\,d\mu\Bigr| \le \varepsilon \quad\text{and}\quad |f_n(x)| \le |f(x)| + 1 \quad \text{for } x \in E.$$

HINT: there exists $\delta > 0$ such that

$$\int_A |f|\,d\mu < \varepsilon/3$$

whenever $\mu(A) < \delta$. There is a set E such that $\mu(X\backslash E) < \delta$ and convergence of f_n to f is uniform on E. Let us now take N such that for all $n \ge N$ one has

$$\sup_{x\in E} |f_n(x) - f(x)| \le \frac{1}{3}\min(1,\varepsilon) \quad\text{and}\quad \Bigl|\int_X (f_n - f)\,d\mu\Bigr| \le \frac{\varepsilon}{3}.$$

Then

$$\begin{aligned}\Bigl|\int_{X\backslash E} f_n\,d\mu\Bigr| &= \Bigl|\int_X f_n\,d\mu - \int_X f\,d\mu + \int_{X\backslash E} f\,d\mu + \int_E (f - f_n)\,d\mu\Bigr| \\ &\le \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{1}{3}\min(1,\varepsilon) \le \varepsilon.\end{aligned}$$

2.12.43. Let μ be a probability measure on a space X and let f_n be μ-measurable functions. Prove that the following conditions are equivalent:

(i) there exists a subsequence f_{n_k} convergent a.e. to 0;

(ii) there exists a sequence of numbers t_n such that

$$\limsup_{n\to\infty} |t_n| > 0 \quad \text{and} \quad \sum_{n=1}^{\infty} t_n f_n(x) \quad \text{converges a.e.};$$

(iii) there exists a sequence of numbers t_n such that

$$\sum_{n=1}^{\infty} |t_n| = \infty \quad \text{and} \quad \sum_{n=1}^{\infty} |t_n f_n(x)| < \infty \quad \text{a.e.}$$

HINT: by Egoroff's theorem (i) yields (ii), (iii). If (iii) is true, then for the set $X_N := \Bigl\{x \in X\colon \ \sum_{n=1}^{\infty} |t_n f_n(x)| \le N\Bigr\}$ we have

$$\sum_{n=1}^{\infty} |t_n| \int_{X_N} |f_n|\, d\mu \le N,$$

whence it follows that

$$\liminf_{n\to\infty} \int_{X_N} |f_n|\, d\mu = 0.$$

This yields (i), since $\mu(X_N) \to 1$, Finally, (ii) implies (i), since $t_n f_n(x) \to 0$ a.e. and it suffices to take n_k with $\liminf_{k\to\infty} |t_{n_k}| > 0$.

2.12.44.° Show that a sequence of measurable functions f_n on a space with a probability measure μ converges almost uniformly (in the sense of Egoroff's theorem) to a measurable function f precisely when

$$\lim_{n\to\infty} \mu\Bigl(\bigcup_{m\ge n} \{x\colon \ |f_m(x) - f_n(x)| \ge \varepsilon\}\Bigr) = 0.$$

2.12.45. Prove the following analog of Egoroff's theorem for spaces with infinite measure: let μ-measurable functions f_n converge μ-a.e. to a function f such that $|f_n| \le g$ μ-a.e., where the function g is integrable with respect to μ; then, for any $\varepsilon > 0$, there exists a set A_ε such that the functions f_n converge to f uniformly on A_ε, and the complement of A_ε has μ-measure less than ε.

HINT: the sets $G := \{g > 1\}$ and $G_k := \bigl\{2^{-k} < g \le 2^{1-k}\bigr\}$ have finite measures by the integrability of g, hence they contain measurable subsets $A \subset G$ and $A_k \subset G_k$ on each of which convergence is uniform and $\mu(G\backslash A) < \varepsilon/2$, $\mu(G_k\backslash A_k) < \varepsilon 4^{-k}$. For A_ε one can take the union of all sets A and A_k with the set of all points x where $f_n(x) = 0$ for all n.

2.12.46. (Tolstoff [**950**]) (i) Let f be a Borel function on $[0,1]^2$, y_0 a fixed point in $[0,1]$ and $\lim_{y\to y_0} f(x,y) = f(x,y_0)$ for any $x \in [0,1]$. Prove that for every $\varepsilon > 0$ there exists a measurable set $A_\varepsilon \subset [0,1]$ of Lebesgue measure $\lambda(A_\varepsilon) > 1 - \varepsilon$ such that $\lim_{y\to y_0} f(x,y) = f(x,y_0)$ uniformly in $x \in A_\varepsilon$.

(ii) Construct a bounded Lebesgue measurable function f on $[0,1]^2$ such that it is Borel in every variable separately and $\lim_{y\to 0} f(x,y) = 0$ for any $x \in [0,1]$, but on no set of positive measure is convergence uniform.

HINT: (i) Let

$$\delta_n(x) := \sup\bigl\{\delta\colon\ |f(x,y) - f(x,y_0)| < 1/n \ \text{ if } |y - y_0| < \delta\bigr\}.$$

By hypothesis, $\delta_n(x) > 0$ for every x. It is readily seen that for fixed $n \in \mathbb{N}$ and $C \geq 0$ the set

$$M(n,C) := \bigl\{(x,y)\colon\ |f(x,y) - f(x,y_0)| \geq 1/n, |y - y_0| < C\bigr\}$$

is Borel. By Proposition 1.10.8 the projection of $M(n,C)$ to the first coordinate axis is a Souslin set and hence is measurable. It is easily verified that this projection is $\{x\colon\ \delta_n(x) < C\}$, which yields the measurability of the function δ_n. Now, given $\varepsilon > 0$, for every n we find a measurable set $A_n \subset [0,1]$ such that $\lambda(A_n) > 1 - \varepsilon 2^{-n}$ and $\delta_n|_{A_n} \geq \gamma_n$, where $\gamma_n > 0$ is some constant. Let $A = \bigcap_{n=1}^{\infty} A_n$. If $n^{-1} < \varepsilon$ and $|y - y_0| < \gamma_n$, then $|f(x,y) - f(x,y_0)| < n^{-1} < \varepsilon$ for all $x \in A \subset A_n$. (ii) There is a partition of $[0,1]$ into disjoint sets E_n with $\lambda^*(E_n) = 1$. Let $f(x, n^{-1}x) = 1$ if $x \in E_n$, $n \in \mathbb{N}$, at all other points let $f = 0$. The function f differs from zero only at the points of a set covered by countably many straight lines of the form $y = nx$. It is clear that f is Lebesgue measurable and Borel in every variable separately. If $\lambda(E) > 0$, then, for any n, E contains points from E_n, hence, for each $\varepsilon > 0$, there exist $x \in E$ and $y < \varepsilon$ with $f(x,y) = 1$.

2.12.47. (Frumkin [**330**]) Let f be a function on $[0,1]^2$ such that, for every fixed t, the function $s \mapsto f(t,s)$ is finite a.e. and measurable. Suppose that $\lim\limits_{t\to 0} f(t,s) = f(0,s)$ for a.e. s. Show that, for each $\delta_1 > 0$, there exists a measurable set $E_{\delta_1} \subset [0,1]$ with the following property: $\lambda(E_{\delta_1}) > 1 - \delta_1$ and, given $\varepsilon > 0$, one can find $\delta > 0$ such that whenever $t < \delta$, the inequality $|f(t,s) - f(0,s)| < \varepsilon$ holds for all s, with the exception of points of some set E_t of measure zero.

2.12.48. (Stampacchia [**904**]) Suppose we are given a sequence of functions f_n on $[0,1]\times[0,1]$ measurable in x and continuous in y. Assume that for every $y \in [0,1]$ the sequence $\{f_n(x,y)\}$ converges for a.e. x and that for a.e. x the sequence of functions $y \mapsto f_n(x,y)$ is equicontinuous. Prove that for every $\varepsilon > 0$ there exists a measurable set $E_\varepsilon \subset [0,1]$ of Lebesgue measure at least $1-\varepsilon$ such that the sequence $\{f_n(x,y)\}$ converges uniformly on the set $E_\varepsilon \times [0,1]$.

2.12.49. Suppose we are given a sequence of numbers $\gamma = \{\gamma_k\}$. For $x \in [0,1]$ let $f_\gamma(x) = 0$ if x is irrational, $f_\gamma(0) = 1$, and $f_\gamma(x) = \gamma_k$ if $x = m/k$ is an irreducible fraction. Prove that the function f_γ is Riemann integrable precisely when $\lim\limits_{k\to\infty} \gamma_k = 0$.

HINT: see Benedetto [**76**, Proposition 3.6, p. 96].

2.12.50.$^\circ$ Let a function f on the real line be periodic with a period $T > 0$ and integrable on intervals. Show that the integrals of f over $[0,T]$ and $[a, a+T]$ coincide for all a.

HINT: the translation invariance of Lebesgue measure yields the claim for simple T-periodic functions.

2.12.51. Construct a set $E \subset [0,1]$ with Lebesgue measure $\alpha \in (0,1)$ such that the integral of the function $|x - c|^{-1}$ over E is infinite for all $c \in [0,1]\backslash E$.

2.12.52. (M.K. Gowurin) Let a function f be Lebesgue integrable on $[0,1]$ and let $\alpha \in (0,1)$. Suppose that the integral of f over every set of measure α is zero. Prove that $f = 0$ almost everywhere.

HINT: show first that

$$\int_0^1 f(x)\,dx = 0,$$

by taking natural numbers n and m such that the number $n - m\alpha$ is nonnegative and does not exceed a given ε; to this end, extend f periodically from $[0,1)$ to $[0,n]$ and observe that the integral of f over $[0, m\alpha]$ is zero; next reduce the claim to the case $\alpha \le 1/2$ by using that $\min(\alpha, 1-\alpha) \le 1/2$; in the latter case observe that if the measure of the set $\{f \ge 0\}$ is at least α, then, by hypothesis and Example 1.12.8, the measure of the set $\{f > 0\}$ equals zero; finally, consider $\{f \le 0\}$.

2.12.53. Suppose that a function f is integrable on $[0,1]$ and $f(x) > 0$ for all x. Show that for each $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\int_A f(x)\,dx \ge \delta$$

for every set A with measure at least ε.

HINT: take $c > 0$ such that the measure of the set $\{f \ge c\}$ is greater than $1 - \varepsilon/2$ and estimate the integral of f over $A \cap \{f \ge c\}$ for sets A of measure ε.

2.12.54. Let $E \subset [0, 2\pi]$ be a set of Lebesgue measure d and let $n \in \mathbb{N}$. Prove the inequality

$$\int_E |\cos(nx)|\,dx \ge \frac{d}{2}\sin\frac{d}{8}.$$

HINT: observe that at all points from E that do not belong to the intervals of length $d/(4n)$ centered at $\pi/(2n) + k\pi/n$, one has the estimate $|\cos(nx)| \ge \sin d/8$, and the sum of measures of these intervals does not exceed $d/2$.

2.12.55. Let $E \subset \mathbb{R}$ be a set of finite Lebesgue measure. Evaluate the limit

$$\lim_{k\to\infty} \int_E (2 - \sin kx)^{-1}\,dx.$$

HINT: $\lambda(E)/\sqrt{3}$; it suffices to consider the case of finitely many intervals; consider first the case $E = [0, b]$; let I be the integral of $(2 - \sin x)^{-1}$ over $[0, 2\pi]$; then for $b \in (0, 2\pi)$ the integral of $n^{-1}(2-\sin x)^{-1}$ over $[0, nb]$ equals $[nb/(2\pi)]I + O(n^{-1})$, where $[r]$ is the integer part of r, which gives in the limit the number $Ib/(2\pi)$.

2.12.56.° Let μ be a bounded nonnegative measure on a σ-algebra $\mathcal{A}$. Prove that the definition of the Lebesgue integral given in the text is equivalent to the following definition. For simple functions we keep the same definition; for bounded measurable f we set

$$\int_X f\,d\mu = \lim_{n\to\infty} \int_X f_n\,d\mu,$$

where $\{f_n\}$ is an arbitrary sequence of simple functions uniformly convergent to f; for nonnegative measurable functions f we set

$$\int_X f\,d\mu = \lim_{n\to\infty} \int_X \min(f, n)\,d\mu,$$

and in the general case we declare f to be integrable if both functions $f^+ = \max(f, 0)$ and $f^- = -\min(f, 0)$ are integrable, and we set

$$\int_X f\,d\mu = \int_X f^+\,d\mu - \int_X f^-\,d\mu.$$

2.12.57.° The purpose of this exercise is to show that our definition of the Lebesgue integral is equivalent to the following definition due to Lebesgue himself. Let μ be a bounded nonnegative measure on a σ-algebra $\mathcal{A}$ and let f be a measurable function. Let us fix $\varepsilon > 0$ and consider the partition P of the real line into intervals $[y_i, y_{i+1})$, $i \in \mathbb{Z}$, $y_i < y_{i+1}$, of lengths not bigger than ε. Let $\delta(P) = \sup |y_{i+1} - y_i|$. Set $I(P) := \sum_{i=-\infty}^{+\infty} y_i \mu\bigl(x\colon\ y_i \le f(x) < y_{i+1}\bigr)$. Suppose that for some ε and P such a series converges (i.e., the series in positive and negative i converge separately). Show that this series converges for any partition and that, for any sequence of partitions P_k with $\delta(P_k) \to 0$, there exists a finite limit $\lim\limits_{k\to\infty} I(P_k)$ independent of our choice of the sequence of partitions, moreover, the function f is integrable in the sense of our definition and its integral equals the above limit. Show that it suffices to consider points $y_i = \varepsilon i$ or $y_i = i/n$, $n \in \mathbb{N}$.

HINT: it is clear that our definition yields the property described in this new definition. If the above-mentioned series converges, then it converges absolutely and hence the function g_P that equals y_i on the set $\{y_i \le f < y_{i+1}\}$ is integrable. Since $|f - g_P| \le \delta(P)$, the function f is integrable and the integrals of g_P approach the integral of f.

2.12.58.° Let f be a bounded function on a space X with a bounded nonnegative measure μ. For every partition of X into disjoint measurable parts $X_1, \ldots, X_n$ we set

$$L(\{X_i\}) = \sum_{i=1}^{n} \inf_{x\in X_i} f(x)\mu(X_i), \quad U(\{X_i\}) = \sum_{i=1}^{n} \sup_{x\in X_i} f(x)\mu(X_i).$$

The lower integral I_* of the function f equals the supremum of the sums $L(\{X_i\})$ over all possible finite partitions, and the upper integral I^* of f equals the infimum of the sums $U(\{X_i\})$ over all possible finite partitions. The function f will be called integrable if $I_* = I^*$. Prove that any function integrable in this sense is μ-measurable and its Lebesgue integral equals $I_* = I^*$. In addition, show that any bounded and μ-measurable function f is integrable in the indicated sense.

HINT: if f is integrable in the indicated sense, then one can find two sequences of simple functions φ_n and ψ_n with $\varphi_n(x) \le f(x) \le \psi_n(x)$ and $\|\varphi_n - \psi_n\|_{L^1(\mu)} \le 1/n$. If f is measurable and bounded, then one can consider the partitions into sets of the form $f^{-1}\bigl((a_i, a_{i+1}]\bigr)$, where $a_{i+1} - a_i = 1/n$ and finitely many intervals $[a_i, a_{i+1})$ cover the range of f.

2.12.59. (MacNeille [**642**], Mikusiński [**690**]) Let $\mathcal{R}$ be an algebra (or semi-algebra) of sets in a space X and let μ be a probability measure on $\mathcal{A} = \sigma(\mathcal{R})$. Prove that the function f is integrable with respect to μ precisely when there exists a sequence of $\mathcal{R}$-simple functions ψ_k (i.e., finite linear combinations of indicators of sets in $\mathcal{R}$) such that

$$\sum_{k=1}^{\infty} \int_X |\psi_k| \, d\mu < \infty$$

and $f(x) = \sum_{k=1}^{\infty} \psi_k(x)$ for every x such that the above series converges absolutely. In addition,

$$\int_X f \, d\mu = \sum_{k=1}^{\infty} \int_X \psi_k \, d\mu.$$

HINT: the hypothesis implies the integrability of f, since by the Fatou theorem the series of $|\psi_k|$ converges a.e. If f is integrable, then there exists a sequence of

$\mathcal{R}$-simple functions φ_k that converges to f a.e. and $\|f - \varphi_k\|_{L^1(\mu)} < 2^{-k-1}$. Then $\|\varphi_k - \varphi_{k+1}\|_{L^1(\mu)} < 2^{-k}$. Let $g_k = \varphi_k - \varphi_{k-1}$. It is clear that $\sum_{k=1}^n g_k \to f$ a.e. and $\sum_{k=1}^\infty |g_k| < \infty$ a.e. Let us consider the set E of measure zero on which the sum of the series of g_k is not equal to f, but the series converges absolutely. If E is empty, then we set $\psi_k = g_k$. If E is not empty, then we can find sets $R_k \in \mathcal{R}$ such that $\sum_{k=1}^\infty \mu(R_k) < \infty$ and every point from E belongs to infinitely many R_k. To this end, for every j we cover E by a sequence of sets $R_{jm} \in \mathcal{R}$ such that the sum of their measures is less than 2^{-j}, and then arrange R_{jm} in a single sequence. Finally, let us form a sequence of functions $g_1, I_{R_1}, -I_{R_1}, g_2, I_{R_2}, -I_{R_2}, \ldots$, according to the rule $\psi_{3k-2} = g_k$, $\psi_{3k-1} = I_{R_k}$, $\psi_{3k} = -I_{R_k}$. If $x \in E$, then the series of $|\psi_k(x)|$ diverges, since it contains infinitely many elements equal to 1. If this series converges, then $x \notin E$ and the series of $|g_k(x)|$ and $I_{R_k}(x)$ converge as well. Hence $f(x) = \sum_{k=1}^\infty g_k(x)$, which equals $\sum_{k=1}^\infty \psi_k(x)$ because $I_{R_k}(x) = 0$ for all sufficiently large k by convergence of the series. It remains to recall that the series of measures of R_k converges.

2.12.60. (F. Riesz) Denote by C_0 the class of all step functions on $[0,1]$, i.e., functions that are constant on intervals from certain finite partitions of $[0,1]$. Let C_1 denote the class of all functions f on $[0,1]$ for which there exists an increasing sequence of functions $f_n \in C_0$ such that $f_n(x) \to f(x)$ a.e. and the Riemann integrals of f_n are uniformly bounded. The limit of the Riemann integrals of f_n is denoted by $L(f)$. Finally, let C_2 denote the class of all differences $f = f_1 - f_2$ with $f_1, f_2 \in C_1$ and let $L(f) = L(f_1) - L(f_2)$. Prove that the class C_2 coincides with the class of Lebesgue integrable functions and that $L(f)$ is the Lebesgue integral of f.

HINT: one implication is obvious and the other one can be found in Riesz, Sz.-Nagy [**809**, Ch. 2].

2.12.61°. Let us define the integral of a bounded measurable function f on $[0,1]$ as follows. First we define the integral of a continuous function g over a closed set E as the difference between the integral of g over $[0,1]$ and the sum of the series of the integrals of g over finitely or countably many disjoint intervals forming $[0,1]\backslash E$. Given a closed set E, the integral over E of any function φ that is continuous on E is defined as the integral over E of its arbitrary continuous extension to $[0,1]$ (it is easily seen that this integral is independent of our choice of extension). Next we take a sequence of closed sets E_n with $\lambda(E_n) \to 1$ such that on each of them f is continuous, and define the integral of f over $[0,1]$ as the limit of the integrals of f over the sets E_n. Prove that this limit exists and equals the Lebesgue integral of f.

2.12.62. A function g on $\mathbb{R}^d$ with values in $[-\infty, +\infty]$ is called lower semicontinuous if, for every $c \in [-\infty, +\infty]$, the set $\{x\colon g(x) > c\}$ is open. Let $E \subset \mathbb{R}^d$ be a measurable set and let a function $f\colon E \to \mathbb{R}^1$ be integrable. Prove that, for any $\varepsilon > 0$, there exists a lower semicontinuous function g on $\mathbb{R}^d$ such that $g(x) \geq f(x)$ for all $x \in E$, $g|_E$ is integrable and the integral of $g - f$ over E does not exceed ε.

HINT: we find $\delta > 0$ with

$$\int_A |f|\, d\lambda < \varepsilon/2$$

whenever $A \subset E$ and $\lambda(A) < \delta$. Let us pick $\delta_n > 0$ such that $\sum_{n=1}^\infty \delta_n < \delta$ and $\sum_{n=1}^\infty \delta_n |q_n| < \varepsilon/2$, where $\{q_n\} = \mathbb{Q}$ is the set of all rational numbers. Let B_n be the ball of radius n centered at the origin and let G_n be an open set containing $E_n := B_n \cap \{x \in E\colon f(x) \geq q_n\}$ such that $\lambda(G_n) < \lambda(E_n) + \delta_n$. Set $g(x) =$

$\sup\{q_n\colon x \in G_n\}$ and $D := \bigcup_{n=1}^{\infty}\bigl((E \cap G_n)\backslash E_n\bigr)$. For any $c \in \mathbb{R}^1$, we have $\{g > c\} = \bigcup_{n\colon q_n>c} G_n$, i.e., g is lower semicontinuous. If $x \in E$ and $r > 0$, then there exists n with $f(x)-r \le q_n \le f(x)$ and $x \in B_n$. Then $x \in E_n$ and hence $g(x) \ge q_n \ge f(x)-r$. Since r is arbitrary, we obtain $g(x) \ge f(x)$. Finally, we show that the integral of $g - f$ over E does not exceed ε. Indeed, let $h := \sum_{n=1}^{\infty} |q_n| I_{(E\cap G_n)\backslash E_n}$. We observe that $g(x) \le f(x) + h(x) + |f(x)|I_D(x)$ for all $x \in E$. This follows from the fact that if $x \in E \cap G_n$, then either $x \in E_n$ and then $q_n \le f(x)$, or $x \notin E_n$ and then $q_n \le h(x)$. It remains to note that the integrals of h and $|f|I_D$ are majorized by $\varepsilon/2$.

2.12.63. (Hahn [**395**]) Let $f \in \mathcal{L}^1[0,1]$, let I be the integral of f, and let $\{\Pi_n\}$ be a decreasing sequence of finite partitions of $[0,1]$ into intervals $J_{n,k}$ $(k \le N_n)$ with $\lambda(J_{n,k}) \le \delta_n \to 0$, where λ is Lebesgue measure. Show that there exist points $\xi_{n,k} \in J_{n,k}$ such that $\bigl|\sum_{k=1}^{N_n} f(\xi_{n,k})\lambda(J_{n,k}) - I\bigr| \to 0$ as $n \to \infty$.

HINT: let us take continuous f_p with $\|f_p - f\|_{L^1} \to 0$ and $\lambda(f_p \ne f) \to 0$. Then we find increasing numbers p_l with $|f_l(t) - f_l(s)| \le 1/l$ for all $|t - s| \le \delta_{p_l}$. If $p_l \le n < p_{l+1}$ (let $p_1 = 1$), then we pick any $\xi_{n,k} \in J_{n,k} \cap \{f_l = f\}$, and if $J_{n,k} \cap \{f_l = f\} = \varnothing$, then we take $\xi_{n,k} \in J_{n,k}$ such that $|f(\xi_{n,k})| \le \inf_{J_{n,k}} |f(t)| + 1$. It remains to observe that the integral of $|f| + |f_l|$ over the set $\{f \ne f_l\}$ approaches zero, and for all $m \ge p_l$, the Riemann sum of f_l corresponding to the partition Π_m differs from the integral of f_l not greater than in $1/l$.

2.12.64. (Darji, Evans [**203**]) Let a function f be integrable on the unit cube $I \subset \mathbb{R}^n$. Show that there exists a sequence $\{x_k\}$ that is everywhere dense in I and has the following property: for every $\varepsilon > 0$, there exists $\delta > 0$ such that for every partition $\mathcal{P}$ of the cube I into finitely many parallelepipeds of the form $[a_i, b_i]\times\cdots\times[a_n, b_n]$ with pairwise disjoint interiors and $|b_i - a_i| < \delta$, one has

$$\Bigl|\sum_{P\in\mathcal{P}} f\bigl(r(P)\bigr)\lambda_n(P) - \int_I f(x)\,dx\Bigr| < \varepsilon,$$

where $r(P)$ is the first element in $\{x_k\}$ belonging to P.

2.12.65. Show that there exists a Borel set in $[0,1]$ such that its indicator function cannot coincide a.e. with the limit of an increasing sequence of nonnegative step functions.

HINT: let E be a Borel set such that the intersections of E and $[0,1]\backslash E$ with all intervals have positive measures. If $\{f_n\}$ is an increasing sequence of nonnegative step functions a.e. convergent to I_E, then there exist an interval I and a number n_1 such that $f_{n_1}(x) \ge 1/2$ for all $x \in I$. Then $I_E(x) \ge 1/2$ a.e. on I, i.e., one has $\lambda(I \cap E) = \lambda(I)$.

2.12.66.$^\circ$ Let f be a measurable function on the real line vanishing outside some interval. Show that if $\varepsilon_n \to 0$, then the functions $x \mapsto f(x + \varepsilon_n)$ converge to f in measure.

HINT: for continuous functions the claim is trivial, in the general case we find a sequence of continuous functions convergent to f in measure. Another solution can be derived from Exercise 4.7.104 in Chapter 4.

2.12.67. Let f be a bounded measurable function on the real line.

(i) Is it true that $f(x + n^{-1}) \to f(x)$ for a.e. x?

(ii) Show that there exists a subsequence $n_k \to \infty$ such that $f(x+n_k^{-1}) \to f(x)$ for a.e. x.

HINT: (i) no; consider the indicator of a compact set $K \subset [0,1]$ constructed as follows. For every n we partition $[0,1]$ into 2^{2^n} intervals $I_{n,k}$ of length $\varepsilon_n = 2^{-2^n}$, from every such interval we delete the interval $U_{n,k}$ of length ε_n^2 that is adjacent to the right endpoint of $I_{n,k}$, and denote the obtained closed set by K_n. Set $K = \bigcap_{n=1}^{\infty} K_n$. Then $\lambda(K) > 0$ and for any $x \in K \cap [0,1)$ there exist an arbitrary large number m with $x + m^{-1} \notin K$. This is verified with the aid of the following elementary assertion: if an interval U of length ε^2 belongs to the interval $[0,\varepsilon]$, then U contains a point of the form n^{-1}, $n \in \mathbb{N}$. For the proof of this assertion, it suffices to consider the smallest $k \in \mathbb{N}$ with $k^{-1} < \varepsilon$; then for some $l \in \mathbb{N}$ we have $(k+l)^{-1} \in U$ because $\varepsilon \le (k-1)^{-1}$, whence $\varepsilon - k^{-1} < \varepsilon^2$ due to the estimate $(k-1)^{-1} - (k-1)^{-2} < k^{-1}$. (ii) It suffices to verify our claim for functions with bounded support; in that case by Exercise 2.12.66 the functions $f(x+1/n)$ converge to f in measure and it remains to choose an a.e. convergent subsequence.

2.12.68.° Let $(X, \mathcal{A}, \mu)$ be a space with a nonnegative measure and let a function $f\colon X \times (a,b) \to \mathbb{R}^1$ be integrable in x for every t and differentiable in t at a fixed point $t_0 \in (a,b)$ for every x. Suppose that there exists a μ-integrable function Φ such that, for each t, there exists a set Z_t such that $\mu(Z_t) = 0$ and

$$|f(x,t) - f(x,t_0)| \le \Phi(x)|t - t_0| \quad \text{if } x \notin Z_t.$$

Show that the integral of $f(x,t)$ with respect to the measure μ is differentiable in t at the point t_0 and

$$\frac{d}{dt}\int_X f(x,t)\,\mu(dx) = \int_X \frac{\partial f(x,t_0)}{\partial t}\,\mu(dx).$$

HINT: for any sequence $\{t_n\}$, the union of the sets Z_{t_n} has measure zero; apply the reasoning from Corollary 2.8.7.

2.12.69. Prove that an arbitrary function $f\colon [0,1] \to \mathbb{R}$ can be written in the form $f(x) = \psi\bigl(\varphi(x)\bigr)$, where $\varphi\colon [0,1] \to [0,1]$ is a Borel function and $\psi\colon [0,1] \to \mathbb{R}$ is measurable with respect to Lebesgue measure.

HINT: writing $x \in [0,1]$ in the form $x = \sum_{n=1}^{\infty} x_n 2^{-n}$, where $x_n = 0$ or 1, we set $\varphi(x) = 2\sum_{n=1}^{\infty} x_n 3^{-n}$; observe that φ maps $[0,1]$ one-to-one to a subset of the Cantor set of measure zero; now ψ can be suitably defined on the range of φ; let $\psi = 0$ outside this range.

2.12.70. Show that almost everywhere convergence on the interval $I = [0,1]$ with Lebesgue measure cannot be defined by a topology, i.e., there exists no topology on the set of all measurable functions on I (or on the set of all continuous functions on I) such that a sequence of functions is convergent in this topology precisely when it converges almost everywhere.

HINT: use that any convergence defined by a topology has the following property: if every subsequence in a sequence $\{f_n\}$ contains a further subsequence convergent to some element f, then $f_n \to f$; find a sequence of continuous functions that converges in measure, but does not converge at any point.

2.12.71. (Marczewski [**651**]) Let μ be a probability measure such that convergence in measure for sequences of measurable functions is equivalent to convergence almost everywhere. Prove that the measure μ is purely atomic.

2.12.72.° Prove that a function f on an interval $[a,b]$ is continuous at a point x precisely when its oscillation at x is zero, where the oscillation at x is defined by the formula

$$\omega_f(x) := \lim_{\varepsilon\to 0} \sup\bigl\{|f(z)-f(y)|\colon\ |z-x|<\varepsilon,\ |y-x|<\varepsilon\bigr\}.$$

2.12.73°. (Baire's theorem) Let f_n be continuous functions on $[a,b]$ such that for every $x\in[a,b]$ there exists a finite limit $f(x)=\lim\limits_{n\to\infty} f_n(x)$. Prove that the set of points of continuity of f is everywhere dense in $[a,b]$.

HINT: apply the Baire category theorem to the sets $\bigl\{x\colon\ \omega_f(x)\ge j^{-1}\bigr\}$.

2.12.74. (i) Construct an example of a sequence of continuous functions f_n on $[0,1]$ such that, for every $x\in[0,1]$, there exists a finite limit $f(x)=\lim\limits_{n\to\infty} f_n(x)$, but the set of points of discontinuity of f is everywhere dense in $[0,1]$.

(ii) Construct an example showing that the function f in (i) may be discontinuous almost everywhere.

2.12.75. Prove that the uniform limit of a sequence of functions of Baire class α or less is also of Baire class α or less.

2.12.76. Prove that if a function φ is continuous on the real line and a function f is of Baire class α or less, then so is the function $\varphi\circ f$.

2.12.77. Prove that if a function f is of Baire class α or less on the plane, then the function $\varphi(x)=f(x,x)$ is of Baire class α or less on the real line.

2.12.78. Prove that the Dirichlet function (the indicator of the set of rational numbers) belongs to the second Baire class, but not to the first one.

2.12.79. Construct a measurable function on $[0,1]$ that cannot be redefined on a set of measure zero to obtain a function from the first Baire class.

HINT: use that all functions in the first Baire class have points of continuity. Consider the indicator function of a positive measure compact set without interior points.

2.12.80. Let a function f on the plane be continuous in every variable separately. Show that at some point f is continuous as a function on the plane.

2.12.81. Let f be a measurable real function on a measure space $(X,\mathcal{A},\mu)$ with a positive measure μ. Prove that there exists a number y such that

$$\int_X \frac{1}{|f(x)-y|}\,\mu(dx)=+\infty.$$

HINT: passing to a subset of X, we may assume that the function f is bounded and the measure μ is finite (if the measure is infinite on some set where f is bounded, then the claim is obvious); hence we assume that $0\le f\le 1$ and that $\mu(X)=1$; the preimage under f of at least one of the intervals $[0,1/2]$ or $[1/2,1]$ has measure at least $1/2$; we denote such an interval by I_1; by induction we construct a sequence of decreasing intervals I_n with $\mu(f^{-1}(I_n))\ge 2^{-n}$; there exists $y\in\bigcap_{n=1}^\infty I_n$; then $\mu\bigl(x\colon\ |f(x)-y|^{-1}\ge 2^n\bigr)\ge 2^{-n}$.

2.12.82°. Let $(X,\mathcal{A},\mu)$ be a measurable space with a finite positive measure μ and let f be a μ-measurable function with values in $\mathbb{R}$ or in $\mathbb{C}$. A point y is called an essential value of f if $\mu\bigl(x\colon\ |f(x)-y|<\varepsilon\bigr)>0$ for each $\varepsilon>0$.

(i) Show that a function f need not assume every essential value and that not every actual value of f is essential.

(ii) Show that the set of all essential values of f has a nonempty intersection with $f(X)$.

(iii) Show that the set of all essential values is closed and coincides with the intersection of the closures of the sets $\widetilde{f}(X)$ over all functions $\widetilde{f}$ a.e. equal to f.

2.12.83. Let μ be a nonnegative measure and let f be a μ-measurable function that has a bounded modification. Such functions are called essentially bounded. The essential supremum $\operatorname{esssup} f$ and essential infimum $\operatorname{essinf} f$ of the function f are defined as follows:

$$\operatorname{esssup} f := \inf\bigl\{M\colon\ f(x) \le M\ \mu\text{-a.e.}\bigr\}, \quad \operatorname{essinf} f := \sup\bigl\{m\colon\ f(x) \ge m\ \mu\text{-a.e.}\bigr\}.$$

A bounded measurable function f on $[a,b]$ is called reduced if, for every interval $(\alpha,\beta) \subset [a,b]$, one has

$$\inf{}_{(\alpha,\beta)} f = \operatorname{essinf}_{[\alpha,\beta]} f, \quad \sup{}_{(\alpha,\beta)} f = \operatorname{esssup}_{[\alpha,\beta]} f.$$

Prove that each bounded measurable function f on $[a,b]$ with Lebesgue measure has a reduced modification.

Hint: construct a version that satisfies the required condition for all rational α and β; observe that this condition is then fulfilled for all α and β.

2.12.84.$^\circ$ Let μ be a probability measure, $\varepsilon_n > 0$, $\sum_{n=1}^\infty \varepsilon_n < \infty$, and let f_n be μ-measurable functions such that

$$\sum_{n=1}^\infty \mu(x\colon\ |f_n(x)| > \varepsilon_n) < \infty.$$

Prove that

$$\sum_{n=1}^\infty |f_n(x)| < \infty \quad \text{a.e.}$$

Hint: let $E = \bigcap_{n=1}^\infty \bigcup_{m=n}^\infty \{x\colon\ |f_m(x)| > \varepsilon_m\}$; since

$$\mu\Bigl(\bigcup_{m=n}^\infty \{|f_m| > \varepsilon_m\}\Bigr) \le \sum_{m=n}^\infty \mu\bigl(\{|f_m| > \varepsilon_m\}\bigr),$$

then $\mu(E) = 0$; if $x \notin E$, then there exists n with $x \notin \{|f_m| > \varepsilon_m\}$ for all $m \ge n$, i.e., $|f_m(x)| \le \varepsilon_m$, which yields convergence of the series.

2.12.85.$^\circ$ Let $f, g\colon [0,1] \to [0,1]$, where f is continuous and g is Riemann integrable. Show that the composition $g \circ f$ may fail to be Riemann integrable.

2.12.86.$^\circ$ Let μ be a nonnegative measure, let $f \in \mathcal{L}^2(\mu) \cap \mathcal{L}^4(\mu)$, and let

$$\int f^2\, d\mu = \int f^3\, d\mu = \int f^4\, d\mu.$$

Prove that $f(x) \in \{0,1\}$ a.e.

Hint: observe that the integral of $(f^2 - f)^2$ vanishes, which yields $f^2 = f$ a.e.

2.12.87.$^\circ$ Let $1 < p < \infty$, $p^{-1} + q^{-1} = 1$. Prove that for all nonnegative a and b one has the inequality $ab \le \frac{a^p}{p} + \frac{b^q}{q}$, where the equality is only possible if $b = a^{p-1}$.

Hint: consider the graph of the function $y = x^{p-1}$ on $[0,a]$ and observe that the area of the region between it and the first coordinate axis equals a^p/p, whereas the area of the region between the graph and the straight line $y = b$ equals b^q/q; use that the sum of the two areas is not less than ab, and the equality is only possible if $b = a^{p-1}$.

2.12.88.° Justify the relation (2.12.8).

2.12.89.° Let $1 < p < \infty$, $p^{-1} + q^{-1} = 1$, $f \in \mathcal{L}^p(\mu)$, $g \in \mathcal{L}^q(\mu)$, and let

$$\int fg\,d\mu = \|f\|_p \|g\|_q > 0.$$

Prove that $g = \operatorname{sign} f \cdot |f|^{p-1}$ a.e.

HINT: conclude from the proof of the Hölder inequality and Exercise 2.12.87 that $|g| = |f|^{p-1}$, whence the claim follows.

2.12.90.° Let μ be a probability measure and let f be a nonnegative μ-integrable function such that $\ln f \in \mathcal{L}^1(\mu)$. Prove that

$$\lim_{p\to 0+} \int \frac{f^p - 1}{p}\,d\mu = \int \ln f\,d\mu.$$

HINT: use the inequality $|t^p - 1|/p \le |t-1| + |\ln t|$ for $t > 0$, $p \in (0,1)$, and the dominated convergence theorem.

2.12.91.° Let μ be a probability measure and let f be a nonnegative μ-integrable function such that $\ln f \in \mathcal{L}^1(\mu)$. Prove that

$$\lim_{p\to 0+} \biggl(\int f^p\,d\mu\biggr)^{1/p} = \exp \int \ln f\,d\mu.$$

HINT: apply the previous exercise.

2.12.92.° Let μ be a probability measure and let $f \in L^1(\mu)$. Prove that

$$1 + \biggl(\int |f|\,d\mu\biggr)^2 \le \biggl(\int \sqrt{1+|f|^2}\,d\mu\biggr)^2 \le \biggl(1 + \int |f|\,d\mu\biggr)^2.$$

HINT: apply Jensen's inequality to the function $\varphi(t) = \sqrt{1+t^2}$ and the estimate $\sqrt{1+|f|^2} \le 1 + |f|$.

2.12.93.° Let $f, g \ge 0$ be integrable functions on a space with a probability measure μ and let $fg \ge 1$. Show that

$$\int f\,d\mu \int g\,d\mu \ge 1.$$

HINT: observe that $\sqrt{f}\sqrt{g} \ge 1$ and apply the Cauchy–Bunyakowsky inequality.

2.12.94.° Let μ be a countably additive measure with values in $[0, +\infty]$ and let $f \in \mathcal{L}^1(\mu)$ be such that $f - 1 \in \mathcal{L}^p(\mu)$ for some $p \in [1, \infty)$. Prove that the measure μ is finite.

HINT: observe that the sets $\{f \le 1/2\}$ and $\{f \ge 1/2\}$ have finite measures due to integrability of $|f-1|^p$ and f.

2.12.95. Let μ be a probability measure, let $\{f_n\} \subset \mathcal{L}^1(\mu)$, and let I_n be the integral of f_n. Suppose that there exists $c > 0$ such that

$$\|f_n - I_n\|_p^p \le c\|f_n\|_1, \quad \forall\, n \in \mathbb{N}.$$

Prove that either

$$\limsup_{n\to\infty} \|f_n\|_1 < \infty \quad \text{and} \quad \liminf_{n\to\infty} |f_n(x)| < \infty \text{ a.e.},$$

or

$$\limsup_{n\to\infty} \|f_n\|_1 = \infty \quad \text{and} \quad \limsup_{n\to\infty} |f_n(x)| = \infty \text{ a.e.}$$

HINT: let $\|f_n\|_1 \to \infty$; if the sequence $\{I_n/\|f_n\|_1^{1/p}\}$ is bounded, then we obtain the uniform boundedness of $\|f_n\|_p^p/\|f_n\|_1$, which by the Hölder inequality yields the uniform boundedness of the numbers $\|f_n\|_p^{p-1}$, hence of the numbers $\|f_n\|_1$, which is a contradiction. Now we may assume that $C_n := I_n/\|f_n\|_1^{1/p} \to +\infty$. Then by Fatou's theorem $\liminf\limits_{n\to\infty}\bigl|f_n(x)/\|f_n\|_1^{1/p} - C_n\bigr| < \infty$ a.e., whence $\limsup\limits_{n\to\infty}|f_n(x)| = \infty$ a.e.

2.12.96.° Let $f \in \mathcal{L}^1[a,b]$ and let

$$\int_a^b t^k f(t)\,dt = 0$$

for all nonnegative integer k. Show that $f = 0$ a.e.

HINT: take a sequence of polynomials p_j that is uniformly bounded on $[a,b]$ and $p_j(t) \to \operatorname{sign} f(t)$ a.e.

2.12.97. (G. Hardy) Let f be a nonnegative measurable function on $[0,+\infty)$ and let $1 \le q < \infty$, $0 < r < \infty$. Show that

$$\int_0^\infty \Bigl(\int_0^t f(s)\,ds\Bigr)^q t^{-r-1}\,dt \le \Bigl(\frac{q}{r}\Bigr)^q \int_0^\infty s^{q-r-1} f(s)^q\,ds.$$

HINT: for $q > 1$ take $p = q/(q-1)$, set $\alpha = (1 - r/q)/p$ and apply the Hölder inequality to the integral of $f(s)s^{\alpha}s^{-\alpha}$ over $[0,t]$ in order to estimate it by the product of the integrals of $f(s)^q s^{\alpha q}$ and $s^{-\alpha p}$ in the corresponding powers.

2.12.98. (P.Yu. Glazyrina) Let $f \ge 0$ be a μ-measurable function. Prove the inequality

$$\int f^p\,d\mu \int f^{s-p}\,d\mu \le \int f^q\,d\mu \int f^{s-q}\,d\mu$$

assuming that p, q, s are real numbers such that $|p - s/2| < |q - s/2|$ and the above integrals exist.

HINT: let $r = (s-2q)/(p-q)$, $t = (s-2q)/(s-p-q)$. Then by our hypothesis $r > 1$, $r^{-1} + t^{-1} = 1$ and $t > 1$. Set $\alpha = q/t$, $\beta = q/r$. Since

$$\alpha t = q,\ \ (p-\alpha)r = (p - q/t)r = (p - q + q/r)r = s - 2q + q = s - q,$$

one has by Hölder's inequality

$$\int f^p\,d\mu \le \Bigl(\int f^{\alpha t}\,d\mu\Bigr)^{1/t}\Bigl(\int f^{(p-\alpha)r}\,d\mu\Bigr)^{1/r} = \Bigl(\int f^q\,d\mu\Bigr)^{1/t}\Bigl(\int f^{s-q}\,d\mu\Bigr)^{1/r}.$$

Similarly, one has

$$\int f^{s-p}\,d\mu \le \Bigl(\int f^q\,d\mu\Bigr)^{1/r}\Bigl(\int f^{s-q}\,d\mu\Bigr)^{1/t}.$$

It remains to multiply the two inequalities.

2.12.99. (Fukuda [**334**], Vakhania, Kvaratskhelia [**971**]) Let μ be a probability measure and let $f \in L^p(\mu)$ be such that $\|f\|_{L^p(\mu)} \le C\|f\|_{L^q(\mu)}$ for some $q \in [1,p)$ and $C \ge 1$. Show that

$$\|f\|_{L^r(\mu)} \le C^{\kappa}\|f\|_{L^s(\mu)}$$

whenever $1 \le s < r \le p$, where $\kappa = 1$ if $q \le s < r \le p$, $\kappa = q(p-s)(s(p-q))^{-1}$ if $s < q < r \le p$, $\kappa = p(q-s)(s(p-q))^{-1}$ if $s < r \le q$.

HINT: the case $q \le s < r \le p$ follows at once by the monotonicity of the function $t \mapsto \|f\|_{L^t(\mu)}$. Let $s < q < r \le p$ and let

$$\alpha = p(q-s)(q(p-s))^{-1}, \quad \beta = s(p-q)(q(p-s))^{-1}.$$

Then $0 < \alpha, \beta < 1$, $\alpha+\beta = 1$. Take $t = p(\alpha q)^{-1}$. Then $t > 1$ and $t = (p-s)(q-s)^{-1}$, $t' = (p-s)(p-q)^{-1}$, $\beta q t' = s$. By Hölder's inequality

$$\|f\|^q_{L^q(\mu)} \le \|f\|^{p/t}_{L^p(\mu)} \|f\|^{\beta q}_{L^s(\mu)} \le C^{p(q-s)/(p-s)} \|f\|^{p(q-s)/(p-s)}_{L^q(\mu)} \|f\|^{s(p-q)/(p-s)}_{L^s(\mu)},$$

which yields $\|f\|_{L^q(\mu)} \le C^{(p(q-s)/(s(p-q))} \|f\|_{L^s\mu)}$. Since $\|f\|_{L^r(\mu)} \le \|f\|_{L^p(\mu)}$, we arrive at the desired estimate. The remaining case is deduced from the considered one.

2.12.100. (i) Let E be a partially ordered real vector space such that if $x \le y$, then $tx \le ty$ for all $t \ge 0$ and $x + z \le y + z$ for all $z \in E$. Suppose that E_0 is a linear subspace in E such that, for each $x \in E$, there exists an element $x_0 \in E_0$ with $x \le x_0$. Let L_0 be a linear function on E_0 such that $L_0(v) \ge 0$ whenever $v \in E_0$ and $v \ge 0$. Prove that L_0 can be extended to a linear function L on E such that $L(x) \ge 0$ for all $x \ge 0$.

(ii) Deduce from (i) the existence of a nonnegative finitely additive function on the class of all subsets of $[0, 1]$ extending Lebesgue measure.

(iii) Deduce from (i) the existence of a generalized limit on the space m of all bounded sequences, i.e., a linear function Λ on m such that $\Lambda(x) \ge 0$ for all $x = (x_n)$ with $x_n \ge 0$ and $\Lambda(x) = \lim\limits_{n\to\infty} x_n$ for all convergent sequences $x = (x_n)$.

HINT: (i) apply the Hahn–Banach theorem 1.12.26 to the function

$$p(x) = \inf\{L_0(v)\colon\ v \in E_0, x \le v\};$$

(ii) take for E the space of all bounded functions on $[0, 1]$ and for E_0 the subspace consisting of measurable functions, define L_0 on E_0 as the Lebesgue integral; (iii) take for E_0 the subspace of all convergent sequences.

2.12.101. (S. Banach) (i) Prove that on the space L of all bounded functions on $[0, 1)$ there exists a linear function Λ with the following properties:

(a) if $f \in L$ is Lebesgue integrable, then $\Lambda(f)$ coincides with the Lebesgue integral of f over $[0, 1)$,

(b) if $f \in L$ and $f \ge 0$, then $\Lambda(f) \ge 0$,

(c) $\Lambda\bigl(f(\,\cdot + s)\bigr) = \Lambda(f)$ for all $f \in L$ and $s \in [0, 1]$, where $f(t+s) = f\bigl(\mathrm{fr}(t+s)\bigr)$, $\mathrm{fr}(s)$ is the fractional part of s.

(ii) Construct a linear function on L that coincides with the integral on the set of all Riemann integrable functions, but differs from the Lebesgue integral at some Lebesgue integrable function.

HINT: (i) consider the function p from Example 1.12.27 on the space L of all bounded functions on the real line with a period 1; on the linear subspace L_0 in L formed by integrable functions we set

$$\Lambda_0(f) = \int_0^1 f\,dx.$$

Show that $\Lambda_0(f) \le p(f)$ by using the equality

$$\int_0^1 f(t+a)\,dt = \int_0^1 f(t)\,dt$$

for periodic functions; extend Λ_0 to a linear function Λ on L with $\Lambda \leq p$ and verify the required properties by using that $p(f) \leq 0$ whenever $f \leq 0$ and that $p\bigl(f(\,\cdot + h)\bigr) = p(f)$. In (ii), a similar reasoning applies.

2.12.102. (S. Banach) Prove that Lebesgue measure on $[0,1]$ can be extended to an additive but not countably additive nonnegative set function ν that is defined on the class of all subsets of $[0,1]$ and has the following invariance property: $\nu(E+h) = \nu(E)$ for all $E \subset (0,1]$ and $h \in (0,1]$, where in the formation of the sum $E+h$ the numbers $e+h>1$ are replaced by $e+h-1$ (in this and the previous example one can deal with the circle and rotations in place of $(0,1]$ and translations).

HINT: consider $\nu(E) = \Lambda(I_E)$, where Λ is the linear function on the space of all bounded functions on $(0,1]$ from Exercise 2.12.101.

2.12.103. Let $f \in \mathcal{L}^1(\mathbb{R}^1)$ and $a > 0$.

(i) Show that the series $\sum\limits_{n=-\infty}^{+\infty} f(n + a^{-1}x)$ converges absolutely for a.e. x.

(ii) Let $g(x) = \sum\limits_{n=-\infty}^{+\infty} f(n+a^{-1}x)$ if the series converges and $g(x) = 0$ otherwise. Show that

$$\int_0^a g(x)\,dx = a\int_{-\infty}^{+\infty} f(x)\,dx.$$

(iii) Show that for a.e. x for each $a > 0$ one has $\lim\limits_{n\to\infty} n^{-a} f(nx) = 0$.

HINT: (i) observe that

$$\sum_{n=-\infty}^{+\infty} \int_0^a |f(n + a^{-1}x)|\,dx = a\int_{-\infty}^{+\infty} |f(x)|\,dx;$$

(ii) use the monotone convergence theorem; (iii) observe that

$$\sum_{n=1}^{\infty} n^{-a} \int_{-\infty}^{+\infty} |f(nx)|\,dx < \infty,$$

by using the change of variable $y = nx$ (see Chapter 3 about the change of variable).

2.12.104. Let $f \in \mathcal{L}^1(\mathbb{R}^1)$. Prove the equality

$$\biggl|\int_{-\infty}^{+\infty} f(x)\,dx\biggr| = \inf\biggl\{\int_{-\infty}^{+\infty} \Bigl|\sum_{i=1}^{n} \alpha_i f(x + x_i)\Bigr|\,dx\biggr\},$$

where inf is taken over all numbers $x_i \in \mathbb{R}^1$, $n \in \mathbb{N}$ and $\alpha_i \geq 0$ with $\alpha_1+\cdots+\alpha_n = 1$.

HINT: let the integral of f be nonnegative; then the right-hand side of the equality to be proven is not less than the left-hand side, since the integral of $\sum\limits_{i=1}^{n} \alpha_i f(x+x_i)$ equals the integral of f; the reverse inequality is easily verified with the aid of the Riemann sums in the case of a continuous function f with bounded support; in the general case one can approximate f in the mean by continuous functions with bounded support.

2.12.105. (Fréchet [**320**], Slutsky [**889**]) Let μ be a probability measure on a space X and let f be a μ-measurable function. We call a number m a median of f if $\mu(f < c) \leq 1/2$ for all $c < m$ and $\mu(f < c) \geq 1/2$ for all $c > m$.

(i) Prove that a median of f exists, but may not be unique.

(ii) Prove that a median is unique if f has a continuous strictly increasing distribution function Φ_f and then $m = \Phi_f^{-1}(1/2)$.

(iii) Suppose that measurable functions f_n converge to f in measure μ. Prove that the set of medians of the functions f_n is bounded and that if m_n is a median of f_n and m is a limit point of $\{m_n\}$, then m is a median of f.

HINT: (i), (ii) take for a median any number in the interval between the numbers $\sup\{c\colon\ \mu(f < c) < 1/2\}$ and $\sup\{c\colon\ \mu(f < c) \le 1/2\}$. (iii) Take an interval $[a, b]$ containing all medians of f; then it is easily verified that for all sufficiently large n all medians of f_n are contained in $[a-1, b+1]$; if $c < m$, but $\mu(f < c) > 1/2$, then there exists $c_1 < c$ such that $\mu(f < c_1) > 1/2$; then, for all sufficiently large n we have $c < m_n$ and $\mu(f_n < c) > 1/2$, which is a contradiction; similarly we verify that $\mu(f < c) \ge 1/2$ for all $c > m$.

2.12.106. Let f be a nonnegative continuous function on $[0, +\infty)$ with the infinite integral over $[0, +\infty)$. Show that there exists $a > 0$ with $\sum_{n=1}^{\infty} f(na) = \infty$.

HINT: see Sadovnichiĭ, Grigoryan, Konyagin [**839**, Ch. 1, §4, Problem 46] and comments in Buczolich [**140**].

2.12.107. (Buczolich, Mauldin) Prove that there exist an open set $E \subset (0, +\infty)$ and intervals I_1 and I_2 in $[1/2, 1)$ such that $\sum_{n=1}^{\infty} I_E(nx) = \infty$ for all $x \in I_1$ and $\sum_{n=1}^{\infty} I_E(nx) < \infty$ for all $x \in I_2$.

HINT: see references and comments in Buczolich [**140**].

2.12.108. Suppose we are given two measurable sets A and B in the circle of length 1 having linear Lebesgue measures α and β, respectively. Let B_φ be the image of the set B under the rotation in the angle φ counter-clockwise. Show that for some φ the set $A \cap B_\varphi$ has measure at least $\alpha\beta$.

HINT: observe that the integral of $\lambda(A \cap B_\varphi)$ in φ equals $\alpha\beta$; see Sadovnichiĭ, Grigoryan, Konyagin [**839**, Ch. 4, §3, Problem 11].

2.12.109. Let f be an integrable complex-valued function on a space X with a probability measure μ. Prove that

$$\int_X f\, d\mu = 0$$

precisely when

$$\int_X |1 + zf(x)|\, dx \ge 1$$

for all complex numbers z.

HINT: if this inequality is fulfilled, then one can use that $\dfrac{|1 + r\exp(i\theta)f(x)| - 1}{r}$ tends to $\mathrm{Re}\big[(\exp(i\theta)f(x)\big]$ as $r \to 0+$ for all $\theta \in \mathbb{R}^1$ and is majorized by $|f(x)|$; one can take θ such that

$$\exp(i\theta)\int_X f\, d\mu = -\Big|\int_X f\, d\mu\Big|.$$

2.12.110. Let $\{f_n\}$ be a sequence of integrable complex-valued functions on $[0, 1]$ such that

$$\lim_{n\to\infty}\int_0^1 |\mathrm{Re} f_n(x)|\, dx = 1, \quad \lim_{n\to\infty}\int_0^1 \big|1 - |f_n(x)|\big|\, dx = 0.$$

Show that

$$\lim_{n\to\infty}\int_0^1 |\mathrm{Im} f_n(x)|\, dx = 0.$$

HINT: see George [**351**, p. 250].

2.12.111. (Kakutani [**481**]) Let f and g be two nonnegative measurable functions on $[0,1]$ having the following property: if the integral of f over some measurable set E is finite, then the integral of g over E is finite as well. Prove that there exist a constant K and a nonnegative integrable function h such that $g(x) \le Kf(x)+h(x)$.

2.12.112.° Suppose that increasing functions f_n converge in measure on the interval $[a,b]$ with Lebesgue measure. Show that they converge almost everywhere.

HINT: there is a subsequence in $\{f_n\}$ that converges almost everywhere on $[a,b]$. It is readily seen that there exists an increasing function f to which this subsequence converges almost everywhere. It remains to verify that $\{f_n\}$ converges to f at every continuity point of f.

2.12.113. (Lovász, Simonovits [**623**]) Suppose we are given lower semicontinuous integrable functions u_1 and u_2 on $\mathbb{R}^n$. Prove that there exist $a, b \in \mathbb{R}^n$ and an affine function $L\colon\ (0,1) \to (0,+\infty)$ such that

$$\int_0^1 u_i\bigl((1-t)a+tb\bigr)L(t)^{n-1}\,dt > 0, \quad i=1,2.$$

HINT: see [**623**] and Kannan, Lovász, Simonovits [**489**].

2.12.114. Suppose that a sequence of convex functions f_n on a ball $U \subset \mathbb{R}^d$ is uniformly bounded. Prove that it contains a subsequence convergent in $L^p(U)$ for all $p \in [1,\infty)$.

HINT: it suffices to show that $\{f_n\}$ is uniformly Lipschitzian on every smaller ball V with the same center. To this end, it is sufficient to show that for every convex function f on an interval $[a,b]$ and every $\delta > 0$, one has $|f'(t)| \le 2\delta^{-1}\sup_{x\in[a,b]}|f(x)|$ for a.e. $t \in [a+\delta, b-\delta]$. This estimate follows easily by the convexity: if $f'(t) > 0$, then $f'(t)(b-t) \le f(b) - f(t)$; the case $f'(t) < 0$ is similar.

2.12.115. Let μ be a probability measure on a measurable space $(X,\mathcal{A})$, let $1 < p < \infty$, and let $f_n \in \mathcal{L}^p(\mu)$ be nonnegative functions such that

$$\|f_n\|_{L^p(\mu)} \le C\|f_n\|_{L^1(\mu)}$$

with some constant C (or, more generally, $\bigl\|\sum_{n=1}^N f_n\bigr\|_{L^p(\mu)} \le C\sum_{n=1}^N \|f_n\|_{L^1(\mu)}$). Prove that the series $\sum_{n=1}^\infty f_n$ converges μ-a.e. if and only if

$$\sum_{n=1}^\infty \int_X f_n\,d\mu < \infty.$$

HINT: in one direction the claim follows by the monotone convergence theorem. Suppose that the series of the integrals of f_n diverges. By Proposition 2.11.7 and the estimate $\Bigl\|\sum_{n=1}^N f_n\Bigr\|_{L^p(\mu)} \le \sum_{n=1}^N \|f_n\|_{L^p(\mu)}$ one has

$$\mu\Bigl(x\colon\ \sum_{n=1}^N f_n(x) \ge \frac{1}{2}\sum_{n=1}^N \|f_n\|_{L^1(\mu)}\Bigr) \ge 2^{-q}\Bigl(\sum_{n=1}^N \|f_n\|_{L^1(\mu)}\Bigr)^q \Bigl\|\sum_{n=1}^N f_n\Bigr\|_{L^p(\mu)}^{-q}$$
$$\ge 2^{-q}\Bigl(\sum_{n=1}^N \|f_n\|_{L^1(\mu)}\Bigr)^q \Bigl(\sum_{n=1}^N \|f_n\|_{L^p(\mu)}\Bigr)^{-q} \ge 2^{-q}C^{-q}.$$

Therefore, $\lim\limits_{N\to\infty}\sum_{n=1}^N f_n(x) = \infty$ on a positive measure set.

2.12.116. (Kadec, Pełczyński [**476**]) Let μ be a probability measure on a measurable space $(X,\mathcal{A})$ and let $p\geq 1$, $\varepsilon>0$. Set

$$M^p_\varepsilon := \Bigl\{f\in\mathcal{L}^p(\mu)\colon\ \mu\bigl(x\colon\ |f(x)|\geq \varepsilon\|f\|_{L^p(\mu)}\bigr)\geq\varepsilon\Bigr\}.$$

(i) Show that $\mathcal{L}^p(\mu)=\bigcup_{\varepsilon>0}M^p_\varepsilon$.

(ii) Suppose that $f\in\mathcal{L}^p(\mu)$, where $p>1$, and that $\|f\|_{L^p(\mu)}\leq C\|f\|_{L^r(\mu)}$, where $r\in(1,p)$. Show that $f\in M^p_\varepsilon$ with $\varepsilon=C^{rp/(p-1)}2^{p/(1-p)}$.

Hint: (i) Let $E_\varepsilon=\bigl\{x\colon\ |f(x)|\geq\varepsilon\|f\|_{L^p(\mu)}\bigr\}$. If $\mu(E_\varepsilon)<\varepsilon$ for all $\varepsilon>0$, then, letting $a:=\|f\|_{L^p(\mu)}$, we obtain $a>0$ and $\mu\bigl(x\colon\ |f(x)|<\varepsilon a\bigr)>1-\varepsilon$, which yields $f=0$ a.e., a contradiction.

(ii) Let $\varepsilon=C^{rp/(p-1)}2^{p/(1-p)}$. If $f\notin M^p_\varepsilon$, then $\mu(E_\varepsilon)<\varepsilon$. Hence by Hölder's inequality

$$\int_X |f|^r\,d\mu\leq\mu(E_\varepsilon)^{(p-1)/p}\|f\|^r_{L^p(\mu)}+\varepsilon^r\|f\|^r_{L^p(\mu)}\leq 2\varepsilon^{(p-1)/p}\|f\|^r_{L^p(\mu)}.$$

Since $\|f\|_{L^r(\mu)}\geq C\|f\|_{L^p(\mu)}$, we obtain the desired bound.

2.12.117. (Sarason [**845**]) Let $(X,\mathcal{A},\mu)$ be a probability space and let $f>0$ be a μ-measurable function such that

$$\int_X f\,d\mu\int_X f^{-1}\,d\mu\leq 1+c^3$$

for some $c\in(0,1/2)$. Let J be the integral of f and let I be the integral of $\ln f$. Show that

$$\int_X |\ln f-\ln J|\,d\mu\leq 8c,\quad \int_X|\ln f-I|\,d\mu\leq 16c.$$

Hint: by scaling we may assume without loss of generality that $J=1$ and thus that the integral of $1/f$ is $1+c^3$. Let $A:=\{x\colon\ (1+c)^{-1}<f(x)<1+c\}$. Observe that $t+t^{-1}\geq 1+c+(1+c)^{-1}$ if $t\geq(1+c)^{-1}$ or $t\leq 1+c$. Since $f+f^{-1}\geq 2$, we obtain

$$2+c^3=\int_X(f+f^{-1})\,d\mu\geq[1+c+(1+c)^{-1}]\mu(X\backslash A)+2\mu(A)=2+c^2(1+c)^{-1}\mu(X\backslash A).$$

Hence $\mu(X\backslash A)\leq c(1+c)\leq 2c$, so $m(A)\geq 1-2c$. Therefore,

$$\int_{X\backslash A}f\,d\mu=1-\int_A f\,d\mu\leq 1-(1+c)^{-1}\mu(A)\leq 1-(1-2c)(1+c)^{-1}\leq 3c,$$

$$\int_{X\backslash A}f^{-1}\,d\mu=1+c^3-\int_A f^{-1}\,d\mu\leq 1+c^3-(1+c)^{-1}\mu(A)\leq 4c.$$

On A we have $|\ln f|<\ln(1+c)\leq c$. Since $|\ln f|\leq f+f^{-1}$ everywhere, we obtain

$$\int_X|\ln f|\,d\mu\leq c+\int_{X\backslash A}(f+f^{-1})\,d\mu\leq 8c.$$

It remains to use the estimate

$$|I|\leq\int_X|\ln f|\,d\mu.$$

CHAPTER 3

Operations on measures and functions

Теряя форму, гибнет красота,
А форма строго требует закона.
В. Солоухин. Венок сонетов

Losing its form, beauty perishes,
and the form demands a law.
V. Solouhin. A wreath of sonnets.

3.1. Decomposition of signed measures

In this section, we consider signed measures. A typical example of a signed measure is the difference of two probability measures. We shall see below that every signed measure on a σ-algebra is the difference of two nonnegative measures. The following theorem enables one in many cases to pass from signed measures to nonnegative ones.

3.1.1. Theorem. *Let μ be a countably additive real-valued measure on a measurable space $(X, \mathcal{A})$. Then, there exist disjoint sets $X^-, X^+ \in \mathcal{A}$ such that $X^- \cup X^+ = X$ and for all $A \in \mathcal{A}$, one has*

$$\mu(A \cap X^-) \le 0 \quad \textit{and} \quad \mu(A \cap X^+) \ge 0.$$

PROOF. A set $E \in \mathcal{A}$ will be called negative if $\mu(A \cap E) \le 0$ for all $A \in \mathcal{A}$. By analogy we define positive sets. Let $\alpha = \inf \mu(E)$, where the infimum is taken over all negative sets. Let E_n be a sequence of negative sets with $\lim\limits_{n\to\infty} \mu(E_n) = \alpha$. It is clear that $X^- := \bigcup_{n=1}^\infty E_n$ is a negative set and that $\mu(X^-) = \alpha$, since $\alpha \le \mu(X^-) \le \mu(E_n)$. We show that $X^+ = X\backslash X^-$ is a positive set. Suppose the contrary. Then, there exists $A_0 \in \mathcal{A}$ such that $A_0 \subset X^+$ and $\mu(A_0) < 0$. The set A_0 cannot be negative, since the set $X^- \cup A_0$ would be negative as well, but $\mu(X^- \cup A_0) < \alpha$, which is impossible. Hence one can find a set $A_1 \subset A_0$ and a number $k_1 \in \mathbb{N}$ such that $A_1 \in \mathcal{A}$, $\mu(A_1) \ge 1/k_1$, and k_1 is the smallest natural number k for which A_0 contains a subset with measure not less than $1/k$. We observe that $\mu(A_0\backslash A_1) < 0$. Repeating the same reasoning for $A_0\backslash A_1$ in place of A_0 we obtain a set A_2 in $\mathcal{A}$ contained in $A_0\backslash A_1$ such that $\mu(A_2) \ge 1/k_2$ with the smallest possible natural k_2. Let us continue this process inductively. We obtain pairwise disjoint sets $A_i \in \mathcal{A}$ with the following property: $A_{n+1} \subset A_0\backslash \bigcup_{i=1}^n A_i$ and $\mu(A_n) \ge 1/k_n$, where k_n is the smallest natural number k such that $A_0\backslash \bigcup_{i=1}^{n-1} A_i$ contains a subset with measure not less than $1/k$. We observe that $k_n \to +\infty$, since otherwise by using that the sets A_n are disjoint we would obtain that $\mu(A_0) = +\infty$. Let $B = A_0 \backslash \bigcup_{i=1}^{\infty} A_i$. Note that $\mu(B) < 0$, since $\mu(A_0) < 0$, $\mu\bigl(\bigcup_{i=1}^{\infty} A_i\bigr) > 0$ and $\bigcup_{i=1}^{\infty} A_i \subset A_0$. Moreover, B is a negative set. Indeed, if $C \subset B$, $C \in \mathcal{A}$ and $\mu(C) > 0$, then there exists a natural number k with $\mu(C) > 1/k$, which for $k_n > k$ contradicts our choice of k_n because $C \subset A_0 \backslash \bigcup_{i=1}^{n} A_i$. Thus, adding B to X^-, we arrive at a contradiction with the definition of α. Hence the set X^+ is positive. □

The decomposition of the space X into the disjoint union $X = X^+ \cup X^-$ constructed in the above theorem is called the *Hahn decomposition*. It is clear that the Hahn decomposition may not be unique, since one can add to X^+ a set all subsets of which have measure zero. However, if $X = \widetilde{X}^+ \cup \widetilde{X}^-$ is another Hahn decomposition, then, for all $A \in \mathcal{A}$, we have

$$\mu(A \cap X^-) = \mu(A \cap \widetilde{X}^-) \quad \text{and} \quad \mu(A \cap X^+) = \mu(A \cap \widetilde{X}^+). \tag{3.1.1}$$

Indeed, any set B in $\mathcal{A}$ belonging to $X^- \cap \widetilde{X}^+$ or to $X^+ \cap \widetilde{X}^-$ has measure zero, since $\mu(B)$ is simultaneously nonnegative and nonpositive.

3.1.2. Corollary. *Under the hypotheses of Theorem* 3.1.1 *let*

$$\mu^+(A) := \mu(A \cap X^+), \quad \mu^-(A) := -\mu(A \cap X^-), \quad A \in \mathcal{A}. \tag{3.1.2}$$

Then μ^+ and μ^- are nonnegative countably additive measures and one has the equality $\mu = \mu^+ - \mu^-$.

It is clear that $\mu(X^+)$ is the maximal value of the measure μ, and $\mu(X^-)$ is its minimal value.

3.1.3. Corollary. *If $\mu\colon\ \mathcal{A} \to \mathbb{R}^1$ is a countably additive measure on a σ-algebra $\mathcal{A}$, then the set of all values of μ is bounded.*

3.1.4. Definition. *The measures μ^+ and μ^- constructed above are called the positive and negative parts of μ, respectively. The measure*

$$|\mu| = \mu^+ + \mu^-$$

is called the total variation of μ. The quantity

$$\|\mu\| = |\mu|(X)$$

is called the variation or the variation norm of μ.

The decomposition $\mu = \mu^+ - \mu^-$ is called the *Jordan* or *Jordan–Hahn decomposition*.

One should not confuse the measure $|\mu|$ with the set function $A \mapsto |\mu(A)|$, which, typically, is not additive (e.g., if $\|\mu\| > \mu(X) = 0$).

We observe that the measures μ^+ and μ^- have the following properties that could be taken for their definitions:

$$\mu^+(A) = \sup\bigl\{\mu(B)\colon\ B \subset A, B \in \mathcal{A}\bigr\},$$

$$\mu^-(A) = \sup\bigl\{-\mu(B)\colon\ B \subset A, B \in \mathcal{A}\bigr\}$$

for all $A \in \mathcal{A}$. In addition,

$$|\mu|(A) = \sup\Bigl\{\sum_{n=1}^{\infty} |\mu(A_n)|\Bigr\}, \tag{3.1.3}$$

where the supremum is taken over all at most countable partitions of A into pairwise disjoint parts from $\mathcal{A}$. One can take only finite partitions and replace sup by max, since the supremum is attained at the partition $A_1 = A \cap X^+$, $A_2 = A \cap X^-$. Note that $\|\mu\|$ does not coincide with the quantity $\sup\bigl\{|\mu(A)|, A \in \mathcal{A}\bigr\}$ if both measures μ^+ and μ^- are nonzero, but one has the inequality

$$\|\mu\| \le 2\sup\bigl\{|\mu(A)|\colon\ A \in \mathcal{A}\bigr\} \le 2\|\mu\|. \tag{3.1.4}$$

All these claims are obvious from the Hahn decomposition.

3.1.5. Remark. It is seen from the proof that Theorem 3.1.1 remains valid in the case where μ is a countably additive set function on $\mathcal{A}$ with values in $(-\infty, +\infty]$. In this case, the measure μ^- is bounded and the measure μ^+ takes values in $[0, +\infty]$. Thus, in the case under consideration, the boundedness of μ is equivalent to the finiteness of $\mu(X)$.

If μ is a signed measure, we set by definition $L^p(\mu) := L^p(|\mu|)$ and $\mathcal{L}^p(\mu) := \mathcal{L}^p(|\mu|)$. For any $f \in \mathcal{L}^1(|\mu|)$ we set

$$\int_X f\,d\mu := \int_X f(x)\,\mu(dx) := \int_X f(x)\,\mu^+(dx) - \int_X f(x)\,\mu^-(dx).$$

Letting ξ be the function equal to 1 on X^+ and -1 on X^-, we obtain

$$\int_X f(x)\,\mu(dx) = \int_X f(x)\xi(x)\,|\mu|(dx).$$

It is clear that with such a definition many assertions proved above about properties of the integral are true in the case of signed measures. In particular, the Lebesgue dominated convergence theorem remains true for signed measures. Certainly, there are assertions that fail for signed measures. For example, the relation $f \le g$ gives no inequality for the integrals. In addition, the Fatou and Beppo Levi theorems fail for signed measures.

3.2. The Radon–Nikodym theorem

Let f be a function integrable with respect to a measure μ (possibly, signed or with values in $[0, +\infty]$) on a measurable space $(X, \mathcal{A})$. Then we obtain the set function

$$\nu(A) = \int_A f\,d\mu. \tag{3.2.1}$$

By the dominated convergence theorem ν is countably additive on $\mathcal{A}$. Indeed, if sets $A_n \in \mathcal{A}$ are pairwise disjoint, then the series $\sum_{n=1}^{\infty} I_{A_n}(x) f(x)$

converges for every x to $I_A(x)f(x)$, since this series may contain only one nonzero element by the disjointness of A_n. In addition,

$$\Bigl|\sum_{n=1}^{N} I_{A_n}(x)f(x)\Bigr| \le I_A(x)|f(x)|.$$

Hence this series can be integrated term-by-term.

We denote ν by $f\cdot\mu$. The function f is called the *density* of the measure ν with respect to μ (or the *Radon–Nikodym density*) and is denoted by the symbol $d\nu/d\mu$. It is clear that the measure ν is absolutely continuous with respect to μ in the sense of the following definition.

3.2.1. Definition. *Let μ and ν be countably additive measures on a measurable space $(X,\mathcal{A})$.*

(i) *The measure ν is called absolutely continuous with respect to μ if $|\nu|(A)=0$ for every set A with $|\mu|(A)=0$. Notation*: $\nu \ll \mu$.

(ii) *The measure ν is called singular with respect to μ if there exists a set $\Omega\in\mathcal{A}$ such that*

$$|\mu|(\Omega)=0 \quad \textit{and} \quad |\nu|(X\backslash\Omega)=0.$$

Notation: $\nu \perp \mu$.

This definition makes sense for measures with values in $[0,+\infty]$, too.

We observe that if a measure ν is singular with respect to μ, then μ is singular with respect to ν, i.e., $\mu\perp\nu$. For this reason, the measures μ and ν are called *mutually singular.* If $\nu\ll\mu$ and $\mu\ll\nu$, then the measures μ and ν are called *equivalent.* Notation: $\mu\sim\nu$.

The following result, called the Radon–Nikodym theorem, is one of the key facts in measure theory.

3.2.2. Theorem. *Let μ and ν be two finite measures on a space $(X,\mathcal{A})$. The measure ν is absolutely continuous with respect to the measure μ precisely when there exists a μ-integrable function f such that ν is given by* (3.2.1).

Proof. Since $\mu=f_1|\mu|$ and $\nu=f_2|\nu|$, where $|f_1(x)|=|f_2(x)|=1$, it suffices to prove the theorem for nonnegative measures μ and ν. Let $\nu\ll\mu$ and let

$$\mathcal{F} := \biggl\{ f\in\mathcal{L}^1(\mu)\colon\ f\ge 0, \int_A f\,d\mu \le \nu(A) \quad \text{for all } A\in\mathcal{A}\biggr\}.$$

Set

$$M := \sup\biggl\{\int_X f\,d\mu\colon\ f\in\mathcal{F}\biggr\}.$$

We show that $\mathcal{F}$ contains a function f on which this supremum is attained. Let us find a sequence of functions $f_n\in\mathcal{F}$ with the integrals approaching M. Let $g_n(x)=\max\bigl(f_1(x),\dots,f_n(x)\bigr)$. We observe that $g_n\in\mathcal{F}$. Indeed, the set $A \in \mathcal{A}$ can be represented in the form $A = \bigcup_{k=1}^{n} A_k$, where $A_k \in \mathcal{A}$ are pairwise disjoint and $g_n(x) = f_k(x)$ for $x \in A_k$. Then

$$\int_A g_n \, d\mu = \sum_{k=1}^{n} \int_{A_k} g_n \, d\mu \le \sum_{k=1}^{n} \nu(A_k) = \nu(A).$$

The sequence $\{g_n\}$ is increasing and the integrals of g_n are bounded by $\nu(X)$. By the monotone convergence theorem the function $f := \lim\limits_{n\to\infty} g_n$ is integrable. It is clear that $f \in \mathcal{F}$ and that the integral of f with respect to the measure μ equals M. We show that f satisfies (3.2.1). The set function

$$\eta(A) := \nu(A) - \int_A f \, d\mu$$

is a nonnegative measure due to our choice of f and is absolutely continuous with respect to μ. We have to show that $\eta = 0$. Suppose that this is not the case. Let us consider the signed measures $\eta - n^{-1}\mu$ and take their Hahn decompositions $X = X_n^+ \cup X_n^-$. Let $X_0^- := \bigcap_{n=1}^{\infty} X_n^-$. Then, by the definition of X_n^-, we have $\eta(X_0^-) \le n^{-1}\mu(X_0^-)$ for all n, whence we obtain $\eta(X_0^-) = 0$. Hence there exists n such that $\eta(X_n^+) > 0$, since otherwise $\eta(X) = \eta(X_n^-)$ for all n and then $\eta(X) = \eta(X_0^-) = 0$. For every measurable set $E \subset X_n^+$, we have $n^{-1}\mu(E) \le \eta(E)$. Hence, letting $h(x) := f(x) + n^{-1} I_{X_n^+}(x)$, we obtain for any $A \in \mathcal{A}$

$$\begin{aligned}\int_A h \, d\mu &= \int_A f \, d\mu + n^{-1}\mu(A \cap X_n^+) \le \int_A f \, d\mu + \eta(A \cap X_n^+) \\ &= \int_{A\backslash X_n^+} f \, d\mu + \nu(A \cap X_n^+) \le \nu(A\backslash X_n^+) + \nu(A \cap X_n^+) = \nu(A).\end{aligned}$$

Thus, $h \in \mathcal{F}$ contrary to the fact that the integral of h with respect to the measure μ is greater than M, since $\mu(X_n^+) > 0$. Hence $\eta = 0$. □

It is clear that the function $d\nu/d\mu$ is determined uniquely up to a set of measure zero, since a function whose integrals over all measurable sets vanish is zero a.e.

An alternative proof of the Radon–Nikodym theorem will be given in Chapter 4 (Example 4.3.3).

We note that if two measures μ and ν are finite and nonnegative and $\nu \ll \mu$, then $\nu \sim \mu$ precisely when $d\nu/d\mu > 0$ a.e. with respect to μ. It is readily verified (Exercise 3.10.32) that if we are given three measures μ_1, μ_2, and μ_3 with $\mu_1 \ll \mu_2$ and $\mu_2 \ll \mu_3$, then $\mu_1 \ll \mu_3$ and

$$d\mu_1/d\mu_3 = (d\mu_1/d\mu_2)(d\mu_2/d\mu_3).$$

The condition for the membership of the Radon–Nikodym density in the space $L^p(\mu)$ can be found in Exercise 4.7.102. Exercise 6.10.72 in Chapter 6 contains a useful assertion about a measurable dependence of the Radon–Nikodym density on a parameter.

By using the Radon–Nikodym theorem one can obtain the following Lebesgue decomposition.

3.2.3. Theorem. *Let μ and ν be two finite measures on a σ-algebra $\mathcal{A}$. Then, there exist a measure μ_0 on $\mathcal{A}$ and a μ-integrable function f such that*

$$\nu = f \cdot \mu + \mu_0, \quad \mu_0 \perp \mu.$$

PROOF. Let us consider the measure $\lambda := |\mu| + |\nu|$. By the Radon–Nikodym theorem $\mu = f_\mu \cdot \lambda$, $\nu = f_\nu \cdot \lambda$, where $f_\mu, f_\nu \in L^1(\lambda)$. Let us set $Y = \{x\colon f_\mu(x) \neq 0\}$. If $x \in Y$ we set $f(x) = f_\nu(x)/f_\mu(x)$. Finally, let $\mu_0(A) := \nu\bigl(A \cap (X\backslash Y)\bigr)$. For the restrictions μ_Y and ν_Y of the measures μ and ν to the set Y we have $\nu_Y = f \cdot \mu_Y$. Hence we obtain the required decomposition. □

It is to be noted that if μ is a finite or σ-finite nonnegative measure on a σ-algebra $\mathcal{A}$ in a space X, then every finite nonnegative measurable function f (not necessarily integrable) defines the σ-finite measure $\nu := f \cdot \mu$ by formula (3.2.1). Indeed, X is the union of the sets $\{x\colon f(x) \leq n\} \cap X_n$, where $\mu(X_n) < \infty$, which are of finite measure. It is clear that in such a form, the Radon–Nikodym theorem remains true for σ-finite measures as well. However, for the measures $\mu(\{0\}) = 1$, $\nu(\{0\}) = \infty$ (or $\mu(\{0\}) = \infty$, $\nu(\{0\}) = 1$) it is no longer true (with finite f); see also Exercise 3.10.31. On the Radon–Nikodym theorem for infinite measures and the problems that arise in this relation, see Halmos [**404**, §31].

3.3. Products of measure spaces

Let $(X_1, \mathcal{A}_1, \mu_1)$ and $(X_2, \mathcal{A}_2, \mu_2)$ be two spaces with finite nonnegative measures. On the space $X_1 \times X_2$ we consider sets of the form $A_1 \times A_2$, where $A_i \in \mathcal{A}_i$, called measurable rectangles. Let $\mu_1{\times}\mu_2(A_1{\times}A_2) := \mu_1(A_1)\mu_2(A_2)$. Extending the function $\mu_1{\times}\mu_2$ by additivity to finite unions of pairwise disjoint measurable rectangles we obtain a finitely additive function on the algebra $\mathcal{R}$ generated by such rectangles. We observe that such an extension of $\mu_1{\times}\mu_2$ to $\mathcal{R}$ is well-defined (is independent of partitions of the set into pairwise disjoint measurable rectangles), which is obvious by the additivity of μ_1 and μ_2. Finally, let $\mathcal{A}_1{\otimes}\mathcal{A}_2$ denote the σ-algebra generated by all measurable rectangles; this σ-algebra is called the product of the σ-algebras $\mathcal{A}_1$ and $\mathcal{A}_2$.

3.3.1. Theorem. *The set function $\mu_1 \times \mu_2$ is countably additive on the algebra generated by all measurable rectangles and uniquely extends to a countably additive measure, denoted by $\mu_1 \otimes \mu_2$, on the Lebesgue completion of this algebra denoted by $\mathcal{A}_1 \overline{\otimes} \mathcal{A}_2$.*

PROOF. Suppose first that $C = \bigcup_{n=1}^\infty C_n$, where

$$C = A \times B, \quad C_n = A_n \times B_n, \quad A,\, A_n \in \mathcal{A}_1, \quad B,\, B_n \in \mathcal{A}_2,$$

and the sets C_n are pairwise disjoint. Let

$$f_n(x) = \mu_2(B_n) \ \text{ if } x \in A_n, \quad f_n(x) = 0 \ \text{ if } x \notin A_n.$$

It is clear that f_n is $\mathcal{A}_1$-measurable and $\sum_{n=1}^{\infty} f_n(x) = \mu_2(B)$ for all $x \in A$. By the monotone convergence theorem we obtain

$$\sum_{n=1}^{\infty} \int_A f_n \, d\mu_1 = \int_A \mu_2(B) \, d\mu_1 = \mu_1 \times \mu_2(C).$$

Since

$$\int_A f_n \, d\mu_1 = \mu_2(B_n)\mu_1(A_n) = \mu_1 \times \mu_2(C_n),$$

our claim is proven in the regarded partial case. Now let $C = \bigcup_{n=1}^{\infty} D_n$ and let $C = \bigcup_{j=1}^{N} C_j$, where C_j are pairwise disjoint measurable rectangles and $D_n = \bigcup_{i=1}^{M_n} D_{n,i}$, where $D_{n,i}$ are pairwise disjoint measurable rectangles as well. Set $D_{n,i,j} = D_{n,i} \cap C_j$. Then $D_{n,i,j}$ are disjoint measurable rectangles and $C_j = \bigcup_n \bigcup_i D_{n,i,j}$, $D_{n,i} = \bigcup_j D_{n,i,j}$. By using our first step we obtain

$$\mu_1 \times \mu_2(C_j) = \sum_n \sum_i \mu(D_{n,i,j}), \quad \mu_1 \times \mu_2(D_{n,i}) = \sum_j \mu(D_{n,i,j}).$$

Since $\mu_1 \times \mu_2(C) = \sum_j \mu_1 \times \mu_2(C_j)$, $\quad \mu_1 \times \mu_2(D_n) = \sum_i \mu_1 \times \mu_2(D_{n,i})$, we obtain $\mu_1 \times \mu_2(C) = \sum_n \mu_1 \times \mu_2(D_n)$ by the previous equality. The assertion about extension follows by the results in §1.5. $\square$

The above-constructed measure $\mu_1 \otimes \mu_2$ is called the product of the measures μ_1 and μ_2. By construction, the measure $\mu_1 \otimes \mu_2$ is complete. Products of measures are called product measures.

It should be noted that the Lebesgue completion of the σ-algebra $\mathcal{A}_1 \otimes \mathcal{A}_2$ generated by all rectangles $A_1 \times A_2$, $A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$, is, typically, larger than this σ-algebra. For example, if $\mathcal{A}_1 = \mathcal{A}_2$ is the Borel σ-algebra of $[0,1]$, and $\mu_1 = \mu_2$ is Lebesgue measure, then $\mathcal{A}_1 \otimes \mathcal{A}_2$ coincides with the Borel σ-algebra of the square (any open set in the square is a countable union of open squares). Obviously, there exist measurable non-Borel sets in the square. It will not help if we replace the Borel σ-algebra of the interval by the σ-algebra of all Lebesgue measurable sets. In that case, as one can see from the following assertion, $\mathcal{A}_1 \otimes \mathcal{A}_2$ will not contain any nonmeasurable subset of the interval regarded as a subset of the square (clearly, such a set has measure zero in the square and belongs to the completion of $\mathcal{A}_1 \otimes \mathcal{A}_2$). Certainly, the measure $\mu_1 \otimes \mu_2$ can be considered on the not necessarily complete σ-algebra $\mathcal{A}_1 \otimes \mathcal{A}_2$.

The next result is a typical application of the monotone class theorem.

3.3.2. Proposition. (i) *Let $(X_1, \mathcal{A}_1)$ and $(X_2, \mathcal{A}_2)$ be two measurable spaces and let $\mathcal{A}_1 \otimes \mathcal{A}_2$ be the σ-algebra generated by all sets $A_1 \times A_2$ with $A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$. Then, for every $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ and every $x_1 \in X_1$, the set*

$$A_{x_1} := \bigl\{ x_2 \in X_2 \colon (x_1, x_2) \in A \bigr\}$$

is contained in $\mathcal{A}_2$. In addition, for every $\mathcal{A}_1 \otimes \mathcal{A}_2$-measurable function f and every $x_1 \in X_1$, the function $x_2 \mapsto f(x_1, x_2)$ is $\mathcal{A}_2$-measurable.

(ii) *For any finite measure ν on $\mathcal{A}_2$, the function $x_1 \mapsto \nu(A_{x_1})$ on X_1 is $\mathcal{A}_1$-measurable.*

PROOF. (i) If A is the product of two sets from $\mathcal{A}_1$ and $\mathcal{A}_2$, then our claim is true. Denote by $\mathcal{E}$ the class of all sets $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ for which it is true. Given sets A^n, one has $\bigl(\bigcup_{n=1}^\infty A^n\bigr)_x = \bigcup_{n=1}^\infty A^n_x$, and similarly for the complements. This shows that the class $\mathcal{E}$ is a σ-algebra. Hence we have $\mathcal{E} = \mathcal{A}_1 \otimes \mathcal{A}_2$. The measurability of the function $x_2 \mapsto f(x_1, x_2)$ follows if we apply the established fact to the sets $\{x_2 \colon f(x_1, x_2) < c\}$.

(ii) The function $f_A(x_1) = \nu(A_{x_1})$ is well-defined according to assertion (i). Denote by $\mathcal{E}$ the class of all sets $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ for which it is $\mathcal{A}_1$-measurable. This class contains all rectangles $A_1 \times A_2$ with $A_i \in \mathcal{A}_i$. Further, $\mathcal{E}$ is a monotone class, which follows by the dominated convergence theorem and the obvious fact that if the sets A^j increase to A, then the sets $A^j_{x_1}$ increase to A_{x_1}. Similarly, one verifies that $\mathcal{E}$ is a σ-additive class, i.e., $\mathcal{E}$ admits countable disjoint unions and $E_1 \backslash E_2 \in \mathcal{E}$ if $E_1, E_2 \in \mathcal{E}$ and $E_2 \subset E_1$. Since the class of all rectangles of the above form is closed with respect to intersections, assertion (ii) of Theorem 1.9.3 yields that the class $\mathcal{E}$ coincides with $\mathcal{A}_1 \otimes \mathcal{A}_2$. □

3.3.3. Corollary. *In the situation of assertion* (ii) *in the above proposition, for any bounded $\mathcal{A}_1 \otimes \mathcal{A}_2$-measurable function f on $X_1 \times X_2$, the following function is well-defined and $\mathcal{A}_1$-measurable:*

$$x_1 \mapsto \int_{X_2} f(x_1, x_2)\, \nu(dx_2).$$

PROOF. It suffices to consider the case where f is the indicator of a set $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$, since every bounded $\mathcal{A}_1 \otimes \mathcal{A}_2$-measurable function can be uniformly approximated by linear combinations of such indicators and the corresponding integrals in ν converge uniformly in x_1. Hence our claim follows from the proposition. □

The product of measures can be constructed by the Carathéodory method: see §3.10(i) below.

By means of the Jordan–Hahn decomposition one defines products of signed measures (this can be done directly, though). Let $\mu = \mu^+ - \mu^-$, $\nu = \nu^+ - \nu^-$, $X = X^+ \cup X^-$, $Y = Y^+ \cup Y^-$ be the Jordan–Hahn decompositions of two measures μ and ν on the spaces X and Y. Set

$$\mu \otimes \nu := \mu^+ \otimes \nu^+ + \mu^- \otimes \nu^- - \mu^+ \otimes \nu^- - \mu^- \otimes \nu^+.$$

Clearly, the measures $\mu^+ \otimes \nu^+ + \mu^- \otimes \nu^-$ and $\mu^+ \otimes \nu^- + \mu^- \otimes \nu^+$ are mutually singular, since the first one is concentrated on the set $(X^+ \times Y^+) \cup (X^- \times Y^-)$ and the second one is concentrated on the set $(X^+ \times Y^-) \cup (X^- \times Y^+)$.

By induction one defines the product of finitely many measures μ_n on the spaces $(X_n, \mathcal{A}_n)$, $n = 1, \ldots, N$. This product is associative, i.e., one has the equality

$$\mu_1 \otimes (\mu_2 \otimes \mu_3) = (\mu_1 \otimes \mu_2) \otimes \mu_3.$$

Finally, let us define the product of two σ-finite nonnegative measures μ and ν on σ-algebras $\mathcal{A}$ and $\mathcal{B}$. Let X be the union of an increasing sequence of sets X_n of finite μ-measure and let Y be the union of an increasing sequence of sets Y_n of finite ν-measure. The formula

$$\mu\otimes\nu(E) = \lim_{n\to\infty} \mu|_{X_n}\otimes\nu|_{Y_n}\bigl(E\cap(X_n\times Y_n)\bigr)$$

defines a σ-finite measure on $\mathcal{A}\otimes\mathcal{B}$.

One could reduce this case to finite measures by choosing finite measures μ_0 and ν_0 such that $\mu = \varrho_\mu\cdot\mu_0$, $\nu = \varrho_\nu\cdot\nu_0$, where ϱ_μ and ϱ_ν are nonnegative measurable functions. Then one can set $\mu\otimes\nu := (\varrho_\mu\varrho_\nu)\cdot\mu_0\otimes\nu_0$. It is readily verified that this gives the same measure as before.

Let us note yet another fact related to products of measurable spaces, which, however, does not involve measures.

3.3.4. Proposition. *Suppose that $(X,\mathcal{A})$ and $(Y,\mathcal{B})$ are measurable spaces and $f\colon X\to\mathbb{R}^1$ and $g\colon Y\to\mathbb{R}^1$ are measurable functions. Then, the mapping $(f,g)\colon X\times Y\to\mathbb{R}^2$ is measurable with respect to $\mathcal{A}\otimes\mathcal{B}$ and $\mathcal{B}(\mathbb{R}^2)$. In particular, the graph of the function f and the sets $\{(x,y)\colon y\le f(x)\}$ and $\{(x,y)\colon y\ge f(x)\}$ belong to $\mathcal{A}\otimes\mathcal{B}(\mathbb{R}^1)$.*

PROOF. Lemma 2.12.5 applies here, but a direct proof is easy. Namely, for every open rectangle $\Pi = I\times J$ the set $\bigl\{(x,y)\colon \bigl(f(x),g(y)\bigr)\in\Pi\bigr\}$ is the product of elements of $\mathcal{A}$ and $\mathcal{B}$ and belongs to $\mathcal{A}\otimes\mathcal{B}$. The class of all sets $E\in\mathcal{B}(\mathbb{R}^2)$ whose preimages with respect to the mapping (f,g) belong to $\mathcal{A}\otimes\mathcal{B}$, is a σ-algebra. Since this class contains all rectangles of the indicated form, it also contains the σ-algebra $\mathcal{B}(\mathbb{R}^2)$ generated by them. In the case where $(Y,\mathcal{B}) = \bigl(\mathbb{R}^1,\mathcal{B}(\mathbb{R}^1)\bigr)$ and $g(y)=y$, we obtain the measurability of the mapping $(x,y)\mapsto\bigl(f(x),y\bigr)$ from $X\times\mathbb{R}^1$ to $\mathbb{R}^2$, which yields the membership in $\mathcal{A}\otimes\mathcal{B}(\mathbb{R}^1)$ of the preimages of Borel sets. For example, the graph of f is the preimage of the straight line $y=x$, and two other sets mentioned in the formulation are the preimages of half-planes. $\square$

Related to this subject are Exercise 3.10.52 and Exercise 3.10.53.

3.4. Fubini's theorem

Suppose that μ and ν are finite nonnegative measures on measurable spaces $(X,\mathcal{A})$ and $(Y,\mathcal{B})$, respectively. For every set $A\subset X\times Y$, we define the sections

$$A_x = \{y\colon (x,y)\in A\},\quad A_y = \{x\colon (x,y)\in A\}.$$

3.4.1. Theorem. *Let a set $A\subset X\times Y$ be measurable with respect to the measure $\mu\otimes\nu$, i.e., belong to $(\mathcal{A}\otimes\mathcal{B})_{\mu\otimes\nu}$. Then, for μ-a.e. x, the set A_x is ν-measurable and the function $x\mapsto\nu(A_x)$ is μ-measurable; similarly, for ν-a.e. y, the set A_y is μ-measurable and the function $y\mapsto\mu(A_y)$ is ν-measurable. In addition, one has*

$$\mu\otimes\nu(A) = \int_X \nu(A_x)\,\mu(dx) = \int_Y \mu(A_y)\,\nu(dy). \tag{3.4.1}$$

PROOF. If $A = B\times C$, where $B \in \mathcal{A}$, $C \in \mathcal{B}$, then our claim is true. Hence it is true for all sets in the algebra $\mathcal{R}$ generated by measurable rectangles. By Proposition 3.3.2(ii), for any $A \in \mathcal{A}\otimes\mathcal{B}$, the functions $x \mapsto \nu(A_x)$ and $y \mapsto \mu(A_y)$ are measurable with respect to $\mathcal{A}$ and $\mathcal{B}$, respectively. Therefore, one has two set functions on $\mathcal{A}\otimes\mathcal{B}$ defined by

$$\zeta_1(A) := \int_X \nu(A_x)\,\mu(dx), \quad \zeta_2(A) := \int_Y \mu(A_y)\,\nu(dy).$$

If we are given pairwise disjoint sets A^n with the union A, then the sets A^n_x are pairwise disjoint and their union is A_x for each x, whence we obtain $\nu(A_x) = \sum_{n=1}^\infty \nu(A^n_x)$. Integrating this series term-by-term against the measure μ by the dominated convergence theorem, we conclude that ζ_1 is countably additive. Similarly, one verifies the countable additivity of ζ_2. The measures ζ_1, ζ_2 and $\mu\otimes\nu$ coincide on the algebra $\mathcal{R}$, hence also on $\mathcal{A}\otimes\mathcal{B}$.

It remains to observe that the theorem is true for every set E of $\mu\otimes\nu$-measure zero. Indeed, there exists a set $\widehat{E} \in \mathcal{A}\otimes\mathcal{B}$ that contains E and has $\mu\otimes\nu$-measure zero. Then $E_x \subset \widehat{E}_x$ and $\nu(\widehat{E}_x) = 0$ for μ-a.e. x by the already-established equality

$$\int_X \nu(\widehat{E}_x)\,\mu(dx) = 0.$$

Similarly, $\mu(E_y) = \mu(\widehat{E}_y) = 0$ for ν-a.e. y. □

3.4.2. Corollary. *The previous theorem is true in the case where μ and ν are σ-finite measures if the set A has finite measure.*

PROOF. Let us write X and Y as $X = \bigcup_{n=1}^\infty X_n$, $Y = \bigcup_{n=1}^\infty Y_n$, where X_n and Y_n are increasing sets of finite measure, then apply the above theorem to $X_n\times Y_n$ and use the monotone convergence theorem. □

3.4.3. Corollary. *Let $Y = \mathbb{R}^1$, let λ be Lebesgue measure on $\mathbb{R}^1$, and let f be a nonnegative integrable function on a measure space $(X, \mathcal{A}, \mu)$ with a σ-finite measure μ. Then*

$$\int_X f\,d\mu = \mu\otimes\lambda\Bigl(\bigl\{(x,y)\colon\ 0 \le y \le f(x)\bigr\}\Bigr). \tag{3.4.2}$$

PROOF. The set $A = \bigl\{(x,y)\colon\ 0 \le y \le f(x)\bigr\}$ is measurable with respect to $\mu\otimes\lambda$ by Proposition 3.3.4. It remains to observe that $\lambda(A_x) = f(x)$. □

We observe that if $\mu\otimes\nu(A) \ge \mu(X)\nu(Y) - \varepsilon\mu(X)$, then (3.4.1) yields the estimate

$$\mu\bigl(x\colon\ \nu(A_x) \ge \nu(Y) - \sqrt{\varepsilon}\bigr) \ge (1 - \sqrt{\varepsilon})\mu(X).$$

Indeed, the integral of the function $x \mapsto \nu(A_x)$ against the measure μ does not exceed the quantity $\nu(Y)\mu(E) + \bigl(\nu(Y) - \sqrt{\varepsilon}\bigr)\bigl(\mu(X) - \mu(E)\bigr)$, where we set $E = \{x\colon\ \nu(A_x) \ge \nu(Y) - \sqrt{\varepsilon}\}$. Hence

$$\nu(Y)\mu(X) - \sqrt{\varepsilon}\mu(X) + \sqrt{\varepsilon}\mu(E) \ge \mu(X)\nu(Y) - \varepsilon\mu(X),$$

whence we obtain $\sqrt{\varepsilon}\mu(E) \ge (\sqrt{\varepsilon} - \varepsilon)\mu(X)$.

The following important result is called *Fubini's theorem.*

3.4.4. Theorem. *Let μ and ν be σ-finite nonnegative measures on the spaces X and Y. Suppose that a function f on $X\times Y$ is integrable with respect to the product measure $\mu\otimes\nu$. Then, the function $y\mapsto f(x,y)$ is integrable with respect to ν for μ-a.e. x, the function $x\mapsto f(x,y)$ is integrable with respect to μ for ν-a.e. y, the functions*

$$x\mapsto \int_Y f(x,y)\,\nu(dy) \quad \textit{and} \quad y\mapsto \int_X f(x,y)\,\mu(dx)$$

are integrable on the corresponding spaces, and one has

$$\int_{X\times Y} f\,d(\mu\otimes\nu) = \int_Y\int_X f(x,y)\,\mu(dx)\,\nu(dy) = \int_X\int_Y f(x,y)\,\nu(dy)\,\mu(dx). \tag{3.4.3}$$

PROOF. It is clear that it suffices to prove the theorem for nonnegative functions f. Let us consider the space $X\times Y\times\mathbb{R}^1$ and the measure $\mu\otimes\nu\otimes\lambda$, where λ is Lebesgue measure. Set

$$A=\bigl\{(x,y,z)\colon\ 0\le z\le f(x,y)\bigr\}.$$

Then by Corollary 3.4.3 we obtain

$$\int_{X\times Y} f\,d(\mu\otimes\nu) = \mu\otimes\nu\otimes\lambda(A).$$

Applying Theorem 3.4.1 and using Corollary 3.4.3 once again, we arrive at the equality

$$\mu\otimes\nu\otimes\lambda(A) = \int_X \nu\otimes\lambda(A_x)\,\mu(dx) = \int_X\biggl(\int_Y f(x,y)\,\nu(dy)\biggr)\,\mu(dx).$$

Note that the measurability of all functions in these equalities is clear from Theorem 3.4.1 and the equality $f(x,y)=\lambda(A_{(x,y)})$. The second equality in (3.4.3) is proved similarly. $\square$

It is suggested in Exercise 3.10.45 to construct examples showing that the existence and equality of the repeated integrals in (3.4.3) does not guarantee the $\mu\otimes\nu$-integrability of the measurable function f. In addition, it may happen that both repeated integrals exist, but are not equal. Finally, there exist measurable functions f such that one of the repeated integrals exists, but the other one does not. However, there is an important special case when the existence of a repeated integral implies the integrability of the corresponding function on the product. This result is called *Tonelli's theorem.*

3.4.5. Theorem. *Let f be a nonnegative $\mu\otimes\nu$-measurable function on $X\times Y$, where μ and ν are σ-finite measures. Then $f\in L^1(\mu\otimes\nu)$ provided that*

$$\int_Y\int_X f(x,y)\,\mu(dx)\,\nu(dy)<\infty.$$

Proof. It suffices to prove our claim for finite measures. Let us set $f_n = \min(f, n)$. The functions f_n are bounded and measurable with respect to $\mu\otimes\nu$, hence are integrable. It is clear that $f_n \to f$ pointwise. By Fubini's theorem applied to f_n one has

$$\int_{X\times Y} f_n \, d(\mu\otimes\nu) = \int_Y \left(\int_X f_n \, d\mu\right) d\nu \le \int_Y \left(\int_X f \, d\mu\right) d\nu,$$

since $f_n(x, y) \le f(x, y)$. By Fatou's theorem f is integrable. □

It is to be noted that the existence of the repeated integrals of a function f on $X\times Y$ does not yield its measurability (Exercise 3.10.50).

Let us give another useful corollary of Fubini's theorem.

3.4.6. Corollary. *Let a function f on $X\times Y$ be measurable with respect to $\mu\otimes\nu$, where both measures are σ-finite. Suppose that for μ-a.e. x, the function $y \mapsto f(x, y)$ is integrable with respect to ν. Then, the function*

$$\Psi\colon\ x \mapsto \int_Y f(x, y)\, \nu(dy)$$

is measurable with respect to μ.

Proof. Suppose first that the measures μ and ν are bounded. Let $f_n(x, y) = f(x, y)$ if $|f(x, y)| \le n$, $f_n(x, y) = n$ if $f(x, y) \ge n$, $f_n(x, y) = -n$ if $f(x, y) \le -n$. Then, the functions f_n are measurable with respect to $\mu\otimes\nu$ and bounded, hence integrable. By Fubini's theorem the functions

$$\Psi_n(x) = \int_Y f_n(x, y)\, \nu(dy)$$

are μ-measurable. Since $f_n \to f$ pointwise and $|f_n| \le |f|$, we obtain by the dominated convergence theorem that $\Psi_n(x) \to \Psi(x)$ for all those x for which the function $y \mapsto |f(x, y)|$ is integrable with respect to ν, i.e., for μ-a.e. x. Therefore, Ψ is a μ-measurable function. In the general case, we find an increasing sequence of measurable sets $X_n\times Y_n \subset X\times Y$ of finite $\mu\otimes\nu$-measure such that the measure $\mu\otimes\nu$ is concentrated on their union. Then we use the already-known assertion for the functions

$$\Phi_n(x) = \int_{Y_n} f(x, y)\, \nu(dy)$$

and observe that $\Phi_n(x) \to \Psi(x)$ for μ-a.e. x by the dominated convergence theorem. □

It is clear that Fubini's theorem is true for signed measures, but Tonelli's theorem is not.

As an application of Fubini's theorem we shall derive a useful identity that expresses the Lebesgue integral over an abstract space in terms of the Riemann integral over $[0, +\infty)$ (in the case $p = 1$ this identity has been verified directly in Theorem 2.9.3).

3.4.7. Theorem. *Let f be a measurable function on a measure space $(X, \mathcal{A}, \mu)$ with a measure μ with values in $[0, +\infty]$. Let $1 \le p < \infty$. The function $|f|^p$ is integrable with respect to the measure μ precisely when the function*

$$t \mapsto t^{p-1}\mu\bigl(x\colon\ |f(x)| > t\bigr)$$

is integrable on $[0, +\infty)$ with respect to Lebesgue measure. In addition, one has

$$\int_X |f|^p \, d\mu = p \int_0^\infty t^{p-1}\mu\bigl(x\colon\ |f(x)| > t\bigr)\, dt. \qquad (3.4.4)$$

PROOF. Let $p = 1$. Suppose that the function f is integrable. Then our claim reduces to the case of a σ-finite measure, since μ is σ-finite on the set $\{f \neq 0\}$. Further, due to the monotone convergence theorem, we may consider only finite measures. Denote by λ Lebesgue measure on $[0, +\infty)$ and set

$$S = \bigl\{(x, y) \in X \times [0, +\infty)\colon\ y \le |f(x)|\bigr\}.$$

The integral of $|f|$ coincides with the measure of the set S with respect to $\mu \otimes \lambda$ by Corollary 3.4.3. We evaluate this measure by Fubini's theorem. For each fixed t, we have

$$S_t = \bigl\{x\colon\ (x, t) \in S\bigr\} = \bigl\{x\colon\ t \le |f(x)|\bigr\}.$$

Since the integral of $\mu(S_t)$ with respect to the argument t over $[0, +\infty)$ equals the integral of $|f|$, we arrive at (3.4.4) with $\bigl(x\colon\ |f(x)| \ge t\bigr)$ in place of $\bigl(x\colon\ |f(x)| > t\bigr)$. However, for almost all t, these two sets have equal μ-measures, since the set of all points t such that $\mu\bigl(x\colon\ |f(x)| = t\bigr) > 0$ is at most countable. Indeed, if it were uncountable, then for some $k \in \mathbb{N}$, one would have an infinite set of points t with $\mu\bigl(x\colon\ |f(x)| = t\bigr) \ge k^{-1}$, which contradicts the integrability of f.

Conversely, if the integral on the right in (3.4.4) is finite, then, for all $t > 0$, the sets $\bigl(x\colon\ |f(x)| > t\bigr)$ have finite measures. Hence, for every natural n, the function $f_n = |f| I_{\{n^{-1} \le |f| \le n\}}$ is integrable. The functions f_n have uniformly bounded integrals due to the estimate

$$\mu\bigl(x\colon\ |f_n(x)| > t\bigr) \le \mu\bigl(x\colon\ |f(x)| > t\bigr)$$

and the case considered above. By Fatou's theorem the function f is integrable. The case $p > 1$ reduces to the case $p = 1$ by the change of variable $t = s^p$ due to the equality $\bigl(x\colon\ |f(x)|^p > t\bigr) = \bigl(x\colon\ |f(x)| > t^{1/p}\bigr)$. Here it suffices to have the change of variable formula for the Riemann integral, but, certainly, an analogous formula for the Lebesgue integral can be applied; see (3.7.6) and a more general assertion in Exercise 5.8.44. $\square$

3.5. Infinite products of measures

Let $(X_\alpha, \mathcal{A}_\alpha, \mu_\alpha)$ be a family of probability spaces, indexed by elements of some infinite set $\mathfrak{A}$. The goal of this section is to define the infinite product of measures μ_α on the space $X = \prod_\alpha X_\alpha$ that consists of all collections

$x = (x_\alpha)_{\alpha\in\mathfrak{A}}$, where $x_\alpha \in X_\alpha$. Let $\bigotimes_\alpha \mathcal{A}_\alpha$ (or just $\bigotimes \mathcal{A}_\alpha$) denote the smallest σ-algebra containing all products of the form $\prod_\alpha A_\alpha$, where $A_\alpha \in \mathcal{A}_\alpha$ and only finitely many sets A_α may differ from X_α. In other words, $\bigotimes_\alpha \mathcal{A}_\alpha$ is the σ-algebra generated by all sets of the form $C\times\prod_{\alpha\notin\{\alpha_1,\ldots,\alpha_n\}} X_\alpha$, where $C \in \mathcal{A}_{\alpha_1}\otimes\cdots\otimes\mathcal{A}_{\alpha_n}$. Sets of such a form are called *cylindrical* or *cylinders*.

We start with countable products of probability measures μ_n on measurable spaces $(X_n, \mathcal{A}_n)$. Let $\mathcal{A} = \bigotimes_{n=1}^\infty \mathcal{A}_n$ be the σ-algebra generated by sets of the form $A_1\times\cdots\times A_n\times X_{n+1}\times X_{n+2}\times\cdots$, where $A_i \in \mathcal{A}_i$. It is clear that $\mathcal{A}$ is the smallest σ-algebra containing all σ-algebras

$$\mathcal{E}_n := \Bigl\{A = C\times X_{n+1}\times X_{n+2}\times\cdots:\ C \in \bigotimes_{i=1}^n \mathcal{A}_i\Bigr\}.$$

The union of all $\mathcal{E}_n$ is an algebra denoted by $\mathcal{A}^0$. On $\mathcal{A}^0$ we have a set function

$$\mu\colon\ A = C\times X_{n+1}\times X_{n+2}\times\cdots \mapsto \mu_1\otimes\cdots\otimes\mu_n(C), \quad A \in \mathcal{E}_n.$$

This set function is well-defined: if A is regarded as an element of $\mathcal{E}_k$ with $k > n$, then the value of $\mu(E)$ is unchanged. This is seen from the equality $\mu_n(X_n) = 1$. By using the already-established countable additivity of finite products we obtain the finite additivity of μ. In fact, μ is countably additive, which is not obvious and is verified in the following theorem.

3.5.1. Theorem. *The set function μ on the algebra $\mathcal{A}^0$ is countably additive and hence uniquely extends to a countably additive measure on the σ-algebra $\mathcal{A}$.*

Proof. Let A_k be decreasing sets in $\mathcal{A}^0$ with the empty intersection. We have to show that $\mu(A_k) \to 0$. We suppose that $\mu(A_k) > \varepsilon > 0$ for all n and arrive at a contradiction by showing that the intersection of the sets A_k is nonempty. Let $\mathcal{A}^n$ denote the algebra of sets in $\prod_{i=n+1}^\infty X_i$ defined by analogy with $\mathcal{A}^0$ and let $\mu^{(n)}$ be the set function on $\mathcal{A}^n$ corresponding to the product of the measures $\mu_{n+1},\ \mu_{n+2},\ldots$ by analogy with μ. By the properties of finite products it follows that, for every set $A \in \mathcal{A}^0$ and every fixed $(x_1,\ldots,x_n) \in \prod_{i=1}^n X_i$, the section

$$A^{x_1,\ldots,x_n} = \Bigl\{(z_{n+1}, z_{n+2},\ldots) \in \prod_{i=n+1}^\infty X_i\colon\ (x_1,\ldots,x_n,z_{n+1},\ldots) \in A\Bigr\}$$

belongs to $\mathcal{A}^n$ and the function

$$(x_1,\ldots,x_n) \mapsto \mu^{(n)}\bigl(A^{x_1,\ldots,x_n}\bigr)$$

is measurable with respect to $\bigotimes_{i=1}^n \mathcal{A}_i$. Denote by B_1^k the set of all points x_1 such that

$$\mu^{(1)}\bigl(A_k^{x_1}\bigr) > \varepsilon/2.$$

Then $B_1^k \in \mathcal{A}_1$ and $\mu_1\bigl(B_1^k\bigr) > \varepsilon/2$, which follows by Fubini's theorem for finite products and the inequality $\mu(A_k) > \varepsilon$. Indeed, $A_k = C_m\times X_{m+1}\times\cdots$

for some m, whence one has $\mu(A_k) = \bigotimes_{i=1}^m \mu_i(C_m)$. By Fubini's theorem we obtain

$$\varepsilon < \mu(A_k) \le \mu_1\bigl(B_1^k\bigr) + \frac{\varepsilon}{2}\mu_1\bigl(X_1\backslash B_1^k\bigr) \le \mu_1\bigl(B_1^k\bigr) + \frac{\varepsilon}{2},$$

which yields the necessary estimate. The sequence of sets B_1^k is decreasing as k is increasing and has the nonempty intersection B_1, since μ_1 is a countably additive measure and $\mu_1\bigl(B_1^k\bigr) > \varepsilon/2$. Let us fix an arbitrary point $x_1 \in B_1$ and repeat the described procedure for the decreasing sets $A_k^{x_1}$ in place of A_k. This is possible, since $\mu^{(1)}\bigl(A_k^{x_1}\bigr) > \varepsilon/2$. We obtain a point $x_2 \in X_2$ such that $\mu^{(2)}(A_k^{x_1,x_2}) > \varepsilon/4$ for all k. We continue this process inductively. After the nth step we obtain a collection $(x_1, \ldots, x_n) \in \prod_{i=1}^n X_i$ such that $\mu^{(n)}\bigl(A_k^{x_1,\ldots,x_n}\bigr) > \varepsilon 2^{-n}$ for all k. Therefore, our construction can be continued, which gives a point $x = (x_1, \ldots, x_n, \ldots)$ belonging to all A_k. Indeed, let us fix k and write A_k as $A_k = C_m \times X_{m+1} \times \cdots$. The set $A_k^{x_1,\ldots,x_m}$ is nonempty, i.e., there exists a point $(z_{m+1}, z_{m+2}, \ldots) \in \prod_{i=m+1}^\infty X_i$ such that $(x_1, \ldots, x_m, z_{m+1}, z_{m+2}, \ldots) \in A_k$. Then $(x_1, \ldots, x_m, x_{m+1}, x_{m+2}, \ldots) \in A_k$, which is obvious from the above representation of A_k. □

We now extend the above result to arbitrary infinite products. This is very simple due to the following lemma. To ease the notation we identify all sets in the product $\prod_{n=1}^\infty X_{\alpha_n}$ of a part of spaces X_α with subsets in the product of all spaces X_α by adding the spaces $X_{\alpha'}$ as factors for all missing indices $\alpha' \in \mathfrak{A}$.

3.5.2. Lemma. *The union of the σ-algebras $\bigotimes_{n=1}^\infty \mathcal{A}_{\alpha_n}$ over all countable subsets $\mathfrak{A}' = \{\alpha_n\} \subset \mathfrak{A}$ coincides with the σ-algebra $\bigotimes_\alpha \mathcal{A}_\alpha$.*

Proof. It is clear that the indicated union (taking into account the above identification) belongs to $\bigotimes_\alpha \mathcal{A}_\alpha$. So, it suffices to observe that it is a σ-algebra. This is seen from the fact that any countable family of sets in this union is determined by an at most countable family of indices, hence belongs to one of the σ-algebras that we consider in the above union. □

It is clear from this lemma that on $\bigotimes_\alpha \mathcal{A}_\alpha$ we have a well-defined countably additive measure μ that to any set A in a σ-algebra $\bigotimes_{n=1}^\infty \mathcal{A}_{\alpha_n}$ assigns its already-defined measure with respect to $\bigotimes_{n=1}^\infty \mu_{\alpha_n}$. The Lebesgue completion of this measure will be called the product of the measures μ_α and denoted by the symbol $\bigotimes_\alpha \mu_\alpha$. It is readily verified that if the whole set of indices $\mathfrak{A}$ is split into two parts $\mathfrak{A}_1$ and $\mathfrak{A}_2$ that yield the products $\mu_1 = \bigotimes_{\alpha\in\mathfrak{A}_1} \mu_\alpha$ and $\mu_2 = \bigotimes_{\alpha\in\mathfrak{A}_2} \mu_\alpha$, then $\mu_1 \otimes \mu_2 = \bigotimes_{\alpha\in\mathfrak{A}} \mu_\alpha$.

We have seen that the product of an arbitrary family of probability measures is countably additive. In the case where these measures have compact approximating classes, this fact can be verified even more simply if we apply the following lemma, which may be of independent interest. This lemma shows that the product measure on the algebra of cylindrical sets has a compact approximating class that consists of countable intersections of finite unions of cylinders with "compact" bases, hence by Theorem 1.4.3 is countably additive.

3.5.3. Lemma. *Suppose that, for every $\alpha \in \mathfrak{A}$, we are given a compact class $\mathcal{K}_\alpha$ of subsets of the space X_α. Then, the class of at most countable intersections of finite unions of finite intersections of cylindrical sets of the form $K_\alpha \times \prod_{\beta\neq\alpha} X_\beta$, $K_\alpha \in \mathcal{K}_\alpha$, is compact as well.*

PROOF. According to Proposition 1.12.4 it suffices to verify the compactness of the class of cylinders of the form $C = K_\alpha \times \prod_{\beta\neq\alpha} X_\beta$, $K_\alpha \in \mathcal{K}_\alpha$. Suppose we have a countable family of such cylinders C_i with bases $K^{(i)}_{\alpha_i} \in \mathcal{K}_{\alpha_i}$. Their intersection has the form $\bigl(\prod_{\alpha\in S} Q_\alpha\bigr) \times \bigl(\prod_{\beta\notin S} X_\beta\bigr)$, where $S = \{\alpha_i\}$, $Q_\alpha = \bigcap_{i\colon\, \alpha_i=\alpha} K^{(i)}_{\alpha_i}$. If this intersection is empty, then so is one of the sets Q_α. By the compactness of the class $\mathcal{K}_\alpha$, there exists n such that $K^{(1)}_\alpha \cap\cdots\cap K^{(n)}_\alpha = \varnothing$. Then $C_1 \cap\cdots\cap C_n = \varnothing$. □

3.5.4. Corollary. *Suppose that the probability space $(X_\alpha, \mathcal{A}_\alpha, \mu_\alpha)$ has a compact approximating class $\mathcal{K}_\alpha$ for every $\alpha \in \mathfrak{A}$. Then, the measure $\bigotimes_{\alpha\in\mathfrak{A}} \mu_\alpha$ on the algebra of cylindrical sets is approximated by the compact class described in Lemma* 3.5.3.

PROOF. For every set $A_1 \times\cdots\times A_n$, where $A_i \in \mathcal{A}_{\alpha_i}$, and every $\varepsilon > 0$, there exist sets $K_i \in \mathcal{K}_{\alpha_i}$ such that $K_i \subset A_i$ and $\mu_{\alpha_i}(A_i\backslash K_i) < \varepsilon/n$. Then we have

$$\mu\Bigl(\Bigl(\prod_{i=1}^n A_i \backslash \prod_{i=1}^n K_i\Bigr) \times \prod_{\alpha\notin\{\alpha_1,\ldots,\alpha_n\}} X_\alpha\Bigr)$$
$$\le \sum_{i=1}^n \bigotimes_{i=1}^n \mu_{\alpha_i}\Bigl((A_i\backslash K_i)\times \prod_{j\le n, j\neq i} X_j\Bigr) = \sum_{i=1}^n \mu_i(A_i\backslash K_i) < \varepsilon.$$

Along with the lemma this yields our assertion because every cylindrical set can be approximated from inside by countable intersections of finite unions of such products, which follows by Corollary 1.5.8. □

3.6. Images of measures under mappings

Suppose we are given two spaces X and Y with σ-algebras $\mathcal{A}$ and $\mathcal{B}$ and an $(\mathcal{A}, \mathcal{B})$-measurable mapping $f\colon X \to Y$. Then, for any bounded (or bounded from below) measure μ on $\mathcal{A}$, the formula

$$\mu\circ f^{-1}\colon\ B \mapsto \mu\Bigl(f^{-1}(B)\Bigr), \quad B \in \mathcal{B},$$

defines a measure on $\mathcal{B}$ called the *image of the measure* μ under the mapping f. The countable additivity of $\mu\circ f^{-1}$ follows by the countable additivity of μ.

3.6.1. Theorem. *Let μ be a nonnegative measure. A $\mathcal{B}$-measurable function g on Y is integrable with respect to the measure $\mu\circ f^{-1}$ precisely when the function $g\circ f$ is integrable with respect to μ. In addition, one has*

$$\int_Y g(y)\,\mu\circ f^{-1}(dy) = \int_X g\bigl(f(x)\bigr)\,\mu(dx). \tag{3.6.1}$$

PROOF. For the indicators of sets in $\mathcal{B}$, formula (3.6.1) is just the definition of the image measure, hence by linearity it extends to simple functions. Next, this formula extends to bounded $\mathcal{B}$-measurable functions, since such functions are uniform limits of simple ones. If g is a nonnegative $\mathcal{B}$-measurable function that is integrable with respect to $\mu\circ f^{-1}$, then for the functions $g_n=\min(g,n)$ equality (3.6.1) is already established. By the monotone convergence theorem, it remains true for g, since the integrals of the functions $g_n\circ f$ against the measure μ are uniformly bounded. Our reasoning also shows the necessity of the μ-integrability of $g\circ f$ for the integrability of $g\geq 0$ with respect to $\mu\circ f^{-1}$. By the linearity of (3.6.1) in g we obtain the general case. □

It is clear that equality (3.6.1) remains true for any function g that is measurable with respect to the Lebesgue completion of the measure $\mu\circ f^{-1}$ and is $\mu\circ f^{-1}$-integrable. This follows from the fact that any such function is equivalent to a $\mathcal{B}$-measurable function. The hypothesis of $\mathcal{B}$-measurability can be replaced by the measurability with respect to the σ-algebra

$$\mathcal{A}^f := \bigl\{E\subset Y\colon\ f^{-1}(E)\in\mathcal{A}\bigr\}$$

if we define the measure $\mu\circ f^{-1}$ on $\mathcal{A}^f$ by the same formula as on $\mathcal{B}$. However, the reader is warned that the σ-algebra $\mathcal{A}^f$ may be strictly larger than the Lebesgue completion of $\mathcal{B}$ with respect to $\mu\circ f^{-1}$. We shall discuss this question in Chapter 7 in the section on perfect measures.

In the case of a signed measure μ equality (3.6.1) remains valid if the function $g\circ f$ is integrable with respect to μ (this is clear from the Jordan decomposition for μ). However, the integrability of g with respect to the measure $\mu\circ f^{-1}$ does not imply the μ-integrability of $g\circ f$ (Exercise 3.10.68).

If we are given a $\mathcal{B}$-measurable real function ψ, then formula (3.6.1) enables us to represent the integral of $\psi\circ f$ as the integral of ψ against the measure $\mu\circ f^{-1}$ on the real line. For example,

$$\int_X |f(x)|^p\,\mu(dx)=\int_{\mathbb{R}}|t|^p\,\mu\circ f^{-1}(dt).$$

Let us introduce the *distribution function* of the function f:

$$\Phi_f(t):=\mu\bigl(x\colon\ f(x)<t\bigr),\quad t\in\mathbb{R}^1. \tag{3.6.2}$$

It is clear that $\Phi_f(t)=\mu\circ f^{-1}\bigl((-\infty,t)\bigr)$, i.e., Φ_f coincides with the distribution function $F_{\mu\circ f^{-1}}$ of the measure $\mu\circ f^{-1}$. In the case where μ is a probability measure, the function Φ_f is increasing, left continuous, has right limits at every point, and

$$\lim_{t\to-\infty}\Phi_f(t)=0,\quad \lim_{t\to\infty}\Phi_f(t)=1.$$

Recalling the definition of the Lebesgue–Stieltjes integral (see formula (2.12.7) in §2.12(vi)), we can write

$$\int_X\psi\bigl(f(x)\bigr)\,\mu(dx)=\int_{\mathbb{R}}\psi(t)\,d\Phi_f(t). \tag{3.6.3}$$

The following interesting observation is due to A.N. Kolmogorov.

3.6.2. Example. Suppose that μ is a probability measure and that f is a μ-measurable function with the continuous distribution function Φ_f. Then, the image of the measure μ under the mapping $\Phi_f \circ f$ is Lebesgue measure λ on $[0,1]$. In other words, $(\mu \circ f^{-1}) \circ \Phi_f^{-1} = \lambda$.

PROOF. We shall verify the second claim, which is equivalent to the first one by the definition of $\mu \circ f^{-1}$. This reduces the general case to the case where μ is an atomless measure on $\mathbb{R}^1$. It suffices to show that $\mu \circ F_\mu^{-1}([0,t]) = t$ for all $t \in [0,1)$, where F_μ is the distribution function of μ. We observe that $F_\mu^{-1}([0,t]) = (-\infty, s]$, where s is the supremum of numbers z such that $\mu\bigl((-\infty, z]\bigr) = t$. If F_μ is not strictly increasing, then the set of such numbers z may be an interval. However, in any case $\mu\bigl((-\infty, s]\bigr) = t$, which proves our assertion. $\square$

In particular, any Borel probability measure μ on the real line without points of positive measure can be transformed into Lebesgue measure on $[0,1]$ by the continuous transformation F_μ. Moreover, it is seen from our reasoning that if F_μ is strictly increasing, i.e., there are no intervals of zero μ-measure, then F_μ is a homeomorphism between $\mathbb{R}^1$ and $(0,1)$. In Chapter 9 such problems are considered in greater detail.

In the study of images of measures one often encounters the problem of measurability of images of sets. We shall later see that this is a rather subtle problem. First we give a simple sufficient condition for measurability.

3.6.3. Lemma. *Let $F\colon \mathbb{R}^n \to \mathbb{R}^n$ be a mapping satisfying the Lipschitz condition, i.e., one has $|F(x) - F(y)| \le L|x-y|$ for all x, $y \in \mathbb{R}^n$, where L is a constant. Then, for every Lebesgue measurable set $A \subset \mathbb{R}^n$, the set $F(A)$ is Lebesgue measurable.*

PROOF. It suffices to prove the lemma for bounded sets. We observe that A can be written as $A = \bigcup_{j=1}^\infty K_j \bigcup B$, where the sets K_j are compact and the set B has measure zero. Since the set $F\bigl(\bigcup_{j=1}^\infty K_j\bigr) = \bigcup_{j=1}^\infty F(K_j)$ is Borel as the union of compact sets $F(K_j)$, it suffices to verify the measurability of $F(B)$. Let $\varepsilon > 0$. We can cover B by a sequence of cubes Q_j with edge lengths r_j and the sum of measures less than ε. By the Lipschitzness of F, the set $F(Q_j)$ is contained in a ball of radius $L\sqrt{n}r_j$, hence in a cube with edge length $2L\sqrt{n}r_j$. So the measure of the union of $F(Q_j)$ does not exceed $\sum_{j=1}^\infty L^n n^{n/2} r_j^n < L^n n^{n/2}\varepsilon$. Thus, $F(B)$ has measure zero. $\square$

3.6.4. Corollary. *Every linear mapping L on $\mathbb{R}^n$ takes Lebesgue measurable sets into Lebesgue measurable sets, and $\lambda_n\bigl(L(A)\bigr) = |\det L|\,\lambda_n(A)$ for any measurable set A of finite measure. The preimage of every Lebesgue measurable set under an invertible linear mapping is Lebesgue measurable.*

PROOF. The assertions about measurability follow by Lemma 3.6.3. If L is degenerate, then the image of $\mathbb{R}^n$ is a proper linear subspace and has measure zero. Let $\det L \neq 0$. It is known from the elementary linear algebra that L can be written as a composition $L = UL_0V$, where U and V are orthogonal linear operators and L_0 is given by a diagonal matrix with strictly positive eigenvalues α_i. Since $|\det L| = \alpha_1 \cdots \alpha_n$ and the mappings U and V preserve Lebesgue measure, it remains to consider the mapping L_0. If A is a cube with edges parallel to the coordinate axes, then the equality $\lambda_n\bigl(L_0(A)\bigr) = \det L_0\, \lambda_n(A)$ is obvious. This equality extends to finite disjoint unions of such cubes, whence one obtains its validity for all measurable sets. $\square$

In Theorem 3.7.1 in the next section we shall derive a change of variable formula for nonlinear mappings.

Lemma 3.6.3 does not extend to arbitrary continuous mappings. In order to consider a counter-example, we define first the *Cantor function*, which is of interest in other respects, too (it will be used below in our discussion of relations between integration and differentiation).

3.6.5. Proposition. *There exists a continuous nondecreasing function C_0 on $[0,1]$ (the Cantor function or the Cantor staircase) such that $C_0(0) = 0$, $C_0(1) = 1$ and $C_0 = (2k-1)2^{-n}$ on the interval $J_{n,k}$ in the complement of the Cantor set C described in Example 1.7.5.*

PROOF. Having defined C_0 as explained on all intervals complementary to C, we obtain a nondecreasing function on $[0,1]\backslash C$. Let $C_0(0) = 0$ and $C_0(x) = \sup\bigl\{C_0(t)\colon\, t \notin C,\ t < x\bigr\}$ for $x \in C$. We obtain a function that assumes all the values of the form $k2^{-n}$. Hence the function C_0 has no jumps and is continuous on $[0,1]$. $\square$

3.6.6. Example. Let $f(x) = \frac{1}{2}\bigl(C_0(x) + x\bigr)$, where C_0 is the Cantor function on $[0,1]$. Then f is a continuous and one-to-one mapping of the interval $[0,1]$ onto itself, and there exists a measure zero set E in the Cantor set C such that $f(E)$ is nonmeasurable with respect to Lebesgue measure.

PROOF. It is clear that f is a continuous and one-to-one mapping of the interval $[0,1]$ onto itself. On every interval complementary to C, the function f has the form $x/2 + \text{const}$ (where the constant depends, of course, on that interval), hence it takes such an interval into an interval of half the length. Therefore, the complement of C is taken to an open set U of measure $1/2$. The set $[0,1]\backslash U$ of measure $1/2$ has a nonmeasurable subset D. It is clear that $E = f^{-1}(D) \subset C$ has measure zero and $f(E) = D$. $\square$

3.6.7. Remark. Let g be the inverse function for the function f in the previous example. Then, the set $g^{-1}(E)$ is nonmeasurable, although E has measure zero and g is a Borel function. This shows that in the definition of a Lebesgue measurable function the requirement of measurability of the preimages of Borel sets does not imply the measurability of preimages of arbitrary Lebesgue measurable sets.

We shall see below that it is the measurability of images of measure zero sets that plays a key role in the problem of measurability of images of general measurable sets.

3.6.8. Definition. *Let* $F\colon (X,\mathcal{A},\mu)\to(Y,\mathcal{B},\nu)$ *be a mapping between measure spaces. We shall say that* F *has Lusin's property* (N) (*or satisfies Lusin's condition* (N)) *with respect to the pair* (μ,ν) *if* $\nu\bigl(F(A)\bigr)=0$ *for every set* $A\in\mathcal{A}$ *with* $\mu(A)=0$.

In the case $(X,\mathcal{A},\mu)=(Y,\mathcal{B},\nu)$ *we shall say that* F *has Lusin's property* (N) *with respect to* μ.

Note that in this definition F is supposed to be defined everywhere.

3.6.9. Theorem. *Let* $F\colon \mathbb{R}^n\to\mathbb{R}^n$ *be a Lebesgue measurable mapping. Then* F *has Lusin's property* (N) *with respect to Lebesgue measure precisely when* F *takes all Lebesgue measurable sets to Lebesgue measurable sets.*

Proof. Let $A\subset\mathbb{R}^n$ be a Lebesgue measurable set. By Lusin's theorem, there exists a sequence of compact sets $K_j\subset A$ such that F is continuous on every K_j and the set $B=A\backslash\bigcup_{j=1}^\infty K_j$ has measure zero. Then the set $\bigcup_{j=1}^\infty F(K_j)$ is Borel. Hence the measurability of $F(A)$ follows from the measurability of $F(B)$ ensured by Lusin's property. The necessity of this property is clear from the fact that if B is a measure zero set and $F(B)$ has positive measure, then $F(B)$ contains a nonmeasurable subset D. Hence $E=B\cap F^{-1}(D)$ has measure zero and the nonmeasurable image. $\square$

Lusin's property (N) is further studied in exercises in Chapter 5 and in Chapter 9.

3.7. Change of variables in $\mathbb{R}^n$

We now derive the change of variables formula for nonlinear mappings on $\mathbb{R}^n$. Suppose that U is an open set in $\mathbb{R}^n$ and a mapping $F\colon U\to\mathbb{R}^n$ is continuously differentiable. The derivative $F'(x)$ (an alternative notation is $DF(x)$) of the mapping F at a point x by definition is a linear mapping on $\mathbb{R}^n$ such that $F(x+h)-F(x)=F'(x)h+o(h)$. The determinant of the matrix of this mapping is called the Jacobian of F at the point x. The Jacobian will be denoted by $JF(x)$. Thus, $JF(x)=\det F'(x)$.

3.7.1. Theorem. *If the mapping* F *is injective on* U, *then, for any measurable set* $A\subset U$ *and any Borel function* $g\in L^1(\mathbb{R}^n)$, *one has the equality*

$$\int_A g\bigl(F(x)\bigr)|JF(x)|\,dx=\int_{F(A)} g(y)\,dy. \tag{3.7.1}$$

Proof. It has been shown that the set $F(A)$ is measurable, since the mapping F is locally Lipschitzian. It is clear that it suffices to prove (3.7.1) in the case where the function g is the indicator of a Borel set B. By the injectivity of F, this reduces to establishing the equality

$$\lambda_n\bigl(F(E)\bigr) = \int_E |JF(x)|\,dx \tag{3.7.2}$$

for all Borel sets $E \subset U$. Let E be a closed cube inside U. Without loss of generality we may assume that $\|F'(x)(h)\| \le \|h\|$ for all $x \in E$ and $h \in \mathbb{R}^n$. Let us fix $\varepsilon \in (0,1)$. By the continuous differentiability of F there exists $\delta > 0$ such that whenever $x, y \in E$ and $\|x - y\| \le \delta$, we have

$$F(y) - F(x) - F'(x)(y - x) = r(x,y), \quad \|r(x,y)\| \le \varepsilon\|y - x\|. \tag{3.7.3}$$

Let us partition E into m^n equal cubes E_j with the diagonal length $d < \delta$. Let x_j be the center of E_j. Set $L_j(x) = F'(x_j)(x - x_j) + F(x_j)$ for $x \in E_j$. Then one can write $\Delta_j := \lambda_n\bigl(F(E_j)\bigr) - \lambda_n\bigl(L_j(E_j)\bigr)$, and in this notation one has

$$\begin{aligned}\lambda_n\bigl(F(E)\bigr) &= \sum_{j=1}^{m^n} \lambda_n\bigl(F(E_j)\bigr) = \sum_{j=1}^{m^n}\Bigl[\lambda_n\bigl(L_j(E_j)\bigr) + \Delta_j\Bigr] \\ &= \sum_{j=1}^{m^n} |\det F'(x_j)|\lambda_n(E_j) + \sum_{j=1}^{m^n}\Delta_j.\end{aligned}$$

It is clear that if m is infinitely increasing, the first sum on the right-hand side of this equality approaches the integral of $|JF|$ over E. Let us estimate Δ_j. By (3.7.3) for all $x \in E_j$ we have

$$\|F(x) - L_j(x)\| \le \varepsilon\|x - x_j\| \le \varepsilon d.$$

Then $F(E_j)$ belongs to the neighborhood of radius εd of the set $L_j(E_j)$. Since we assume that L_j is Lipschitzian with the constant 1, we obtain by Fubini's theorem that, denoting by C_n the number of all faces of the n-dimensional cube, the measure of the εd-neighborhood of the set $L_j(E_j)$ differs from the measure of $L_j(E_j)$ not greater than in $2C_n\varepsilon\lambda_n(E_j)$. Thus,

$$\Delta_j = \lambda_n\bigl(F(E_j)\bigr) - \lambda_n\bigl(L_j(E_j)\bigr) \le 2C_n\varepsilon\lambda_n(E_j),$$

whence we have

$$\sum_{j=1}^{m^n}\Delta_j \le 2C_n\lambda_n(E)\varepsilon.$$

Let us now show that for some constant K_n one has

$$\sum_{j=1}^{m^n}\Delta_j \ge -K_n\lambda_n(E)\sqrt{\varepsilon}.$$

To this end, we shall prove the estimate

$$\lambda_n\bigl(L_j(E_j)\bigr) - \lambda_n\bigl(F(E_j)\bigr) \le K_n\sqrt{\varepsilon}\lambda_n(E_j), \quad j = 1, \ldots, m^n. \tag{3.7.4}$$

If $|\det F'(x_j)| \le \sqrt{\varepsilon}$, then (3.7.4) is fulfilled with $K_n = 1$. Let us consider the case where $|\det F'(x_j)| > \sqrt{\varepsilon}$. Then the operator $F'(x_j)$ has the inverse G_j, and

$$\|G_j(h)\| \le \varepsilon^{-1/2}\|h\|, \quad \forall\, h \in \mathbb{R}^n. \tag{3.7.5}$$

Indeed, $F'(x_j) = TL$, where T is an orthogonal operator and the operator L has an orthonormal eigenbasis with positive eigenvalues $\alpha_1, \ldots, \alpha_n$. Due to our assumption we have $\alpha_i \le 1$. Hence $\alpha_i > \sqrt{\varepsilon}$, whence it follows that $\alpha_i^{-1} < \varepsilon^{-1/2}$, which proves (3.7.5). By (3.7.3) and (3.7.5) we conclude that $F(E_j)$ contains $L_j(Q_j)$, where Q_j is the cube with the same center as E_j and the diameter $(1-\sqrt{\varepsilon})d$. Indeed, let $y \in E_j$. We may assume that $\delta > 0$ is so small that $\big\|\big(I - G_jF'(x)\big)h\big\| \le \|h\|/2$ whenever $x \in E_j$ and $\|h\| \le 1$. Such a choice is possible, since the mapping F' is continuous and $G_jF'(x_j) = I$. The equation $F(x) = L_j(y)$ is equivalent to the equation $x - G_jF(x) + G_jL_j(y) = x$. By the above-established estimate we obtain that the mapping

$$\Psi(x) = x - G_jF(x) + G_jL_j(y)$$

satisfies the Lipschitz condition with the constant $1/2$. We observe that

$$\begin{aligned}\Psi(x) &= x - G_jF(x) + y - x_j + G_jF(x_j)\\ &= y + (x - x_j) + G_j\big(F(x_j) - F(x)\big) = y + G_j\big(r(x, x_j)\big).\end{aligned}$$

Hence $\|\Psi(x) - y\| \le \varepsilon^{-1/2}\varepsilon\|x - x_j\|$ and so $\Psi(x) \in E_j$. Thus, the mapping $\Psi\colon E_j \to E_j$ is a contraction. It is well known that there exists $x \in E_j$ with $\Psi(x) = x$, i.e., $F(x) = L_j(y)$. Therefore, in the case under consideration we obtain

$$\begin{aligned}\lambda_n\big(L_j(E_j)\big) - \lambda_n\big(F(E_j)\big) &\le \lambda_n\big(L_j(E_j)\big) - \lambda_n\big(L_j(Q_j)\big)\\ &= |\det F'(x_j)|\big[\lambda_n(E_j) - \lambda_n(Q_j)\big]\\ &= |\det F'(x_j)|\big(1 - (1-\sqrt{\varepsilon})^n\big)\lambda_n(E_j),\end{aligned}$$

which yields (3.7.4). Thus, formula (3.7.1) is established for cubes. The general case easily follows from this. □

3.7.2. Corollary. *Let F be a strictly increasing continuously differentiable function on a bounded or unbounded interval (a, b). Then, for any Borel function g integrable on $\big(F(a), F(b)\big)$, one has*

$$\int_a^b g\big(F(t)\big)F'(t)\,dt = \int_{F(a)}^{F(b)} g(s)\,ds. \tag{3.7.6}$$

In Chapter 5 we prove a change of variable formula for a broader class of functions F.

One can easily see from the proof of Theorem 3.7.1 that the following Sard inequality is true (in fact, it is true under broader hypotheses, see Chapter 5).

3.7.3. Proposition. *Let $F\colon U \to \mathbb{R}^n$ be a continuously differentiable mapping. Then, for any measurable set A, one has*

$$\lambda_n\bigl(F(A)\bigr) \le \int_A |JF(x)|\, dx. \tag{3.7.7}$$

It follows by (3.7.7) that the image of the set $\bigl\{x\colon\ JF(x) = 0\bigr\}$ under the mapping F (called the set of critical values of F) has measure zero. This assertion is the simplest case of Sard's theorem. We observe that if we prove first that the set of critical values has measure zero, then inequality (3.7.7) can be easily derived from the statement of the theorem, without looking at its proof. To this end, we consider the integral over the set, where $JF \neq 0$ and apply the inverse function theorem, which asserts that every point x with $JF(x) \neq 0$ has a neighborhood where F is injective.

Finally, let us observe that according to (3.6.1), formula (3.7.1) can be restated as the equality $(|JF| \cdot \lambda_n|_U) \circ F^{-1} = \lambda_n|_{F(U)}$, where λ_n is Lebesgue measure. Therefore, if $|JF(x)| > 0$, we obtain the equality

$$\lambda_n|_U \circ F^{-1} = \varrho \cdot \lambda_n|_{F(U)}, \quad \text{where } \varrho(x) = \bigl|JF\bigl(F^{-1}(x)\bigr)\bigr|^{-1}.$$

Indeed, for any bounded measurable function g on U, one has

$$\begin{aligned}\int_U g\bigl(F(x)\bigr)\, dx &= \int_U g\bigl(F(x)\bigr)\bigl|JF\bigl(F^{-1}F(x)\bigr)\bigr|^{-1} |JF(x)|\, dx\\ &= \int_{F(U)} g(y)\bigl|JF\bigl(F^{-1}(y)\bigr)\bigr|^{-1}\, dy.\end{aligned}$$

3.8. The Fourier transform

In this section, we consider the Fourier transform of functions and measures: one of the most efficient tools in analysis.

3.8.1. Definition. (i) *The Fourier transform of a function $f \in \mathcal{L}^1(\mathbb{R}^n)$ (possibly complex-valued) is the complex-valued function*

$$\widehat{f}(y) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-i(y,x)} f(x)\, dx.$$

The Fourier transform of an element $f \in L^1(\mathbb{R}^n)$ is the function $\widehat{f}$ for an arbitrary representative of the equivalence class of f.

(ii) *The characteristic functional (or the characteristic function) of a bounded Borel measure μ on $\mathbb{R}^n$ is the complex function*

$$\widetilde{\mu}(y) = \int_{\mathbb{R}^n} e^{i(y,x)}\, \mu(dx).$$

The necessity to distinguish versions of an integrable function when considering Fourier transforms will be clear below, when we discuss the recovery of the value of f at a given point from the function $\widehat{f}$. It is clear that if the measure μ is given by a density f with respect to Lebesgue measure, then its characteristic functional coincides up to a constant factor with the Fourier transform of its density with the reversed argument. The above definition is consistent with that adopted in probability theory of the characteristic functional of a probability measure, which is also applicable in infinite-dimensional spaces. On the other hand, our choice of a constant in the definition of the Fourier transform of functions yields the unitary operator on $L^2(\mathbb{R}^n)$ (see (3.8.3)). Finally, the minus sign in the exponent is just a tradition. We shall see below that changing the sign in the exponent we arrive at the inverse transform.

In some cases, one can explicitly evaluate Fourier transforms. Let us consider one of the most important examples.

3.8.2. Example. Let $\alpha > 0$. Then

$$\frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \exp[-i(y,x)] \exp[-\alpha|x|^2]\, dx = \frac{1}{(2\alpha)^{n/2}} \exp\Bigl[-\frac{1}{4\alpha}|y|^2\Bigr].$$

PROOF. The evaluation of this integral by Fubini's theorem reduces to the one-dimensional case, where by the obvious change of variable it suffices to consider the case $\alpha = 1/2$. In that case, both sides of the equality to be proven are analytic functions of y, equal at $y = it$, $t \in \mathbb{R}$, which follows by Exercise 3.10.47. Hence these functions coincide at all $y \in \mathbb{R}$. $\square$

3.8.3. Definition. *A function $\varphi\colon \mathbb{R}^n \to \mathbb{C}$ is called positive definite if, for all $y_i \in \mathbb{R}^n$, $c_i \in \mathbb{C}$, $i = 1, \ldots, k$, one has $\sum_{i,j=1}^k c_i \overline{c_j} \varphi(y_i - y_j) \geq 0$.*

It follows by the above example that the function $\exp(-\beta|y|^2)$ on $\mathbb{R}^n$ is positive definite for all $\beta \geq 0$. We observe that the function

$$p_\sigma(x) = \frac{1}{(2\pi\sigma)^{n/2}} \exp\Bigl(-\frac{|x|^2}{2\sigma}\Bigr)$$

for any $\sigma > 0$ has the integral 1. A probability measure with density p_σ has the characteristic functional $\exp(-\sigma|y|^2/2)$. The probability measure with density p_1 is called the standard Gaussian measure on $\mathbb{R}^n$. The theory of Gaussian measures is presented in the book Bogachev [**105**].

Properties of positive definite functions are discussed below in §3.10(v).

3.8.4. Proposition. (i) *The Fourier transform of any integrable function f is a bounded uniformly continuous function and $\lim\limits_{|y|\to\infty} \widehat{f}(y) = 0$.*

(ii) *The characteristic functional of any bounded measure μ is a uniformly continuous bounded function. If the measure μ is nonnegative, then the function $\widetilde{\mu}$ is positive definite.*

PROOF. (i) It is clear that $|\widehat{f}(y)| \leq (2\pi)^{-n/2}\|f\|_{L^1}$. If f is the indicator of a cube with edges parallel to the coordinate axes, then $\widehat{f}$ is easily evaluated by Fubini's theorem, and the claim is true. So this claim is true for linear combinations of the indicators of such cubes. Now it remains to take a sequence f_j of such linear combinations that converges to f in $L^1(\mathbb{R}^n)$, and observe that the functions $\widehat{f_j}$ converge uniformly to $\widehat{f}$.

(ii) The first assertion is proved similarly to (i). The second one follows by the equality

$$\sum_{i,j=1}^{k} c_i\overline{c_j}\widetilde{\mu}(y_i - y_j) = \int_{\mathbb{R}^n}\Bigl|\sum_{j=1}^{k} c_j e^{i(y_j,x)}\Bigr|^2 \mu(dx),$$

which is readily verified. □

Let us consider several other useful properties of the Fourier transform.

3.8.5. Proposition. *Let f be a continuously differentiable and integrable function on $\mathbb{R}^n$ and let its partial derivative $\partial_{x_j} f$ be integrable. Then*

$$\widehat{\partial_{x_j} f}(y) = iy_j\widehat{f}(y).$$

PROOF. If f has bounded support, then this equality follows by the integration by parts formula. In order to reduce to this the general case, it suffices to take a sequence of smooth functions ζ_k on $\mathbb{R}^n$ with the following properties: $0 \le \zeta_k \le 1$, $\sup_k |\partial_{x_j}\zeta_k| \le C$, $\zeta_k(x) = 1$ if $|x| \le k$. Then the functions $\zeta_k f$ converge in $L^1(\mathbb{R}^n)$ to f, and the functions $\partial_{x_j}(\zeta_k f)$ converge to $\partial_{x_j} f$, since $f\partial_{x_j}\zeta_k \to 0$ in $L^1(\mathbb{R}^n)$ by the dominated convergence theorem. □

It follows that if f is a smooth function with bounded support, then its Fourier transform decreases at infinity faster than any power.

3.8.6. Proposition. *If two bounded Borel measures have equal Fourier transforms, then they coincide. In particular, two integrable functions with equal Fourier transforms are equal almost everywhere.*

PROOF. It suffices to show that any bounded measure μ with the identically zero Fourier transform equals zero. In turn, it suffices to prove that every bounded continuous function f has the zero integral with respect to the measure μ (see Exercise 3.10.29). We may assume that $\|\mu\| \le 1$ and $|f| \le 1$. Let $\varepsilon \in (0,1)$. We take a continuous function f_0 with bounded support such that $|f_0| \le 1$ and

$$\int_{\mathbb{R}^n} |f(x) - f_0(x)|\,|\mu|(dx) \le \varepsilon.$$

Next we find a cube $K = [-\pi k, \pi k]^n$, $k \in \mathbb{N}$, containing the support of f_0 such that $|\mu|(\mathbb{R}^n\backslash K) < \varepsilon$. By the Weierstrass theorem, there exists a function g of the form $g(x) = \sum_{j=1}^m c_j \exp[i(y_j, x)]$, where y_j are vectors with coordinates of the form l/k, such that $|f_0(x) - g(x)| < \varepsilon$ for all $x \in K$. By the periodicity of g we have $|g(x)| \le 1 + \varepsilon \le 2$ for all $x \in \mathbb{R}^n$. The integral of g against the measure μ vanishes by the equality $\widetilde{\mu} = 0$. Finally, we obtain

$$\begin{aligned}\Bigl|\int_{\mathbb{R}^n} f\,d\mu\Bigr| &\le \varepsilon + \Bigl|\int_{\mathbb{R}^n} f_0\,d\mu\Bigr| \le \varepsilon + \Bigl|\int_{\mathbb{R}^n} [f_0 - g]\,d\mu\Bigr| \\ &\le 2\varepsilon + \int_{\mathbb{R}^n\backslash K} |g|\,d|\mu| \le 4\varepsilon.\end{aligned}$$

Since ε is arbitrary, our claim is proven. Note that one could also apply Theorem 2.12.9, by taking for $\mathcal{H}_0$ the algebra of linear combinations of the functions $\sin(y,x)$ and $\cos(y,x)$, and for $\mathcal{H}$ the space of bounded Borel functions having the zero integral with respect to the measure μ. The second assertion follows by the first one, since we obtain the equality almost everywhere of the considered functions with the reversed arguments. □

3.8.7. Corollary. *A bounded Borel measure on $\mathbb{R}^n$ is invariant under the mapping $x \mapsto -x$ precisely when $\widetilde{\mu}$ is a real function. In particular, an integrable function is symmetric or even (i.e., $f(x) = f(-x)$ a.e.) precisely when its Fourier transform is real.*

PROOF. The necessity of the indicated condition is obvious, since $\sin x$ is an odd function. The sufficiency is clear from the fact that the characteristic functional of the measure ν that is the image of μ under the central symmetry equals the complex conjugated function of $\widetilde{\mu}$, i.e., coincides with that function, since it is real. The coincidence of the characteristic functionals yields the equality of the measures. □

It is natural to ask how one can recover a function f from its Fourier transform determining the function up to a modification. For this purpose one uses the *inverse Fourier transform.* For any integrable function f, the inverse Fourier transform is defined by the formula

$$\check{f}(x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i(y,x)} f(y)\, dy.$$

We shall see that if the direct Fourier transform of f is integrable, then its inverse transform gives the initial function f. In fact, this is true even without the assumption of integrability of $\widehat{f}$ if one defines the inverse Fourier transform for generalized functions (distributions). We shall not do this, but only prove a sufficient condition for recovering a function at a given point from its Fourier transform, and then we prove the Parseval equality, upon which the definition of the Fourier transform of generalized functions is based.

3.8.8. Theorem. *Suppose that a function f is integrable on the real line and that at some point x it satisfies the Dini condition: the function*

$$t \mapsto [f(x+t) - f(x)]/t$$

is integrable in some neighborhood of the origin. Then the following inversion formula is true:

$$f(x) = \lim_{R\to+\infty} \frac{1}{\sqrt{2\pi}} \int_{-R}^{R} e^{ixy} \widehat{f}(y)\, dy. \tag{3.8.1}$$

In particular, this formula is true at all points of differentiability of f.

PROOF. Set

$$J_R := \frac{1}{\sqrt{2\pi}} \int_{-R}^{R} e^{ixy} \widehat{f}(y)\, dy,$$

where $R > 0$. By using Fubini's theorem and the change of variable $z = t + x$ we obtain

$$\begin{aligned} J_R &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} f(z)\int_{-R}^{R} e^{iy(x-z)}\,dy\,dz \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} f(z)\frac{2\sin\bigl(R(x-z)\bigr)}{x-z}\,dz = \frac{1}{\pi}\int_{-\infty}^{+\infty} f(t+x)\frac{\sin(Rt)}{t}\,dt. \end{aligned}$$

It is known from the elementary calculus that

$$\lim_{T\to+\infty}\int_{-T}^{T}\frac{\sin t}{t}\,dt = \pi.$$

Let $\varepsilon > 0$. Since the integral of $\sin(Rt)/t$ over $[-T, T]$ equals the integral of $\sin t/t$ over $[-RT, RT]$, there exists $T_1 > 1$ such that for all $T > T_1$ and $R > 1$ one has

$$\Bigl|\frac{f(x)}{\pi}\int_{-T}^{T}\frac{\sin(Rt)}{t}\,dt - f(x)\Bigr| < \frac{\varepsilon}{3}.$$

By the integrability of f, there exists $T_2 > T_1$ such that

$$\int_{\{|t|\ge T_2\}}\frac{|f(x+t)|}{|t|}\,dt \le \int_{|t|\ge T_2}|f(x+t)|\,dt < \varepsilon.$$

By our hypothesis, the function $\varphi(t) = [f(x+t) - f(x)]/t$ is integrable over $[-T_2, T_2]$. Hence the Fourier transform of the function $\varphi I_{[-T_2,T_2]}$ tends to zero at the infinity. Therefore, there exists $R_1 > 1$ such that for all $R > R_1$ one has

$$\Bigl|\int_{-T_2}^{T_2}\sin(Rt)\frac{f(x+t)-f(x)}{t}\,dt\Bigr| < \frac{\varepsilon}{3}.$$

Taking into account the three estimates above we obtain for all $R > R_1$

$$\begin{aligned} |J_R - f(x)| &\le \Bigl|J_R - \frac{f(x)}{\pi}\int_{-T_2}^{T_2}\frac{\sin(Rt)}{t}\,dt\Bigr| \\ &+ \Bigl|\frac{f(x)}{\pi}\int_{-T_2}^{T_2}\frac{\sin(Rt)}{t}\,dt - f(x)\Bigr| \le \Bigl|J_R - \frac{f(x)}{\pi}\int_{-T_2}^{T_2}\frac{\sin(Rt)}{t}\,dt\Bigr| + \frac{\varepsilon}{3} \\ &= \frac{1}{\pi}\Bigl|\int_{-\infty}^{+\infty} f(t+x)\frac{\sin(Rt)}{t}\,dt - \int_{-T_2}^{T_2} f(x)\frac{\sin(Rt)}{t}\,dt\Bigr| + \frac{\varepsilon}{3} \\ &\qquad\le \frac{1}{\pi}\Bigl|\int_{-T_2}^{T_2}[f(t+x)-f(x)]\frac{\sin(Rt)}{t}\,dt\Bigr| \\ &\qquad+ \frac{1}{\pi}\Bigl|\int_{\{|t|\ge T_2\}} f(t+x)\frac{\sin(Rt)}{t}\,dt\Bigr| + \frac{\varepsilon}{3} < \frac{\varepsilon}{\pi} + \frac{\varepsilon}{\pi} + \frac{\varepsilon}{3} < \varepsilon. \end{aligned}$$

The theorem is proven. □

3.8.9. Corollary. *Let f be an infinitely differentiable function on $\mathbb{R}^n$ with bounded support. Then*

$$f(x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i(y,x)} \widehat{f}(y)\, dy. \tag{3.8.2}$$

PROOF. We recall that the function $\widehat{f}$ decreases at infinity faster than any power, hence it is integrable. So in the case $n = 1$ equality (3.8.2) follows by (3.8.1). The case $n > 1$ follows by Fubini's theorem. In order to simplify notation we consider the case $n = 2$. Then, for any fixed x_2, we have

$$f(x_1, x_2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ix_1 y_1} g_1(y_1, x_2)\, dy_1,$$

where $y_1 \mapsto g_1(y_1, x_2)$ is the Fourier transform of the function of a single variable $x_1 \mapsto f(x_1, x_2)$. For every fixed y_1, the function $x_2 \mapsto g_1(y_1, x_2)$ is infinitely differentiable and has bounded support. Hence

$$\begin{aligned} g_1(y_1, x_2) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ix_2 y_2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-iy_2 z_2} g_1(y_1, z_2)\, dz_2\, dy_2 \\ &= \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{+\infty} e^{ix_2 y_2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-iy_1 z_1 - iy_2 z_2} f(z_1, z_2)\, dz_1\, dz_2\, dy_2, \end{aligned}$$

which yields (3.8.2). $\square$

3.8.10. Theorem. *For all $\varphi, \psi \in \mathcal{L}^1(\mathbb{R}^n)$, one has*

$$\int_{\mathbb{R}^n} \widehat{\varphi}\overline{\psi}\, dx = \int_{\mathbb{R}^n} \varphi\overline{\check{\psi}}\, dx, \quad \int_{\mathbb{R}^n} \check{\psi}\overline{\varphi}\, dx = \int_{\mathbb{R}^n} \psi\overline{\widehat{\varphi}}\, dx.$$

PROOF. We recall that $\widehat{\varphi}$ and $\check{\psi}$ are bounded functions. By applying Fubini's theorem to the equality

$$\int_{\mathbb{R}^n} \widehat{\varphi}\overline{\psi}\, dx = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-i(x,y)} \varphi(y)\overline{\psi(x)}\, dy\, dx,$$

we obtain the first formula and the second one is similar. $\square$

3.8.11. Corollary. *Let $\varphi \in \mathcal{L}^1(\mathbb{R}^n)$. Then, for every infinitely differentiable function ψ with bounded support, the following Parseval equality is true*:

$$\int_{\mathbb{R}^n} \varphi(x)\overline{\psi(x)}\, dx = \int_{\mathbb{R}^n} \widehat{\varphi}(y)\overline{\widehat{\psi}(y)}\, dy. \tag{3.8.3}$$

PROOF. As noted above, the function $f := \widehat{\psi}$ decreases faster than any power and is integrable. It remains to apply the inversion formula $\psi = \check{f}$. $\square$

The Parseval equality enables one to define the Fourier transform on L^2 (see Exercise 3.10.76).

3.8.12. Corollary. *Let $f \in \mathcal{L}^1(\mathbb{R}^n)$ and $\widehat{f} \in \mathcal{L}^1(\mathbb{R}^n)$. Then f has a continuous modification f_0 and*

$$f_0(x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i(y,x)} \widehat{f}(y)\, dy, \quad \forall\, x \in \mathbb{R}^n.$$

PROOF. By hypothesis, the function $g := \widehat{f}$ is integrable. Hence its inverse Fourier transform f_0 is continuous. Let us verify that $f = f_0$ a.e. To this end, it suffices to show that, for each smooth real function φ with bounded support, one has

$$\int_{\mathbb{R}^n} f\varphi\, dx = \int_{\mathbb{R}^n} f_0\varphi\, dx.$$

By the Parseval equality we have

$$\int f\varphi\, dx = \int \widehat{f}\,\overline{\widehat{\varphi}}\, dx.$$

On the other hand,

$$\int g\overline{\widehat{\varphi}}\, dx = \int f_0\varphi\, dx,$$

whence the assertion follows. □

Fubini's theorem can also be applied to the product of two bounded Borel measures μ and ν on $\mathbb{R}^n$. This gives the following assertion.

3.8.13. Proposition. *Let μ and ν be two bounded Borel measures on $\mathbb{R}^n$. Then one has*

$$\int_{\mathbb{R}^n} \widetilde{\mu}(y)\, \nu(dy) = \int_{\mathbb{R}^n} \widetilde{\nu}(x)\, \mu(dx). \tag{3.8.4}$$

3.8.14. Corollary. *Let μ and ν be two Borel probability measures on $\mathbb{R}^n$. If the function $\widetilde{\nu}$ is real, then*

$$\mu\bigl(x\colon\ \widetilde{\nu}(x) \le t\bigr) \le \frac{1}{1-t} \int_{\mathbb{R}^n} \bigl[1 - \widetilde{\mu}(y)\bigr]\, \nu(dy), \quad \forall\, t \in (0,1), \tag{3.8.5}$$

where the right-hand side is real.

PROOF. The left-hand side equals $\mu\bigl(x\colon\ 1 - \widetilde{\nu}(x) \ge 1 - t\bigr)$, which by Chebyshev's inequality is majorized by

$$\frac{1}{1-t} \int_{\mathbb{R}^n} \bigl[1 - \widetilde{\nu}(x)\bigr]\, \mu(dx).$$

Now we apply (3.8.4), which also shows that the right-hand side of (3.8.5) is real. □

It should be emphasized that the function $\widetilde{\mu}$ itself may not be real; it is only claimed that its integral against the measure ν is real.

3.8.15. Corollary. *For any Borel probability measure μ on $\mathbb{R}^n$ one has*

$$\mu\bigl(x\colon\ |x|\ge t\bigr)\le \frac{\sqrt{e}}{\sqrt{e}-1}\int_{\mathbb{R}^n}\bigl[1-\widetilde{\mu}(y/t)\bigr]\,\gamma(dy),\quad \forall\, t>0, \tag{3.8.6}$$

where γ is the standard Gaussian measure on $\mathbb{R}^n$.

Proof. We know that $\widetilde{\gamma}(x)=\exp\bigl(-|x|^2/2\bigr)$. Let γ_t be the image of γ under the mapping $x\mapsto x/t$. Then $\widetilde{\gamma_t}(x)=\exp\bigl(-t^{-2}|x|^2/2\bigr)$. Therefore, by (3.8.5), we obtain

$$\mu\bigl(x\colon\ |x|\ge t\bigr)=\mu\bigl(x\colon\ \widetilde{\gamma_t}(x)\le e^{-1/2}\bigr)\le \frac{1}{1-e^{-1/2}}\int_{\mathbb{R}^n}[1-\widetilde{\mu}(y)]\,\gamma_t(dy).$$

The right-hand side of this inequality equals the right-hand side of (3.8.6) by the definition of γ_t. □

3.8.16. Corollary. *Let $r>0$ and let μ be a probability measure on $\mathbb{R}^n$. Then one has*

$$\mu\bigl(x\colon\ |x|\ge r^{-2}\bigr)\le 6nr^2+3\sup_{|z|\le r}|1-\widetilde{\mu}(z)|. \tag{3.8.7}$$

Proof. The left-hand side of (3.8.7) is majorized by the integral of the function $3|1-\widetilde{\mu}(r^2y)|$ against the measure γ, since $\sqrt{e}(\sqrt{e}-1)^{-1}<3$. The integral over the ball of radius r^{-1} is majorized by $3\sup_{|z|\le r}|1-\widetilde{\mu}(z)|$, as $|r^2y|\le r$ if $|y|\le r^{-1}$. By Chebyshev's inequality one has

$$\gamma\bigl(y\colon\ |y|>r^{-1}\bigr)\le r^2\int_{\mathbb{R}^n}|y|^2\,\gamma(dy)=nr^2.$$

It remains to observe that $|1-\widetilde{\mu}|\le 2$. □

3.9. Convolution

In this section, we apply Fubini's theorem and Hölder's inequality to convolutions of integrable functions.

3.9.1. Lemma. *Let a function f on $\mathbb{R}^n$ be Lebesgue measurable. Then, the function $(x,y)\mapsto f(x-y)$ is Lebesgue measurable on $\mathbb{R}^{2n}$.*

Proof. Set $g(x,y)=f(x-y)$ and consider the invertible linear transformation $F\colon (x,y)\mapsto (x-y,y)$. Then $g(x,y)=f_0\bigl(F(x,y)\bigr)$, where the function $f_0(x,y)=f(x)$ is Lebesgue measurable on $\mathbb{R}^{2n}$. By Corollary 3.6.4 the function g is measurable as well. □

3.9.2. Theorem. (i) *Let $f,\,g\in\mathcal{L}^1(\mathbb{R}^n)$. Then the function*

$$f*g(x)=\int_{\mathbb{R}^n}f(x-y)g(y)\,dy, \tag{3.9.1}$$

called the convolution of f and g, is defined for almost all x and is integrable. In addition,

$$\|f*g\|_{L^1(\mathbb{R}^n)}\le \|f\|_{L^1(\mathbb{R}^n)}\|g\|_{L^1(\mathbb{R}^n)}. \tag{3.9.2}$$

Moreover, $f*g=g*f$ almost everywhere.

(ii) *Let* $f \in \mathcal{L}^\infty(\mathbb{R}^n)$, $g \in \mathcal{L}^1(\mathbb{R}^n)$. *Then the function*

$$f * g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\,dy$$

is defined for all x *and*

$$\|f * g\|_{L^\infty(\mathbb{R}^n)} \le \|f\|_{L^\infty(\mathbb{R}^n)} \|g\|_{L^1(\mathbb{R}^n)}. \tag{3.9.3}$$

In addition, $f * g(x) = g * f(x)$.

PROOF. (i) We know that the function $\psi\colon (x,y) \mapsto |f(x-y)g(y)|$ is measurable on $\mathbb{R}^{2n}$. Since

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(x-y)|\,|g(y)|\,dx\,dy = \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}|f(z)|\,dz\right)|g(y)|\,dy < \infty,$$

it follows by Theorem 3.4.5 that the function ψ is integrable on $\mathbb{R}^{2n}$ and

$$\|\psi\|_{L^1(\mathbb{R}^{2n})} \le \|f\|_{L^1(\mathbb{R}^n)}\|g\|_{L^1(\mathbb{R}^n)}.$$

By Fubini's theorem the function

$$\varphi\colon\ x \mapsto \int \psi(x,y)\,dy$$

is defined for almost all x and is integrable. Hence the function $f * g$ is integrable as well, for $|f * g(x)| \le \varphi(x)$, and the measurability of $f * g$ follows by Lemma 3.9.1 and the assertion about measurability in Fubini's theorem. For all x such that the function $f(x-y)g(y)$ is integrable in y, the change of variable $z = x - y$ yields the equality $f * g(x) = g * f(x)$.

Assertion (ii) is obvious, since the function $y \mapsto g(x-y)$ is integrable for all x. □

3.9.3. Corollary. *If* $f, g \in \mathcal{L}^1(\mathbb{R}^n)$, *then* $\widehat{f * g}(y) = (2\pi)^{n/2}\widehat{f}(y)\widehat{g}(y)$.

PROOF. We already know that $f * g \in L^1(\mathbb{R}^n)$. By Fubini's theorem we have

$$\begin{aligned}(2\pi)^{-n/2}&\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i(y,x)} f(x-z)g(z)\,dzdx \\ &= (2\pi)^{-n/2}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i(y,u)}e^{-i(y,z)} f(u)g(z)\,dzdu,\end{aligned}$$

whence the desired formula follows. □

The next theorem generalizes the previous one and contains the important Young inequality.

3.9.4. Theorem. *Suppose that*

$$1 \le p \le q \le \infty, \quad \frac{1}{q} = \frac{1}{r} + \frac{1}{p} - 1.$$

Then, for any functions $f \in \mathcal{L}^p(\mathbb{R}^n)$ and $g \in \mathcal{L}^r(\mathbb{R}^n)$, the function $f * g$ is defined almost everywhere (everywhere if $q = \infty$), belongs to $\mathcal{L}^q(\mathbb{R}^n)$ and one has $f * g = g * f$ almost everywhere and

$$\|f * g\|_{L^q(\mathbb{R}^n)} \le \|f\|_{L^p(\mathbb{R}^n)} \|g\|_{L^r(\mathbb{R}^n)}. \tag{3.9.4}$$

PROOF. Let us consider the case $1 < p < q$, $r < q$. By Lemma 3.9.1 and Fubini's theorem, for almost every x, the function $y \mapsto f(x-y)g(y)$ is measurable. Then, for each fixed x with such a property, we can consider the function

$$|f(x-y)g(y)| = \Bigl(|f(x-y)|^p |g(y)|^r\Bigr)^{1/q} |f(x-y)|^{1-p/q} |g(y)|^{1-r/q}$$

of y and apply the generalized Hölder inequality with the exponents

$$p_1 = q, \quad p_2 = \frac{r}{1-r/q}, \quad p_3 = \frac{p}{1-p/q},$$

since $p_1^{-1} + p_2^{-1} + p_3^{-1} = 1$. Indeed,

$$\frac{1}{q} + \frac{q-r}{rq} + \frac{q-p}{pq} = \frac{pq + rq - rp}{rpq} = \frac{1}{r} + \frac{1}{p} - \frac{1}{q}.$$

Therefore,

$$|f * g(x)| \le \|f\|_p^{1-p/q} \|g\|_r^{1-r/q} \biggl(\int_{\mathbb{R}^n} |f(x-y)|^p |g(y)|^r \, dy\biggr)^{1/q}.$$

Thus, the function $y \mapsto f(x-y)g(y)$ is integrable for all points x such that it is measurable and the function $|f|^p * |g|^r$ is defined, i.e., for almost all x according to the previous theorem. One has $f * g(x) = g * f(x)$, which is proved by the same change of variable as in the previous theorem. Similarly, we obtain that the function $f * g$ is measurable. Finally, we have

$$\|f * g\|_q^q \le \|f\|_p^{q-p} \|g\|_r^{q-r} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(x-y)|^p |g(y)|^r \, dy \, dx = \|f\|_p^q \|g\|_r^q.$$

The remaining cases $1 = p < q = r$ and $p = q$, $r = 1$ follow by the previous theorem and Hölder's inequality applied to the function $y \mapsto f(x-y)g(y)$ for any fixed x. In particular, if $q = \infty$, then the integral of $|f(x-y)g(y)|$ in y is estimated by $\|f\|_p \|g\|_r$ due to Hölder's inequality, since in that case we have $p^{-1} + r^{-1} = 1$. □

3.9.5. Corollary. *Let $g \in \mathcal{L}^1(\mathbb{R}^n)$ and let a function f be bounded and continuous. Then, the function $f * g$ is bounded and continuous as well. If, in addition, f has continuous and bounded derivatives up to order k, then $f * g$ also does and*

$$\partial_{x_{i_1}} \ldots \partial_{x_{i_m}} (f * g) = (\partial_{x_{i_1}} \ldots \partial_{x_{i_m}} f) * g$$

for all $m \le k$.

PROOF. The continuity of $f * g$ follows by the dominated convergence theorem. If f has bounded and continuous partial derivatives, then by the theorem on differentiation of the Lebesgue integral with respect to a parameter (see Corollary 2.8.7) we obtain that the function $f * g$ has partial derivatives as well and $\partial_{x_i}(f * g) = \partial_{x_i} f * g$, moreover, these partial derivatives are continuous and bounded. By induction, the assertion extends to higher-order derivatives. $\square$

3.9.6. Corollary. *Let $f \in \mathcal{L}^p(\mathbb{R}^n)$, $g \in \mathcal{L}^q(\mathbb{R}^n)$, $p^{-1} + q^{-1} = 1$. Then, the function $f * g$ defined by equality* (3.9.1) *is continuous and bounded.*

PROOF. For any fixed x, the function $y \mapsto f(x - y)$ belongs to $\mathcal{L}^p(\mathbb{R}^n)$, hence by Hölder's inequality the integral in (3.9.1) exists for every x and is a bounded function. For any $f \in C_0^\infty(\mathbb{R}^n)$ the continuity of $f * g$ is trivial. In the general case, given $p < \infty$ we take a sequence of functions $f_j \in C_0^\infty(\mathbb{R}^n)$ convergent to f in $L^p(\mathbb{R}^n)$ (it suffices to approximate first the indicators of cubes, see §4.2 in Chapter 4). By the estimate

$$|f_j * g(x) - f * g(x)| \le \|f_j - f\|_{L^p(\mathbb{R}^n)} \|g\|_{L^q(\mathbb{R}^n)}, \quad \forall\, x \in \mathbb{R}^n,$$

the functions $f_j * g$ converge uniformly on $\mathbb{R}^n$ to $f * g$. If $p = \infty$, then $q = 1$ and a similar reasoning applies. $\square$

3.9.7. Example. Let A and B be two sets of positive Lebesgue measure in $\mathbb{R}^n$. Then, the set

$$A + B := \{a + b\colon\ a \in A, b \in B\}$$

contains an open ball.

PROOF. It suffices to consider bounded sets. By the continuity of $I_A * I_B$, the set

$$U = \bigl\{x\colon\ I_A * I_B(x) > 0\bigr\}$$

is open. The integral of $I_A * I_B$ equals the product of the measures of A and B and hence is not zero. Therefore, U is nonempty. Finally, $U \subset A + B$, since, for any $x \in U$, there exists $y \in B$ such that $x - y \in A$ (otherwise $I_A(x - y) I_B(y) = 0$ for all y and then $I_A * I_B(x) = 0$), whence we obtain the inclusion $x = x - y + y \in A + B$. $\square$

Exercise 3.10.98 contains a more general result.

Apart from convolutions of functions, one can consider convolutions of measures.

3.9.8. Definition. *Let μ and ν be two bounded Borel measures on $\mathbb{R}^n$. Their convolution $\mu * \nu$ is defined as the measure on $\mathbb{R}^n$ that is the image of the measure $\mu \otimes \nu$ on $\mathbb{R}^n \times \mathbb{R}^n$ under the mapping $(x, y) \mapsto x + y$.*

It follows by definition and Fubini's theorem that, for any $B \in \mathcal{B}(\mathbb{R}^n)$, one has the equality

$$\mu * \nu(B) = \int_{\mathbb{R}^n \times \mathbb{R}^n} I_B(x+y)\mu(dx)\,\nu(dy) \tag{3.9.5}$$
$$= \int_{\mathbb{R}^n} \mu(B-y)\,\nu(dy) = \int_{\mathbb{R}^n} \nu(B-x)\,\mu(dx).$$

The right-hand side of this equality can be taken for the definition of convolution. We note that the function $x \mapsto \mu(B-x)$ is Borel for every $B \in \mathcal{B}(\mathbb{R}^n)$. This follows by Proposition 3.3.2.

It is clear that $\mu * \nu = \nu * \mu$ and that $\widetilde{\mu * \nu} = \widetilde{\mu}\widetilde{\nu}$, since

$$\int_{\mathbb{R}^n} e^{i(y,x)} \mu * \nu(dx) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{iy(u+v)}\,\mu(du)\,\nu(dv),$$

which yields the stated equality by Fubini's theorem.

Finally, let us consider the convolution of a function and a measure. The proof of the following assertion is similar to the above reasoning and is delegated to Exercise 3.10.99. If μ is absolutely continuous, then this result is covered by the Young inequality with $r = 1$, $p = q$.

3.9.9. Proposition. *Let f be a Borel function in $\mathcal{L}^p(\mathbb{R}^n)$ and let μ be a bounded Borel measure on $\mathbb{R}^n$. The function*

$$f * \mu(x) := \int_{\mathbb{R}^n} f(x-y)\,\mu(dy)$$

is defined for almost all x with respect to Lebesgue measure and

$$\|f * \mu\|_{L^p(\mathbb{R}^n)} \le \|f\|_{L^p(\mathbb{R}^n)} \|\mu\|.$$

3.9.10. Example. Let μ and ν be probability measures on a measurable space $(X, \mathcal{A})$ such that $\nu \ll \mu$ and let σ be a probability measure on a measurable space $(Y, \mathcal{B})$. Suppose that $T\colon X \times Y \to Z$ be a measurable mapping with values in a measurable space $(Z, \mathcal{E})$. Then

$$\nu_{\sigma,T} := (\nu \otimes \sigma) \circ T^{-1} \ll \mu_{\sigma,T} := (\mu \otimes \sigma) \circ T^{-1}$$

and

$$\int_Z \Bigl|\frac{d\nu_{\sigma,T}}{d\mu_{\sigma,T}}\Bigr|^p d\mu_{\sigma,T} \le \int_X \Bigl|\frac{d\nu}{d\mu}\Bigr|^p d\mu$$

for any $p \in [1, \infty)$ such that $d\nu/d\mu \in L^p(\mu)$.

In particular, if $X = Y = Z = \mathbb{R}^n$ and $T(x,y) = x+y$, one obtains

$$\int_{\mathbb{R}^n} \Bigl|\frac{d(\nu * \sigma)}{d(\mu * \sigma)}\Bigr|^p d(\mu * \sigma) \le \int_{\mathbb{R}^n} \Bigl|\frac{d\nu}{d\mu}\Bigr|^p d\mu.$$

PROOF. It is obvious that $\nu \otimes \sigma \ll \mu \otimes \sigma$ and $d(\nu \otimes \sigma)/d(\mu \otimes \sigma) = f$, where $f := d\nu/d\mu$ is regarded as a function on $X \times Y$. Hence $\nu_{\sigma,T} \ll \mu_{\sigma,T}$. Let

$g := d\nu_{\sigma,T}/d\mu_{\sigma,T}$ and $q = p/(p-1)$. For every function $\varphi \in \mathcal{L}^\infty(\mu_{\sigma,T})$, one has by Hölder's inequality

$$\int_Z \varphi g\, d\mu_{\sigma,T} = \int_Z \varphi\, d\nu_{\sigma,T} = \int_{X\times Y} \varphi \circ T\, d(\nu\otimes\sigma) = \int_{X\times Y} \varphi \circ T f\, d(\mu\otimes\sigma)$$
$$\leq \|f\|_{L^p(\mu)} \Bigl(\int_Y \int_X \bigl|\varphi\bigl(T(x,y)\bigr)\bigr|^q \,\mu(dx)\,\sigma(dy)\Bigr)^{1/q} = \|f\|_{L^p(\mu)} \|\varphi\|_{L^q(\mu_{\sigma,T})},$$

which by Example 2.11.6 yields the desired inequality. In the case where $X = Y = Z = \mathbb{R}^n$ and $T(x,y) = x + y$ we have $(\mu\otimes\sigma)\circ T^{-1} = \mu * \sigma$ and similarly for ν. For an alternative proof of a more general fact, see Exercise 10.10.93 in Chapter 10. □

3.10. Supplements and exercises

3.10(i). On Fubini's theorem and products of σ-algebras

In applications of Fubini's theorem one should not forget that it deals with sets in products of spaces (and with functions on them) which are known in advance to be measurable with respect to the product measure. There exists a Lebesgue nonmeasurable set in the unit square such that all intersections of this set with the straight lines parallel to the coordinate axes consist of at most one point (see Exercise 3.10.49). It is suggested in Exercise 3.10.50 that the reader construct an example of a nonmeasurable nonnegative function on the square such that the repeated integrals exist and vanish. Finally, Exercise 3.10.51 provides an example of a bounded function (the indicator of a set) such that one of the repeated integrals equals 0 and the other one equals 1. However, the construction essentially uses the continuum hypothesis. Moreover, Friedman [**328**] proved that it is consistent with the standard set theory with the axiom of choice (ZFC) that if, for a bounded (not necessarily measurable) function f on the square both repeated integrals exist, then they are equal. The existence of the repeated integrals means that, for a.e. x, the function $f(x,y)$ is integrable in y, the function

$$\int f(x,y)\,dy$$

is integrable in x, and the same is true when we consider the variables in the reversed order.

There exist rather exotic measurable sets, too. According to Fubini's theorem, for any set A of measure 1 in the square $[0,1]\times[0,1]$, almost every section by the straight line parallel to the first coordinate axis has the linear measure 1. The surprising Example 1.12.25, due to Nikodym, shows that in this statement it is essential to consider a priori fixed axes: there exists a set of full measure in the plane such that through every point of this set one can pass a straight line meeting this set at the given point.

It is to be noted that the product of nonnegative measures μ and ν can be defined in such a way that the initial equality $\mu\otimes\nu(A\times B) = \mu(A)\nu(B)$ will not be obvious and will require a justification, but the measures may not be finite or σ-finite. This approach is based on Carathéodory outer measures (see §1.12). Suppose we are given two Carathéodory outer measures μ^* and ν^* in the sense of Definition 1.11.1 (i.e., they are not necessarily generated by the usual measures). Let μ and ν denote their restrictions to the σ-algebras $\mathfrak{M}_{\mu^*}$ and $\mathfrak{M}_{\nu^*}$ (which are known to be countably additive measures). First we define the set function $\mu^*\times\nu^*$ on the class of all subsets in $X\times Y$ by the formula

$$\mu^*\times\nu^*(E) = \inf\Bigl\{\sum_{i=1}^{\infty}\mu(A_i)\nu(B_i)\Bigr\},$$

where inf is taken over all $A_i \in \mathfrak{M}_{\mu^*}$, $B_i \in \mathfrak{M}_{\nu^*}$ with $E \subset \bigcup_{i=1}^{\infty}(A_i\times B_i)$. Then the following theorem can be proved (see, e.g., Bruckner, Bruckner, Thomson [**136**, Theorem 6.2]).

3.10.1. Theorem. *The set function $\mu^*\times\nu^*$ is a regular Carathéodory outer measure on $X\times Y$, and for all $A \in \mathfrak{M}_{\mu^*}$ and $B \in \mathfrak{M}_{\nu^*}$, we have $A\times B \in \mathfrak{M}_{\mu^*\times\nu^*}$ and $\mu^*\times\nu^*(A\times B) = \mu^*(A)\nu^*(B)$.*

If a function is integrable with respect to such a product measure, then it vanishes outside some set on which the product measure is σ-finite, hence integration of this function reduces to integration with respect to a product of two σ-finite measures. In particular, Fubini's theorem is true in this setting. However, without additional assumptions such as σ-finiteness any further development of this approach is not very fruitful. For example, Tonelli's theorem may fail here (Exercises 3.10.58, 3.10.64, 3.10.65, 3.10.66, and 3.10.67 demonstrate the subtleties arising here; see also Falconer, Mauldin [**278**]).

In most of applications, Fubini's theorem is applied to measures that are defined on product spaces equipped with products of σ-algebras. However, in some cases, a product space possesses other natural σ-algebras. For example, if X and Y are two topological spaces equipped with their Borel σ-algebras $\mathcal{B}(X)$ and $\mathcal{B}(Y)$, then the space $X\times Y$ has the product topology, hence it can be equipped with the corresponding Borel σ-algebra $\mathcal{B}(X\times Y)$, which may be strictly larger than $\mathcal{B}(X)\otimes\mathcal{B}(Y)$. Such problems are addressed in Chapter 6 and Chapter 7. Here we only discuss the case where X and Y are nonempty sets equipped with the σ-algebras of all subsets; these σ-algebras are denoted by $\mathcal{P}(X)$ and $\mathcal{P}(Y)$. Clearly, these σ-algebras coincide with the Borel σ-algebras corresponding to the discrete metrics on X and Y, i.e., the distances between all distinct points are 1. Is it true that $\mathcal{P}(X)\otimes\mathcal{P}(Y) = \mathcal{P}(X\times Y)$? We shall see in §6.4 that the answer is "no" if the cardinality of X and Y is greater than $\mathfrak{c}$. The situation is more complicated if X and Y are of uncountable cardinality less than or equal to $\mathfrak{c}$. The following result was obtained in Rao [**784**].

3.10.2. Proposition. *Let Ω be a set of cardinality corresponding to the first uncountable ordinal ω_1 and let $\mathcal{P}(\Omega)$ be the set of all its subsets. Then*

$$\mathcal{P}(\Omega)\otimes\mathcal{P}(\Omega) = \mathcal{P}(\Omega\times\Omega).$$

Under the continuum hypothesis, the σ-algebra generated by all products $A\times B$, $A, B \subset [0,1]$, coincides with the class of all sets in $[0,1]\times[0,1]$.

PROOF. We may deal with the ordinal interval $\Omega = [0,\omega_1)$ equipped with its natural order $\leq$. Any function on Ω with values in $[0,1]$ is $\bigl(\mathcal{P}(\Omega),\mathcal{B}([0,1])\bigr)$-measurable, hence its graph belongs to $\mathcal{P}(\Omega)\otimes\mathcal{B}([0,1])$. Since one can embed Ω into $[0,1]$, the graph of any mapping from Ω to Ω belongs to $\mathcal{P}(\Omega)\otimes\mathcal{P}(\Omega)$. This yields that $\mathcal{P}(\Omega)\otimes\mathcal{P}(\Omega)$ contains every set $E \in \Omega\times\Omega$ such that all sections $E_x := \{y\colon\ (x,y)\in E\}$, $x\in\Omega$, are at most countable. The same is true for any set E such that all sections $E_y := \{x\colon\ (x,y)\in E\}$, $y\in\Omega$, are at most countable. The sets $\{\alpha\colon\ \alpha\leq\alpha_0\}$ are at most countable for all $\alpha_0<\omega_1$. Hence $\mathcal{P}(\Omega)\otimes\mathcal{P}(\Omega)$ contains every subset of the set $\{(\alpha,\beta)\colon\ \alpha\leq\beta\}$ and every subset of the set $\{(\alpha,\beta)\colon\ \beta\leq\alpha\}$. This proves our claim, since the union of the two indicated sets is $\Omega\times\Omega$. See also Kharazishvili [**511**, p. 201]; Mauldin [**659**]. $\square$

Now we see how this result along with Fubini's theorem yields a shorter proof of Theorem 1.12.40. Moreover, the following fact established in Banach, Kuratowski [**57**] is true.

3.10.3. Corollary. *There exists a countable family of sets $A_n\subset\Omega$ such that the σ-algebra $\sigma(\{A_n\})$ contains all singletons, but carries no nonzero measure vanishing on all singletons.*

In particular, under the continuum hypothesis, there exists a countable family of sets $A_n\subset[0,1]$ such that Lebesgue measure cannot be extended to a countably additive measure on the σ-algebra generated by all Borel sets and all sets A_n.

PROOF. We recall that $\Omega=[0,\omega_1)$ is well-ordered and that for any $\beta\in\Omega$, the set $\{\alpha\colon\ \alpha\leq\beta\}$ is at most countable. By Exercise 3.10.38 and the above proposition, the set $M := \{(\alpha,\beta)\colon\ \alpha\leq\beta\}$ is contained in the σ-algebra generated by some countable collection of products $A_i\times A_j$. We can consider Ω as a subset of $[0,1]$ and add to $\{A_n\}$ all sets $\Omega\cap(r,s)$ with rational r, s. Hence we obtain countably many sets, again denoted by A_n, such that $\sigma(\{A_n\})$ contains all singletons in Ω. Suppose that μ is a probability measure on the σ-algebra $\mathcal{A}=\sigma(\{A_n\})$. Then M is measurable with respect to $\mu\otimes\mu$. This leads to a contradiction because by Fubini's theorem the set M and its complement have $\mu\otimes\mu$-measure zero. Indeed, all horizontal sections of the set M and all vertical sections of its complement are at most countable. Finally, under the continuum hypothesis, there is a one-to-one correspondence between Ω and $[0,1]$. $\square$

3.10(ii). Steiner's symmetrization

In this section, we consider an interesting transformation of sets that preserves Lebesgue measure λ_n. Let $a, b \in \mathbb{R}^n$ and $|a| = 1$. The straight line $L_a(b)$ having the direction vector a and passing through the point b is determined by the equality $L_a(b) = \{b + ta\colon\ t \in \mathbb{R}\}$. Let Π_a denote the orthogonal complement of the straight line $\mathbb{R}a$.

3.10.4. Definition. *For every set A in $\mathbb{R}^n$, Steiner's symmetrization of A with respect to the hyperplane Π_a is the set*

$$S_a(A) := \bigcup_{b\in\Pi_a, A\cap L_a(b)\neq\varnothing} \Bigl\{b + ta\colon\ |t| \le \frac{1}{2}\lambda_1^*\bigl(A \cap L_a(b)\bigr)\Bigr\},$$

where λ_1 is the natural Lebesgue measure on the straight line $L_a(b)$.

For example, let a be the vector e_2 in $\mathbb{R}^2$ and let A be the set under the graph of a nonnegative measurable function f on $[0, 1]$. The symmetrization S_a takes A to the set bounded by the graphs of the functions $f/2$ and $-f/2$, since for $b \in \Pi_a = \mathbb{R}e_1$ the section of A by the line $L_a(b)$ is an interval of length $f(b)$. By Fubini's theorem, it is clear that A and $S_a(A)$ have equal areas.

In the general case, on the set $\Omega_A := \{b \in \Pi_a\colon\ L_a(b) \cap A \neq \varnothing\}$ we define the function $f(b) = \lambda_1^*\bigl(A \cap L_a(b)\bigr)$. Then $S_a(A)$ is the set between the graphs of the functions $f/2$ and $-f/2$ on the set Ω_A. If A is measurable, then Fubini's theorem yields that Ω_A is measurable with respect to the natural Lebesgue measure λ_{Π_a} on the $(n-1)$-dimensional subspace Π_a and the function f is measurable on Ω_A. This shows the measurability of $S_a(A)$. In addition, for λ_{Π_a}-almost all $b \in \Omega_A$, the set $A \cap L_a(b)$ is measurable with respect to λ_1.

The diameter of a nonempty set A is the number $\operatorname{diam} A$ equal the supremum of the distances between points in the set A; $\operatorname{diam}\varnothing := 0$.

3.10.5. Proposition. *For any set A, we have* $\operatorname{diam} S_a(A) \le \operatorname{diam} A$. *If the set A is measurable, then $\lambda_n\bigl(S_a(A)\bigr) = \lambda_n(A)$.*

Proof. Since the closure of A has the same diameter as A, we may assume in the first assertion that A is closed. Moreover, we may assume that A is bounded (otherwise the claim is obvious). We take $\varepsilon > 0$ and choose $x, y \in S_a(A)$ with $\operatorname{diam} S_a(A) \le |x - y| + \varepsilon$. Set $b = x - (x, a)a$, $c = y - (y, a)a$. Then $b, c \in \Pi_a$. Let

$$m_b = \inf\{t\colon\ b + ta \in A\}, \quad M_b = \sup\{t\colon\ b + ta \in A\},$$
$$m_c = \inf\{t\colon\ c + ta \in A\}, \quad M_c = \sup\{t\colon\ c + ta \in A\}.$$

We may assume that $M_c - m_b \ge M_b - m_c$. Then

$$M_c - m_b \ge \frac{M_b - m_b}{2} + \frac{M_c - m_c}{2} \ge \frac{1}{2}\lambda_1\bigl(A \cap L_a(b)\bigr) + \frac{1}{2}\lambda_1\bigl(A \cap L_a(c)\bigr).$$

We observe that $|(x, a)| \le \lambda_1\bigl(A \cap L_a(b)\bigr)/2$. This follows by the definition of $S_a(A)$, since $x = b + (x, a)a \in S_a(A)$. Similarly, $|(y, a)| \le \lambda_1\bigl(A \cap L_a(c)\bigr)/2$.

Therefore, $M_c - m_b \geq |(x,a)| + |(y,a)| \geq |(x-y,a)|$, whence we have

$$\begin{aligned}|\operatorname{diam} S_a(A) - \varepsilon|^2 &\leq |x-y|^2 = |b-c|^2 + |(x-y,a)|^2 \\ &\leq |b-c|^2 + |M_c - m_b|^2 = |(b+m_b a) - (c+M_c a)|^2 \leq (\operatorname{diam} A)^2\end{aligned}$$

because $b + m_b a, c + M_c a \in A$ by the assumption that A is closed. Since ε is arbitrary, we obtain $\operatorname{diam} S_a(A) \leq \operatorname{diam} A$.

In the proof of the second assertion we may assume, by the rotational invariance of Lebesgue measure, that $a = e_n = (0, \ldots, 0, 1)$. Then we have $\Pi_a = \mathbb{R}^{n-1}$. The measurability of $S(A)$ has already been justified. By Fubini's theorem, the function $f(b) = \lambda_1\bigl(A \cap L_a(b)\bigr)$ is measurable on $\mathbb{R}^{n-1}$, and its integral equals the measure of A. The same integral is obtained by evaluating the measure of $S_a(A)$ by Fubini's theorem, since, for each $b \in \mathbb{R}^{n-1}$ such that $L_a(b) \cap A \neq \varnothing$, the section of the set $S_a(A)$ by the straight line $b + \mathbb{R}e_n$ is an interval of length $f(b)$. □

The next result shows that among the sets of a given diameter, the ball has the maximal volume. This is not obvious because a set of diameter 1 need not be contained in a ball of diameter 1. For example, a triangle of diameter 1 may not be covered by a disc of diameter 1.

3.10.6. Corollary. *For any set $A \subset \mathbb{R}^n$, one has*

$$\lambda_n^*(A) \leq \lambda_n(U)\Bigl(\frac{\operatorname{diam} A}{2}\Bigr)^n, \tag{3.10.1}$$

where U is the unit ball.

PROOF. It suffices to consider closed sets, since the closure of a set has the same diameter. We shall assume that A is bounded. Let us take the standard basis $e_1, \ldots, e_n$ and consider the consecutive symmetrizations $A_1 = S_{e_1}(A), \ldots, A_n = S_{e_n}(A_{n-1})$. We know that $\lambda_n(A_n) = \lambda_n(A)$ and $\operatorname{diam} A_n \leq \operatorname{diam} A$. Hence it suffices to show that (3.10.1) is true for A_n. If we show that A_n is centrally symmetric, then (3.10.1) will be a trivial consequence of the fact that A_n is contained in a ball of radius $\operatorname{diam} A_n/2$. Indeed, in that case for any $x \in A_n$, we have $-x \in A_n$, whence we obtain $|x| \leq \operatorname{diam} A_n/2$.

It remains to show that A_n is centrally symmetric. To this end, we verify that A_n is symmetric with respect to the hyperplanes Π_{e_j}. It is clear that A_1 is symmetric with respect to Π_{e_1}. Suppose that $1 \leq k < n$ and A_k is symmetric with respect to Π_{e_j}, $j \leq k$. The set $A_{k+1} = S_{e_{k+1}}(A_k)$ is symmetric with respect to $\Pi_{e_{k+1}}$. Let $j \leq k$ and let R_j be the reflection with respect to Π_{e_j}. Let us take $b \in \Pi_{e_{k+1}}$. By using that $R_j(A_k) = A_k$ we obtain

$$\lambda_1\bigl(A_k \cap L_{e_{k+1}}(b)\bigr) = \lambda_1\bigl(A_k \cap L_{e_{k+1}}(R_j(b))\bigr).$$

This yields the equality

$$\{t\colon\ b + te_{k+1} \in A_{k+1}\} = \{t\colon\ R_j(b) + te_{k+1} \in A_{k+1}\}.$$

Hence $R_j(A_{k+1}) = A_{k+1}$, i.e., A_{k+1} is symmetric with respect to Π_{e_j}. By induction we obtain our claim. □

Melnikov [**679**] proved that the above result remains valid for an arbitrary (not necessarily Euclidean) finite-dimensional normed space, and his proof of the following theorem is very elementary (only Fubini's theorem is used) and is almost as short as the above reasoning.

3.10.7. Theorem. *Suppose that a set A in the space $\mathbb{R}^n$ equipped with some norm p has diameter 2 with respect to the norm p. Then the inequality $\lambda_n^*(A) \le \lambda_n(U)$ holds, where U is the unit ball in the norm p.*

Close to Steiner's symmetrization is the concept of a symmetric rearrangement of a set or function. The symmetric rearrangement of a measurable set $A \subset \mathbb{R}^n$ is the set $A^* \subset \mathbb{R}^n$ that is the open ball with the center at the origin and the volume equal to that of A. The symmetric rearrangement of a function I_A is the function I_{A^*}, denoted by I_A^*. Now, for an arbitrary measurable function f on $\mathbb{R}^n$, its measurable rearrangement is defined by the formula

$$f^*(x) = \int_0^\infty I^*_{\{|f|>t\}}(x)\,dt.$$

It is clear that the function f^* is a function of $|x|$. In Exercise 3.10.102, an equivalent definition of the rearrangement of a function is given, according to which the rearrangement is a function on the real line equimeasurable with the given function on $\mathbb{R}^n$. A concise exposition of the basic properties of symmetric rearrangements is given in the book Lieb, Loss [**612**]. So here we only mention without proof several key facts. For any $t > 0$, one has the equality

$$\bigl\{x\colon\ f^*(x) > t\bigr\} = \bigl\{x\colon\ |f(x)| > t\bigr\}^*.$$

Hence, for Lebesgue measure λ_n, we obtain

$$\lambda_n\bigl(x\colon\ f^*(x) > t\bigr) = \lambda_n\bigl(x\colon\ |f(x)| > t\bigr).$$

This equality yields $\|f^*\|_{L^p} = \|f\|_p$. In addition, $\|f^* - g^*\|_{L^p} \le \|f - g\|_p$. The last inequality is a special case of a more general fact. Namely, let Ψ be a nonnegative convex function on the real line such that $\Psi(0) = 0$ and let f and g be nonnegative measurable functions on $\mathbb{R}^n$ with bounded support. Then

$$\int_{\mathbb{R}^n} \Psi\bigl(f^*(x) - g^*(x)\bigr)\,dx \le \int_{\mathbb{R}^n} \Psi\bigl(f(x) - g(x)\bigr)\,dx.$$

For all nonnegative measurable functions with bounded support one has

$$\int_{\mathbb{R}^n} f(x)g(x)\,dx \le \int_{\mathbb{R}^n} f^*(x)g^*(x)\,dx.$$

The following deep result is due to F. Riesz. For all nonnegative measurable functions f, g, h on $\mathbb{R}^n$, one has

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)g(x-y)h(y)\,dx\,dy \le \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f^*(x)g^*(x-y)h^*(y)\,dx\,dy.$$

The above cited book contains proofs, references, and other related interesting results.

In Busemann, Petty [**154**], the following question was raised. Let B be a unit ball centered at the origin in $\mathbb{R}^n$ and let K be a centrally symmetric convex set. Suppose that for every $(n-1)$-dimensional linear subspace L in $\mathbb{R}^n$, one has $\lambda_{n-1}(K\cap L) < \lambda_{n-1}(B\cap L)$. Is it true that $\lambda_n(K) < \lambda_n(B)$? It turned out that this is true if $n \le 3$, but for $n \ge 4$ this is false; see Gardner [**341**], Gardner, Koldobsky, Schlumprecht [**343**], Zhang [**1049**], [**1050**], Larman, Rogers [**571**]).

3.10(iii). Hausdorff measures

In this subsection, we discuss an interesting class of measures containing Lebesgue measure: Hausdorff measures. As above, let $\operatorname{diam} C$ denote the diameter of a set C. We recall that the Gamma-function is defined by the formula
$$\Gamma(s) = \int_0^\infty e^{-x}x^{s-1}\,dx, \quad s>0.$$
Set $\alpha(s) = \pi^{s/2}/\Gamma(1+s/2)$. Then $\alpha(n)$ is the volume of the unit ball in $\mathbb{R}^n$ (see Exercise 3.10.83).

3.10.8. Definition. *Let $s\in[0,+\infty)$ and let $\delta\in(0,+\infty)$. For any set $A\subset\mathbb{R}^d$, let*
$$H^s_\delta(A) = \inf\biggl\{\sum_{i=1}^\infty \alpha(s)\Bigl(\frac{\operatorname{diam} C_j}{2}\Bigr)^s\colon\ A\subset\bigcup_{j=1}^\infty C_j,\ \operatorname{diam} C_j\le\delta\biggr\},$$
$$H^s(A) = \lim_{\delta\to 0} H^s_\delta(A) = \sup_{\delta>0} H^s_\delta(A).$$

We note that the second equality in the definition of H^s is fulfilled, since $H^s_\delta \ge H^s_{\delta'}$ whenever $0<\delta<\delta'$.

It is clear that H^s_δ is the Carathéodory outer measure corresponding to the set function $\tau(C) = \alpha(s)2^{-s}(\operatorname{diam} C)^s$ on the family of all sets of diameter at most δ (see Example 1.11.5). Hence the set function H^s_δ is countably subadditive. We observe that in the definition of H^s_δ one could use only closed sets, since the diameter of the closure of C equals that of C.

3.10.9. Proposition. *The set function H^s is a regular Carathéodory outer measure, and all Borel sets are measurable with respect to H^s. In addition, the function H^s is invariant with respect to translations and orthogonal linear operators.*

Proof. The countable subadditivity of H^s follows by the countable subadditivity of H^s_δ for $\delta>0$. Let $A,B\subset\mathbb{R}^n$ and $\operatorname{dist}(A,B)>0$. We pick a positive number $\delta<\operatorname{dist}(A,B)/4$ and take sets C_j that cover $A\cup B$ and have diameters at most δ. This cover falls into a cover of A by some of the sets C_j (which are denoted again by C_j) and a cover of B by sets C'_j such that

$\bigl(\bigcup_{j=1}^\infty C_j\bigr)\bigcap\bigl(\bigcup_{j=1}^\infty C_j'\bigr)=\varnothing$. Hence

$$H_\delta^s(A)+H_\delta^s(B)\le\sum_{j=1}^\infty \alpha(s)2^{-s}(\operatorname{diam} C_j)^s+\sum_{j=1}^\infty \alpha(s)2^{-s}(\operatorname{diam} C_j')^s,$$

whence we obtain that $H_\delta^s(A)+H_\delta^s(B)\le H_\delta^s(A\cup B)$, which yields the estimate $H^s(A)+H^s(B)\le H^s(A\cup B)$ as $\delta\to 0$. By the countable subadditivity we arrive at the equality $H^s(A\cup B)=H^s(A)+H^s(B)$. According to Theorem 1.11.10 all Borel sets are H^s-measurable.

If $H^s(A)<\infty$, then, for every $k\in\mathbb{N}$, one can find a cover of A by closed sets C_j^k with diameters at most k^{-1} and

$$\sum_{j=1}^\infty \alpha(s)2^{-s}(\operatorname{diam} C_j^k)^s\le H_{1/k}^s(A)+k^{-1}.$$

The set $B=\bigcap_{k=1}^\infty\bigcup_{j=1}^\infty C_j^k$ is Borel and

$$H_{1/k}^s(B)\le\sum_{j=1}^\infty \alpha(s)2^{-s}(\operatorname{diam} C_j^k)^s\le H_{1/k}^s(A)+k^{-1},$$

whence one has $H^s(B)\le H^s(A)\le H^s(A)$. The last claim is obvious. □

We shall call H^s the s-dimensional Hausdorff measure. It is clear that

$$H^s(\lambda A)=\lambda^s H^s(A),\quad \forall\,\lambda>0.$$

In addition, $H^0(A)$ is just the cardinality of the set A (finite or infinite).

It is easily verified (Exercise 3.10.103), that if $s<t$ and $H^s(A)<\infty$, then $H^t(A)=0$. If $H_\delta^s(A)=0$ for some $\delta>0$, then $H^s(A)=0$.

If A is a bounded set in $\mathbb{R}^n$, then A is contained in some cube with the edge length C and can be covered by $(C/r)^n$ cubes with the edge length r. Hence it can also be covered by $n^{n/2}(C/\delta)^n$ balls of diameter δ. Therefore, $H^n(A)<\infty$ (it is shown below that H^n is Lebesgue outer measure). It is also clear that $H^s(A)=0$ for $s>n$.

The Hausdorff dimension of A is defined as the number

$$\dim_H(A):=\inf\bigl\{s\in[0,+\infty)\colon\ H^s(A)=0\bigr\}.$$

3.10.10. Lemma. *If $s=n=1$, then the set functions H^1 and H_δ^1 are equal for all $\delta>0$ and coincide with Lebesgue outer measure.*

Proof. If a set A is covered by closed sets C_j of diameter at most δ, then its outer measure does not exceed the sum of diameters of C_j, whence $\lambda_1^*(A)\le H_\delta^1(A)$. On the other hand, A can be covered by a sequence of disjoint intervals C_j with diameters less than δ such that the sum of diameters is as close to the outer measure of A as we wish. Hence $\lambda_1^*(A)\ge H_\delta^1(A)$. □

3.10.11. Proposition. *If $s=n$, then the set function H^n coincides with Lebesgue outer measure.*

PROOF. By the regularity of both outer measures, it suffices to verify their equality on all Borel sets. Thus, we may deal further with the measures H^n and λ_n on Borel sets. According to Exercise 1.12.74, the invariance with respect to translations yields the equality $H^n = c\lambda_n$ for some $c > 0$. We show that $c \leq 1$. Otherwise for the open unit ball U we have $H^n(U) > \lambda_n(U)$. Let us pick $\delta > 0$ with $H^n_\delta(U) > \lambda_n(U)$. It follows by Theorem 1.7.4 that there exist disjoint balls $U_j \subset U$ with radii at most δ such that $\lambda_n\bigl(U\backslash\bigcup_{j=1}^\infty U_j\bigr) = 0$. Then

$$H^n_\delta\Bigl(U\backslash\bigcup_{j=1}^\infty U_j\Bigr) \leq H^n\Bigl(U\backslash\bigcup_{j=1}^\infty U_j\Bigr) = 0.$$

Hence

$$H^n_\delta(U) = \sum_{j=1}^\infty H^n_\delta(U_j) \leq \sum_{j=1}^\infty \lambda_n(U_j) = \lambda_n(U).$$

This contradiction shows that $c \leq 1$. On the other hand, according to inequality (3.10.1), if U is covered by closed sets C_j of diameter at most δ, then

$$\lambda_n(U) \leq \sum_{j=1}^\infty \lambda_n(C_j) \leq \sum_{j=1}^\infty \alpha(n)2^{-n}(\operatorname{diam} C_j)^n$$

and hence $\lambda_n(U) \leq H^n_\delta(U) \leq H^n(U)$. □

It is proposed in Exercise 3.10.104 that the reader construct sets B_α in the interval $[0,1]$ with $H^\alpha(B_\alpha) = 1$ for all $\alpha \in (0,1)$ and show that the Cantor set has a finite positive H^α-measure for $\alpha = \ln 2/\ln 3$.

3.10.12. Lemma. *Let a mapping $f\colon \mathbb{R}^n \to \mathbb{R}^m$ satisfy the Lipschitz condition with the constant Λ, i.e., $|f(x)-f(y)| \leq \Lambda|x-y|$ for all $x,y \in \mathbb{R}^n$. Then, for every $s \geq 0$ and every $A \subset \mathbb{R}^n$, we have $H^s\bigl(f(A)\bigr) \leq \Lambda^s H^s(A)$.*

PROOF. We may assume that $\Lambda > 0$, otherwise the claim is obvious. Suppose that A is covered by sets C_j of diameter at most $\delta > 0$. Then $\operatorname{diam} f(C_j) \leq \Lambda \operatorname{diam} C_j \leq \Lambda\delta$ and the sets $f(C_j)$ cover $f(A)$. Hence

$$H^s_{\Lambda\delta}\bigl(f(A)\bigr) \leq \Lambda^s \sum_{j=1}^\infty \alpha(s)2^{-s}(\operatorname{diam} C_j)^s,$$

so $H^s_{\Lambda\delta}\bigl(f(A)\bigr) \leq \Lambda^s H^s_\delta(A)$. Letting $\delta \to 0$ we obtain our assertion. □

In particular, orthogonal projections do not increase Hausdorff measures.

3.10.13. Corollary. *Let A be a set in $\mathbb{R}^n$ of positive outer measure and let $f\colon \mathbb{R}^n \to \mathbb{R}^m$. Let us denote by $G(f,A)$ the graph of f on A, i.e., $G(f,A) = \bigl\{\bigl(x,f(x)\bigr), x \in A\bigr\}$. Then, the Hausdorff dimension of $G(f,A)$ is not less than n, and in the case where f is Lipschitzian, it is exactly n.*

PROOF. By the above lemma the Hausdorff dimension does not increase under projection, and the projection of the set $G(f,A)$ to $\mathbb{R}^n$ is the set A, which by our hypothesis has the Hausdorff dimension n. If f is Lipschitzian, then $G(f,A)$ is the image of A under the Lipschitzian mapping $x \mapsto \bigl(x, f(x)\bigr)$, whence by the equality $H^s(\mathbb{R}^n) = 0$ for $s > n$ and the lemma we obtain the second assertion. □

Certain generalizations of Hausdorff measures on general metric spaces will be considered in Chapter 7.

3.10(iv). Decomposition of set functions

It is shown in this subsection that any additive set function can be decomposed in the sum of a countably additive measure and an additive set function without countably additive components. Let X be a nonempty set.

3.10.14. Theorem. *Let $\mathcal{R}$ be a ring of subsets of a space X and let $m\colon \mathcal{R} \to [0,+\infty]$ be a function with the following property of superadditivity:*

$$m(A_1 \cup \cdots \cup A_n) \ge m(A_1) + \cdots + m(A_n)$$

for all disjoint $A_1, \ldots, A_n \in \mathcal{R}$.

(i) *For all $A \in \mathcal{R}$ we set*

$$m_{\rm add}(A) := \inf\Bigl\{\sum_{j=1}^n m(A_j)\colon \quad A = \bigcup_{j=1}^n A_j, A_j \in \mathcal{R},\ A_j \textit{ are disjoint}\Bigr\}.$$

Then $m_{\rm add}$ is an additive set function, $m_{\rm add} \le m$, and $m_{\rm add} \ge \nu$ for each additive set function $\nu\colon \mathcal{R} \to [0,+\infty]$ such that $\nu \le m$.

(ii) *Set*

$$m_\sigma(A) := \inf\Bigl\{\sum_{j=1}^\infty m(A_j)\colon \quad A = \bigcup_{j=1}^\infty A_j,\ A_j \in \mathcal{R},\ A_j \textit{ are disjoint}\Bigr\},\ A \in \mathcal{R}.$$

Then m_σ is a countably additive set function, $m_\sigma \le m$, and $m_\sigma \ge \nu$ for each countably additive set function $\nu\colon \mathcal{R} \to [0,+\infty]$ such that $\nu \le m$.

PROOF. (i) Let $E_1, E_2 \in \mathcal{R}$, $E_1 \cap E_2 = \varnothing$. We show that $m_{\rm add}(E_1 \cup E_2) \le m_{\rm add}(E_1) + m_{\rm add}(E_2)$. We may assume that the right-hand side is finite. Let us fix $\varepsilon > 0$ and find disjoint sets $E_1^1, \ldots, E_1^k \in \mathcal{R}$ and disjoint sets $E_2^1, \ldots, E_2^n \in \mathcal{R}$ with $E_1 = \bigcup_{i=1}^k E_1^i$, $E_2 = \bigcup_{j=1}^n E_2^j$, and

$$\sum_{i=1}^k m(E_1^i) < m_{\rm add}(E_1) + \varepsilon,\ \sum_{j=1}^n m(E_2^j) < m_{\rm add}(E_2) + \varepsilon.$$

Then E_1^i and E_2^j are disjoint, hence

$$m_{\rm add}(E_1 \cup E_2) \le \sum_{i=1}^k m(E_1^i) + \sum_{j=1}^n m(E_2^j) < m_{\rm add}(E_1) + m_{\rm add}(E_2) + 2\varepsilon.$$

It remains to use that ε is arbitrary. Let us establish the opposite inequality. Now we may assume that $m_{\rm add}(E_1 \cup E_2) < \infty$. For any fixed $\varepsilon > 0$, we write $E_1 \cup E_2$ as the disjoint union of sets $A_j \in \mathcal{R}$, $j = 1, \ldots, n$, such that

$\sum_{j=1}^{n} m(A_j) < m_{\mathrm{add}}(E_1 \cup E_2) + \varepsilon$. Then we have $E_1^j := E_1 \cap A_j \in \mathcal{R}$, $E_2^j := E_2 \cap A_j \in \mathcal{R}$ and by the superadditivity of m we obtain

$$m_{\mathrm{add}}(E_1 \cup E_2) + \varepsilon > \sum_{j=1}^{n} m(A_j) \geq \sum_{j=1}^{n} [m(E_1^j) + m(E_2^j)] \geq m_{\mathrm{add}}(E_1) + m_{\mathrm{add}}(E_2).$$

Finally, if $\nu\colon \mathcal{R} \to [0, +\infty]$ is an additive set function and $\nu \leq m$, then, for any disjoint sets $E_1, \ldots, E_n \in \mathcal{R}$, we have

$$\sum_{j=1}^{n} m(E_j) \geq \sum_{j=1}^{n} \nu(E_j) = \nu(E),$$

whence one has $m_{\mathrm{add}} \geq \nu$.

The proof of (ii) is analogous. Given a countable collection of disjoint sets $E_n \in \mathcal{R}$, in order to obtain the estimate $m_\sigma\Bigl(\bigcup_{n=1}^{\infty} E_n\Bigr) \leq \sum_{n=1}^{\infty} m_\sigma(E_n)$, we fix ε and take partitions of E_n into sets $E_n^j \in \mathcal{R}$ such that $\sum_{j=1}^{\infty} m(E_n^j) < m_\sigma(E_n) + \varepsilon 2^{-n}$. For the proof of the opposite estimate, we observe that the finite superadditivity obviously implies the countable superadditivity: $m\Bigl(\bigcup_{j=1}^{\infty} A_j\Bigr) \geq \sum_{j=1}^{\infty} m(A_j)$ for disjoint $A_j \in \mathcal{R}$ with union in $\mathcal{R}$. □

In the situation of the above theorem, we shall call m purely superadditive if $m_{\mathrm{add}} = 0$ and purely additive if $m = m_{\mathrm{add}}$ and $m_\sigma = 0$.

3.10.15. Corollary. *Suppose that the function m in the above theorem assumes only finite values. Then $m = m_0 + m_1 + m_\sigma$, where the set function $m_0 := m - m_{\mathrm{add}}$ is purely superadditive and the set function $m_1 := m_{\mathrm{add}} - m_\sigma$ is purely additive. If $m = m_0' + m_1' + m_2$, where $m_0' \geq 0$ is purely superadditive, $m_1' \geq 0$ is purely additive and $m_2 \geq 0$ is countably additive, then $m_0' = m_0$, $m_1' = m_1$, $m_2 = m_\sigma$.*

Proof. If m_0 is not purely superadditive, i.e., $(m_0)_{\mathrm{add}} \neq 0$, then one has $m_{\mathrm{add}} + (m_0)_{\mathrm{add}} \leq m$. Since the function $m_{\mathrm{add}} + (m_0)_{\mathrm{add}}$ is additive, one has $m_{\mathrm{add}} \geq m_{\mathrm{add}} + (m_0)_{\mathrm{add}}$. Since m assumes only finite values, the function m_{add} also does. Hence $(m_0)_{\mathrm{add}} = 0$, which is a contradiction. Similarly, we verify that m_1 is purely additive. If m_0', m_1' and m_2 are functions with the properties mentioned in the formulation, then one readily verifies that

$$m_{\mathrm{add}} = (m_0')_{\mathrm{add}} + (m_1')_{\mathrm{add}} + (m_2)_{\mathrm{add}} = m_1' + m_2$$

and $m_\sigma = (m_{\mathrm{add}})_\sigma = (m_1')_\sigma + m_2$. Hence $m_0' = m_0$, $m_1' = m_1$, $m_2 = m_\sigma$. □

In particular, every nonnegative real additive set function m on a ring $\mathcal{R}$ can be written in the form $m = m_1 + m_2$, where m_2 is countably additive and m_1 is purely finitely additive, i.e., there exists no nonzero countably additive measure majorized by m_1. We note that the set function on $\mathbb{N}$ in Example 1.12.28 is a nonzero purely additive function.

Earlier we considered total variations of measures. This concept is meaningful for general set functions, too. Let $\mathcal{F}$ be some class of subsets of a space X containing some nonempty set. For a function m on $\mathcal{F}$ with values in the extended real line we set

$$v(m)(A) = \sup\Bigl\{\sum_{j=1}^{n} |m(A_j)|\colon\ n \in \mathbb{N}, A_j \in \mathcal{F} \text{ are disjoint and } A_j \subset A\Bigr\}.$$

If no such A_j exist, then we set $v(m)(A) = 0$. We shall call $v(m)$ the total variation of m. The function $v(m)$ is defined on all sets $A \subset X$ and takes values in $[0, +\infty]$. We observe that if $\varnothing \in \mathcal{F}$ and $m(\varnothing) = 0$, then in the definition of $v(m)$ one can take countable unions. It is clear that $v(m)$ is superadditive and $m \le v(m)$ on $\mathcal{F}$. In a similar way we define the total variation of a set function m on $\mathcal{F}$ with values in a normed space Y: in the definition of $v(m)$, the quantities $|m(A_j)|$ should mean $\|m(A_j)\|_Y$. For every $E \in \mathcal{F}$ set

$$m^+(E) = \sup_{F\in\mathcal{F}, F\subset E} m(F), \quad m^-(E) = -\inf_{F\in\mathcal{F}, F\subset E} m(F).$$

3.10.16. Proposition. *Let $\mathcal{R}$ be a ring of subsets of X and let m be an additive set function on $\mathcal{R}$ with values in $(-\infty, \infty]$. Then, the function $v(m)\colon \mathcal{R} \to [0, +\infty]$ is additive and $m^+ = \bigl(v(m) + m\bigr)/2$.*

The proof is left as Exercise 3.10.91.

3.10.17. Corollary. *If in the situation of Proposition* 3.10.16 *the function $v(m)$ is finite on $\mathcal{R}$, then $m = m^+ - m^-$, where m^+ and m^- are finite nonnegative additive set functions on $\mathcal{R}$.*

This decomposition of m is called the Jordan decomposition.

3.10(v). Properties of positive definite functions

In Chapter 7 (§7.13) we shall prove Bochner's theorem, according to which the class of all positive definite continuous functions on $\mathbb{R}^n$ coincides with the family of the characteristic functionals of bounded nonnegative Borel measures. In this subsection, we establish some general properties of positive definite functions.

3.10.18. Proposition. *Let φ be a positive definite function on $\mathbb{R}^n$. Then*:

(i) $\varphi(0) \ge 0$;
(ii) $\varphi(-y) = \overline{\varphi(y)}$ *and* $|\varphi(y)| \le \varphi(0)$;
(iii) *the functions* $\overline{\varphi}$ *and* $\operatorname{Re}\varphi$ *are positive definite*;
(iv) $|\varphi(y) - \varphi(z)|^2 \le 2\varphi(0)[\varphi(0) - \operatorname{Re}\varphi(y - z)]$;
(v) *the sums and products of positive definite functions are positive definite*; *in addition,* $\exp\varphi$ *is a positive definite function.*

PROOF. Assertion (i) is obtained by letting $i = 1$, $c_1 = 1$. The first claim in (ii) is seen from the inequality

$$|c_1|^2\varphi(0) + |c_2|^2\varphi(0) + c_1\overline{c_2}\varphi(y) + c_2\overline{c_1}\varphi(-y) \ge 0$$

for all $c_1, c_2 \in \mathbb{C}$, since if $\varphi(-y) \neq \overline{\varphi(y)}$, then one can pick c_1 and c_2 such that we obtain a number with a nonzero imaginary part. The second claim in (ii) follows from the first one by taking complex numbers c_1 and c_2 such that $|c_1| = c_2 = 1$ and $c_1\varphi(y) = -|\varphi(y)|$. Assertion (v) and the positive definiteness of $\overline{\varphi}$ are obvious from the definition. Hence the function $\mathrm{Re}\,\varphi$ is positive definite as well. The proof of (iv) is Exercise 3.10.92. $\square$

3.10.19. Lemma. *If φ is a measurable positive definite function on $\mathbb{R}^n$, then, for every Lebesgue integrable nonnegative function f, one has*

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \varphi(x-y)f(x)f(y)\,dx\,dy \geq 0. \tag{3.10.2}$$

If the function f is even, then

$$\int_{\mathbb{R}^n} \varphi(x) f{*}f(x)\,dx \geq 0. \tag{3.10.3}$$

In particular, for all $\alpha > 0$ we have

$$\int_{\mathbb{R}^n} \varphi(x)\exp(-\alpha|x|^2)\,dx \geq 0. \tag{3.10.4}$$

Proof. Let $k \geq 2$. Then, for all vectors $y_j \in \mathbb{R}^n$, $j = 1, \ldots, k$, we have $k\varphi(0) + \sum_{i\neq j}\varphi(y_i - y_j) \geq 0$. By using the boundedness and measurability of φ we can integrate this inequality with respect to the measure $f(y_1)\cdots f(y_k)\,dy_1\cdots dy_k$. Denoting the integral of f against Lebesgue measure by $I(f)$ and assuming that $I(f) > 0$, we obtain

$$k\varphi(0)I(f)^k + k(k-1)I(f)^{k-2}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\varphi(x-y)f(x)f(y)\,dx\,dy \geq 0.$$

Dividing by $k(k-1)I(f)^k$ and letting k to the infinity, we arrive at the required inequality. If the function f is even, then the left-hand side of (3.10.2) equals the left-hand side of (3.10.3). Finally, the function $g(x) = \exp(-\alpha|x|^2)$ can be written as $f * f$, where $f(x) = c\exp(-2\alpha|x|^2)$ and c is a positive number. This follows by the equalities $\widehat{g}(y) = (2\alpha)^{-n/2}\exp[-|y|^2/(4\alpha)]$ and $\widehat{f * f} = (2\pi)^{n/2}(\widehat{f})^2$. $\square$

3.10.20. Theorem. *Let φ be a Lebesgue measurable positive definite function on $\mathbb{R}^n$. Then φ coincides almost everywhere with a continuous positive definite function.*

Proof. Suppose first that the function φ is integrable. Let $f = \widehat{\varphi}$. The function f is bounded and continuous. We show that $f \geq 0$. Let us consider the functions

$$p_t(x) = (2\pi t)^{-n/2}\exp[-|x|^2/(2t)], \quad t > 0.$$

We observe that for every fixed x, the function $z \mapsto \exp[i(z,x)]$ equals the characteristic functional of Dirac's measure at the point x, hence is positive definite (certainly, this fact can be verified directly). Therefore, the function $z \mapsto \varphi(z)\exp[i(z,x)]$ is positive definite too. By the Parseval equality, Example 3.8.2 and (3.10.4), we obtain

$$\begin{aligned} p_t * f(x) &= \int_{\mathbb{R}^n} f(y)p_t(x-y)\,dy \\ &= (2\pi)^{-n/2}\int_{\mathbb{R}^n} \varphi(z)\exp[-i(z,x)]\exp[-t|z|^2/2]\,dz \ge 0. \end{aligned}$$

By the continuity of f we have $f * p_{1/k}(x) \to f(x)$. Hence $f \ge 0$. Let us now show that the function f is integrable. To this end, we take a sequence of functions $\psi_k(x) = \exp[-k^{-1}|x|^2/2]$ and observe that the above equality with $x = 0$ and $t = k$ yields

$$\int_{\mathbb{R}^n} f(x)\psi_k(x)\,dx = \pi^{n/2}\int_{\mathbb{R}^n} \varphi(x)p_{1/k}(x)\,dx \le \pi^{n/2}\varphi(0)$$

because p_t is a probability density. Since $\psi_k(x) \to 1$ for each x, by Fatou's theorem the function f is integrable. According to Corollary 3.8.12, the inverse Fourier transform of f equals φ a.e.

In the general case, the function $\varphi(x)\exp(-|x|^2)$ is positive definite (as the product of two positive definite functions) and integrable. We have shown that it coincides almost everywhere with a continuous function. Hence the function φ has a continuous modification ψ. We show that ψ is a positive definite function. Indeed, by the continuity one has $\psi(x) = \lim\limits_{t\to 0}\psi * p_t(x)$ for each x. However, $\psi * p_t(x) = \varphi * p_t(x)$ for all x and $t > 0$. It remains to note that $\varphi * p_t$ is a positive definite function. Indeed,

$$\varphi * p_t(x) = \lim_{\varepsilon\to 0}\varphi_\varepsilon * p_t(x),$$

where $\varphi_\varepsilon(x) = \varphi(x)\exp(-\varepsilon|x|^2)$. We already know that φ_ε coincides almost everywhere with the Fourier transform of some nonnegative integrable function g_ε. Hence $\varphi_\varepsilon * p_t$ is the Fourier transform of the nonnegative function $(2\pi)^{n/2}g_\varepsilon\widehat{p_t}$, i.e., is positive definite. Thus, ψ is a continuous positive definite function, almost everywhere equal to φ. □

This theorem does not mean, of course, that a measurable positive definite function is automatically continuous. For example, if $\varphi(0) = 1$ and $\varphi(x) = 0$ for $x \ne 0$, then φ is a discontinuous Borel positive definite function.

The reader is warned that there exist positive definite functions on the real line that are not Lebesgue measurable (Exercise 3.10.116).

3.10(vi). The Brunn–Minkowski inequality and its applications

In this subsection, we consider several classical inequalities, in which the ideas of measure theory, geometry, and analysis are interlacing in an elegant way.

3.10.21. Theorem. *Suppose that u, v, w are nonnegative Lebesgue integrable functions on $\mathbb{R}^n$ such that, for some $t \in [0,1]$, one has*

$$w\bigl(tx + (1-t)y\bigr) \geq u(x)^t v(y)^{1-t}, \quad \forall\, x, y \in \mathbb{R}^n. \tag{3.10.5}$$

Then

$$\int_{\mathbb{R}^n} w(x)\, dx \geq \biggl(\int_{\mathbb{R}^n} u(x)\, dx\biggr)^t \biggl(\int_{\mathbb{R}^n} v(y)\, dy\biggr)^{1-t}. \tag{3.10.6}$$

PROOF. It suffices to consider the case $n = 1$. The multidimensional case reduces to the one-dimensional case by Fubini's theorem. To this end, one considers the functions

$$w_1(x') = \int_{-\infty}^{+\infty} w(x', x_n)\, dx_n, \quad x' \in \mathbb{R}^{n-1},$$

and similarly defined u_1, v_1, where functions on $\mathbb{R}^n$ are written as functions on $\mathbb{R}^{n-1} \times \mathbb{R}^1$. Then the functions w_1, u_1, and v_1 satisfy the conditions of the theorem as well. Indeed,

$$\int_{-\infty}^{+\infty} w\bigl(tx' + (1-t)y', x_n\bigr)\, dx_n$$
$$\geq \biggl(\int_{-\infty}^{+\infty} u(x', x_n)\, dx_n\biggr)^t \biggl(\int_{-\infty}^{+\infty} v(y', y_n)\, dy_n\biggr)^{1-t}$$

by the one-dimensional case, since for fixed $x', y' \in \mathbb{R}^{n-1}$ we have

$$w\bigl(tx' + (1-t)y', tx_n + (1-t)y_n\bigr) \geq u(x', x_n)^t v(y', y_n)^{1-t}.$$

Thus, we shall deal with $n = 1$. In addition, it suffices to consider bounded functions u and v because one can first establish our inequality for the cut-off functions $\min(u, N)$ and $\min(v, N)$, which also satisfy our conditions. By the homogeneity we may pass to the case $\sup u = \sup v = 1$ (if one of these functions vanishes almost everywhere, then the assertion is trivial). For any $s \in [0,1]$ let

$$A(s) := \bigl\{x\colon\ u(x) \geq s\bigr\},\ B(s) := \bigl\{x\colon\ v(x) \geq s\bigr\},\ C(s) := \bigl\{x\colon\ w(x) \geq s\bigr\}.$$

Then, denoting Lebesgue measure by λ_1, we obtain by Theorem 2.9.3 that

$$\int u(x)\, dx = \int_0^1 \lambda_1\bigl(A(s)\bigr)\, ds,\ \int v(x)\, dx = \int_0^1 \lambda_1\bigl(B(s)\bigr)\, ds,$$
$$\int w(x)\, dx = \int_0^1 \lambda_1\bigl(C(s)\bigr)\, ds.$$

It follows by our hypothesis that $tA(s) + (1-t)B(s) \subset C(s)$ for all $s \in (0,1)$. This yields the estimate

$$t\lambda_1\bigl(A(s)\bigr) + (1-t)\lambda_1\bigl(B(s)\bigr) \leq \lambda_1\bigl(C(s)\bigr). \tag{3.10.7}$$

Indeed, it suffices to verify that, for arbitrary compact sets $K \subset tA(s)$ and $K' \subset (1-t)B(s)$, we have $\lambda_1(K) + \lambda_1(K') \leq \lambda_1(K + K')$. Due to the translation invariance of Lebesgue measure, we may assume that the point 0 is the supremum of K and the infimum of K'. Then $K \cup K' \subset K + K'$, hence $\lambda_1(K) + \lambda_1(K') = \lambda_1(K \cup K') \le \lambda_1(K + K')$. Estimate (3.10.7) is established. By this estimate we finally obtain

$$\begin{aligned}\int w(x)\,dx &= \int_0^1 \lambda_1\bigl(C(s)\bigr)\,ds \\ &\ge t\int_0^1 \lambda_1\bigl(A(s)\bigr)\,ds + (1-t)\int_0^1 \lambda_1\bigl(B(s)\bigr)\,ds \\ &= t\int u(x)\,dx + (1-t)\int v(y)\,dy \ge \Bigl(\int u(x)\,dx\Bigr)^t \Bigl(\int v(y)\,dy\Bigr)^{1-t},\end{aligned}$$

where the concavity of ln (or Exercise 2.12.87) is used. □

3.10.22. Corollary. *Let f and g be two nonnegative integrable Borel functions on $\mathbb{R}^n$ and let $\alpha \in (0,1)$. Set*

$$h(f,g)(x) := \sup_{y\in\mathbb{R}^n} f\Bigl(\frac{x-y}{\alpha}\Bigr)^{\alpha} g\Bigl(\frac{y}{1-\alpha}\Bigr)^{1-\alpha}.$$

Then $h(f,g)$ is a measurable function and one has

$$\int_{\mathbb{R}^n} h(f,g)(x)\,dx \ge \Bigl(\int_{\mathbb{R}^n} f(x)\,dx\Bigr)^{\alpha} \Bigl(\int_{\mathbb{R}^n} g(x)\,dx\Bigr)^{1-\alpha}. \tag{3.10.8}$$

PROOF. For all $x, z \in \mathbb{R}^n$ and $y = (1-\alpha)z$ we have

$$h(f,g)\bigl(\alpha x + (1-\alpha)z\bigr) \ge f\Bigl(\frac{\alpha x + (1-\alpha)z - y}{\alpha}\Bigr)^{\alpha} g\Bigl(\frac{y}{1-\alpha}\Bigr)^{1-\alpha},$$

which equals $f(x)^{\alpha} g(z)^{1-\alpha}$. In order to apply the above theorem, it remains to observe that the measurability of $h(f,g)$ follows by Corollary 2.12.8. If the function $h(f,g)$ is not integrable, then our inequality is trivial. □

We recall that, for any nonempty Borel sets A, B in $\mathbb{R}^n$ and any numbers $\alpha, \beta > 0$, the set $\alpha A + \beta B := \{\alpha a + \beta b, a \in A, b \in B\}$ is Souslin, hence measurable.

3.10.23. Corollary. *Suppose that μ is a probability measure on $\mathbb{R}^n$ with a density ϱ and there exists $\alpha \in (0,1)$ such that*

$$\varrho\bigl(\alpha x + (1-\alpha)y\bigr) \ge \varrho(x)^{\alpha}\varrho(y)^{1-\alpha}, \quad \forall\, x, y \in \mathbb{R}^n.$$

Then, for all nonempty Borel sets A and B, one has the inequality

$$\mu\bigl(\alpha A + (1-\alpha)B\bigr) \ge \mu(A)^{\alpha}\mu(B)^{1-\alpha}. \tag{3.10.9}$$

PROOF. Let

$$u = \varrho I_A,\ v = \varrho I_B,\ w = \varrho I_{\alpha A + (1-\alpha)B}.$$

Let $x \in A$, $y \in B$. Then $\alpha x + (1-\alpha)y \in \alpha A + (1-\alpha)B$, hence

$$w\bigl(\alpha x + (1-\alpha)y\bigr) = \varrho\bigl(\alpha x + (1-\alpha)y\bigr) \ge \varrho(x)^{\alpha}\varrho(y)^{1-\alpha} = u(x)^{\alpha}v(y)^{1-\alpha}.$$

In all other cases $u(x)^{\alpha}v(y)^{1-\alpha} = 0$. It remains to apply Theorem 3.10.21. □

A function V defined on a convex set $D(V) \subset \mathbb{R}^n$ is called convex if it is convex on the intersections of $D(V)$ with all straight lines. It is clear that the condition in the above corollary is fulfilled if the density of μ has the form $\varrho(x) = e^{-V(x)}$, where V is a convex function on $\mathbb{R}^n$. For example, one can take a function $V(x) = Q(x) + c$, where Q is a quadratic form with positive eigenvalues and $c \in \mathbb{R}^1$. A more general example: $V(x) = \theta\bigl(Q(x)\bigr) + c$, where θ is an increasing convex function on $[0, +\infty)$.

The next result is the classical Brunn–Minkowski inequality.

3.10.24. Theorem. *Let λ_n be Lebesgue measure on $\mathbb{R}^n$. Then, for all nonempty Borel sets $A, B \subset \mathbb{R}^n$, one has*

$$\lambda_n(A+B)^{1/n} \geq \lambda_n(A)^{1/n} + \lambda_n(B)^{1/n}. \tag{3.10.10}$$

PROOF. We shall assume that both sets have positive measures because otherwise the assertion is trivial. Let us consider the sets $A_0 = \lambda_n(A)^{-1/n}A$ and $B_0 = \lambda_n(B)^{-1/n}B$ and apply inequality (3.10.5) to the functions $u = I_{A_0}$, $v = I_{B_0}$, $w = I_{tA_0+(1-t)B_0}$ and the number

$$t = \frac{\lambda_n(A)^{1/n}}{\lambda_n(A)^{1/n} + \lambda_n(B)^{1/n}}.$$

Then $\lambda_n(A_0) = \lambda_n(B_0) = 1$, and we obtain the inequality

$$\lambda_n\bigl(tA_0 + (1-t)B_0\bigr) \geq \lambda_n(A_0)^t \lambda_n(B_0)^{1-t} = 1,$$

the left-hand side of which equals $\bigl(\lambda_n(A)^{1/n} + \lambda_n(B)^{1/n}\bigr)^{-n}\lambda_n(A+B)$, whence we obtain (3.10.10). $\square$

We note that the simple one-dimensional case of the Brunn–Minkowski inequality was obtained and used in the proof of Theorem 3.10.21. One more useful convexity inequality is given by the following theorem due to Anderson [**24**].

3.10.25. Theorem. *Let A be a bounded centrally symmetric convex set in $\mathbb{R}^n$ and let f be a nonnegative locally integrable function on $\mathbb{R}^n$ such that $f(x) = f(-x)$ and, for all $c > 0$, the sets $\{x\colon f(x) \geq c\}$ are convex. Then, for every $h \in \mathbb{R}^n$ and every $t \in [0, 1]$, one has*

$$\int_A f(x+th)\,dx \geq \int_A f(x+h)\,dx. \tag{3.10.11}$$

PROOF. Set $B_s(z) = \{x\colon f(x) \geq z\} \cap (A - sh)$, $z \geq 0$, $s \in [-1, 1]$. Then, by Theorem 2.9.3, one has

$$\int_A f(x+th)\,dx = \int_{A-th} f(x)\,dx = \int_0^\infty \lambda_n\bigl(B_t(z)\bigr)\,dz.$$

Hence Anderson's inequality reduces to the following inequality for measures of sets:

$$\lambda_n\bigl(B_t(z)\bigr) \geq \lambda_n\bigl(B_1(z)\bigr), \quad \forall\, z > 0. \tag{3.10.12}$$

Let us set $\alpha = (t+1)/2$ and observe that

$$\alpha B_1(z) + (1-\alpha)B_{-1}(z) \subset B_t(z).$$

Indeed, if $x \in A-h$, $f(x) \geq z$, $y \in A+h$, $f(y) \geq z$, then $\alpha x+(1-\alpha)y \in A-th$ and $f\bigl(\alpha x + (1-\alpha)y\bigr) \geq z$ by the convexity of A, the equality $2\alpha - 1 = t$ and the convexity of $\{f \geq z\}$. This inclusion and the Brunn–Minkowski inequality yield

$$\lambda_n\bigl(B_t(z)\bigr)^{1/n} \geq \alpha\lambda_n\bigl(B_1(z)\bigr)^{1/n} + (1-\alpha)\lambda_n\bigl(B_{-1}(z)\bigr)^{1/n}.$$

The sets $B_1(z)$ and $B_{-1}(z)$ are the images of each other under the central symmetry, hence have equal measures, which yields (3.10.12). □

3.10.26. Definition. *A Borel probability measure on* $\mathbb{R}^n$ *is called convex or logarithmically concave if, for all nonempty Borel sets* A *and* B *and all* $\alpha \in [0,1]$*, one has*

$$\mu\bigl(\alpha A + (1-\alpha)B\bigr) \geq \mu(A)^{\alpha}\mu(B)^{1-\alpha}.$$

3.10.27. Theorem. (i) *A probability measure* μ *on* $\mathbb{R}^n$ *with a density* ϱ *is convex precisely when there exists a convex function* V *with the domain of definition* $D(V) \subset \mathbb{R}^n$ *such that* $\varrho = \exp(-V)$ *on* $D(V)$ *and* $\varrho = 0$ *outside* $D(V)$. (ii) *A probability measure* μ *on* $\mathbb{R}^n$ *is convex precisely when it is the image of some absolutely continuous convex measure on* $\mathbb{R}^k$*, where* $k \leq n$*, under an affine mapping.*

A proof is given in Borell [**116**]. For a recent survey on the Brunn–Minkowski inequality, see Gardner [**342**].

3.10(vii). Mixed volumes

Let A and B be bounded nonempty convex Borel sets in $\mathbb{R}^n$. The function $\lambda_n(\alpha A + \beta B)$ of two variables $\alpha, \beta > 0$, where λ_n is Lebesgue measure, is a polynomial of the form

$$\lambda_n(\alpha A + \beta B) = \sum_{k=0}^{n} \alpha^{n-k}\beta^k C_n^k v_{n-k,k}(A,B),$$

where the coefficients $v_{n-k,k}(A,B)$ are independent of α, β (see Burago, Zalgaller [**143**, Ch. 4]). These coefficients are called Minkowski's mixed volumes. Note that one has $v_{n,0}(A,B) = \lambda_n(A)$, $v_{0,n}(A,B) = \lambda_n(B)$.

Let us establish the following Minkowski inequality for mixed volumes.

3.10.28. Theorem. *Let* A *and* B *be two convex compact sets of positive measure in* $\mathbb{R}^n$*. Then*

$$v_{n-1,1}(A,B)^n \geq \lambda_n(A)\lambda_n(B)^{n-1},$$

where the equality is only possible if A *and* B *are homothetic.*

PROOF. Let $B_t = (1-t)A + tB$. By the Brunn–Minkowski inequality, the function $\lambda_n(B_t)^{1/n}$ is convex. Hence the nonnegative function

$$F(t) = \lambda_n(B_t)^{1/n} - (1-t)\lambda_n(A)^{1/n} - t\lambda_n(B)^{1/n}$$

is convex on $[0,1]$. One has $F(0) = F(1) = 0$. Hence $F'(0) \geq 0$ and $F'(0) = 0$ precisely when $F = 0$. By the formula

$$\lambda_n(B_t) = \sum_{k=0}^{n} (1-t)^{n-k} t^k \frac{n!}{(n-k)!k!} v_{n-k,k}(A,B)$$

we deduce that

$$F'(0) = [v_{n-1,1}(A,B) - \lambda_n(A)]\lambda_n(A)^{(1-n)/n} + \lambda_n(A)^{1/n} - \lambda_n(B)^{1/n},$$

whence the desired inequality follows. The equality is only possible if $F = 0$, i.e., if one has the equality in the Brunn–Minkowski inequality, which implies that A and B are homothetic (see Hadwiger [**392**, Ch. V]). □

Regarding mixed volumes, see Burago, Zalgaller [**143**].

3.10(viii). The Radon transform

Let us make a remark on the Radon transform. Suppose we are given an integrable function f on $\mathbb{R}^2$ such that its restrictions to all straight lines are integrable. Denote by $\mathcal{L}$ the set of all straight lines in $\mathbb{R}^2$. Every element $L \in \mathcal{L}$ is determined by a pair (x, e), where x is a point in L and e is a directing unit vector (certainly, some pairs must be identified). The Radon transform of the function f is the function $\mathcal{R}(f)$ on $\mathcal{L}$ defined by the equality

$$\mathcal{R}(f)(L) := \int_L f\, ds,$$

where we integrate the restriction of f to L with respect to the natural Lebesgue measure on L. The question arises whether one can recover the function f from $\mathcal{R}(f)$. In fact, we even have two questions: is the transformation $\mathcal{R}$ injective and how can one effectively recover f from $\mathcal{R}(f)$? This problem was solved positively in Radon [**779**] (where several earlier related works by other authors were cited). Analogous problems arise in the case of multidimensional spaces and nonlinear manifolds, when one has to obtain some information about a function on the basis of knowledge of its integrals over a given family of surfaces. Several decades after Radon's work this problem acquired considerable importance in applied sciences in relation to computer tomography. At present, intensive investigations continue in this field, see Helgason [**419**] and Natterer [**708**].

Knowing the integrals of a function over all straight lines, we can find the integral of f over every half-space. For example, the integral over the half-space $\{x \leq c\}$ is obtained by integrating over $(-\infty, c]$ the integral of f over the vertical line passing through the point x of the real axis (in fact, it suffices to know the integral of f over almost every line with a given direction). This shows that $\mathcal{R}$ is injective because a finite measure that vanishes on all half-spaces is zero. However, the established uniqueness gives no effective recovery procedure. Explicit inversion formulae can be found in [**419**]. The Radon transform is closely connected with the Fourier transform. Indeed, let $(x, y) = s\omega$, where $s \in \mathbb{R}^1$ and ω is a unit vector. Evaluating the Fourier transform in the new coordinates with the first basis vector ω, we obtain

$$\widehat{f}(s\omega) = (2\pi)^{-1/2} \int \exp(-ist)\mathcal{R}(f)(\omega, t)\, dt,$$

where $\mathcal{R}(f)(\omega, t)$ is the integral of f over the line $\{u \in \mathbb{R}^2 \colon (u, \omega) = t\}$. Hence f can be obtained as the inverse Fourier transform of the right-hand side. However, the above-mentioned inversion formulae do not employ Fourier transforms. On a closely related problem of an explicit recovery of a measure from its values on the half-spaces, see Kostelyanec, Rešetnyak [**543**], Hačaturov [**390**]. Zalcman [**1047**] constructed an example of a non-integrable real analytic function f on $\mathbb{R}^2$ which has a zero integral over every straight line. According to Boman [**110**], there exist a function $f \in C_0^\infty(\mathbb{R}^2)$ that is not identically zero and a positive smooth function $(x, L) \mapsto \varrho_L(x)$, where $x \in \mathbb{R}^2$ and $L \in \mathcal{L}$ (the set of pairs (x, L) has a natural structure of a smooth manifold), such that the integral of $f\varrho_L$ over L vanishes for all $L \in \mathcal{L}$.

Exercises

3.10.29.° Let μ be a signed Borel measure on $\mathbb{R}^n$ that is bounded on bounded sets. Prove that if every continuous function with bounded support has the zero integral with respect to the measure μ, then $\mu = 0$.

Hint: $\mu(U) = 0$ for every bounded open set U, since the function I_U is the pointwise limit of a uniformly bounded sequence of continuous functions f_j vanishing outside U (consider the compact sets $K_j = \bigl\{x \in U_n \colon \operatorname{dist}(x, \partial U) \geq j^{-1}\bigr\}$ and take continuous functions f_j such that $f_j = 1$ on K_j, $f_j = 0$ outside U and $0 \leq f_j \leq 1$).

3.10.30. Let $\mathcal{A}$ be the algebra of all finite subsets of $\mathbb{R}$ and their complements. If A is finite, then we set

$$\mu(A) := \operatorname{Card}\bigl(A \cap (-\infty, 0]\bigr) - \operatorname{Card}\bigl(A \cap (0, +\infty)\bigr),$$

where $\operatorname{Card}(M)$ is the cardinality of M, and if the complement of A is finite, then we set $\mu(A) := -\mu(\mathbb{R}^1 \backslash A)$. Show that μ is a countably additive signed measure on the algebra $\mathcal{A}$, but μ has no countably additive extensions to the σ-algebra $\sigma(\mathcal{A})$ (even if we admit measures with values in $[-\infty, +\infty)$ or $(-\infty, +\infty]$).

Hint: see Dudley [**251**] or Wise, Hall [**1022**, Example 4.17]. The countable additivity is verified directly. The absence of countably additive extensions to $\sigma(\mathcal{A})$ follows from the fact that the range of μ on $\mathcal{A}$ is not bounded from below (nor from above).

3.10.31. (i) Let μ be a finite nonnegative measure on a σ-algebra $\mathcal{A}$ in a space X and let ν be a countably additive measure on $\mathcal{A}$ with values in $[0, +\infty]$ such that $\nu \ll \mu$. Show that there exists a set $S \in \mathcal{A}$ such that the measure $\nu|_S$ assumes only the values 0 and $+\infty$ and the measure $\nu|_{X \backslash S}$ is σ-finite.

(ii) Deduce from (i) that, given σ-finite measures $\mu \geq 0$ and $\nu \geq 0$ with $\nu \ll \mu$ on a σ-algebra $\mathcal{A}$, for every sub-σ-algebra $\mathcal{B} \subset \mathcal{A}$, there is a $\mathcal{B}$-measurable function

ξ such that $\nu|_B = \xi \cdot \mu|_B$ for every $B \in \mathcal{B}$ with $\mu(B) + \nu(B) < \infty$. Show that this is not true for all $B \in \mathcal{B}$ in the case where μ is Lebesgue measure on $\mathbb{R}^1$, $\nu = \varrho \cdot \mu$ is a probability measure, and $\mathcal{B}$ is generated by all singletons.

HINT: consider the class $\mathcal{S}$ of all sets in $\mathcal{A}$ that have no subsets of finite nonzero ν-measure; observe that any set of infinite ν-measure in $\mathcal{S}$ has positive μ-measure and show that there exists a set $S \in \mathcal{A}$ such that $X \backslash S$ contains no sets in $\mathcal{S}$ of infinite ν-measure; verify that the measure $\nu|_{X \backslash S}$ is σ-finite by using that μ does not vanish on sets of positive ν-measure. See also Vestrup [**976**, §9.2].

3.10.32.° Suppose we are given three bounded measures μ_1, μ_2, and μ_3 on a σ-algebra $\mathcal{A}$ such that $\mu_1 \ll \mu_2$ and $\mu_2 \ll \mu_3$. Show that one has $\mu_1 \ll \mu_3$ and $d\mu_1/d\mu_3 = (d\mu_1/d\mu_2)(d\mu_2/d\mu_3)$.

3.10.33.° Let μ and ν be two probability measures on a σ-algebra $\mathcal{A}$ such that for some $\alpha \in (0,1)$, one has $\|\alpha\mu - (1-\alpha)\nu\| = 1$. Prove that $\mu \perp \nu$.

HINT: let $\mu = f \cdot \sigma$, $\nu = g \cdot \sigma$, where $\sigma = (\mu + \nu)/2$. Then the integral of $|\alpha f - (1-\alpha)g|$ against the measure σ equals 1, which is possible only if $fg = 0$ σ-a.e., since the integral of $\alpha f + (1-\alpha)g$ equals 1.

3.10.34.° Let μ and ν be two probability measures such that $\nu \ll \mu$. Show that if a sequence of μ-measurable functions f_n converges in measure μ to a function f, then it converges to f in measure ν as well.

3.10.35.° Let μ and ν be two probability measures and let f_n, $n \in \mathbb{N}$, and f be $\mu\otimes\nu$-measurable functions such that for μ-a.e. fixed x the functions $f_n(\,\cdot\,, x)$ converge to $f(\,\cdot\,, x)$ in measure ν. Show that the functions f_n converge to f in measure $\mu\otimes\nu$.

HINT: use Fubini's theorem to show that the integrals of $|f_n - f|/(|f - f_n| + 1)$ with respect to $\mu\otimes\nu$ tend to zero.

3.10.36. Suppose that a sequence of measures μ_n on a measurable space $(X, \mathcal{A})$ converges in variation to a measure μ and a sequence of measures ν_n converges in variation to a measure ν. Let $\nu_n = \nu_n^{ac} + \nu_n^s$, $\nu = \nu^{ac} + \nu^s$, where $\nu_n^{ac} \ll \mu_n$, $\nu_n^s \perp \mu_n$, $\nu^{ac} \ll \mu$, $\nu^s \perp \mu$. Prove that $\mathcal{A}$-measurable versions of the Radon–Nikodym densities $d\nu_n^{ac}/d\mu_n$ converge to $d\nu^{ac}/d\mu$ in measure $|\mu|$. In particular, if $\mu_n \ll \mu$ and $\nu_n \ll \mu_n$, then $d\nu_n/d\mu_n \to d\nu/d\mu$ in measure $|\mu|$.

HINT: let $\sigma := |\mu| + |\nu| + \sum_{n=1}^\infty 2^{-n}(|\mu_n| + |\nu_n|)(\|\mu_n\| + \|\nu_n\|)^{-1}$; one has $\mu_n = f_n \cdot \sigma$, $\mu = f \cdot \sigma$, $\nu_n = g_n \cdot \sigma$, $\nu = g \cdot \sigma$, where f_n, g_n, f, g are $\mathcal{A}$-measurable functions from $\mathcal{L}^1(\sigma)$. Clearly, $f_n \to f$ and $g_n \to g$ in $L^1(\sigma)$, hence in measure σ. This yields convergence of the functions $I_{\{f_n \neq 0\}} g_n/f_n$ to $I_{\{f\neq 0\}} g/f$ in measure σ, hence in measure $|\mu|$. These functions serve as the aforementioned Radon–Nikodym densities.

3.10.37. (Nikodym [**717**]) Let μ be a bounded nonnegative measure on a σ-algebra $\mathcal{A}$ in a space X, let G be a nonmeasurable set. Let $\sigma(\mathcal{A}\cup G)$ be the σ-algebra generated by $\mathcal{A}$ and G, and let $\underline{G}$ and $\widetilde{G}$ be a measurable kernel and a measurable envelope of G. Denote by γ_1 and γ_2 the Radon–Nikodym densities of the measures $A \mapsto \mu(A \cap \underline{G})$ and $A \mapsto \mu(A \cap \widetilde{G})$ with respect to μ. Let γ be a μ-measurable function such that $\gamma_1 \le \gamma \le \gamma_2$. Show that the formula

$$\nu(E) = \int_A \gamma(x)\,\mu(dx) + \int_B \bigl(1 - \gamma(x)\bigr)\,\mu(dx),$$

where $E = (A\cap G)\cup\bigl(B\cap(X\backslash G)\bigr)$, $A, B \in \mathcal{A}$, defines a countably additive extension of μ to $\sigma(\mathcal{A}\cup G)$ and that every countably additive extension of μ to $\sigma(\mathcal{A}\cup G)$ has such a form.

3.10.38.° Let $(X,\mathcal{A})$ and $(Y,\mathcal{B})$ be two measurable spaces. Show that every set in $\mathcal{A}\otimes\mathcal{B}$ is contained in the σ-algebra generated by sets $A_n\times B_n$ for some at most countable collections $\{A_n\}\subset\mathcal{A}$ and $\{B_n\}\subset\mathcal{B}$.

HINT: see Problem 1.12.54.

3.10.39.° Let $(X,\mathcal{A})$ and $(Y,\mathcal{B})$ be two measurable spaces and let a mapping $f\colon A\to Y$ be $(\mathcal{A},\mathcal{B})$-measurable. Show that the mapping $\varphi\colon x\mapsto\bigl(x,f(x)\bigr)$ from X to $X\times Y$ is $(\mathcal{A},\mathcal{A}\otimes\mathcal{B})$-measurable. Deduce from this that, given a measurable space $(Z,\mathcal{E})$ and a mapping $g\colon X\times Y\to Z$ measurable with respect to the pair $(\mathcal{A}\otimes\mathcal{B},\mathcal{E})$, the mapping $x\mapsto g\bigl(x,f(x)\bigr)$ from X to Z is $(\mathcal{A},\mathcal{E})$-measurable.

HINT: the first claim is seen from the fact that $\varphi^{-1}(A\times B)=A\cap f^{-1}(B)\in\mathcal{A}$ for all $A\in\mathcal{A}$, $B\in\mathcal{B}$, and $\mathcal{A}\otimes\mathcal{B}$ is generated by the products $A\times B$. The second claim readily follows from this.

3.10.40. Let $T=\{(x,y)\in[0,1]^2\colon\ x-y\in\mathbb{Q}\}$. Show that T has measure zero, but meets every set of the form $A\times B$, where A and B are sets of positive measure in $[0,1]$. See also Exercise 3.10.63.

HINT: use that $A-B$ contains an interval.

3.10.41.° Suppose that a function f on $[0,1]^2$ is Lebesgue measurable and that, for a.e. x and a.e. y, the functions $z\mapsto f(x,z)$ and $z\mapsto f(z,y)$ are constant. Show that $f=c$ a.e. for some constant c.

HINT: otherwise there is a number r such that the measures of the sets $\{f<r\}$ and $\{f\geq r\}$ are positive. By hypothesis and Fubini's theorem, these sets contain horizontal and vertical unit intervals and hence meet, which is a contradiction.

3.10.42. Let μ and ν be finite nonnegative measures on measurable spaces $(X,\mathcal{A})$ and $(Y,\mathcal{B})$, $A\subset X$, $B\subset Y$. Prove the equality $(\mu\otimes\nu)^*(A\times B)=\mu^*(A)\nu^*(B)$.

HINT: by considering measurable envelopes one obtains

$$(\mu\otimes\nu)^*(A\times B)\leq\mu^*(A)\nu^*(B).$$

If $\mu^*(A)\nu^*(B)=0$, then the claim is obvious. The general case reduces easily to the case $\mu^*(A)=\nu^*(B)=1$; if $(\mu\otimes\nu)^*(A\times B)<1$, then there exists $E\in\mathcal{A}\otimes\mathcal{B}$ with $A\times B\subset E$ and $\mu\otimes\nu(E)<1$. By Fubini's theorem there exists $y\in Y$ with $\mu(E_y)<1$, and it remains to observe that $A\subset E_y$, whence $\mu^*(A)<1$, which is a contradiction. One could also use Theorem 1.12.14 and extend the measures μ and ν to the sets A and B in such a way that the extensions equal $\mu^*(A)$ and $\nu^*(B)$ on A and B, respectively.

3.10.43. Let $(X,\mathcal{A})$ and $(Y,\mathcal{B})$ be measurable spaces. Show that, for every $E\in\mathcal{A}\otimes\mathcal{B}$, the family of sections $E_x=\{y\in Y\colon\ (x,y)\in E\}$ contains at most continuum of distinct sets.

HINT: by Exercise 3.10.38, the set E belongs to the σ-algebra generated by sets $A_n\times B_n$ for some at most countable collections $\{A_n\}\subset\mathcal{A}$ and $\{B_n\}\subset\mathcal{B}$; for every $x\in X$, we consider the sequence $\{I_{A_n}(x)\}$ and verify that if $I_{A_n}(x_1)=I_{A_n}(x_2)$ for all n, then $E_{x_1}=E_{x_2}$; hence the cardinality of the family of distinct sections of E does not exceed the cardinality of the family of all sequences of 0 and 1.

3.10.44. Let $(X, \mathcal{A})$ be a measurable space of cardinality greater than that of the continuum. Show that the diagonal $D = \{(x,x), x \in X\}$ does not belong to the σ-algebra $\mathcal{A}\otimes\mathcal{A}$.

HINT: use Exercise 3.10.43.

3.10.45.° Construct examples showing that (a) the existence and equality of the repeated integrals in (3.4.3) do not guarantee the $\mu\otimes\nu$-integrability of a measurable function f; (b) it may occur that both repeated integrals exist for some measurable function f, but are not equal; (c) there exists a measurable function f such that one of the repeated integrals exists, but the other one does not.

3.10.46. (Minkowski's inequality for integrals) Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be spaces with nonnegative σ-finite measures and let f be an $\mathcal{A}\otimes\mathcal{B}$-measurable function. Prove that whenever $1 \le p < q < \infty$ one has

$$\int_Y \biggl(\int_X |f(x,y)|^p \,\mu(dx)\biggr)^{q/p} \nu(dy) \le \biggl(\int_X \biggl(\int_Y |f(x,y)|^q \,\nu(dy)\biggr)^{p/q} \mu(dx)\biggr)^{q/p}.$$

HINT: it suffices to consider the case $p = 1$, $q > 1$; then the integral on the left can be written by Fubini's theorem as

$$\int_X \int_Y \biggl(\int_X |f(x,y)|\,\mu(dx)\biggr)^{q-1} |f(z,y)|\,\nu(dy)\,\mu(dz),$$

which by Hölder's inequality with the exponents $q/(q-1)$ and q (applied to the inner integral against ν) is majorized by

$$\int_X \biggl[\int_Y \biggl(\int_X |f(x,y)|\,\mu(dx)\biggr)^q \nu(dy)\biggr]^{(q-1)/q} \biggl[\int_Y |f(z,y)|^q \,\nu(dy)\biggr]^{1/q} \mu(dz)$$
$$= \biggl[\int_Y \biggl(\int_X |f(x,y)|\,\mu(dx)\biggr)^q \nu(dy)\biggr]^{(q-1)/q} \int_X \biggl[\int_Y |f(z,y)|^q \,\nu(dy)\biggr]^{1/q} \mu(dz).$$

3.10.47.° Prove the equalities

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\Bigl(-\frac{1}{2}t^2\Bigr)\,dt = 1, \quad \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 \exp\Bigl(-\frac{1}{2}t^2\Bigr)\,dt = 1.$$

HINT: evaluate the integral

$$\int\!\!\int \exp(-x^2 - y^2)\,dx\,dy$$

in two ways: by Fubini's theorem and in polar coordinates. The second equality can be derived from the integration by parts formula, since the derivative of $\exp(-t^2/2)$ is $-t\exp(-t^2/2)$.

3.10.48.° Let $e_1, \ldots, e_n$ be a basis in $\mathbb{R}^n$. Prove that a Lebesgue measurable set $A \subset \mathbb{R}^n$ has measure zero precisely when it can be written in the following form: $A = A_1 \cup \cdots \cup A_n$, where the sets A_j are measurable and, for every index j and every $x \in \mathbb{R}^n$, the set $\{t \in \mathbb{R}\colon\ x + te_j \in A_j\}$ has measure zero on the real line (in other words, the sections of A_j by the straight lines parallel to e_j have zero linear measures).

HINT: the sufficiency of the above condition is clear from Fubini's theorem. In the proof of necessity we may assume that $\{e_j\}$ is a standard basis and use induction on n. By Fubini's theorem, the set B of all points $y \in \mathbb{R}^{n-1}$ such that the set $\{t \in \mathbb{R}\colon\ y + te_n \in A\}$ is not measurable or has nonzero measure, has measure zero in $\mathbb{R}^{n-1}$. For A_n we take $A\cap\bigl((\mathbb{R}^{n-1}\backslash B)\times\mathbb{R}e_n\bigr)$, and represent B in the form $B_1\cup\cdots\cup B_{n-1}$, where all sections of B_j by the straight lines parallel to e_j have zero linear measures. Finally, let $A_j := A\cap(B_j\times\mathbb{R}e_n)$ for $j\le n-1$.

3.10.49. (Sierpiński [**872**]) (i) Show that in the plane (or in the unit square) there exists a Lebesgue nonmeasurable set that meets every straight line parallel to one of the coordinate axes in at most one point.

(ii) Show that in the plane there is a nonmeasurable set whose intersection with every straight line has at most two points.

HINT: (i) use that the family of compacts of positive measure in the square has cardinality $\mathfrak{c}$ of the continuum and write it in the form $\{K_\alpha, \alpha<\omega(\mathfrak{c})\}$, where α are ordinal numbers and $\omega(\mathfrak{c})$ is the smallest ordinal number of cardinality of the continuum; construct the required set A by transfinite induction by choosing in every K_α a point (x_α, y_α) as follows: if points $(x_\beta, y_\beta)\in K_\beta$ are already chosen for $\beta<\alpha<\omega(\mathfrak{c})$ such that no two of them belong to a straight line parallel to one of the coordinate axes, then $K_\alpha\backslash\bigcup_{\beta<\alpha}\{(x_\beta, y_\beta)\}$ contains a point (x_α, y_α) such that the straight lines $x_\alpha\times\mathbb{R}$ and $\mathbb{R}^1\times y_\alpha$ contain no points from $\bigcup_{\beta<\alpha}\{(x_\beta, y_\beta)\}$ (otherwise K_α would have measure zero by Fubini's theorem, since the cardinality of the set $\{\beta<\alpha\}$ is than $\mathfrak{c}$); finally, let $A=\{(x_\alpha, y_\alpha), \alpha<\omega(\mathfrak{c})\}$. Example (ii) is similar, see the cited paper.

3.10.50. Show that there exists a bounded nonnegative function f on the square $[0,1]\times[0,1]$ such that it is not Lebesgue measurable, but the repeated integrals

$$\int_0^1\int_0^1 f(x,y)\,dxdy \quad\text{and}\quad \int_0^1\int_0^1 f(x,y)\,dydx$$

exist and vanish.

HINT: use the previous exercise.

3.10.51. (Sierpiński [**873**]) (i) Assuming the continuum hypothesis construct a set $S\subset[0,1]^2$ such that all its vertical sections are at most countable and all its horizontal sections have at most countable complements. Observe that the repeated integrals of I_S exist and are different.

(ii) Without use of the continuum hypothesis construct a measurable space X with a probability measure μ and a set $S\in X^2$ such that the repeated integrals

$$\int_X\int_X I_S(x,y)\,\mu(dx)\mu(dy) \quad\text{and}\quad \int_X\int_X I_S(x,y)\,\mu(dy)\mu(dx)$$

exist and are not equal.

(iii) Under the continuum hypothesis construct a set $E\subset[0,1]^2$ such that its indicator function I_E is measurable in every variable separately, the function

$$x\mapsto\int_0^1 I_E(x,y)\,dy$$

is measurable, but the function

$$y\mapsto\int_0^1 I_E(x,y)\,dx$$

is not.

HINT: (i) by means of the continuum hypothesis one can find a linear ordering of $[0,1]$ such that every point is preceded by at most countably many elements. Let S be the class of all pairs $(x,y)\in[0,1]^2$ such that x precedes y. (ii) Take for X

the set of all ordinal numbers smaller than the first uncountable ordinal number, consider the σ-algebra $\mathcal{A}$ of all sets that are either at most countable or have at most countable complements, and define the measure μ on $\mathcal{A}$ as follows: $\mu(A) = 0$ if A is at most countable and $\mu(A) = 1$ otherwise. Let S be the set of all pairs (x, y) such that $x \le y$. (iii) Take a nonmeasurable set $D \subset [0,1]$ and consider $E := S \cap ([0,1] \times D)$. The first function above is zero and the second one is I_D.

3.10.52.° Prove that the graph of a measurable real function on a measure space $(X, \mathcal{A}, \mu)$ with a finite measure μ has measure zero with respect to $\mu\otimes\lambda$, where λ is Lebesgue measure.

HINT: the claim reduces to the case of bounded f; then, for every n, the graph of f is covered by a finite collection of sets of the form

$$f^{-1}\bigl([r_i - n^{-1}, r_i + n^{-1})\times[r_i - n^{-1}, r_i + n^{-1})\bigr),$$

and the measure of their union is at most $2\|\mu\|n^{-1}$. An alternative reasoning: use that the graph is measurable and apply Fubini's theorem.

3.10.53.° Let $(X, \mathcal{A}_X)$ and $(Y, \mathcal{A}_Y)$ be measurable spaces and let $f\colon X \to Y$ be a mapping. Construct examples showing that:

(i) even if f is $(\mathcal{A}_X, \mathcal{A}_Y)$-measurable, its graph may not belong to $\mathcal{A}_X\otimes\mathcal{A}_Y$;

(ii) the graph f may belong to $\mathcal{A}_X\otimes\mathcal{A}_Y$ without f being measurable.

Prove that if the set $\{(y, y), y \in Y\}$ belongs to $\mathcal{A}_Y\otimes\mathcal{A}_Y$, then the graph of any $(\mathcal{A}_X, \mathcal{A}_Y)$-measurable mapping belongs to $\mathcal{A}_X\otimes\mathcal{A}_Y$.

HINT: (i) consider the identity mapping from $[0,1]$ with the σ-algebra generated by singletons to the same space; (ii) consider the identity mapping from $[0,1]$ with the standard Borel σ-algebra to $[0,1]$ with the σ-algebra of all Lebesgue measurable sets. The last claim follows by the measurability of the mapping $(x, y) \mapsto \bigl(f(x), y\bigr)$ with respect to the pair $(\mathcal{A}_X\otimes\mathcal{A}_Y, \mathcal{A}_Y\otimes\mathcal{A}_Y)$. See also Corollary 6.10.10 in Chapter 6.

3.10.54. Show that under the continuum hypothesis the plane can be covered by countably many graphs of functions $y = y(x)$ and $x = x(y)$. In particular, there exists a nonmeasurable graph among them.

HINT: consider the set S from Exercise 3.10.51(i); for every y, there exists an at most countable set of points $g_n(y)$ with $\bigl(g_n(y), y\bigr) \in S$, for every x, there exists an at most countable set of points $f_n(x)$ with $\bigl(x, f_n(x)\bigr) \notin S$. If $(x, y) \in S$, then (x, y) belongs to the graph of $x = g_n(y)$ for some n, and if $(x, y) \notin S$, then (x, y) belongs to the graph of $y = f_n(x)$ for some n.

3.10.55. (Fichtenholz [**291**]) There exists a measurable function f on $[0,1]^2$ such that f is not integrable, but for all measurable sets $A, B \subset [0,1]$, the repeated integrals

$$\int_A\int_B f(x,y)\,dxdy \quad\text{and}\quad \int_B\int_A f(x,y)\,dydx$$

exist, are finite and equal.

3.10.56. Let f be a Riemann integrable function on $[0,1]^2$.

(i) Prove that for almost every $x \in [0,1]$, the function $y \mapsto f(x, y)$ is Riemann integrable and the function $\varphi\colon x \mapsto \varphi(x)$, where $\varphi(x)$ equals the Riemann integral

$$\int_0^1 f(x,y)\,dy$$

if it exists and the lower Riemann integral otherwise, is Riemann integrable.

(ii) Prove that if at all points x where the Riemann integral in y does not exist, we redefine φ to be zero, then the obtained function may not be Riemann integrable (although it remains Lebesgue integrable and its Lebesgue integral is unchanged).

HINT: see Zorich [**1053**, Ch. XI, §4].

3.10.57. (Fichtenholz [**285**], Lichtenstein [**611**]) Let f be a bounded function on the square $[0,1]\times[0,1]$ such that, for every fixed y, the function $x \mapsto f(x,y)$ is Riemann integrable, and, for every fixed x, the function $y \mapsto f(x,y)$ is Lebesgue integrable.

(i) Prove that the function

$$F_1(x) = \int_0^1 f(x,y)\,dy$$

is Riemann integrable, the function

$$F_2(y) = \int_0^1 f(x,y)\,dx$$

is Lebesgue integrable, and their respective integrals are equal.

(ii) Prove that if the function $y \mapsto f(x,y)$ also is Riemann integrable for every x, then the repeated Riemann integrals of f exist and are equal. Note, however, that in this situation f may not be Lebesgue integrable over the square.

HINT: the function $F_2(y)$ is the pointwise limit of the functions

$$S_n(y) = n^{-1}\sum_{k=1}^{n} f(k/n, y),$$

hence is measurable; let J be its Lebesgue integral; for any partition of $[0,1]$ into finitely many intervals $[a_i, a_{i+1})$, $1 \le i \le n$, and any choice of points $x_i \in [a_i, a_{i+1})$, the functions $T_n(y) = \sum\limits_{i=1}^{n} f(x_i, y)(a_{i+1}-a_i)$ converge to $F_2(y)$ as $\max(a_{i+1}-a_i) \to 0$, hence by the dominated convergence theorem one has

$$\lim_{n\to\infty}\sum_{i=1}^{n} F_1(x_i)(a_{i+1}-a_i) = \lim_{n\to\infty}\int_0^1 T_n(y)\,dy = J;$$

thus, F_1 is Riemann integrable and J is its integral; the last claim follows from the already-proven facts. The indicator of the set from Exercise 3.10.49 gives an example of a nonmeasurable function with the required properties.

3.10.58. Let $X = Y = [0,1]$, let λ^* be Lebesgue outer measure, and let $\nu^*(A)$ be the cardinality of a set A. Show that the diagonal D of the square $[0,1]^2$ is measurable with respect to $\lambda^*\times\nu^*$ in the sense of Theorem 3.10.1, but the repeated integrals of I_D against $d\nu^* d\lambda^*$ and $d\lambda^* d\nu^*$ equal, respectively, 1 and 0.

HINT: for the verification of measurability use that by Theorem 3.10.1 all open rectangles are measurable.

3.10.59. (i) (Davies [**206**]) Let $E \subset \mathbb{R}^2$ be a Lebesgue measurable set of finite measure. Then, there exists a family L of straight lines in $\mathbb{R}^2$ such that the union of all these lines is measurable and has the same measure as E and every point E belongs to at least one line from L. A multidimensional analog is obtained in Falconer [**276**].

(ii) (Csőrnyei [**195**]) Prove that the assertion analogous to (i) is true for every σ-finite Borel measure on the plane.

3.10.60. (Falconer [**276**]) Let A be a set of Lebesgue measure zero in $\mathbb{R}^n$ and let $1 < k < n$. Denote by $G_{n,k}$ the space of all k-dimensional linear subspaces in $\mathbb{R}^n$ equipped with its natural measure (see Federer [**282**]; for the purposes of this exercise it suffices to embed $G_{n,k}$ into $\mathbb{R}^{kn}$ and consider the corresponding measure). Prove that, for almost all $\Pi \in G_{n,k}$, all sections of A by the planes parallel to Π have k-dimensional measure zero.

3.10.61. (Talagrand [**931**, p. 115]) Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be probability spaces and let $E \in \mathcal{A}\otimes\mathcal{B}$, $\mu\otimes\nu(E) = \varepsilon > 0$. Show that there exists a set $A \in \mathcal{A}$ with the following property: $\mu(A) > 0$ and for every $k \in \mathbb{N}$ there exists $\varepsilon_k > 0$ such that $\nu\bigl(\bigcap_{i=1}^k E_{x_i}\bigr) \geq \varepsilon_k$ for all $x_1, \ldots, x_k \in A$, where $E_x := \bigl\{y\colon\ (x, y) \in E\bigr\}$.

3.10.62. (Erdős, Oxtoby [**271**]) Let $(X_1, \mathcal{A}_1, \mu_1)$ and $(X_2, \mathcal{A}_2, \mu_2)$ be probability spaces with atomless measures. Show that there exists a set $A \in \mathcal{A}_1\otimes\mathcal{A}_2$ such that $\mu_1\otimes\mu_2(A) > 0$ and if $A_i \in \mathcal{A}_i$ and $\mu_1(A_1)\mu_2(A_2) > 0$, then $\mu_1\otimes\mu_2\bigl((A_1\times A_2)\backslash A\bigr) > 0$.

3.10.63. (i) (Brodskiĭ [**130**], Eggleston [**264**]) Let a set $E \subset [0, 1]\times[0, 1]$ have Lebesgue measure 1. Prove that there exist a nonempty perfect set $P \subset [0, 1]$ and a compact set $K \subset [0, 1]$ of positive measure such that $P\times K \subset E$.

(ii) (Davies [**208**]) Suppose that every union of less than $\mathfrak{c}$ Lebesgue measure zero sets has measure zero (which holds, e.g., under the continuum hypothesis or Martin's axiom). Prove that every measurable set $E \subset [0, 1]^2$ of Lebesgue measure 1 contains a product-set $X\times Y$ such that X and Y in $[0, 1]$ have outer measure 1.

3.10.64. Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be measure spaces, where μ and ν take values in $[0, +\infty]$. Denote by λ_{max} the measure corresponding to the Carathéodory outer measure generated by the set function $\tau(A\times B) = \mu(A)\nu(B)$ on the class of all sets $A\times B$, where $A \in \mathcal{A}$, $B \in \mathcal{B}$. Let Λ be the domain of definition of λ_{max} according to the Carathéodory construction. Let λ_{min} denote the set function on Λ with values in $[0, +\infty]$ defined by the formula

$$\lambda_{min}(L) = \sup\bigl\{\lambda_{max}\bigl(L \cap (A\times B)\bigr)\colon\ A \in \mathcal{A},\ \mu(A) < \infty,\ B \in \mathcal{B},\ \nu(B) < \infty\bigr\}.$$

(i) Show that $\mathcal{A}\otimes\mathcal{B} \in \Lambda$ and $\lambda_{max}(A\times B) = \mu(A)\nu(B)$ for all $A \in \mathcal{A}$, $B \in \mathcal{B}$.

(ii) Show that $\lambda_{min}(A\times B) = \mu(A)\nu(B)$ if $A \in \mathcal{A}$, $B \in \mathcal{B}$ and $\mu(A)\nu(B) < \infty$.

(iii) Show that $\lambda_{min}(E) = \lambda_{max}(E)$ if $\lambda_{max}(E) < \infty$.

(iv) Let λ be a measure on $\mathcal{A}\otimes\mathcal{B}$ with values in $[0, +\infty]$ such that $\lambda(A\times B) = \mu(A)\nu(B)$ for all $A \in \mathcal{A}$, $B \in \mathcal{B}$. Show that $\lambda_{min}(E) \leq \lambda(E) \leq \lambda_{max}(E)$ for all $E \in \mathcal{A}\otimes\mathcal{B}$.

(v) Show that the measures λ_{min} and λ_{max} possess equal collections of integrable functions and the corresponding integrals coincide.

Hint: see, e.g., Fremlin [**327**, §251].

3.10.65. Let μ, ν, λ_{min}, and λ_{max} be the same as in Exercise 3.10.64. Show that the following conditions are equivalent: (i) $\lambda_{min} = \lambda_{max}$, (ii) λ_{max} is semifinite, (iii) λ_{max} is locally determined.

3.10.66. Let μ, ν, λ_{min}, and λ_{max} be the same as in Exercise 3.10.64.

(i) Let μ and ν be decomposable measures. Prove that the measure λ_{min} is decomposable.

(ii) Show that there exist a Maharam measure μ and a probability measure ν such that the measure λ_{min} is not Maharam.

Hint: see Fremlin [**327**, 251N, 254U].

3.10.67. (Luther [**639**]) Let $X = Y = [0,1]$, let $\mathcal{A} = \mathcal{B}([0,1])$, and let the measure $\mu = \nu$ with values in $[0,+\infty]$ be defined as follows: we fix a non-Borel set E; then every point x is assigned the measure 2 or 1 depending on whether x belongs to E or not, finally, the measure extends naturally to all Borel sets (in particular, all infinite sets obtain infinite measures). Let π be the Carathéodory extension of the measure $\mu\otimes\nu$. Prove that the measure π is semifinite, $\mu = \nu$ is semifinite and complete, but for the diagonal D in $[0,1]\times[0,1]$ the function $\nu(D_x) = I_E(x) + 1$ is not measurable with respect to μ.

3.10.68.° Construct a signed bounded measure μ on $\mathbb{N}$, a mapping $f\colon \mathbb{N} \to \mathbb{N}$ and a function g on $\mathbb{N}$ such that $\mu\circ f^{-1} = 0$, but the function $g\circ f$ is not integrable with respect to μ (although g is integrable against the measure $\mu\circ f^{-1}$).

HINT: let $\mu(2n) = n^{-2}$, $\mu(2n-1) = -n^{-2}$, $f(2n) = f(2n-1) = n$, $g(n) = n$.

3.10.69. Let $f \in \mathcal{L}^1(\mathbb{R}^1)$. Prove that the function $f(x-x^{-1})$ is integrable and one has

$$\int_{-\infty}^{+\infty} f(x - x^{-1})\,dx = \int_{-\infty}^{+\infty} f(x)\,dx.$$

HINT: change the variable $y = -x^{-1}$ on the left and observe that the integral on the left equals half of the integral of the function $f(x-x^{-1})(1+x^{-2})$, then use the change of variable $z = x - x^{-1}$, which gives the integral on the right.

3.10.70. Prove that there exists a continuous function f on $[0,1]$ that is constant on no interval, but $f(x)$ is a rational number for a.e. x.

HINT: let μ be a probability measure on $[0,1]$ concentrated on the set of all rational numbers. It is easily verified that there exists a continuous function $f\colon [0,1] \to [0,1]$ such that $\mu = \lambda\circ f^{-1}$ (in §9.7 a considerably more general fact is established). Hence the set F of all continuous functions $f\colon [0,1] \to [0,1]$ such that $\mu = \lambda\circ f^{-1}$ is nonempty. This set is closed in the space $C[0,1]$ of all continuous functions, which is complete with the metric $d(\varphi,\psi) = \sup|\varphi(t) - \psi(t)|$. Hence F itself is a complete metric space with the above metric. If F contains no function that is nonconstant on every interval, then F is the union of a countable family of sets F_n each of which consists of functions assuming some rational value r on some interval (p,q) with rational endpoints. By Baire's theorem (Exercise 1.12.83), there exists F_n containing a ball U with some center f_0 and some radius $d > 0$. This leads to a contradiction, since one can find in U a function $\psi \in F$ nonconstant on (p,q). To this end, it suffices to find a continuous function $\psi\colon [0,1] \to [0,1]$ such that $\psi(t) = f_0(t)$ for $t \notin [p-\delta, p+\delta]$ for sufficiently small $\delta > 0$, $|\psi(t) - f_0(t)| < d$ for all other t, $\psi(p) < r$, and such that ψ transforms Lebesgue measure λ on $[p-\delta, p+\delta]$ to the measure $\lambda|_{[p-\delta,p+\delta]} \circ f_0^{-1}$.

3.10.71.° Let E be a set of finite measure on the real line and let $\alpha_n \to +\infty$. Prove that

$$\lim_{n\to\infty} \int_E (\sin\alpha_n t)^2\,dt = \lambda(E)/2.$$

HINT: $2(\sin\alpha_n t)^2 = 1 - \cos 2\alpha_n t$, the integral of $\cos(2\alpha_n t)I_E$ tends to zero.

3.10.72.° Let a sequence of real numbers α_n be such that $f(x) := \lim\limits_{n\to\infty} \sin(\alpha_n x)$ exists on a set E of positive measure. Prove that $\{\alpha_n\}$ has a finite limit.

HINT: consider the case where the measure E is finite and $\{\alpha_n\}$ has two finite limit points α and β and observe that the functions $\sin\alpha x$ and $\sin\beta x$ cannot coincide on an uncountable set; show that $\{\alpha_n\}$ cannot tend to $+\infty$ or $-\infty$ because then $f = 0$ a.e. on E, since the integral of $g(x)\sin(\alpha_n x)$ approaches zero for every integrable function g; now the limit of the integrals of $(\sin\alpha_n x)^2$ over E must vanish, but this limit is $\lambda(E)/2$.

3.10.73.° Prove that there exists a Lebesgue measurable one-to-one mapping f of the real line onto itself such that the inverse mapping is not Lebesgue measurable.

HINT: the complement to the Cantor set C can be transformed onto $[0,\infty)$ by an injective Borel mapping, and C can be mapped injectively onto $(-\infty, 0)$ such that some compact part of C is taken onto a nonmeasurable set. Since C has measure zero, one obtains a measurable mapping.

3.10.74. Prove that there exists a Borel one-to-one function $f\colon [0,1] \to [0,1]$ such that $f(x) = x$ for all x, with the exception of points of a countable set, but the inverse function is discontinuous at all points of $(0,1]$.

HINT: see Sun [**922**, Example 27].

3.10.75. (Aleksandrov [**14**], Ivanov [**451**]) Let K be a compact set in $\mathbb{R}^n$ such that the intersection of K with every straight line is a finite union of intervals (possibly degenerate). Prove the Jordan measurability of K, i.e., the equality $\lambda_n(\partial K) = 0$, where λ_n is Lebesgue measure.

3.10.76.° Let $f \in \mathcal{L}^2(\mathbb{R}^n)$, where we consider the space of complex-valued functions. Let $f_j(x) = f(x)$ if $|x_i| \le j$, $i = 1, \ldots, n$, $f_j(x) = 0$ at all other points.

(i) (**Plancherel's theorem**) Show that the sequence of functions $\widehat{f_j}$ converges in $L^2(\mathbb{R}^n)$ to some function, called the Fourier transform of f in $L^2(\mathbb{R}^n)$ and denoted by $\widehat{f}$.

(ii) Show that the mapping $f \mapsto \widehat{f}$ is a bijection of $L^2(\mathbb{R}^n)$ and

$$\int_{\mathbb{R}^n} f(x)\overline{g(x)}\,dx = \int_{\mathbb{R}^n} \widehat{f}(x)\overline{\widehat{g}(x)}\,dx \quad \text{for all } f, g \in L^2(\mathbb{R}^n).$$

(iii) Show that the Fourier transform defined in (i) is uniquely determined by the property that on $L^2(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$ it coincides with the previously defined Fourier transform and satisfies the equality in (ii).

(iv) Show that there exists a sequence $j_k \to \infty$ such that $\widehat{f_{j_k}}(x) \to \widehat{f}(x)$ a.e.

HINT: use the Parseval equality and completeness of L^2. It is to be noted that in (iv) one actually has a.e. convergence for the whole sequence (see, e.g., Fremlin [**327**, §286U]).

3.10.77.° The Laplace transform of a complex-valued function $f \in L^2[0,+\infty)$ is defined by

$$Lf(s) = \int_0^\infty e^{-st} f(t)\,dt, \quad s > 0.$$

Show that $Lf \in L^2[0,+\infty)$ and that $\|Lf\|_2 \le \sqrt{\pi}\|f\|_2$.

HINT: suppose first that f vanishes in a neighborhood of the origin. By the Cauchy–Bunyakowsky inequality

$$|Lf(s)|^2 \le \int_0^\infty e^{-st}|f(t)|^2 t^{1/2}\,dt \int_0^\infty e^{-st} t^{-1/2}\,dt = \sqrt{\pi} s^{-1/2} \int_0^\infty e^{-st}|f(t)|^2 t^{1/2}\,dt.$$

Integrating this inequality in s over $[0,+\infty)$, interchanging the order of integration and using that the integral of $e^{-st}t^{1/2}s^{-1/2}$ in s is equal to π, we find that $\|Lf\|_2^2 \le \pi\|f\|_2^2$. The general case follows by approximation.

3.10.78.° Give an example of a function $f \in L^1(\mathbb{R}^1)$ such that its Fourier transform is neither in $L^1(\mathbb{R}^1)$ nor in $L^2(\mathbb{R}^1)$, and an example of a function g in $L^2(\mathbb{R}^1)$ such that its Fourier transform does not belong to $L^1(\mathbb{R}^1)$.

3.10.79. Find a uniformly continuous function f on $\mathbb{R}^1$ that satisfies the condition $\lim\limits_{|x|\to\infty} f(x) = 0$, but is not the Fourier transform of a function from $L^1(\mathbb{R}^1)$.

Hint: consider the odd function equal to $1/\ln x$ for $x > 2$; see Stein, Weiss [**908**]. The very existence of functions with the required properties can be established without constructing concrete examples, e.g., by using the Banach inverse mapping theorem that states that the inverse operator for a continuous linear bijection $T\colon X \to Y$ of Banach spaces is continuous: we take $X = L^1(\mathbb{R}^1)$ and the space Y of continuous complex functions tending to zero at infinity equipped with the sup-norm, next we find smooth even functions f_j such that $0 \le f_j \le I_{[-1,1]}$, $f_j(x) \to f(x) = I_{[-1,1]}(x)$. The sequence of functions $\varphi_j = \widehat{f_j}$ is not bounded in L^1 because $\widehat{f} \notin L^1$. However, the sequence of functions $\widehat{\varphi_j} = f_j$ is bounded in Y.

3.10.80.° For f in the complex space $\mathcal{L}^2(\mathbb{R}^1)$ we set

$$\mathcal{H}_\varepsilon f(x) = \frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{y}{y^2+\varepsilon^2} f(x-y)\,dy.$$

Show that there exists the limit $\mathcal{H}_0 f := \lim\limits_{\varepsilon\to 0} \mathcal{H}_\varepsilon f$ in $L^2(\mathbb{R}^1)$ as $\varepsilon \to 0$; then $\mathcal{H}_\varepsilon f$ is called the Hilbert transform of f. In addition, one has $\mathcal{H}_0 = \mathcal{F}^{-1}\mathcal{M}\mathcal{F}$, where $\mathcal{F}$ is the Fourier transform in $L^2(\mathbb{R}^1)$ and $\mathcal{M}g(x) = i(2\pi)^{-1/2}(\mathrm{sign} x)g(x)$.

Hint: let $g_\varepsilon(y) = \pi^{-1}y/(y^2+\varepsilon^2)$, then $\mathcal{F}\mathcal{H}_\varepsilon f = \widehat{g_\varepsilon}\widehat{f}$; use that $\mathcal{F}$ is an isometry of $L^2(\mathbb{R}^1)$ and $\widehat{g_\varepsilon}(x) = i(2\pi)^{-1/2}(\mathrm{sign} x)\exp(-|\varepsilon x|)$.

3.10.81. Suppose that $f \in \mathcal{L}^1(\mathbb{R}^1)$, $\varphi \in \mathcal{L}^\infty(\mathbb{R}^1)$ and that, for some $\beta > 0$ and all x, we have $\varphi(x+\beta) = -\varphi(x)$ (e.g., $\varphi(x) = \sin x$, $\beta = \pi$). Show that

$$\lim_{n\to\infty}\int_{-\infty}^{+\infty} f(x)\varphi(nx)\,dx = 0.$$

Hint: observe that it suffices to prove the claim for functions f that are finite linear combinations of the indicators of intervals, which reduces everything to the case where f is the indicator of the interval $[0,a]$. We have

$$\int_0^a \varphi(nx)\,dx = \frac{1}{n}\int_0^{na}\varphi(y)\,dy.$$

The right-hand side is $O(1/n)$ because the integral of φ over every interval of length 2β vanishes, which is easily seen from the equality of the integrals of $\varphi(x)$ and $-\varphi(x+\beta)$ over $[T, T+\beta]$.

3.10.82. Let us define the standard surface measure σ_{n-1} on the unit sphere S^{n-1} in $\mathbb{R}^n$ by the equality

$$\sigma_{n-1}(B) := n\lambda_n\bigl(x\colon\ 0 < |x| \le 1, x/|x| \in B\bigr), \quad B \in \mathcal{B}(S^{n-1}).$$

Show that σ_{n-1} is a unique Borel measure on S^{n-1} that satisfies the equality

$$r^{n-1}dr\otimes\sigma_{n-1} = \lambda_n \circ \Phi^{-1},$$

where $\Phi\colon \mathbb{R}^n\backslash\{0\} \to (0,\infty)\times S^{n-1}$, $\Phi(x) = \bigl(|x|, x/|x|\bigr)$. In particular, if f is integrable over $\mathbb{R}^n$, then one has

$$\int_{\mathbb{R}^n} f(x)\,dx = \int_0^\infty \int_{S^{n-1}} r^{n-1} f(ry)\,\sigma_{n-1}(dy)\,dr.$$

HINT: verify the equality of the measures $r^{n-1}dr\otimes\sigma_{n-1}$ and $\lambda_n\circ\Phi^{-1}$ on all sets of the form $(a,b]\times E$, where $E\in\mathcal{B}(S^{n-1})$.

3.10.83.° (i) Show that $\sigma_{n-1}(S^{n-1}) = 2\pi^{n/2}/\Gamma(n/2)$.

(ii) Let c_k be the volume of a ball of radius 1 in $\mathbb{R}^k$. Show that

$$c_n = \pi^{n/2}/\Gamma(1+n/2), \quad c_{2k} = \pi^k/k!, \quad c_{2k+1} = 2^{2k+1}k!\pi^k/(2k+1)!.$$

HINT: the answers in (i) and (ii) are easily deduced one from the other. In order to get (ii), apply Fubini's theorem, which gives the relation $c_n = c_{n-1}b_n$, where b_n is the integral of $(1-x^2)^{(n-1)/2}$ over $[-1,1]$ or the doubled integral of $\sin^n\theta$ over $[0,\pi/2]$.

3.10.84. (Schechtman, Schlumprecht, Zinn [**850**]) Let σ be a probability measure on the unit sphere S in $\mathbb{R}^n$ that is proportional to the standard surface measure and let ν be a probability measure on $(0,+\infty)$. Let us consider the measure $\mu=\nu\otimes\sigma$ on $\mathbb{R}^n$ (more precisely, μ is the image of $\nu\otimes\sigma$ under the mapping $(t,y)\mapsto ty$). Let $\mathcal{U}_n$ be the group of all orthogonal matrices $n\times n$ with its natural Borel σ-algebra and a Borel probability measure m with the following property: for each Borel set $B\subset\mathcal{U}_n$ and each $U\in\mathcal{U}_n$, letting L_U and R_U be the left and right multiplications in $\mathcal{U}_n$ by U, we have $m\bigl(L_U(B)\bigr) = m\bigl(R_U(B)\bigr) = m(B)$ (the existence of such a measure – Haar's measure – is proved in Chapter 9). Prove that, for all centrally symmetric convex Borel sets A and B in $\mathbb{R}^n$, one has the inequality

$$\int_{\mathcal{U}_n} \mu\bigl(A\cap U(B)\bigr)\,m(dU) \ge \mu(A)\mu(B).$$

In particular, if B is spherically symmetric, then $\mu(A\cap B)\ge\mu(A)\mu(B)$. These inequalities are true for any probability measure μ with a spherically symmetric density.

HINT: verify that, for every $\psi\in S$, the image of the measure m under the mapping $U\mapsto U\psi$ coincides with σ according to Exercise 9.12.56 in Chapter 9; show that

$$\mu(A) = \int_S \nu(A_\varphi)\,\sigma(d\varphi), \quad \mu(B) = \int_S \nu(B_\psi)\,\sigma(d\psi),$$

$$\int_{\mathcal{U}_n} \mu\bigl(A\cap U(B)\bigr)\,m(dU) = \int_S\int_S \nu(A_\varphi\cap B_\psi)\,\sigma(d\varphi)\,\sigma(d\psi),$$

where $A_\varphi = \{r>0\colon r\varphi\in A\}$; finally, one has $\nu(A_\varphi\cap B_\psi)\ge\nu(A_\varphi)\nu(B_\psi)$, since $A_\varphi\cap B_\psi$ is either A_φ or B_ψ.

3.10.85. (**Sard's theorem**) Let $U\subset\mathbb{R}^n$ be open and let $F\colon U\to\mathbb{R}^n$ be continuously differentiable. Prove that the image of the set of all points where the derivative of F is not invertible has measure zero.

HINT: a more general result can be derived from Theorem 5.8.29.

3.10.86. Let f be a continuously differentiable function on $\mathbb{R}^n$ that vanishes outside a cube Q and let

$$\int_Q f(x)\,dx = 0.$$

Show that there exist continuously differentiable functions $f_1, \dots, f_n$ on $\mathbb{R}^n$ such that $f_i = 0$ outside Q and $f = \sum_{i=1}^n \partial_{x_i} f_i$.

HINT: it suffices to prove the claim for the cube $[0,1]^n$. Use induction on n. If the claim is true for n, then, given a function f of the argument $x = (y,t)$, $y \in \mathbb{R}^n$, $t \in \mathbb{R}^1$, we set

$$g(y) = \int_{-\infty}^{\infty} f(y,t)\, dt.$$

The integral of g vanishes, hence $g = \sum_{i=1}^n \partial_{y_i} g_i$, where the functions g_i on $\mathbb{R}^n$ are continuously differentiable and vanish outside $[0,1]^n$. Let

$$f_{n+1}(y,t) := \int_{-\infty}^{t} [f(y,s) - \zeta(s) g(y)]\, ds, \;\; f_i(y,t) := g_i(y)\zeta(t), \;\; i \le n,$$

where ζ is a smooth function with support in $[0,1]$ and the integral 1. It is verified directly that we obtain the required functions.

3.10.87. Let U be a closed ball in $\mathbb{R}^n$ and let $F\colon\, U \to \mathbb{R}^n$ be a mapping that is infinitely differentiable in a neighborhood of U. Suppose that $y \notin F(\partial U)$, where ∂U is the boundary of U. Let W be a cube containing y in its interior and not meeting $F(\partial U)$, and let ϱ be a nonnegative smooth function vanishing outside W and having the integral 1. Show that the quantity defined by the following formula and called the degree of the mapping F on U at the point y is independent of our choice of a function ϱ with the stated properties:

$$d(F,U;y) := \int_U \varrho\bigl(F(x)\bigr) JF(x)\, dx, \quad JF = \det F'.$$

HINT: use Exercise 3.10.86; if a smooth function g has support in W and its integral vanishes, then the integral of $\partial_{x_i} g\bigl(F(x)\bigr) JF(x)$ over U vanishes by the integration by parts formula. For example, in the case $n = 2$ we have $\partial_{x_1} g\bigl(F(x)\bigr) JF(x) = \partial_{x_1}(g \circ F)(x)\partial_{x_2} F_2(x) - \partial_{x_2}(g \circ F)(x)\partial_{x_1} F_2(x)$, where $F = (F_1, F_2)$; in the general case, see Dunford, Schwartz [**256**, Lemma in §12, Ch. V].

3.10.88. Show that if the point y in the previous exercise is such that $F^{-1}(y) = \{x_1, \dots, x_k\}$, where $JF(x_i) \ne 0$, then $d(F,U;y) = \sum_{i=1}^k \operatorname{sign} JF(x_i)$.

HINT: use the inverse function theorem and the change of variables formula for a sufficiently small neighborhood W.

3.10.89. (i) Show that in Exercise 3.10.87 the number $d(F,U;y)$ is an integer for all $y \notin F(\partial U)$ and that this number is locally constant as a function of y. Deduce that the degree of the mapping at y is unchanged if one replaces F with F_1 with $\|F(x) - F_1(x)\| + |JF(x) - JF_1(x)| \le \varepsilon$, where $\varepsilon > 0$ is sufficiently small. (ii) Let $F\colon\, U \to U$ be continuous. Prove that there exists $x \in U$ with $F(x) = x$.

HINT: (i) use Sard's theorem, the inverse function theorem, and the previous exercise. (ii) If F is infinitely differentiable, but has no fixed points, then for $G(x) = x - F(x)$ we have $d(G,U;0) = 0$ contrary to (i), since for $G_t(x) := x - tF(x)$, $0 \le t \le 1$, we have $0 \notin G_t(\partial U)$, $d(G_0,U;0) = 1$. For continuous F, we find smooth $F_k\colon\, U \to U$ uniformly convergent to F. There exists x_k with $F_k(x_k) = x_k$. A limit point of $\{x_k\}$ is a fixed point of F.

3.10.90. (Faber, Mycielski [**274**]) (i) Let $P \subset \mathbb{R}^n$ be a compact set that is a finite union of compact n-dimensional simplexes and let $f\colon\, P \to \mathbb{R}$ be a smooth function in a neighborhood of P such that f vanishes outside P. Show that

$$\int_P \det\Bigl(\frac{\partial^2 f}{\partial x_i \partial x_j}\Bigr)_{i,j\le n} dx = 0.$$

Construct an example showing that an analogous assertion for a ball P may fail.

(ii) Let $B \subset \mathbb{R}^n$ be a compact set and let $F\colon B \to \mathbb{R}^n$ be a smooth mapping in a neighborhood of B such that $F(\partial B)$ has measure zero and the connected complement. Show that

$$\int_B \det\bigl(F'(x)\bigr)\, dx = 0.$$

3.10.91. Prove Proposition 3.10.16.

3.10.92. Prove that if a function ψ is positive definite, then

$$|\psi(y) - \psi(z)|^2 \le 2\psi(0)[\psi(0) - \operatorname{Re}\psi(y-z)].$$

3.10.93. Prove that if a function ψ on $\mathbb{R}^n$ is positive definite and continuous at the origin, then it is continuous everywhere.

HINT: apply the previous exercise.

3.10.94. Prove that a complex function φ equals the characteristic functional of a nonnegative absolutely continuous measure precisely when there exists a complex function $\psi \in \mathcal{L}^2(\mathbb{R}^n)$ such that

$$\varphi(x) = \int_{\mathbb{R}^n} \psi(x+y)\overline{\psi(y)}\, dy.$$

HINT: if $f \in L^1(\mathbb{R}^n)$ and $f \ge 0$, then $h := \sqrt{f} \in L^2(\mathbb{R}^n)$, whence we have $\check{f} = (2\pi)^{-n/2}\check{h} * \check{h}$, and $\check{h}(-x) = \overline{\check{h}(x)}$; the converse is proven similarly, taking into account that $|\widehat{g}|^2 \in L^1(\mathbb{R}^n)$ and $|\widehat{g}|^2 \ge 0$.

3.10.95. Let μ be a probability measure on the real line with the characteristic functional $\widetilde{\mu}$ and let $F_\mu(t) := \mu\bigl((-\infty, t)\bigr)$.

(i) Prove that, for every t, the limit

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \exp(-its)\widetilde{\mu}(s)\, ds$$

exists and equals the jump of the function F_μ at the point t.

(ii) Let $\{t_j\}$ be all points of discontinuity of F_μ and let d_j be the size of the jump at t_j. Prove the equality

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \bigl|\widetilde{\mu}(s)\bigr|^2\, ds = \sum_{j=1}^{\infty} d_j^2.$$

Deduce that a necessary and sufficient condition for the continuity of F_μ is that the limit on the left be zero.

HINT: see Lukacs [**628**, §§3.2, 3.3].

3.10.96.° Let f be a Lebesgue integrable function on $\mathbb{R}^n$ such that, for every orthogonal linear operator U on $\mathbb{R}^n$, the functions f and $f \circ U$ coincide almost everywhere. Prove that there exists a function g on $[0,\infty)$ such that $f(x) = g(|x|)$ for almost all x.

HINT: let $\varrho_\varepsilon(y) = \varepsilon^{-n}\psi(|y|/\varepsilon)$, where ψ is a smooth function on the real line with bounded support such that $\psi(|y|)$ has the integral 1; verify that the smooth functions $f * \varrho_\varepsilon(x)$ are spherically invariant and hence $f * \varrho_\varepsilon(x) = g_\varepsilon(|x|)$ for some functions g_ε on $[0, +\infty)$. Now one can use the fact (see Theorem 4.2.4 in Chapter 4) that the functions $f * \varrho_{\varepsilon_k}$ converge to f almost everywhere for a suitable sequence $\varepsilon_k \to 0$, which gives convergence of the functions g_{ε_k} almost everywhere on $[0, +\infty)$ to some function g. See also Exercise 9.12.42 in Chapter 9.

3.10.97. Prove that a bounded Borel measure on $\mathbb{R}^n$ is spherically invariant precisely when its characteristic functional is a function of $|x|$.

3.10.98. Let A and B be two sets of positive measure in $\mathbb{R}^n$ and let C be a set in $\mathbb{R}^{2n}$ that coincides with the set $A \times B$ up to a measure zero set. Show that the set $D := \{x + y\colon\ x, y \in \mathbb{R}^n, (x, y) \in C\}$ coincides up to a measure zero set with a set that contains an open ball.

HINT: deduce from the equality $I_C(x, y) = I_A(x)I_B(y)$ a.e. that for a.e. x we have the equality

$$I_A * I_B(x) = \int I_C(x - y, y)\, dy;$$

if such a point x belongs to the nonempty open set $U = \{I_A * I_B > 0\}$, then $x \in D$.

3.10.99. Prove Proposition 3.9.9.

3.10.100. Let $f \in \mathcal{L}^1(\mathbb{R}^1)$. Prove the equalities

$$\biggl|\int_{-\infty}^{+\infty} f(x)\, dx\biggr| = \lim_{T\to+\infty} \int_{-\infty}^{+\infty} \biggl|(2T)^{-1} \int_{-T}^{T} f(x + t)\, dt\biggr|\, dx,$$

$$\int_0^1 \Bigl|\sum_{n=-\infty}^{\infty} f(x + n)\Bigr|\, dx = \lim_{N\to\infty} \int_{-\infty}^{+\infty} \Bigl|(2N + 1)^{-1} \sum_{n=-N}^{N} f(x + n)\Bigr|\, dx.$$

HINT: if f has support in the interval $[-k, k]$, then the first equality is verified directly. Indeed, let $T > k$. The integration in x on the right in the first equality is taken in fact over $[-T - k, T + k]$, and for all $x \in [-T + k, T - k]$ the absolute value of the integral of $f(x + t)$ in t over $[-T, T]$ equals the absolute value of the integral of f, whereas the integral over the interval of length $2k$ multiplied by T^{-1} approaches zero as $T \to +\infty$. The general case reduces to this special one by means of approximations of f by functions with bounded support due to the observation that on the right in the equality to be proven one has the integral of $|f * \psi_T|$, where $\psi_T = (2T)^{-1} I_{[-T,T]}$, and that $\|\psi_T\|_{L^1} = 1$. The second equality is verified in much the same way.

3.10.101. Let $(X, \mathcal{A}, \mu)$ be a probability space and let ν be a bounded nonnegative measure on $\mathcal{A}$. Prove that, for every $\varepsilon > 0$, the family $\mathcal{A}_\varepsilon := \{A \in \mathcal{A}\colon\ \mu(A) \le \varepsilon\}$ contains a set A_ε such that $\nu(A_\varepsilon)$ is maximal in the following sense: if $B \in \mathcal{A}_\varepsilon$ and $\mu(B) \le \mu(A_\varepsilon)$, then $\nu(A_\varepsilon) \ge \nu(B)$.

HINT: Rao [**788**, Proposition 7, p. 266].

3.10.102. Let (X, μ) be a space with a nonnegative measure μ and let f be a μ-measurable function. The nonincreasing rearrangement of the function f is the function f^* on $[0, +\infty)$ with values in $[0, +\infty]$ defined by the equality

$$f^*(t) = \inf\bigl\{s \ge 0\colon\ \mu\bigl(x\colon\ |f(x)| > s\bigr) \le t\bigr\}, \quad \text{where } \inf\varnothing = +\infty.$$

(i) Show that if f assumes finitely many values $0 < c_1 < \cdots < c_n$ on measurable sets $A_0, A_1, \ldots, A_n$ and $0 < \mu(A_i) < \infty$ if $1 \le i \le n$, then

$$f^*(t) = \sum_{j=1}^{n} c_j I_{[\mu(B_{n-j}),\mu(B_{n+1-j}))}(t) = \sum_{j=1}^{n} b_j I_{[0,\mu(B_j))}(t),$$

where $B_j = A_{n+1-j} \cup \cdots \cup A_n$, $B_0 = \varnothing$, $b_j = c_{n+1-j} - c_{n-j}$, $c_0 = 0$.

(ii) Show that $f^*(t) = \sup\bigl\{s \ge 0\colon\ \mu\bigl(x\colon\ |f(x)| > s\bigr) > t\bigr\}$.

(iii) Show that if measurable functions f_n monotonically increase to $|f|$, then the functions f_n^* monotonically increase to f^*.

(iv) Show that the functions f and f^* are equimeasurable, i.e., one has

$$\mu\bigl(x\colon\ |f(x)| > s\bigr) = \lambda\bigl(t\colon\ f^*(t) > s\bigr),$$

where λ is Lebesgue measure.

(v) Prove the following Hardy and Littlewood inequality:

$$\int_X |fg|\, d\mu \le \int_0^\infty f^*(t) g^*(t)\, dt,$$

where f and g are measurable functions.

Hint: see Hardy, Littlewood, Polya [**408**, Ch. X].

3.10.103. Let us consider the measures H_δ^s and H^s from §3.10(iii). Verify that if $s < t$ and $H^s(A) < \infty$, then $H^t(A) = 0$, and if $H_\delta^s(A) = 0$ for some $\delta > 0$, then $H^s(A) = 0$.

3.10.104. (i) Show that, for every $\alpha \in (0,1)$, there exists a set $B_\alpha \subset [0,1]$ with the Hausdorff measure of order α equal to 1.

(ii) Show that for the Cantor set C and $\alpha = \ln 2/\ln 3$ we have $0 < H^\alpha(C) < \infty$.

Hint: see Federer [**282**, 2.10.29], Falconer [**277**, §2.3].

3.10.105. Let H^s be the Hausdorff measure on $\mathbb{R}^n$. Prove that the H^s-measure of every Borel set $B \subset \mathbb{R}^n$ equals the supremum of the H^s-measures of compact subsets of B.

Hint: if $H^s(B) < \infty$, then this is a common property of Borel measures on the space $\mathbb{R}^n$, and if $H^s(B) = \infty$, then, for any $C > 0$, one can find $\delta > 0$ with $H_\delta^s(B) > C$; in B we find a bounded set B' with $H_\delta^s(B') > C$, next in B' we find a compact set K with $H_\delta^s(K) > C$, which yields $H^s(K) > C$.

3.10.106. Let H^s be the Hausdorff measure on $\mathbb{R}^n$ and let $K \subset \mathbb{R}^n$ be a compact set with $H^s(K) = \infty$. Prove that there exists a compact set $C \subset K$ with $0 < H^s(C) < \infty$.

Hint: see Federer [**282**, Theorem 2.10.47].

3.10.107. (Erdős, Taylor [**272**]) Let A_n be Lebesgue measurable sets in $[0,1]$ with $\lambda(A_n) \ge \varepsilon > 0$ for all $n \in \mathbb{N}$. Show that, for every continuous monotonically increasing function φ with $\varphi(0) = 0$ and $\lim\limits_{t\to 0+} \varphi(t)/t = +\infty$, there exists a subsequence n_k such that the set $\bigcap_{k=1}^\infty A_{n_k}$ has infinite measure with respect to the Hausdorff measure generated by the function φ.

3.10.108. (Darst [**204**]) Prove that there exist an infinitely differentiable function f on the real line and a set Z of Lebesgue measure zero such that the set $f^{-1}(Z)$ is not Lebesgue measurable.

3.10.109. (Kaufman, Rickert [**497**]) (i) Let μ be a complex measure with $\|\mu\| = 1$ (see the definition before Proposition 3.10.16). Prove that there exists a measurable set E such that $|\mu(E)| \geq 1/\pi$.

(ii) Prove that in (i) one can pick a set E with $|\mu(E)| > 1/\pi$ precisely when the Radon–Nikodym density f of the measure μ with respect to $|\mu|$ satisfies the equality

$$\int f(t)^k \, |\mu|(dt) = 0$$

for all $k \in \{-1, 1, -2n, 2n\}$, $n \in \mathbb{N}$.

(iii) Let μ be a measure with values in $\mathbb{R}^n$ such that $\|\mu\| = 1$. Prove that there exists a measurable set E such that

$$|\mu(E)| \geq \Gamma(n/2)\Big(2\sqrt{\pi}\Gamma\big((n+1)/2\big)\Big)^{-1}.$$

3.10.110. (i) Suppose that the values of two Borel probability measures μ and ν on $\mathbb{R}^n$ coincide on every half-space of the form $\{x\colon\ (x, y) \leq c\}$, $y \in \mathbb{R}^n$, $c \in \mathbb{R}^1$. Prove that $\mu = \nu$. Prove the same for open half-spaces.

(ii) (Pták, Tkadlec [**771**]) Suppose that the values of two Borel probability measures μ and ν on $\mathbb{R}^n$ coincide on every open ball with the origin at the boundary. Prove that $\mu = \nu$.

(iii) Prove the analog of (ii) for closed balls.

Hint: in the case $n = 1$ the assertion is trivial, since the values of μ and ν coincide on all intervals $(a, b]$. Hence in the case $n > 1$ the measures μ and ν have equal images under the mappings $\pi_y\colon\ x \mapsto (x, y)$, whence by the change of variables formula we have

$$\widetilde{\mu}(y) = \int_{\mathbb{R}^1} \exp(it)\, \mu \circ \pi_y^{-1}(dt) = \int_{\mathbb{R}^1} \exp(it)\, \nu \circ \pi_y^{-1}(dt) = \widetilde{\nu}(y).$$

(ii) Let $f(x) = x/|x|^2$, $|x| > 0$, $f(0) = 0$; then $\mu \circ f^{-1}\big(f(U)\big) = \nu \circ f^{-1}\big(f(U)\big)$ for every open ball U with the origin at the boundary, i.e., the values of the measures $\mu \circ f^{-1}$ and $\nu \circ f^{-1}$ coincide on every open half-space whose closure does not contain the origin. Hence $\mu \circ f^{-1} = \nu \circ f^{-1}$, whence one has $\mu = \nu$. (iii) Observe that $\mu(0) = \nu(0)$ and use the same reasoning.

3.10.111.° Let a function Φ be strictly increasing and continuous on $[0, 1]$. Prove that for every bounded Borel function f one has

$$\int_0^1 f(x)\, d\Phi(x) = \int_{\Phi(0)}^{\Phi(1)} f\big(\Phi^{-1}(y)\big)\, dy$$

with the Lebesgue–Stieltjes integral on the left and the Lebesgue integral on the right.

3.10.112. Let μ be a Borel (possibly signed) measure on $[0, 1]$ with the following property: if continuous functions f_n are uniformly bounded and converge to zero almost everywhere with respect to Lebesgue measure λ, then

$$\int f_n \, d\mu \to 0.$$

Prove that $\mu \ll \lambda$.

Hint: let K be a compact set with $\lambda(K) = 0$. Let us take a uniformly bounded sequence of continuous functions f_n convergent to I_K almost everywhere with respect to the measure $|\mu| + \lambda$. Then $f_n \to 0$ λ-a.e. and $f_n \to I_K$ μ-a.e., which yields

$$\mu(K) = \lim_{n\to\infty} \int f_n \, d\mu = 0.$$

3.10.113. (i) Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be complete probability spaces, let $A \subset X$ be a set that is not measurable with respect to μ, and let $B \subset Y$ be a set such that $A \times B$ is measurable with respect to $\mu \otimes \nu$. Prove that $\nu(B) = 0$.

(ii) Let $(X_n, \mathcal{A}_n, \mu_n)$, where $n \in \mathbb{N}$, be complete probability spaces and let sets $A_n \subset X_n$ be such that $\prod_{n=1}^{\infty} A_n$ is measurable with respect to $\bigotimes_{n=1}^{\infty} \mu_n$. Prove that either every A_n is measurable with respect to μ_n or $\mu\bigl(\prod_{n=1}^{\infty} A_n\bigr) = 0$ and then $\lim\limits_{n\to\infty} \prod_{i=1}^{n} \mu_i^*(A_i) = 0$.

Hint: (i) by Fubini's theorem the set C of all points y such that $(A \times B)_y$ is not measurable with respect to μ, has ν-measure zero. In addition, $B \subset C$, since one has $(A \times B)_y = A$ for all $y \in B$. (ii) If among the sets A_n there are nonmeasurable ones and their product has a nonzero measure, then by (i) the product of all non-measurable sets A_n is measurable. Hence we may assume that all the sets A_n are nonmeasurable. Their product has measure zero, since by (i) the product of all A_n with $n > 1$ has measure zero. Then we obtain $\lim\limits_{n\to\infty} \prod_{i=1}^{n} \mu_i^*(A_i) = 0$. Indeed, by Theorem 1.12.14, there exist probability measures ν_n on the σ-algebras $\mathcal{A}_n'$ obtained by adding the sets A_n to $\mathcal{A}_n$ such that $\nu_n(A_n) = \mu_n^*(A_n)$ and $\nu_n|_{\mathcal{A}_n} = \mu_n$. Let us consider the measure $\nu := \bigotimes_{n=1}^{\infty} \nu_n$ on $\bigotimes_{n=1}^{\infty} \mathcal{A}_n'$. There exists a set $E \in \bigotimes_{n=1}^{\infty} \mathcal{A}_n$ such that $\mu(E) = 0$ and $\prod_{n=1}^{\infty} A_n \subset E$. Then $\nu(E) = \mu(E) = 0$, since ν coincides with μ on $\bigotimes_{n=1}^{\infty} \mathcal{A}_n$. Hence $\prod_{n=1}^{\infty} \nu_n(A_n) = \nu\bigl(\prod_{n=1}^{\infty} A_n\bigr) = 0$.

3.10.114. Let $(X_\alpha, \mathcal{A}_\alpha, \mu_\alpha)$, where $\alpha \in \Lambda$ and $\Lambda \neq \varnothing$, be measurable spaces with complete probability measures and let $E_\alpha \subset X_\alpha$ be such that $E = \prod_{\alpha\in\Lambda} E_\alpha$ is measurable with respect to $\bigotimes_\alpha \mu_\alpha$, but does not belong to $\bigotimes_\alpha \mathcal{A}_\alpha$. Prove that $\prod_{\alpha\in\Lambda} \mu_\alpha^*(E_\alpha) = 0$, i.e., there exists an at most countable family of indices α_n such that the product of numbers $\mu_{\alpha_n}^*(A_{\alpha_n})$ diverges to zero.

Hint: Let $\Lambda_1 = \{\alpha\colon\ \mu_\alpha^*(E_\alpha) = 1\}$, $\Lambda_2 = \Lambda \backslash \Lambda_1$. If Λ_2 is uncountable, then, for some $q < 1$, there exist infinitely many indices α with $\mu_\alpha^*(E_\alpha) < q$, which proves the assertion. Let Λ_2 be finite or countable. Let $\Pi_1 = \prod_{\alpha\in\Lambda_1} E_\alpha$, $\Pi_2 = \prod_{\alpha\in\Lambda_2} E_\alpha$. We may assume that $E_\alpha \neq X_\alpha$ for all α. The same reasoning as in assertion (ii) in the previous exercise shows that Π_1 cannot have measure zero with respect to $\pi_1 := \bigotimes_{\alpha\in\Lambda_1} \mu_\alpha$. Hence by assertion (i) in the previous exercise the set Π_2 is measurable. If its measure equals zero with respect to $\pi_2 := \bigotimes_{\alpha\in\Lambda_2} \mu_\alpha$, then, by the previous exercise, the product of $\mu_\alpha(E_\alpha)$ with $\alpha \in \Lambda_2$ diverges to zero. If one has $\pi_2(\Pi_2) > 0$, then all sets E_α, $\alpha \in \Lambda_2$, are measurable, and the set Π_1 is π_1-measurable. As it has already been noted, $\pi_1(\Pi_1) > 0$, whence it follows that Λ_1 is at most countable. Indeed, otherwise Π_1 would not contain nonempty sets from $\bigotimes_{\alpha\in\Lambda_1} \mathcal{A}_\alpha$, since such sets depend only on countably many indices and $E_\alpha \neq X_\alpha$. Then, by the previous exercise, whenever $\alpha \in \Lambda_1$, the set E_α is μ_α-measurable, which leads to a contradiction by the completeness of the measures μ_α.

3.10.115.° Let μ be a Borel probability measure with a density ϱ on $\mathbb{R}^2$. (i) Show that the distribution of $f(x, y) = x + y$ on $(\mathbb{R}^2, \mu)$ has the density

$$\varrho_1(t) = \int_{-\infty}^{+\infty} \varrho(t - s, s) \, ds.$$

(ii) Show that the distribution of $g(x,y)=x/y$ on $(\mathbb{R}^2,\mu)$ has the density

$$\varrho_2(t)=\int_{-\infty}^{+\infty}|s|\varrho(ts,s)\,ds.$$

HINT: for every bounded Borel function φ, by using the change of variables $x+y=t$, $y=s$ one has

$$\int_{-\infty}^{+\infty}\varphi(t)\varrho_1(t)\,dt=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\varrho(x+y)\varrho(x,y)\,dx\,dy=\int\varphi(t)\,\mu\circ f^{-1}(dt).$$

For ϱ_2 the proof is similar.

3.10.116. Let $\varphi(x)=\exp\bigl(il(x)\bigr)$, where l is a nonmeasurable additive function on the real line (such a function is easily constructed by using a Hamel basis). Show that φ is positive definite and $\varphi(0)=1$.

HINT: Let $c_j\in\mathbb{C}$, $x_j\in\mathbb{R}^1$ and $a_j:=c_j\exp\bigl(il(x_j)\bigr)$. Then we obtain the equality $c_j\overline{c_k}\varphi(x_j-x_k)=a_j\overline{a_k}$, since $\varphi(x_j-x_k)=\exp\bigl(il(x_j)\bigr)\exp\bigl(-il(x_k)\bigr)$.

3.10.117. (i) Let μ be a probability measure on $\mathbb{R}^n$. Prove that

$$0\le 1-\operatorname{Re}\widetilde{\mu}(2y)\le 4\bigl(1-\operatorname{Re}\widetilde{\mu}(y)\bigr),\quad y\in\mathbb{R}^n.$$

(ii) Show that if $\widetilde{\mu}(y)=1$ in some neighborhood of the origin, then μ is Dirac's measure at the origin.

HINT: (i) observe that $1-\cos 2t=2(1-\cos^2 t)\le 4(1-\cos t)$; derive from (i) that $\widetilde{\mu}(y)=1$ for all y.

3.10.118. (Gneiting [**364**]) Let $E\subset\mathbb{R}$ be a closed set symmetric about the origin and let $0\in E$. Show that there exist probability measures μ and ν on $\mathbb{R}$ such that $\widetilde{\mu}(t)=\widetilde{\nu}(t)$ for all $t\in E$ and $\widetilde{\mu}(t)\ne\widetilde{\nu}(t)$ for all $t\notin E$.

3.10.119. Let μ and ν be two Borel probability measures on the real line. Prove that

$$\iint (x+y)^2\,\mu(dx)\,\nu(dy)<\infty\quad\text{precisely when}\quad\int x^2\,\mu(dx)+\int y^2\,\nu(dy)<\infty.$$

HINT: if the double integral is finite, then there exists y such that

$$\int (x+y)^2\,\mu(dx)<\infty,$$

whence the μ-integrability of x^2 follows.

3.10.120. (Gromov [**381**]) Suppose that in $\mathbb{R}^n$ we are given $k\le n+1$ balls $B(x_i,r_i)$ with the centers x_i and radii r_i and k balls $B(y_i,r_i)$ with the centers y_i and radii r_i such that $|x_i-x_j|\ge|y_i-y_j|$ for all i,j. Then the following inequality holds: $\lambda_n\bigl(\bigcap_{i=1}^k B(x_i,r_i)\bigr)\le\lambda_n\bigl(\bigcap_{i=1}^k B(y_i,r_i)\bigr)$, where λ_n is Lebesgue measure.

As far as I know, the following question raised in the 1950s by several authors (M. Kneser, E.T. Poulsen, and H. Hadwiger; see Meyer, Reisner, Schmuckenschläger [**685**]) remains open: suppose that in $\mathbb{R}^n$ we are given k balls $B(x_i,r)$ of radius r centered at the points $x_1,\dots,x_k$ and k balls $B(y_i,r)$ of radius r centered at the points $y_1,\dots,y_k$ such that $|x_i-x_j|\le|y_i-y_j|$ for all i,j; is it true that $\lambda_n\bigl(\bigcup_{i=1}^k B(x_i,r)\bigr)\le\lambda_n\bigl(\bigcup_{i=1}^k B(y_i,r)\bigr)$?

3.10.121. (i) Let $(X_i, \mathcal{A}_i, \mu_i)$, $i = 1, \ldots, n$, be measurable spaces with nonnegative σ-finite measures and let f_i be nonnegative $\bigotimes_{i=1}^n \mu_i$-measurable functions on $\prod_{i=1}^n X_i$ such that f_i is independent of the ith variable. Prove the inequality

$$\left(\int f_1 \cdots f_n \, d\mu_1 \cdots d\mu_n\right)^{n-1} \leq \prod_{i=1}^n \int f_i^{n-1} \prod_{j \neq i} d\mu_j.$$

(ii) Let E be a Borel set in $\mathbb{R}^3$ and let E_i be its orthogonal projection to the coordinate plane $x_i = 0$. Prove the inequality $\lambda_3(E)^2 \leq \lambda_2(E_1)\lambda_2(E_2)\lambda_2(E_3)$.

Hint: (i) use induction on n; let

$$g_i = \int f_i^{n-1} \, d\mu_1, \; I_i = \int f_i^{n-1} \prod_{j \neq i} d\mu_j,$$

and let I be the integral of $f_1 \cdots f_n$ with respect to $\mu_1 \cdots \mu_n$. By applying the generalized Hölder inequality and the usual Hölder inequality with exponents $p = n - 1$ and $q = (n-1)/(n-2)$, we have

$$\begin{aligned} I &\leq \int f_1 g_2^{1/(n-1)} \cdots g_n^{1/(n-1)} \, d\mu_2 \cdots d\mu_n \\ &\leq I_1^{1/(n-1)} \left(\int g_2^{1/(n-2)} \cdots g_n^{1/(n-2)} \, d\mu_2 \cdots d\mu_n\right)^{(n-2)/(n-1)}. \end{aligned}$$

It remains to use the inductive hypothesis and the fact that

$$I_i = \int g_i \prod_{j \geq 2, j \neq i} d\mu_j.$$

(ii) Observe that $I_E(x_1, x_2, x_3) \leq I_{E_3}(x_1, x_2) I_{E_1}(x_2, x_3) I_{E_2}(x_1, x_3)$.

3.10.122. (i) (T. Carleman) Suppose we are given a sequence of numbers σ_n with $\sum_{n=1}^\infty \sigma_{2n}^{-1/(2n)} = \infty$. Prove that two probability measures μ and ν on the real line coincide if they have equal moments

$$\int_{-\infty}^{+\infty} t^n \, \mu(dt) = \int_{-\infty}^{+\infty} t^n \, \nu(dt) = \sigma_n, \quad \forall\, n \in \mathbb{N}.$$

(ii) Prove that for all n one has

$$\int_0^\infty x^n \exp\bigl(-x^{1/4}\bigr) \sin\bigl(x^{1/4}\bigr) \, dx = 0.$$

Deduce the existence of two different probability measures on the real line with equal moments for all n.

(iii) (M.G. Krein) Show that a probability density ϱ on the real line is not uniquely determined by its moments in the class of all probability measures precisely when the function $(1 + x^2)^{-1} \min\bigl(\ln \varrho(x), 0\bigr)$ has a finite integral over $\mathbb{R}^1$.

Hint: see Ahiezer [**5**].

3.10.123. Let f and g be nonnegative Lebesgue measurable functions on $\mathbb{R}^n$ and let the mapping $f * g$ with values in $[0, +\infty]$ be defined as follows: $f * g(x)$ is the integral of the function $y \mapsto f(x - y)g(y)$ if it is integrable and $f * g(x) = +\infty$ otherwise. Show that $f * g$ is Borel measurable.

Hint: observe that $f * g(x) = \lim\limits_{n \to \infty} \min(f, n) * \bigl(\min(g, n) I_{[-n,n]}\bigr)$.

3.10.124. Let B be an open ball in $\mathbb{R}^n$ and let $f\colon B \to \mathbb{R}$ be a measurable function such that

$$\int_B \int_B \frac{|f(x)-f(y)|}{|x-y|^{n+1}}\,dx\,dy < \infty.$$

Prove that $f = c$ a.e., where c is a constant.

HINT: it is clear that f is integrable on B; the assertion reduces to the case of a smooth function, since letting $f_\varepsilon := f * g_\varepsilon$, $g_\varepsilon(x) = \varepsilon^{-n} g(x/\varepsilon)$, we obtain that f_ε satisfies the above condition in a smaller ball. The function

$$|f(x) - f(y) - f'(y)(x-y)|/|x-y|^{n+1}$$

in the case of smooth f is integrable on $B{\times}B$ by Taylor's formula. Hence the function $|f'(y)(x-y)|/|x-y|^{n+1}$ is integrable as well. If f is not constant, then there exists a point y such that $f'(y) \neq 0$ and the function $x \mapsto |f'(y)(x-y)|/|x-y|^{n+1}$ is integrable on B, which is false (we may assume that $y = 0$ and consider the polar coordinates). A proof based on the theory of Sobolev spaces is given in Brezis [**126**].

3.10.125. (Kolmogorov [**531**]) Let E be a Lebesgue measurable set on the real line. Let $L(E)$ be the supremum of lengths of the intervals onto which E can be mapped by means of a nonexpanding (i.e., Lipschitzian with the constant 1) mapping. Show that $L(E)$ coincides with Lebesgue measure of E.

HINT: let $f(x) = \lambda\bigl(E \cap (-\infty, x)\bigr)$. Then f is nonexpanding and $f(E) = [0, \lambda(E)]$, whence one has $L(E) \geq \lambda(E)$. The reverse inequality follows by considering the covers of E by sequences of disjoint intervals.

CHAPTER 4

The spaces L^p and spaces of measures

When communicating our knowledge to other people, we do one of the three things: either, being well aware of the subject, we extract from it for other persons only that what we take for the most essential; or we rush to present everything what we know; or, finally, we communicate not only what we know, but also what we do not know.

N.I. Pirogov. Letters from Heidelberg.

4.1. The spaces L^p

In this section, we study certain normed spaces of integrable functions. We recall that a linear space L over the field of real or complex numbers equipped with a function $x \mapsto \|x\|_L \geq 0$ is called a normed space with the norm $\|\cdot\|_L$ if:

(i) $\|x\|_L = 0$ precisely when $x = 0$;

(ii) $\|\lambda x\|_L = |\lambda|\, \|x\|_L$ for all $x \in L$ and all scalars λ;

(iii) $\|x + y\|_L \leq \|x\|_L + \|y\|_L$ for all x, $y \in L$.

If only conditions (ii) and (iii) are fulfilled, then $\|\cdot\|_L$ is called a seminorm. For example, the identically zero function is a seminorm (but not a norm if the space L differs from zero). It is easily verified that the normed space L equipped with the function $d(x, y) := \|x - y\|_L$ is a metric space. If this metric space is complete (i.e., every fundamental sequence has a limit), then the normed space L is called complete. Complete normed spaces are called *Banach spaces* in honor of the outstanding Polish mathematician Stephan Banach.

Let $(X, \mathcal{A}, \mu)$ be a measure space with a nonnegative measure μ (possibly with values in $[0, +\infty]$) and let $p \in [1, +\infty)$. As in §2.11 above, we denote by $\mathcal{L}^p(\mu)$ the class of all μ-measurable functions f such that $|f|^p$ is a μ-integrable function. In order to turn these classes into normed spaces with the integral norms, one has to identify μ-equivalent functions (without such an identification the norms defined below do not satisfy condition (i) above, and the classes $\mathcal{L}^p(\mu)$ are not linear spaces, as explained in §2.11). The sets $\mathcal{L}^p(\mu)$ are equipped with their natural equivalence relation: $f \sim g$ if $f = g$ μ-a.e., as already mentioned in §2.11.

Denote by $L^p(\mu)$ the factor-space of $\mathcal{L}^p(\mu)$ with respect to this equivalence relation. Thus, $L^p(\mu)$ is the space of equivalence classes of μ-measurable functions f such that $|f|^p$ is integrable. In the case of Lebesgue measure on $\mathbb{R}^n$ we use the notation $L^p(\mathbb{R}^n)$, and in the case of a subset $E \subset \mathbb{R}^n$ the notation $L^p(E)$. In place of $L^p([a,b])$ and $L^p\big([a,+\infty)\big)$ we write $L^p[a,b]$ and $L^p[a,+\infty)$.

It is customary to speak of $L^p(\mu)$ as the space of all functions integrable of order p, which is formally incorrect, but convenient. Certainly, it is meant that functions equal almost everywhere are regarded as the same element. The Minkowski inequality yields that the function $\|\cdot\|_p$ (see §2.11) defines a norm on $L^p(\mu)$.

The same notation is employed for complex-valued functions, but we shall always give a special note when considering complex spaces.

In a special way one defines the spaces $\mathcal{L}^\infty(\mu)$ and $L^\infty(\mu)$. The set $\mathcal{L}^\infty(\mu)$ consists of bounded everywhere defined μ-measurable functions. Let $L^\infty(\mu)$ denote the factor-space of $\mathcal{L}^\infty(\mu)$ with respect to the equivalence relation introduced above. However, one cannot take for a norm on $L^\infty(\mu)$ the function $\sup_{x\in X}|f(x)|$ with an arbitrary representative f of the equivalence class, since unlike the integral norm, the sup-norm depends on the choice of such a representative. For this reason the norm $\|\cdot\|_\infty$ on $L^\infty(\mu)$ is introduced as follows:

$$\|f\|_\infty := \|f\|_{L^\infty(\mu)} := \inf_{\widehat{f}\sim f}\sup_{x\in X}|\widehat{f}(x)|,$$

where inf is taken over all representatives of the equivalence class of f. On the space $\mathcal{L}^\infty(\mu)$ we thus obtain the seminorm $\|\cdot\|_\infty$. It is to be noted that the same seminorm can be written as

$$\|f\|_\infty := \operatorname{esssup}_{x\in X}|f(x)| := \inf_{\Omega\colon\, \mu(X\backslash\Omega)=0}\ \sup_{x\in\Omega}|f(x)|, \quad f\in\mathcal{L}^\infty(\mu).$$

The quantity $\operatorname{esssup}_{x\in X}|f(x)|$ is also called the essential supremum of the function $|f|$. Thus, $\|f\|_\infty = \operatorname{esssup}_{x\in X}|\widehat{f}(x)|$, where $\widehat{f}$ is an arbitrary representative of the equivalence class of f.

4.1.1. Lemma. *For all $\lambda\in\mathbb{R}^1$, $f,g\in\mathcal{L}^p(\mu)$, we have*

$$\|\lambda f\|_p = |\lambda|\,\|f\|_p, \quad \|f+g\|_p \le \|f\|_p + \|g\|_p.$$

PROOF. If $f\in\mathcal{L}^p(\mu)$ and $\lambda\in\mathbb{R}^1$, then $\lambda f\in\mathcal{L}^p(\mu)$ and $\|\lambda f\|_p = |\lambda|\,\|f\|_p$. Let $g\in\mathcal{L}^p(\mu)$. For $p=\infty$ the inequality $\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty$ is obvious. For $p\in[1,+\infty)$ we apply the Minkowski inequality from §2.11. □

If the space X contains a nonempty set of measure zero, then the function $\|\cdot\|_p$ is not a norm on the linear space of finite everywhere defined functions from $\mathcal{L}^p(\mu)$, since it vanishes at the indicator of that set.

For every $f\in L^p(\mu)$, let $\|f\|_p = \|\widetilde{f}\|_p$, where $\widetilde{f}$ is an arbitrary representative of the equivalence class of f. Clearly, $\|f\|_p$ does not depend on our choice of such a representative.

The space $L^p(\mu)$ has a natural structure of a linear space: the sum of two equivalence classes with representatives f and g is the equivalence class of $f+g$. It is clear that this definition does not depend on our choice of representatives in the classes containing f and g. The multiplication by scalars is defined analogously. One may ask whether instead of passing to the factor-space we could simply choose a representative in every equivalence class in such a way that pointwise sums and multiplication by constants would correspond to the above-defined operations on equivalence classes. This turns out to be possible only for $p = \infty$ (see Theorem 10.5.4 on liftings and Exercise 10.10.53 in Chapter 10).

4.1.2. Corollary. *The function $\|\cdot\|_p$ is a norm on the space $L^p(\mu)$.*

4.1.3. Theorem. *The spaces $L^p(\mu)$ are complete, i.e., are Banach spaces.*

Proof. Suppose first that the measure μ is finite. Let a sequence $\{f_n\}$ be fundamental in the norm $\|\cdot\|_p$. We shall also denote by f_n arbitrary representatives of equivalence classes and deal further with individual functions. In the case $p = \infty$ we set $\varepsilon_{n,k} = \|f_n - f_k\|_\infty$ and obtain the set

$$\Omega = \bigcap_{n,k} \bigl\{x\colon\ |f_n(x) - f_k(x)| \le \varepsilon_{n,k}\bigr\}$$

of full measure. The sequence $\{f_n\}$ is uniformly fundamental on Ω and hence is uniformly convergent. Let $p < \infty$. By Chebyshev's inequality, one has

$$\mu\Bigl(x\colon\ |f_n(x) - f_k(x)| \ge c\Bigr) \le c^{-p}\|f_n - f_k\|_p^p$$

which yields that the sequence $\{f_n\}$ is fundamental in measure, hence converges in measure to some function f. We observe that the fundamentality in the norm $\|\cdot\|_p$ implies the boundedness in this norm. Hence by Fatou's theorem with convergence in measure (see Theorem 2.8.5), one has the inclusion $f \in \mathcal{L}^p(\mu)$. Let us show that $\|f - f_n\|_p \to 0$. Let $\varepsilon > 0$. We pick a number N such that $\|f_n - f_k\|_p < \varepsilon$ for $n,\, k \ge N$. For every fixed $k \ge N$, the sequence $|f_n - f_k|$ converges in measure to $|f - f_k|$ as $n \to \infty$. This follows by the estimate $\Bigl||f_n - f_k| - |f - f_k|\Bigr| \le |f_n - f|$. Applying Fatou's theorem once again, we obtain $\|f - f_k\|_p \le \varepsilon$. The case of an infinite measure reduces at once to the case of a σ-finite measure, which in turn reduces easily to the case of a finite measure, as explained in §2.6. □

We note that the spaces $L^p(\mu)$ can also be considered for $0 < p < 1$, but they have no natural norms, although can be equipped with metrics (see Exercise 4.7.62).

Finally, if μ is a signed measure, then for all $p \ge 0$ we set by definition $L^p(\mu) := L^p(|\mu|)$ and $\mathcal{L}^p(\mu) := \mathcal{L}^p(|\mu|)$.

4.2. Approximations in L^p

It is useful to be able to approximate functions from L^p by functions from more narrow classes. First we prove an elementary general result that is frequently used as a first step in constructing finer approximations.

We recall that a metric space is called separable if it contains a countable everywhere dense subset.

4.2.1. Lemma. *The set of all simple functions is everywhere dense in every space $L^p(\mu)$, $1 \le p < \infty$.*

PROOF. Let $f \in L^p(\mu)$ and $p < \infty$. By the dominated convergence theorem, the functions $f_n = fI_{\{-n \le f \le n\}}$ converge to f in $L^p(\mu)$. Hence it suffices to approximate bounded functions in $L^p(\mu)$. In the case of a finite measure it suffices to approximate bounded functions by simple ones. In the general case, we need an intermediate step: we approximate any bounded function $f \in L^p(\mu)$ by functions of the form $fI_{\{n^{-1} \le |f|\}}$ with some $n \in \mathbb{N}$, which is also possible by the dominated convergence theorem. Now everything reduces to the case of a finite measure because the measure of the set where our new function is not zero is finite. $\square$

The set of measurable functions with finitely many values (such functions are simple in the case of a finite measure) is everywhere dense in $L^\infty(\mu)$, which is proved by the method explained in §2.1.

In §4.7(vi), we present additional results on approximations in L^p for general measures. In many cases simple functions can be approximated by functions from various other classes (not necessarily simple). For example, in the case where μ is a Borel measure on $\mathbb{R}^n$ that is bounded on bounded sets, every measurable set of finite μ-measure can be approximated (in the sense of measure of the symmetric difference) by sets from the algebra generated by cubes with edges parallel to the coordinate axes. This means that linear combinations of the indicators of sets in this algebra are dense in $L^p(\mu)$ with $p < \infty$ (e.g., in the case $n = 1$, the set of step functions is dense in $L^p(\mu)$). In turn, every such function is easily approximated in $L^p(\mu)$ by continuous functions with bounded support (it suffices to approximate the indicator of every open cube K, which is easily done by taking continuous functions equal to 0 outside K, equal to 1 in a close smaller cube and having a range in $[0, 1]$). Finally, continuous functions with bounded support are uniformly approximated by smooth functions. This yields the following conclusion.

4.2.2. Corollary. *Let a nonnegative Borel measure μ on $\mathbb{R}^n$ be bounded on bounded sets. Then, the class $C_0^\infty(\mathbb{R}^n)$ of smooth functions with bounded support is everywhere dense in $L^p(\mu)$, $1 \le p < \infty$. In particular, the spaces $L^p(\mu)$, $1 \le p < \infty$, are separable.*

In the case of Lebesgue measure (and some other measures) a very efficient method of approximation of functions is based on the use of convolution. Let ϱ be a function integrable over $\mathbb{R}^n$ such that

$$\int_{\mathbb{R}^n} \varrho(x)\, dx = 1.$$

Set $\varrho_\varepsilon(x) = \varepsilon^{-n}\varrho(x/\varepsilon)$, $\varepsilon > 0$.

4.2.3. Lemma. *Let $f \in \mathcal{L}^p(\mathbb{R}^n)$, $1 \le p < \infty$. Then, the mapping*

$$T_f\colon \mathbb{R}^n \to L^p(\mathbb{R}^n), \quad T_f(v)(x) = f(x+v),$$

is continuous and bounded.

Proof. For any $v \in \mathbb{R}^n$, we have

$$\|T_f(v)\|_p^p = \int_{\mathbb{R}^n} |f(x+v)|^p \, dx = \|f\|_p^p.$$

If the function f is continuous and vanishes outside some ball, then we have as $v_j \to v$

$$\|T_f(v_j) - T_f(v)\|_p^p = \int_{\mathbb{R}^n} |f(x+v_j) - f(x+v)|^p \, dx \to 0,$$

since the functions $x \mapsto f(x+v_j)$ vanish outside some ball and uniformly converge to the function $x \mapsto f(x+v)$. In the general case, there exists a sequence of continuous functions f_k with bounded support convergent to f in $L^p(\mathbb{R}^n)$. As shown above, the mappings T_{f_k} are continuous. They converge to T_f uniformly on $\mathbb{R}^n$, since

$$\begin{aligned}\|T_f(v) - T_{f_k}(v)\|_p^p &= \int_{\mathbb{R}^n} |f(x+v) - f_k(x+v)|^p \, dx \\ &= \int_{\mathbb{R}^n} |f(x) - f_k(x)|^p \, dx = \|f - f_k\|_p^p.\end{aligned}$$

Hence the mapping T_f is continuous as well. $\square$

4.2.4. Theorem. *Let $f \in \mathcal{L}^p(\mathbb{R}^n)$, $1 \le p < \infty$. Then one has*

$$\lim_{\varepsilon \to 0} \|f * \varrho_\varepsilon - f\|_p = 0.$$

In particular, on every ball, the functions $f * \varrho_\varepsilon$ converge to f in measure.

Proof. Let

$$G(y) = \int_{\mathbb{R}^n} |f(x) - f(x-y)|^p \, dx.$$

By Lemma 4.2.3, the function G is bounded and $G(\varepsilon y) \to 0$ for all y as $\varepsilon \to 0$. We have by Hölder's inequality

$$\begin{aligned}\|f * \varrho_\varepsilon - f\|_p^p &= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} [f(x) - f(x - \varepsilon y)] \varrho(y) \, dy \right|^p dx \\ &\le \|\varrho\|_{L^1(\mathbb{R}^n)}^{p-1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(x) - f(x-\varepsilon y)|^p |\varrho(y)| \, dy \, dx \\ &= \|\varrho\|_{L^1(\mathbb{R}^n)}^{p-1} \int_{\mathbb{R}^n} G(\varepsilon y) |\varrho(y)| \, dy.\end{aligned}$$

By the dominated convergence theorem, the right-hand side of this estimate tends to zero as $\varepsilon \to 0$. $\square$

4.2.5. Corollary. *If f is a bounded measurable function, then, on every ball, the functions $f * \varrho_\varepsilon$ converge to f in the mean and in measure.*

PROOF. If f vanishes outside some ball, then the theorem applies. We may assume that $|f| \leq 1$. Denote by B_j the ball of radius j centered at the origin. Suppose we are given a ball $B = B_k$ and $\delta > 0$. Set $f_j = f I_{B_j}$. We find m such that the integral of ϱ over $\mathbb{R}^n \backslash B_m$ is less than $\delta/4$. For $j \geq m+k$ and all $\varepsilon \in [0,1]$, we have $f_j(x+\varepsilon y) = f(x+\varepsilon y)$ if $x \in B$, $y \in B_m$. Hence

$$\|f - f * \varrho_\varepsilon\|_{L^1(B)} = \|f_j - f * \varrho_\varepsilon\|_{L^1(B)}$$
$$\leq \|f_j - f_j * \varrho_\varepsilon\|_{L^1(B)} + \|(f_j - f) * \varrho_\varepsilon\|_{L^1(B)} \leq \|f_j - f_j * \varrho_\varepsilon\|_{L^1(B)} + \frac{\delta}{2}.$$

It remains to apply the theorem to the function f_j. □

Convergence in measure yields the existence of a sequence $\varepsilon_k \to 0$ for which one has convergence almost everywhere. Under some additional assumptions on ϱ, one has convergence almost everywhere as $\varepsilon \to 0$ (see Chapter 5).

By choosing for ϱ a smooth function with bounded support and unit integral, we obtain constructive approximations of functions in $L^p(\mathbb{R}^n)$ by smooth functions with bounded derivatives (see Corollary 3.9.5).

Completing this section, we observe that there exist bounded measures μ such that the spaces $L^p(\mu)$ are not separable. As an example we mention the product of the continuum copies of the unit interval with Lebesgue measure. In this case, the family of all coordinate functions has cardinality of the continuum and the mutual distance between these functions in $L^1(\mu)$ is one and the same positive number. Hence one has the continuum of disjoint balls and no countable everywhere dense sets exist. The spaces $L^\infty(\mu)$ are nonseparable (excepting trivial cases) even for nice measures. For example, the space $L^\infty[0,1]$, where the interval is equipped with Lebesgue measure, is nonseparable because the distance between the functions $I_{[0,\alpha]}$ and $I_{[0,\beta]}$ with $0 < \alpha < \beta \leq 1$ equals 1.

4.3. The Hilbert space L^2

Let μ be a measure with values in $[0,+\infty]$. The space $L^2(\mu)$ is distinguished among other $L^p(\mu)$ by the property that it is Euclidean: its norm is generated by the inner product

$$(f,g) = \int_X fg\, d\mu.$$

It is clear that $fg \in L^1(\mu)$ whenever $f, g \in L^2(\mu)$, since $|fg| \leq f^2 + g^2$. In the case of the complex space $L^2(\mu)$ the inner product is given by the formula

$$(f,g) = \int_X f\overline{g}\, d\mu.$$

In order not to forget the complex conjugation over g, it is useful to remember that the inner product in $\mathbb{C}$ is given by the expression $z_1\overline{z}_2$, but not by $z_1 z_2$, which at $z_1 = z_2$ may be negative.

We recall that a linear space L is called Euclidean if it is equipped with an inner product, i.e., a function $(\,\cdot\,,\,\cdot\,)$ on $L\times L$ with the following properties:

1) $(x,x)\ge 0$ and $(x,x)=0$ precisely when $x=0$;

2) $(x,y)=(y,x)$ in the case of real L and $(x,y)=\overline{(y,x)}$ in the case of complex L;

3) the function $x\mapsto (x,y)$ is linear for every fixed vector y.

Every Euclidean space L has the following natural norm:

$$\|x\| = \sqrt{(x,x)}.$$

The fact that this is a norm indeed is easily verified by means of the following Cauchy–Bunyakowsky (or Cauchy–Bunyakowsky–Schwarz) inequality:

$$|(x,y)| \le \|x\|\,\|y\|. \tag{4.3.1}$$

In turn, for the proof of (4.3.1) it suffices to observe that the discriminant of the nonnegative second-order polynomial $t\mapsto (x+ty, x+ty)$ is nonpositive (in the complex case one can replace x by θx with $|\theta|=1$ such that $(\theta x, y)$ is real).

Two vectors x and y in a Euclidean space are called orthogonal, which is denoted by $x\perp y$, if $(x,y)=0$.

A Euclidean space that is complete with respect to its natural norm is called a *Hilbert space* in honor of the outstanding German mathematician David Hilbert. Thus, $L^2(\mu)$ is a Hilbert space. It is shown below that every infinite-dimensional separable Hilbert space is isomorphic to $L^2[0,1]$. Finite-dimensional Euclidean spaces are isomorphic to spaces $L^2(\mu)$ as well, but in that case one should take measures μ concentrated at finite sets.

4.3.1. Proposition. *Let H_0 be a closed linear subspace in a Hilbert space H. Then $H_0^\perp := \{x\in H\colon\ x\perp h\ \forall\, h\in H_0\}$ is a closed linear subspace in H and $H = H_0\oplus H_0^\perp$. Hence for every $h\in H$, there is a unique vector $h_0\in H_0$ with $h-h_0\in H_0^\perp$. In addition,*

$$\|h-h_0\| = \inf\bigl\{\|h-x\|\colon\ x\in H_0\bigr\}.$$

PROOF. Let us set $d=\inf\bigl\{\|h-x\|\colon\ x\in H_0\bigr\}$. Then, for any $n\in\mathbb{N}$, there exists a vector $x_n\in H_0$ such that $\|h-x_n\|^2\le d^2+n^{-1}$. We show that the sequence $\{x_n\}$ is fundamental. To this end, it suffices to observe that

$$\|x_n - x_k\| \le \frac{1}{\sqrt{n}} + \frac{1}{\sqrt{k}}.$$

Indeed, there exists a scalar t such that $h-\bigl(x_n+t(x_k-x_n)\bigr)\perp x_n-x_k$. Set $p=x_n+t(x_k-x_n)$. Then

$$\|h-p\|\le\|h-x_n\|,\quad \|h-p\|\le\|h-x_k\|.$$

It remains to apply the estimate

$$\|x_n-x_k\| \le \|x_n-p\| + \|x_k-p\| \le \frac{1}{\sqrt{n}} + \frac{1}{\sqrt{k}},$$

which follows from the equality (the Pythagorean theorem)

$$\|h - x_n\|^2 = \|h - p\|^2 + \|x_n - p\|^2$$

and the estimates $\|h - p\|^2 \geq d^2$, $\|h - x_n\|^2 \leq d^2 + n^{-1}$, and analogous relations for k. Since H is complete and H_0 is closed, the sequence $\{x_n\}$ converges to some element $h_0 \in H_0$. One has $\|h - h_0\|^2 \leq d^2$, whence we obtain $\|h - h_0\| = d$. Clearly, $h - h_0 \perp x$ for all $x \in H_0$, since otherwise one can take the vector $p = h_0 + (h_0 - h, x)x$, which gives the estimate $\|h - p\| < \|h - h_0\|$.

It is easily seen that $H_0^\perp$ is a closed linear subspace. If a vector $h_0' \in H_0$ is such that $h - h_0' \in H_0^\perp$, then $h_0 - h_0' \perp h_0 - h_0'$, hence $h_0' = h_0$. This shows that $H = H_0 \oplus H_0^\perp$. □

The vector h_0 constructed in the previous proposition is called the orthogonal projection of the vector h to the subspace H_0. As a corollary we obtain the Riesz theorem on the representation of linear functionals on Hilbert spaces. This theorem yields a natural isomorphism between a Hilbert space H and its dual H^*, i.e., the space of continuous linear functions on H.

4.3.2. Corollary. *Let f be a continuous linear function on a Hilbert space H. Then, there exists a unique vector v such that*

$$f(x) = (x, v) \quad \textit{for all } x \in H.$$

PROOF. By the continuity and linearity of f the set $H_0 = \{x\colon\ f(x) = 0\}$ is a closed linear subspace in H. For the identically zero functional our claim is trivial, so we assume that there is a vector u such that $f(u) = 1$. Let u_0 be the orthogonal projection of u to H_0 and let $v = \|u - u_0\|^{-2}(u - u_0)$. We show that $f(x) = (x, v)$ for all $x \in H$. Indeed, $x = f(x)u + z$, where $z = x - f(x)u \in H_0$, i.e., $z \perp u - u_0$. Hence $(x, v) = f(x)(u, v) = f(x)$ because

$$(u, v) = \|u - u_0\|^{-2}(u, u - u_0) = \|u - u_0\|^{-2}(u - u_0, u - u_0) = 1$$

by the orthogonality of $u - u_0$ and u_0. □

Riesz's theorem can be used for an alternative proof of the Radon–Nikodym theorem.

4.3.3. Example. Let μ and ν be two finite nonnegative measures on a measurable space $(X, \mathcal{A})$ and let $\nu \ll \mu$. Let us consider the measure $\lambda = \mu + \nu$. Then, every function ψ that is integrable with respect to λ is integrable with respect to μ and its integral against the measure μ does not change if one redefines ψ on a set of λ-measure zero. In addition,

$$\int_X |\psi|\, d\mu \leq \int_X |\psi|\, d\lambda.$$

Therefore, the linear function

$$L(\varphi) = \int_X \varphi\, d\mu$$

is well-defined on $L^2(\lambda)$ (is independent of our choice of a representative of φ) and, by the Cauchy–Bunyakowsky inequality, one has

$$|L(\varphi)| \le \int_X |\varphi|\, d\lambda \le \|1\|_{L^2(\lambda)} \|\varphi\|_{L^2(\lambda)}.$$

The estimate $|L(\varphi_1 - \varphi_2)| \le \|1\|_{L^2(\lambda)} \|\varphi_1 - \varphi_2\|_{L^2(\lambda)}$ yields the continuity of L. By the Riesz theorem, there exists an $\mathcal{A}$-measurable function $\psi \in \mathcal{L}^2(\lambda)$ such that

$$\int_X \varphi\, d\mu = \int_X \psi\varphi\, d\lambda \quad \text{for all } \varphi \in L^2(\lambda). \tag{4.3.2}$$

Therefore, $\mu = \psi\lambda$, $\nu = (1-\psi)\lambda$, since one can take $\varphi = I_A$, $A \in \mathcal{A}$. We show that the function $(1-\psi)/\psi$ serves as the Radon–Nikodym derivative $d\nu/d\mu$. Let $\Omega = \{x\colon\ \psi(x) \le 0\}$. Then Ω belongs to $\mathcal{A}$. Substituting in (4.3.2) the function $\varphi = I_\Omega$, we obtain

$$\mu(\Omega) = \int_\Omega \psi\, d\lambda \le 0,$$

whence $\mu(\Omega) = 0$. Let $\Omega_1 = \{x\colon\ \psi(x) > 1\}$. By using that $\mu(\Omega_1) \le \lambda(\Omega_1)$, we obtain in a similar way that the set Ω_1 has μ-measure zero, since

$$\mu(\Omega_1) = \int_{\Omega_1} \psi\, d\lambda > \lambda(\Omega_1).$$

Then the function f defined by the equality

$$f(x) = \frac{1-\psi(x)}{\psi(x)} \quad \text{if } x \notin \Omega, \qquad f(x) = 0 \quad \text{if } x \in \Omega,$$

is nonnegative and $\mathcal{A}$-measurable. We observe that the function f is integrable with respect to the measure μ. Indeed, the functions $f_n = fI_{\{\psi \ge 1/n\}}$ are bounded and increase pointwise to f such that

$$\int_X f_n\, d\mu = \int_X I_{\{\psi \ge 1/n\}}(1-\psi)\, d\lambda = \int_X I_{\{\psi \ge 1/n\}}\, d\nu \le \nu(X).$$

Hence the monotone convergence theorem applies. In addition, we obtain convergence of $\{f_n\}$ to f in $L^1(\mu)$. Finally, for every $A \in \mathcal{A}$, we have $I_A I_{\{\psi \ge 1/n\}} \to I_A$ μ-a.e., hence ν-a.e. (here we use the absolute continuity of ν with respect to μ). Hence

$$\nu(A) = \lim_{n\to\infty} \int_X I_A I_{\{\psi \ge 1/n\}}\, d\nu = \lim_{n\to\infty} \int_X I_A I_{\{\psi \ge 1/n\}} f\, d\mu = \int_A f\, d\mu$$

by convergence of $\{f_n\}$ to f in $L^1(\mu)$.

We now turn to orthonormal bases.

4.3.4. Corollary. *There exists a family of mutually orthogonal unit vectors e_α in $L^2(\mu)$ such that every element f in $L^2(\mu)$ is the sum of the following series convergent in $L^2(\mu)$:*

$$f = \sum_\alpha c_\alpha e_\alpha, \tag{4.3.3}$$

where at most countably many coefficients c_α may be nonzero. In addition, one has

$$c_\alpha = (f, e_\alpha), \quad \|f\|^2 = \sum_\alpha |c_\alpha|^2. \tag{4.3.4}$$

The family $\{e_\alpha\}$ is called an orthonormal basis of the space $L^2(\mu)$. If $L^2(\mu)$ is separable, then its orthonormal basis is finite or countable.

PROOF. Suppose first that $L^2(\mu)$ has a countable everywhere dense set $\{f_n\}$. Let $\|f_1\| > 0$ and let $e_1 = f_1/\|f_1\|$. We pick the first vector f_{i_2} that is linearly independent of e_1 and denote by g_2 the orthogonal projection of f_{i_2} to the linear span of e_1. Set $e_2 = (f_{i_2} - g_2)/\|f_{i_2} - g_2\|$. We continue the described process by induction. Suppose that we have already constructed a finite family $e_1, \ldots, e_n$ of mutually orthogonal unit vectors. If the linear span L_n of these vectors contains $\{f_n\}$, then it coincides with $L^2(\mu)$ because otherwise we could find a nonzero vector h orthogonal to all f_n, but such a vector is not approximated by the elements f_n due to the relation

$$\|h - f_n\|^2 = \|h\|^2 + \|f_n\|^2 \geq \|h\|^2.$$

If L_n does not contain $\{f_n\}$, then we take the first vector $f_{i_{n+1}} \notin L_n$, denote by g_{n+1} the orthogonal projection of $f_{i_{n+1}}$ to L_n (which exists, since L_n is finite-dimensional) and set

$$e_{n+1} = (f_{i_{n+1}} - g_{n+1})/\|f_{i_{n+1}} - g_{n+1}\|.$$

As a result we obtain either a finite basis or an orthonormal sequence $\{e_n\}$, the linear span L of which coincides with the linear span of $\{f_n\}$. Let us show that, for all $f \in L^2(\mu)$, the series $\sum_{n=1}^\infty (f, e_n)e_n$ converges to f. Let $\varepsilon > 0$. There is a function f_n satisfying the inequality $\|f - f_n\| < \varepsilon$. We pick N such that f_n is contained in the linear span of $e_1, \ldots, e_N$. Let $k \geq N$. It is easily seen that the vector $h_k = f - \sum_{i=1}^k (f, e_i)e_i$ is orthogonal to the vectors e_i, $i \leq k$. By the Pythagorean theorem, $\|f - f_n\|^2 = \|h_k\|^2 + \|h_k - f_n\|^2$, whence $\|h_k\| < \varepsilon$. This shows that the sums $\sum_{i=1}^k (f, e_i)e_i$ converge to f in $L^2(\mu)$.

If the space $L^2(\mu)$ has no countable everywhere dense sets, then the existence of an orthonormal basis is established by means of Zorn's lemma. Let us consider the set $\mathcal{M}$ consisting of all orthonormal systems. We have the following natural partial order on $\mathcal{M}$: $U \leq V$, i.e., the orthonormal system U is majorized by the orthonormal system V if U is a subset of V. It is clear that $U \leq U$ and that $U \leq W$ if $U \leq V$ and $V \leq W$. In addition, $U = V$ if $U \leq V$ and $V \leq U$. Suppose that $\mathcal{M}_0$ is a linearly ordered part of $\mathcal{M}$ (i.e., every two elements in $\mathcal{M}_0$ are comparable). Then the system formed by all vectors belonging to systems in $\mathcal{M}_0$ is orthonormal. Indeed, if a vector u comes from a system U and a vector v comes from a system V, then one of the two systems is contained in the other (for example, $U \subset V$) and hence $u \perp v$. By Zorn's lemma, there exists a maximal orthonormal system $\{e_\alpha\}$, i.e., a system such that there is no unit vector orthogonal to all its vectors. It follows by Proposition 4.3.1 that the linear span of the vectors e_α is everywhere dense in $L^2(\mu)$ (otherwise one could find a unit vector orthogonal to its closure).

Now let $f \in L^2(\mu)$. There exists a sequence of finite linear combinations of the vectors e_α convergent to f. Hence f belongs to the closure of the linear span of an at most countable collection $\{e_{\alpha_n}\}$. By the first step we obtain $f = \sum_{n=1}^\infty (f, e_{\alpha_n}) e_{\alpha_n}$. It is clear that our reasoning applies to any Hilbert space. □

It is seen from the proof that in the separable case most important for applications an orthonormal basis is obtained by means of the orthogonalization of an arbitrary sequence with a dense linear span. If $\{e_\alpha\}$ is an orthonormal basis in $L^2(\mu)$, then the numbers $c_\alpha = (\varphi, e_\alpha)$ are called the Fourier coefficients of the function $\varphi \in L^2(\mu)$. By using an orthonormal basis every separable infinite-dimensional Hilbert space can be identified with the space l^2 of all sequences $x = (x_n)$ with $\sum_{n=1}^\infty |x_n|^2 < \infty$, where in the real case $(x, y) := \sum_{n=1}^\infty x_n y_n$. Thus, all such spaces turn out to be isomorphic to the space $L^2[0,1]$ (an isomorphism of Hilbert spaces is a linear bijection preserving the inner product). An obvious corollary of the completeness of $L^2(\mu)$ is the following Riesz–Fischer theorem.

4.3.5. Theorem. *For any orthonormal system $\{\varphi_n\}$ in $L^2(\mu)$ and any sequence $\{c_n\} \in l^2$, the series $\sum_{n=1}^\infty c_n \varphi_n$ converges in $L^2(\mu)$.*

The reader will easily derive the following simple, but important result.

4.3.6. Theorem. *Let $\{\varphi_n\}$ be an orthonormal sequence in $L^2(\mu)$. Then, for all $f \in L^2(\mu)$, the following Bessel inequality holds*:

$$\sum_{n=1}^\infty |(f, \varphi_n)|^2 \le \|f\|^2_{L^2(\mu)}.$$

If f belongs to the closed linear span of $\{\varphi_n\}$ (and only for such f), then one has the Parseval equality

$$\sum_{n=1}^\infty |(f, \varphi_n)|^2 = \|f\|^2_{L^2(\mu)}.$$

In particular, this equality is true if $\{\varphi_n\}$ is an orthonormal basis.

It is easily seen that the above results are true for complex functions as well. In the following example we consider real spaces.

4.3.7. Example. (i) The sequence $1/\sqrt{2\pi}$, $\cos(nx)/\sqrt{\pi}$, $\sin(nx)/\sqrt{\pi}$, where $n \in \mathbb{N}$, is an orthonormal basis in $L^2[0, 2\pi]$ (in the complex case an orthonormal basis is formed by the functions $\exp(inx)/\sqrt{2\pi}$, $n \in \mathbb{Z}$).

(ii) The orthogonalization of the functions $1, x, x^2, \ldots$ in $L^2[-1,1]$ leads to the Legendre polynomials $L_n(x) = c_n \frac{d^n}{dx^n}(x^2-1)^n$, where c_n are normalization constants and $L_0 = 1$.

(iii) In the space $L^2(\gamma)$, where γ is the standard Gaussian measure on the real line with density $\exp(-x^2/2)/\sqrt{2\pi}$, an orthonormal basis is formed by the Chebyshev–Hermite polynomials

$$H_n(x) = \frac{(-1)^n}{\sqrt{n!}} e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}.$$

(iv) The functions $(2\pi)^{-1/4} H_n(x) \exp(-x^2/4)$ form an orthonormal basis in the space $L^2(\mathbb{R}^1)$.

Proof. (i) It is easily verified that the trigonometric system is orthogonal, and its completeness, i.e., the fact that its linear span is dense, follows, for example, from the Weierstrass theorem, which enables one to approximate uniformly any continuous 2π-periodic function by linear combinations of trigonometric functions (see Zorich [**1053**, Ch. XVI, §4]). (ii) The completeness of the Legendre system also follows by the Weierstrass theorem, and the indicated formula for them is left as Exercise 4.7.47. (iii) The fact that the Chebyshev–Hermite polynomials are orthonormal is verified by means of the integration by parts formula. Since H_n has the degree n, it follows that exactly these polynomials (up to a sign) are obtained after the orthogonalization of x^n. The completeness of $\{H_n\}$ in $L^2(\gamma)$ is proved as follows. Let $f \in L^2(\gamma)$ and $f \perp x^n$ for all n. The function

$$\varphi(z) = \int_{-\infty}^{+\infty} \exp(izx) f(x) \exp(-x^2/2)\, dx$$

is holomorph in the complex plane (it can be differentiated in z by the dominated convergence theorem). Then $\varphi^{(n)}(0) = 0$ for all $n = 0, 1, , \ldots$, whence $\varphi(z) = 0$ for all z. Therefore, $f(x) \exp(-x^2/2) = 0$ a.e. Finally, (iv) follows from (iii). $\square$

If $\{\varphi_n\}$ is an orthonormal basis in $L^2(\mu)$, then for all $\varphi \in L^2(\mu)$ the series $\varphi = \sum_{n=1}^{\infty} (\varphi, \varphi_n)\varphi_n$, called orthogonal, converge in $L^2(\mu)$. It is natural to ask about their convergence almost everywhere. By the Riesz theorem one can find a subsequence of partial sums convergent almost everywhere. However, the whole series may not converge almost everywhere. It was shown by L. Carleson that in the case of the trigonometric system in $L^2[0, 2\pi]$ one has convergence almost everywhere for all $\varphi \in L^2[0, 2\pi]$ (later R.A. Hunt extended Carleson's theorem to $L^p[0, 2\pi]$ with $p > 1$). A detailed proof can be read in Arias de Reyna [**36**], Jørboe, Mejlbro [**471**], Lacey [**564**], and Mozzochi [**702**]. On the other hand, the Fourier series with respect to the trigonometric system can be considered for functions $\varphi \in L^1[0, 2\pi]$. Set

$$a_n := \frac{1}{\pi} \int_0^{2\pi} \varphi(x)\, \cos nx\, dx, \quad b_n := \frac{1}{\pi} \int_0^{2\pi} \varphi(x)\, \sin nx\, dx. \tag{4.3.5}$$

Then, the formal series

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos nx + b_n \sin nx]$$

is called the Fourier series of the function φ with respect to the trigonometric system. A.N. Kolmogorov showed that there exists a function $\varphi \in \mathcal{L}^1[0, 2\pi]$ such that its Fourier series with respect to the trigonometric system diverges at every point. We shall see in Chapter 5 that if one is summing such a series not in the usual sense, but in the Cesàro or Abel sense (see below), then its sum coincides almost everywhere with φ. In the study of convergence of trigonometric Fourier series the following representation of partial sums is useful, which is obtained by the identity

$$\frac{1}{2} + \cos z + \cos 2z + \cdots + \cos kz = \frac{\sin \frac{2k+1}{2} z}{2 \sin \frac{z}{2}}$$

and elementary calculations:

$$\begin{aligned} S_n(x) :&= \frac{a_0}{2} + \sum_{k=1}^{n} [a_k \cos kx + b_k \sin kx] \\ &= \frac{1}{\pi} \int_0^{2\pi} \varphi(t) \Big[\frac{1}{2} + \sum_{k=1}^{n} \cos k(t-x) \Big] \, dt \qquad (4.3.6) \\ &= \frac{1}{\pi} \int_0^{2\pi} \varphi(t) \frac{\sin \frac{2k+1}{2}(t-x)}{2 \sin \frac{t-x}{2}} \, dt. \end{aligned}$$

This formula is a basis of several sufficient conditions for pointwise convergence of Fourier series (for example, Dini's condition, Exercise 4.7.68). For improving convergence of series the Cesàro method of summation is frequently used. Given a series with the terms α_n and the partial sums $s_n = \sum_{k=1}^{n} \alpha_k$, one considers the sequence $\sigma_n := (s_1 + \cdots + s_n)/n$. If the series $\sum_{n=1}^{\infty} \alpha_n$ converges to a number s, then the sequence σ_n converges to s as well, but the described transformation may produce a convergent sequence from a divergent series (for example, $\alpha_n = (-1)^n$). One more method of summation of series is called Abel's summation. Let us consider the power series $S(r) := \sum_{n=1}^{\infty} \alpha_n r^n$ for $r \in (0, 1)$. If the sums $S(r)$ are defined and have a finite limit s as $r \to 1$, then s is called the sum of the series $\sum \alpha_n$ in Abel's sense. If a series is Cesàro summable to a number s, then it is summable to s in Abel's sense (Exercise 4.7.51). When applied to the Fourier series of φ, the Cesàro summation leads, by virtue of the equality $\sum_{k=0}^{n-1} \sin(2k+1)z = \sin^2 nz / \sin z$, to the following Fejér sums (see Theorem 5.8.5):

$$\sigma_n(x) := \frac{S_0(x) + \cdots + S_n(x)}{n} = \int_0^{2\pi} \varphi(x+z) \Phi_n(z) \, dz, \qquad (4.3.7)$$

where the function

$$\Phi_n(z) = \frac{1}{2\pi n} \Big(\sin \frac{nz}{2} / \sin \frac{z}{2} \Big)^2$$

is called the Fejér kernel. Regarding trigonometric and orthogonal series, see Ahiezer [**4**], Bary [**66**], Edwards [**263**], Garsia [**346**], Hardy, Rogosinski [**409**], Kashin, Saakian [**495**], Olevskiĭ [**730**], Suetin [**920**], and Zygmund [**1055**], where one can find additional references.

4.4. Duality of the spaces L^p

The norm of a linear function Ψ on a normed space E is defined by the equality $\|\Psi\| = \sup_{\|v\|\le 1} |\Psi(v)|$. If $\|\Psi\| < \infty$, then Ψ is called bounded. Note that Ψ is bounded if and only if it is continuous. Indeed, $|\Psi(u) - \Psi(v)| = |\Psi(u-v)| \le \|\Psi\|\|u-v\|$; on the other hand, the continuity of Ψ implies its boundedness on some ball centered at the origin, hence on the unit ball. The space E^* of all continuous linear functions on E is called the dual to E. It is easily verified that E^* is complete with respect to the above norm. The general form of a continuous linear function on L^p is described by the following theorem due to F. Riesz. We recall that we often identify the elements of $L^p(\mu)$ with their representatives from $\mathcal{L}^p(\mu)$.

4.4.1. Theorem. *Suppose that a nonnegative measure μ on a σ-algebra $\mathcal{A}$ in a space X is finite or σ-finite and that $1 \le p < \infty$. Then, the general form of a continuous linear function on $L^p(\mu)$ is given by the formula*

$$\Psi(f) = \int_X fg\, d\mu, \tag{4.4.1}$$

where $g \in L^q(\mu)$, $p^{-1} + q^{-1} = 1$. In addition, $\|\Psi\| = \|g\|_q$.

Proof. Let $p > 1$ and $g \in L^q(\mu)$. By Hölder's inequality, the right-hand side of equality (4.4.1) gives a linear function Ψ on $L^p(\mu)$ and $|\Psi(f)| \le \|f\|_p\|g\|_q$, whence we obtain the continuity of Ψ and the estimate $\|\Psi\| \le \|g\|_q$. If $\|g\|_q = 0$, then $\Psi = 0$. In the case $\|g\|_q > 0$ we set $f = \operatorname{sign} g\, |g|^{q/p}/\|g\|_q^{q/p}$. Then $\|f\|_p = 1$ and

$$\Psi(f) = \|g\|_q^{-q/p} \int_X |g|^q\, d\mu = \|g\|_q^{-q/p}\|g\|_q^q = \|g\|_q.$$

Therefore, $\|\Psi\| = \|g\|_q$. For $p = 1$ we obtain $q = \infty$. In this case, one has the obvious inequality $\|\Psi\| \le \|g\|_\infty$. On the other hand, in the case of a nonzero measure μ (for $\mu = 0$ the assertion is trivial), for every $\varepsilon > 0$, the set $E := \{x\colon |g(x)| \ge \|g\|_\infty - \varepsilon\}$ has positive measure, which enables one to construct a nonnegative function f with $\|f\|_1 = 1$ that vanishes outside E. Then $\Psi(f \operatorname{sign} g) \ge \|g\|_\infty - \varepsilon$. Since $\|f \operatorname{sign} g\|_1 = 1$, we obtain $\|\Psi\| \ge \|g\|_\infty$.

Suppose now that Ψ is a continuous linear function on $L^p(\mu)$. Suppose first that the measure μ is finite. Set

$$\nu(A) = \Psi(I_A), \quad A \in \mathcal{A}.$$

If sets A_n in $\mathcal{A}$ are pairwise disjoint and their union is A, then the series $\sum_{n=1}^\infty I_{A_n}$ converges in $L^p(\mu)$ to I_A. This follows by the dominated convergence theorem because $\sum_{n=1}^N I_{A_n}(x) \to I_A(x)$ for each x and we have $\bigl|\sum_{n=1}^N I_{A_n}(x)\bigr| \le |I_A(x)|$. Hence ν is a countably additive measure. Since $\|I_A\|_p = \mu(A)^{1/p}$, the estimate

$$|\nu(A)| \le \|\Psi\|\mu(A)^{1/p}$$

yields the absolute continuity of ν with respect to μ. By the Radon–Nikodym theorem, there exists an integrable function g such that

$$\Psi(I_A) = \int_A g\,d\mu, \quad \forall\, A \in \mathcal{A}.$$

This means that equality (4.4.1) is valid for all simple functions f. Since any bounded measurable function is the uniform limit of a sequence of simple functions, (4.4.1) remains true for all bounded measurable functions f. Let us show that $g \in L^q(\mu)$. Indeed, let $q < \infty$ and $A_n = \{|g| \le n\}$. Let us set $f_n = |g|^{q/p} I_{A_n} \operatorname{sign} g$. Then f_n is a bounded measurable function, hence

$$\int_{A_n} |g|^q\,d\mu = \Psi(f_n) \le \|\Psi\|\,\|f_n\|_p = \|\Psi\| \left(\int_{A_n} |g|^q\,d\mu\right)^{1/p}.$$

Therefore, $\|gI_{A_n}\|_q \le \|\Psi\|$. By Fatou's theorem, $g \in L^q(\mu)$ and $\|g\|_q \le \|\Psi\|$. If $q = \infty$, then the set $A := \{x\colon\, g(x) > \|\Psi\|\}$ has measure zero because otherwise $\Psi(I_A/\mu(A)) > \|\Psi\|$. Similarly, the set $A := \{x\colon\, g(x) < -\|\Psi\|\}$ has measure zero. It remains to observe that the continuous linear functional generated by the function g on $L^p(\mu)$ coincides with Ψ on the everywhere dense set of simple functions, whence we obtain the equality of both functionals on all of $L^p(\mu)$. The case of a σ-finite measure is readily deduced from the proven assertion. □

This theorem does not extend to the case $p = \infty$. For example, on the space $L^\infty[0,1]$, where $[0,1]$ is equipped with Lebesgue measure, there exists a continuous linear function Ψ that cannot be represented in the form of (4.4.1). To this end, we define Ψ on the space $C[0,1]$ of continuous functions with the norm $\|f\| = \sup|f(t)|$ by the formula $\Psi(f) = f(0)$ and extend Ψ to a continuous linear function on $L^\infty[0,1]$ by the Hahn–Banach theorem 1.12.26. It is clear that even on continuous functions Ψ cannot be represented by formula (4.4.1). In fact, even without constructing concrete examples, the existence of such a function Ψ follows by the fact that $L^\infty[0,1]$ is nonseparable and the space $L^1[0,1]$ is separable. Exercise 4.7.87 outlines another method of proof of Theorem 4.4.1 for arbitrary infinite measures in the case $1 < p < \infty$. However, for $p = 1$ the above formulation of the theorem does not extend to arbitrary measures: it suffices to consider the measure μ on the class of all sets in $[0,1]$ that equals zero on the empty set and is infinite on all nonempty sets. Then only the identically zero function is integrable and the dual of $L^1(\mu)$ is $\{0\}$. Yet, in this example the unique continuous linear function on $L^1(\mu)$ is represented in the form of (4.4.1). Exercise 4.7.89 contains a construction of an example of a continuous linear function on $L^1(\mu)$ that does not admit representation (4.4.1). Exercise 4.7.93 deals with the dual to $L^1(\mu)$ for infinite measures. The above proof yields the following assertion.

4.4.2. Proposition. *Let μ be a finite nonnegative measure. A continuous linear function Ψ on $L^\infty(\mu)$ has the form (4.4.1), where $g \in L^1(\mu)$, precisely when the set function $A \mapsto \Psi(I_A)$ is countably additive.*

We recall the well-known Banach–Steinhaus theorem (also called "the uniform boundedness principle"), which we formulate along with its corollary.

4.4.3. Theorem. (i) *Let E be a Banach space and let a set $M \subset E^*$ be such that* $\sup_{l\in M} |l(x)| < \infty$ *for all $x \in E$ (for real E this is equivalent to the condition* $\sup_{l\in M} l(x) < \infty$ *for all $x \in E$). Then M is norm bounded.*

(ii) *A set in a normed space is bounded if every continuous linear function is bounded on it.*

PROOF. (i) Let us consider the sets

$$E_n := \bigl\{x \in E\colon\ |l(x)| \le n \text{ for all } l \in M\bigr\}.$$

Since M consists of continuous functions, the sets E_n are closed. By hypothesis, their union is E. Therefore, by the Baire category theorem (see Exercise 1.12.83), there is n such that E_n contains a closed ball $B(z, r)$ of radius $r > 0$ centered at a point z. Since the family M is uniformly bounded on $B(z, r)$ and consists of linear functions, it is uniformly bounded on the ball $B(0, r)$, hence on the ball $B(0, 1)$.

(ii) It is readily verified that the space E^* of continuous linear functions on a normed space E is a Banach space with the norm

$$\|f\| := \sup_{\|x\|\le 1} |f(x)|.$$

Every vector $x \in E$ generates a continuous linear function F_x on E^* by the formula $F_x(l) := l(x)$. One has $\|F_x\| = \|x\|$ because $|F_x(l)| \le \|l\|\|x\|$ $\|x\| = \sup_{f\in E^*, \|f\|\le 1} |f(x)|$ according to a simple corollary of the Hahn–Banach theorem: the functional $tx \mapsto t\|x\|$ on the line $\mathbb{R}^1 x$ can be extended to an element $f \in X^*$ of unit norm. It remains to apply assertion (i) to the functionals F_x, where x runs through the given set. $\square$

Applying this theorem to the spaces L^p (for definiteness, real), we arrive at the following result.

4.4.4. Proposition. *Let μ be a nonnegative finite or σ-finite measure on a space X. A set $\mathcal{F}$ is bounded in $L^p(\mu)$, where $p \in [1, +\infty)$, precisely when*

$$\sup_{f\in\mathcal{F}} \int_X fg\,d\mu < \infty \quad \textit{for all } g \in L^{p/(p-1)}(\mu).$$

4.4.5. Corollary. *Let μ be a nonnegative finite or σ-finite measure and let $p^{-1} + q^{-1} = 1$, where $1 < p < \infty$. Suppose that a measurable function f is such that $fg \in L^1(\mu)$ for all $g \in L^q(\mu)$. Then $f \in L^p(\mu)$.*

PROOF. Set $f_n(x) = f(x)$ if $|f(x)| \le n$ and $f_n(x) = 0$ if $|f(x)| > n$. For all $g \in L^q(\mu)$, we have $|f_n g| \le |fg|$ and $fg \in L^1(\mu)$. Hence the integrals of $f_n g$ converge to the integral of fg. This yields the uniform boundedness of integrals of $|f_n|^p$, hence $f \in L^p(\mu)$. $\square$

The next theorem strengthens the uniform boundedness principle for L^1. The previous proposition says that a set $\mathcal{F} \subset L^1(\mu)$ is bounded if

$$\sup_{f\in\mathcal{F}} \int_X fg\,d\mu < \infty \quad \text{for every } g \in \mathcal{L}^\infty(\mu).$$

It turns out that the boundedness is guaranteed by a yet weaker condition: it suffices to take for g only the indicators. As usual, we consider the real case (in the complex case one has to consider the absolute values of the integrals).

4.4.6. Theorem. *A family $\mathcal{F} \subset L^1(\mu)$, where the measure μ takes values in $[0, +\infty]$, is norm bounded in $L^1(\mu)$ precisely when for every $A \in \mathcal{A}$ one has*

$$\sup_{f\in\mathcal{F}} \biggl|\int_A f\,d\mu\biggr| < \infty.$$

PROOF. Suppose first that the measure μ is finite. The necessity of the above condition is obvious. Its sufficiency will be established if we show that

$$\sup_{A\in\mathcal{A}} \sup_{f\in\mathcal{F}} \biggl|\int_A f\,d\mu\biggr| < \infty.$$

Suppose that this is not true, i.e., there exist two sequences $A_n \in \mathcal{A}$ and $\{f_n\} \subset \mathcal{F}$ with

$$\biggl|\int_{A_n} f_n\,d\mu\biggr| > n.$$

We show that this leads to a contradiction. The idea of our reasoning is to apply Baire's category theorem to the complete metric space $\mathcal{A}/\mu$ (see §1.12(iii)). According to this theorem, if $\mathcal{A}/\mu = \bigcup_{n=1}^\infty M_n$, where M_n are closed sets, then at least one of the sets M_n contains a ball of positive radius. Set

$$M_n = \biggl\{A \in \mathcal{A}\colon\ \biggl|\int_A f_i\,d\mu\biggr| \le n, \quad \forall\, i\biggr\}.$$

Here we identify sets in $\mathcal{A}$ with their equivalence classes. It is clear that the sets M_n are closed in $\mathcal{A}/\mu$ and their union is $\mathcal{A}/\mu$. By Baire's theorem, there exist m, $\varepsilon > 0$, and $B \in \mathcal{A}$ such that for all i one has

$$\biggl|\int_A f_i\,d\mu\biggr| \le m \quad \text{whenever } \mu(A \triangle B) \le \varepsilon. \tag{4.4.2}$$

According to Theorem 1.12.9 we can decompose X into measurable sets $X_1, \ldots, X_k$ such that $\mu(X_j) \le \varepsilon$ for all $i = j, \ldots, p$ and the sets $X_{p+1}, \ldots, X_k$ are atoms with measures greater than ε. On any atom the function f_i coincides a.e. with a constant, hence there exists $C > 0$ such that, for all $j = 1, \ldots, k-p$ and all i, one has

$$\int_{X_{p+j}} |f_i|\,d\mu = \biggl|\int_{X_{p+j}} f_i\,d\mu\biggr| \le C.$$

Now let A be an arbitrary set in $\mathcal{A}$ and let $A_j = A \cap X_j$. For all $j = 1, \ldots, k-p$ we have for each i

$$\left|\int_{A_{p+j}} f_i \, d\mu\right| \le \int_{X_{p+j}} |f_i| \, d\mu \le C.$$

Let $j = 1, \ldots, p$. We observe that

$$A_j = (B \cup A_j)\backslash(B\backslash A_j). \tag{4.4.3}$$

Since $B \bigtriangleup (B \cup A_j) = A_j \backslash B$ and $B \bigtriangleup (B\backslash A_j) = B \cap A_j$, one has

$$\mu\bigl(B \bigtriangleup (B \cup A_j)\bigr) \le \mu(A_j) \le \varepsilon, \quad \mu\bigl(B \bigtriangleup (B\backslash A_j)\bigr) \le \mu(A_j) \le \varepsilon.$$

According to (4.4.3) and (4.4.2), for all i and $j = 1, \ldots, p$, we obtain

$$\begin{aligned} \left|\int_{A_j} f_i \, d\mu\right| &= \left|\int_{B\cup A_j} f_i \, d\mu - \int_{B\backslash A_j} f_i \, d\mu\right| \\ &\le \left|\int_{B\cup A_j} f_i \, d\mu\right| + \left|\int_{B\backslash A_j} f_i \, d\mu\right| \le 2m. \end{aligned}$$

Thus,

$$\left|\int_{A} f_i \, d\mu\right| \le 2mp + C(k-p) \quad \text{for all } i \text{ and } A \in \mathcal{A}.$$

In particular, this estimate is true for $A = A_i$, which is a contradiction.

It remains to reduce the general case to the case of a bounded measure. We observe that the measure μ is σ-finite on the set

$$X_0 = \bigcup_{n=1}^{\infty} \bigl\{x \colon\ |f_n(x)| \ne 0\bigr\}.$$

Thus, $X_0 = \bigcup_{n=1}^{\infty} X_n$, where $\mu(X_n) < \infty$ and the sets X_n are pairwise disjoint. We replace the measure μ by the finite measure $\mu_0 = \varrho \cdot \mu$, where $\varrho = 2^{-n}\bigl(1 + \mu(X_n)\bigr)^{-1}$ on X_n and $\varrho = 0$ outside X_0. Set $g_n = f_n/\varrho$. Then, for any $A \in \mathcal{A}$, we have

$$\int_A g_n \, d\mu_0 = \int_A f_n \, d\mu.$$

One has $\|g_n\|_{L^1(\mu_0)} = \|f_n\|_{L^1(\mu)}$. Thus, the functions g_n on $(X, \mathcal{A}, \mu_0)$ satisfy the same conditions as the functions f_n on $(X, \mathcal{A}, \mu)$. By the first step we conclude that the theorem is true in the general case. $\square$

4.5. Uniform integrability

In this section, we discuss the property of uniform integrability, which is closely connected with the property of absolute continuity and limit theorems for integrals.

Let $(X, \mathcal{A}, \mu)$ be a measure space with a nonnegative measure μ (finite or with values in $[0, +\infty]$).

4.5.1. Definition. *A set of functions $\mathcal{F} \subset \mathcal{L}^1(\mu)$ (or $\mathcal{F} \subset L^1(\mu)$) is called uniformly integrable if*

$$\lim_{C\to+\infty} \sup_{f\in\mathcal{F}} \int_{\{|f|>C\}} |f|\, d\mu = 0. \tag{4.5.1}$$

A set consisting of a single integrable function is uniformly integrable by the absolute continuity of the Lebesgue integral. Hence, for any integrable function f_0, the set of all measurable functions f with $|f| \le |f_0|$ is uniformly integrable. It is clear that in the case of an infinite measure, a bounded measurable function may not be integrable, although (4.5.1) is fulfilled for such functions. In the literature, one can encounter other definitions of uniform integrability that are equivalent to the one above in the case of bounded measures, but, in some respects, may be more natural for infinite measures (see Theorem 4.7.20(v) and Exercise 4.7.82).

4.5.2. Definition. *A family of functions $\mathcal{F} \subset \mathcal{L}^1(\mu)$ (or $\mathcal{F} \subset L^1(\mu)$) has uniformly absolutely continuous integrals if, for every $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$\int_A |f|\, d\mu < \varepsilon \quad \text{for all } f \in \mathcal{F} \text{ if } \mu(A) < \delta.$$

4.5.3. Proposition. *Let μ be a finite measure. A set $\mathcal{F}$ of μ-integrable functions is uniformly integrable precisely when it is bounded in $L^1(\mu)$ and has uniformly absolutely continuous integrals. If the measure μ is atomless, then the uniform integrability is equivalent to the uniform absolute continuity of integrals.*

Proof. Suppose that $\mathcal{F}$ is uniformly integrable. Let $\varepsilon > 0$. We can find $C > 0$ such that

$$\int_{\{|f|>C\}} |f|\, d\mu < \frac{\varepsilon}{2}, \quad \forall\, f \in \mathcal{F}.$$

Set $\delta = \varepsilon(2C)^{-1}$. Let $\mu(A) < \delta$. Then, for all $f \in \mathcal{F}$, we have

$$\int_A |f|\, d\mu = \int_{A\cap\{|f|\le C\}} |f|\, d\mu + \int_{A\cap\{|f|>C\}} |f|\, d\mu \le \frac{C\varepsilon}{2C} + \frac{\varepsilon}{2} = \varepsilon.$$

In addition,

$$\int_X |f|\, d\mu \le C\mu(X) + \int_{\{|f|>C\}} |f|\, d\mu < C\mu(X) + \frac{\varepsilon}{2}.$$

Suppose now that a set $\mathcal{F}$ is bounded in $L^1(\mu)$ and has uniformly absolutely continuous integrals. Let $\varepsilon > 0$. We take δ from the definition of the uniform absolute continuity of integrals and observe that by Chebyshev's inequality, there exists $C_1 > 0$ such that

$$\mu\bigl(\{|f| > C\}\bigr) \le C^{-1}\|f\|_{L^1(\mu)} < \delta$$

for all $f \in \mathcal{F}$ and $C > C_1$. Finally, if μ is atomless, then the uniform absolute continuity of integrals yields the boundedness in $L^1(\mu)$ because, for $\varepsilon = 1$, the space can be partitioned into finitely many (say, $N(\delta)$) sets with measures less than the corresponding δ, which gives $\|f\|_{L^1(\mu)} \le N(\delta)$ for all $f \in \mathcal{F}$. □

If μ has an atom (say, is the probability measure at the point 0), then the uniform absolute continuity of integrals does not imply the boundedness in $L^1(\mu)$, since the values of functions in $\mathcal{F}$ at this atom may be as large as we wish.

The next important result is called the Lebesgue–Vitali theorem.

4.5.4. Theorem. *Let μ be a finite measure. Suppose that f is a μ-measurable function and $\{f_n\}$ is a sequence of μ-integrable functions. Then, the following assertions are equivalent*:

(i) *the sequence $\{f_n\}$ converges to f in measure and is uniformly integrable*;

(ii) *the function f is integrable and the sequence $\{f_n\}$ converges to f in the space $L^1(\mu)$.*

PROOF. Suppose that condition (i) is fulfilled. Then the set $\{f_n\}$ is bounded in $L^1(\mu)$. By Fatou's theorem applied to the functions $|f_n|$, the function f is integrable. For the proof of convergence of $\{f_n\}$ to f in $L^1(\mu)$, it suffices to show that each subsequence $\{g_n\}$ in $\{f_n\}$ contains a subsequence $\{g_{n_k}\}$ convergent to f in $L^1(\mu)$. For $\{g_{n_k}\}$ we take a subsequence $\{g_n\}$ convergent to f almost everywhere, which is possible by the Riesz theorem. Let $\varepsilon > 0$. By Proposition 4.5.3, there exists $\delta > 0$ such that

$$\int_A |f_n| \, d\mu \le \varepsilon$$

for any n and any set A with $\mu(A) < \delta$. Applying Fatou's theorem, we obtain

$$\int_A |f| \, d\mu \le \varepsilon$$

whenever $\mu(A) < \delta$. By Egoroff's theorem, there exists a set A with $\mu(A) < \delta$ such that convergence of $\{g_{n_k}\}$ to f on $X\backslash A$ is uniform. Let N be such that $\sup_{X\backslash A} |g_{n_k} - f| \le \varepsilon$ for $k \ge N$. Then

$$\int_X |g_{n_k} - f| \, d\mu \le \varepsilon\mu(X) + \int_A |g_{n_k}| \, d\mu + \int_A |f| \, d\mu \le \varepsilon(2 + \mu(X)),$$

whence we obtain convergence of $\{g_{n_k}\}$ to f in $L^1(\mu)$.

If condition (ii) is fulfilled, then the sequence $\{f_n\}$ is bounded in $L^1(\mu)$ and converges in measure to f. In view of Proposition 4.5.3, it remains to observe that the sequence $\{f_n\}$ has uniformly absolutely continuous integrals. This follows by the estimate

$$\int_A |f_n| \, d\mu \le \int_A |f_n - f| \, d\mu + \int_A |f| \, d\mu$$

and the absolute continuity of the Lebesgue integral. □

Now we can transfer this theorem to infinite measures; the proof of the following corollary is left as Exercise 4.7.67.

4.5.5. Corollary. *Let μ be a measure with values in $[0, +\infty]$ and let functions $f_n, f \in \mathcal{L}^1(\mu)$ be such that $f_n(x) \to f(x)$ a.e. Then convergence of $\{f_n\}$ to f in $L^1(\mu)$ is equivalent to the following:*

$$\lim_{\mu(E)\to 0} \sup_n \int_E |f_n| \, d\mu = 0$$

and, for every $\varepsilon > 0$, there exists a measurable set X_ε such that $\mu(X_\varepsilon) < \infty$ and

$$\sup_n \int_{X \setminus X_\varepsilon} |f_n| \, d\mu < \varepsilon.$$

4.5.6. Theorem. *Suppose that a measure μ is finite or takes values in $[0, +\infty]$ and a sequence of μ-integrable functions f_n is such that for every set $A \in \mathcal{A}$, the sequence*

$$\int_A f_n \, d\mu$$

has a finite limit. Then, the sequence $\{f_n\}$ is bounded in $L^1(\mu)$ and has uniformly absolutely continuous integrals (in the case of a finite measure, it is uniformly integrable). In addition, there exists an integrable function f such that the above limit coincides with

$$\int_A f \, d\mu$$

for every set $A \in \mathcal{A}$.

PROOF. First we observe that the general case, as in Theorem 4.4.6, reduces to the case of a finite measure. Indeed, as in the cited theorem, the measure μ is σ-finite on the set $X_0 = \bigcup_{n=1}^\infty \{x\colon |f_n(x)| \neq 0\}$. Thus, $X_0 = \bigcup_{k=1}^\infty X_k$, where $\mu(X_k) < \infty$ and X_k are pairwise disjoint. Now we replace the measure μ by the finite measure $\mu_0 = \varrho \cdot \mu$, where

$$\varrho = 2^{-k}\bigl(1 + \mu(X_k)\bigr)^{-1} \text{ on } X_k \text{ and } \varrho = 0 \text{ outside } X_0.$$

Set $g_n = f_n/\varrho$. Then, for every $A \in \mathcal{A}$, we have

$$\int_A g_n \, d\mu_0 = \int_A f_n \, d\mu.$$

One has $\|g_n\|_{L^1(\mu_0)} = \|f_n\|_{L^1(\mu)}$. Hence the functions g_n on (X, μ_0) satisfy the same conditions as the functions f_n on (X, μ). So, if we prove our claim for g_n, then we obtain the theorem in the general case. In particular, if $g \in L^1(\mu_0)$ and

$$\int_A g \, d\mu_0 = \lim_{n\to\infty} \int_A g_n \, d\mu_0,$$

then the function $f = g\varrho$ can be taken for f. Thus, we assume that the measure μ is bounded. By Theorem 4.4.6, the sequence $\{f_n\}$ is bounded in $L^1(\mu)$. We show that the functions f_n have uniformly absolutely continuous integrals. As in the proof of Theorem 4.4.6, we consider the complete metric space $\mathcal{A}/\mu$ and, given $\varepsilon > 0$, we set

$$M_{k,m} = \left\{ A \in \mathcal{A}\colon \ \left| \int_A (f_k - f_m)\, d\mu \right| \le \varepsilon \right\}, \quad k, m \in \mathbb{N}.$$

The corresponding sets of equivalence classes in $\mathcal{A}/\mu$ will be denoted by $M_{k,m}$ as well. It is clear that these sets are closed in $\mathcal{A}/\mu$. Therefore, the sets $M_n = \bigcap_{k,m\ge n} M_{k,m}$ are closed. By the hypothesis of the theorem, one has

$$\mathcal{A}/\mu = \bigcup_{n=1}^{\infty} M_n,$$

since, for every $A \in \mathcal{A}$, the integrals of the functions f_k over A differ in at most ε for all sufficiently large k. By Baire's theorem (see Exercise 1.12.83) some M_n contains a ball, i.e., there exist $B \in \mathcal{A}$ and $r > 0$ such that for all $k, m \ge n$ we have

$$\left| \int_A (f_k - f_m)\, d\mu \right| \le \varepsilon \quad \text{if } \mu(A \triangle B) \le r. \tag{4.5.2}$$

Let us take a positive number $\delta < r$ such that, whenever $\mu(A) \le \delta$, one has

$$\left| \int_A f_j \, d\mu \right| \le \varepsilon, \quad j = 1, \ldots, n.$$

We observe that

$$\mu\bigl(B \triangle (A \cup B)\bigr) = \mu(A\backslash B) \le \delta < r, \tag{4.5.3}$$

$$\mu\bigl(B \triangle (B\backslash A)\bigr) = \mu(A \cap B) \le \delta < r. \tag{4.5.4}$$

For all $j > n$ we have

$$\begin{aligned} \int_A f_j \, d\mu &= \int_A f_n \, d\mu + \int_A (f_j - f_n)\, d\mu \\ &= \int_A f_n \, d\mu + \int_{A\cup B} (f_j - f_n)\, d\mu - \int_{B\backslash A} (f_j - f_n)\, d\mu. \end{aligned}$$

By (4.5.2), (4.5.3), (4.5.4) we obtain

$$\left| \int_A f_j \, d\mu \right| \le 3\varepsilon$$

for all j, whence the uniform absolute continuity of $\{f_n\}$ follows. In the case of a bounded original measure, we obtain the uniform integrability according to Proposition 4.5.3.

Now let us consider the set function

$$\nu(A) = \lim_{n\to\infty} \int_A f_n \, d\mu, \quad A \in \mathcal{A}.$$

Let us show that ν is a countably additive measure that is absolutely continuous with respect to μ. By the additivity of the integral we have the finite additivity of ν. Let $A_n \in \mathcal{A}$, $A_{n+1} \subset A_n$ and $\bigcap_{n=1}^\infty A_n = \varnothing$. Let $\varepsilon > 0$. One can take $\delta > 0$ such that

$$\left|\int_B f_n \, d\mu\right| \leq \varepsilon \quad \text{for all } n \text{ if } \mu(B) < \delta.$$

Next we pick N such that $\mu(A_n) < \delta$ for all $n \geq N$. Then, for all $n \geq N$, we obtain $|\nu(A_n)| \leq \varepsilon$, whence the countable additivity of ν follows (see Proposition 1.3.3). The absolute continuity of ν with respect to μ is obvious. By the Radon–Nikodym theorem $\nu = f \cdot \mu$, where $f \in L^1(\mu)$. □

4.5.7. Corollary. *If in the situation of the above theorem the functions f_n converge a.e., then their limit coincides a.e. with f and*

$$\lim_{n\to\infty} \|f_n - f\|_{L^1(\mu)} = 0.$$

Proof. The assertion reduces to the case of a σ-finite measure and subsequently to the case of a bounded measure as we usually do. In the latter case, letting $g(x) := \lim\limits_{n\to\infty} f_n(x)$, by the uniform integrability we obtain $\lim\limits_{n\to\infty} \|f_n - g\|_{L^1(\mu)} = 0$, whence $g(x) = f(x)$ a.e. □

It is the right place to remark that according to a nice theorem due to Fichtenholz, if integrable functions f and f_n, $n \in \mathbb{N}$, on the interval $[a, b]$ are such that

$$\lim_{n\to\infty} \int_U f_n \, dx = \int_U f \, dx$$

for every open set $U \subset [a, b]$, then this equality is true for every measurable set in $[a, b]$. Generalizations of this theorem are discussed in §8.10(x).

4.5.8. Corollary. *Suppose that a measure μ on the σ-algebra of all sets in a countable space $X = \{x_k\}$ is finite or takes values in $[0, +\infty]$ and that a sequence of functions $f_n \in L^1(\mu)$ is such that for every $A \subset X$ there exists a finite limit of the integrals of f_n over A. Then, the sequence $\{f_n\}$ converges in $L^1(\mu)$. In particular, if, for every n, we are given an absolutely convergent series $\sum_{j=1}^\infty \alpha_{n,j}$ such that, for every $A \subset \mathbb{N}$, there exists a finite limit $\lim\limits_{n\to\infty} \sum_{j\in A} \alpha_{n,j}$, then, there exists an absolutely convergent series with the general term α_j such that*

$$\lim_{n\to\infty} \sum_{j=1}^{\infty} |\alpha_{n,j} - \alpha_j| = 0.$$

Proof. Let us consider the measure μ on $\mathbb{N}$ that assigns the value 1 to every point. Then absolutely convergent series become functions in $L^1(\mu)$, and convergence on one-element sets in A becomes pointwise convergence. Therefore, we obtain not only convergence of the integrals on every set, but pointwise convergence as well, whence we obtain convergence in L^1. □

We now establish a useful criterion of uniform integrability, which is due to Ch.-J. de la Vallée Poussin. When applied to a family consisting of a single function it yields a useful "improvement of integrability".

4.5.9. Theorem. *Let μ be a finite nonnegative measure. A family $\mathcal{F}$ of μ-integrable functions is uniformly integrable if and only if there exists a nonnegative increasing function G on $[0,+\infty)$ such that*

$$\lim_{t\to+\infty}\frac{G(t)}{t}=\infty \quad \textit{and} \quad \sup_{f\in\mathcal{F}}\int G\bigl(|f(x)|\bigr)\,\mu(dx)<\infty. \tag{4.5.5}$$

In such a case, one can choose a convex increasing function G.

PROOF. Let condition (4.5.5) be fulfilled and let M majorize the integrals of the functions $G\circ|f|$, $f\in\mathcal{F}$. Given $\varepsilon>0$, we find a number C such that $G(t)/t\geq M/\varepsilon$ if $t\geq C$. Then, for every $f\in\mathcal{F}$, we have the inequality $|f(x)|\leq\varepsilon G\bigl(|f(x)|\bigr)/M$ whenever $|f(x)|\geq C$. Therefore,

$$\int_{\{|f|\geq C\}}|f|\,d\mu\leq\frac{\varepsilon}{M}\int_{\{|f|\geq C\}}G\circ|f|\,d\mu\leq\frac{\varepsilon}{M}M=\varepsilon.$$

Thus, the family $\mathcal{F}$ is uniformly integrable.

Let us prove the converse. The function G will be obtained in the form

$$G(t)=\int_0^t g(s)\,ds,$$

where g is an increasing nonnegative step function tending to $+\infty$ as $t\to+\infty$ and assuming the values α_n on the intervals $(n,n+1]$, where $n=0,1,\ldots$. In order to pick appropriate numbers α_n, we set for every $f\in\mathcal{F}$

$$\mu_n(f)=\mu\bigl(x\colon\ |f(x)|>n\bigr).$$

By the uniform integrability of $\mathcal{F}$, there exists a sequence of natural numbers C_n increasing to infinity such that

$$\sup_{f\in\mathcal{F}}\int_{\{|f|\geq C_n\}}|f|\,d\mu\leq 2^{-n}. \tag{4.5.6}$$

We observe that

$$\int_{\{|f|\geq C_n\}}|f|\,d\mu\geq\sum_{j=C_n}^{\infty}j\mu\bigl(x\colon\ j<|f(x)|\leq j+1\bigr)\geq\sum_{k=C_n}^{\infty}\mu_k(f).$$

It follows by (4.5.6) that

$$\sum_{n=1}^{\infty}\sum_{k=C_n}^{\infty}\mu_k(f)\leq 1 \quad \text{for all } f\in\mathcal{F}.$$

Now let $\alpha_n=0$ if $n<C_1$. If $n\geq C_1$, we set

$$\alpha_n=\max\{k\in\mathbb{N}\colon\ C_k\leq n\}.$$

It is clear that $\alpha_n \to +\infty$. For any $f \in \mathcal{F}$, we have

$$\int G\bigl(|f(x)|\bigr)\,\mu(dx)$$
$$\le \alpha_1\mu\bigl(x\colon\ 1<|f(x)|\le 2\bigr) + (\alpha_1+\alpha_2)\mu\bigl(x\colon\ 2<|f(x)|\le 3\bigr)+\dots$$
$$=\sum_{n=1}^{\infty}\alpha_n\mu_n(f)=\sum_{n=1}^{\infty}\sum_{k=C_n}^{\infty}\mu_k(f).$$

It remains to note that the function G is nonnegative, increasing, convex, and $G(t)/t \to +\infty$ as $t \to +\infty$. □

4.5.10. Example. Let μ be a finite nonnegative measure. A family $\mathcal{F}$ of μ-integrable functions is uniformly integrable provided that

$$\sup_{f\in\mathcal{F}}\int |f|\ln|f|\,d\mu<\infty,$$

where we set $0\ln 0 := 0$. In order to apply the criterion of de la Vallée Poussin, we take the function $G(t) = t\ln t$ for $t \ge 1$, $G(t) = 0$ for $t < 1$, and observe that $G(|f|) \le |f|\ln|f| + 1$. Another sufficient condition: for some $p > 1$ one has

$$\sup_{f\in\mathcal{F}}\int |f|^p\,d\mu<\infty.$$

4.6. Convergence of measures

There are several modes of convergence of measures, frequently used in applications. The principal ones are convergence in variation, setwise convergence, and, in the case where the space X is topological, weak convergence. In this section, we discuss the first two modes of convergence.

Let $(X,\mathcal{A})$ be a space with a σ-algebra and let $\mathcal{M}(X,\mathcal{A})$ be the space of all real countably additive measures on $\mathcal{A}$. It is clear that this is a linear space. We observe that the variation (see Definition 3.1.4) is a norm on $\mathcal{M}(X,\mathcal{A})$. This is obvious from expression (3.1.3) for $\|\mu\|$.

4.6.1. Theorem. *The space $\mathcal{M}(X,\mathcal{A})$ with the norm $\mu\mapsto\|\mu\|$ is a Banach space.*

Proof. If a sequence of measures μ_n in the space $\mathcal{M}(X,\mathcal{A})$ is fundamental in variation, then, for every $A\in\mathcal{A}$, the sequence $\{\mu_n(A)\}$ is fundamental and hence has some limit $\mu(A)$. Let us show that the set function $A\mapsto\mu(A)$ is countably additive and $\|\mu_n-\mu\|\to 0$. The additivity of μ is obvious from the additivity of the measures μ_n. We observe that

$$\lim_{n\to\infty}\sup\bigl\{|\mu(A)-\mu_n(A)|\colon\ A\in\mathcal{A}\bigr\}=0. \tag{4.6.1}$$

Indeed, let $\varepsilon>0$ and let n_0 be such that $\|\mu_n-\mu_k\|\le\varepsilon$ for all $n,k\ge n_0$. Let $A\in\mathcal{A}$. We pick $k\ge n_0$ such that $|\mu(A)-\mu_k(A)|\le\varepsilon$. Then, for all $n\ge n_0$,

we obtain

$$\begin{aligned}|\mu(A) - \mu_n(A)| &\le |\mu(A) - \mu_k(A)| + |\mu_k(A) - \mu_n(A)| \\ &\le \varepsilon + \|\mu_k - \mu_n\| \le 2\varepsilon,\end{aligned}$$

which proves (4.6.1). Now let $\{A_i\}$ be a sequence of pairwise disjoint sets in $\mathcal{A}$ and let $\varepsilon > 0$. We find n_0 such that

$$\sup\bigl\{|\mu(A) - \mu_n(A)|\colon\ A \in \mathcal{A}\bigr\} \le \varepsilon \quad \text{for all } n \ge n_0.$$

Next we find k_0 such that

$$\Bigl|\mu_{n_0}\Bigl(\bigcup_{i=k+1}^{\infty} A_i\Bigr)\Bigr| \le \varepsilon \quad \text{for all } k \ge k_0.$$

Then $\bigl|\mu\bigl(\bigcup_{i=k+1}^{\infty} A_i\bigr)\bigr| \le 2\varepsilon$ for all $k \ge k_0$. By the additivity of μ we finally obtain

$$\Bigl|\mu\Bigl(\bigcup_{i=1}^{\infty} A_i\Bigr) - \sum_{i=1}^{k} \mu(A_i)\Bigr| = \Bigl|\mu\Bigl(\bigcup_{i=k+1}^{\infty} A_i\Bigr)\Bigr| \le 2\varepsilon,$$

which gives the countable additivity of μ. Finally, relation (4.6.1) yields that $\|\mu - \mu_n\| \to 0$. □

It should be noted that $\mathcal{M}(X, \mathcal{A})$ can also be equipped with the norm

$$\mu \mapsto \sup_{A \in \mathcal{A}} |\mu(A)|,$$

equivalent to the variation norm (see (3.1.4)).

We now turn to setwise convergence of measures. This is a weaker mode of convergence than convergence in variation. For example, the sequence of measures μ_n on $[0, 2\pi]$ given by the densities $\sin nx$ with respect to Lebesgue measure converges on every measurable set to zero. This follows by the Riemann–Lebesgue theorem, according to which

$$\lim_{n\to\infty} \int_0^{2\pi} f(x) \sin nx \, dx = 0$$

for every integrable function f (Exercise 4.7.79).

4.6.2. Definition. *Let M be a family of real measures on a σ-algebra $\mathcal{A}$. This family is called uniformly countably additive if, for every sequence of pairwise disjoint sets A_i, the series $\sum_{i=1}^{\infty} \mu(A_i)$ converges uniformly in $\mu \in M$, i.e., for every $\varepsilon > 0$, there exists n_ε such that $\bigl|\sum_{i=n}^{\infty} \mu(A_i)\bigr| < \varepsilon$ for all $n \ge n_\varepsilon$ and all $\mu \in M$.*

The next important result unifies two remarkable facts in measure theory: the Nikodym convergence theorem and the Vitali–Lebesgue–Hahn–Saks theorem.

4.6.3. Theorem. *Let a sequence of measures μ_n in the space $\mathcal{M}(X,\mathcal{A})$ be such that $\lim\limits_{n\to\infty}\mu_n(A)$ exists and is finite for every set $A\in\mathcal{A}$. Then*:

(i) *the formula $\mu(A)=\lim\limits_{n\to\infty}\mu_n(A)$ defines a measure $\mu\in\mathcal{M}(X,\mathcal{A})$;*

(ii) *there exist a nonnegative measure $\nu\in\mathcal{M}(X,\mathcal{A})$ and a bounded nondecreasing nonnegative function α on $[0,+\infty)$ such that $\lim\limits_{t\to 0}\alpha(t)=0$ and*

$$\sup_n|\mu_n(A)|\le\alpha\bigl(\nu(A)\bigr),\quad \forall\, A\in\mathcal{A}. \tag{4.6.2}$$

In particular, $\sup_n\|\mu_n\|<\infty$ *and the sequence $\{\mu_n\}$ is uniformly countably additive*;

(iii) *if a nonnegative measure $\lambda\in\mathcal{M}(X,\mathcal{A})$ is such that $\mu_n\ll\lambda$ for all n, then*

$$\lim_{t\to 0}\sup\bigl\{\mu_n(A)\colon\ A\in\mathcal{A},\, \lambda(A)\le t,\, n\in\mathbb{N}\bigr\}=0.$$

Proof. Let $\nu=\sum_{n=1}^\infty c_n|\mu_n|$, where $c_n=2^{-n}(1+\|\mu_n\|)^{-1}$. It is clear that $\mu_n\ll\nu$ for all n. By the Radon–Nikodym theorem $\mu_n=f_n\cdot\nu$, where $f_n\in L^1(\nu)$. One has $\|\mu_n\|=\|f_n\|_{L^1(\nu)}$. By Theorem 4.5.6, the sequence $\{f_n\}$ is bounded in $L^1(\nu)$ and there exists a function $f\in L^1(\nu)$ such that

$$\lim_{n\to\infty}\int_A f_n\,d\nu=\int_A f\,d\nu,\quad \forall\, A\in\mathcal{A}.$$

Letting $\mu=f\cdot\nu$ we obtain a measure with the property mentioned in (i). According to Theorem 4.5.6, the functions f_n have uniformly absolutely continuous integrals, whence it follows that

$$\alpha(t)=\sup\biggl\{\int_A|f_n|\,d\nu\colon\ A\in\mathcal{A},\nu(A)\le t, n\in\mathbb{N}\biggr\}$$

tends to zero as $t\to 0$. It is clear that α is a nonnegative nondecreasing bounded function. Hence assertion (ii) is proven. The uniform countable additivity of μ_n follows by (ii). Finally, for the proof of (iii) it suffices to observe that the previous reasoning applies to λ in place of ν. □

4.6.4. Corollary. *Let measures $\mu_n\in\mathcal{M}(X,\mathcal{A})$ be such that for every set $A\in\mathcal{A}$ one has $\sup_n|\mu_n(A)|<\infty$. Then $\sup_n\|\mu_n\|<\infty$.*

Proof. If our claim is false, we can pass to a subsequence and assume that $\|\mu_n\|\ge n$. The measures $\mu_n/\sqrt{n}$ converge to zero at every set in $\mathcal{A}$. Hence $\sup_n\|\mu_n/\sqrt{n}\|<\infty$, which is a contradiction. □

Some conditions that are equivalent to the uniform countable additivity are collected in the following lemma.

4.6.5. Lemma. *Let M be a family of bounded measures on a σ-algebra $\mathcal{A}$. The following conditions are equivalent*:

(i) *the family M is uniformly countably additive*;

(ii) *one has $\lim\limits_{i\to\infty}\sup_{\mu\in M}|\mu(A_i)|=0$ for every sequence of pairwise disjoint sets $A_i\in\mathcal{A}$*;

(iii) *for every decreasing sequence of sets $A_i \in \mathcal{A}$ with $\bigcap_{i=1}^\infty A_i = \varnothing$, one has $\lim\limits_{i\to\infty} \mu(A_i) = 0$ uniformly in $\mu \in M$;*

(iv) *if a bounded nonnegative measure ν is such that $\mu_n \ll \nu$ for all n, then*

$$\lim_{t\to 0} \sup\bigl\{\mu(A)\colon\ \mu \in M,\ A \in \mathcal{A},\ \nu(A) \le t\bigr\} = 0.$$

PROOF. The equivalence of conditions (i) and (iii) is verified exactly as in the case of a single measure taking into account that, for any sequences of increasing sets A_i, the sets $A_{i+1}\backslash A_i$ are disjoint. It is clear that (i) yields (ii). In addition, (iv) yields (ii) and (iii). Let us verify that (ii) implies (iv). If this is not the case, there exists a bounded nonnegative measure ν with respect to which all the measures μ_n are absolutely continuous such that, for some $c > 0$ for every $\varepsilon > 0$, there exist an index m_ε and a set $A_\varepsilon \in \mathcal{A}$ with $\nu(A_\varepsilon) < \varepsilon$ and $|\mu_{m_\varepsilon}(A_\varepsilon)| > c$. We construct disjoint sets $B_i \in \mathcal{A}$ and indices k_i with $|\mu_{k_i}(B_i)| > c/2$, which will lead to a contradiction with (ii). To this end, we set $B_{1,1} = A_1$ and $k_1 = m_1$. Next we find $\varepsilon_1 > 0$ such that $|\mu_{k_1}|(E) < c/4$ for all $E \in \mathcal{A}$ with $\nu(E) < \delta$. Let $k_2 := m_{\varepsilon_1}$, $B_{2,1} := B_{1,1}\backslash A_{\varepsilon_1}$, $B_{2,2} := A_{\varepsilon_1}$. Then $|\mu_{k_1}(B_{2,1})| > c - c/4$. Suppose that for every $i \le n$, we have already found indices k_i and sets $B_{i,j}$ with $j = 1, \ldots, i$, such that $B_{i,j} \subset B_{i-1,j}$ if $j \le i-1$, $B_{i,j} \cap B_{i,k} = \varnothing$ if $j \ne k$, and

$$|\mu_{k_j}(B_{i,j})| > c - c/4 - \cdots - c/4^i$$

if $j \le i$. One can take $\varepsilon_n > 0$ such that $|\mu_{k_i}|(E) < c/4^{n+1}$ for all $i \le n$ whenever $\nu(E) < \varepsilon_n$. Finally, we set

$$k_{n+1} := m_{\varepsilon_n}, \quad B_{n+1,n+1} := A_{\varepsilon_n}, \quad B_{n+1,j} := B_{n,j}\backslash A_{\varepsilon_n}.$$

The sets $B_i := \bigcap_{n=1}^\infty B_{n,i}$ are the required ones. $\square$

An interesting generalization of this lemma is given in Theorem 4.7.27.

The proof of Theorem 4.6.3 gives in fact a stronger assertion (obtained by Saks [**841**]), namely, that the conclusion of the theorem remains true if one has convergence of $\mu_n(E)$ for all sets E from some class S of sets that is a second category set in the space $\mathcal{A}/\nu$, where ν is a nonnegative finite measure such that $\mu_n \ll \nu$, $\mu \ll \nu$. As already noted, Fichtenholz [**288**], [**290**] proved that if the integrals of functions $f_n \in L^1[0,1]$ over every open set converge to zero, then the integrals over every measurable set converge to zero as well (generalizations of this result to topological spaces are given in Chapter 8). G.M. Fichtenholz raised the question about a characterization of classes S of sets with the property that convergence to zero of integrals over the sets in S yields convergence to zero of integrals over all measurable sets. This problem was studied in Gowurin [**376**], where it was shown that S may even be a first category set in the metric space of all measurable sets in $[0,1]$ (see Exercises 4.7.134, 4.7.135).

4.7. Supplements and exercises

4.7(i). The spaces L^p and the space of measures as structures

We recall that an upper bound of a set F in a partially ordered set $(E, \leq)$ is an element $m \in E$ such that $f \leq m$ for all $f \in F$ (regarding partially ordered sets, see §1.12(vi)). An upper bound m is called a supremum of F if $m \leq \widetilde{m}$ for every other upper bound $\widetilde{m}$ of the set F. By analogy one defines the terms *lower bound* and *infimum.* A partially ordered set $(E, \leq)$ is called a *structure* or a *lattice* if every pair of elements $x, y \in E$ has a supremum denoted by $x \vee y$, and an infimum denoted by $x \wedge y$. A supremum is unique provided that the relations $x \leq y$ and $y \leq x$ yield that $x = y$. A structure E is called complete if every subset of E with an upper bound has a supremum. If this condition is fulfilled for all countable subsets, then E is called a σ-complete structure. A supremum of a set F in a lattice E is denoted by $\bigvee F$.

The set $\mathcal{L}^0(\mu)$ of real μ-measurable functions is a structure with its natural ordering: $f \leq g$ if $f(x) \leq g(x)$ μ-a.e. For $f \vee g$ and $f \wedge g$ one takes $\max(f, g)$ and $\min(f, g)$, respectively. It is clear that the classes of real functions $\mathcal{L}^p(\mu)$, $p \in (0, \infty]$, and the corresponding spaces $L^p(\mu)$ of equivalence classes are structures with the same ordering. Note that the relations $f \leq g$ and $g \leq f$ imply the equality $f = g$ in the classes $L^p(\mu)$ unlike the classes $\mathcal{L}^p(\mu)$.

4.7.1. Theorem. *Let $(X, \mathcal{A}, \mu)$ be a measure space with a σ-finite measure μ. Then, the sets $\mathcal{L}^0(\mu)$ and $L^0(\mu)$ are complete structures with the above-mentioned ordering. In addition, if a set $\mathcal{F} \subset \mathcal{L}^0(\mu)$ has an upper bound h, then there exists an at most countable set $\{f_n\} \subset \mathcal{F}$ such that*

$$f \leq \sup_n f_n \leq h \quad \textit{for all } f \in \mathcal{F}.$$

PROOF. It suffices to consider finite measures. The first claim is a corollary of the last one, which we now prove. Suppose first that there exists a number M such that $0 \leq f \leq M$ for all $f \in \mathcal{F}$. Let us add to $\mathcal{F}$ all functions of the form $\max(f_{\alpha_1}, \ldots, f_{\alpha_k})$, where $f_{\alpha_i} \in \mathcal{F}$. The obtained family is denoted by $\mathcal{G}$. It is clear that $\max(g_1, \ldots, g_k) \in \mathcal{G}$ for all $g_i \in \mathcal{G}$. Any upper bound of the family $\mathcal{F}$ is an upper bound for $\mathcal{G}$. Hence it suffices to prove our claim for $\mathcal{G}$. The integrals of functions in $\mathcal{G}$ have a finite supremum I. We can assume that the family $\mathcal{G}$ is infinite. Let us take a sequence of functions $g_n \in \mathcal{G}$ the integrals of which approach I. One can assume that $g_n(x) \leq g_{n+1}(x)$, passing to the sequence $g_n' = \max(g_n, g_{n-1}')$, $g_1' = g_1$. Set $g^*(x) = \lim\limits_{n\to\infty} g_n(x) = \sup_n g_n(x)$. Then the integral of g^* equals I. Let us show that $g(x) \leq g^*(x)$ a.e. for all $g \in \mathcal{G}$ (then $g^* \leq h$ for any upper bound h of the family $\mathcal{G}$). Indeed, otherwise there exists $g \in \mathcal{G}$ with $g(x) > g^*(x)$ on a set E of positive measure. Then

$$\int_E g\,d\mu \geq \int_E g^*\,d\mu + \varepsilon, \quad \text{where } \varepsilon > 0.$$

Let us take n such that

$$\int_X g_n\,d\mu > I - \varepsilon.$$

Letting $\psi := \max(g_n, g) \in \mathcal{G}$, we have

$$\int_X \psi\,d\mu \geq \int_{X\backslash E} g_n\,d\mu + \int_E g^*\,d\mu + \varepsilon \geq \int_X g_n\,d\mu + \varepsilon > I$$

contrary to the definition of I. In the case where the functions in $\mathcal{F}$ are nonnegative, it suffices to apply, for every fixed n, the above-proven assertion to the family of functions $\min(n, f)$, $f \in \mathcal{F}$. It is clear that it suffices to have the estimate $f \geq C$, $f \in \mathcal{F}$, for some C. Finally, in the general case, we fix $f_0 \in \mathcal{F}$ and partition X into disjoint sets $X_k := \{x\colon\, k < f_0(x) \leq k+1\}$, $k \in \mathbb{Z}$. On every X_k our claim is true, since one can apply what we have already proven to the family $\max(f, f_0)$, $f \in \mathcal{F}$. If $f_{k,n}$, $n \in \mathbb{N}$, is a sequence in $\mathcal{F}$ corresponding to the set X_k, then the countable family of functions $f_{k,n}$, $k, n \in \mathbb{N}$, is the required one for the whole X. □

4.7.2. Corollary. *The sets $\mathcal{L}^p(\mu)$ and $L^p(\mu)$, where the measure μ is σ-finite and $p \in [0, +\infty]$, are complete structures with the above-mentioned ordering. In addition, if a set $\mathcal{F} \subset \mathcal{L}^p(\mu)$ has an upper bound h, then its supremum in $\mathcal{L}^p(\mu)$ coincides with the supremum in $\mathcal{L}^0(\mu)$, and there exists an at most countable set $\{f_n\} \subset \mathcal{F}$ such that $f \leq \sup_n f_n \leq h$ for all $f \in \mathcal{F}$.*

PROOF. The case $p = 0$ has already been considered. This case and Fatou's theorem yield the assertion for $p \in (0, +\infty)$. The assertion for $p = \infty$ follows directly from the assertion for $p = 0$. □

4.7.3. Corollary. *Let μ be a finite nonnegative measure on a space $(X, \mathcal{A})$ and let A_t, $t \in T$, be a family of measurable sets. Then, it contains an at most countable subfamily $\{A_{t_n}\}$ such that $\mu\bigl(A_t \backslash \bigcup_{n=1}^\infty A_{t_n}\bigr) = 0$ for every t.*

PROOF. The function 1 majorizes the indicators of A_t. By the above theorem, there exists an at most countable family of indices t_n such that, for every t, we have $I_{A_t} \leq \sup_n I_{A_{t_n}}$ a.e. Hence a.e. point x from A_t is contained in $\bigcup_{n=1}^\infty A_{t_n}$. □

It is to be noted that a supremum $\bigvee \mathcal{F}$ of a set $\mathcal{F}$ in $\mathcal{L}^p(\mu)$ may not coincide with the function $\sup_{f\in\mathcal{F}} f(x)$ defined pointwise. For example, let F be a set in $[0,1]$. For $t \in F$, we set $f_t(s) = 1$ if $s = t$, $f_t(s) = 0$ if $s \neq t$, where $s \in [0,1]$. Then $\sup_{t\in F} f_t(s) = I_F(s)$, although the identically zero function is a supremum of the family $\{f_t\}$ in $\mathcal{L}^p[0,1]$. If F is not measurable, then the function $\sup_{t\in F} f_t(s) = I_F(s)$ is nonmeasurable as well. As an example of an incomplete structure one can indicate the space $C[0,1]$ of continuous functions with its natural order $f \leq g$. In this structure, the set of all continuous functions vanishing on $[0,1/2)$ and majorized by 1 on $[1/2,1]$ has an upper bound 1, but it has no supremum. If the set of all measurable functions on $[0,1]$ is equipped with the partial order corresponding to the inequality $f(x) \leq g(x)$ for each x (in place of the comparison almost everywhere used above), then we also obtain an incomplete structure.

It is worth mentioning that the above results do not extend to arbitrary infinite measures, although there exist non-σ-finite measures for which they are true (see Exercise 4.7.91).

As an application of the above results we prove the following useful assertion from Halmos, Savage [**405**].

4.7.4. Theorem. *Let μ_t, $t \in T$, be a family of probability measures on a σ-algebra $\mathcal{A}$ absolutely continuous with respect to some fixed probability measure μ on $\mathcal{A}$. Then, there exists an at most countable set of indices t_n such that all measures μ_t are absolutely continuous with respect to the probability measure $\sum_{n=1}^{\infty} 2^{-n}\mu_{t_n}$.*

PROOF. By hypothesis, $\mu_t = f_t \cdot \mu$, where $f_t \in L^1(\mu)$. Let us consider μ-measurable sets $X_t = \bigl\{x\colon\ f_t(x) \neq 0\bigr\}$ and apply Theorem 4.7.1 to the family of indicators I_{X_t} (they are majorized by the function 1). By the cited theorem, there exists an at most countable family of indices t_n such that, for every t, we have $I_{X_t}(x) \leq \sup_n I_{X_{t_n}}(x)$ μ-a.e. This means that on the set $\bigl\{x\colon\ \sum_{n=1}^{\infty} 2^{-n} f_{t_n}(x) = 0\bigr\}$ we have $f_t(x) = 0$ for μ-a.e. x. Therefore, the measure μ_t is absolutely continuous with respect to the probability measure $\sum_{n=1}^{\infty} 2^{-n}\mu_{t_n}$. □

Let us now show that the space $\mathcal{M}(X,\mathcal{A})$ of all bounded signed measures on $\mathcal{A}$ is a complete structure. One has the following natural partial order on $\mathcal{M}(X,\mathcal{A})$: $\mu \leq \nu$ if and only if $\mu(A) \leq \nu(A)$ for all $A \in \mathcal{A}$.

For any $\mu,\nu \in \mathcal{M}(X,\mathcal{A})$, we set

$$\mu \vee \nu := \mu + (\nu - \mu)^+, \quad \mu \wedge \nu := \mu - (\nu - \mu)^-.$$

If μ and ν are given by densities f and g with respect to some nonnegative measure λ (for example, $\lambda = |\mu| + |\nu|$), then

$$\mu \vee \nu = \max(f,g) \cdot \lambda, \quad \mu \wedge \nu = \min(f,g) \cdot \lambda.$$

It is readily seen that $\mu \vee \nu$ is the minimal measure majorizing μ and ν. Indeed, if a measure η is such that $\mu \leq \eta$ and $\nu \leq \eta$, then we take a nonnegative measure λ such that $\mu = f \cdot \lambda$, $\nu = g \cdot \lambda$, $\eta = h \cdot \lambda$. One has $h \geq f$ and $h \geq g$ λ-a.e., whence $h \geq \max(f,g)$ λ-a.e. Thus, $\mathcal{M}(X,\mathcal{A})$ is a structure. It is obvious that suprema and infima of subsets of $\mathcal{M}(X,\mathcal{A})$ are uniquely defined.

4.7.5. Theorem. *The structure $\mathcal{M}(X,\mathcal{A})$ is complete.*

PROOF. Suppose that a set $M \subset \mathcal{M}(X, \mathcal{A})$ is majorized by a measure μ. Let us show that M has a supremum (which is uniquely defined in $\mathcal{M}(X, \mathcal{A})$). Suppose first that all measures in M are nonnegative. Then, for each $m \in M$, we have $m \ll \mu$ and by the Radon–Nikodym theorem $m = f_m \cdot \mu$, where $f_m \in L^1(\mu)$. The condition $m \leq \mu$ means that $f_m \leq 1$ μ-a.e., i.e., the family $\{f_m\}$ is majorized by the function 1 and by the above results has a supremum f in $L^1(\mu)$. It is clear that the measure $f \cdot \mu$ is the supremum of M. The case where there exists a measure μ_0 such that $\mu_0 \leq m$ for all $m \in M$, reduces to the above-considered situation, since the set $M - \mu_0$ consists of nonnegative measures and is majorized by the measure $\mu - \mu_0$. If ν is the supremum of $M - \mu_0$, then $\nu + \mu_0$ is the supremum of M. Let us consider the general case and fix $m_0 \in M$. The set $M_0 = \{m \vee m_0, m \in M\}$ consists of measures majorizing the measure m_0. In addition, $m \vee m_0 \leq \mu$ for all $m \in M$, since $m_0 \leq \mu$ and $m \leq \mu$. As we have established, M_0 has a supremum ν. Let us show that ν is the supremum of M. Indeed, $m \leq m \vee m_0 \leq \nu$ for all $m \in M$. Suppose that η is a measure such that $m \leq \eta$ for all $m \in M$. In particular, $m_0 \leq \eta$, whence $\eta \vee m_0 = \eta$. Then $m \vee m_0 \leq \eta \vee m_0 = \eta$ for all $m \in M$, i.e., η is an upper bound for M_0, whence we obtain $\nu \leq \eta$. Thus, ν is the smallest upper bound, i.e., it is the supremum. □

4.7(ii). The weak topology in L^p

In applications one frequently uses elementary properties of the weak topology in the space L^p, which we briefly discuss here. We recall that a sequence of vectors x_n in a normed space E is called weakly convergent to a vector x if $l(x_n) \to l(x)$ for all $l \in E^*$, where E^* is the space of all continuous linear functions on E. If, for every $l \in E^*$, the sequence $l(x_n)$ is fundamental, then $\{x_n\}$ is called weakly fundamental. This convergence can be described by means of the so-called *weak topology* on E, in which the open sets are all possible unions of sets of the form

$$U(a, l_1, \ldots, l_n, \varepsilon_1, \ldots, \varepsilon_n) = \bigl\{x\colon\ |l_1(x-a)| < \varepsilon_1, \ldots, |l_n(x-a)| < \varepsilon_n\bigr\},$$
$$a \in E,\ l_i \in E^*,\ \varepsilon_i > 0,\ n \in \mathbb{N},$$

and also the empty set. It is seen from the definition that in any infinite-dimensional space E, every nonempty set that is open in the weak topology contains an infinite-dimensional affine subspace, for $U(0, l_1, \ldots, l_n, \varepsilon_1, \ldots, \varepsilon_n)$ contains the intersection of the hyperplanes $l_i^{-1}(0)$. Hence such a set is not bounded, whence we conclude that in any infinite-dimensional space E the weak topology is strictly weaker than the topology generated by the norm. However, it may occur that the collections of convergent (countable) sequences are the same in the weak topology and norm topology. As an example we mention the space l^1 of all real sequences $x = (x_n)$ with finite norm $\|x\| = \sum_{n=1}^{\infty} |x_n|$. This space can be regarded as the space $L^1(\mathbb{N}, \nu)$, where ν is the measure on $\mathbb{N}$ assigning the value 1 to every point. The fact that weak convergence of a sequence in l^1 yields norm convergence is clear from

Corollary 4.5.8. However, in every space $L^p[a,b]$, $1 \le p \le \infty$, one can find a sequence that converges weakly, but not in the norm. For example, if $\{e_n\}$ is an orthonormal basis in $L^2[a,b]$, then $e_n \to 0$ in the weak topology, but there is no norm convergence.

It is worth noting that the weak topology is a special case of the topology $\sigma(E,F)$, where E is a linear space (not necessarily normed) and F is some linear space of linear functions on E separating the points in E (i.e., for every $x \neq 0$, there exists $l \in F$ with $l(x) \neq 0$). The topology $\sigma(E,F)$ is called the topology generated by the duality with F and is defined by means of the same sets $U(a, l_1, \ldots, l_n, \varepsilon_1, \ldots, \varepsilon_n)$ as above, with the only difference that now $l_i \in F$. Letting $F = E^*$ in the case of a normed space E we arrive at the weak topology. It is readily verified that if a linear function l on E is continuous in the topology $\sigma(E,F)$, then $l \in F$ (details can be found in Schaefer [**848**, Ch. IV]). Thus, the dual (the set of all continuous linear functions) to the space E with the topology $\sigma(E,F)$ is exactly F. In particular, in spite of the fact that the weak topology of a normed space is weaker than the norm topology, it yields the same collection of continuous linear functions.

Let μ be a nonnegative (possibly infinite) measure on the space $(\Omega, \mathcal{A})$. By the Banach–Steinhaus theorem (see §4.4) we obtain the following result.

4.7.6. Proposition. *Every weakly convergent sequence in $L^p(\mu)$ is norm bounded.*

We know that any continuous linear function on $L^p(\mu)$ with $1 < p < \infty$ is generated by an element of $L^q(\mu)$, where $q = p/(p-1)$ (we have considered above the case of a finite or σ-finite measure, and the case of an arbitrary measure is considered in Exercise 4.7.87). Hence convergence of a sequence of functions f_n to a function f in the weak topology of $L^p(\mu)$, $1 < p < \infty$, is merely the relation

$$\lim_{n\to\infty} \int_\Omega f_n g \, d\mu = \int_\Omega f g \, d\mu, \quad \forall\, g \in L^q(\mu).$$

The properties of the weak topology in L^1 and L^p with $p > 1$ differ considerably. Here we give several results in the case $p > 1$; the case $p = 1$ will be considered separately.

It follows by the above results that the spaces $L^p(\mu)$ with $1 < p < \infty$ are reflexive in the sense of the following definition.

4.7.7. Definition. *A Banach space E is called reflexive if, for every continuous linear functional f on E^*, there exists a vector $v \in X$ such that $f(l) = l(v)$ for all $l \in X^*$.*

The reflexivity of a space E is written concisely as the equality $E^{**} = E$. The reader is warned that this equality is not the same as the existence of an isometry between E^{**} and E!

4.7.8. Theorem. *Either of the following conditions is equivalent to the reflexivity of a Banach space E:*

(i) *the closed unit ball in the space E is compact in the weak topology;*

(ii) *every continuous linear functional on E attains its maximum on the closed unit ball.*

See Diestel [**222**] for a proof.

4.7.9. Corollary. *In the spaces $L^p(\mu)$ with $1 < p < \infty$, all closed balls are compact in the weak topology. In addition, every norm bounded sequence of functions f_n contains a subsequence that converges in the weak topology to some function $f \in L^p(\mu)$.*

We note that for separable spaces $L^p(\mu)$, the last assertion has a trivial proof: one takes a countable everywhere dense set of functions g_i in $L^q(\mu)$ and picks a subsequence f_{n_k} such that the integrals of $f_{n_k} g_i$ converge for each i. The general case can be reduced to this one (passing to the σ-algebra generated by $\{f_n\}$), but it is simpler to apply the following Eberlein–Šmulian theorem (a proof can be found in Dunford, Schwartz [**256**, Ch. V, §6]), which we shall also use in the case $p = 1$.

4.7.10. Theorem. *Let A be a set in a Banach space E. Then, the following conditions are equivalent:* (i) *the closure of A in the weak topology is compact;* (ii) *every sequence in A has a subsequence that converges weakly in E;* (iii) *every infinite sequence in A has a limit point in E in the weak topology* (*i.e., a point every neighborhood of which contains infinitely many points of this sequence*).

One more useful general result about weak convergence is the following Krein–Milman theorem (see Dunford, Schwartz [**256**, Ch. V, §6]).

4.7.11. Theorem. *Suppose that a set A in a Banach space E is compact in the weak topology. Then, the closed convex envelope of A* (*i.e., the intersection of all closed convex sets containing A*) *is compact in the weak topology.*

The next result characterizes weak convergence in L^p for sequences convergent almost everywhere or in measure. We emphasize, however, that weak convergence in L^p does not yield convergence in measure.

4.7.12. Proposition. *Let $1 < p < \infty$ and let functions $f_n \in \mathcal{L}^p(\mu)$ converge almost everywhere* (*or in measure*) *to a function f. Then, a necessary and sufficient condition for convergence of $\{f_n\}$ to f in the weak topology of $L^p(\mu)$ is the boundedness of $\{f_n\}$ in the norm of $L^p(\mu)$.*

PROOF. The boundedness in the norm follows by weak convergence. Let $\{f_n\}$ be bounded in $L^p(\mu)$. By Exercise 4.7.76 it suffices to verify convergence of the integrals of $f_n g$ to the integral of fg for every simple μ-integrable function g (the function g is nonzero only on a set of finite measure). This convergence takes place indeed by convergence of $f_n g$ to fg almost everywhere (or in measure), since all these functions are nonzero only on a set of finite measure and are uniformly integrable due to the boundedness of $\{f_n g\}$ in $L^p(\mu)$. $\square$

In the case $p = 1$, almost everywhere convergence and norm boundedness do not yield weak convergence. Indeed, otherwise we would obtain the uniform integrability of f_n, hence convergence in the norm, but it is easily seen that the functions $f_n(x) = nI_{(0,1/n]}(x)$ have unit norms in $L^1[0,1]$ and converge pointwise to zero.

In connection with the above proposition, see also Proposition 4.7.30.

For $p = 1$ weak convergence in $L^1(\mu)$ along with almost everywhere convergence yield convergence in the norm by Corollary 4.5.7. For $p > 1$ this is not true (Exercise 4.7.78).

One more interesting property of weak convergence in L^p is given in Corollary 4.7.16 below.

Another important special case of a topology of the form $\sigma(E, F)$ is the weak* topology on the dual space E^* of a normed space E. This topology is denoted by $\sigma(E^*, E)$ and defined as the topology on E^* generated by the duality with the space E regarded as the space of linear functions on E^*: every element $x \in E$ generates a linear function on E^* by the formula $l \mapsto l(x)$. Convergence of functionals in the weak* topology is merely convergence at every vector in E. For a reflexive Banach space E, the weak* topology on E^* coincides with the weak topology of the Banach space E^*. An important property of the weak* topology is expressed by the following Banach–Alaoglu theorem (see, e.g., Dunford, Schwartz [**256**, Ch. V, §4]).

4.7.13. Theorem. *Let E be a normed space. Then, the closed balls in E^* are compact in the weak* topology.*

If E is separable, then the closed balls in E^* are metrizable compacts in the weak* topology. In this case, every bounded sequence in E^* contains a weakly* convergent subsequence (of course, the last claim can be easily proved directly by choosing a subsequence that converges at every element of a countable everywhere dense set). However, in the general case this is not true. For example, if $E = l^\infty$, then the sequence of functionals $l_n \in E^*$ defined by $l_n(x) = x_n$ has no weakly* convergent subsequences (otherwise such a subsequence would be weakly* convergent to zero, which is impossible). Thus, for the weak* topology (unlike the weak topology) compactness is not equivalent to sequential compactness.

4.7(iii). Uniform convexity of L^p

4.7.14. Definition. *A normed space E with the norm $\|\cdot\|$ is called uniformly convex if, for every $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$\text{whenever } \|x\| = 1, \ \|y\| = 1 \text{ and } \left\|\frac{x+y}{2}\right\| \geq 1 - \delta, \text{ one has } \|x - y\| \leq \varepsilon.$$

Let μ be a nonnegative measure (possibly with values in $[0,+\infty]$) on a measurable space $(X,\mathcal{A})$.

4.7.15. Theorem. *For $1<p<\infty$, the spaces $L^p(\mu)$ are uniformly convex.*

PROOF. We observe that, for every $\varepsilon>0$, there exists $\delta=\delta(p,\varepsilon)>0$ such that, for all $a,b\in\mathbb{R}$, we have

$$\varepsilon^p(|a|^p+|b|^p)\le 4|a-b|^p \Rightarrow \Bigl|\frac{a+b}{2}\Bigr|^p\le(1-\delta)\frac{|a|^p+|b|^p}{2}. \qquad (4.7.1)$$

Indeed, it suffices to show that such δ exists for all real numbers a and b with $1\le a^2+b^2\le 2$, since for every nonzero vector (a,b) in the plane one can find $t>0$ such that the vector (ta,tb) belongs to the indicated ring, and both inequalities in (4.7.1) are then multiplied by t^p. By the compactness argument it is clear that in the absence of a required δ, there exists a vector (a,b) such that

$$1\le a^2+b^2\le 2,\ \varepsilon^p(|a|^p+|b|^p)\le 4|a-b|^p,\ \Bigl|\frac{a+b}{2}\Bigr|^p\ge\frac{|a|^p+|b|^p}{2}.$$

The last inequality is only possible if $a=b$, which is obvious from the consideration of the graph of the function $|x|^p$ with $p>1$. Now the first two of the foregoing inequalities are impossible. This contradiction proves (4.7.1).

Let $\varepsilon>0$ and let functions f and g have unit norms in $L^p(\mu)$ and satisfy the inequality $\|f-g\|_{L^p(\mu)}\ge\varepsilon$. Let us consider the set

$$\Omega=\Bigl\{x\colon\ \varepsilon^p(|f(x)|^p+|g(x)|^p)\le 4|f(x)-g(x)|^p\Bigr\}.$$

By (4.7.1) we obtain

$$\Bigl|\frac{f(x)+g(x)}{2}\Bigr|^p\le(1-\delta)\frac{|f(x)|^p+|g(x)|^p}{2},\quad \forall\,x\in\Omega. \qquad (4.7.2)$$

It is clear that

$$\int_{X\backslash\Omega}|f-g|^p\,d\mu\le\frac{\varepsilon^p}{4}\int_X\bigl[|f|^p+|g|^p\bigr]\,d\mu\le\frac{\varepsilon^p}{2},$$

whence one has

$$\int_\Omega|f-g|^p\,d\mu\ge\frac{\varepsilon^p}{2}.$$

Taking into account the estimate $\bigl(|f|^p+|g|^p\bigr)/2-|(f+g)/2|^p\ge 0$ and inequality (4.7.2) we obtain

$$\begin{aligned}\int_X\biggl(\frac{|f|^p+|g|^p}{2}-\Bigl|\frac{f+g}{2}\Bigr|^p\biggr)\,d\mu&\ge\int_\Omega\biggl(\frac{|f|^p+|g|^p}{2}-\Bigl|\frac{f+g}{2}\Bigr|^p\biggr)\,d\mu\\ &\ge\delta\int_\Omega\frac{|f|^p+|g|^p}{2}\,d\mu\ge\delta 2^{-p-1}\int_\Omega|f-g|^p\,d\mu\ge\frac{\delta}{4}\frac{\varepsilon^p}{2^p}.\end{aligned}$$

Therefore,

$$\int_X\Bigl|\frac{f+g}{2}\Bigr|^p\,d\mu\le 1-\frac{\delta}{4}\frac{\varepsilon^p}{2^p},$$

which means the uniform convexity of $L^p(\mu)$. The theorem is proven. □

4.7.16. Corollary. *Suppose that a sequence of functions f_n converges weakly to a function f in $L^p(\mu)$, where $1 < p < \infty$. Assume, in addition, that*

$$\lim_{n\to\infty} \|f_n\|_{L^p(\mu)} = \|f\|_{L^p(\mu)}.$$

Then $\lim\limits_{n\to\infty} \|f_n - f\|_{L^p(\mu)} = 0$.

PROOF. If we have no norm convergence, then, passing to a subsequence, we may assume that $\|f - f_n\|_{L^p(\mu)} \geq \varepsilon > 0$. In addition, we may assume that the functions f_n have unit norms. By the uniform convexity of $L^p(\mu)$, there exists $\delta > 0$ such that $\|f_n + f\|_{L^p(\mu)} \leq 2(1-\delta)$ for all n. Let $q^{-1} + p^{-1} = 1$. There is $g \in L^q(\mu)$ with $\|g\|_{L^q(\mu)} = 1$ and

$$\int fg\,d\mu = 1.$$

Then

$$\int_X \frac{f_n + f}{2} g\,d\mu \to 1,$$

which leads to a contradiction, since by Hölder's inequality we obtain

$$\int_X \frac{f_n + f}{2} g\,d\mu \leq \Bigl\|\frac{f_n + f}{2}\Bigr\|_{L^p(\mu)} \leq 1 - \delta.$$

It is seen from the proof that the established property holds for all uniformly convex spaces. □

This corollary fails for $p = 1$ (Exercise 4.7.80).

4.7.17. Corollary. *For any $p \in (1, +\infty)$, the space $L^p(\mu)$ has the Banach–Saks property, i.e., every norm bounded sequence $\{f_n\}$ in $L^p(\mu)$ contains a subsequence $\{f_{n_k}\}$ such that the sequence $\dfrac{f_{n_1} + \cdots + f_{n_k}}{k}$ converges in the norm.*

PROOF. All uniformly convex Banach spaces have the Banach–Saks property: see Diestel [**222**, Ch. 3, §7]. □

The Banach–Saks property implies the reflexivity of a Banach space E by Theorem 4.7.8. Hence $L^1[0,1]$ does not have this property (which is also obvious from the consideration of $nI_{[0,1/n]}$). A partial compensation is given by Theorem 4.7.24.

4.7(iv). Uniform integrability and weak compactness in L^1

In this subsection, we consider only nonnegative measures on a measurable space $(X, \mathcal{A})$.

4.7.18. Theorem. *Let μ be a finite measure and let $\mathcal{F}$ be some set of μ-integrable functions. The set $\mathcal{F}$ is uniformly integrable precisely when it has compact closure in the weak topology of $L^1(\mu)$.*

PROOF. Let $\mathcal{F}$ be uniformly integrable. Then it is bounded in $L^1(\mu)$. Denote by $\mathcal{H}$ the closure of $\mathcal{F}$ in the space $\bigl(L^\infty(\mu)\bigr)^*$ equipped with the weak* topology $\sigma\bigl(L^\infty(\mu)^*, L^\infty(\mu)\bigr)$. By Theorem 4.7.13, the set $\mathcal{H}$ is compact. Since $L^1(\mu)$ is linearly isometric to a subspace in $L^\infty(\mu)^*$ (we recall that every Banach space E is isometric to a subspace in E^{**} under the natural embedding into this space, see the proof of Theorem 4.4.3), the topology $\sigma\bigl(L^\infty(\mu)^*, L^\infty(\mu)\bigr)$ induces on $L^1(\mu)$ the topology $\sigma\bigl(L^1(\mu), L^\infty(\mu)\bigr)$. Let us show that $\mathcal{H}\subset L^1(\mu)$. By construction, every element $F\in\mathcal{H}$ is a continuous linear functional on $L^\infty(\mu)$ that equals the limit of some net of functionals

$$F_\alpha(g)=\int_X f_\alpha g\,d\mu,\quad g\in L^\infty(\mu),$$

where $f_\alpha\in\mathcal{F}$, i.e., there is a partially ordered set Λ such that, for each $\alpha,\beta\in\Lambda$, there exists $\gamma\in\Lambda$ with $\alpha\le\gamma$ and $\beta\le\gamma$, and, for every $g\in L^\infty(\mu)$ and $\varepsilon>0$, there exists $\gamma\in\Lambda$ with $|F_\alpha(g)-F(g)|<\varepsilon$ for all $\alpha\ge\gamma$. The set $\mathcal{F}$ has uniformly absolutely continuous integrals (Proposition 4.5.3). Hence, for any $\varepsilon>0$, there exists $\delta>0$ such that

$$F(I_A)\le\limsup_\alpha F_\alpha(I_A)\le\limsup_\alpha\int_A|f_\alpha|\,d\mu<\varepsilon\quad\text{whenever } \mu(A)<\delta.$$

According to Proposition 4.4.2, the functional F is generated by a function $f\in L^1(\mu)$. Suppose that $\mathcal{F}$ has compact closure in the weak topology, but is not uniformly integrable. Then, there are $\eta>0$ and a sequence $\{f_n\}\subset\mathcal{F}$ such that

$$\int_{\{|f_n|>n\}}|f_n|\,d\mu\ge\eta$$

for all $n\ge1$. By the Eberlein–Šmulian theorem 4.7.10, the sequence $\{f_n\}$ contains a subsequence $\{f_{n_k}\}$ convergent to some function $f\in L^1(\mu)$ in the weak topology $\sigma(L^1,L^\infty)$. In particular, for every μ-measurable set A we have

$$\lim_{k\to\infty}\int_A f_{n_k}\,d\mu=\int_A f\,d\mu,$$

which leads to a contradiction with Theorem 4.5.6. □

4.7.19. Corollary. *Suppose that $\{f_n\}$ is a uniformly integrable sequence on a space with a finite measure μ. Then, there exists a subsequence f_{n_k} that converges in the weak topology of $L^1(\mu)$ to some function $f\in L^1(\mu)$, i.e., one has*

$$\lim_{n\to\infty}\int f_{n_k}g\,d\mu=\int fg\,d\mu,\quad\forall\, g\in L^\infty(\mu).$$

PROOF. As shown above, the sequence $\{f_n\}$ has compact closure in the weak topology. By the Eberlein–Šmulian theorem, it contains a weakly convergent subsequence.

Let us give an alternative reasoning that employs the weak compactness of balls in L^2. Set $f_{n,k}:=f_nI_{\{|f_n|\le k\}}$, $n,k\in\mathbb{N}$. For any fixed k, the sequence of functions $\{f_{n,k}\}$ is bounded in $L^2(\mu)$, hence contains a subsequence that is weakly convergent in $L^2(\mu)$. By the standard diagonal argument one can obtain a sequence $\{n_j\}$ such that, for every k, the functions $f_{n_j,k}$ converge weakly in $L^2(\mu)$ to some function $g_k \in L^2(\mu)$: one takes a subsequence $\{n_{1,j}\}$ for $k = 1$, a subsequence $\{n_{2,j}\} \subset \{n_{1,j}\}$ for $k = 2$ and so on; then one takes $n_j := n_{j,j}$. We observe that

$$\begin{aligned}\|g_k - g_l\|_{L^1(\mu)} &= \int (g_k - g_l)\operatorname{sign}(g_k - g_l)\, d\mu \\ &= \lim_{j\to\infty} \int (f_{n_j,k} - f_{n_j,l})\operatorname{sign}(g_k - g_l)\, d\mu \\ &\le \liminf_{j\to\infty} \|f_{n_j,k} - f_{n_j,l}\|_{L^1(\mu)} \to 0\end{aligned}$$

as $k, l \to \infty$ by the uniform integrability of $\{f_n\}$. Hence the functions g_k converge in $L^1(\mu)$ to some function f. The sequence $\{f_{n_j}\}$ converges to f weakly in $L^1(\mu)$. Indeed, for every bounded measurable function g and every $\varepsilon > 0$, there exists a number k such that $\|f_n - f_{n,k}\|_{L^1(\mu)} < \varepsilon\bigl(\sup|g(x)| + 1\bigr)^{-1}$ for all n (which is possible by the uniform integrability) and the integral of $|g(f - g_k)|$ does not exceed ε, next we find a number j_1 such that for all $j \ge j_1$ the integral of $|g(g_k - f_{n_j,k})|$ does not exceed ε. It remains to use the fact that the integral of $|g(f_{n_j,k} - f_{n_j})|$ does not exceed ε. □

4.7.20. Theorem. *Let $(X, \mathcal{A}, \mu)$ be a measure space, where the measure μ takes values in $[0, +\infty]$, and let $\mathcal{F} \subset L^1(\mu)$. The following conditions are equivalent*:

(i) *the closure of $\mathcal{F}$ in the weak topology of $L^1(\mu)$ is compact*;

(ii) *$\mathcal{F}$ is norm bounded and the measures $f\cdot\mu$, where $f \in \mathcal{F}$, are uniformly countably additive in the sense of Definition* 4.6.2;

(iii) *the closure of the set $\{|f|\colon\ f \in \mathcal{F}\}$ in the weak topology of $L^1(\mu)$ is compact*;

(iv) *$\mathcal{F}$ is norm bounded, the functions in $\mathcal{F}$ have uniformly absolutely continuous integrals and, for every $\varepsilon > 0$, there exists a measurable set X_ε such that $\mu(X_\varepsilon) < \infty$ and*

$$\int_{X\setminus X_\varepsilon} |f|\, d\mu < \varepsilon \quad \textit{for all } f \in \mathcal{F};$$

(v) *for every $\varepsilon > 0$, there exists a μ-integrable function g such that*

$$\int_{\{|f|>g\}} |f|\, d\mu \le \varepsilon \quad \textit{for all } f \in \mathcal{F}.$$

Proof. For bounded measures the equivalence of the listed conditions follows by Theorem 4.7.18, Proposition 4.5.3, and Lemma 4.6.5. It is clear from the Eberlein–Šmulian theorem and the definition of the uniform countable additivity that it suffices to prove the equivalence of (i)–(iii) for countable sets $\mathcal{F} = \{f_n\}$. Hence the general case reduces at once to the case where the measure μ is σ-finite because there exists a set $X_0 \in \mathcal{A}$ on which our measure is σ-finite and all functions f_n vanish outside X_0 (see Proposition 2.6.2). Next we find a finite measure μ_0 such that

$$\mu(A) = \int_A \varrho \, d\mu_0, \quad A \in \mathcal{A},$$

where $\varrho > 0$ is a measurable function. Now everything reduces to the finite measure μ_0 and the functions $g_n = f_n/\varrho$. Indeed, the sequence of functions $g_{n_k} \in L^1(\mu_0)$ weakly converges to g in $L^1(\mu_0)$ precisely when the sequence g_{n_k}/ϱ weakly converges to g/ϱ in $L^1(\mu)$. The situation with the absolute values of these functions is analogous. Condition (ii) for the functions f_n and the measure μ is equivalent to the same condition for the functions g_n and the measure μ_0. It is seen from the above reasoning that condition (iv) implies (i)–(iii) in the general case, too. We now verify that (iv) follows from (i)–(iii) for infinite measures. It is clear that due to the already-established facts for finite measures, we have only to verify the second condition in (iv). If it is not fulfilled, then, for some $\varepsilon > 0$, one can find a sequence of increasing measurable sets X_n and a sequence of functions $f_n \in \mathcal{F}$ such that $f_n = 0$ outside the set $Y = \bigcup_{n=1}^\infty X_n$, $\mu(X_n) > n$ and

$$\int_{X \backslash X_n} |f_n| \, d\mu \geq \varepsilon.$$

We consider the measures $\mu_n := f_n \cdot \mu$ and obtain a contradiction with Lemma 4.6.5. The equivalence of (v) to all other conditions follows from Exercise 4.7.82. □

In the case where a finite measure μ has no atoms, the norm boundedness of $\mathcal{F}$ in condition (iv) follows by the uniform absolute continuity (Proposition 4.5.3).

4.7.21. Corollary. *Let μ be a bounded nonnegative measure and let a set $M \subset L^1(\mu)$ be norm bounded. The closure of M in the weak topology is compact if and only if for every sequence of μ-measurable sets A_n such that $A_{n+1} \subset A_n$ and $\bigcap_{n=1}^\infty A_n = \varnothing$, one has*

$$\lim_{n\to\infty} \sup_{f \in M} \int_{A_n} |f| \, d\mu = 0.$$

Proof. This condition is necessary by condition (v) in the theorem. It is sufficient by condition (ii) and Lemma 4.6.5. □

If the measure μ is separable, then the weak topology on weakly compact sets in $L^1(\mu)$ is metrizable (Exercise 4.7.148).

Unlike the case $p \in (1, +\infty)$, in general, the spaces $L^1(\mu)$ do not have the property that any bounded sequence contains a weakly convergent subsequence (see Corollary 4.7.9 and Exercise 4.7.77). The next assertion gives partial compensation.

4.7.22. Lemma. *Let $(X, \mathcal{A}, \mu)$ be a space with a finite nonnegative measure, let $\{f_n\} \subset L^1(\mu)$, and let $\sup_n \|f_n\|_{L^1(\mu)} < \infty$. Then, for every $\varepsilon > 0$, one can find a measurable set E_ε, a number $\delta > 0$, and an infinite set $S \subset \mathbb{N}$ such that $\mu(E_\varepsilon) < \varepsilon$ and, for any set $A \subset X \backslash E_\varepsilon$ with $\mu(A) < \delta$, one has*

$$\int_A |f_k|\, d\mu < \varepsilon, \quad \forall\, k \in S.$$

PROOF. Suppose the contrary. Then, for some $\varepsilon > 0$, whatever is our choice of a set E with $\mu(E) < \varepsilon$, a number $\delta > 0$, and an infinite set $S \subset \mathbb{N}$, there exist $A \subset X \backslash E$ and $k \in S$ such that $\mu(A) < \delta$ and

$$\int_A |f_k|\, d\mu \geq \varepsilon.$$

Let us show that, for every set C with $\mu(C) < \varepsilon$ and every infinite part $S \subset \mathbb{N}$, there exist a set $A \subset X \backslash C$ and an infinite subset $T \subset S$ such that $\mu(A \cup C) < \varepsilon$ and

$$\int_A |f_k|\, d\mu \geq \varepsilon, \quad \forall\, k \in T.$$

To this end, we set $S_1 = S$ and take a positive number $\delta_1 < \bigl(\varepsilon - \mu(C)\bigr)/2$. Next we find $B_1 \subset X \backslash C$ with $\mu(B_1) < \delta_1$ and $k_1 \in S_1$ such that

$$\int_{B_1} |f_{k_1}|\, d\mu \geq \varepsilon.$$

We continue this process inductively so as $\delta_i \leq \delta_{i-1}/2$ and

$$S_i := \{k \in S_{i-1} \colon\ k > k_{i-1}\}.$$

Letting $A = \bigcup_{i=1}^\infty B_i$, $T = \{k_i\}$, we obtain the required objects.

By using the established fact we shall arrive at a contradiction. To this end, we describe one more inductive construction: let us apply the above fact to $C = \varnothing$ and $S = \mathbb{N}$. We obtain sets $A_1 \subset X$ and $T_1 \subset \mathbb{N}$ such that $\mu(A_1) < \varepsilon$ and

$$\int_{A_1} |f_k|\, d\mu \geq \varepsilon, \quad \forall\, k \in T_1.$$

Next we apply our auxiliary result to $C = A_1$ and $S = T_1$, which yields an infinite part $T_2 \subset T_1$ and a set $A_2 \subset X \backslash A_1$ such that $\mu(A_1 \cup A_2) < \varepsilon$ and

$$\int_{A_1 \cup A_2} |f_k|\, d\mu = \int_{A_1} |f_k|\, d\mu + \int_{A_2} |f_k|\, d\mu \geq 2\varepsilon, \quad \forall\, k \in T_2.$$

Next we deal with $C = A_1 \cup A_2$ and $S = T_2$. Let $N > \varepsilon^{-1} \sup_n \|f_n\|_{L^1(\mu)}$. After N steps we obtain disjoint sets $A_1, \ldots, A_N$ and a number k such that the integral of $|f_k|$ over $A_1 \cup \cdots \cup A_N$ is greater than $\|f_k\|_{L^1(\mu)}$, which is impossible. The possibility of continuation of our inductive construction is provided by the property that $\mu(A_1 \cup \cdots \cup A_n) < \varepsilon$ at all previous steps. $\square$

Now we are able to prove Gaposhkin's theorem on subsequences that converge "almost weakly in L^1".

4.7.23. Theorem. *Let μ be a finite nonnegative measure on a measurable space $(X, \mathcal{A}, \mu)$, let $\{f_n\} \subset \mathcal{L}^1(\mu)$, and let $\sup_n \|f_n\|_{L^1(\mu)} < \infty$. Then, one can find a subsequence $\{n_k\}$ and a function $f \in L^1(\mu)$ such that $\{f_{n_k}\}$ converges to f almost weakly in $L^1(\mu)$ in the following sense: for every $\varepsilon > 0$, there exists a measurable set X_ε such that $\mu(X \backslash X_\varepsilon) < \varepsilon$ and the functions $f_n|_{X_\varepsilon}$ converge to $f|_{X_\varepsilon}$ in the weak topology of the space $L^1(\mu|_{X_\varepsilon})$.*

Proof. We apply the above lemma to construct a subsequence f_{n_j} such that there exist sets Y_n with $\mu(X \backslash Y_n) < 2^{-n}$ on each of which the sequence $\{f_{n_j}\}$ has uniformly absolutely continuous integrals. Then it will contain a further subsequence that is weakly convergent in L^1 on every set Y_n. For Y_n we take the set $X \backslash \bigcup_{k=1}^\infty E_{\varepsilon(n,k)}$, where $\varepsilon(n, k) > 0$ is chosen as follows:

$$\varepsilon(n, k) = \min\bigl(2^{-k-n}, \delta(n, k-1)\bigr),$$

and the number $\delta(n, k-1)$ corresponds to $\varepsilon(n, k-1)$ according to the lemma, where $\varepsilon(n, 1) = 2^{-n}$. By the lemma we have an infinite part $\mathcal{F}_n \subset \{f_n\}$ with uniformly absolutely continuous integrals on Y_n. Moreover, it is clear from our reasoning that these parts can be chosen in such a way that we have $\mathcal{F}_{n+1} \subset \mathcal{F}_n$, whence one easily obtains the existence of a subsequence with uniformly absolutely continuous integrals on every Y_n. □

Let us consider one more remarkable property of bounded sequences in L^1, established by Komlós [**538**]. In Chapter 10, where the proof of the first part of the following theorem is given, some additional results can be found.

4.7.24. Theorem. *Let μ be a finite nonnegative measure on a space X, let $\{f_n\} \subset L^1(\mu)$, and let*

$$\sup_n \|f_n\|_{L^1(\mu)} < \infty.$$

Then, one can find a subsequence $\{g_n\} \subset \{f_n\}$ and a function $g \in L^1(\mu)$ such that, for every sequence $\{h_n\} \subset \{g_n\}$, the arithmetic means $(h_1 + \cdots + h_n)/n$ converge almost everywhere to g.

One can also obtain the following: for every $\varepsilon > 0$, there exists a set X_ε such that $\mu(X \backslash X_\varepsilon) < \varepsilon$ and the functions $(h_1 + \cdots + h_n)/n$ converge to g in the norm of $L^1(X_\varepsilon, \mu)$.

Proof. The most difficult part of Komlós's theorem is the existence of a subsequence with the arithmetic means of all subsequences convergent a.e. to some function $g \in L^1(\mu)$. This part will be proved in Chapter 10 (see §10.10) by using the techniques of conditional expectations discussed there. If this part is already known, then we apply it to the subsequence $\{f_{n_k}\}$, obtained in Theorem 4.7.23, that converges almost weakly in $L^1(\mu)$ to some function f. It is clear that the arithmetic means of any subsequence $\{h_n\}$ in $\{f_{n_k}\}$ converge in the same sense to the same limit f. It remains to observe that if these arithmetic means converge almost everywhere to a function g, then $f = g$ a.e. Indeed, the fact that the sequence of functions $n^{-1}(h_1 + \cdots + h_n)$ converges almost weakly in $L^1(\mu)$ yields that, given $\varepsilon > 0$, there exists a set X_ε such that $\mu(X\backslash X_\varepsilon) < \varepsilon$ and on X_ε this sequence is uniformly integrable. By the Lebesgue–Vitali theorem, it converges to g on X_ε in the norm of $L^1(X_\varepsilon, \mu)$, hence in the weak topology. Therefore, $f = g$ a.e. on X_ε, whence one has the equality $f = g$ a.e. on X. In addition, we obtain convergence in $L^1(X_\varepsilon, \mu)$. $\square$

4.7(v). The topology of setwise convergence of measures

Setwise convergence of measures, considered in Theorem 4.6.3, can be defined by means of a topology. Namely, this convergence is exactly convergence in the topology $\sigma(\mathcal{M}, \mathcal{F})$, where $\mathcal{M} = \mathcal{M}(X, \mathcal{A})$ is the space of all bounded countably additive measures on $\mathcal{A}$ and $\mathcal{F}$ is the linear space of all simple $\mathcal{A}$-measurable functions. A fundamental system of neighborhoods of a point μ_0 in this topology consists of all sets of the form

$$W_{A_1,\ldots,A_n,\varepsilon}(\mu_0) = \bigl\{\mu \in \mathcal{M}(X, \mathcal{A})\colon\ |\mu(A_i) - \mu_0(A_i)| < \varepsilon,\ i = 1, \ldots, n\bigr\},$$

where $A_i \in \mathcal{A}$ and $\varepsilon > 0$ (see §4.7(ii) about the definition of this topology). If the σ-algebra $\mathcal{A}$ is infinite, then the topology $\sigma(\mathcal{M}, \mathcal{F})$ is not generated by any norm (Exercise 4.7.115). One more natural topology on $\mathcal{M}$ is generated by the duality with the space $B(X, \mathcal{A})$ of bounded $\mathcal{A}$-measurable functions, i.e., this is the topology $\sigma\bigl(\mathcal{M}, B(X, \mathcal{A})\bigr)$. If the σ-algebra $\mathcal{A}$ is infinite, then this topology is strictly stronger than the topology $\sigma(\mathcal{M}, \mathcal{F})$. But, as it follows from Theorem 4.6.3, *for countable sequences* convergence in the topology $\sigma(\mathcal{M}, \mathcal{F})$ is equivalent to convergence in the topology $\sigma\bigl(\mathcal{M}, B(X, \mathcal{A})\bigr)$ (for the proof one should also use that every function in $B(X, \mathcal{A})$ is uniformly approximated by simple functions).

Finally, since $\mathcal{M}$ is a Banach space, one can consider the usual weak topology $\sigma(\mathcal{M}, \mathcal{M}^*)$ of a Banach space (see §4.7(ii)), which in nontrivial cases is strictly stronger than the topology $\sigma(\mathcal{M}, \mathcal{F})$, but is strictly weaker than the topology generated by the variation norm (Exercise 4.7.116). We shall now see that convergence of *countable sequences* in the topology $\sigma(\mathcal{M}, \mathcal{M}^*)$ is the same as in the topology of setwise convergence. In addition, both topologies possess the same families of compact sets.

4.7.25. Theorem. *For every set $M \subset \mathcal{M}(X, \mathcal{A})$ the following conditions are equivalent.* (i) *The set M has compact closure in the topology $\sigma(\mathcal{M}, \mathcal{M}^*)$.*

(ii) *The set M is bounded in the variation norm and there is a nonnegative measure $\nu \in \mathcal{M}(X, \mathcal{A})$ (a probability if $M \neq \{0\}$) such that the family M is uniformly ν-continuous, i.e., for every $\varepsilon > 0$, there is $\delta > 0$ with the property that*

$$|\mu(A)| \le \varepsilon \quad \textit{for all } \mu \in M \textit{ whenever } A \in \mathcal{A} \textit{ and } \nu(A) \le \delta.$$

In this case, all measures from M are absolutely continuous with respect to ν, the closure of the set $\{d\mu/d\nu\colon\ \mu \in M\}$ is compact in the weak topology of $L^1(\nu)$, and ν can be found in the form $\sum_{n=1}^\infty c_n|\mu_n|$ with some finite or countable collection $\{\mu_n\} \subset M$ and suitable numbers $c_n > 0$.

(iii) *The set M is bounded in the variation norm and uniformly countably additive.*

(iv) *The set M has compact closure in the topology of setwise convergence. This is also equivalent to the compactness of its closure in the topology of convergence on every bounded $\mathcal{A}$-measurable function.*

(v) *Every sequence in M has a subsequence convergent on every set in $\mathcal{A}$.*

PROOF. First we observe that, for every nonnegative measure ν on $\mathcal{A}$, the space $L^1(\nu)$ is embedded as a closed linear subspace in $\mathcal{M}(X,\mathcal{A})$ if we identify $f \in L^1(\nu)$ with the measure $f \cdot \nu$. With this identification, the topology $\sigma(\mathcal{M},\mathcal{M}^*)$ induces on $L^1(\nu)$ the topology $\sigma(L^1, L^\infty)$. This follows by the Hahn–Banach theorem (or by the fact that $(L^1(\nu))^* = L^\infty(\nu)$).

Let (i) be fulfilled. We show first that, for every $\varepsilon > 0$, there exist $\delta > 0$ and a finite collection $\mu_1, \ldots, \mu_n \in M$ such that

$$|\mu(A)| \le \varepsilon \quad \text{for all } \mu \in M \text{ whenever } A \in \mathcal{A} \text{ and } |\mu_i|(A) \le \delta \text{ for all } i \le n.$$

Suppose the contrary. Then by induction one can construct a sequence of measures μ_n in M and a sequence of sets A_n in $\mathcal{A}$ such that

$$|\mu_{n+1}(A_n)| \ge \varepsilon, \quad |\mu_i|(A_n) \le 2^{-n}, \quad \forall\, i \le n.$$

Let $\mu = \sum_{n=1}^\infty 2^{-n}\|\mu_n\|^{-1}|\mu_n|$. Then $\mu_n = f_n \cdot \mu$, $f_n \in L^1(\mu)$. It is clear by the remark made above that the sequence $\{f_n\}$ has compact closure in the weak topology $\sigma\bigl(L^1(\mu), L^\infty(\mu)\bigr)$. By Theorem 4.7.18 and Proposition 4.5.3 this sequence has uniformly absolutely continuous integrals, which leads to a contradiction, since $\mu(A_n) \le n2^{-n} + \sum_{i=n+1}^\infty 2^{-i} \to 0$ and $\mu_{n+1}(A_n) \ge \varepsilon$. Thus, our claim is proved.

Now, for every n, we find a number $\delta_n > 0$ and measures $\mu_1^n, \ldots, \mu_{k_n}^n$ corresponding to $\varepsilon = n^{-1}$. Let us take numbers $c_{n,j} > 0$ such that the measure $\nu = \sum_{n=1}^\infty \sum_{j=1}^{k_n} c_{n,j}|\mu_j^n|$ be a probability (if all the measures μ_j^n are zero, then M consists of the zero measure). Let $\varepsilon > 0$. Pick n such that $n^{-1} < \varepsilon$. There is $\delta > 0$ such that $|\mu_j^n|(A) \le \delta_n$ for all $j = 1, \ldots, k_n$, whenever $\nu(A) \le \delta$. Then, by our construction, $|\mu(A)| \le n^{-1} < \varepsilon$. Thus, one has (ii).

Let (ii) be fulfilled. If (iii) does not hold, then, for some ε, there exist increasing numbers n_k and measures $\mu_k \in M$ such that $\bigl|\sum_{j=n_k}^\infty \mu_k(A_j)\bigr| \ge \varepsilon$ for all k. Since $\mu_k = f_k \cdot \nu$, where $f_k \in L^1(\nu)$, we arrive at a contradiction with the fact that, according to Theorem 4.7.18 and Proposition 4.5.3, the functions f_k have uniformly absolutely continuous integrals.

Let (iii) be fulfilled. Let us show that every sequence $\{\mu_n\} \subset M$ contains a subsequence convergent in the topology $\sigma(\mathcal{M},\mathcal{M}^*)$. Then, by the Eberlein–Šmulian theorem, we obtain (i), which yields (iv), since the topology $\sigma(\mathcal{M},\mathcal{M}^*)$ is stronger than the topology of setwise convergence. Let us fix a nonnegative measure ν with $\mu_n = f_n \cdot \nu$, $f_n \in L^1(\nu)$. According to what has already been proven, it suffices to show that the measures μ_n are uniformly ν-continuous. But this follows at once by Lemma 4.6.5. Since the topology of convergence on bounded $\mathcal{A}$-measurable functions is weaker than $\sigma(\mathcal{M},\mathcal{M}^*)$, it has the same compact sets.

Let (iv) be given. The topology of setwise convergence and the topology of convergence on bounded $\mathcal{A}$-measurable functions coincide on M, since M is bounded in variation and every bounded $\mathcal{A}$-measurable function is uniformly approximated by simple functions. Suppose we are given a sequence $\{\mu_n\}$ in M. We take a probability measure ν on $\mathcal{A}$ such that $\mu_n = f_n \cdot \nu$, where $f_n \in L^1(\nu)$. Taking into account that any continuous linear functional on $L^1(\nu)$ is generated by a bounded $\mathcal{A}$-measurable function, we obtain that the set $\{f_n\}$ has compact closure in the weak topology of $L^1(\nu)$. By the Eberlein–Šmulian theorem this yields (v).

Finally, the implication (v)$\Rightarrow$(i) follows by the Eberlein–Šmulian theorem. Indeed, suppose we have a sequence of measures $\mu_n \in M$. As above, we take a measure $\nu \geq 0$ such that $\mu_n = f_n \cdot \nu$, $f_n \in L^1(\nu)$. It is clear that M is bounded in variation. Then, by (v), $\{f_n\}$ contains a subsequence that is weakly convergent in $L^1(\nu)$. It is seen from the observation made at the beginning of the proof that the corresponding subsequence of measures in $\{\mu_n\}$ converges in the topology $\sigma(\mathcal{M}, \mathcal{M}^*)$. □

One more condition of compactness in the topology of setwise convergence is given in Exercise 4.7.130.

4.7.26. Corollary. *A sequence of measures $\mu_n \in \mathcal{M}(X, \mathcal{A})$ converges in the topology $\sigma(\mathcal{M}, \mathcal{M}^*)$ precisely when it converges on every set in $\mathcal{A}$.*

We observe that if the measure ν in assertion (ii) of the above theorem has no atoms, then the boundedness of M in variation follows automatically by the uniform ν-continuity. Indeed, for every $\varepsilon > 0$, we find $\delta > 0$ such that $|\mu(E)| \leq \varepsilon$ if $\nu(E) < \delta$, $E \in \mathcal{A}$. It is clear that $|\mu|(E) \leq 2\varepsilon$, since $|\mu(E')| \leq \varepsilon$ for all $E' \subset E$, $E' \in \mathcal{A}$. It remains to observe that the whole space can be partitioned into finitely many parts with measures less than δ (see Theorem 1.12.9). Therefore, if all measures in M have no atoms, then in (ii) we need not require the boundedness in variation. In the general case this is not possible. For example, if X consists of the single point 0 and $\delta(0) = 1$, then the measures $n\delta$ are uniformly δ-continuous and uniformly countably additive, but are not uniformly bounded.

We recall once again that on more general sets of measures all three topologies considered in the above theorem are distinct.

In connection with the Vitali–Lebesgue–Hahn–Saks theorem and Lemma 4.6.5 one can naturally ask whether it would be enough to verify the required conditions only for sets in some algebra generating $\mathcal{A}$ in place of the whole $\mathcal{A}$. For example, dealing with a cube $\mathbb{R}^n$, for such an algebra it would be nice to take the algebra of elementary sets. Simple examples show that this may be impossible for some of the conditions that are equivalent in the case of a σ-algebra. More surprising is the following result, found by Areshkin [**33**] for nonnegative measures, extended by V.N. Aleksjuk to signed measures and given here with the proof borrowed from Areshkin, Aleksjuk, Klimkin [**34**].

Let $\mathfrak{R}$ be a ring of subsets in a space X and let $\mathfrak{S}$ be the generated σ-ring.

4.7.27. Theorem. *Suppose we are given a family of countably additive measures μ_α, $\alpha\in\Lambda$, of bounded variation on $\mathfrak{S}$. The following conditions are equivalent.*

(i) *The measures μ_α are uniformly additive on $\mathfrak{R}$ in the following sense: for every sequence of pairwise disjoint sets R_n in $\mathfrak{R}$, one has*

$$\lim_{n\to\infty}\sum_{k=n}^{\infty}\mu_\alpha(R_k)=0 \quad \textit{uniformly with respect to } \alpha\in\Lambda.$$

(ii) *For every sequence $\{\mu_{\alpha_n}\}\subset\{\mu_\alpha\}$ and every sequence of pairwise disjoint sets $R_n\in\mathfrak{R}$, one has*

$$\lim_{n\to\infty}\mu_{\alpha_n}(R_n)=0.$$

(iii) *The family $\{\mu_\alpha\}$ is equicontinuous on $\mathfrak{R}$ in the following sense: for every sequence of sets $R_n\in\mathfrak{R}$ with $R_{n+1}\subset R_n$ and $\bigcap_{n=1}^\infty R_n=\varnothing$, one has $\lim\limits_{n\to\infty}\mu_\alpha(R_n)=0$ uniformly in $\alpha\in\Lambda$.*

(iv) *Conditions* (i)–(iii) (*or any of these conditions*) *are fulfilled on $\mathfrak{S}$.*

Proof. The equivalence of conditions (i)–(iii) in the case where $\mathfrak{S}$ is a σ-algebra has already been established (see Lemma 4.6.5). The case of a σ-ring is analogous (in fact, this can be proven by elementary reasoning without any category arguments). In particular, the equivalence of (i) and (ii) for a ring is verified in Exercise 4.7.136. The equivalence of (i) and (iii) is obvious. We now show that (ii) yields (iv). Suppose that this is not the case. Say, let (ii) be false for $\mathfrak{S}$ in place of $\mathfrak{R}$. Then, there exist measures μ_n in the given family and disjoint sets $S_n\in\mathfrak{S}$ such that $|\mu_n|(S_n)\ge\varepsilon>0$. According to Exercise 4.7.137(ii), there exist sets $R_n\in\mathfrak{R}$ such that

$$|\mu_k|(S_n\bigtriangleup R_n)<\varepsilon 2^{-n}/4,\quad k\in\mathbb{N}. \tag{4.7.3}$$

Then $|\mu_n|(R_n)\ge 3\varepsilon/4$. Let $E_1=R_1$, $E_n=R_n\backslash\bigcup_{i=1}^{n-1}R_i$. The sets E_n are disjoint. For distinct k and j by the disjointness of S_k and S_j we have

$$R_k\cap R_j\subset(S_k\bigtriangleup R_k)\cup(S_j\bigtriangleup R_j),$$

whence

$$|\mu_n|(R_k\cap R_j)\le|\mu_n|(S_k\bigtriangleup R_k)+|\mu_n|(S_j\bigtriangleup R_j).$$

Hence $|\mu_n|\bigl(R_k\backslash(R_1\cup\cdots\cup R_{k-1})\bigr)<\varepsilon/2$, whence $|\mu_n|(E_k\bigtriangleup R_k)<\varepsilon/2$. Thus, $|\mu_n|(E_n)\ge\varepsilon/4$, which leads to a contradiction with (ii) for $\mathfrak{R}$. $\square$

4.7(vi). Norm compactness and approximations in L^p

Let $(X,\mathcal{A},\mu)$ be a space with a nonnegative measure (possibly with values in $[0,+\infty]$) and let Π be the set of all finite collections $\pi=\{E_1,\ldots,E_n\}$ of disjoint sets of finite positive measure. The set Π is partially ordered by the relation $\pi_1\le\pi_2$ defined as follows: every set in π_1 up to a measure zero set is a union of sets in π_2. For every $\pi_1,\pi_2\in\Pi$, there exists $\pi_3\in\Pi$ with $\pi_1\le\pi_3$, $\pi_2\le\pi_3$, i.e., Π is a directed set and one can consider nets of functions indexed by elements of Π. For any function f that is integrable on all sets of finite μ-measure we set

$$\mathbb{E}^\pi f(x) := \frac{1}{\mu(E_i)} \int_{E_i} f\,d\mu \quad \text{if } x \in E_i, \quad \mathbb{E}^\pi f(x) = 0 \text{ if } x \notin \textstyle\bigcup_{i=1}^n E_i.$$

It is clear that

$$\mathbb{E}^\pi f(x) = \sum_{j=1}^n \mu(E_j)^{-1} \biggl(\int_{E_j} f\,d\mu\biggr) I_{E_j}(x).$$

Note that in the case of a probability measure, $\mathbb{E}^\pi f$ is the conditional expectation of f with respect to the finite σ-algebra generated by the partition π (see Chapter 10 about this concept).

The following criterion of compactness is due to M. Riesz [**810**].

4.7.28. Theorem. *Let μ be a countably additive measure on a space X with values in $[0,+\infty]$ and let $1 \le p < \infty$. A set $K \subset L^p(\mu)$ has compact closure in the norm of $L^p(\mu)$ precisely when it is bounded and*

$$\lim_\pi \sup_{f\in K} \|\mathbb{E}^\pi f - f\|_{L^p(\mu)} = 0. \tag{4.7.4}$$

In particular, if the measure μ is finite and $F \subset L^1(\mu)$ is a bounded set, then F is norm compact in $L^p(\mu)$ if and only if, for every $\varepsilon > 0$, there exists a finite partition π of X into disjoint sets of positive measure such that, for every function $f \in F$, one has

$$\|f - \mathbb{E}^\pi f\|_{L^1(\mu)} \le \varepsilon. \tag{4.7.5}$$

PROOF. By Hölder's inequality we have

$$\biggl|\int_{E_j} f\,d\mu\biggr|^p \le \mu(E_j)^{p-1} \int_{E_j} |f|^p\,d\mu,$$

which yields that $\|\mathbb{E}^\pi f\|_{L^p(\mu)} \le \|f\|_{L^p(\mu)}$ for all $f \in L^p(\mu)$. For any simple integrable function f that is constant on disjoint sets $E_1, \ldots, E_n$, one has $\mathbb{E}^\pi f = f$ whenever $\pi \ge \pi_0$, $\pi_0 = \{E_1, \ldots, E_n\}$. The necessity of the above condition is easily derived from this. Indeed, if K has compact closure, then, given $\varepsilon > 0$, one can find functions $f_1, \ldots, f_m$ forming an $\varepsilon/4$-net in K, i.e., every point in K lies at a distance at most $\varepsilon/4$ from some of the points f_j. Next we find simple functions $\varphi_j \in L^p(\mu)$ with $\|f_j - \varphi_j\|_{L^p(\mu)} < \varepsilon/4$. Let us take a collection $\pi_0 = (A_1, \ldots, A_n) \in \Pi$ on the elements of which all functions φ_j are constant. Let $\pi \ge \pi_0$. For every $f \in K$, we find j with $\|f - \varphi_j\|_{L^p(\mu)} < \varepsilon/2$. On account of the equality $\mathbb{E}^\pi \varphi_j = \varphi_j$ we obtain

$$\begin{aligned} \|f - \mathbb{E}^\pi f\|_{L^p(\mu)} &\le \|f - \varphi_j\|_{L^p(\mu)} + \|\varphi_j - \mathbb{E}^\pi \varphi_j\|_{L^p(\mu)} \\ &\quad + \|\mathbb{E}^\pi \varphi_j - \mathbb{E}^\pi f\|_{L^p(\mu)} \le 2\|f - \varphi_j\|_{L^p(\mu)} < \varepsilon. \end{aligned}$$

It is clear that in the case where the measure μ is finite, one can take for π finite partitions of X into disjoint sets of positive measure. The sufficiency of the above conditions follows from the fact that $\mathbb{E}^\pi\bigl(L^p(\mu)\bigr)$ are finite-dimensional linear subspaces, hence their bounded subsets have compact closure. $\square$

The operators $\mathbb{E}^\pi$ constructed above are linear and continuous on $L^p(\mu)$ and have finite-dimensional ranges, on which they are the identity mappings. So it is appropriate to call them finite-dimensional projections (in the case $p = 2$ they are orthogonal projections). A useful property of such projections is that they provide simultaneous approximations by simple functions for all functions from a given compact, and not only approximations of every individual function as was the case in §4.2. Yet, these projections still depend on a given compact, but in the case of a separable $L^p(\mu)$ one can easily get rid of this dependence. Namely, assuming for simplicity that $\mu(X) < \infty$, let us take a countable family of measurable sets A_j such that finite linear combinations of their indicators are dense in $L^p(\mu)$ (which is possible due to the separability of $L^p(\mu)$). Now let us consider the partitions π_n generated by $A_1, \ldots, A_n$; the elements of π_n are disjoint finite intersections of the sets A_i, $i \le n$, and their complements. It is clear from the above proof that $\mathbb{E}^{\pi_n} f \to f$ uniformly in f from any compact set in $L^p(\mu)$. Another method of approximation in a separable space $L^p(\mu)$ employs Schauder bases. We recall that a Schauder basis in a Banach space Z is a sequence of vectors e_n such that, for every $x \in Z$, there exists a unique sequence of numbers x_n with $x = \lim\limits_{n\to\infty} \sum_{j=1}^n x_j e_j$. It is known that every separable $L^p(\mu)$ has a Schauder basis; this is clear from Corollary 9.12.27 in Chapter 9 on isomorphisms of the spaces L^p if we observe that in $l^p = L^p(\mathbb{N}, \nu)$, where $\nu(n) = 1$ for all n, a natural Schauder basis consists of the functions $h_n = I_{\{n\}}$, and in $L^p[0,1]$ a Schauder basis is formed by the Haar functions (Exercise 4.7.59).

Let $\mu \ge 0$ be a finite measure on a measurable space $(X, \mathcal{A})$, let $f \in L^1(\mu)$, and let A be a set of positive μ-measure. The quantity

$$\overline{\text{osc}} f|_A := \mu(A)^{-1} \int_A \left| f(x) - \mu(A)^{-1} \int_A f(y)\, \mu(dy) \right| \mu(dx)$$

is called the average oscillation of the function f on A.

4.7.29. Theorem. *Suppose that a set F in $L^1(\mu)$ has compact closure in the weak topology. Then, the closure of F is compact in the norm of $L^1(\mu)$ precisely when F satisfies the following condition* (G): *for every $\varepsilon > 0$ and every set A of positive μ-measure, there exists a finite collection of sets $A_1, \ldots, A_n \subset A$ of positive measure such that every function $f \in F$ has the average oscillation less than ε on at least one of the sets A_j.*

Proof. If the closure of F is norm compact, then it is weakly compact and (4.7.5) is fulfilled. It is clear that for any $f \in F$ estimate (4.7.5) yields that f has the average oscillation less than ε on at least one of the sets A_j.

Conversely, suppose that condition (G) is fulfilled. One can assume that μ is a probability measure. First we observe that, for every fixed function $h \in L^1(\mu)$, the set $F + h = \{f + h \colon f \in F\}$ satisfies condition (G) as well. Indeed, let $\varepsilon > 0$ and $\mu(A) > 0$. It is clear that there exists a set $B \subset A$ of positive measure such that the function h is uniformly bounded on B. Next we find a simple function g such that $\sup_{x\in X} |h(x)I_B(x) - g(x)| < \varepsilon/4$. The intersection of B with at least one of the finitely many sets on which g is constant is a set C of positive measure. Since F satisfies condition (G), there exists a finite collection of sets $C_j \subset C$ of positive measure such that every function $f \in F$ has the average oscillation less than $\varepsilon/2$ on at least one of these sets, say, C_m. It remains to observe that since g is constant on C_m and $|h(x) - g(x)| < \varepsilon/4$ on $C_m \subset B$, one has

$$\int_{C_m} \left|(f+h) - \int_{C_m} (f+h)\,d\mu\right| d\mu$$
$$\le \int_{C_m} \left|(f+g) - \int_{C_m} (f+g)\,d\mu\right| d\mu + \int_{C_m} \left|(h-g) - \int_{C_m} (h-g)\,d\mu\right| d\mu$$
$$\le \int_{C_m} \left|f - \int_{C_m} f\,d\mu\right| d\mu + 2\mu(C_m) \sup_{x\in C_m} |h(x) - g(x)| < \varepsilon\mu(C_m).$$

Suppose now that the closure of F is not norm compact. Then, there exists a weakly convergent sequence $\{f_n\} \subset F$ without norm convergent subsequences. According to what we have already proved, one can shift the set F and assume that $\{f_n\}$ weakly converges to 0. Moreover, passing to a subsequence, one may also assume that $\{|f_n|\}$ weakly converges to some function f. It is clear that $f \ge 0$ a.e. and $\alpha := \|f\|_{L^1(\mu)} > 0$ because otherwise we would obtain norm convergence. Let $\varepsilon := \alpha/4$ and $A = \{x\colon\ f(x) \ge 3\alpha/4\}$. Then $\mu(A) > 0$. Suppose now that $A_1, \ldots, A_k$ are arbitrary subsets of A of positive measure. We show that our sequence contains a function f_N whose average oscillation is greater than ε on every A_j. To this end, by using weak convergence of $\{f_n\}$ to 0 and weak convergence of $\{|f_n|\}$ to f, we pick N such that

$$\left|\int_{A_j} f_N\,d\mu\right| < \varepsilon\mu(A_j), \quad \left|\int_{A_j} f\,d\mu - \int_{A_j} |f_N|\,d\mu\right| < \varepsilon\mu(A_j), \quad j = 1, \ldots, k.$$

Then, for every A_j, we obtain

$$\mu(A_j)^{-1} \int_{A_j} \left|f_N - \mu(A_j)^{-1} \int_{A_j} f_N\,d\mu\right| d\mu$$
$$\ge \mu(A_j)^{-1} \int_{A_j} |f_N|\,d\mu - \mu(A_j)^{-1} \left|\int_{A_j} f_N\,d\mu\right|$$
$$\ge \mu(A_j)^{-1} \int_{A_j} f\,d\mu - \varepsilon - \varepsilon \ge \varepsilon,$$

since one has the inequality

$$\int_{A_j} f\,d\mu \ge 3\varepsilon\mu(A_j)$$

due to the estimate $f \ge 3\varepsilon$ on $A_j \subset A$. Thus, we arrive at a contradiction with condition (G). $\square$

Exercise 4.7.129 gives a compactness criterion for the space $L^0(\mu)$ of all measurable functions with the topology of convergence in measure.

4.7(vii). Certain conditions of convergence in L^p

We shall prove several useful results linking diverse modes of convergence in L^p. A result of this type has already been given in Corollary 4.7.16. The next one is taken from Brézis, Lieb [**127**].

4.7.30. Proposition. *Let μ be a measure with values in $[0,+\infty]$. Suppose that a sequence $\{f_n\} \subset \mathcal{L}^p(\mu)$, where $0 < p < \infty$ converges almost everywhere to a function f and $\sup_n \|f_n\|_{L^p(\mu)} < \infty$. Then*

$$\lim_{n\to\infty} \big\| |f_n|^p - |f_n - f|^p - |f|^p \big\|_{L^1(\mu)} = 0, \tag{4.7.6}$$

$$\lim_{n\to\infty} \Big(\|f_n\|^p_{L^p(\mu)} - \|f_n - f\|^p_{L^p(\mu)} \Big) = \|f\|^p_{L^p(\mu)}. \tag{4.7.7}$$

If, in addition, $\|f_n\|_{L^p(\mu)} \to \|f\|_{L^p(\mu)}$, then $\|f_n - f\|_{L^p(\mu)} \to 0$.

PROOF. It is easily verified that, for every $\varepsilon > 0$, there exists a number $C(p,\varepsilon) > 0$ such that

$$\big| |a+b|^p - |a|^p \big| \le \varepsilon |a|^p + C(p,\varepsilon)|b|^p, \quad \forall\, a,b \in \mathbb{R}. \tag{4.7.8}$$

Set $g_{n,\varepsilon} = \max\Big(\big||f_n|^p - |f_n - f|^p - |f|^p\big| - \varepsilon|f_n - f|^p, 0\Big)$. Then $\lim\limits_{n\to\infty} g_{n,\varepsilon}(x) = 0$ a.e. By (4.7.8) with $a = f_n - f$ and $b = f$ we have

$$\begin{aligned} g_{n,\varepsilon} &\le \max\Big(\big||f_n|^p - |f - f_n|^p\big| + |f|^p - \varepsilon|f_n - f|^p, 0\Big) \\ &\le \max\Big(\varepsilon|f_n - f|^p + C(p,\varepsilon)|f|^p + |f|^p - \varepsilon|f_n - f|^p, 0\Big) \\ &\le [C(p,\varepsilon)+1]|f|^p. \end{aligned}$$

By the dominated convergence theorem we obtain that, for every fixed $\varepsilon > 0$, the integrals of $g_{n,\varepsilon}$ converge to zero as $n \to \infty$. Therefore, there exists N such that $\|g_{n,\varepsilon}\|_{L^1(\mu)} \le \varepsilon$ for all $n \ge N$. Then, as one can easily verify, for all $n \ge N$, we have

$$\int \big||f_n|^p - |f_n - f|^p - |f|^p\big|\, d\mu \le \varepsilon \|f_n - f\|^p_{L^p(\mu)} + \varepsilon.$$

By the uniform boundedness of $\|f_n\|_{L^p(\mu)}$ we obtain convergence of the sequence of functions $|f_n|^p - |f_n - f|^p - |f|^p$ to zero in $L^1(\mu)$, which yields convergence of their integrals to zero. □

4.7.31. Proposition. *Let μ be a probability measure and let*

$$\{\xi_n\} \subset \mathcal{L}^1(\mu), \quad \|\xi_n\|_{L^1(\mu)} \le C, \quad \forall\, n \in \mathbb{N}.$$

Suppose that, for every fixed integer $k \ge 0$, the functions

$$\xi_{n,k}(x) := \xi_n(x) I_{[-k,k]}\big(\xi_n(x)\big)$$

weakly converge in $L^2(\mu)$ to a function η_k as $n \to \infty$. Then, there exists a function $\eta \in \mathcal{L}^1(\mu)$ such that

$$\lim_{k\to\infty} \eta_k(x) = \eta(x) \text{ a.e. and } \lim_{k\to\infty} \|\eta_k - \eta\|_{L^1(\mu)} = 0.$$

PROOF. Let $\eta_0 = 0$ and $\zeta_n := \eta_n - \eta_{n-1}$. Then $\eta_n = \sum_{k=1}^n \zeta_k$. We show that

$$\sum_{k=1}^{\infty} \|\zeta_k\|_{L^1(\mu)} \le C + 1. \tag{4.7.9}$$

By Fatou's theorem, this yields a.e. convergence of the series $\sum_{k=1}^{\infty} |\zeta_k(x)|$ which gives a.e. convergence of the sequence $\eta_n(x)$. In addition, convergence of the series $\sum_{k=1}^{\infty} |\zeta_k|$ in $L^1(\mu)$ shows that the sequence $\{\eta_n\}$ is fundamental in $L^1(\mu)$ and hence converges in $L^1(\mu)$ to the same function to which it converges almost everywhere. For the proof of (4.7.9) it suffices to obtain the estimate

$$\sum_{k=1}^{\infty} \liminf_{n\to\infty} \|\xi_{n,k} - \xi_{n,k-1}\|_{L^1(\mu)} \le C + 1, \tag{4.7.10}$$

since the general term of the series in (4.7.9) is majorized by the general term of the series in (4.7.10) due to Exercise 4.7.85 and the fact that the functions $\xi_{n,k} - \xi_{n,k-1}$ weakly converge to $\eta_k - \eta_{k-1} = \zeta_k$ as $n \to \infty$. Let us fix $N \in \mathbb{N}$. It is clear that there exists $m = m(N) \in \mathbb{N}$ such that

$$\sum_{k=1}^{N} \liminf_{n\to\infty} \|\xi_{n,k} - \xi_{n,k-1}\|_{L^1(\mu)} \le \sum_{k=1}^{N} \|\xi_{m,k} - \xi_{m,k-1}\|_{L^1(\mu)} + 1. \tag{4.7.11}$$

The right-hand side of (4.7.11) is majorized by $\|\xi_m\|_{L^1(\mu)} + 1$. Indeed, we have $|\xi_m(x)| = \sum_{k=1}^{\infty} |\xi_{m,k}(x) - \xi_{m,k-1}(x)|$, since whenever $|\xi_m(x)| > 0$, there exists an integer number $k = k(x) \ge 0$ such that $k < |\xi_m(x)| \le k+1$, which yields $\xi_{m,j}(x) = 0$ for all $j \le k$ and $\xi_{m,j}(x) = \xi_m(x)$ for all $j \ge k+1$. □

The proof of the following result can be found in Saadoune, Valadier [**837**]. It is instructive to compare it with Theorem 4.7.23.

4.7.32. Theorem. *Let μ be a probability measure on a space $(X, \mathcal{A})$ and let $\{f_n\}$ be a sequence of μ-measurable functions. Then, there exist a subsequence $\{f_{n_k}\}$ and a measurable set E such that $\{f_{n_k}\}$ converges in measure on E, but, for every set $A \subset X\backslash E$ of positive measure, $\{f_{n_k}\}$ contains no sequences convergent in measure on A.*

The next result is obtained in Visintin [**978**].

4.7.33. Theorem. *Let μ be a σ-finite measure on a space X and let a sequence $\{f_n\}$ converge to f in the weak topology of $L^1(\mu)$. If, for a.e. x, the point $f(x)$ is extreme in the closed convex envelope of the sequence $\{f_n(x)\}$, then $\lim\limits_{n\to\infty} \|f - f_n\|_{L^1(\mu)} = 0$.*

4.7(viii). Hellinger's integral and Hellinger's distance

Let μ and ν be two probability measures on a space $(X, \mathcal{A})$. Let us take some finite or σ-finite nonnegative measure λ on $\mathcal{A}$ such that $\mu \ll \lambda$ and $\nu \ll \lambda$. For example, one can take $\lambda = \mu + \nu$ or $\lambda = (\mu + \nu)/2$.

4.7.34. Definition. *Let $\alpha \in (0,1)$. Hellinger's integral of the order α of the pair of measures μ and ν is the quantity*

$$H_\alpha(\mu,\nu) := \int_X \Bigl(\frac{d\mu}{d\lambda}\Bigr)^\alpha \Bigl(\frac{d\nu}{d\lambda}\Bigr)^{1-\alpha} d\lambda.$$

4.7.35. Lemma. *The quantity $H_\alpha(\mu,\nu)$ is independent of our choice of a measure λ with respect to which μ and ν are absolutely continuous. In addition, one has*

$$0 \le H_\alpha(\mu,\nu) = H_{1-\alpha}(\nu,\mu) \le 1. \tag{4.7.12}$$

PROOF. The estimate $H_\alpha(\mu,\nu) \le 1$ follows by Hölder's inequality:

$$H_\alpha(\mu,\nu) \le \biggl(\int_X \frac{d\mu}{d\lambda}\, d\lambda\biggr)^\alpha \biggl(\int_X \frac{d\nu}{d\lambda}\, d\lambda\biggr)^{1-\alpha} = 1.$$

The equality in (4.7.12) is obvious from the definition. Let us consider the measure $\lambda_0 = \mu+\nu$. Then $\lambda_0 \ll \lambda$ for any measure λ, with respect to which μ and ν are absolutely continuous. Therefore, $d\mu/d\lambda = (d\mu/d\lambda_0)(d\lambda_0/d\lambda)$, $d\nu/d\lambda = (d\nu/d\lambda_0)(d\lambda_0/d\lambda)$. Hence one has

$$\int_X \Bigl(\frac{d\mu}{d\lambda}\Bigr)^\alpha \Bigl(\frac{d\nu}{d\lambda}\Bigr)^{1-\alpha} d\lambda = \int_X \Bigl(\frac{d\mu}{d\lambda_0}\Bigr)^\alpha \Bigl(\frac{d\nu}{d\lambda_0}\Bigr)^{1-\alpha} \frac{d\lambda_0}{d\lambda}\, d\lambda,$$

which proves that Hellinger's integral is independent of our choice of λ. $\square$

We observe that if $\mu = \mu_0 + \mu'$, where $\mu_0 \ll \nu$ and $\mu' \perp \nu$, then letting $\lambda = \nu + \mu'$, we obtain

$$H_\alpha(\mu,\nu) = \int_X \Bigl(\frac{d\mu_0}{d\nu}\Bigr)^\alpha d\nu.$$

Hellinger's integral of the order $1/2$ is most frequently used. Let us set $H(\mu,\nu) := H_{1/2}(\mu,\nu)$. It is clear that $H(\mu,\nu) = H(\nu,\mu)$. Let

$$r_2(\mu,\nu) := \Bigl(1 - H(\mu,\nu)\Bigr)^{1/2}. \tag{4.7.13}$$

By using a measure λ with respect to which μ and ν are absolutely continuous, one can write

$$r_2(\mu,\nu)^2 = \frac{1}{2}\int_X \Bigl|\sqrt{d\mu/d\lambda} - \sqrt{d\mu/d\lambda}\Bigr|^2 d\lambda. \tag{4.7.14}$$

4.7.36. Lemma. *The function r_2 given by equality* (4.7.13) (*or* (4.7.14)) *is a metric on the set of all probability measures on $\mathcal{A}$.*

PROOF. The equality $r_2(\mu,\nu) = r_2(\nu,\mu)$ is obvious. If $r_2(\mu,\nu) = 0$, then, letting $\lambda = \mu+\nu$, we observe that the inner product of the functions $\sqrt{d\mu/d\lambda}$ and $\sqrt{d\nu/d\lambda}$ in $L^2(\lambda)$ equals 1. Since these functions have unit norms, they are proportional, whence it follows that they coincide λ-almost everywhere. Hence $\mu = \nu$. The triangle inequality for r_2 follows by the triangle inequality in $L^2(\lambda)$ taking into account the fact that for any three measures μ, ν, and η one can find a common dominating measure λ (for example, their sum). $\square$

The metric r_2 is called Hellinger's distance (metric). As we shall now see, Hellinger's integral is connected with the variation distance.

4.7.37. Theorem. *For arbitrary probability measures μ and ν on $(X,\mathcal{A})$ the following inequalities are true*:

$$2[1-H(\mu,\nu)] \le \|\mu-\nu\| \le 2\sqrt{1-H(\mu,\nu)^2}, \tag{4.7.15}$$

$$2r_2^2(\mu,\nu) \le \|\mu-\nu\| \le \sqrt{8}r_2(\mu,\nu), \tag{4.7.16}$$

$$2[1-H_\alpha(\mu,\nu)] \le \|\mu-\nu\| \le c_\alpha\sqrt{1-H_\alpha(\mu,\nu)}, \quad \alpha\in(0,1). \tag{4.7.17}$$

PROOF. Inequality (4.7.16) follows from (4.7.15) by definition and the estimate $1+H(\mu,\nu)\le 2$. Let $f=d\mu/d\lambda$, $g=d\nu/d\lambda$, where $\lambda=\mu+\nu$. For the proof of the first inequality in (4.7.15), it suffices to sum the inequality

$$1-H(\mu,\nu)=\int_X \sqrt{f}(\sqrt{f}-\sqrt{g})\,d\lambda \le \int_{\{f\ge g\}} |f-g|\,d\lambda$$

and the symmetric inequality

$$1-H(\mu,\nu)\le \int_{\{f\le g\}} |g-f|\,d\lambda.$$

The same reasoning proves the first inequality in (4.7.17). The second inequality in (4.7.15) is deduced from the Cauchy–Bunyakowsky inequality (see (2.11.3)) as follows:

$$\begin{aligned}\int_X |f-g|\,d\lambda &= \int_X |\sqrt{f}-\sqrt{g}|(\sqrt{f}+\sqrt{g})\,d\lambda\\ &\le \biggl(2-2\int_X \sqrt{fg}\,d\lambda\biggr)^{1/2}\biggl(2+2\int_X \sqrt{fg}\,d\lambda\biggr)^{1/2}.\end{aligned}$$

In order to obtain the second inequality in (4.7.17), we observe that, for any $\alpha\in(0,1/2)$, one can take $p=p(\alpha)=(2\alpha)^{-1}>1$ and then $k_\alpha>0$ such that $1-s^{1/p}\ge k_\alpha(1-s)$ for all $s\in[0,1]$. Then by Hölder's inequality applied to the measure $g\cdot\lambda$, on account of the equality $p\alpha=1/2$ we obtain

$$\int_X f^\alpha g^{1-\alpha}\,d\lambda \le \biggl(\int_X f^{1/2}g^{1/2}\,d\lambda\biggr)^{1/p},$$

whence

$$1-\int_X f^\alpha g^{1-\alpha}\,d\lambda \ge k_\alpha\biggl(1-\int_X f^{1/2}g^{1/2}\,d\lambda\biggr).$$

Due to (4.7.15) this leads to (4.7.17) with $c_\alpha=\sqrt{8}k_\alpha$. □

Hellinger's integral $H_\alpha(\mu,\nu)$ can also be considered for $\alpha>1$, however, this expression may be infinite. The case where it is finite for $\alpha=2$ was considered by Hellinger [**420**], which became a starting point of the study of the concepts in this subsection. An abstract definition of Hellinger's integral for $\alpha=2$ is this. Let a measure ν on a space $(X,\mathcal{A})$ be absolutely continuous with respect to a probability measure μ on $(X,\mathcal{A})$ and let $f=d\nu/d\mu$. The supremum of the sums $\sum_{k=1}^n \nu(A_k)^2/\mu(A_k)$ over all finite partitions of the space into disjoint measurable sets of positive μ-measure is called Hellinger's integral and denoted by

$$\int \frac{\nu^2(dx)}{\mu(dx)}.$$

This quantity is finite if and only if $f \in L^2(\mu)$ and in that case it coincides with $\|f\|^2_{L^2(\mu)}$ (see Exercise 4.7.102). According to the same exercise, the membership of f in $L^p(\mu)$ with some $p > 1$ is characterized by the boundedness of analogous sums $\sum_{k=1}^n \nu(A_k)^p \mu(A_k)^{1-p}$.

Finally, let us point out a relation between Hellinger's distance $H(\mu, \nu)$ and Kullback's divergence defined by the following formula in the case of equivalent probability measures μ and ν:

$$K(\mu, \nu) := \int_X \ln(d\mu/d\nu)\, d\mu = \int_X \frac{d\mu}{d\lambda} \ln \frac{d\mu/d\lambda}{d\nu/d\lambda}\, d\lambda.$$

Here, as above, λ is an arbitrary probability measure with $\mu \ll \lambda$ and $\nu \ll \lambda$, for example, $\lambda = (\mu + \nu)/2$; it is easily seen that the corresponding expression is independent of our choice of λ, so that one can also take $\lambda = \nu$, which shows that $K(\mu, \nu)$ equals the entropy of $d\mu/d\nu$ with respect to the measure ν. According to (2.12.23) we have $K(\mu, \nu) \geq 0$, where $K(\mu, \nu)$ may be infinite. We observe that $K(\mu, \nu)$ may not be symmetric.

4.7.38. Proposition. *For any equivalent probability measures μ and ν we have*

$$r_2(\mu, \nu)^2 \leq K(\mu, \nu) \quad \textit{and} \quad \|\mu - \nu\|^2 \leq 2K(\mu, \nu).$$

PROOF. Let $f = d\nu/d\mu$. Since $\ln(1 + x) \leq x$, one has the estimate $\ln f = 2 \ln(1 + \sqrt{f} - 1) \leq 2(\sqrt{f} - 1)$, i.e., $\ln f^{-1} \geq 2 - 2\sqrt{f}$, which gives the first inequality after integrating with respect to the measure μ. The second one follows by Theorem 2.12.24 (with the constant 4 in place of 2 it follows from the first inequality). □

4.7(ix). Additive set functions

Let $\mathcal{A}$ be a σ-algebra of subsets in a space X and let $ba(\mathcal{A})$ be the space of all finitely additive bounded functions $m\colon \mathcal{A} \to \mathbb{R}^1$ equipped with the norm $\|m\|_1 := |m|(X)$, where for every $A \subset \mathcal{A}$ we set

$$|m|(A) := \sup\Bigl\{\sum_{i=1}^n |m(A_i)|\Bigr\},$$

where sup is taken over all finite partitions of A into disjoint sets $A_i \in \mathcal{A}$. It is readily verified that $ba(X, \mathcal{A})$ is a Banach space with the norm $\|\cdot\|_1$. Let $B(X, \mathcal{A})$ be the space of all $\mathcal{A}$-measurable bounded functions with the norm $\|f\|_\infty := \sup_{x \in X} |f(x)|$. The integral of a function $f \in B(X, \mathcal{A})$ with respect to the set function $m \in ba(X, \mathcal{A})$ is defined as follows: for a simple function $f = \sum_{i=1}^{n} c_i I_{A_i}$, where the sets A_i are disjoint, we set

$$\int_X f\, dm := \sum_{i=1}^{n} c_i m(A_i).$$

This integral is linear and is estimated in the absolute value by $\|f\|_\infty \|m\|_1$. Now the integral extends by continuity to all functions $f \in B(X, \mathcal{A})$ with the preservation of the indicated estimate and linearity. Simple details of verification along with the proof of the following assertion are left to the reader as Exercise 4.7.121.

4.7.39. Proposition. *The space $ba(X, \mathcal{A})$ can be identified with the dual space to $B(X, \mathcal{A})$ by the mapping $m \mapsto l_m$, where*

$$l_m(f) = \int_X f\, dm \quad \textit{and} \quad \|m\|_1 = \|l_m\|.$$

Let us mention the following lemma due to Rosenthal [**824**]; its proof is delegated to Exercise 4.7.122.

4.7.40. Lemma. *Let $\{m_n\} \subset ba(X, \mathcal{A})$ be a uniformly bounded sequence. Then, for every $\varepsilon > 0$ and every sequence of disjoint sets $A_i \subset \mathcal{A}$, there exists a sequence of indices k_n such that $|m_{k_n}|\bigl(\bigcup_{j \neq n} A_{k_j}\bigr) < \varepsilon$ for all n.*

Finally, let us mention the Phillips lemma [**752**] (Exercise 4.7.123).

4.7.41. Lemma. *Let $\mathcal{A}$ be the σ-algebra of all subsets in $\mathbb{N}$ and let $\{m_n\} \subset ba(\mathbb{N}, \mathcal{A})$ be such that $\lim\limits_{n\to\infty} m_n(A) = 0$ for all $A \subset \mathbb{N}$. Then*

$$\lim_{n\to\infty} \sum_{j=1}^{\infty} |m_n(\{j\})| = 0,$$

where $\{j\}$ is the set consisting of a single element j.

Exercises

4.7.42.° Let $f \in L^p(\mathbb{R}^1)$ and $f \in L^q(\mathbb{R}^1)$, where $p \le q$. Prove that $f \in L^r(\mathbb{R}^1)$ for all $r \in [p, q]$.

HINT: consider the sets $\{|f| \le 1\}$ and $\{|f| \ge 1\}$.

4.7.43.° Let f be a bounded measurable function on a space with a nonnegative measure μ. Prove that $\|f\|_{L^\infty(\mu)} = \inf\bigl\{a \ge 0\colon\ \mu\bigl(x\colon\ |f(x)| > a\bigr) = 0\bigr\}$.

4.7.44.° Show that $\|f\|_{L^\infty(\mu)} = \lim\limits_{p\to\infty} \|f\|_{L^p(\mu)}$ if the measure μ is bounded and $f \in L^\infty(\mu)$.

HINT: verify the assertion for simple functions, approximate f uniformly by a sequence of simple functions f_j and observe that $\|f - f_j\|_{L^p(\mu)}$ is majorized by $\|1\|_{L^p(\mu)} \|f - f_j\|_{L^\infty(\mu)}$.

4.7.45.° Let μ be a probability measure and let f be a measurable function such that $\sup_{p\ge 1}\|f\|_{L^p(\mu)}<\infty$. Prove that $f\in L^\infty(\mu)$.

HINT: use Chebyshev's inequality.

4.7.46.° Let $A\subset \mathbb{R}^1$ be a set of positive Lebesgue measure. Prove that the spaces L^p on the set A equipped with Lebesgue measure are infinite-dimensional.

HINT: construct a countable sequence of pairwise disjoint intervals whose intersections with A have positive measures.

4.7.47.° Prove the formula for the Legendre polynomials in Example 4.3.7.

4.7.48.° Prove that the functions $\sqrt{2/\pi}\sin nt$, $n\in\mathbb{N}$, form an orthonormal basis in $L^2[0,\pi]$. Prove the same for the functions $\sqrt{1/\pi}$, $\sqrt{2/\pi}\cos nt$, $n\in\mathbb{N}$.

HINT: it is verified directly that both systems are orthonormal. If the first system is not complete, then there is a nontrivial function $g\in L^2[0,\pi]$ orthogonal to it. Let $h(t)=g(t)$ if $t\in[0,\pi]$, $h(t)=-g(-t)$ if $t\in[-\pi,0]$. Then h is orthogonal to all $\sin nt$ in $L^2[-\pi,\pi]$. Since h is an odd function, one has $h\perp\cos nt$ for all $n=0,1,\ldots$, hence $h=0$ a.e.

4.7.49.° Let μ be the measure on $(0,+\infty)$ with density e^{-x} with respect to Lebesgue measure. Prove that the Lagguere polynomials obtained by the orthogonalization of the functions $1,x,x^2,\ldots$, form an orthonormal basis in $L^2(\mu)$.

HINT: if $g\in L^2(\mu)$ and $c>1/2$, then the function $g(x)\exp(-cx)$ is μ-integrable, which yields the analyticity of the Fourier transform of $g(x)e^{-x}$ in a strip.

4.7.50.° (i) Let $(X,\mathcal{A},\mu)$ and $(Y,\mathcal{B},\nu)$ be two probability spaces. Suppose that for some $p\in[1,+\infty)$ sets $F\subset L^p(\mu)$ and $G\subset L^p(\nu)$ are everywhere dense. Show that the set of linear combinations of products fg, where $f\in F$, $g\in G$, is everywhere dense in $L^p(\mu\otimes\nu)$. Prove that if $\{f_n\}$ and $\{g_n\}$ are orthonormal bases in $L^2(\mu)$ and $L^2(\nu)$, respectively, then $\{f_ng_k\}$ is an orthonormal basis in $L^2(\mu\otimes\nu)$.

(ii) Let $(X_\alpha,\mathcal{A}_\alpha,\mu_\alpha)$ be a family of probability spaces. Suppose that for some $p\in[1,+\infty)$ and every α, we are given an everywhere dense set $F_\alpha\subset L^p(\mu_\alpha)$. Show that the set of linear combinations of products $f_{\alpha_1}\cdots f_{\alpha_n}$, where $f_{\alpha_i}\in F_{\alpha_i}$, is everywhere dense in $L^p(\bigotimes_\alpha\mu_\alpha)$. Deduce that if, for every α, we have an orthonormal basis $\{f_{\alpha,\beta}\}$ in $L^2(\mu_\alpha)$, then the elements $f_{\alpha_1,\beta_1}\cdots f_{\alpha_n,\beta_n}$, where the indices α_i are distinct, form an orthonormal basis in $L^2(\bigotimes_\alpha\mu_\alpha)$.

HINT: (i) observe that the set of simple functions is dense in $L^p(\mu\otimes\nu)$, hence the set of linear combinations of indicators of measurable rectangles is dense as well. Given $A\in\mathcal{A}$ and $B\in\mathcal{B}$, we can find sequences $\{f_n\}\subset F$ and $\{g_n\}\subset G$ convergent to I_A and I_B in the corresponding L^p-norms. It follows by the equality $f_ng_n-I_AI_B=(f_n-I_A)g_n+I_A(g_n-I_B)$ and Fubini's theorem that $f_ng_n\to I_AI_B$ in $L^p(\mu\otimes\nu)$. Applying this assertion in the case $p=2$ and noting that the elements f_ng_k have unit norms and are mutually orthogonal, we obtain the second claim. The reasoning in (ii) is much the same.

4.7.51.° Prove that if a series is Cesàro summable to a number s, then it is summable to s in the sense of Abel (see §4.3).

4.7.52. Let $\{\varphi_n\}$ be an orthonormal basis in $L^2[0,1]$.

(i) Prove that there exist numbers c_n, $n\ge 2$, such that the sums $\sum_{n=2}^N c_n\varphi_n(x)$ converge to φ_1 in measure.

(ii) Prove that, for every $\varepsilon > 0$, there exists a set E_ε with measure greater than $1-\varepsilon$ such that the linear span of the functions φ_n, $n \geq 2$, is everywhere dense in $L^2(E_\varepsilon)$, where E_ε is equipped with Lebesgue measure.

(iii) Prove that there exists a positive bounded measurable function θ such that the linear span of the functions $\theta\varphi_n$, $n \geq 2$, is everywhere dense in the space $L^2[0,1]$.

HINT: (i) it suffices to show that, for every fixed k, the set of finite linear combinations of the functions φ_n, $n \geq k$, is everywhere dense in the space $L^0[0,1]$ with the metric defining convergence in measure. Otherwise $L^0[0,1]$ would contain a linear subspace of finite codimension closed in the indicated metric, which is impossible by Exercise 4.7.61. (ii) Applying (i) and the Riesz and Egoroff theorems one can find a set E_ε with measure greater than $1-\varepsilon$ on which φ_1 is the uniform limit of a sequence of finite linear combinations of the functions φ_n, $n \geq 2$. Then E_ε is the required set, since otherwise one could find a function $g \in L^2(E_\varepsilon)$ with

$$\int_{E_\varepsilon} g\varphi_n \, dx = 0$$

for all $n \geq 2$. Since φ_1 on E_ε is the uniform limit of linear combinations of φ_n, $n \geq 2$, we obtain

$$\int_{E_\varepsilon} g\varphi_1 \, dx = 0,$$

i.e., letting $g = 0$ outside E_ε we obtain a function that is orthogonal to all φ_n, whence $g = 0$ a.e. (iii) There is a positive bounded function θ such that the function φ_1/θ does not belong to $L^2[0,1]$. If we had a function $g \in L^2[0,1]$ orthogonal to all $\theta\varphi_n$, $n \geq 2$, then we would obtain $g\theta = c\varphi_1$ for some number c. Then $c = 0$ due to our choice of θ, whence $g = 0$ a.e.

4.7.53.° Let $\sum_{n=1}^\infty \alpha_n^2 = \infty$. Prove that there exist numbers β_n such that $\sum_{n=1}^\infty \beta_n^2 < \infty$ and $\sum_{n=1}^\infty \alpha_n\beta_n = \infty$.

4.7.54.° Let $\alpha_n \geq 0$ and $\sum_{n=1}^\infty \alpha_n = \infty$. Prove that there exist numbers $c_n \geq 0$ such that $\sum_{n=1}^\infty \alpha_n c_n = \infty$ and $\sum_{n=1}^\infty \alpha_n c_n^2 < \infty$.

HINT: in the case of a bounded sequence α_n one can partition $\mathbb{N}$ into finite intervals I_k with $2^{k-1} \leq \sum_{i\in I_k} \alpha_i < 2^k$ and for $n \in I_k$ take $c_n = 2^{-k}$; for an increasing sequence $\{\alpha_{n_k}\}$ take $c_{n_k} = \alpha_{n_k}^{-1} k^{-1}$.

4.7.55.° Let $A \subset \mathbb{R}^1$ be a set of infinite Lebesgue measure. Prove that there exists a function $f \in L^2(\mathbb{R}^1)$ that is not integrable on A.

HINT: denote by α_n the measure of the set $A \cap [n, n+1)$, $n \in \mathbb{Z}$, apply Exercise 4.7.54 and let $f = c_n$ on the above set.

4.7.56.° Let $f \in \mathcal{L}^1(\mathbb{R})$, $f > 0$. Prove that $1/f \notin \mathcal{L}^1(\mathbb{R})$.

HINT: apply the Cauchy–Bunyakowsky inequality to $f^{-1/2} f^{1/2}$.

4.7.57.° Prove that the set of nonnegative functions is closed and nowhere dense in the space $L^1[0,1]$.

4.7.58. (Müntz's theorem) Suppose we are given a sequence of real numbers $p_i > -1/2$ with $\lim\limits_{i\to\infty} p_i = +\infty$. Prove that $\sum_{i\colon p_i\neq 0} 1/p_i = \infty$ precisely when the linear span of the functions x^{p_i} is everywhere dense in $L^2[0,1]$.

HINT: see Ahiezer [**4**, Ch. 1].

4.7.59. Prove that the Haar functions h_n form a Schauder basis in $L^p[0,1]$ for all $p \in [1,+\infty)$. The Haar functions h_n are defined as follows: for all $n \geq 1$ and $1 \leq i \leq 2^n$ we set $h_{2^n+i}(t) = I_{[(2i-2)/2^{n+1},(2i-1)/2^{n+1}]}(t) - I_{((2i-1)/2^{n+1},2i/2^{n+1}]}(t)$.

HINT: see Kashin, Saakyan [**495**, Ch. 3].

4.7.60.° Let μ be a finite nonnegative measure on a space X. For $f, g \in L^0(\mu)$, we set

$$d_0(f,g) := \int_X \frac{|f-g|}{1+|f-g|}\,d\mu, \quad d_1(f,g) := \int_X \min(|f-g|,1)\,d\mu.$$

Prove that d_0 and d_1 are metrics, with respect to which $L^0(\mu)$ is complete, and that a sequence converges in one of these metrics precisely when it converges in measure (similarly for fundamental sequences).

HINT: the triangle inequality follows from the triangle inequality for the metrics $|t-s|/(1+|t-s|)$ and $\min(|t-s|,1)$ on the real line. By Chebyshev's inequality, one has $\mu(|f-g| \geq \varepsilon) = \mu\bigl(|f-g|/(1+|f-g|) \geq \varepsilon/(1+\varepsilon)\bigr) \leq d_0(f,g)/\varepsilon$. Finally, $d_0(f,g) \leq \varepsilon\mu(X) + \mu(|f-g| \geq \varepsilon)$. For d_1 one has a similar estimate.

4.7.61. (Nikodym [**719**]) Prove that on the space $L^0[0,1]$ of all Lebesgue measurable functions equipped with the metric

$$d(f,g) = \int_0^1 |f-g|/(1+|f-g|)\,dx$$

corresponding to convergence in measure, there exists no continuous linear function except for the identically zero one. Extend this assertion to the case of an arbitrary atomless probability measure.

HINT: if L is such a function, then the set $V := L^{-1}(-1,1)$ is not the whole space and contains some ball U with the center 0 and radius $r > 0$ with respect to the above metric. The set V is convex and hence contains the convex envelope of U. A contradiction is due to the fact that the convex envelope of U equals $L^0[0,1]$. Indeed, let f be an arbitrary measurable function. Then, for every n, we have $f = (f_1 + \cdots + f_n)/n$, where $f_k(t) = nf(t)I_{[(k-1)/n,k/n)}(t)$. It is clear from the definition of the metric d that if $n^{-1} < r$, then all the functions f_k belong to U.

4.7.62. Let μ be a nonnegative measure, $0 < p < 1$, and let $L^p(\mu)$ be the set of all equivalence classes of μ-measurable functions f such that $|f|^p \in L^1(\mu)$.

(i) Prove that the function

$$d_p(f,g) := \int |f-g|^p\,d\mu$$

is a complete metric on the space $L^p(\mu)$.

(ii) Prove that $L^p(\mu)$ is a linear space such that the operations of addition and multiplication by real numbers are continuous on $L^p(\mu)$ with the metric d_p (i.e., $L^p(\mu)$ is a complete metrizable topological vector space).

(iii) Prove that in the case where μ is Lebesgue measure on $[a,b]$, there is no nonzero linear function on the space $L^p(\mu)$ continuous with respect to the metric d_p. In particular, convergence in the metric d_p cannot be described by any norm.

4.7.63.° Show that a probability measure μ on a σ-algebra $\mathcal{A}$ is separable if and only if all spaces $L^p(\mu)$, $p \in (0,+\infty)$, are separable, and the separability of either of these spaces is sufficient.

HINT: use that the set of simple functions is everywhere dense in each of these spaces and that a subspace of a separable metric space is separable.

4.7.64. Let $\mathcal{A}$ be a countably generated σ-algebra (i.e., generated by a countable family of sets) and let μ_t, $t \in T$, be some family of probability measures on $\mathcal{A}$. Prove that this family is separable in the variation norm precisely when there exists a probability measure μ on $\mathcal{A}$ such that $\mu_t \ll \mu$ for all $t \in T$.

HINT: in the case of a countably generated σ-algebra the space $L^1(\mu)$ is separable; if a sequence of measures μ_{t_n} is everywhere dense in a given family of measures in the variation norm, then one can take the measure $\mu = \sum_{n=1}^{\infty} 2^{-n}\mu_{t_n}$.

4.7.65.° Let μ be a probability measure and let $f \in L^p(\mu)$. Show that the function $\theta\colon\ r \mapsto \ln \|f\|_{L^r(\mu)}^r$ is convex on $[1,p]$, i.e., $\theta\bigl(tr+(1-t)s\bigr) \le t\theta(r)+(1-t)\theta(s)$ for all $0<t<1$ and $r,s \in [1,p]$.

HINT: apply Hölder's inequality with the exponents $1/t$ and $1/(1-t)$.

4.7.66. Let ψ be a positive function on $[1,+\infty)$ increasing to the infinity. Prove that there exists a positive measurable function f on $[0,1]$ such that $\|f\|_p \le \psi(p)$ for all $p \ge 1$ and $\lim\limits_{p\to\infty} \|f\|_p = \infty$.

HINT: see George [**351**, p. 261].

4.7.67. Prove Corollary 4.5.5.

4.7.68.° Suppose that a function $f \in \mathcal{L}^1[0,2\pi]$ satisfies Dini's condition at a point x (see Theorem 3.8.8). Prove that its Fourier series at x converges to $f(x)$.

HINT: apply formula (4.3.6).

4.7.69. (W. Orlicz) Let $\{e_n\}$ be an orthonormal basis in the space $L^2[a,b]$. (i) Prove that

$$\sum_{n=1}^{\infty} \int_A |e_n(x)|^2\, dx = \infty$$

for every set $A \subset [a,b]$ of positive measure. (ii) Prove that $\sum_{n=1}^{\infty} |e_n(x)|^2 = \infty$ a.e.

HINT: (i) take an infinite orthonormal basis $\{\varphi_k\}$ in the space $L^2(A)$ by Exercise 4.7.46, show that $(I_A e_n, I_A e_n) = \sum_{k=1}^{\infty}(e_n,\varphi_k)^2$ by using the relations $I_A e_n = \sum_{k=1}^{\infty}(I_A e_n,\varphi_k)\varphi_k$, $I_A\varphi_k = \varphi_k$. (ii) Apply (i) to the sets $\bigl\{x\colon\ \sum_{n=1}^{\infty} |e_n(x)|^2 \le M\bigr\}$.

4.7.70. Let $\mathfrak{R}$ be a semiring in a σ-algebra $\mathcal{A}$ with a probability measure μ. Show that the set of linear combinations of the indicator functions of sets in $\mathfrak{R}$ is everywhere dense in $L^1(\mu)$ precisely when, for every $A \in \mathcal{A}$ and $\varepsilon > 0$, there exists a set B that is a union of finitely many sets in $\mathfrak{R}$ such that $\mu(A \triangle B) < \varepsilon$.

4.7.71. Suppose that a sequence of μ-integrable functions f_n (where μ takes values in $[0,+\infty]$) converges almost everywhere to a function f and that there exist integrable functions g_n such that $|f_n| \le g_n$ almost everywhere. Prove that if the sequence $\{g_n\}$ converges in $L^1(\mu)$ (or the measure μ is finite and $\{g_n\}$ is uniformly integrable), then f is integrable and $\{f_n\}$ converges to f in $L^1(\mu)$.

HINT: in the case of a finite measure we observe that the sequence $\{f_n\}$ is uniformly integrable; the general case reduces to the case of a σ-finite measure μ, then to the case of a finite measure μ_0 with a positive density ϱ with respect to μ. Alternatively, one can apply Young's theorem 2.8.8.

4.7.72.° Let $(X,\mathcal{A},\mu)$ be a probability space and let integrable functions f_n converge in measure to an integrable function f such that

$$\lim_{n\to\infty} \int_X \sqrt{1+f_n^2}\, d\mu = \int_X \sqrt{1+f^2}\, d\mu.$$

Prove that $f_n \to f$ in $L^1(\mu)$.

HINT: apply Young's theorem 2.8.8 and the estimate $|f_n| \le \sqrt{1+f_n^2}$.

4.7.73. (Klei, Miyara [**522**]) Let $(X, \mathcal{A}, \mu)$ be a probability space and let M be a norm bounded set in $L^1(\mu)$. The modulus of uniform integrability of M is the function

$$\eta(M,\varepsilon) := \sup\biggl\{ \int_A |f|\,d\mu\colon\ f \in M, A \in \mathcal{A}, \mu(A) \le \varepsilon \biggr\}.$$

Set $\eta(M) := \lim\limits_{\varepsilon\to 0} \eta(M,\varepsilon)$. It is clear that the equality $\eta(M)=0$ is equivalent to the uniform integrability of M. Let $f_n \in L^1(\mu)$, $f_n \ge 0$, be such that the sequence of the integrals of f_n is convergent. Prove that

$$\int_X \liminf_{n\to\infty} f_n\,d\mu \le \lim_{n\to\infty} \int_X f_n\,d\mu - \eta(\{f_n\}).$$

Show that under the above conditions the equality occurs precisely when $\{f_n\}$ contains a subsequence convergent a.e. to the function $\liminf\limits_{n\to\infty} f_n$.

4.7.74. (Farrell [**279**]) (i) Let $(X, \mathcal{A}, \mu)$ be a probability space and let $\mathcal{F}$ be an algebra of bounded measurable functions such that, for every measurable set A, there exists $f \in \mathcal{F}$ with $f > 0$ a.e. on A and $f \le 0$ a.e. on $X\backslash A$. Prove that for all $p \in [1,\infty)$ the algebra $\mathcal{F}$ is dense in $L^p(\mu)$. Moreover, the same is true if the hypothesis is fulfilled for every set A in some family $\mathcal{E} \subset \mathcal{A}$ with the property that the linear space generated by I_E, $E \in \mathcal{E}$, is dense in $L^1(\mu)$.

(ii) Let μ be a Borel probability measure on the real line and let f be a strictly increasing bounded function. Show that the algebra of functions generated by f and 1 is dense in $L^p(\mu)$, $1 \le p < \infty$.

HINT: (i) let $A \in \mathcal{A}$, let $f \in \mathcal{F}$ be the corresponding function, and let $|f| < N$; there is a uniformly bounded sequence of polynomials P_n such that $\lim\limits_{n\to\infty} P_n(t) = 1$ for all $t \in (0,N]$ and $\lim\limits_{n\to\infty} P_n(t) = 0$ for all $t \in [-N,0]$; then $P_n \circ f \in \mathcal{F}$, $P_n \circ f \to I_A$ a.e. and in $L^p(\mu)$. Hence every simple function belongs to the closure of $\mathcal{F}$. In the case of the more general assumption involving $\mathcal{E}$, the above reasoning shows that the closure of $\mathcal{F}$ in $L^p(\mu)$ contains all functions of the form $\max\bigl(-N, \min(g,N)\bigr)$, where g is a linear combination of indicators of sets in $\mathcal{E}$, $N \in \mathbb{N}$. Take a sequence $\{g_k\}$ of such linear combinations convergent in $L^1(\mu)$ to a bounded function φ. Then the functions $\max\bigl(-N, \min(g_k,N)\bigr)$ with $N > \sup|\varphi(x)|$ converge to φ in $L^p(\mu)$. Assertion (ii) follows by applying (i) to the family of rays.

4.7.75. (G. Hardy) Let $f \in L^p(0,+\infty)$, where $p > 1$. Show that the functions

$$\varphi(x) = \frac{1}{x}\int_0^x f(t)\,dt, \quad \psi(x) = \int_x^\infty \frac{f(t)}{t}\,dt$$

belong to $L^p(0,+\infty)$ as well.

HINT: see Titchmarsh [**947**, p. 405].

4.7.76.° Let G be an everywhere dense set in $L^q(\mu)$, $p^{-1}+q^{-1}=1$, $q>1$, and let a sequence $\{f_n\}$ be bounded in the norm of $L^p(\mu)$. Prove that this sequence weakly converges to $f \in L^p(\mu)$ precisely when the integrals of $f_n g$ converge to the integral of fg for every $g \in G$.

4.7.77° Give an example of a sequence of functions $f \in L^1[0,1]$ that is bounded in the norm of $L^1[0,1]$ and converges a.e. to 0, but has no subsequence convergent in the weak topology of $L^1[0,1]$.

HINT: consider the functions $f_n(t) = nI_{[0,1/n]}$.

4.7.78° Let $1 < p < \infty$. Construct an example of a sequence of functions f_n that weakly converges to zero in the space $L^p[0,1]$ and converges to zero almost everywhere on $[0,1]$, but does not converge in the norm of $L^p[0,1]$.

HINT: consider $f_n(x) = n^{1/p} I_{[0,1/n]}(x)$; use Exercise 4.7.76 applied to the set $G = L^\infty[0,1]$.

4.7.79° (i) (Riemann–Lebesgue theorem) Show that

$$\lim_{n\to\infty} \int_0^{2\pi} f(x) \sin nx \, dx = 0$$

for every Lebesgue integrable function f.

(ii) Let μ be a probability measure and let $\{\varphi_n\}$ be an orthonormal system in $L^2(\mu)$ such that $|\varphi_n| \le M$, where M is a number. Show that

$$\lim_{n\to\infty} \int f\varphi_n \, d\mu = 0$$

for every μ-integrable function f.

HINT: (i) observe that this is true for piecewise constant functions, then approximate f by such functions. Alternatively, one can refer to Proposition 3.8.4. (ii) For bounded functions f the assertion follows by Bessel's inequality, in the general case we approximate f in $L^1(\mu)$ by bounded functions.

4.7.80. Give an example of a sequence of nonnegative functions f_n that weakly converges in $L^1[0,1]$ to a function f and $\|f_n\|_{L^1} \to \|f\|_{L^1}$, but $\{f_n\}$ does not converge in the norm of $L^1[0,1]$.

HINT: consider the functions $f_n(x) = 1 + \sin(2\pi nx)$ and $f(x) = 1$.

4.7.81. Show that there exists a sequence of positive continuous functions f_n on $[0,1]$ and a continuous function f such that for all $a, b \in [0,1]$ one has

$$\lim_{n\to\infty} \int_a^b f_n(t) \, dt = \int_a^b f(t) \, dt,$$

but there is a measurable set E such that the integrals of f_n over E do not converge.

HINT: see Example 8.2.12 and the subsequent note.

4.7.82. Let μ be a measure with values in $[0, +\infty]$ on a space $(X, \mathcal{A})$. The following terminology is used in the books Hunt [**448**] and Bauer [**70**]: a set M in $\mathcal{L}^1(\mu)$ (or in $L^1(\mu)$) is called uniformly integrable if

$$\forall\, \varepsilon > 0 \ \exists\, g \in \mathcal{L}^1(\mu)\colon \quad \int_{\{|f|>g\}} |f| \, d\mu \le \varepsilon, \quad \forall\, f \in M. \tag{4.7.18}$$

With such a definition, any integrable function is uniformly integrable.

(i) Show that (4.7.18) yields the existence of a measurable set E such that the measure μ on E is σ-finite and every function $f \in M$ vanishes a.e. outside E.

(ii) Show that for finite measures (4.7.18) is equivalent to the uniform integrability in our sense.

(iii) Show that (4.7.18) is equivalent to the following property: the set M is bounded in the norm of $L^1(\mu)$ and, for every $\varepsilon > 0$, there exist a nonnegative integrable function h and a number $\delta > 0$ such that, whenever $A \in \mathcal{A}$ and

$$\int_A h\,d\mu \le \delta,$$

one has

$$\int_A |f|\,d\mu \le \varepsilon \quad \text{for all } f \in M.$$

(iv) Let the measure μ be σ-finite and let $h > 0$ be a μ-integrable function. Show that (4.7.18) is equivalent to the property that, for every $\varepsilon > 0$, there exists $C > 0$ such that

$$\int_{\{|f|>Ch\}} |f|\,d\mu \le \varepsilon, \quad \forall\, f \in M.$$

In addition, (4.7.18) is equivalent to the following: the set M is bounded in the norm of $L^1(\mu)$ and, for every $\varepsilon > 0$, there exists a number $\delta > 0$ such that if $A \in \mathcal{A}$ and

$$\int_A h\,d\mu \le \delta,$$

then

$$\int_A |f|\,d\mu \le \varepsilon \quad \text{for all } f \in M.$$

(v) Prove that (4.7.18) is equivalent to the following: M is bounded in $L^1(\mu)$, the functions in M have uniformly absolutely continuous integrals and, for every $\varepsilon > 0$, there exists a measurable set X_ε such that $\mu(X_\varepsilon) < \infty$ and

$$\int_{X\setminus X_\varepsilon} |f|\,d\mu \le \varepsilon \quad \text{for all } f \in M.$$

Hint: (i) take functions g_n corresponding to $\varepsilon_n = n^{-1}$ and the set $E = \bigcup_{n=1}^\infty \{g_n > 0\}$. (ii) Use the uniform integrability of g. (iii) In order to deduce (4.7.18) from (iii), observe that every function $f \in M$ vanishes a.e. on the set $\{h = 0\}$, hence one can pass to the space $X_0 := \{h > 0\}$ with the finite measure $\nu := h \cdot \mu$; the functions f/h, where $f \in M$, belong to $L^1(\nu)$ and have uniformly absolutely continuous integrals (with respect to ν), therefore, they form a uniformly integrable set in $L^1(\nu)$. This shows that for g one can take Ch with some C. The same reasoning proves (iv), and (v) reduces easily to the case of a finite measure.

4.7.83. Let $0 < p < q < \infty$ and let μ be a countably additive measure with values in $[0, +\infty]$. (i) Prove that $L^p(\mu) \not\subset L^q(\mu)$ precisely when there exist sets of arbitrarily small positive μ-measure. (ii) Prove that $L^q(\mu) \not\subset L^p(\mu)$ precisely when there exist sets of arbitrarily large finite μ-measure.

Hint: (i) observe that if a series of $c_n > 0$ converges, then one can find b_n increasing to $+\infty$ such that the series of $c_n b_n^p$ converges and the series of $c_n b_n^q$ diverges; (ii) is similar; see Romero [**819**], Subramanian [**918**], and also Miamee [**687**].

4.7.84. Let f and g be integrable on $[0,1]$ and let $|f(x)| \le g(x)$. Prove that there exists a sequence of integrable functions f_n such that, for every measurable set $E \subset [0,1]$, one has

$$\lim_{n\to\infty} \int_E f_n\,dx = \int_E f\,dx, \quad \lim_{n\to\infty} \int_E |f_n|\,dx = \int_E g\,dx.$$

Hint: see Zaanen [**1043**, 45.6].

4.7.85.° Suppose that functions f_n weakly converge in $L^p(\mu)$ to a function f, where $p \geq 1$. Show that $\|f\|_p \leq \liminf\limits_{n\to\infty} \|f_n\|_p$.

4.7.86.° Let $1 < p < \infty$, $p^{-1} + q^{-1} = 1$, let $\mu \geq 0$ be a σ-finite measure, and let Ψ be a continuous linear function on $L^p(\mu)$. Let $f \in L^p(\mu)$ be a function such that $\|f\|_p = 1$ and $\Psi(f) = \|\Psi\| > 0$. Prove that Ψ is given by the function $g = \mathrm{sign} f \cdot |f|^{p-1} \in L^q(\mu)$ by formula (4.4.1) and that g is a unique function generating Ψ.

Hint: take $g \in L^q(\mu)$ generating Ψ by formula (4.4.1) and observe that

$$\int fg\,d\mu = \Psi(f) = \|\Psi\| = \|g\|_q = \|f\|_p\|g\|_q,$$

whence the assertion follows by Exercise 2.12.89.

4.7.87. Let μ be a countably additive measure on a σ-algebra with values in $[0, +\infty]$. (i) Show that for any nonzero continuous linear function Ψ on $L^p(\mu)$ with $1 < p < \infty$, there exists $f \in L^p(\mu)$ with $\|f\|_p = 1$ and $\Psi(f) = \|\Psi\|$.

(ii) Prove that in the case $1 < p < \infty$ the dual to $L^p(\mu)$ can be identified with $L^q(\mu)$, $q = p/(p-1)$, in the same sense as in Theorem 4.4.1.

(iii) Extend the assertion of Exercise 4.7.86 to the case of an arbitrary (not necessarily σ-finite) countably additive measure with values in $[0, +\infty]$.

Hint: (i) use the Banach–Saks property (which follows by the uniform convexity of $L^p(\mu)$) or the reflexivity of uniformly convex spaces. (ii) If Ψ is a continuous linear function on $L^p(\mu)$ and $\|\Psi\| = 1$, then by (i) one has $f \in L^p(\mu)$ with $\|f\|_p = 1$ and $\Psi(f) = 1$. Then $g = \mathrm{sign}(f)|f|^{p-1} \in L^q(\mu)$ and $\|g\|_q = 1$. Next one verifies that $\Psi^{-1}(0) = L$, where

$$L := \Bigl\{h\colon \int hg\,d\mu = 0\Bigr\}.$$

To this end, we observe that if one has $h \in L\backslash\Psi^{-1}(0)$, then one can take a measurable set Ω outside of which f and h vanish and the restriction of the measure μ to Ω is a σ-finite measure. The restriction of Ψ to $L^p(\Omega, \mu)$ is a continuous linear functional with unit norm, hence, by Exercise 4.7.86, it is given by the function g, which yields $\Psi^{-1}(0) \cap L^p(\Omega, \mu) = L \cap L^p(\Omega, \mu)$.

4.7.88. Let μ be a nonnegative measure, $1 \leq p \leq \infty$, and let L be a linear function on $L^p(\mu)$ such that $L(f) \geq 0$ whenever $f \geq 0$. Prove the continuity of L.

Hint: if L is discontinuous, then there exists a sequence f_n such that $\|f_n\|_p \to 0$ and $L(f_n) \geq 1$. One may assume that $\|f_n\|_p \leq 4^{-n}$, passing to a subsequence. Let $p < \infty$. The series $\sum_{n=1}^\infty 2^{np}|f_n|^p$ converges a.e. to an integrable function g. Then $G := g^{1/p} \in L^p(\mu)$ and, for every k, we have $\sum_{n=1}^k |f_n| = \sum_{n=1}^\infty 2^{-n}2^n|f_n| \leq \bigl(\sum_{n=1}^\infty 2^{-np}\bigr)^{1/p'} G$, whence the uniform boundedness of the numbers $\sum_{n=1}^k L(|f_n|)$ follows, which leads to a contradiction. In the case $p = \infty$ the reasoning is similar.

4.7.89. Construct an example of a countably additive measure μ with values in $[0, +\infty]$ defined on a σ-algebra $\mathcal{A}$ such that there exists a continuous linear function Ψ on $L^1(\mu)$ that cannot be written in the form indicated in Theorem 4.4.1.

Hint: let $X = [0, 1]$ be equipped with the σ-algebra $\mathcal{A}$ of all sets that are either at most countable or have at most countable complements; let μ be the counting measure on $\mathcal{A}$, i.e., $\mu(A)$ is the cardinality of A; then every function $f \in L^1(\mu)$ is nonzero on an at most countable set $\{t_n\}$ and the functional $f \mapsto \sum_{n\colon\, t_n \leq 1/2} f(t_n)$

is continuous, but it is not generated by any function from $L^\infty(\mu)$, since such a function would coincide with $I_{[0,1/2]}$, which is not μ-measurable; see also Federer [**282**, Example 2.5.11].

4.7.90. (i) Construct a space $(X,\mathcal{A},\mu)$ with a countably additive measure μ with values in $[0,+\infty]$ and an $\mathcal{A}$-measurable function f that belongs to no $L^p(\mu)$ with $p\in[1,+\infty)$, but $fg\in L^1(\mu)$ for every function $g\in\bigcup_{q\ge1}L^q(\mu)$.

(ii) Show that if a space $(X,\mathcal{A},\mu)$ with a countably additive measure μ with values in $[0,+\infty]$ and an $\mathcal{A}$-measurable function f are such that μ is σ-finite on the set $\{f\neq0\}$ and $fg\in L^1(\mu)$ for every function $g\in L^q(\mu)$, where $1<q\le\infty$, then $f\in L^p(\mu)$, where $p^{-1}+q^{-1}=1$.

(iii) Let a measure μ on a measurable space $(X,\mathcal{A})$ be semifinite in the sense of Exercise 1.12.132, let f be an $\mathcal{A}$-measurable function, and let $p^{-1}+q^{-1}=1$, where $1\le p<\infty$. Suppose that $fg\in L^1(\mu)$ for every function $g\in L^q(\mu)$. Show that $f\in L^p(\mu)$.

HINT: (i) consider the measure μ assigning $+\infty$ to every nonempty set in $[0,1]$ and $f=1$; (ii) apply Corollary 4.4.5; (iii) show that $\mu(|f|\ge c)<\infty$ for all $c>0$; to this end, prove that assuming the contrary and using that the measure is semifinite, one can find a function $g\in L^q(\mu)$ such that $gI_{\{|f|\ge c\}}$ does not belong to $L^1(\mu)$.

4.7.91. (Segal [**861**]) (i) Let μ be a measure with values in $[0,+\infty]$. Prove that μ is semifinite precisely when the embedding $L^\infty(\mu)\to\bigl(L^1(\mu)\bigr)^*$ is injective.

(ii) Let μ be a semifinite measure. Prove that μ is Maharam (or localizable) in the sense of Exercise 1.12.134 precisely when, for every $L\in L^1(\mu)^*$, there exists a unique element $g_L\in L^\infty(\mu)$ with

$$L(f)=\int fg_L\,d\mu\quad\text{for all } f\in L^1(\mu).$$

In this case, $L\mapsto g_L$ is an isometry between $L^1(\mu)^*$ and $L^\infty(\mu)$.

HINT: (i) if μ is semifinite and $f,g\in L^\infty(\mu)$ are not equal, then there exists a set of finite positive measure on which f and g differ; conversely, if there is a measurable set E without subsets of finite positive measure, then all functions fI_E, $f\in L^\infty(\mu)$, generate the zero functional on $L^1(\mu)$. (ii) See Fremlin [**322**, Ch. 6], Rao [**788**, p. 288], Zaanen [**1043**].

4.7.92. Let $X=\mathbb{R}^2$, $\mu(A)=+\infty$ if A is uncountable, $\mu(A)=\delta_0(A)$ if A is at most countable, where δ_0 is Dirac's measure at the origin. Show that μ is a countably additive measure on the σ-algebra of all sets in $\mathbb{R}^2$ with values in $[0,+\infty]$ that is neither localizable nor semifinite. Verify that $L^1(\mu)=L^p(\mu)\neq L^\infty(\mu)$ for all $p\in[1,+\infty)$ and $\|f\|_{L^p(\mu)}=|f(0)|$ for all $f\in\mathcal{L}^p(\mu)$.

4.7.93. Let $(X,\mathcal{A},\mu)$ be a space with a complete countably additive measure μ with values in $[0,+\infty]$. Denote by $\mathcal{N}_{loc}(\mu)$ the class of locally zero sets, i.e., sets E such that $\mu(E\cap A)=0$ for all $A\in\mathcal{A}$ with $\mu(A)<\infty$. Next, denote by $L^\infty_{loc}(\mu)$ the class of all μ-measurable functions f with $\|f\|_{\infty,loc}<\infty$, where we set $\|f\|_{\infty,loc}=\inf\bigl\{a\colon\ \{x\colon\ |f(x)|>a\}\in\mathcal{N}_{loc}(\mu)\bigr\}$ and identify functions that are not equal only on a set from $\mathcal{N}_{loc}(\mu)$.

(i) Prove that $L^\infty_{loc}(\mu)$ is a Banach space with the norm $\|\cdot\|_{\infty,loc}$.

(ii) Prove that for all $f\in L^\infty_{loc}(\mu)$ one has

$$\|f\|_{\infty,loc}=\sup\biggl\{\biggl|\int_X fg\,d\mu\biggr|,\|g\|_{L^1(\mu)}=1\biggr\},$$

and the mapping $L^\infty_{loc}(\mu) \to \bigl(L^1(\mu)\bigr)^*$ is injective and preserves the distances.

(iii) Let $\mathcal{P}$ be the class of all simple μ-integrable functions and let a μ-measurable function f be such that $fg \in L^1(\mu)$ for all $g \in \mathcal{P}$ and

$$\sup\biggl\{\biggl|\int_X fg\,d\mu\biggr|\colon\ g \in \mathcal{P}, \|g\|_{L^1(\mu)} = 1\biggr\} < \infty.$$

Prove that $f \in L^\infty_{loc}(\mu)$.

(iv) Let a measure μ be decomposable in the sense of Exercise 1.12.131. Prove that every continuous linear functional on $L^1(\mu)$ is generated by a function from the class $L^\infty_{loc}(\mu)$, i.e., $\bigl(L^1(\mu)\bigr)^*$ is naturally isomorphic to $L^\infty_{loc}(\mu)$.

4.7.94. Let μ and γ be the measures with values in $[0, +\infty]$ defined in Exercise 1.12.137 and let

$$l(f) = \int f\,d\gamma, \quad f \in L^1(\mu).$$

Prove that l is a continuous linear functional on $L^1(\mu)$, but there is no function $g \in L^\infty_{loc}(\mu)$ such that

$$l(f) = \int fg\,d\mu \quad \text{for all } f \in L^1(\mu).$$

4.7.95.° Let $f_n, f \in L^\infty[a,b]$. Prove that the following conditions are equivalent:
(i) one has

$$\int_a^b f_n(x)g(x)\,dx \to \int_a^b f(x)g(x)\,dx, \quad \forall\, g \in L^1[a,b];$$

(ii) one has $\sup_n \|f_n\|_{L^\infty} < \infty$ and

$$\int_a^z f_n(x)\,dx \to \int_a^z f(x)\,dx, \quad \forall\, z \in [a,b].$$

Hint: use the Banach–Steinhaus theorem and the fact that the linear space generated by the indicators of intervals is dense in $L^1[a,b]$.

4.7.96.° Let f be a measurable function on the real line with a period 1.
(i) Prove that if $f \in L^1[0,1]$, then

$$\lim_{n\to\infty} \int_0^1 g(x)f(nx)\,dx = \int_0^1 g(x)\,dx \int_0^1 f(x)\,dx \tag{4.7.19}$$

for all $g \in C[0,1]$ (where $n \in \mathbb{N}$).

(ii) Prove that if f is bounded, then the above relation is true for all $g \in L^1[0,1]$.

Hint: subtracting from the function f its integral over $[0,1]$, we may assume that this integral vanishes; then observe that

$$\int_0^1 f(nx)\,dx = 0$$

for all $n \in \mathbb{N}$ and derive that

$$\int_0^z f(nx)\,dx = n^{-1}\int_0^{nz} f(y)\,dy \to 0, \quad \forall\, z \in [0,1].$$

Finally, observe that (4.7.19) for smooth g follows by the integration by parts formula; in the general case, we consider suitable approximations (uniform for continuous g and in $L^1[0,1]$ for integrable g).

4.7.97° Let f be a bounded measurable function on the real line with a period 1. Show that if a sequence of functions $f(nx)$ has a subsequence convergent on a set of positive measure, then f a.e. equals some constant.

HINT: apply the previous exercise.

4.7.98. Prove that the functions $|\sin \pi nx|$ converge weakly in $L^2[0,1]$ and find their limit.

HINT: $\{|\sin \pi nx|\}$ converges weakly to $2/\pi$ by Exercise 4.7.96.

4.7.99. Suppose a sequence of functions f_n converges weakly in $L^1[0,1]$ to a function f. Is it true that the functions $|f_n|$ converge weakly to $|f|$?

HINT: no; see Exercise 4.7.98.

4.7.100. Prove that for every irrational number α, there exist infinitely many rational numbers p/q, where p, q are integers, such that $|\alpha - p/q| < q^{-2}$.

HINT: consider $n+1$ numbers $0, \alpha - [\alpha], \ldots, n - [n\alpha]$, where $[x]$ is the integer part of x, and n intervals $\bigl[j/n, (j+1)/n\bigr)$, $j = 0, 1, \ldots, n-1$. Then, one of these intervals contains at least two of the above numbers, say, $n_1\alpha - [n_1\alpha]$ and $n_2\alpha - [n_2\alpha]$, $n_1 < n_2$. Set $q = n_2 - n_1$, $p = [n_2\alpha] - [n_1\alpha]$. Then $q \le n$ and

$$|q\alpha - p| = |n_2\alpha - [n_2\alpha] - n_1\alpha + [n_1\alpha]| < 1/n,$$

i.e., $|\alpha - p/q| < (nq)^{-1} \le q^{-2}$. Suppose that the regarded collection of rational numbers consists of only finitely many numbers $p_1/q_1, \ldots, p_m/q_m$. Letting $\varepsilon = \min_{i \le m} |\alpha - p_i/q_i|$, we pick n such that $1/n < \varepsilon$. Then, as shown above, we can find p/q with $q \le n$ and $|\alpha - p/q| < (nq)^{-1}$, which is estimated by ε as well as by q^{-2}, contrary to our choice of ε.

4.7.101. Let f be a measurable function on $[0,1)$, extended periodically to the whole real line and having the integral $I(f)$ over $[0,1]$. For every $n \in \mathbb{N}$, we consider the Riemannian sum

$$S_n f(x) := n^{-1} \sum_{k=0}^{n} f(x + k/n), \quad x \in [0,1).$$

(i) Prove that $\|S_n f\|_{L^p[0,1)} \le \|f\|_{L^p[0,1)}$ for all $f \in L^p[0,1)$, $p \in [1,\infty)$, and that $\|I(f) - S_n f\|_{L^p[0,1)} \to 0$ as $n \to \infty$.

(ii) Show that, for every function $f \in L^1[0,1)$, there exists a sequence $n_m \to \infty$ such that $S_{n_m} f(x) \to I(f)$ for almost all $x \in [0,1)$ (in fact, one can take $n_m = 2^m$, see Example 10.3.18 in Chapter 10).

(iii) Give an example of an integrable function f with a period 1 such that $S_n f(x) \to I(f)$ only on a measure zero set. Verify that if $f(x) = x^{-r}$ for $x \in (0,1)$, where $r \in (1/2, 1)$, then one has the equality $\limsup_{n\to\infty} S_n f(x) = +\infty$ almost everywhere.

(iv) Show that in (iii) one can take for f the indicator of an open set.

HINT: (i) use that

$$\int_0^1 |f(x+h)|\,dx = \int_0^1 |f(x)|\,dx$$

if f has a period 1; then verify convergence for continuous function; (ii) use the Riesz theorem; (iii) use Exercise 4.7.100 and observe that $S_q f(\alpha) \ge q^{2r-1}$; (iv) see Besicovitch [**84**], Rudin [**833**].

4.7.102. Let μ be a probability measure and let $f \in L^1(\mu)$. Prove that f belongs to $L^p(\mu)$ with some $p \in (1, \infty)$ precisely when there exists $C > 0$ such that

$$\sum_{k=1}^{n} \mu(A_k)^{1-p} \Bigl| \int_{A_k} f \, d\mu \Bigr|^p \le C$$

for every finite partition of the space into disjoint measurable sets A_k of positive measure. In addition, the smallest possible constant C equals $\|f\|_p^p$.

HINT: if $f \in L^p(\mu)$, then the left-hand side of the above inequality is estimated by $\|f\|_p^p$ by Hölder's inequality. Conversely, if there exists such a number C, then the assertion reduces to $f \ge 0$ (by considering separately the sets where $f \ge 0$ and $f < 0$). The corresponding estimate is true for every function $f_N = \min(f, N)$. By choosing for A_k the set $\{c_k \le f_N < c_k + \varepsilon\}$ with a sufficiently small $\varepsilon > 0$, one can obtain in the left-hand side of our inequality the values that are arbitrarily close to $\|f_N\|_p^p$; hence $\|f_N\|_p^p \le C$ for all N, whence $\|f\|_p^p \le C$.

4.7.103. Let $f \in \mathcal{L}^1[0,1]$ and

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

Prove that $f \in \mathcal{L}^p[0,1]$ with some $p \in (1, +\infty)$ precisely when there exists $C > 0$ such that

$$\sum_{k=1}^{n} \frac{|F(x_k) - F(x_{k-1})|^p}{(x_k - x_{k-1})^{p-1}} \le C$$

for every finite partition $0 = x_0 < x_1 < \cdots < x_n = 1$, and the smallest possible C coincides with $\|f\|_p^p$.

HINT: if $f \in \mathcal{L}^p[0,1]$, then the above estimate is a special case of Exercise 4.7.102; on the other hand, this estimate shows that

$$\Bigl| \int_0^1 f g \, dx \Bigr| \le C^{1/p} \|g\|_q$$

for every function g that equals c_k on $[x_k, x_{k+1})$, which follows by Hölder's inequality. By the Riesz theorem, there exists a function $f_0 \in \mathcal{L}^p[0,1]$ with

$$\int_0^1 f_0 g \, dx = \int_0^1 f g \, dx$$

for all g of the indicated form; then $f = f_0$ a.e.

4.7.104.° Let $f \in \mathcal{L}^1(\mathbb{R}^1)$. (i) Show that if $\varepsilon_n \to 0$, then

$$\lim_{n\to\infty} \int_{-\infty}^{+\infty} |f(x + \varepsilon_n) - f(x)| \, dx = 0.$$

(ii) Show that

$$\lim_{|t|\to\infty} \int_{-\infty}^{+\infty} |f(x + t) - f(x)| \, dx = 2 \int_{-\infty}^{+\infty} |f(x)| \, dx.$$

(iii) Let $f_n \to f$ in $L^1(\mathbb{R}^1)$ and $a_n \to a$ in $\mathbb{R}^1$. Show that

$$\lim_{n\to\infty} \int_{-\infty}^{+\infty} |f_n(x + a_n) - f(x + a)| \, dx = 0,$$

$$\lim_{n\to\infty} \int_{-\infty}^{+\infty} |f_n(x + a_n)| \, dx = \int_{-\infty}^{+\infty} |f(x + a)| \, dx.$$

HINT: in (i) and (ii) consider first $f \in C_0^\infty(\mathbb{R}^1)$, then take $f_j \in C_0^\infty(\mathbb{R}^1)$ convergent to f in $L^1(\mathbb{R}^1)$; (iii) apply (i) to $\varepsilon_n = a_n - a$ and use the translation invariance of Lebesgue measure.

4.7.105.° (Young [**1035**]) Suppose that integrable functions f_n on a space with a finite measure μ converge a.e. to a function f and

$$\int_E f_n \, d\mu \to 0 \quad \text{as } n \to \infty,\ \mu(E) \to 0.$$

Prove that f is integrable.

HINT: observe that this condition implies the uniform absolute continuity of the integrals of f_n.

4.7.106. Construct a sequence $f_n \in \mathcal{L}^1[0,1]$ with $\|f_n\|_{L^1[0,1]} \le 1$ that is uniformly integrable on no set E of positive measure (in particular, the closure of this sequence in the weak topology of $L^1(E)$ is not compact).

HINT: see Ball, Murat [**48**].

4.7.107.° Let $(X, \mathcal{A}, \mu)$ be a probability space. Prove that a set $F \subset L^1(\mu)$ is uniformly integrable precisely when

$$\lim_{M\to+\infty} \sup_{f\in F} \int_X \max(|f| - M, 0) \, d\mu = 0. \tag{4.7.20}$$

HINT: (4.7.20) yields

$$\lim_{M\to+\infty} \sup_{f\in F} \int_{\{|f|\ge 2M\}} \max(|f| - M, 0) \, d\mu = 0,$$

hence $\lim\limits_{M\to+\infty} \sup_{f\in F} M\mu(|f| \ge 2M) = 0$. Therefore,

$$\lim_{M\to+\infty} \sup_{f\in F} \int_{\{|f|\ge 2M\}} |f| \, d\mu = 0.$$

It is clear that the uniform integrability yields (4.7.20).

4.7.108. (see Bourgain [**120**]) Show that a set $F \subset L^1[0,1]$ has compact closure in the weak topology if and only if, for every $\varepsilon > 0$, there exists a number C such that, for every function $f \in F$, there is a measurable set $S_f \subset [0,1]$ such that

$$\int_{S_f} |f(t)| \, dt \le \varepsilon \quad \text{and } |f(t)| \le C \text{ for all } t \in [0,1]\backslash S_f.$$

HINT: observe that F with the indicated property is bounded and uniformly integrable.

4.7.109. Let A be a nonempty set. Suppose that for every $n \in \mathbb{N}$ and $\alpha \in A$, we are given a function $f_{n,\alpha} \in L^2[0,1]$ such that, for every function g in $L^2[0,1]$, one has $\lim\limits_{n\to\infty} (f_{n,\alpha}, g) = 0$ uniformly in $\alpha \in A$. Prove that, for every $g \in L^2[0,1]$ and $\varepsilon > 0$, there exists N such that for every interval $I \subset [0,1]$ one has

$$\left| \int_I g(x) f_{n,\alpha}(x) \, dx \right| < \varepsilon, \quad \forall\, n \ge N, \alpha \in A.$$

Prove the analogous assertion for functions on a cube in $\mathbb{R}^n$.

HINT: by the Hahn–Banach theorem we obtain $\sup_{n,\alpha}\|f_{n,\alpha}\|_2<\infty$; hence by the Cauchy–Bunyakowsky inequality and the absolute continuity of the Lebesgue integral, there exists $\delta>0$ such that

$$\left|\int_I g(x)f_{n,\alpha}(x)\,dx\right|<\varepsilon/4$$

for every set I with measure less than δ; next we partition $[0,1]$ into equal intervals $J_1,\ldots,J_k$ of length less than δ and take N such that

$$\left|\int_{J_i} f_{n,\alpha}g\,dx\right|<\varepsilon/(2k)\quad\text{for all } i=1,\ldots,k,\ n\geq N \text{ and } \alpha\in A;$$

the integral of $f_{n,\alpha}g$ over any interval I is the sum of $m\leq k$ integrals over intervals J_i and two integrals over intervals of length less than δ. The case of a cube is similar (cf. Gaposhkin [**338**, Lemma 1.4.1]).

4.7.110.° Let μ be a finite nonnegative measure and let $1\leq p<\infty$. Prove that a set $K\subset L^p(\mu)$ has compact closure in $L^p(\mu)$ precisely when the set $\{|f|^p\colon\ f\in K\}$ is uniformly integrable and every sequence in K contains a subsequence convergent in measure.

HINT: use the Lebesgue–Vitali theorem.

4.7.111. Let $1\leq p<\infty$ and let $\mathcal{K}$ be a bounded set in $L^p(\mathbb{R}^n)$.

(i) (A.N. Kolmogorov; for $p=1$, A.N. Tulaikov) Prove that the closure of $\mathcal{K}$ in $L^p(\mathbb{R}^n)$ is compact precisely when the following conditions are fulfilled:

(a) one has

$$\sup_{f\in\mathcal{K}}\lim_{C\to\infty}\int_{|x|>C}|f(x)|^p\,dx=0,$$

(b) for every $\varepsilon>0$, there exists $r>0$ such that $\sup_{f\in\mathcal{K}}\|f-S_rf\|_p\leq\varepsilon$, where S_rf is Steklov's function defined by the equality

$$S_rf(x):=\lambda_n\bigl(B(x,r)\bigr)^{-1}\int_{B(x,r)}f(y)\,dy,$$

$B(x,r)$ is the ball of radius r centered at x.

(ii) (M. Riesz) Show that the compactness of the closure of $\mathcal{K}$ is equivalent also to condition (a) combined with

(b') one has

$$\sup_{f\in\mathcal{K}}\lim_{h\to0}\int_{\mathbb{R}^n}|f(x+h)-f(x)|^p\,dx=0.$$

(iii) (V.N. Sudakov) Show that conditions (a) and (b) (or (a) and (b')) yield the boundedness of $\mathcal{K}$ in $L^p(\mathbb{R}^n)$, hence there is no need to require boundedness in advance.

HINT: (i) if $\mathcal{K}$ has compact closure, then $\mathcal{K}$ is bounded and, for every $\varepsilon>0$, has a finite ε-net (a set whose ε-neighborhood contains $\mathcal{K}$); hence the necessity of (a) and (b) follows from the fact that both conditions are fulfilled for every single function f. For the proof of sufficiency we observe that $S_r(\mathcal{K})$ has compact closure. Indeed, S_r is the operator of convolution with the bounded function $g=I_{B(0,r)}/\lambda_n\bigl(B(0,r)\bigr)$. For any $\delta>0$, one has a function $g_\delta\in C_0^\infty(\mathbb{R}^n)$ with $\|g_\delta-g\|_1\leq\delta$, which by Young's inequality reduces everything to the operator of convolution with g_δ. Then, the functions $g_\delta*f$, $f\in\mathcal{K}$, are equicontinuous on balls, whence one can easily obtain that every sequence in this set has a subsequence convergent in L^p. Condition (b') yields (b), hence (ii) follows from (i). Finally, (iii) is verified in Sudakov [**919**] by means of the following reasoning: if a linear operator S is compact (or has a compact power) and 1 is not its eigenvalue, then $I - S$ is invertible, hence the estimate $\|f - Sf\| \leq 1$, $f \in \mathcal{K}$, yields the boundedness of $\mathcal{K}$. In our case the verification reduces to proving that if an integrable function f with support in a ball U agrees on U with $S_r f$, then $f = 0$.

4.7.112.° Let μ be a signed measure on a measurable space $(X, \mathcal{A})$ such that $\mu(X) = 0$. Prove that $\|\mu\| = 2\sup_{A\in\mathcal{A}} |\mu(A)|$. In particular, for probability measures μ_1 and μ_2, we have $\|\mu_1 - \mu_2\| = 2\sup_{A\in\mathcal{A}} |(\mu_1 - \mu_2)(A)|$.

HINT: use that $\|\mu\| = \mu(X^+) - \mu(X^-)$ and $\mu(X^+) = -\mu(X^-)$, where $X = X^+ \cup X^-$ is the Hahn decomposition.

4.7.113. Construct a sequence of bounded signed countably additive measures on some algebra such that this sequence is uniformly bounded on every set in this algebra, but is not bounded in the variation norm.

HINT: consider the algebra of finite subsets of $\mathbb{N}$ and their complements and take the measures $\mu_n(A) = \sum_{n_k\in A\cap[1,\ldots,n]} c_{n_k}$, where c_n are terms of a convergent series that is not absolutely convergent.

4.7.114. Find a sequence of nonnegative countably additive measures that has a finite limit on every set in some algebra $\mathcal{A}$, but does not converge on some set in $\sigma(\mathcal{A})$.

HINT: consider the measures $f_n \cdot \lambda$, where λ is Lebesgue measure on $[0,1]$ and f_n are the functions from Exercise 4.7.81.

4.7.115. Prove that if a σ-algebra $\mathcal{A}$ is infinite, then the topology of convergence of measures on all sets in $\mathcal{A}$ cannot be generated by a norm.

HINT: use that the dual to the space of measures with the topology of setwise convergence coincides with the linear space L of simple functions; the dual to a Banach space is Banach; if $\mathcal{A}$ is infinite, then L cannot be complete with respect to a norm q, since for all $A_n \in \mathcal{A}$, the function $\sum_{n=1}^{\infty} 2^{-n} q(I_{A_n})^{-1} I_{A_n}$ belongs to L, which is impossible because there exist sets A_n such that this function assumes countably many values.

4.7.116. Let $\mathcal{A}$ be the Borel σ-algebra of $[0,1]$. Show that on the space $\mathcal{M}$ of all countably additive measures on $\mathcal{A}$ all three topologies considered in §4.7(v), i.e., the topology of setwise convergence, the topology generated by the duality with the space of all bounded $\mathcal{A}$-measurable functions, and the topology $\sigma(\mathcal{M}, \mathcal{M}^*)$, are distinct, although the collections of convergent countable sequences in these topologies are the same.

HINT: the dual spaces to $\mathcal{M}$ with the first two topologies are identified, respectively, with the space of all simple functions and the space of all bounded $\mathcal{A}$-measurable functions, but these two spaces are distinct for any infinite σ-algebra. If one takes a non-Borel Souslin set A, then the functional $\mu \mapsto \mu(A)$ belongs to $\mathcal{M}^*$, but is not generated by any $\mathcal{A}$-measurable function.

4.7.117. Let $\mathcal{A}$ be an algebra of sets and let $\{\mu_n\}$ be a uniformly countably additive sequence of bounded measures on the generated σ-algebra $\sigma(\mathcal{A})$. Prove that if, for every $A \in \mathcal{A}$, there exists a finite limit $\lim\limits_{n\to\infty} \mu_n(A)$, then the same is true for every $A \in \sigma(A)$.

4.7.118. (Drewnowski [**237**]) (i) Let $\mathcal{A}$ be a σ-algebra and let $\mu\colon\ \mathcal{A}\to\mathbb{R}^1$ be a bounded additive function. Suppose that $A_n\in\mathcal{A}$ are disjoint sets. Prove that there exists a sequence $\{n_k\}$ such that μ is countably additive on the σ-algebra generated by $\{A_{n_k}\}$.

(ii) Show that if in (i) we are given a sequence of bounded additive functions μ_i on $\mathcal{A}$, then one can choose a common sequence $\{n_k\}$ for all μ_i.

Hint: see Drewnowski [**237**], Swartz [**924**, §2.2].

4.7.119. (P. Antosik and J. Mikusiński) Suppose that for all $i,j\in\mathbb{N}$ we have numbers x_{ij} such that, for every j, there exists a finite limit $x_j=\lim\limits_{i\to\infty}x_{ij}$, and that every sequence of natural numbers m_j possesses a subsequence $\{k_j\}$ such that the sequence $\sum_{j=1}^\infty x_{ik_j}$ converges to a finite limit as $i\to\infty$. Prove that $x_j=\lim\limits_{i\to\infty}x_{ij}$ uniformly in $j\in\mathbb{N}$, $\lim\limits_{j\to\infty}x_{ij}=0$ uniformly in $i\in\mathbb{N}$, and $\lim\limits_{j\to\infty}x_{jj}=0$.

Hint: see Swartz [**924**, §2.8].

4.7.120. (i) Deduce Theorem 4.6.3 from Exercise 4.7.119.

(ii) Prove that Corollary 4.6.4 remains valid in the case where μ_n is a bounded finitely additive set function on a σ-algebra $\mathcal{A}$.

Hint: (ii) use Exercise 4.7.118; see Diestel [**223**, p. 80].

4.7.121. Prove Proposition 4.7.39.

4.7.122. Prove Lemma 4.7.40.

Hint: we may assume that $|m_n|\bigl(\bigcup_j A_j\bigr)\le 1$; let us partition $\mathbb{N}$ into infinitely many disjoint infinite parts Σ_p. If there is p such that for every $k\in\Sigma_p$ one has $|m_k|\bigl(\bigcup_{j\in\Sigma_p\backslash\{k\}}A_j\bigr)<\varepsilon$, then Σ_p is a required subsequence. If there is no such p, then for every p there is $k_p\in\Sigma_p$ with $|m_{k_p}|\bigl(\bigcup_{j\in\Sigma_p\backslash\{k_p\}}A_j\bigr)\ge\varepsilon$. Since

$$\bigcup_{j\in\Sigma_p\backslash\{p\}}A_j\subset\Bigl(\bigcup_n A_n\Bigr)\backslash\Bigl(\bigcup_n A_{k_n}\Bigr)\quad\text{and}\quad |m_{k_p}|\Bigl(\bigcup_n A_n\Bigr)\le 1,$$

one has $|m_{k_p}|\bigl(\bigcup_n A_{k_n}\bigr)\le 1-\varepsilon$ for all p. Let us pass to $m_n'=m_{k_n}$ and $A_n'=A_{k_n}$. Now $|m_n'|\bigl(\bigcup_n A_n'\bigr)\le 1-\varepsilon$. We repeat the described step. If we still have no required Σ_p, then we obtain a subsequence k_n with $|m_{k_p}'|\bigl(\bigcup_n A_{k_n}'\bigr)\le 1-2\varepsilon$ for all p. In finitely many steps we obtain a desired subsequence. One could also use Exercise 4.7.118.

4.7.123. Prove Lemma 4.7.41.

Hint: use Exercise 4.7.120 and 4.7.122 and suppose the contrary; see Diestel [**223**, p. 83].

4.7.124. (Kaczmarz, Nikliborc [**474**]) Let φ be a continuous even function on the real line with the following properties (α): $\varphi(t)>0$ if $t\ne 0$ and there exist A and a such that $\varphi(t)\ge A$ if $|t|\ge a$. Let μ be a probability measure on $(X,\mathcal{A})$ and let f_n be μ-measurable functions.

(i) Suppose that

$$\int_X\varphi(f_n-f_m)\,d\mu\to 0\quad\text{as } n,m\to\infty.$$

Prove that there exists a μ-measurable function f such that

$$\int_X\varphi(f-f_n)\,d\mu\to 0.$$

(ii) Let φ satisfy the following additional condition (β): there is N such that $\varphi(t+s) \le N\varphi(t) + N\varphi(s)$. Suppose that the functions $\varphi \circ f_n$ are integrable. Show that in (i) one has

$$\int_X \varphi(f_n)\,d\mu \to \int_X \varphi(f)\,d\mu.$$

(iii) Suppose that the functions f_n converge a.e. to some function f and there exists a function φ with the properties (α) and (β) and finite integrals $\varphi \circ f_n$. Show that there exists a continuous even function ψ with the properties (α) and (β) such that $\lim\limits_{|t|\to\infty} \psi(t)/\varphi(t) = 0$ and

$$\int_X \psi(f - f_n)\,d\mu \to 0.$$

In particular, since one can always take a bounded function for φ, there exists an unbounded function ψ with the aforementioned properties.

4.7.125. Let $\varphi\colon\ [0,\infty) \to [0,\infty)$ with $\varphi(0) = 0$ be either an increasing concave function or a convex function with $\varphi(2x) \le C\varphi(x)$. Let $(X, \mathcal{A}, \mu)$ be a probability space and let measurable functions f_n converge in measure to f. Suppose that $\varphi \circ |f|, \varphi \circ |f_n| \in L^1(\mu)$ and

$$\int_X \varphi \circ |f_n|\,d\mu \to \int_X \varphi \circ |f|\,d\mu.$$

Prove that

$$\int_X \varphi \circ |f_n - f|\,d\mu \to 0$$

and that the functions $\varphi \circ |f_n|$ are uniformly integrable.

HINT: the uniform integrability follows by Theorem 2.8.9; one has $\varphi(x+y) \le C_1[\varphi(x) + \varphi(y)]$, $C_1 = \max(C/2, 1)$; then $\varphi \circ |f_n - f| \le C_1[\varphi \circ |f_n| + \varphi \circ |f|]$. The second case is analogous.

4.7.126. Let μ be a nonnegative measure and let $\varphi\colon\ [0,+\infty) \to [0,+\infty)$ be a continuous increasing convex function such that $\varphi(0) = 0$, $\varphi(x) > 0$ if $x > 0$. For any measurable function f, we set

$$\|f\|_\varphi := \inf\Bigl\{\alpha > 0\colon\ \int \varphi(|f|/\alpha)\,d\mu \le 1\Bigr\}$$

and denote by $\mathcal{L}^\varphi(\mu)$ the set of all f with $\|f\|_\varphi < \infty$. Show that:

(i) $\mathcal{L}^\varphi$ is closed under sums and multiplication by scalars and the corresponding linear space $L^\varphi(\mu)$ of the equivalence classes is complete with respect to the norm $\|\cdot\|_\varphi$ (the Orlicz space); (ii) if f and g are equimeasurable, then $\|f\|_\varphi = \|g\|_\varphi$.

HINT: see Krasnosel'skiĭ, Rutickiĭ [**546**], Rao [**788**].

4.7.127. Let μ be a finite nonnegative measure. For every measurable function f, we set

$$f^*(t) = \inf\bigl\{s \ge 0\colon\ \mu\bigl(x\colon\ |f(x)| > s\bigr) \le t\bigr\}$$

and for all $p, q \in [1, \infty)$ we define the Lorentz space $L^{p,q}(\mu)$ as the set of all equivalence classes of measurable functions f such that

$$\int_0^\infty t^{1/p-1}[f^*(t)]^q\,dt < \infty.$$

Show that $L^{p,p}(\mu) = L^p(\mu)$. On Lorentz classes, see Stein, Weiss [**908**], Nielsen [**714**], Zaanen [**1043**].

4.7.128° (Tagamlickiĭ [**930**]) Let μ be a probability measure and let a sequence of μ-integrable functions f_n converge in measure to a function f. Prove the equivalence of the following conditions:

(i) $f \in L^1(\mu)$ and $f_n \to f$ in $L^1(\mu)$;

(ii) for every subsequence $\{f_{n_k}\}$, there exists a function $\varphi \in L^1(\mu)$ such that, for infinitely many values k, one has $|f_{n_k}(x)| \le \varphi(x)$ a.e.

Hint: if $f_n \to f$ in $L^1(\mu)$, then $|f_{k_n}| \le |f| + \sum_{j=1}^{\infty} |f_{k_j} - f|$, where k_j is chosen in such a way that $\|f_{k_j} - f\|_{L^1(\mu)} \le 2^{-j}$; for a subsequence the reasoning is similar; (i) follows from (ii) by the dominated convergence theorem.

4.7.129. (Fréchet [**316**], Veress [**974**]) Let μ be a probability measure on a space X and let M be some set of μ-measurable functions. Prove the equivalence of the following conditions:

(i) the set M has compact closure in the metric of convergence in measure (Exercise 4.7.60);

(ii) every sequence in M contains an a.e. convergent subsequence;

(iii) for every $\varepsilon > 0$ and $\alpha > 0$, there exists a finite collection of measurable functions $\psi_1, \ldots, \psi_n$ such that, for every function $f \in M$, one can find an index $i \le n$ with $\mu\bigl(x\colon\ |f(x) - \psi_i(x)| \ge \varepsilon\bigr) < \alpha$;

(iv) for every $\varepsilon > 0$, there exist a number $C > 0$ and a finite partition of the space into disjoint measurable parts $E_1, \ldots, E_n$ such that, for every function $f \in M$, there exists a measurable set E_f with the following properties:

$$\mu(E_f) < \varepsilon, \quad \sup_{x \in X \setminus E_f} |f(x)| < C, \quad \sup_{x,y \in E_i \setminus E_f} |f(x) - f(y)| < \varepsilon$$

for all $f \in M$ and $i = 1, \ldots, n$.

Hint: see Dunford, Schwartz [**256**, Theorem IV.11.1].

4.7.130. Let $\mathcal{A}$ be a σ-algebra of subsets of a space X. Prove that a set M in the space of all bounded measures on $\mathcal{A}$ has compact closure in the topology of setwise convergence precisely when for every uniformly bounded sequence of $\mathcal{A}$-measurable functions f_n converging pointwise to 0, one has the equality

$$\lim_{n\to\infty} \int_X f_n \, d\mu = 0$$

uniformly in $\mu \in M$.

Hint: if M is compact, then we apply Theorem 4.7.25(ii) and Egoroff's theorem. If the above condition is fulfilled, then condition (ii) in Lemma 4.6.5 is satisfied, so Theorem 4.7.25(i) applies.

4.7.131. (Areshkin [**29**]) Suppose that bounded countably additive signed measures μ_n on a σ-algebra $\mathcal{A}$ in a space X converge to a measure μ on every set in $\mathcal{A}$. Let $X = X^+ \cup X^-$, $X = X_n^+ \cup X_n^-$ be the Hahn decompositions for μ and μ_n. Prove that the measures $|\mu_n|$ converge to $|\mu|$ on every set in $\mathcal{A}$ precisely when

$$\lim_{n\to\infty} \mu_n(X^+ \cap X_n^-) = \lim_{n\to\infty} \mu_n(X^- \cap X_n^+) = 0.$$

4.7.132. (Areshkin [**31**]) Suppose that bounded nonnegative countably additive measures μ_n on a σ-algebra $\mathcal{A}$ in a space X converge to a measure μ on every set in $\mathcal{A}$ and that we are given $\mathcal{A}$-measurable functions f_n and f.

(i) Suppose that the functions f_n converge to f μ-a.e. Prove that, for every $\delta > 0$, one has $\lim\limits_{n\to\infty} \mu_n\bigl(x\colon\ |f(x) - f_n(x)| \ge \delta\bigr) = 0$.

(ii) Suppose that for every $\delta > 0$, one has $\lim\limits_{n\to\infty} \mu_n\bigl(x\colon |f(x) - f_n(x)| \ge \delta\bigr) = 0$ and that the functions f_n are uniformly bounded. Prove that

$$\lim_{n\to\infty} \int_X f_n\, d\mu_n = \int_X f\, d\mu. \tag{4.7.21}$$

(iii) Suppose that $f_n(x) \to f(x)$ μ-a.e., $f_n \in L^1(\mu_n)$ and that, for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\biggl|\int_E f_n\, d\mu_n\biggr| < \varepsilon \quad \text{whenever } E \in \mathcal{A} \text{ and } \mu_n(E) < \delta.$$

Prove that $f \in L^1(\mu)$ and (4.7.21) holds.

(iv) Deduce from (ii) that if the functions f_n are nonnegative and converge μ-a.e. to f, then

$$\int_X f\, d\mu \le \liminf_{n\to\infty} \int_X f_n\, d\mu_n.$$

(v) Suppose that the functions f_n converge μ-a.e. to f and that there exist $\mathcal{A}$-measurable functions g_n convergent μ-a.e. to a function g such that $|f_n| \le g_n$, $g_n \in L^1(\mu_n)$, $g \in L^1(\mu)$, and

$$\int_X g\, d\mu = \lim_{n\to\infty} \int_X g_n\, d\mu_n.$$

Deduce from (iv) that (4.7.21) holds.

Hint: in (i)–(iii) use Egoroff's theorem and the uniform absolute continuity of the measures μ_n. In (iv) consider $\min(f_n, k)$ with fixed k and let $k \to \infty$. In (v) consider the nonnegative functions $g_n - f_n$ and $g_n + f_n$.

4.7.133. (Areshkin, Klimkin [**35**]) Suppose that a sequence of measures μ_n on a σ-algebra $\mathcal{A}$ converges on every set in $\mathcal{A}$ to a measure μ and let $\mathcal{A}$-measurable functions f_n converge pointwise to a function f, where $f_n \in \mathcal{L}^1(\mu_n)$. Prove that the following conditions are equivalent:

(a) $f \in \mathcal{L}^1(\mu)$ and

$$\int_A f\, d\mu = \lim_{n\to\infty} \int_A f_n\, d\mu_n \quad \text{for every } A \in \mathcal{A};$$

(b) for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\biggl|\int_A f_n\, d\mu_n\biggr| \le \varepsilon \quad \text{whenever } A \in \mathcal{A} \text{ and } |\mu_n|(A) \le \delta.$$

Hint: take a probability measure ν such that $\mu_n = g_n \cdot \nu$, $\mu = g \cdot \nu$. If (a) is fulfilled, then we can use the uniform ν-integrability of $\{f_n g_n\}$ and Egoroff's theorem for ν. If we have (b), then one can use the uniform ν-integrability of $\{g_n\}$.

4.7.134. (Gowurin [**376**]) Let $\mathcal{X}$ be the space of all equivalence classes of Lebesgue measurable sets in $[0,1]$ equipped with the metric $d(A,B) = \lambda(A \triangle B)$, where λ is Lebesgue measure. Let $S(E_0, r) = \{E \in \mathcal{X}\colon d(E, E_0) = r\}$ be the sphere of radius $r \in (0,1)$ with the center $E_0 \in \mathcal{X}$. Suppose that this sphere does not contain the element corresponding to the empty set. Prove that if $f_n \in L^1[0,1]$ and

$$\lim_{n\to\infty} \int_E f_n\, dx = 0 \quad \text{for all } E \in S(E_0, r),$$

then the same is true for every measurable set $E \subset [0,1]$.

4.7.135. (S. Saks, see [**376**]) Prove that the class of all open sets is a first category set (a countable union of nowhere dense sets) in the space $\mathcal{X}$ from the previous exercise.

HINT: let $\{U_n\}$ be all intervals (open, semi-open or closed) in $[0,1]$ with rational endpoints. Every open set in $[0,1]$ is a finite or countable union of disjoint intervals in $\{U_n\}$. For fixed $k \in \mathbb{N}$, we consider the class M_k of all open sets $U \subset [0,1]$ such that there exist $U_{n_1}, \ldots, U_{n_k} \subset U$ with $\lambda\bigl(U \backslash \bigcup_{i=1}^k U_{n_i}\bigr) \le \lambda(U)/4$. The set M_k is nowhere dense in $\mathcal{X}$. Indeed, it is easily verified that given an open ball $B(C,r) \subset \mathcal{X}$ of radius $r > 0$ with the center C that is represented as a finite union of $p = 2q \ge 8k$ equal intervals $U_{m_1}, \ldots, U_{m_p}$ with disjoint closures, one can find $\delta > 0$ such that the ball $B(C,\delta)$ does not meet M_k. To this end, we take $\delta < \lambda(C)/4$ smaller than the minimal distance between the intervals $U_{m_1}, \ldots, U_{m_p}$ constituting C. If $U \in M_k$ belongs to this ball, then we take intervals $U_{n_1}, \ldots, U_{n_k} \subset U$ such that the measure of their union is at least $3\lambda(U)/4$. Clearly, each U_{n_i} cannot meet more than one U_{m_j}, since otherwise U would contain an interval of length greater than δ contrary to the estimate $\lambda(U \triangle C) < \delta$. Therefore, more than q intervals U_{m_j} do not meet $\bigcup_{i=1}^k U_{n_i}$, which shows that $\lambda\bigl(C \triangle \bigcup_{i=1}^k U_{n_i}\bigr) \ge \lambda(C)/2$. Hence $\lambda(C \triangle U) \ge \lambda(C)/4$, a contradiction. Clearly, every open ball in $\mathcal{X}$ contains a point C of the indicated form with a sufficiently large p.

4.7.136. Verify the equivalence of (i) and (ii) in Theorem 4.7.27.

HINT: Let (i) be fulfilled, but (ii) not. Then there exist disjoint sets $R_n \in \mathfrak{R}$ and measures μ_n in the given family with $|\mu_n(R_n)| \ge \varepsilon > 0$. Set $S_n = \bigcup_{k=n}^\infty R_k$. Take an increasing sequence of indices n_k such that $|\mu_{n_k}|(S_{n_{k+1}}) < \varepsilon/2$. Hence

$$\Bigl|\sum_{i=k}^\infty \mu_{n_k}(R_{n_i})\Bigr| \ge |\mu_{n_k}(R_{n_k})| - |\mu_{n_k}|\Bigl(\bigcup_{i=k+1}^\infty R_{n_i}\Bigr) > \varepsilon/2,$$

contrary to condition (i). Conversely, if (ii) is fulfilled, but (i) is not, then there exist disjoint $R_n \in \mathfrak{R}$ and $\varepsilon > 0$ such that, for each k, there is a number $n(k)$ with $\bigl|\sum_{j=k}^\infty \mu_{n(k)}(R_j)\bigr| > \varepsilon$. By using that $|\mu_n|\bigl(\bigcup_{j=m}^\infty R_j\bigr) \to 0$ as $m \to \infty$, we pick strictly increasing numbers m_k and p_k such that one has $m_k < p_k < m_{k+1}$ and $\bigl|\mu_{n_k}\bigl(\bigcup_{j=m_k}^{p_k-1} R_j\bigr)\bigr| > \varepsilon/2$, which contradicts (ii).

4.7.137. Let μ_n be real measures of bounded variation on the σ-ring $\mathfrak{S}$ generated by a ring $\mathfrak{R}$. Suppose that $\lim\limits_{n\to\infty} \mu_n(R_n) = 0$ for every infinite sequence of disjoint sets $R_n \in \mathfrak{R}$.

(i) Let $A_k = \bigcup_{j=1}^\infty A_j^k$, $B_k = \bigcup_{j=1}^\infty B_j^k$, where $A_j^k, B_j^k \in \mathfrak{R}$, and let the sets $E_k = A_k \backslash B_k$ be pairwise disjoint. Prove that $\lim\limits_{n\to\infty} |\mu_n|(E_n) = 0$.

(ii) Prove that, for every $S \in \mathfrak{S}$ and $\varepsilon > 0$, there exists a set R of the form $R = \bigcup_{j=1}^\infty R_j$ with $R_j \in \mathfrak{R}$ such that $|\mu_n|(S \triangle R) < \varepsilon$ for all n.

HINT: (i) otherwise we may assume that $|\mu_n|(E_n) \ge \varepsilon > 0$. The sets A_j^k can be made disjoint for every fixed k. The same can be done with the sets B_j^k. By Exercise 4.7.136, there exist indices p_k such that

$$|\mu_n|\Bigl(\bigcup_{j=p_k+1}^\infty A_j^k\Bigr) < \varepsilon 2^{-k}/8, \quad |\mu_n|\Bigl(\bigcup_{j=p_k+1}^\infty B_j^k\Bigr) < \varepsilon 2^{-k}/8 \quad \text{for all } n.$$

Let $C_k = \bigl(\bigcup_{j=1}^{p_k} A_j^k\bigr)\backslash\bigl(\bigcup_{j=1}^{p_k} B_j^k\bigr)$. Then

$$C_k \in \mathfrak{R}, \quad C_k \triangle E_k \subset \Bigl(\bigcup_{j=p_k+1}^{\infty} A_j^k\Bigr)\backslash\Bigl(\bigcup_{j=p_k+1}^{\infty} B_j^k\Bigr),$$

whence $|\mu_n|(C_k \triangle E_k) < \varepsilon 2^{-k}/4$ for all n and k. It is clear by the definition of C_k that $C_i \cap C_j \subset (C_i \triangle E_i) \cup (C_j \triangle E_j)$. Therefore, for all n, i, j one has

$$|\mu_n|(C_i \cap C_j) < \frac{\varepsilon}{4}(2^{-i} + 2^{-j}). \tag{4.7.22}$$

Let us consider the sets $D_n = C_n \backslash \bigcup_{j=1}^{n-1} C_j$, where $C_0 = \varnothing$. Then

$$C_n \triangle D_n = \bigcup_{j=1}^{n-1} (C_n \cap C_j)$$

and by (4.7.22) we obtain $|\mu_n|(C_n \triangle D_n) < \varepsilon/2$. Hence $|\mu_n|(E_n \triangle D_n) < 3\varepsilon/4$. Therefore, $|\mu_n|(D_n) > \varepsilon/4$, which leads to a contradiction.

(ii) Let $\nu_n = |\mu_1| + \cdots + |\mu_n|$. One can find sets $E_n \in \mathfrak{R}$ with

$$\nu_n(S \triangle E_n) < \varepsilon 2^{-n}/4, \quad n \in \mathbb{N}.$$

Let $D_n = \bigcup_{j=n}^{\infty} E_j$. The sets D_n are decreasing to $\varnothing$, and the sets $D_n \backslash D_{n+1}$ are disjoint and have the form indicated in (i). It is easy to deduce from assertion (i) that there exists p such that $|\mu_n|(D_p \backslash D_n) < \varepsilon/2$ for all $n > p$. Indeed, otherwise we find numbers $p_1 < n_1 < p_2 < n_2 < \ldots$ such that $|\mu_{n_j}|(D_{p_j} \backslash D_{n_j}) \geq \varepsilon/2$, which contradicts (i). If $n \leq p$, then we have $|\mu_n|(S \triangle D_p) < \varepsilon/4$. If $n > p$, then we obtain

$$|\mu_n|(S \triangle D_p) \leq |\mu_n|(S \triangle D_n) + |\mu_n|(D_n \triangle D_p) < \varepsilon.$$

The set D_p has the required form.

4.7.138. (Dubrovskiĭ [**249**]). Let $\{\varphi_\alpha\}$ be a uniformly bounded family of countably additive measures on a σ-algebra $\mathcal{M}$ dependent on the parameter α from some set A. For every sequence of disjoint sets $E_n \in \mathcal{M}$ we let $\delta(\{E_n\}) = \lim\limits_{n\to\infty} \Bigl[\sup\limits_{\alpha \in A} |\varphi_\alpha|\bigl(\bigcup_{k=n+1}^{\infty} E_k\bigr)\Bigr]$. Denote by Δ the supremum of the numbers $\delta(\{E_n\})$ over all possible sequences of the indicated type. Suppose that there exists a nonnegative measure μ on $\mathcal{M}$ such that $\varphi_\alpha \ll \mu$ for all α. Set $f_\alpha := d\varphi_\alpha/d\mu$. Prove that Δ coincides with the quantity

$$\lim_{N\to\infty} \Bigl[\sup_{\alpha \in A} \int_{\{|f_\alpha| > N\}} \bigl[|f_\alpha| - N\bigr]\, d\mu\Bigr].$$

In particular, the latter is independent of μ.

4.7.139. (M.N. Bobynin, E.H. Gohman) Suppose $\mathcal{A}$ is a σ-algebra, $A_n \in \mathcal{A}$, $A_{n+1} \subset A_n$ and $\bigcap_{n=1}^{\infty} A_n = \varnothing$. Let μ_n be measures on $\mathcal{A}$ (possibly signed or complex-valued) such that $\mu_n(A_n) \neq 0$ for all n. Prove that there exists a set $A \in \mathcal{A}$ such that one has $|\mu_n(A)| > \frac{1}{5}|\mu_n(A_n)|$ for infinitely many indices n.

Hint: see Bobynin [**100**, Lemma 1].

4.7.140. Let μ and ν be bounded measures on a σ-algebra $\mathcal{A}$. Show that

$$\mu \vee \nu(A) = \sup\{\mu(B) + \nu(A\backslash B)\colon\ B \in \mathcal{A}, B \subset A\}, \quad \forall\, A \in \mathcal{A},$$

$$\mu \wedge \nu(A) = \inf\{\mu(B) + \nu(A\backslash B)\colon\ B \in \mathcal{A}, B \subset A\}, \quad \forall\, A \in \mathcal{A}.$$

HINT: let $\mu = f \cdot \lambda$, $\nu = g \cdot \lambda$, where λ is a nonnegative measure; then the integral of $\max(f, g)$ over A with respect to λ equals the sum of the integral of f over $A \cap \{f \geq g\}$ and the integral of g over $A \cap \{f < g\}$; similarly for $\mu \wedge \nu$.

4.7.141. Let μ be a nonnegative measure and let $f, g \in L^p(\mu)$, $1 < p < \infty$. Show that the function

$$F(t) = \int |f + tg|^p \, d\mu$$

is differentiable and

$$F'(0) = p \int |f|^{p-2} f g \, d\mu.$$

4.7.142. Let $1 < p < \infty$. Show that, for every $\varepsilon > 0$, there exists $\delta > 0$ such that if $f, g \in L^p[0,1]$, $\|g\|_p = 1$, $\|f\|_p \leq \delta$, and

$$\int f(x) \, dx = 0,$$

then

$$\iint |f(x) + g(y)|^p \, dx \, dy \leq 1 + \varepsilon \|f\|_p.$$

HINT: see Fremlin [**327**, §273M].

4.7.143. (Carlen, Loss [**167**]) Let μ be a probability measure on a space X and let $u \in L^2(\mu)$ have unit $L^2(\mu)$-norm and zero integral.

(i) Prove that for every $\alpha \in [0,1]$ and $p \geq 2$, letting $f = \alpha u + \sqrt{1-\alpha^2}$, one has

$$\|f\|^p_{L^p(\mu)} \leq (1-\alpha)^{p/2} + \frac{\alpha^2 p(p-1)}{2} \|f\|^{p-2}_{L^p(\mu)} \|u\|^2_{L^p(\mu)},$$

provided that $u \in L^p(\mu)$.

(ii) Let $u^2 \ln(u^2) \in L^1(\mu)$. Prove that

$$\int_X f^2 \ln(f^2) \, d\mu \leq 2\alpha^2 + \alpha^4 + \alpha^2 \int_X u^2 \ln(u^2) \, d\mu.$$

4.7.144. (i) (Clarkson [**183**]) Prove the following inequalities for $f, g \in L^p(\mu)$:

$$\Bigl\|\frac{f+g}{2}\Bigr\|_p^p + \Bigl\|\frac{f-g}{2}\Bigr\|_p^p \leq \frac{1}{2}\|f\|_p^p + \frac{1}{2}\|g\|_p^p, \ 2 \leq p < \infty,$$

$$\Bigl\|\frac{f+g}{2}\Bigr\|_p^{p'} + \Bigl\|\frac{f-g}{2}\Bigr\|_p^{p'} \leq \Bigl[\frac{1}{2}\|f\|_p^p + \frac{1}{2}\|g\|_p^p\Bigr]^{\frac{1}{p-1}}, \ 1 < p \leq 2, p' = \frac{p}{p-1}.$$

(ii) (Hanner [**407**]) Prove the following inequalities for $f, g \in L^p(\mu)$, where $1 \leq p \leq 2$:

$$\|f+g\|_p^p + \|f-g\|_p^p \geq (\|f\|_p + \|g\|_p)^p + \bigl|\|f\|_p - \|g\|_p\bigr|^p,$$

$$(\|f+g\|_p + \|f-g\|_p)^p + \bigl|\|f+g\|_p - \|f-g\|_p\bigr|^p \leq 2^p(\|f\|_p + \|g\|_p)^p.$$

Prove the reversed inequalities in the case $2 \leq p < \infty$.

HINT: (i) see Sobolev [**893**, Ch. III, §7], where one can find a generalization, and Hewitt, Stromberg [**431**, Ch. 4, §15]; (ii) see Lieb, Loss [**612**, §2.5].

4.7.145. (Douglas [**235**]) Suppose that $(X, \mathcal{A})$ is a measurable space, $\mathcal{M}^+(\mathcal{A})$ is the set of all finite nonnegative measures on $\mathcal{A}$, $\mathcal{F}$ is some linear space of real $\mathcal{A}$-measurable functions. Let $\mu \in \mathcal{M}^+(\mathcal{A})$ and $\mathcal{F} \subset \mathcal{L}^1(\mu)$. Set

$$E^\mu := \Bigl\{\nu \in \mathcal{M}^+(\mathcal{A}) \colon \ \mathcal{F} \subset \mathcal{L}^1(\nu), \int f \, d\nu = \int f \, d\mu \quad \text{for all } f \in \mathcal{F}\Bigr\}.$$

(i) Prove that $\mathcal{F}$ is dense in $L^1(\mu)$ precisely when μ is an extreme point in E^μ, i.e., there are no measures $\mu_1, \mu_2 \in E^\mu$ and $t \in (0,1)$ such that one has $\mu_1 \neq \mu$ and $\mu = t\mu_1 + (1-t)\mu_2$.

(ii) Let $\mathcal{B}$ be a sub-σ-algebra in $\mathcal{A}$. Prove that μ is an extreme point in the set of all measures $\nu \in \mathcal{M}^+(\mathcal{A})$ such that $\nu|_{\mathcal{B}} = \mu|_{\mathcal{B}}$ precisely when, for every $A \in \mathcal{A}$, there exists $B \in \mathcal{B}$ with $\mu(A \triangle B) = 0$.

HINT: (i) if $\mathcal{F}$ is not dense, then there exists $g \in L^\infty(\mu)$ with $0 < \|g\| \leq 1$ and

$$\int gf \, d\mu = 0 \quad \text{for all } f \in \mathcal{F}.$$

Then $\mu = (\mu_1 + \mu_2)/2$, where $\mu_1 := (1+g) \cdot \mu \in E^\mu$, $\mu_2 := (1-g) \cdot \mu \in E^\mu$. Conversely, let $\mu = t\mu_1 + (1-t)\mu_2$, where $t \in (0,1)$, $\mu_i \in E^\mu$. Then $\mu_i \ll \mu$, hence $\mu_i = g_i \cdot \mu$, $g_i \in L^1(\mu)$. Since the integrals of the functions $(g_1 - 1)f$, where $f \in \mathcal{F}$, against the measure μ vanish, it is easily verified that $g_1 - 1$ belongs to the closure of $\mathcal{F}$ only in the case $g_1 = 1$. Then $g_2 = 1$. (ii) One can take for $\mathcal{F}$ the space of all bounded $\mathcal{B}$-measurable functions. It is dense in $L^1(\mu)$ precisely when $\mathcal{B}$ is dense in the measure algebra $\mathcal{A}/\mu$.

4.7.146. Let X be an infinite-dimensional normed space. (i) Prove that the weak topology on any ball is strictly weaker than the norm topology. (ii) Prove that X with the weak topology is not metrizable.

HINT: (i) every weak neighborhood of the center meets the sphere. (ii) If a metric d generates the weak topology, then the balls $\{x\colon d(x,0) < n^{-1}\}$ contain neighborhoods $U(0, l_{n,1}, \ldots, l_{n,k_n}, \varepsilon_n)$ with $l_{n,i} \in X^*$. Hence X^* is the linear span of all $l_{n,i}$, which is impossible because X^* is a Banach space.

4.7.147. Prove that every weakly compact set in l^1 is norm compact.
HINT: apply the results of §4.7(iv).

4.7.148. (i) Let μ be a separable finite nonnegative measure. Show that every uniformly integrable subset of $L^1(\mu)$ is metrizable in the weak topology. In particular, every weakly compact subset of $L^1(\mu)$ is metrizable in the weak topology.

(ii) Let $\mathcal{A}$ be a countably generated σ-algebra. Show that every compact subset of the space $\mathcal{M}$ of all bounded measures on $\mathcal{A}$ with the setwise convergence topology is metrizable.

HINT: (i) M has compact closure K in the weak topology; there is a countable family $\{\varphi_n\} \subset \mathcal{L}^\infty(\mu)$ with the following property: if $f, g \in \mathcal{L}^1(\mu)$ are such that the integrals of $f\varphi_n$ and $g\varphi_n$ are equal for all n, then $f = g$ a.e. The functions

$$f \mapsto \int f\varphi_n \, d\mu$$

are continuous on K in the weak topology and separate the points. Hence K is metrizable (see Exercise 6.10.24 in Chapter 6). (ii) The same reasoning applies with the functions $\mu \mapsto \mu(A_n)$ on $\mathcal{M}$, where a countable family $\{A_n\}$ generates $\mathcal{A}$.

4.7.149. (i) Let $f \in \mathcal{L}^2(\mathbb{R}^1)$. Show that the set $\mathcal{F}$ of all functions of the form $\sum_{k=1}^n c_k f(x + \delta_k)$, where $n \in \mathbb{N}$, $c_k, \delta_k \in \mathbb{R}^1$, is everywhere dense in $L^2(\mathbb{R}^1)$ precisely when the set of zeros of the Fourier transform of f has measure zero.

(ii) Let $f \in \mathcal{L}^1(\mathbb{R}^1)$. Show that the set $\mathcal{F}$ indicated in (i) is everywhere dense in $L^1(\mathbb{R}^1)$ precisely when the Fourier transform of the function f does not vanish.

(iii) (Segal [**860**]) Show that if $1 < p < 2$, then the a.e. positivity of the Fourier transform of a function $f \in L^p(\mathbb{R}^1)$ does not imply that the set indicated in (i) is everywhere dense in $L^p(\mathbb{R}^1)$.

HINT: (i) observe that the Fourier transform of the indicated sum is the function $\sum_{k=1}^n c_k \exp(-i\delta_k x)\widehat{f}(x)$. If $f \in \mathcal{L}^2(\mathbb{R}^1)$ and $\widehat{f} = 0$ on a compact set A of positive measure, then the inverse Fourier transform of the function I_A is orthogonal to $\mathcal{F}$. Let $\widehat{f} \neq 0$ a.e. If $\mathcal{F}$ is not dense, then there exists a nontrivial function $g \in \mathcal{L}^2(\mathbb{R}^1)$ that is orthogonal to all shifts $f(\,\cdot - y)$, $y \in \mathbb{R}^1$. By the Parseval equality for the Fourier transform in L^2, the Fourier transform of the function $\widehat{f}\overline{\widehat{g}}$ vanishes, hence $g = 0$, which is a contradiction. In (ii), a similar reasoning applies.

4.7.150.° Suppose that a sequence of functions $f_n \in L^1(\mu)$ converges weakly to a function f and a sequence of functions $g_n \in L^1(\mu)$ converges weakly to a function g and $|f_n(x)| \le g_n(x)$ for all n. Show that $|f(x)| \le g(x)$ a.e. Construct an example demonstrating that the estimates $|f_n(x)| \le |g_n(x)|$ do not imply that $|f(x)| \le |g(x)|$ a.e.

HINT: for any measurable set A, one has

$$\int_A |f|\, d\mu = \lim_{n\to\infty} \int_A f_n \mathrm{sign} f\, d\mu,$$

which is estimated by the integral of gI_A. To construct an example take $[0,1]$ with Lebesgue measure, set $f_n = 1$ and choose a sequence of functions g_n with $|g_n(x)| = 1$ that is weakly convergent to zero.

4.7.151.° Let μ be a bounded nonnegative Borel measure on an open cube in V in $\mathbb{R}^n$. Show that the set $C_0^\infty(V)$ of infinitely differentiable functions with support in V is everywhere dense in $L^p(\mu)$, $1 \le p < \infty$.

HINT: it suffices to approximate the indicators of cubes $K \subset V$; given $\varepsilon > 0$ there are a closed cube $Q \subset K$ and an open cube U with $K \subset U \subset \overline{U} \subset V$ and $\mu(U\backslash Q) < \varepsilon$. Take $f \in C_0^\infty(V)$ such that $0 \le f \le 1$, $f|_Q = 1$, $f = 0$ outside U.

4.7.152. Let μ be a probability measure and let M be a convex set in $L^1(\mu)$ that consists of probability densities and is closed with respect to convergence in measure. Show that M is compact in the weak topology.

HINT: it suffices to show that M is uniformly integrable. If not, by Corollary 4.7.21 one can find decreasing measurable sets A_n with empty intersection and functions f_n in M such that, for some $\alpha > 0$, the integral of f_n over A_n is greater than α for every n. By the Komlós theorem we obtain a sequence $S_k := (f_{n_1} + \cdots + f_{n_k})/k$ that converges a.e. to some f. Then $f \in M$ by hypothesis, hence $S_k \to f$ in $L^1(\mu)$. The integral of S_k over A_{n_k} is greater than or equal to α. Since the integrals of f over A_n tend to zero, one arrives at a contradiction.

4.7.153. Let $(X, \mathcal{A}, \mu)$ be a probability space and let $\{f_n\}$ be a sequence of probability densities convergent μ-a.e. to a function f. Let $\Lambda \in L^\infty(\mu)^*$ be a limit point of $\{f_n\}$ in the $*$-weak topology of $L^\infty(\mu)^*$ (which exists by the Banach–Alaoglu theorem). Then Λ corresponds to a nonnegative additive set function Λ_0 on $\mathcal{A}$. Show that $\Lambda_0 = f \cdot \mu + \Lambda_a$, where Λ_a is a nonnegative additive function on $\mathcal{A}$ without σ-additive component.

HINT: we know that $\Lambda_0 = \Lambda_a + \nu$, where ν is a nonnegative σ-additive measure on $\mathcal{A}$ and Λ_a is a nonnegative additive set function on $\mathcal{A}$ without σ-additive component. Clearly, $\nu \ll \mu$, hence $\nu = \varrho \cdot \mu$, where $\varrho \ge 0$ is μ-integrable. For any

$A \in \mathcal{A}$, we have

$$\int_A f\,d\mu \le \liminf_{n\to\infty} \int_A f_n\,d\mu \le \Lambda_0(A),$$

i.e., $f \cdot \mu \le \Lambda_0$. On the other hand, By Egoroff's theorem, given $\varepsilon > 0$ we can find a set $E \subset X$ such that $[(f+\varrho)\cdot\mu](X\backslash E) < \varepsilon$ and $\{f_n\}$ converges to f uniformly on E. Hence for every set $A \in \mathcal{A}$ contained in E one has

$$\lim_{n\to\infty} \int_A f_n\,d\mu = \int_A f\,d\mu.$$

Note that the left-hand side equals $\Lambda_0(A)$. Hence the restriction of Λ_a to E coincides with the measure $f\cdot\mu - \varrho\cdot\mu$, which means that this restriction vanishes. Thus, $f(x) = \varrho(x)$ for μ-a.e. $x \in E$. This yields that $f = \varrho$ μ-a.e.

4.7.154. Construct probability densities f_n on $[0,1]$ with Lebesgue measure λ that converge to 0 in measure but where the constant function 1 belongs to the closure of $\{f_n\}$ in the weak topology $\sigma(L^1, L^\infty)$. In particular, in the previous exercise, one cannot replace convergence almost everywhere by convergence in measure.

HINT: for every $n \in \mathbb{N}$, we partition $[0,1]$ into 4^n equal intervals $J_{n,k}$. Let $c_{n,m} \in [0, 4^n]$ be such that $c_{n,m+1} - c_{n,m} = 8^{-n}$, $c_{n,1} = 0$, $m \le (32)^n + 1$. For each n, denote by $\mathcal{F}_n$ the collections of all functions f on $[0,1]$ that are constant on each $J_{n,k}$, assume only values $c_{n,m}$, have integral 1, and satisfy the condition $\lambda(\{f > 0\}) \le 2^{-n}$. Clearly, $\mathcal{F}_n$ is finite. Next we write the functions from all $\mathcal{F}_n$ in a single sequence $\{f_n\}$ such that the elements of $\mathcal{F}_{n+1}$ follow the elements of $\mathcal{F}_n$. By construction, $f_n \to 0$ in measure. Let us show that every neighborhood U of 1 in the topology $\sigma(L^1, L^\infty)$ contains a function from $\{f_n\}$ distinct from 1. We may assume that

$$U = \Bigl\{\varphi\colon \Bigl|\int_0^1 \psi_i(\varphi - 1)\,dx\Bigr| < \varepsilon,\ i = 1, \ldots, n\Bigr\},$$

where the functions ψ_i assume finitely many values. This can be easily reduced to the case where each ψ_i is the indicator function of a measurable set A_i of positive measure and the sets A_i are pairwise disjoint. In that case, in each A_i we pick a density point a_i, i.e., letting $\Delta_i = [a_i - \delta, a_i + \delta]$, one has $\lim\limits_{\delta\to 0} \lambda(A_i \cap \Delta_i)/\lambda(\Delta_i) = 1$ (see Chapter 5). We can assume that $\varepsilon < 1/2$ and $n > 1$. Let us take $\delta < \varepsilon n^{-1}/2$ such that the intervals Δ_i are disjoint and $\lambda(A_i \cap \Delta_i) > (1 - \varepsilon/4)\lambda(\Delta_i)$. Next we observe that each Δ_i can be replaced by a slightly smaller interval $E_i \subset \Delta_i$ such that E_i is a finite union of some of the intervals $J_{m,k}$, where $2^{-m} < \varepsilon(4n)^{-1}$ and m is common for all $i = 1, \ldots, n$, and $\lambda(A_i \cap E_i) > (1 - \varepsilon/4)\lambda(E_i)$. For every i, one can find $c_i \in \{c_{m,1,}, \ldots, c_{m,8^m+1}\}$ such that $|c_i\lambda(E_i \cap A_i) - \lambda(A_i)| < \varepsilon(4n)^{-1}$. This is possible because $\lambda(E_i \cap A_i) > 2\lambda(E_i) > 4^{1-m}$, $\lambda(A_i)/\lambda(E_i \cap A_i) < 4^{m-1}$, $\varepsilon(4n)^{-1}/\lambda(E_i \cap A_i) > 2^{-m}4^{1-m} > 8^{-m}$. Finally, let $f = c_1 I_{E_1} + \cdots + c_n I_{E_n}$. Clearly, $f \in \{f_n\}$. We show that $f \in U$. We have the estimates $c_i\lambda(E_i) < 2c_i\lambda(A_i \cap E_i) \le \varepsilon(2n)^{-1} + 2\lambda(A_i)$. Note that for every $j \ne i$ one has $\lambda(E_j \cap A_i) < \varepsilon\lambda(E_j)/4$, since $A_j \cap A_i = \varnothing$ and $\lambda(E_j \cap A_j) > (1 - \varepsilon/4)\lambda(E_j)$. Therefore, we arrive at the estimates $c_j\lambda(E_j \cap A_i) < \varepsilon c_j\lambda(E_j)/4 < \varepsilon/(8n) + \varepsilon\lambda(A_j)/2$. This gives the inequality $|c_i\lambda(E_i \cap A_i) - \lambda(A_i)| + \sum_{j \ne i} c_j\lambda(E_j \cap A_i) < \varepsilon$. Thus, $f \in U$.

CHAPTER 5

Connections between the integral and derivative

> All those who wrote on the theory of functions of a real variable know well how difficult it is to be simultaneously rigorous and brief in such matters.
>
> *N.N. Lusin.*

5.1. Differentiability of functions on the real line

Let us recall that a function f defined in a neighborhood of a point $x \in \mathbb{R}^1$ is called differentiable at this point if there exists a finite limit

$$\lim_{h\to 0} \frac{f(x+h)-f(x)}{h},$$

which is called the derivative of f at the point x and denoted by $f'(x)$. The developments of mathematical analysis, in particular, the integration theory, are closely connected with the problem of recovering a function from its derivative. The fundamental theorem of calculus — the Newton–Leibniz formula — recovers a function f on $[a,b]$ from its derivative f':

$$f(x) = f(a) + \int_a^x f'(y)\,dy. \tag{5.1.1}$$

For continuously differentiable functions f the integral in formula (5.1.1) exists in Riemann's sense, hence there is no problem in interpreting this identity. The problems do appear when one attempts to extend the Newton–Leibniz formula to broader classes of functions. There are essentially three problems: in what sense the derivative exists, in what sense it is integrable, and, finally, if it exists in a certain sense and is integrable, then is equality (5.1.1) true? In order to show the character of potential difficulties, we consider several examples. First we construct a function f that is differentiable at every point of the real line, but f' is not Lebesgue integrable on $[0,1]$.

5.1.1. Example. Let

$$f(x) = \begin{cases} x^2 \sin\frac{1}{x^2} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

The function f is everywhere differentiable, but the function f' is not Lebesgue integrable on $[0,1]$.

PROOF. The differentiability of f outside the origin is obvious and the equality $f'(0) = 0$ follows from the definition by the boundedness of sine. One has

$$f'(x) = 2x \sin\frac{1}{x^2} - 2\frac{1}{x}\cos\frac{1}{x^2}$$

if $x \neq 0$. It suffices to show that the function

$$\psi(x) = \frac{1}{x}\cos\frac{1}{x^2}$$

is not Lebesgue integrable on $[0,1]$. Suppose the contrary. Then the function $\frac{1}{x}\cos\frac{1}{2x^2}$ is integrable as well, which is verified by using the change of variable $y = \sqrt{2}x$. Therefore, the function

$$\varphi(x) = \frac{1}{x}\cos^2\frac{1}{2x^2}$$

is integrable, too. Since $\psi(x) = 2\varphi(x) - x^{-1}$, we obtain the integrability of x^{-1}, which is a contradiction. □

The function f' in the above example is integrable in the improper Riemann sense. However, it is now easy to destroy this property as well. Let us take a compact set $K \subset [0,1]$ of positive Lebesgue measure without inner points (see Example 1.7.6). The set $[0,1]\backslash K$ has the form $\bigcup_{n=1}^{\infty}(a_n, b_n)$, where the intervals (a_n, b_n) are pairwise disjoint. Let us take a differentiable function θ such that $\theta(x) = 1$ if $x \leq 1/2$ and $\theta(x) = 0$ if $x \geq 1$. Set $g(x) = \theta(x)f(x)$ if $x \geq 0$ and $g(x) = 0$ if $x < 0$. We observe that $g'(0) = g'(1) = 0$ and $|g(x)| \leq C\min\{x^2, (1-x)^2\}$ for some C.

5.1.2. Example. We define a function F by the formula

$$F(x) = \sum_{n=1}^{\infty}(b_n - a_n)^2 g\Bigl(\frac{x - a_n}{b_n - a_n}\Bigr).$$

The function F is everywhere differentiable and its derivative F' is not Lebesgue integrable on $[0,1]$ and is discontinuous at every point of the set K. In particular, F' has no improper Riemann integral on $[0,1]$.

PROOF. It is clear that the series defining the function F converges uniformly because the function g is bounded. It suffices to show that $F'(x) = 0$ at every point $x \in K$, since on the interval (a_n, b_n) the function F equals the function $(b_n - a_n)^2 g\bigl(x - a_n/(b_n - a_n)\bigr)$. By our construction, $F(x) = 0$ if $x \in K$. Let $h > 0$. If $x + h \in K$, then $F(x+h) - F(x) = 0$. If $x + h \notin K$, then we can find an interval (a_n, b_n) containing $x + h$. Then $x + h - a_n < h$ and hence

$$\Bigl|\frac{F(x+h) - F(x)}{h}\Bigr| = \Bigl|\frac{F(x+h)}{h}\Bigr| = (b_n - a_n)^2\frac{1}{h}\Bigl|g\Bigl(\frac{x+h-a_n}{b_n - a_n}\Bigr)\Bigr|$$
$$\leq \frac{(b_n - a_n)^2}{h}C\frac{h^2}{(b_n - a_n)^2} = Ch,$$

which tends to zero as $h \to 0$. The case $h < 0$ is similar. It is obvious that the function F' is unbounded in the right neighborhood of the point a_n, since F on (a_n, b_n) is an affine transformation of g on $(0, 1)$. Therefore, F' is discontinuous at every point in the closure of $\{a_n\}$. This closure coincides with K due to the absence of inner points of K. □

One can construct an everywhere differentiable function f such that its derivative is discontinuous almost everywhere (Exercise 5.8.119). However, f' cannot be discontinuous everywhere (Exercise 5.8.37). Finally, we observe that if f' exists everywhere and is finite, then it cannot be non-integrable on every interval, since there exists an interval where it is bounded (Exercise 5.8.37).

Thus, neither the Lebesgue integral nor the improper Riemann integral solve the problem of recovering an everywhere differentiable function from its derivative. In §5.7, we consider a more general (non-absolute) integral solving this problem (although not constructively). We remark, however, that in the applications of the theory of integration, much more typical is the problem of recovering functions that have derivatives only almost everywhere. Certainly, without additional assumptions, this is impossible. For example, the above-considered Cantor function (Proposition 3.6.5) has a zero derivative almost everywhere, but is not constant. Lebesgue described the class of all functions that are almost everywhere differentiable and can be recovered from their derivatives by means of the Newton–Leibniz formula for the Lebesgue integral. It turned out that these are precisely the absolutely continuous functions. Before discussing such functions, we shall consider a broader class of functions, which also are differentiable almost everywhere, but may not be indefinite integrals.

In the study of derivatives it is useful to consider the so called derivates of a function f that take values on the extended real line and are defined by the following equalities:

$$D^+ f(x) = \limsup_{h \to +0} \frac{f(x+h) - f(x)}{h},$$
$$D_+ f(x) = \liminf_{h \to +0} \frac{f(x+h) - f(x)}{h},$$
$$D^- f(x) = \limsup_{h \to -0} \frac{f(x+h) - f(x)}{h},$$
$$D_- f(x) = \liminf_{h \to -0} \frac{f(x+h) - f(x)}{h}.$$

If $D^+ f(x) = D_+ f(x)$, then we say that the function f has the right derivative $f'_+(x) := D^+ f(x) = D_+ f(x)$ at the point x, and if $D^- f(x) = D_- f(x)$, then we say that f has the left derivative $f'_-(x) := D^- f(x) = D_- f(x)$ at the point x. It is clear that the existence of a finite derivative of f at the point x is equivalent to the equality and finiteness at this point of the right and left derivatives.

The upper and lower derivatives $\overline{D}f(x)$ and $\underline{D}f(x)$ are defined, respectively, as the supremum and infimum of the ratio $[f(x+h)-f(x)]/h$ as $h \to 0$, $h \neq 0$.

5.1.3. Lemma. *For any function f on the interval $[a,b]$, the set of all points at which the right and left derivatives of f exist, but are not equal, is finite or countable.*

PROOF. Let $D := \{x\colon f'_-(x) < f'_+(x)\}$ and let $\{r_n\}$ be the set of all rational numbers. For any $x \in D$, there exists the smallest k such that $f'_-(x) < r_k < f'_+(x)$. Furthermore, there exists the smallest m such that $r_m < x$ and for all $t \in (r_m, x)$ one has
$$\frac{f(t)-f(x)}{t-x} < r_k.$$
Finally, there exists the smallest number n such that $r_n > x$ and for all $t \in (x, r_n)$ one has
$$\frac{f(t)-f(x)}{t-x} > r_k.$$
According to our choice of m and n we obtain
$$f(t)-f(x) > r_k(t-x) \quad \text{if } t \neq x \text{ and } t \in (r_m, r_n). \tag{5.1.2}$$
Thus, to every point $x \in D$ we associate a triple of natural numbers (k, m, n). Note that to distinct points we associate different triples. Indeed, suppose that to points x and y there corresponds one and the same triple (k, m, n). Taking $t = y$ in (5.1.2), we obtain $f(y) - f(x) > r_k(y-x)$. If in (5.1.2) in place of x we take y and set $t = x$, then we obtain the opposite inequality. Thus, D is at most countable. In a similar manner one verifies that the set $\{f'_+ < f'_-\}$ is at most countable. □

Completing this section we formulate the following remarkable theorem due to N. Lusin (see the proof in Bruckner [**135**, Ch. 8]; Lusin [**632**], [**633**], [**635**]; Saks [**840**, Ch. VII, §2]).

5.1.4. Theorem. *Let f be a measurable a.e. finite function on $[0,1]$. Then, there exists a continuous function F on $[0,1]$ such that F is differentiable a.e. and $F'(x) = f(x)$ a.e.*

5.2. Functions of bounded variation

5.2.1. Definition. *A function f on a set $T \subset \mathbb{R}^1$ is of bounded variation if one has*
$$V(f,T) := \sup \sum_{i=1}^{n} |f(t_{i+1}) - f(t_i)| < \infty,$$
where sup *is taken over all collections $t_1 \le t_2 \le \cdots \le t_{n+1}$ in T. The quantity $V(f,T)$ is called the variation of f on T. If $T = [a,b]$, then we set $V_a^b(f) := V(f,[a,b])$.*

If a function f is of bounded variation, then it is bounded and for any $t_0 \in T$ one has

$$\sup_{t\in T} |f(t)| \le |f(t_0)| + V(f, T).$$

We shall be mainly interested in the case where T is an interval $[a, b]$ or (a, b) (possibly unbounded).

The simplest example of a function of bounded variation is an increasing function f on $[a, b]$ (in the case of an unbounded interval it is additionally required that the limits at the endpoints be finite). Indeed, we have $V_a^b(f) = V(f, [a, b]) = f(b) - f(a)$. It is clear that any decreasing function is of bounded variation as well. The space $BV[a, b]$ of all functions of bounded variation is linear. In addition,

$$V_a^b(\alpha f + \beta g) \le |\alpha| V_a^b(f) + |\beta| V_a^b(g) \tag{5.2.1}$$

for any two functions f and g of bounded variation and arbitrary real numbers α and β. This is obvious from the estimate

$$\bigl|\alpha f(t_{i+1}) + \beta g(t_{i+1}) - \alpha f(t_i) - \beta g(t_i)\bigr| \\ \le |\alpha|\, |f(t_{i+1}) - f(t_i)| + |\beta|\, |g(t_{i+1}) - g(t_i)|.$$

Hence the difference of two increasing functions is a function of bounded variation. The converse is true as well.

5.2.2. Proposition. *Let f be a function of bounded variation on $[a, b]$. Then*:

(i) *the functions $V\colon\ x \mapsto V(f, [a, x])$ and $U\colon\ x \mapsto V(x) - f(x)$ are nondecreasing on $[a, b]$*;

(ii) *the function V is continuous at a point $x_0 \in [a, b]$ if and only if the function f is continuous at this point*;

(iii) *for every $c \in (a, b)$, one has*

$$V(f, [a, b]) = V(f, [a, c]) + V(f, [c, b]). \tag{5.2.2}$$

Proof. If we add a new point to a partition of $[a, b]$, the corresponding sum of the absolute values of the increments of the function does not decrease. Hence in the calculation of $V_a^b(f)$ we can consider only the partitions containing the point c. Then

$$V(f, [a, b]) = \sup\Bigl[\sum_{i=1}^{k} |f(t_{i+1}) - f(t_i)| + \sum_{i=k+1}^{n} |f(t_{i+1}) - f(t_i)|\Bigr],$$

where sup is taken over all partitions with $t_{k+1} = c$. This equality gives (5.2.2), whence it follows that V is a nondecreasing function. The function $U = V - f$ is nondecreasing as well, since whenever $x \ge y$ we have

$$V(x) - V(y) = V_y^x(f) \ge |f(x) - f(y)| \ge f(x) - f(y).$$

Then $|V(x) - V(y)| \ge |f(x) - f(y)|$, whence the continuity of f at every point of continuity of V follows at once. It remains to verify the continuity of V at the points x where f is continuous. Let $\varepsilon > 0$. We find $\delta_0 > 0$ with $|f(x+h)-f(x)| \le \varepsilon/2$ whenever $|h| \le \delta_0$. By definition, there exist partitions $a = t_1 \le \cdots \le t_{n+1} = x$ and $x = s_1 \le \cdots \le s_{n+1} = b$ such that

$$\Bigl|V(f,[a,x]) - \sum_{i=1}^{n} |f(t_{i+1}) - f(t_i)|\Bigr| + \Bigl|V(f,[x,b]) - \sum_{i=1}^{n} |f(s_{i+1}) - f(s_i)|\Bigr| \le \frac{\varepsilon}{2}.$$

Let $|h| < \delta := \min(\delta_0, x - t_n, s_2 - x)$ and $h > 0$. Then

$$\begin{aligned} V(x+h) - V(x) = V_x^{x+h}(f) &= V(f,[x,b]) - V(f,[x+h,b]) \\ &\le \sum_{i=1}^{n} |f(s_{i+1}) - f(s_i)| + \frac{\varepsilon}{2} - V(f,[x+h,b]) \\ &\le |f(x) - f(x+h)| + |f(x+h) - f(s_2)| \\ &\quad + \sum_{i=2}^{n} |f(s_{i+1}) - f(s_i)| + \frac{\varepsilon}{2} - V(f,[x+h,b]) \\ &\le |f(x) - f(x+h)| + \frac{\varepsilon}{2} \le \varepsilon, \end{aligned}$$

since $|f(x+h)-f(s_2)| + \sum_{i=2}^{n} |f(s_{i+1})-f(s_i)| \le V_{x+h}^{b}(f)$. A similar estimate holds for $h < 0$. □

The variation of a function f may not be an additive set function. For example, $V(f,[0,1]) = 1 > V(f,[0,1)) = 0$ if $f(x) = 0$ on $[0,1)$ and $f(1) = 1$.

5.2.3. Corollary. *A continuous function of bounded variation is the difference of two continuous nondecreasing functions.*

5.2.4. Corollary. *Every function of bounded variation has at most countably many points of discontinuity.*

Proof. By the above proposition, it is sufficient to consider a nondecreasing function f. In this case, the points of discontinuity are exactly the points x such that $\lim\limits_{h\to 0+} f(x-h) < \lim\limits_{h\to 0+} f(x+h)$. It is clear that they are at most countably many. □

In the proof of the following important theorem we employ a technical lemma, which can be easily obtained from considerably more general results in §5.5. In order not to break our order of exposition, we give a direct proof of the necessary lemma.

5.2.5. Lemma. *Let E be a set in $(0,1)$. Suppose that we are given some family $\mathcal{I}$ of open intervals such that for every $x \in E$ and every $\delta > 0$, it contains an interval $(x, x+h)$ with $h < \delta$. Then, for every $\varepsilon > 0$, one can find a finite subfamily of disjoint intervals $I_1, \ldots, I_k$ in this family such that*

$$\lambda\Bigl(\bigcup\nolimits_{j=1}^{k} I_j\Bigr) < \lambda^*(E) + \varepsilon \quad \textit{and} \quad \lambda^*\Bigl(E \bigcap \bigcup\nolimits_{j=1}^{k} I_j\Bigr) > \lambda^*(E) - \varepsilon.$$

In addition, given an open set U containing E, such intervals can be taken inside U.

PROOF. We find an open set $G \supset E$ such that $\lambda(G) < \lambda^*(E) + \varepsilon$. If we are given an open set $U \supset E$, then we take G in U. Deleting from $\mathcal{I}$ all the intervals not contained in G, one can assume from the very beginning that the intervals of $\mathcal{I}$ are in U. Hence the measure of their union does not exceed $\lambda^*(E) + \varepsilon$. Let E_n be the set of all points $x \in E$ such that $\mathcal{I}$ contains an interval $(x, x+h)$ with $h > 1/n$. Since E is the union of the increasing sets E_n, there is n with $\lambda^*(E_n) > \lambda^*(E) - \varepsilon/2$. Let $\delta = \varepsilon/(2n+2)$. Let a_1 be the infimum of E_n. Let us take a point $x_1 \in E_n$ in $[a_1, a_1 + \delta]$. Let $I_1 = (x_1, x_1 + h_1) \in \mathcal{I}$ be an interval with $h_1 > 1/n$. If the set $E_n \cap (x_1 + h_1, 1)$ is nonempty, then let a_2 be its infimum. Let us take a point $x_2 \in E_n$ in $[a_2, a_2 + \delta]$ and find $I_2 = (x_2, x_2 + h_2) \in \mathcal{I}$ with $h_2 > 1/n$. Continuing this process, we obtain $k \le n$ intervals $I_j = (x_j, x_j + h_j)$ with $h_j > 1/n$ such that there are no points of E_n on the right from $x_k + h_k$ and $x_j \in [a_j, a_j + \delta]$, where a_j is the infimum of $E_n \cap (x_{j-1} + h_{j-1}, 1)$. It is clear that the points in E_n that are not covered by $\bigcup_{j=1}^k I_j$, are contained in the union of the intervals $[a_j, a_j + \delta]$, $j = 1, \ldots, k$. Hence the outer measure of the set of such points does not exceed $n\delta < \varepsilon/2$. Therefore, by the subadditivity of outer measure

$$\lambda^*\bigl(E \bigcap \textstyle\bigcup_{j=1}^k I_j\bigr) \ge \lambda^*(E_n) - \lambda^*\bigl(E_n \backslash \bigcup_{j=1}^k I_j\bigr) > \lambda^*(E) - \varepsilon.$$

Finally, one has $\lambda\bigl(\bigcup_{j=1}^k I_j\bigr) \le \lambda(G) < \lambda^*(E) + \varepsilon$. □

5.2.6. Theorem. *Let f be a function of bounded variation on $[a,b]$. Then f has a finite derivative almost everywhere on $[a,b]$.*

PROOF. It suffices to give a proof for a nondecreasing function f. Let $S = \{x\colon D_+f(x) < D^+f(x)\}$. Let us show that $\lambda(S) = 0$. To this end, it is sufficient to show that for every pair of rational numbers $u < v$, the set

$$S(u,v) = \bigl\{x\colon D_+f(x) < u < v < D^+f(x)\bigr\}$$

has measure zero. Suppose that $\lambda^*\bigl(S(u,v)\bigr) = c > 0$. Every point x in the set $S(u,v)$ is the left endpoint of arbitrarily small intervals $(x, x+h)$ with the property that $f(x+h) - f(x) < hu$. By Lemma 5.2.5, for fixed $\varepsilon > 0$, there exists a finite collection of pairwise disjoint intervals $(x_i, x_i + h_i)$ such that for their union U one has the estimates

$$\lambda^*\bigl(U \cap S(u,v)\bigr) > c - \varepsilon, \quad \lambda(U) = \sum_i h_i < c + \varepsilon.$$

It is clear that $\sum_i \bigl[f(x_i + h_i) - f(x_i)\bigr] < \sum_i h_i u < u(c+\varepsilon)$. On the other hand, every point $y \in U \cap S(u,v)$ is the left endpoint of arbitrarily small intervals $(y, y+r)$ with $f(y+r) - f(y) > rv$. Hence by Lemma 5.2.5 one can find a finite collection of pairwise disjoint intervals $(y_j, y_j + r_j)$ in U such that for their union W one has

$$\lambda^*\bigl(W \cap S(u,v)\bigr) > \lambda^*\bigl(U \cap S(u,v)\bigr) - \varepsilon > c - 2\varepsilon.$$

Then $\sum_j \bigl[f(y_j + r_j) - f(y_j)\bigr] > v \sum_j r_j > v(c - 2\varepsilon)$. Since f is nondecreasing and every interval $(y_j, y_j + r_j)$ belongs to one of the intervals $(x_i, x_i + h_i)$, we obtain the following estimate:

$$\sum_j \bigl[f(y_j + r_j) - f(y_j)\bigr] \le \sum_i \bigl[f(x_i + h_i) - f(x_i)\bigr].$$

Hence $v(c - 2\varepsilon) < u(c + \varepsilon)$. Since $\varepsilon > 0$ is arbitrary, we obtain $v \le u$, which is a contradiction. Therefore, $c = 0$, and the right derivative of f exists almost everywhere. One proves similarly that the left derivative of f exists almost everywhere. The set E of all points x with $f'(x) = +\infty$ has measure zero. Indeed, let $\varepsilon > 0$ and $N \in \mathbb{N}$. There exists $h(x) > 0$ such that $f(x + h) - f(x) > Nh$ whenever $0 < h < h(x)$. By Lemma 5.2.5, there is a finite collection of disjoint intervals $(x_i, x_i + h_i)$, where $x_i \in E$ and $h_i = h(x_i)$, the sum of lengths of which, denoted by L, is at least $\lambda^*(E) - \varepsilon$. The intervals $\bigl(f(x_i), f(x_i+h_i)\bigr)$ are disjoint and the sum of their lengths is not less than NL. Hence we obtain $\lambda^*(E) \le \varepsilon + L \le \varepsilon + V(f, [a, b])/N$. Thus, $\lambda(E) = 0$. Now the assertion follows by Lemma 5.1.3. □

5.2.7. Corollary. *Every nondecreasing function f on a closed interval $[a, b]$ has a finite derivative f' almost everywhere on $[a, b]$, the function f' is integrable on $[a, b]$ and*

$$\int_a^b f'(x)\,dx \le f(b) - f(a). \tag{5.2.3}$$

PROOF. Set $f(x) = f(b)$ if $x \ge b$. Let $f_n(x) = h_n^{-1}[f(x + h_n) - f(x)]$, $h_n = n^{-1}$. Then $f_n \ge 0$ and $f_n(x) \to f'(x)$ a.e. In addition,

$$\begin{aligned}\int_a^b f_n(x)\,dx &= \frac{1}{h_n}\int_{a+h_n}^{b+h_n} f(y)\,dy - \frac{1}{h_n}\int_a^b f(x)\,dx \\ &= \frac{1}{h_n}\int_b^{b+h_n} f(x)\,dx - \frac{1}{h_n}\int_a^{a+h_n} f(x)\,dx \le f(b) - f(a),\end{aligned}$$

since $f = b$ on $[b, b + h_n]$ and $f \ge f(a)$ on $[a, a + h_n]$. It remains to apply Fatou's theorem. □

This corollary yields the integrability of the derivative of every function of bounded variation. Cantor's function C_0 (see Example 3.6.5) shows that in (5.2.3) there might be no equality even for continuous functions. Indeed, $C_0'(x) = 0$ almost everywhere, but $C_0(x) \ne \text{const}$. In the next section, we consider a subclass of the space of functions of bounded variation with an equality in (5.2.3).

We note an interesting result due to Fubini [**332**], the proof of which is delegated to Exercise 5.8.42.

5.2.8. Proposition. *Let f_n be nondecreasing functions on $[a, b]$ such that the series $f(x) = \sum_{n=1}^{\infty} f_n(x)$ converges for all $x \in [a, b]$. Then*

$$f'(x) = \sum_{n=1}^{\infty} f_n'(x) \quad \textit{a.e.}$$

5.3. Absolutely continuous functions

In this section, we consider functions on bounded intervals.

5.3.1. Definition. *A function f on an interval $[a,b]$ is called absolutely continuous if, for every $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$\sum_{i=1}^{n} |f(b_i) - f(a_i)| < \varepsilon$$

for every finite collection of pairwise disjoint intervals (a_i, b_i) in $[a,b]$ with $\sum_{i=1}^{n} |b_i - a_i| < \delta$.

Let $AC[a,b]$ denote the class of all absolutely continuous functions on the interval $[a,b]$.

It is obvious from the definition that any absolutely continuous function is uniformly continuous. The converse is not true: for example, the function f on $[0,1]$ that equals n^{-1} at $(2n)^{-2}$, vanishes at $(2n+1)^{-2}$ and is linearly interpolated between these points is not absolutely continuous. This is clear from divergence of the series $\sum_{n=1}^{\infty} \bigl|f\bigl((2n)^{-1}\bigr)\bigr|$ and convergence to zero of the sequence of sums $\sum_{n=m}^{\infty} \bigl[(2n)^{-1} - (2n+1)^{-1}\bigr]$.

5.3.2. Lemma. *Let functions $f_1, \ldots, f_n$ be absolutely continuous on the interval $[a,b]$ and let a function φ be defined and satisfy the Lipschitz condition on a set $U \subset \mathbb{R}^n$. Suppose that $\bigl(f_1(x), \ldots, f_n(x)\bigr) \in U$ for all $x \in [a,b]$. Then the function $\varphi(f_1, \ldots, f_n)$ is absolutely continuous on the interval $[a,b]$.*

PROOF. By hypothesis, for some $C > 0$ and all $x, y \in U$ we have

$$|\varphi(x) - \varphi(y)| \le C\|x - y\|.$$

In addition, given $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\sum_{i=1}^{k} |f_j(b_i) - f_j(a_i)| < \varepsilon n^{-1}(C+1)^{-1}, \quad j = 1, \ldots, n,$$

for every collection of pairwise disjoint intervals $(a_1, b_1), \ldots, (a_k, b_k)$ in $[a,b]$ with $\sum_{i=1}^{k} |b_i - a_i| < \delta$. Now the estimate

$$\sum_{i=1}^{k} \bigl|\varphi\bigl(f_1(b_i), \ldots, f_n(b_i)\bigr) - \varphi\bigl(f_1(a_i), \ldots, f_n(a_i)\bigr)\bigr|$$
$$\le \sum_{i=1}^{k} C\biggl(\sum_{j=1}^{n} |f_j(b_i) - f_j(a_i)|^2\biggr)^{1/2} \le C\sum_{i=1}^{k}\sum_{j=1}^{n} |f_j(b_i) - f_j(a_i)| < \varepsilon$$

proves our claim. □

5.3.3. Corollary. *If functions f and g are absolutely continuous on $[a,b]$, then so are fg and $f+g$, and if $g \ge c > 0$, then f/g is absolutely continuous.*

5.3.4. Proposition. *Every function f that is absolutely continuous on the interval $[a,b]$ is of bounded variation on this interval.*

PROOF. We take δ corresponding to $\varepsilon = 1$ in the definition of absolutely continuous functions. Next we pick a natural number $M > |b-a|\delta^{-1}$. Suppose we are given a partition $a = t_1 \le \cdots \le t_n = b$. We add to the points t_i all points of the form $s_j = a + (b-a)jM^{-1}$, $j = 0, \ldots, M$. The elements of this new partition are denoted by z_i, $i = 1, \ldots, k$. Then

$$\begin{aligned}\sum_{i=1}^{n-1} |f(t_{i+1}) - f(t_i)| &\le \sum_{i=1}^{k-1} |f(z_{i+1}) - f(z_i)| \\ &= \sum_{j=1}^{M} \sum_{i\colon\, z_{i+1}\in(s_{j-1}, s_j]} |f(z_{i+1}) - f(z_i)| \le M,\end{aligned}$$

since the sum of lengths of the intervals (z_i, z_{i+1}) with $z_{i+1} \in (s_{j-1}, s_j]$ does not exceed $s_j - s_{j-1} = |b-a|M^{-1} < \delta$. Thus, $V(f, [a,b]) \le M$. $\square$

5.3.5. Corollary. *Let a function f be absolutely continuous on $[a,b]$. Then the function $V\colon\, x \mapsto V(f, [a,x])$ is absolutely continuous as well, hence f is the difference of the nondecreasing absolutely continuous functions V and $V - f$.*

PROOF. Let $\varepsilon > 0$. We find $\delta > 0$ such that the sum of the absolute values of the increments of f on every finite collection of disjoint intervals (a_i, b_i) of total length less than δ is estimated by $\varepsilon/2$. It remains to observe that the sum of the absolute values of the increments of V on the intervals (a_i, b_i) is estimated by ε. Indeed, suppose we are given such a collection of k intervals (a_i, b_i). For every i, one can find a partition of $[a_i, b_i]$ by points $a_i = t_1^i \le \cdots \le t_{N_i}^i = b_i$ such that

$$V(f, [a_i, b_i]) < \sum_{j=1}^{N_i - 1} |f(t_{j+1}^i) - f(t_j^i)| + \varepsilon 4^{-i}.$$

Then

$$\sum_{i=1}^{k} |V(b_i) - V(a_i)| = \sum_{i=1}^{k} V(f, [a_i, b_i]) < \sum_{i=1}^{k} \sum_{j=1}^{N_i - 1} |f(t_{j+1}^i) - f(t_j^i)| + \frac{\varepsilon}{2} < \varepsilon,$$

since the intervals (t_j^i, t_{j+1}^i) are pairwise disjoint and the sum of their lengths does not exceed δ. $\square$

For every Lebesgue integrable function f on $[a,b]$ and any constant C, one can consider the function

$$F(x) = C + \int_a^x f(t)\, dt,$$

which is called an *indefinite integral* of f. It turns out that the functions of such a form are precisely the absolutely continuous functions.

5.3.6. Theorem. *A function f is absolutely continuous on $[a,b]$ if and only if there exists an integrable function g on $[a,b]$ such that*

$$f(x) = f(a) + \int_a^x g(y)\,dy, \quad \forall\, x \in [a,b]. \tag{5.3.1}$$

PROOF. If f has form (5.3.1), then by the absolute continuity of the Lebesgue integral, for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\int_D |g(x)|\,dx < \varepsilon$$

for any set D of measure less than δ. It remains to observe that

$$\sum_{i=1}^n |f(b_i) - f(a_i)| = \sum_{i=1}^n \Bigl|\int_{a_i}^{b_i} g(x)\,dx\Bigr| \le \int_U |g(x)|\,dx < \varepsilon$$

for any union $U = \bigcup_{i=1}^n [a_i, b_i]$ of pairwise disjoint intervals of total length less than δ.

Let us prove the converse assertion. It suffices to prove it for nondecreasing functions f because by Corollary 5.3.5 the function f is the difference of nondecreasing absolutely continuous functions. According to Theorem 1.8.1, there exists a nonnegative Borel measure μ on $[a,b]$ such that $f(x) = \mu\bigl([a,x)\bigr)$ for all $x \in [a,b]$. Now it is sufficient to show that the measure μ is given by an integrable density g with respect to Lebesgue measure λ, which by the Radon–Nikodym theorem is equivalent to the absolute continuity of the measure μ with respect to Lebesgue measure. Let E be a Borel set of Lebesgue measure zero in $[a,b]$. We have to verify that $\mu(E) = 0$. Let us fix $\varepsilon > 0$. By hypothesis, there exists $\delta > 0$ such that the sum of the absolute values of the increments of f on any disjoint intervals of the total length less than δ is estimated by ε. By Theorem 1.4.8, there exists an open set U containing E such that $\mu(U \backslash E) < \varepsilon$. Making U smaller, one can ensure the estimate $\lambda(U) < \delta$. The set U is the finite or countable union of pairwise disjoint intervals (a_i, b_i). By the choice of δ, for every finite union of (a_i, b_i), we have

$$\mu\Bigl(\bigcup_{i=1}^n (a_i, b_i)\Bigr) = \sum_{i=1}^n |f(b_i) - f(a_i)| < \varepsilon,$$

whence by the countable additivity of μ we obtain $\mu(U) < \varepsilon$. Therefore, $\mu(E) < 2\varepsilon$ and hence $\mu(E) = 0$. $\square$

5.3.7. Corollary. *If* (5.3.1) *is fulfilled, then*

$$V(f, [a,b]) = \int_a^b |g(x)|\,dx. \tag{5.3.2}$$

PROOF. Since for every interval $[s,t] \subset [a,b]$ one has

$$|f(t) - f(s)| = \Bigl|\int_s^t g(x)\,dx\Bigr| \le \int_s^t |g(x)|\,dx,$$

we obtain

$$V(f,[a,b]) \leq \int_a^b |g(x)|\,dx.$$

Let us prove the reverse inequality. We may assume that $f(a)=0$. Let us fix $\varepsilon>0$. By using the absolute continuity of the Lebesgue integral, we find $\delta>0$ such that

$$\int_D |g(x)|\,dx < \frac{1}{8}\varepsilon$$

for every set D of measure less than δ. Set

$$\Omega_+ = \{x\colon\ g(x)\geq 0\},\quad \Omega_- = \{x\colon\ g(x)<0\}.$$

Then we find finitely many pairwise disjoint intervals (a_1,b_1), ...,(a_n,b_n) in $[a,b]$ such that

$$\lambda\Bigl(\Omega_+ \bigtriangleup \bigcup_{i=1}^n (a_i,b_i)\Bigr) < \delta. \tag{5.3.3}$$

Next we choose in $[a,b]\backslash \bigcup\limits_{i=1}^n (a_i,b_i)$ a finite collection of pairwise disjoint intervals (c_1,d_1), ...,(c_k,d_k) such that

$$\lambda\Bigl(\Omega_- \bigtriangleup \bigcup_{i=1}^k (c_i,d_i)\Bigr) < \delta.$$

Set $\Delta_i = (a_i,b_i)\backslash\{g>0\}$. Then

$$\begin{aligned} f(b_i)-f(a_i) &= \int_{a_i}^{b_i} g(x)\,dx = \int_{a_i}^{b_i} |g(x)|\,dx + \int_{a_i}^{b_i} \Bigl[g(x)-|g(x)|\Bigr]\,dx \\ &= \int_{a_i}^{b_i} |g(x)|\,dx - 2\int_{\Delta_i} |g(x)|\,dx. \end{aligned}$$

On account of estimate (5.3.3), which, in particular, shows that the sum of measures of the sets Δ_i is less than δ, we obtain

$$\sum_{i=1}^n |f(b_i)-f(a_i)| \geq \sum_{i=1}^n \int_{a_i}^{b_i} |g(x)|\,dx - \frac{1}{4}\varepsilon \geq \int_{\Omega_+} |g(x)|\,dx - \frac{1}{2}\varepsilon.$$

Similarly, we obtain

$$\sum_{i=1}^k |f(d_i)-f(c_i)| \geq \int_{\Omega_-} |g(x)|\,dx - \frac{1}{2}\varepsilon.$$

Thus,

$$V(f,[a,b]) \geq \sum_{i=1}^n |f(b_i)-f(a_i)| + \sum_{i=1}^k |f(d_i)-f(c_i)| \geq \int_a^b |g(x)|\,dx - \varepsilon,$$

which completes the proof. $\square$

5.4. The Newton–Leibniz formula

5.4.1. Lemma. *Let f be an integrable function on $[a,b]$ such that*

$$\int_a^x f(t)\,dt = 0, \quad \forall\, x \in [a,b].$$

Then $f = 0$ almost everywhere.

Proof. It follows by our hypothesis that the integral of f over every interval in $[a,b]$ is zero, whence we obtain that the integrals of f over finite unions of intervals vanish. Let us show that the integral of f over the set $\Omega = \{x\colon\ f(x) > 0\}$ vanishes as well. Indeed, let $\varepsilon > 0$. By the absolute continuity of the Lebesgue integral there exists $\delta > 0$ such that

$$\int_D |f|\,dx < \varepsilon$$

for every set D of measure less than δ. We find a set A that is finite union of intervals with $\lambda(\Omega \triangle A) < \delta$. Then

$$\int_\Omega f(x)\,dx \le \int_A f(x)\,dx + \int_{\Omega\triangle A} |f(x)|\,dx \le \int_A f(x)\,dx + \varepsilon = \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, the left-hand side of this inequality vanishes, i.e., Ω has measure zero. Similarly, the set $\{f < 0\}$ has measure zero. An alternative reasoning is this: the Borel measure $\mu := f \cdot \lambda$ vanishes on all intervals, hence on the σ-algebra generated by them, i.e., is zero on the Borel σ-algebra. In other words, the integrals of f over all Borel sets vanish, which means that $f = 0$ a.e. □

5.4.2. Theorem. *Let a function f be integrable on $[a,b]$. Then*

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x) \quad \textit{almost everywhere on } [a,b].$$

Proof. Set $f(x) = 0$ if $x \notin [a,b]$. Let

$$F(x) = \int_a^x f(t)\,dt.$$

Suppose first that $|f(x)| \le M < \infty$. Let $h_n \to 0$. As shown above, the function F is absolutely continuous, therefore, is of bounded variation and is almost everywhere differentiable on $[a,b]$. Then for a.e. $x \in [a,b]$ we have $\lim\limits_{n\to\infty} h_n^{-1}[F(x+h_n) - F(x)] = F'(x)$. Since

$$\Bigl|\frac{F(x+h_n) - F(x)}{h_n}\Bigr| = \Bigl|\frac{1}{h_n}\int_x^{x+h_n} f(t)\,dt\Bigr| \le M,$$

we obtain by the monotone convergence theorem

$$\lim_{n\to\infty}\int_a^x \frac{F(y+h_n) - F(y)}{h_n}\,dy = \int_a^x F'(y)\,dy$$

for every $x \in [a, b]$. We observe that

$$\int_a^x \frac{F(y+h_n) - F(y)}{h_n}\,dy = \frac{1}{h_n}\int_{a+h_n}^{x+h_n} F(y)\,dy - \frac{1}{h_n}\int_a^x F(y)\,dy$$
$$= \frac{1}{h_n}\int_{x}^{x+h_n} F(y)\,dy - \frac{1}{h_n}\int_a^{a+h_n} F(y)\,dy,$$

which approaches $F(x) - F(a)$ as $n \to \infty$ by the continuity of F. Hence

$$F(x) = F(x) - F(a) = \int_a^x F'(y)\,dy,$$

i.e., one has

$$\int_a^x \bigl[F'(y) - f(y)\bigr]\,dy = 0, \quad \forall\, x \in [a, b].$$

By Lemma 5.4.1 this means that $F'(x) - f(x) = 0$ a.e. on $[a, b]$.

We proceed to the general case. We may assume that $f \geq 0$ because f is the difference of two nonnegative integrable functions. Let $f_n = \min(f, n)$. Since $f - f_n \geq 0$, the function

$$\int_a^x \bigl(f(t) - f_n(t)\bigr)\,dt$$

is nondecreasing, therefore, its derivative exists almost everywhere and is nonnegative. Thus,

$$\frac{d}{dx}\int_a^x f(y)\,dy \geq \frac{d}{dx}\int_a^x f_n(y)\,dy \quad \text{a.e.}$$

By the boundedness of f_n and the previous step, we obtain $F'(x) \geq f_n(x)$ a.e. Hence $F'(x) \geq f(x)$ a.e., whence we obtain

$$\int_a^b F'(x)\,dx \geq \int_a^b f(x)\,dx.$$

On the other hand, by Corollary 5.2.7 we have

$$\int_a^b F'(x)\,dx \leq F(b) - F(a) = \int_a^b f(x)\,dx,$$

whence it follows that

$$\int_a^b \bigl[F'(x) - f(x)\bigr]\,dx = 0,$$

which is only possible if $F'(x) - f(x) = 0$ a.e. because $F'(x) - f(x) \geq 0$ a.e. as shown above. □

The Newton–Leibniz formula yields the following integration by parts formula.

5.4.3. Corollary. *Let f and g be two absolutely continuous functions on the interval $[a, b]$. Then*

$$\int_a^b f'(x)g(x)\,dx = f(b)g(b) - f(a)g(a) - \int_a^b f(x)g'(x)\,dx. \tag{5.4.1}$$

PROOF. Since the function fg is absolutely continuous, the Newton–Leibniz formula applies and it remains to observe that $(fg)' = f'g + fg'$ almost everywhere (i.e., at all points where f and g are differentiable). □

A related result is found in Exercise 5.8.43.

One more useful corollary of the Newton–Leibniz formula is the change of variables formula for absolutely continuous transformations.

5.4.4. Corollary. *Let φ be a monotone absolutely continuous function on the interval $[c, d]$ and let $F([c, d]) \subset [a, b]$. Then, for every function f that is Lebesgue integrable on the interval $[a, b]$, the function $f(\varphi)\varphi'$ is integrable on $[c, d]$ and one has*

$$\int_{\varphi(c)}^{\varphi(d)} f(x)\,dx = \int_c^d f\bigl(\varphi(y)\bigr)\varphi'(y)\,dy. \tag{5.4.2}$$

This assertion remains true for unbounded intervals of the form $(-\infty, d]$, $[c, +\infty)$, $(-\infty, +\infty)$.

PROOF. We may assume that φ is increasing and $a = \varphi(c)$, $b = \varphi(d)$. Since the function φ' is integrable on $[c, d]$, we obtain the finite nonnegative Borel measure $\mu = \varphi' \cdot \lambda$, where λ is Lebesgue measure on $[c, d]$. Denote by ν the Borel measure $\mu \circ \varphi^{-1}$ on $[a, b]$, i.e., $\nu(B) = \mu\bigl(\varphi^{-1}(B)\bigr)$. By the general change of variables formula (see Theorem 3.6.1) equality (5.4.2) for all Borel measurable integrable functions f is equivalent to the equality of the measure ν to Lebesgue measure λ_1 on $[a, b]$. Hence, for the proof in the case of Borel measurable f, it suffices to establish the equality $\nu([\alpha, \beta]) = \lambda_1([\alpha, \beta])$ for every interval $[\alpha, \beta]$ in $[a, b]$ (see Corollary 2.7.4). There exists an interval $[\gamma, \delta] \subset [c, d]$ such that $[\gamma, \delta] = \varphi^{-1}([\alpha, \beta])$ and $\varphi(\gamma) = \alpha$, $\varphi(\delta) = \beta$. It remains to observe that

$$\nu([\alpha, \beta]) = \mu([\gamma, \delta]) = \int_\gamma^\delta \varphi'(y)\,dy = \varphi(\delta) - \varphi(\gamma) = \beta - \alpha.$$

In order to extend the established equality from Borel measurable functions to arbitrary Lebesgue integrable ones, it suffices to verify that the measure ν is absolutely continuous, i.e., for every set E of Lebesgue measure zero, the set $\varphi^{-1}(E)$ in $[c, d]$ has μ-measure zero. This is equivalent to saying that the set $\varphi^{-1}(E) \bigcap \{y\colon\ \varphi'(y) > 0\}$ has Lebesgue measure zero. Since E is covered by a Borel set of Lebesgue measure zero, one can deal with the case where E itself is Borel. Then it remains to apply equality (5.4.2) to the function $f = I_E$. The case of an unbounded interval follows from the considered case. □

It is worth noting that there is an alternative justification of the change of variable formula. To this end, as is clear from the above reasoning, it suffices to establish (5.4.2) for continuous f. In that case, the required formula follows at once from the Newton–Leibniz formula applied to the function $F(\varphi)$, where

$$F(x) = \int_a^x f(t)\,dt.$$

By the continuous differentiability of F this function is absolutely continuous and $(F\circ\varphi)'(x) = F'\bigl(\varphi(x)\bigr)\varphi'(x)$ almost everywhere on $[c,d]$. However, for discontinuous f such a justification, although possible, requires some extra work because the equality $(F\circ\varphi)' = F'(\varphi)\varphi'$ may not hold at all points of differentiability of φ. G.M. Fichtenholz showed (see Exercise 5.8.86) that formula (5.4.2) remains true without the hypothesis of the absolute continuity of φ if the composition $F\circ\varphi$ is absolutely continuous (in our case this condition is fulfilled automatically by Exercise 5.8.59); see also Morse [**698**]. Finally, it is worth noting that formula (5.4.2) is true for not necessarily monotone functions φ if it is known additionally that the function $f(\varphi)\varphi'$ is integrable; however, unlike the above situation this does not hold automatically (see Exercise 5.8.122).

As a corollary of the established facts we obtain the following Lebesgue decomposition of monotone functions.

5.4.5. Proposition. *Let F be a nondecreasing left continuous function on the interval $[a,b]$. Then $F = F_{\rm ac} + F_{\rm sing}$, where $F_{\rm ac}$ is an absolutely continuous nondecreasing function and $F_{\rm sing}$ is a nondecreasing left continuous function with $F'_{\rm sing}(t) = 0$ a.e. In addition, $F_{\rm sing} = F_{\rm a} + F_{\rm c}$, where $F_{\rm c}$ is a continuous nondecreasing function and $F_{\rm a}$ is a nondecreasing jump function, i.e., $F_{\rm a}$ is constant on the intervals on which there are no jumps.*

PROOF. We know that F' exists a.e., is integrable and

$$F(y) - F(x) \geq \int_x^y F'(t)\,dt \quad \text{if } a \leq x \leq y \leq b.$$

Hence the function

$$F_{\rm sing}(x) := F(x) - \int_a^x F'(t)\,dt$$

is increasing. It is clear that $F'_{\rm sing}(x) = 0$ a.e. Let

$$F_{\rm ac}(x) := \int_a^x F'(t)\,dt.$$

The function $F_{\rm sing}$ has at most countably many points of discontinuity t_n. The size of the jump at t_n is denoted by h_n. Let $F_{\rm a}(t) := \sum_{n:\ t_n<t} h_n$. It is verified directly that this is an increasing left continuous function and that the function $F_{\rm c} := F_{\rm sing} - F_{\rm a}$ is increasing and continuous. $\square$

One can look at the Lebesgue decomposition from another point of view (which also gives a different justification). If F is the distribution function of a bounded nonnegative Borel measure μ on $[a,b]$ (any left continuous increasing function has such a form), then the Lebesgue decomposition for measures yields the equality $\mu = \mu_{\rm ac} + \mu_{\rm sing}$, where the measure $\mu_{\rm ac}$ is absolutely continuous and the measure $\mu_{\rm sing}$ is singular with respect to Lebesgue measure. Then $F_{\rm ac}$ and $F_{\rm sing}$ are the corresponding distribution functions. The singular measure $\mu_{\rm sing}$ possesses a purely atomic component (concentrated on a countable set) and an atomless component, which gives the corresponding decomposition of $F_{\rm sing}$.

An application of the Newton–Leibniz formula to differentiation with respect to a parameter is found in Exercise 5.8.135.

5.5. Covering theorems

In this section, we discuss several important theorems that enable one to choose in covers of sets by intervals, balls or cubes disjoint subcovers up to sets of measure zero. First we prove Vitali's theorem for the real line with Lebesgue measure λ.

5.5.1. Theorem. *Let $E \subset \mathbb{R}^1$ be an arbitrary set. Suppose that for every $x \in E$ and $\varepsilon > 0$, we are given a closed interval $I(x,\varepsilon) \ni x$ of positive length less than ε. Then, there exists an at most countable set of disjoint closed intervals $I_j = I(x_j, \varepsilon_j)$ such that $\lambda\bigl(E\backslash \bigcup_{j=1}^\infty I_j\bigr) = 0$.*

PROOF. Suppose first that the set E is bounded. We may assume that $E \subset (0,1)$. Deleting the intervals of the given cover not belonging to $(0,1)$, we arrive at the situation where all given closed intervals are contained in $(0,1)$. The collection of all these intervals is denoted by S. We find an interval $I_1 \in S$ such that

$$\lambda(I_1) > \frac{1}{2}\sup\bigl\{\lambda(J)\colon\ J \in S\bigr\}.$$

Denote by S_1 the collection of all intervals remaining after deletion from S all the intervals meeting I_1 (in particular, I_1 itself). We find an interval $I_2 \in S_1$ with

$$\lambda(I_2) > \frac{1}{2}\sup\bigl\{\lambda(J)\colon\ J \in S_1\bigr\}.$$

Let us continue this process inductively: if the class S_n of closed intervals is not empty, then we find an interval $I_{n+1} \in S_n$ such that

$$\lambda(I_{n+1}) > \frac{1}{2}\sup\bigl\{\lambda(J)\colon\ J \in S_n\bigr\}.$$

Deleting from S_n all intervals that have nonempty intersections with I_n, we obtain the class S_{n+1}. It is clear that as a result we obtain a finite or countable set of pairwise disjoint intervals of the initial cover. In particular, $\sum_{j=1}^\infty \lambda(I_j) \le 1$. Let us show that $\lambda\bigl(E\backslash \bigcup_{j=1}^\infty I_j\bigr) = 0$. Since

$$E\backslash \bigcup_{j=1}^\infty I_j \subset E_n := E\backslash \bigcup_{j=1}^n I_j,$$

it suffices to verify that $\lambda^*(E_n) \to 0$. To this end, let T_j denote the interval with the same center as I_j and length $\lambda(T_j) = 5\lambda(I_j)$. Convergence of the series with the general term $\lambda(I_j)$ yields that $\sum_{j=n}^{\infty} \lambda(T_j) \to 0$ as $n \to \infty$. Thus, it remains to verify that the set $\bigcup_{j=n}^{\infty} T_j$ covers E_n. We observe that the intervals of the family S_n cover the set E_n because every point $x \in E_n$ is contained in some interval from S that does not meet the closed set $I_1 \cup \cdots \cup I_n$. Every interval $I \in S$ has been deleted at some step k, since

$$\sup\bigl\{\lambda(J)\colon\ J \in S_n\bigr\} < 2\lambda(I_{n+1}) \to 0.$$

According to our construction, for this index k we have $I \cap I_k \neq \varnothing$ and $I \in S_{k-1} \backslash S_k$. Hence

$$\lambda(I) \le \sup\bigl\{\lambda(J)\colon\ J \in S_{k-1}\bigr\} < 2\lambda(I_k).$$

Therefore, $I \subset T_k$. It follows that all intervals in the family S_n are covered by the set $T_n \cup T_{n+1} \cup \cdots$, hence this union contains E_n.

In the case of an unbounded set E we find can subcovers of the bounded sets $E \cap (k, k+1)$ by intervals from $(k, k+1)$, and deal with each of these intersections separately. □

Let us generalize Vitali's theorem to the multidimensional case. It is natural to ask what sets can be taken in place of intervals. This is a rather subtle question and we do not discuss it, see Guzmán [**386**]. Lebesgue measure on $\mathbb{R}^n$ will be denoted by λ for notational simplicity.

5.5.2. Theorem. *Let $E \subset \mathbb{R}^n$ be an arbitrary set. Suppose that for every point $x \in E$ and every $\varepsilon > 0$, we are given a closed ball $B_{x,\varepsilon} \ni x$ of positive diameter less than ε. Then, this family of balls contains an at most countable subfamily of pairwise disjoint balls B_k such that*

$$\lambda\Bigl(E \backslash \bigcup_{k=1}^{\infty} B_k\Bigr) = 0.$$

The same is true if in place of balls we are given closed cubes with edges parallel to the coordinate axes.

Proof. We shall follow the same plan as in the previous theorem. First we consider the case where E is bounded. Hence, without loss of generality we may assume that all balls of our cover, denoted by S, are contained in some ball. As in the one-dimensional case, we define inductively a sequence of pairwise disjoint balls B_k according to the formula

$$\lambda(B_k) > \frac{1}{2}\sup\bigl\{\lambda(B)\colon\ B \in S_{k-1}\bigr\}, \quad k > 1, \qquad (5.5.1)$$
$$S_{k-1} := \bigl\{B \in S\colon\ B \cap (B_1 \cup \cdots \cup B_{k-1}) = \varnothing\bigr\}.$$

For B_1 we take a ball of measure greater than $\frac{1}{2}\sup\{\lambda(B)\colon\ B \in S\}$. If this inductive process is finite, then we obtain a finite collection of balls covering E. So we assume that we obtain an infinite sequence of pairwise disjoint balls B_k.

Let us prove that the outer measure of the set $E\backslash\bigcup_{k=1}^m B_k$ tends to zero as $m\to\infty$. We show that

$$\lambda^*\Bigl(E\backslash\bigcup_{k=1}^m B_k\Bigr)\le(1+2^{1+1/n})^n\sum_{k=m+1}^\infty\lambda(B_k).$$

The right-hand side of this inequality approaches zero by convergence of the series of measures of disjoint subsets of a ball. The desired estimate will be established if we prove that $E\backslash\bigcup_{k=1}^m B_k$ is contained in the union of the sets T_m, T_{m+1}, ..., where

$$T_j=\Bigl\{\text{the union of all } B\in S_j\colon\ \lambda(B)\le 2\lambda(B_{j+1}), B\cap B_{j+1}\ne\varnothing\Bigr\}.$$

Indeed, $\lambda^*(T_j)\le(1+2^{1+1/n})^n\lambda(B_{j+1})$, since T_j is contained in the ball with the same center as B_{j+1} and the radius $(1+2^{1+1/n})r_{j+1}$, where r_{j+1} is the radius of B_{j+1}. This is seen from the fact that the radius of a ball B does not exceed $2^{1/n}r_{j+1}$ if its measure is not greater than $2\lambda(B_{j+1})$ (inflating a ball q times increases its volume q^n times).

Now we verify the inclusion $E\backslash\bigcup_{k=1}^m B_k\subset\bigcup_{j=m}^\infty T_j$. Let $x\in E\backslash\bigcup_{k=1}^m B_k$. Since the union of the balls $B_1,\ldots,B_m$ is closed, there exists a neighborhood of the point x that has no common points with that union. Therefore, there exists a ball $B\in S$ such that $x\in B$ and $B\cap B_k=\varnothing$, $k=1,\ldots,m$. By the construction of the balls B_k we have $\lambda(B_{m+1})\ge\frac12\lambda(B)$. If $B\cap B_{m+1}$ is nonempty, then $B\subset T_m$. Otherwise we take the smallest number $l>m$ such that $B\cap B_l$ is nonempty. Such a number exists because otherwise in view of (5.5.1) the measure B_k could not approach zero. Then $B\cap B_{l-1}=\varnothing$, whence one has $\lambda(B_l)\ge\frac12\lambda(B)$. Hence $B\in T_{l-1}$, which proves the required inclusion.

In the case of an unbounded set E, we partition $\mathbb{R}^n$ into cubes of unit volume with pairwise disjoint interiors Q_j and apply the previous step to every intersection $E\cap Q_j$ and its subcover obtained by deleting all balls of the initial cover not contained in Q_j. In the case of cubes in place of balls the reasoning is similar. □

Some generalizations of this theorem and related results are given in §5.8(i) and Exercise 5.8.88.

By a similar reasoning one proves the following assertion, in which a weaker conclusion is compensated by less restrictive assumptions on the initial cover.

5.5.3. Proposition. *Suppose that a measurable set E in $\mathbb{R}^n$ is covered by a family of closed balls with positive and uniformly bounded radii. Then this cover contains an at most countable family of disjoint balls B_k such that*

$$\lambda\Bigl(\bigcup_{k=1}^\infty B_k\Bigr)\ge(1+2^{1+1/n})^{-n}\lambda(E).$$

PROOF. The balls B_k are constructed in the same manner as in Theorem 5.5.2 (independently of whether E is bounded or not). If the series of their measures diverges, then our estimate is obvious. If this series converges, then we take closed balls V_k with the same centers as B_k and radii multiplied by $1+2^{1+1/n}$. It remains to show that $E \subset \bigcup_{k=1}^{\infty} V_k$. To this end, we take a ball B in the original cover and verify that $B \subset \bigcup_{k=1}^{\infty} V_k$. If the ball B is in $\{B_k\}$, then this is obvious. Otherwise, for some of the constructed balls B_j we have

$$B_j \cap B \neq \varnothing \quad \text{and} \quad \lambda(B_j) \geq \frac{1}{2}\lambda(B).$$

Indeed, if the constructed sequence is infinite, then $\lambda(B_k) \to 0$ and we take the first l with $\lambda(B_l) < \frac{1}{2}\lambda(B)$. Then B meets at least one of the balls $B_1, \ldots, B_{l-1}$ because otherwise we would obtain a contradiction with the choice of B_l. Thus, B meets B_j for some $j \leq l-1$. Note that the radius of B does not exceed $2^{1/n}r$, where r is the radius of B_j, since otherwise $\lambda(B_j) < \frac{1}{2}\lambda(B)$ contrary to the choice of l. Hence B belongs to V_j. If the sequence of balls B_k is finite and there is no number l with $\lambda(B_l) < \frac{1}{2}\lambda(B)$, then $\lambda(B_j) \geq \frac{1}{2}\lambda(B)$ for all j. But in this case B meets one of the constructed balls B_j, since otherwise our construction of the sequence of balls could not be completed. As above, we obtain that $B \subset V_j$. □

As an application of the covering theorems we prove the following useful assertion.

5.5.4. Proposition. *Let f be a function on the real line and let E be a measurable set such that at every point of E the function f is differentiable. Then*

$$\lambda\bigl(f(E)\bigr) \leq \int_E |f'(x)|\,dx. \tag{5.5.2}$$

In particular, the function f on E has Lusin's property (N). *If for all $x \in E$ we have $|f'(x)| \leq L$, then $\lambda\bigl(f(E)\bigr) \leq L\lambda(E)$.*

PROOF. It is clear that it is sufficient to consider the case where E is contained in $[0,1]$. In addition, we observe that it suffices to prove the last assertion. Indeed, if the function $|f'|$ is integrable over the set E, then, given $\varepsilon > 0$, we partition $[0,\infty)$ into disjoint intervals $I_j = [L_j, L_{j+1})$ of length ε and let $E_j := \bigl\{x \in E\colon\ f'(x) \in I_j\bigr\}$. Then one has

$$\lambda\bigl(f(E_j)\bigr) \leq L_{j+1}\lambda(E_j) \leq \int_{E_j} |f'(x)|\,dx + \varepsilon\lambda(E_j),$$

which after summing in j gives estimate (5.5.2) with the extra summand ε on the right.

Thus, we assume further that $|f'(x)| \leq L$ for all $x \in E$. Let $\varepsilon > 0$. There is an open set U containing E such that $\lambda(U) < \lambda(E) + \varepsilon$. For every $x \in E$, there exists $h_x > 0$ such that $|f(x+h) - f(x)| \leq (L+\varepsilon)|h|$ whenever $|h| \leq h_x$. If $f'(x) > 0$, then h_x can be chosen with the property that $f(x+h) \geq f(x)$ for all $h \in [0, h_x]$. If $f'(x) < 0$, then we choose h_x such that $f(x-h) \geq f(x)$ for all

$h \in [0, h_x]$, and if $f'(x) = 0$, then we take h_x such that $|f(x+h) - f(x)| \le \varepsilon/2$ whenever $|h| \le h_x$. Finally, making h_x smaller in all the three cases we obtain the inclusion $(x - h_x, x + h_x) \subset U$. Therefore, to every point $f(x)$, where $x \in E$, we associate a system of intervals of the form $[f(x), f(y)]$ or $[f(y), f(x)]$, shrinking to x, such that $|f(y) - f(x)| \le (L + \varepsilon)|y - x|$. This family contains an at most countable subfamily of disjoint intervals I_j with the endpoints $f(x_j)$ and $f(y_j)$ which cover $f(E)$ up to a set of measure zero. We observe that the intervals Δ_j with the endpoints x_j and y_j are disjoint and contained in U, in addition, $|I_j| \le (L + \varepsilon)|\Delta_j|$. Therefore,

$$\lambda^*\bigl(f(E)\bigr) \le (L + \varepsilon)\lambda(U) \le (L + \varepsilon)\lambda(E) + \varepsilon(L + \varepsilon).$$

Letting $\varepsilon \to 0$, we obtain $\lambda^*\bigl(f(E)\bigr) \le L\lambda(E)$. In particular, this shows that f on E has property (N) because our reasoning applies to subsets of E. Finally, $f(E)$ is measurable (which is not obvious in advance). Indeed, $E = N \cup S$, where N is a set of measure zero and S is the union of a sequence of compact sets S_n. The sets $f(S_n)$ are compact by the continuity of f on E and $f(N)$ has measure zero. □

5.6. The maximal function

Let f be a measurable function on $\mathbb{R}^n$ that is integrable on every ball. Denote by $B(x, r)$ the closed ball of radius r centered at x, and by λ Lebesgue measure on $\mathbb{R}^n$ (omitting the index n for simplicity). Set

$$Mf(x) = \sup_{r>0} \frac{1}{\lambda\bigl(B(x, r)\bigr)} \int_{B(x,r)} |f(y)|\, dy.$$

The function Mf is called the maximal function for f. This function plays an important role in analysis, in particular, in the theory of singular integrals. The function Mf may equal $+\infty$ at certain points or even everywhere. In addition, even for a bounded integrable function f, the function Mf may not be integrable. For example, if f is the indicator of $[0, 1]$, then $Mf(x) = (2x)^{-1}$ if $x > 1$. The following theorem describes basic properties of the maximal function.

5.6.1. Theorem. (i) *If $f \in L^p(\mathbb{R}^n)$ for some $p \ge 1$, then the function Mf is almost everywhere finite.*

(ii) *If $f \in L^1(\mathbb{R}^n)$, then for every $t > 0$ one has*

$$\lambda\bigl(x\colon\ (Mf)(x) > t\bigr) \le \frac{C_n}{t} \int_{\mathbb{R}^n} |f(y)|\, dy, \tag{5.6.1}$$

where C_n depends only on n.

(iii) *If $f \in L^p(\mathbb{R}^n)$, where $1 < p \le \infty$, then $Mf \in L^p(\mathbb{R}^n)$ and*

$$\|Mf\|_{L^p} \le C_{n,p}\|f\|_{L^p},$$

where $C_{n,p}$ depends only on n and p.

PROOF. First we prove assertion (ii). Let $E_t = \bigl\{x\colon\ (Mf)(x) > t\bigr\}$. By the definition of Mf, for every $x \in E_t$, there exists a ball B_x centered at x such that

$$\int_{B_x} |f(y)|\,dy > t\lambda(B_x).$$

When x runs through the set E_t, the family of all balls B_x covers E_t. One has $\lambda(B_x) < t^{-1}\|f\|_{L^1}$, i.e., the radii of these balls are uniformly bounded. By Proposition 5.5.3, there exists an at most countable subfamily of pairwise disjoint balls B_{x_k} in this family such that $\sum_{k=1}^\infty \lambda(B_{x_k}) \ge C\lambda(E_t)$, where C is some constant that depends only on n. By the choice of the balls B_{x_k} we obtain

$$\int_{\bigcup_{k=1}^\infty B_{x_k}} |f(y)|\,dy > t\sum_{k=1}^\infty \lambda(B_{x_k}) \ge Ct\lambda(E_t).$$

The left-hand side of this inequality does not exceed $\|f\|_{L^1}$.

In the case $p = 1$ assertion (i) follows from what we have already proved, and in the case $p = \infty$ it is obvious. In the proof of (i) for $p \in (1,\infty)$ we set $f_1 = fI_{\{|f|>1\}}$ and $f_2 = fI_{\{|f|\le 1\}}$. Then $|f| \le f_1 + 1$ and $Mf \le Mf_1 + 1$, which reduces the assertion to the case where the function f is integrable over the whole space, since $f_1 \in L^1(\mathbb{R}^n)$.

Let us prove assertion (iii). In the case $p = \infty$ the function Mf is bounded and one can take $C_{n,p} = 1$. Let $1 < p < \infty$. We take the function g that coincides with f if $|f(x)| \ge t/2$ and equals 0 otherwise. It is clear that $g \in L^1(\mathbb{R}^n)$ and $|f(x)| \le |g(x)| + t/2$, whence $(Mf)(x) \le (Mg)(x) + t/2$. Hence

$$E_t := \bigl\{x\colon\ (Mf)(x) > t\bigr\} \subset \bigl\{x\colon\ (Mg)(x) > t/2\bigr\},$$

which by assertion (ii) yields the estimate

$$\lambda(E_t) \le \lambda\bigl(x\colon\ (Mg)(x) > t/2\bigr) \le \frac{2C_n}{t}\|g\|_{L^1} = \frac{2C_n}{t}\int_{\{|f|>t/2\}} |f(y)|\,dy. \tag{5.6.2}$$

Let $F(t) = \lambda(E_t)$. According to Theorem 3.4.7 and (5.6.2) we have

$$\begin{aligned}\int_{\mathbb{R}^n} |(Mf)(x)|^p\,dx &= p\int_0^\infty t^{p-1}F(t)\,dt \\ &\le p\int_0^\infty t^{p-1}\biggl(\frac{2C_n}{t}\int_{\{|f|>t/2\}} |f(y)|\,dy\biggr)\,dt.\end{aligned}$$

The double integral on the right-hand side of this inequality is evaluated by Fubini's theorem. To this end, it suffices to observe that the integration in t with fixed y yields

$$2C_np|f(y)|\int_0^{2|f(y)|} t^{p-2}\,dt = 2C_np|f(y)|\frac{|2f(y)|^{p-1}}{p-1} = 2^pC_n\frac{p}{p-1}|f(y)|^p.$$

Thus, the indicated double integral equals $2^pC_n\frac{p}{p-1}\|f\|^p_{L^p}$. By Theorem 3.4.7 we obtain $Mf \in L^p(\mathbb{R}^n)$. In addition, the required inequality is true. $\square$

We have seen that any integrable function f for almost all x is recovered from its averages over the intervals $[x-h, x+h]$ by means of the formula

$$f(x) = \lim_{h\to 0} \frac{1}{2h} \int_{x-h}^{x+h} f(y)\, dy.$$

This formula has an important multidimensional generalization, which we now consider.

5.6.2. Theorem. *Suppose a function f is integrable on every ball in $\mathbb{R}^n$. Then, for almost all x, one has*

$$\lim_{r\to 0} \frac{1}{\lambda\bigl(B(x,r)\bigr)} \int_{B(x,r)} |f(y) - f(x)|\, dy = 0. \tag{5.6.3}$$

For every point x with such a property, called a Lebesgue point of f, one has

$$\lim_{r\to 0} \frac{1}{\lambda\bigl(B(x,r)\bigr)} \int_{B(x,r)} f(y)\, dy = f(x). \tag{5.6.4}$$

Proof. First we prove that (5.6.4) holds for almost all x. We may assume that f vanishes outside some ball U (the general case follows in an obvious way). Then $f \in L^1(\mathbb{R}^n)$. We are going to prove that the set Ω of all points $x \in U$ for which there exists a sequence of radii $r_k = r_k(x) \to 0$ such that the quantities

$$\frac{1}{\lambda\bigl(B(x,r_k)\bigr)} \int_{B(x,r_k)} f(y)\, dy$$

do not converge to $f(x)$, has measure zero. We show that for every natural number m, the set E of all points $x \in U$ such that

$$\limsup_{k\to\infty} \frac{1}{\lambda\bigl(B\bigl(x,r_k(x)\bigr)\bigr)} \int_{B(x,r_k(x))} f(y)\, dy \geq f(x) + \frac{1}{m} \tag{5.6.5}$$

has measure zero. Let $\varepsilon > 0$. There exists a continuous function g such that $\|f-g\|_{L^1} < \varepsilon$. Since for the continuous function g equality (5.6.4) is true for every x, we see that, for every $x \in E$, relation (5.6.5) is true for the function $f_1 = f - g$ in place of f. The set E is contained in the set $E_0 = \bigl\{x \in U\colon\ (Mf_1)(x) \geq f_1(x) + m^{-1}\bigr\}$. The measure of E_0 is estimated by means of Chebyshev's inequality and (5.6.1) as follows:

$$\begin{aligned}
\lambda(E_0) &\leq \lambda\Bigl(x \in E_0\colon\ f_1(x) \leq -\frac{1}{2m}\Bigr) + \lambda\Bigl(x \in E_0\colon\ f_1(x) \geq -\frac{1}{2m}\Bigr) \\
&\leq 2m\|f_1\|_{L^1} + \lambda\Bigl(x \in E_0\colon\ (Mf_1)(x) \geq \frac{1}{2m}\Bigr) \\
&\leq 2m\|f_1\|_{L^1} + 2mC_n\|f_1\|_{L^1} \leq 2m\varepsilon(C_n+1).
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we obtain $\lambda(E) = 0$. Replacing f by $-f$ we obtain that the set of all points x where

$$\liminf_{k\to\infty} \frac{1}{\lambda\bigl(B\bigl(x,r_k(x)\bigr)\bigr)} \int_{B(x,r_k(x))} f(y)\, dy \leq f(x) - \frac{1}{m}$$

has measure zero as well. Since m is arbitrary, we obtain that $\lambda(\Omega) = 0$ as required.

According to the previous step, for every $c \in \mathbb{R}^1$, there is a set E_c of measure zero such that for all $x \notin E_c$ one has

$$\lim_{r\to 0} \frac{1}{\lambda\bigl(B(x,r)\bigr)} \int_{B(x,r)} |f(y) - c|\, dy = |f(x) - c|. \tag{5.6.6}$$

Let $\{c_j\}$ be all rational numbers and $E = \bigcup_{j=1}^\infty E_{c_j}$. Let $x \notin E$. Then (5.6.6) is fulfilled for x and all rational c. The estimate $\bigl||f(y)-c|-|f(y)-k|\bigr| \le |c-k|$ yields that equality (5.6.6) remains valid for all real c. Letting $c = f(x)$, we complete the proof. $\square$

The set of all Lebesgue points of the function f is called its *Lebesgue set.*

We note that by the above results, for any function f that is integrable on every ball, one has $Mf(x) \ge |f(x)|$ a.e.

The established properties of the maximal function and the differentiation theorem remain valid for many other families of sets in place of balls. For example, Theorem 5.6.2 is true if $B(x,r)$ is the cube with the edge r and center x.

5.6.3. Corollary. *Let $\mathcal{K}$ be some family of measurable sets in $\mathbb{R}^n$ satisfying the following condition: there is a number $c > 0$ such that for every $K \in \mathcal{K}$, there exists an open ball $K(0,r)$ for which $K \subset K(0,r)$ and $\lambda(K) \ge c\lambda\bigl(K(0,r)\bigr)$. Let a function f be integrable on every ball in $\mathbb{R}^n$. Then, for every point x in the Lebesgue set of f, one has*

$$\lim_{K\in\mathcal{K},\lambda(K)\to 0} \frac{1}{\lambda(K)} \int_K |f(x-y) - f(x)|\, dy = 0. \tag{5.6.7}$$

In particular,

$$\lim_{K\in\mathcal{K},\lambda(K)\to 0} \frac{1}{\lambda(K)} \int_{K+x} f(y)\, dy = f(x). \tag{5.6.8}$$

PROOF. The first assertion is clear from the estimate

$$\frac{1}{\lambda(K)} \int_K |f(x-y) - f(x)|\, dy \le \frac{1}{c\lambda\bigl(K(0,r)\bigr)} \int_{K(0,r)} |f(x-y) - f(x)|\, dy.$$

For the proof of the last assertion we observe that

$$\frac{1}{\lambda(K)} \int_K f(x-y)\, dy - f(x) = \frac{1}{\lambda(K)} \int_K [f(x-y) - f(x)]\, dy$$

and that the family of sets $\{-K\colon\ K \in \mathcal{K}\}$ satisfies the same conditions as the family $\mathcal{K}$. $\square$

However, one cannot replace cubes by arbitrary parallelepipeds with edges parallel to the coordinate axes. More precisely, the following assertion is true. Let $\mathcal{R}_0$ be the family of all centrally symmetric parallelepipeds with edges parallel to the coordinate axes, and let $\mathcal{R}$ be the family of all centrally symmetric parallelepipeds.

5.6.4. Theorem. (i) *There is a function* $f \in \mathcal{L}^1(\mathbb{R}^2)$ *such that*

$$\limsup_{R_0\in\mathcal{R}_0, \operatorname{diam}(R_0)\to 0} \frac{1}{\lambda(R_0)} \int_{R_0} f(x-y)\,dy = +\infty \quad \text{for a.e. } x.$$

(ii) *There is a compact set* $K \subset \mathbb{R}^2$ *of positive measure such that*

$$\liminf_{R\in\mathcal{R}, \operatorname{diam}(R)\to 0} \frac{1}{\lambda(R)} \int_{R} I_K(x-y)\,dy = 0 \quad \text{for all } x.$$

(iii) *If* $f \in \mathcal{L}^p(\mathbb{R}^n)$*, where* $p > 1$*, then*

$$\lim_{R_0\in\mathcal{R}_0, \operatorname{diam}(R_0)\to 0} \frac{1}{\lambda(R_0)} \int_{R_0} f(x-y)\,dy = f(x) \quad \text{for a.e. } x.$$

Proofs can be found in the books Guzmán [**386**], Stein [**906**], which give a thorough discussion of related matters.

5.7. The Henstock–Kurzweil integral

We recall that the Riemann integral of a function f on $[a,b]$ is defined as the limit of the sums $\sum_{i=1}^n f(c_i)(x_i - x_{i-1})$, which must exist as the parameter $\delta := \max_i |x_i - x_{i-1}|$ approaches zero, where arbitrary finite partitions of the interval $[a,b]$ by consequent points x_i and arbitrary points $c_i \in [x_i, x_{i+1}]$ are admissible. This freedom in the choice of the partitioning points x_i and points c_i considerably restricts the class of functions for which the above limit exists. For example, if we allow only partitions into equal intervals and c_i are their centers, then the class of functions "integrable" in such a sense will be considerably broader than the Riemannian one. However, such a straightforward generalization does not lead to a fruitful theory. A more fruitful approach was developed in the works of Kurzweil, Henstock, McShane, and other researchers. In this section, we discuss the principal definitions and results in this direction. A considerably more detailed exposition is found in Gordon [**373**], Swartz [**925**], and other books mentioned in the bibliographical comments.

5.7.1. Definition. *Let* $\delta(\cdot)$ *be a positive function on* $[a,b]$. (i) *A tagged interval is a pair* $(x,[c,d])$*, where* $[c,d] \subset [a,b]$*,* $c < d$ *and* $x \in [c,d]$*. A free tagged interval is a pair* $(x,[c,d])$*, where* $[c,d] \subset [a,b]$ *and* $x \in [a,b]$ (*i.e., here we do not require the inclusion* $x \in [c,d]$).

(ii) *A tagged interval* $(x,[c,d])$ *is subordinate to the function* δ *if we have* $[c,d] \subset \bigl(x-\delta(x), x+\delta(x)\bigr)$*. Similarly, we define the subordination of free tagged intervals.*

The number x is called the tag of the interval $[c,d]$. We consider finite collections $\mathcal{P} = \{(x_i,[c_i,d_i]), i = 1,\ldots,n\}$ that consist of tagged intervals $[c_i,d_i]$ that pairwise have no common inner points. Such intervals will be called non-overlapping. A collection $\mathcal{P}$ is called subordinate to the function δ if every tagged interval in $\mathcal{P}$ is subordinate to δ. If $[a,b] = \bigcup_{j=1}^n [c_j,d_j]$, then $\mathcal{P}$ is called a tagged partition. An analogous terminology is introduced for collections of free tagged intervals; such collections will be denoted by $\widehat{\mathcal{P}}$ to distinguish them from collections $\mathcal{P}$. A collection of non-overlapping free tagged intervals with the union $[a,b]$ will be called a free tagged partition.

5.7.2. Lemma. *For an arbitrary positive function δ, there exists a tagged partition of $[a,b]$ subordinate to δ.*

Proof. Let M be the set of all points $x\in(a,b]$ such that our claim is true for the interval $[a,x]$. Since $\delta(a)>0$, one has $\bigl(a,a+\delta(a)\bigr)\subset M$. The nonempty set M has the supremum m. It is clear that $m\in M$. Indeed, $\delta(m)>0$, hence there exists a point $x\in M$ with $x>m-\delta(m)$. Therefore, to the tagged partition of the interval $[a,x]$ subordinate to δ, one can add the pair $(m,[x,m])$. Finally, we observe that $m=b$. Otherwise there exists a point $x\in(m,b)$ with $x-m<\delta(m)$, which yields that $x\in M$ because to any tagged partition of $[a,m]$ that is subordinate to δ one can add the pair $(m,[m,x])$. □

If f is a function on $[a,b]$, then, to every collection $\mathcal{P}$ of non-overlapping tagged intervals, we associate the sums

$$I(f,\mathcal{P}):=\sum_{i=1}^n f(x_i)(d_i-c_i),\quad \mu(\mathcal{P}):=\sum_{i=1}^n (d_i-c_i).$$

The analogous sums $I(f,\widehat{\mathcal{P}})$ and $\mu(\widehat{\mathcal{P}})$ are associated to all free collections $\widehat{P}$. If $\mathcal{P}$ is a partition of the interval, then $I(f,\mathcal{P})$ is a Riemannian sum; the numbers $I(f,\widehat{\mathcal{P}})$ are called generalized Riemannian sums.

5.7.3. Definition. (i) *A function f on $[a,b]$ is called Henstock–Kurzweil integrable if there exists a number I with the following property: for every $\varepsilon>0$, there exists a function $\delta\colon [a,b]\to(0,+\infty)$ such that $|I(f,\mathcal{P})-I|<\varepsilon$ for every tagged partition $\mathcal{P}$ of the interval $[a,b]$ that is subordinate to the function δ. The number I is called the Henstock–Kurzweil integral of the function f and denoted by the symbol*

$$(\mathcal{HK})\int_a^b f.$$

The function f is called Henstock–Kurzweil integrable on a measurable set $E\subset[a,b]$ if the function fI_E is Henstock–Kurzweil integrable.

(ii) *A function f on $[a,b]$ is called McShane integrable if in* (i) *in place of $\mathcal{P}$ one can take free tagged partitions $\widehat{\mathcal{P}}$. The corresponding number I is called the McShane integral of the function f.*

It is clear that the McShane integrability yields the Henstock–Kurzweil integrability and the two integrals are equal, since any tagged partition is a free tagged partition. As we shall later see, the McShane integral coincides with the Lebesgue integral, which by Lemma 5.7.2 gives a description of the Lebesgue integral by means of Riemannian sums (though, a non-constructive description; cf. Exercise 2.12.63). The Henstock–Kurzweil integral on an interval is more general than the Lebesgue integral, but coincides with the latter for nonnegative functions (or functions bounded from one side). It is clear that all Riemann integrable functions are Henstock–Kurzweil integrable because for δ one can take positive constants. Unlike the Lebesgue integral, the Henstock–Kurzweil integral contains the improper Riemann integral. Similarly to the Riemann definition, the definitions of the Henstock–Kurzweil and McShane integrals can be formulated without explicitly mentioning the values of integrals. The proof of the following lemma is delegated to Exercise 5.8.101.

5.7.4. Lemma. *A function f is Henstock–Kurzweil integrable on $[a,b]$ precisely when for every $\varepsilon > 0$, there exists a positive function δ on $[a,b]$ such that $|I(f,\mathcal{P}_1) - I(f,\mathcal{P}_2)| < \varepsilon$ for all tagged partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ subordinate to δ. The same assertion with free tagged partitions in place of tagged partitions is true for the McShane integral.*

Let us consider the following illuminating example of evaluation of the Henstock–Kurzweil integral, where one can see the role of functions δ.

5.7.5. Example. If $f = 0$ a.e. on $[a,b]$, then the function f is McShane and Henstock–Kurzweil integrable and both integrals equal zero.

PROOF. Let $\varepsilon > 0$. Set

$$E_1 = \bigl\{x\colon\ 0 < |f(x)| < 1\bigr\},\ E_n = \bigl\{x\colon\ n-1 \le |f(x)| < n\bigr\},\ n > 1.$$

The set E_n has measure zero and possesses a neighborhood U_n of measure less than $\varepsilon n^{-1}2^{-n}$. The sets E_n are disjoint. Let

$$\delta(x) = \begin{cases} 1 & \text{if } x \in [a,b]\backslash \bigcup_{n=1}^{\infty} E_n, \\ \operatorname{dist}(x,[a,b]\backslash U_n) & \text{if } x \in E_n. \end{cases}$$

Suppose that $\widehat{\mathcal{P}}$ is a free tagged partition of $[a,b]$ subordinate to δ. By $\widehat{\mathcal{P}}_n$ we denote the subcollection in $\mathcal{P}$ consisting of the pairs $(x,[c,d])$ with $x \in E_n$. It is clear that one has $[c,d] \subset U_n$ for every such pair, since the numbers $|x-c|$ and $|x-d|$ are less than $\delta(x)$, whereas $\delta(x) \ge |x-y|$ for all $y \in [a,b]\backslash U_n$. Then

$$|I(f,\widehat{\mathcal{P}})| \le \sum_{n=1}^{\infty} |I(f,\widehat{\mathcal{P}}_n)| < \sum_{n=1}^{\infty} n\lambda(U_n) < \sum_{n=1}^{\infty} \varepsilon 2^{-n} = \varepsilon.$$

Hence the McShane and Henstock–Kurzweil integrals of f vanish. □

Note that f in the definition of the McShane and Henstock–Kurzweil integrals must be defined everywhere (not just almost everywhere), but this example and the following proposition show that redefinitions of f on measure zero sets do not affect the respective integrabilities.

The proof of the following simple technical assertion is left as Exercise 5.8.102.

5.7.6. Proposition. (i) *If f is Henstock–Kurzweil integrable on $[a,b]$, then f is integrable in the same sense on every interval $[\alpha,\beta]\subset[a,b]$.*

(ii) *If f is Henstock–Kurzweil integrable on $[a,c]$ and $[c,b]$ for some point $c\in(a,b)$, then f is integrable in the same sense on $[a,b]$ and*

$$(\mathcal{HK})\int_a^b f=(\mathcal{HK})\int_a^c f+(\mathcal{HK})\int_c^b f.$$

(iii) *The set $\mathcal{L}_{HK}[a,b]$ of all functions on $[a,b]$ that are Henstock–Kurzweil integrable is a linear space, on which the Henstock–Kurzweil integral is linear.*

(iv) *If $f,g\in\mathcal{L}_{HK}[a,b]$ and $f\le g$ a.e., then*

$$(\mathcal{HK})\int_a^b f\le(\mathcal{HK})\int_a^b g.$$

It is clear from this result that if f is Henstock–Kurzweil integrable on $[a,b]$, then we obtain the following function on $[a,b]$:

$$F(x)=(\mathcal{HK})\int_a^x f. \tag{5.7.1}$$

We know that the function $f(x)=x^2\sin(x^{-4})$, $f(0)=0$, on the real line is differentiable at every point, but f' is not Lebesgue integrable on $[0,1]$. The following important result shows that the function f' is Henstock–Kurzweil integrable on every interval $[a,b]$.

5.7.7. Theorem. *Let f be a continuous function on $[a,b]$ that is differentiable at all points, with the exception of points of some at most countable set $C=\{c_n\}$. Then the function f' (assigned, for example, the zero value at the points from C) is Henstock–Kurzweil integrable on $[a,b]$ and*

$$(\mathcal{HK})\int_a^z f'=f(z)-f(a),\quad \forall\, z\in[a,b].$$

PROOF. Let $\varepsilon>0$. We define the function δ as follows: if $x\notin C$, then, by the differentiability of f at x, there exists $\delta(x)>0$ such that

$$|f(u)-f(x)-f'(x)(u-x)|\le\varepsilon|u-x|, \tag{5.7.2}$$
$$\forall\, u\in\bigl(x-\delta(x),x+\delta(x)\bigr)\cap[a,b].$$

If $x=c_n$, then, by the continuity of f, one has $\delta(x)>0$ such that

$$|f(u)-f(v)|<\varepsilon 2^{-n},\ \forall\, u,v\in\bigl(x-\delta(x),x+\delta(x)\bigr)\cap[a,b]. \tag{5.7.3}$$

Let $\mathcal{P}=\{(x_i,[a_i,b_i]),i\le n\}$ be a tagged partition of $[a,b]$ subordinate to δ, J_0 the collection of all indices i with $x_i\in C$, J_1 the collection of the remaining indices i, and let $\mathcal{P}_0$ and $\mathcal{P}_1$ be the subcollections in $\mathcal{P}$, corresponding to J_0 and J_1. Then, for all $i\in J_1$, we obtain from (5.7.2) that

$$|f(b_i)-f(a_i)-f'(x_i)(b_i-a_i)|\le\varepsilon(b_i-a_i).$$

We observe that $\sum_{i\in J_0}|f(b_i)-f(a_i)|\le 2\varepsilon$. If all x_i with $i\in J_0$ are distinct, then this follows at once from (5.7.3), and one even has ε in place of 2ε. In the general case, multiple x_i may only occur as the endpoints of two adjacent intervals $[a_i, b_i]$ and $[a_j, b_j]$ with $b_i = a_j = x_i = x_j$. Hence on account of our setting $f' = 0$ on C, we obtain

$$\begin{aligned}\bigl|I(f',\mathcal{P}) - [f(b) - f(a)]\bigr| \le \Bigl|I(f',\mathcal{P}_1) - \sum_{i\in J_1}[f(b_i) - f(a_i)]\Bigr| \\ + \sum_{i\in J_0}|f(b_i) - f(a_i)| \le \varepsilon(b-a) + 2\varepsilon,\end{aligned}$$

which proves our claim for $z = b$, hence for all $z \in [a, b]$. □

In particular, the Henstock–Kurzweil integral (unlike the Lebesgue one) solves the problem of recovering any everywhere differentiable function f from f', although not at all as constructively as the Lebesgue integral does for absolutely continuous functions f (for example, in the above theorem the function δ is constructed by using the function f which we want to "recover"). We shall state without proof (which can be read in Gordon [**373**, Ch. 9]) a theorem, which shows, in particular, that the Henstock–Kurzweil integral contains the improper Riemann integral.

5.7.8. Theorem. *Let a function f be defined on $[a, b]$ and Henstock–Kurzweil integrable on every interval $[c, d]$, where $c > a$, $d < b$, such that the integrals*

$$(\mathcal{HK})\int_c^d f$$

have a finite limit as $c \to a$, $d \to b$. Then the Henstock–Kurzweil integral of the function f on the interval $[a, b]$ exists and equals the indicated limit.

Below we shall need the following lemma, which is frequently used in the theory of the Henstock–Kurzweil integral. Its proof is delegated to Exercise 5.8.103.

5.7.9. Lemma. *Suppose that a function f on $[a, b]$ is Henstock–Kurzweil integrable and F is defined by (5.7.1). Let $\varepsilon > 0$ and let δ be a positive function such that $|I(f,\mathcal{P}) - F(b)| < \varepsilon$ for every tagged partition $\mathcal{P}$ of $[a, b]$ subordinate to δ. Then, for every finite collection $\mathcal{P}_0 = \{(x_i, [c_i, d_i]),\ i = 1, \ldots, n\}$ of non-overlapping tagged intervals subordinate to δ, one has*

$$\Bigl|I(f,\mathcal{P}_0) - \sum_{i=1}^n [F(d_i) - F(c_i)]\Bigr| \le \varepsilon,$$

$$\sum_{i=1}^n \bigl|f(x_i)(d_i - c_i) - [F(d_i) - F(c_i)]\bigr| \le 2\varepsilon.$$

The next important theorem shows, in particular, the measurability of all Henstock–Kurzweil integrable functions, which is not obvious from the definition.

5.7.10. Theorem. *Let a function f on $[a,b]$ be Henstock–Kurzweil integrable and let the function F be defined by equality* (5.7.1). *Then F is continuous on $[a,b]$ and almost everywhere has the derivative $F'(x) = f(x)$. In particular, the function f is measurable.*

PROOF. Let $c \in [a,b]$ and $\varepsilon > 0$. We take a positive function δ corresponding to ε in the definition of the integral. Let

$$\eta := \min\Bigl(\delta(c), \varepsilon\bigl(1+|f(c)|\bigr)^{-1}\Bigr).$$

If $|x-c| < \eta$, then the pair $(c,[x,c])$ is subordinate to δ. By the second estimate in Lemma 5.7.9 we obtain

$$|F(c)-F(x)| \le |F(c)-F(x)-f(c)(c-x)| + |f(c)(c-x)| < 3\varepsilon.$$

The continuity of F is proven. We now prove that $D^+F(x) = f(x)$ almost everywhere. Other derivates are considered similarly. Set

$$A := \bigl\{x \in [a,b)\colon\ D^+F(x) \ne f(x)\bigr\}.$$

For every $x \in A$, there exists $r(x) > 0$ with the following property: for each $h > 0$, there exists $y_{x,h} \in (x,x+h) \cap [a,b)$ with

$$|F(y_{x,h}) - F(x) - f(x)(y_{x,h}-x)| \ge r(x)(y_{x,h}-x).$$

Let $A_n = \{x \in A\colon\ r(x) \ge 1/n\}$. It suffices to verify that $\lambda^*(A_n) = 0$ for all $n \in \mathbb{N}$. Let us fix $\alpha > 0$ and find a positive function δ corresponding to the number $\varepsilon = \alpha(4n)^{-1}$ in the definition of the Henstock–Kurzweil integral. Since the intervals $[x, y_{x,h}]$, where $x \in A_n$ and $0 < h < \delta(x)$, cover A_n in the sense of Vitali, one can choose a finite collection of disjoint intervals $[c_i,d_i]$, $1 \le i \le k$, such that

$$\lambda^*(A_n) < \sum_{i=1}^k (d_i - c_i) + \alpha/2.$$

By construction, the collection of tagged intervals $(c_i,[c_i,d_i])$ is subordinate to the function δ and

$$|F(d_i) - F(c_i) - f(c_i)(d_i-c_i)| \ge r(c_i)(d_i-c_i).$$

On account of the established estimates and Lemma 5.7.9 we obtain

$$\begin{aligned}\sum_{i=1}^k (d_i-c_i) &\le \sum_{i=1}^k \frac{1}{r(c_i)}\bigl|F(d_i)-F(c_i)-f(c_i)(d_i-c_i)\bigr| \\ &\le n\sum_{i=1}^k \bigl|f(c_i)(d_i-c_i) - [F(d_i)-F(c_i)]\bigr| \le n\frac{2\alpha}{4n} = \frac{\alpha}{2}.\end{aligned}$$

Hence $\lambda^*(A_n) < \alpha$. Finally, we obtain $\lambda(A_n) = 0$. □

5.7.11. Corollary. *If a function f on $[a,b]$ is Henstock–Kurzweil integrable and is bounded from above or below, then it is Lebesgue integrable.*

PROOF. We may assume that $f \geq 0$. The function F defined by formula (5.7.1) is increasing. Therefore, almost everywhere it has the derivative F' that is Lebesgue integrable. The above theorem yields that $F'(x) = f(x)$ almost everywhere. Thus, f is Lebesgue integrable. $\square$

5.7.12. Corollary. *If a function f is Henstock–Kurzweil integrable on every measurable set $E \subset [a, b]$, then it is Lebesgue integrable.*

PROOF. The function f is measurable by the above theorem. By our hypothesis the functions $fI_{\{f\geq 0\}}$ and $fI_{\{f<0\}}$ are Henstock–Kurzweil integrable. According to the previous corollary these functions are Lebesgue integrable, which yields our assertion. $\square$

It is interesting that if a function is Henstock–Kurzweil integrable, then one can always choose a measurable function δ in Definition 5.7.3 (see Gordon [**373**, Theorem 9.24]).

We observe that so far it is not obvious that the simultaneous existence of the Henstock–Kurzweil and Lebesgue integrals implies their equality. One way to establish this equality is to compare both integrals with the McShane integral, to the consideration of which we now proceed.

5.7.13. Lemma. *If a function f on $[a, b]$ is McShane integrable, then so is the function $|f|$.*

PROOF. Given $\varepsilon > 0$, we choose a positive function δ such that

$$|I(f, \widehat{\mathcal{P}}) - I(f, \widehat{\mathcal{P}}')| < \varepsilon$$

for any free tagged partitions $\widehat{\mathcal{P}}$ and $\widehat{\mathcal{P}}'$ subordinate to δ. Let

$$\widehat{\mathcal{P}}_1 = \{(x_i, I_i), i \leq N_1\} \quad \text{and} \quad \widehat{\mathcal{P}}_2 = \{(y_j, K_j), j \leq N_2\}$$

be free tagged partitions subordinate to δ. We take nondegenerate intervals of the form $I_i \cap K_j$ and obtain free tagged partitions

$$\widehat{\mathcal{P}}_1' = \{(x_i, I_i \cap K_j), i \leq N_1, j \leq N_2\},$$
$$\widehat{\mathcal{P}}_2' = \{(y_j, I_i \cap K_j), i \leq N_1, j \leq N_2\}$$

subordinate to δ. One has $I(|f|, \widehat{\mathcal{P}}_1') = I(|f|, \widehat{\mathcal{P}}_1)$, $\quad I(|f|, \widehat{\mathcal{P}}_2') = I(|f|, \widehat{\mathcal{P}}_2)$. Therefore,

$$\begin{aligned}\Bigl|I(|f|, \widehat{\mathcal{P}}_1) - I(|f|, \widehat{\mathcal{P}}_2)\Bigr| &= \Bigl|I(|f|, \widehat{\mathcal{P}}_1') - I(|f|, \widehat{\mathcal{P}}_2')\Bigr| \\ &\leq \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \bigl||f(x_i)| - |f(y_j)|\bigr| \lambda(I_i \cap K_j) \\ &\leq \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} |f(x_i) - f(y_j)| \lambda(I_i \cap K_j).\end{aligned}$$

Finally, we observe that the right-hand side of this inequality is less than ε. Indeed, let us consider the two free tagged partitions $\widehat{\mathcal{P}}$ and $\widehat{\mathcal{P}}'$ subordinate to δ and defined as follows: if $f(x_i) \geq f(y_j)$, then we include $(x_i, I_i \cap K_j)$ in $\widehat{\mathcal{P}}$ and $(y_j, I_i \cap K_j)$ in $\widehat{\mathcal{P}}'$; if $f(x_i) < f(y_j)$, then we include $(y_j, I_i \cap K_j)$ in $\widehat{\mathcal{P}}$ and $(x_i, I_i \cap K_j)$ in $\widehat{\mathcal{P}}'$. Then one has

$$I(f, \widehat{\mathcal{P}}) - I(f, \widehat{\mathcal{P}}') = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} |f(x_i) - f(y_j)| \lambda(I_i \cap K_j),$$

which completes the proof. □

5.7.14. Theorem. *A function f on $[a,b]$ is McShane integrable precisely when it is Lebesgue integrable. In this case, both integrals coincide.*

PROOF. (i) We may assume that $[a,b] = [0,1]$. Let f be Lebesgue integrable on $[0,1]$ and let $0 < \varepsilon < 1$. We find a positive number $\eta < \varepsilon/3$ such that the Lebesgue integral of $|f|$ over a set A is less than ε whenever $\lambda(A) < \eta$. Let

$$E_n = \bigl\{x\colon\ (n-1)\varepsilon/4 < f(x) \leq \varepsilon n/4\bigr\}, \quad n \in \mathbb{Z}.$$

The sets E_n are measurable, disjoint and cover $[0,1]$. We find an open set $U_n \supset E_n$ with $\lambda(U_n \backslash E_n) < \eta 2^{-|n|}(3|n|+3)^{-1}$. Let us consider the function δ defined as follows: if $x \in E_n$, then

$$\delta(x) = \operatorname{dist}(x, [0,1] \backslash U_n).$$

Suppose that $\widehat{\mathcal{P}} = \{(x_i, [a_i, b_i]), i \leq k\}$ is a free tagged partition of $[0,1]$ subordinate to δ. For every i, there is a unique number n_i with $x_i \in E_{n_i}$. Set $A_i = [a_i, b_i] \cap E_{n_i}$, $B_i = [a_i, b_i] \backslash E_{n_i}$. If $x \in A_i$, then $|f(x_i) - f(x)| \leq \varepsilon/4$. Further, for every integer n, let $J_n := \{j\colon\ n_j = n\}$ and $C_n := \bigcup_{i \in J_n} B_i$. Then the definition of δ yields

$$C_n = \bigcup_{i \in J_n} ([a_i, b_i] \backslash E_n) \subset U_n \backslash E_n,$$

which on account of the inclusion $x_i \in E_n$ for all $i \in J_n$ yields

$$\sum_{i=1}^{k} |f(x_i)| \lambda(B_i) \leq \sum_{n=-\infty}^{\infty} \frac{1}{4} \varepsilon(|n|+1) \lambda(C_n) \leq \sum_{n=-\infty}^{\infty} \frac{1}{4} \varepsilon(|n|+1) \lambda(U_n \backslash E_n) < \frac{\varepsilon}{4}.$$

Finally, since the set $C = \bigcup_{n \in \mathbb{Z}} C_n$ has measure at most $\sum_{n \in \mathbb{Z}} \lambda(U_n \backslash C_n) < \eta$, one has

$$\sum_{i=1}^{k} \int_{B_i} |f(t)| \, dt \leq \int_C |f(t)| \, dt < \frac{\varepsilon}{3}.$$

It remains to apply the following estimate for sums and the Lebesgue integrals, where we use in addition that the integral of $|f(x_i) - f(x)|$ over A_i does not exceed $\varepsilon\lambda(A_i)/4$:

$$\Big|I(f,\widehat{\mathcal{P}}) - \int_0^1 f(x)\,dx\Big| = \Big|\sum_{i=1}^k \int_{a_i}^{b_i} [f(x_i) - f(x)]\,dx\Big|$$
$$\le \sum_{i=1}^k \int_{A_i} |f(x_i) - f(x)|\,dx + \sum_{i=1}^k |f(x_i)|\lambda(B_i) + \sum_{i=1}^k \int_{B_i} |f(x)|\,dx < \varepsilon.$$

Thus, the McShane integral of the function f exists and equals its Lebesgue integral.

(ii) Let f be McShane integrable. By Lemma 5.7.13, the function $|f|$ is integrable in the same sense. Then f is Lebesgue integrable because f and $|f|$ are Henstock–Kurzweil integrable. $\square$

A natural question arises regarding what happens if in the definition of the Henstock–Kurzweil integral we admit general measurable sets in place of tagged intervals. It turns out that this also leads to the Lebesgue integral (see Exercise 5.8.132).

5.8. Supplements and exercises

5.8(i). Covering theorems

The following interesting covering theorem is due to A.S. Besicovitch.

5.8.1. Theorem. *For every $n \in \mathbb{N}$, there exists a number $N_n \in \mathbb{N}$ such that for every collection $\mathcal{F}$ of nondegenerate closed balls in $\mathbb{R}^n$ with uniformly bounded radii, one can find subcollections $\mathcal{F}_1, \ldots, \mathcal{F}_{N_n} \subset \mathcal{F}$, each of which consists of at most countably many disjoint balls such that the set of centers of all balls in $\mathcal{F}$ is covered by the balls from $\mathcal{F}_1 \cup \cdots \cup \mathcal{F}_{N_n}$.*

PROOF. Balls in $\mathcal{F}$ are denoted by $B(a,r)$; let A be the set of their centers. Suppose first that A is bounded. Let $R = \sup\{r\colon B(a,r) \in \mathcal{F}\}$. We can find $B_1 = B(a_1,r_1) \in \mathcal{F}$ with $r_1 > 3R/4$. The balls B_j, $j > 1$, are chosen inductively as follows. Let $A_j = A \backslash \bigcup_{i=1}^{j-1} B_i$. If the set A_j is empty, then our construction is completed and, letting $J := j-1$, we obtain J balls $B_1, \ldots, B_J$. If A_j is nonempty, then we choose $B_j = B(a_j,r_j) \in \mathcal{F}$ such that

$$a_j \in A_j \quad \text{and} \quad r_j > \frac{3}{4}\sup\bigl\{r\colon B(a,r) \in \mathcal{F}, a \in A_j\bigr\}.$$

In the case of an infinite sequence of balls B_j we set $J = \infty$. Note the following properties of the constructed objects:

(a) if $j > i$, then $r_j \le 4r_i/3$,
(b) the balls $B(a_j, r_j/3)$ are disjoint and if $J = \infty$, then $r_j \to 0$,
(c) $A \subset \bigcup_{j=1}^{J} B(a_j, r_j)$.

Property (a) follows by the definition of r_i and the inclusion $a_j \in A_j \subset A_i$. Now (b) is seen from the fact that if $j > i$, then $a_j \notin B_i$, whence

$$|a_i - a_j| > r_i > \frac{r_i}{3} + \frac{r_j}{3}$$

according to (a). By the boundedness of A we obtain $r_j \to 0$ in the case of an infinite sequence. Finally, (c) is obvious if $J < \infty$. If $J = \infty$ and $B(a, r) \in \mathcal{F}$, then there exists r_j with $r_j < 3r/4$, whence $a \in \bigcup_{i=1}^{j-1} B_i$ by our construction of r_j.

We fix $k > 1$ and let

$$I_k := \{j\colon\ j < k, B_j \cap B_k \ne \varnothing\}, \quad M_k := I_k \cap \{j\colon\ r_j \le 3r_k\}.$$

Let us show that

$$\operatorname{Card}(M_k) \le 20^n. \tag{5.8.1}$$

Indeed, if $j \in M_k$ and $x \in B(a_j, r_j/3)$, then the balls B_j and B_k have a nonempty intersection and $r_j \le 3r_k$, which yields

$$|x - a_k| \le |x - a_j| + |a_j - a_k| \le \frac{r_j}{3} + r_j + r_k \le 5r_k,$$

i.e., $B(a_j, r_j/3) \subset B(a_k, 5r_k)$. By the disjointness of $B(a_j, r_j/3)$ and property (a) we obtain

$$\begin{aligned}\lambda_n\bigl(B(a_k, 5r_k)\bigr) &\ge \sum_{j\in M_k} \lambda_n\bigl(B(a_j, r_j/3)\bigr) = \sum_{j\in M_k} C_n r_j^n 3^{-n} \\ &\ge \sum_{j\in M_k} C_n r_k^n 4^{-n} = \operatorname{Card}(M_k) C_n r_k^n 4^{-n},\end{aligned}$$

where $\lambda_n\bigl(B(a, r)\bigr) = C_n r^n$. Hence the obtained estimates yield the inequality $5^n \ge \operatorname{Card}(M_k) 4^{-n}$.

Now we estimate the cardinality of $I_k \backslash M_k$. Let us consider two distinct elements $i, j \in I_k \backslash M_k$. Then $1 \le i, j < k$, $B_i \cap B_k \ne \varnothing$, $B_j \cap B_k \ne \varnothing$, $r_i > 3r_k$, $r_j > 3r_k$. One has $|a_i| \le r_i + r_k$ and $|a_j| \le r_j + r_k$. Let $\theta \in [0, \pi]$ be the angle between $a_i - a_k$ and $a_j - a_k$. Our next step is to prove the estimate

$$\theta \ge \theta_0 := \arccos 61/64 > 0. \tag{5.8.2}$$

For notational simplicity, we shall assume that $a_k = 0$. Then $0 = a_k \notin B_i \cup B_j$ and $r_i < |a_i|$, $r_j < |a_j|$. In addition, we can assume that $|a_i| \le |a_j|$. Hence

$$3r_k < r_i < |a_i| \le r_i + r_k, \quad 3r_k < r_j < |a_j| \le r_j + r_k, \quad |a_i| \le |a_j|.$$

We observe that if $\cos\theta > 5/6$, then $a_i \in B_j$. Indeed, if we have $|a_i - a_j| \ge |a_j|$, then

$$\cos\theta = \frac{|a_i|^2 + |a_j|^2 - |a_i - a_j|^2}{2|a_i||a_j|} \le \frac{|a_i|}{2|a_j|} \le \frac{1}{2} < \frac{5}{6}.$$

If $|a_i - a_j| \le |a_j|$, but $a_i \notin B_j$, then $r_j < |a_i - a_j|$ and hence

$$\cos\theta = \frac{|a_i|^2 + |a_j|^2 - |a_i - a_j|^2}{2|a_i||a_j|} \le \frac{|a_i|}{2|a_j|} + \frac{(|a_j| - |a_i - a_j|)(|a_j| + |a_i - a_j|)}{2|a_i||a_j|}$$
$$\le \frac{1}{2} + \frac{r_j + r_k - r_j}{r_i} \le \frac{5}{6},$$

where we used the inequality $|a_j| + |a_i - a_j| \le 2|a_j|$.

We now prove the following assertion:

$$0 \le |a_i - a_j| + |a_i| - |a_j| \le \frac{8}{3}(1 - \cos\theta)|a_j| \quad \text{if } a_i \in B_j. \tag{5.8.3}$$

Indeed, since $a_i \in B_j$, one has $i < j$. Hence $a_j \notin B_i$ and one has $|a_i - a_j| > r_i$. Then (keeping our convention that $|a_i| \le |a_j|$)

$$0 \le \frac{|a_i - a_j| + |a_i| - |a_j|}{|a_j|} \le \frac{|a_i - a_j| + |a_i| - |a_j|}{|a_j|} \frac{|a_i - a_j| - |a_i| + |a_j|}{|a_i - a_j|}$$
$$= \frac{|a_i - a_j|^2 - (|a_j| - |a_i|)^2}{|a_j||a_i - a_j|} = \frac{2|a_i|(1 - \cos\theta)}{|a_i - a_j|}$$
$$\le \frac{2(r_i + r_k)(1 - \cos\theta)}{r_i} \le \frac{8}{3}(1 - \cos\theta).$$

Now we arrive at (5.8.2). Indeed, if $\cos\theta \le 5/6$, then $\cos\theta \le 61/64$. If $\cos\theta > 5/6$, then, according to what we have shown, $a_i \in B_j$. Then $i < j$ and hence $a_j \notin B_i$. Thus, $r_i < |a_i - a_j| \le r_j$. In addition, $r_j \le 4r_i/3$. Therefore, by the estimate $r_j > 3r_k$, we obtain

$$|a_i - a_j| + |a_i| - |a_j| \ge r_i + r_i - r_j - r_k \ge \frac{r_j}{2} - r_k \ge \frac{1}{8}(r_j + r_k) \ge \frac{1}{8}|a_j|,$$

which yields $|a_j|/8 \le 8(1 - \cos\theta)|a_j|/3$ by (5.8.3). Hence $\cos\theta \le 61/64$.

It follows that there exists a number $K_n \in \mathbb{N}$, depending only on n, such that

$$\mathrm{Card}(I_k \backslash M_k) \le K_n. \tag{5.8.4}$$

Indeed, let us fix $\delta > 0$ such that if x is a vector with $|x| = 1$ and $y, z \in B(x, \delta)$, then the angle between y and z is less than $\theta_0 = \arccos 61/64$. Let K_n be the smallest natural number among numbers l such that the unit sphere can be covered by l balls of radius δ with centers in this sphere. Then the sphere ∂B_k can be covered by K_n balls of radius δr_k with centers in ∂B_k. According to (5.8.2), for all distinct $i, j \in I_k \backslash M_k$, the angle between a_i and a_j (we assume that $a_k = 0$) is at most θ_0, whence it is seen that the rays generated by the vectors a_i and a_j cannot meet one and the same ball of radius δ and center in ∂B_k. In particular, they cannot meet one and the same ball from the above taken cover by K_n balls. This yields (5.8.4).

Now we set $L_n = 20^n + K_n + 1$, $N_n = 2L_n$. Then

$$\mathrm{Card}(I_k) = \mathrm{Card}(M_k) + \mathrm{Card}(I_k \backslash M_k) \le 20^n + K_n < L_n.$$

Let us make our choice of $\mathcal{F}_i$. We define a mapping

$$\sigma\colon \{1,2,\ldots\} \to \{1,\ldots,L_n\}$$

as follows: $\sigma(i) = i$ if $1 \le i \le L_n$. If $k \ge L_n$, we define $\sigma(k+1)$ as follows: as noted above,

$$\mathrm{Card}\{j\colon\ 1 \le j \le k, B_j \cap B_{k+1} \neq \varnothing\} < L_n,$$

i.e., there exists the smallest number $l \in \{1,\ldots,L_n\}$ with $B_{k+1} \cap B_j = \varnothing$ for all $j \in \{1,\ldots,k\}$ such that $\sigma(j) = l$. We set $\sigma(k+1) = l$. Finally, let

$$\mathcal{F}_j := \{B_i\colon\ \sigma(i) = j\}, \quad j \le L_n.$$

It is clear from the definition of σ that every collection $\mathcal{F}_j$ consists of disjoint balls. It is easily seen that every ball B_i belongs to some collection $\mathcal{F}_j$, whence one has

$$A \subset \bigcup_{j=1}^{J} B_j = \bigcup_{j=1}^{L_n} \bigcup_{B\in\mathcal{F}_j} B.$$

It remains to consider the case of an unbounded set A. Let

$$A_l = A \cap \{x\colon\ 6R(l-1) \le |x| < 6Rl\},$$

and let $\mathcal{F}^l$ denote the family of all balls in $\mathcal{F}$ with the centers in A_l. As we have proved, for every l, there exists an at most countable subcollection $\mathcal{F}_j^l$, $j \le L_n$, of disjoint balls such that their union covers A_l. Since the radii of all balls do not exceed R, no ball in the collection $\mathcal{F}^l$ can meet a ball in the collection $\mathcal{F}^{l+2}$. It remains to take, for every $j \le L_n$, the collection $\mathcal{F}_j = \bigcup_{l=1}^\infty \mathcal{F}_j^{2l-1}$ and the collection $\mathcal{F}_j' = \bigcup_{l=1}^\infty \mathcal{F}_j^{2l}$, which completes our proof. □

5.8.2. Corollary. *Let $\mathfrak{m}$ be a Carathéodory outer measure on $\mathbb{R}^n$ such that $\mathcal{B}(\mathbb{R}^n) \subset \mathfrak{M}_\mathfrak{m}$. Suppose that $\mathcal{F}$ is a collection of nondegenerate closed balls, the set of centers of which is denoted by A, such that $\mathfrak{m}(A) < \infty$ and, for every $a \in A$ and every $\varepsilon > 0$, $\mathcal{F}$ contains a ball $K(a,r)$ with $r < \varepsilon$. Then, for every nonempty open set $U \subset \mathbb{R}^n$, one can find an at most countable collection of disjoint balls $B_j \in \mathcal{F}$ such that*

$$\bigcup_{j=1}^\infty B_j \subset U \quad \textit{and} \quad \mathfrak{m}\Bigl((A\cap U)\backslash \bigcup_{j=1}^\infty B_j\Bigr) = 0.$$

PROOF. Let N_n be the constant from the above theorem. We fix a number $\alpha \in (1 - 1/N_n, 1)$. Let us show that $\mathcal{F}$ contains a finite collection of disjoint balls $B_1,\ldots,B_{k_1}$ with the following property:

$$\bigcup_{j=1}^{k_1} B_j \subset U, \quad \mathfrak{m}\Bigl((A\cap U)\backslash \bigcup_{j=1}^{k_1} B_j\Bigr) \le \alpha\mathfrak{m}(A\cap U). \tag{5.8.5}$$

To this end, we denote by $\mathcal{F}^1$ the part of $\mathcal{F}$ consisting of the balls of radius at most 1 contained in U. By the Besicovitch theorem, there exist collections $\mathcal{F}^1_1, \ldots, \mathcal{F}^1_{N_n}$ each of which consists of disjoint balls from $\mathcal{F}^1$ such that

$$A \cap U \subset \bigcup_{j=1}^{N_n} \bigcup_{B \in \mathcal{F}^1_j} B.$$

Hence

$$\mathfrak{m}(A \cap U) \le \sum_{j=1}^{N_n} \mathfrak{m}\Bigl((A \cap U) \cap \Bigl(\bigcup_{B \in \mathcal{F}^1_j} B\Bigr)\Bigr).$$

So, there exists $j \in \{1, \ldots, N_n\}$ with

$$\mathfrak{m}\Bigl((A \cap U) \cap \Bigl(\bigcup_{B \in \mathcal{F}^1_j} B\Bigr)\Bigr) \ge \frac{1}{N_n}\mathfrak{m}(A \cap U).$$

Therefore, there exists a finite collection $B_1, \ldots, B_{k_1} \in \mathcal{F}^1_j$ such that

$$\mathfrak{m}\Bigl((A \cap U) \cap \Bigl(\bigcup_{i=1}^{k_1} B_i\Bigr)\Bigr) \ge (1-\alpha)\mathfrak{m}(A \cap U),$$

which yields (5.8.5), since

$$\mathfrak{m}(A \cap U) = \mathfrak{m}\Bigl((A \cap U) \cap \Bigl(\bigcup_{i=1}^{k_1} B_i\Bigr)\Bigr) + \mathfrak{m}\Bigl((A \cap U)\backslash\Bigl(\bigcup_{i=1}^{k_1} B_i\Bigr)\Bigr)$$

by the $\mathfrak{m}$-measurability of the sets B_i.

Now we set $U_2 := U \backslash \bigcup_{j=1}^{k_1} B_j$ and consider the family $\mathcal{F}^2$ of all balls in $\mathcal{F}$ contained in U_2 with radius at most 1. The set U_2 is open. Hence there exists a finite collection of disjoint balls $B_{k_1+1}, \ldots, B_{k_2}$ from $\mathcal{F}^2$ with $\bigcup_{j=k_1+1}^{k_2} B_j \subset U_2$ and

$$\mathfrak{m}\Bigl((A \cap U)\backslash \bigcup_{j=1}^{k_2} B_j\Bigr) = \mathfrak{m}\Bigl((A \cap U_2)\backslash \bigcup_{j=k_1+1}^{k_2} B_j\Bigr) \le \alpha\mathfrak{m}(A \cap U_2) \le \alpha^2\mathfrak{m}(A \cap U).$$

By induction, we obtain a sequence of disjoint balls B_j in $\mathcal{F}$ such that

$$\mathfrak{m}\Bigl((A \cap U)\backslash \bigcup_{j=1}^{k_p} B_j\Bigr) \le \alpha^p\mathfrak{m}(A \cap U).$$

Since $\mathfrak{m}(A) < \infty$ and $\alpha^p \to 0$, we obtain the required collection. □

We observe that the set A may not be $\mathfrak{m}$-measurable.

5.8.3. Corollary. *Let $\mathfrak{m}$ be a Carathéodory outer measure on $\mathbb{R}^n$ such that $\mathcal{B}(\mathbb{R}^n) \subset \mathfrak{M}_{\mathfrak{m}}$. Then, for every nonempty open set $U \subset \mathbb{R}^n$ such that $\mathfrak{m}(U) < \infty$, there exists an at most countable collection of nondegenerate open balls $B_j \subset U$ with the pairwise disjoint closures such that* $\mathfrak{m}\bigl(U \backslash \bigcup_{j=1}^{\infty} B_j\bigr) = 0$.

PROOF. For every point $a \in U$, we take all closed balls $B(a,r) \subset U$ with $r > 0$ and $\mathfrak{m}\bigl(\partial B(a,r)\bigr) = 0$. By the countable additivity of $\mathfrak{m}$ on $\mathfrak{M}_{\mathfrak{m}}$ and the condition that $\mathfrak{m}(U) < \infty$, the continuum of sets $\partial B(a,r)$, $r > 0$, contains at most countably many sets of positive measure. Therefore, one can find numbers $r_j(a) \to 0$ with $\mathfrak{m}\bigl(\partial B\bigl(a, r_j(a)\bigr)\bigr) = 0$. It remains to observe that the set of centers of our balls coincides with U and apply the previous corollary. $\square$

Important applications of covering theorems are connected with differentiation of measures (see §5.8(iii) below).

5.8(ii). Density points and Lebesgue points

Let A be a measurable set in $\mathbb{R}^n$ equipped with Lebesgue measure λ_n. A point $x \in \mathbb{R}^n$ is called a *density point* (or a *point of density*) of A if

$$\lim_{r\to 0} \frac{\lambda_n\bigl(A \cap B(x,r)\bigr)}{\lambda_n\bigl(B(x,r)\bigr)} = 1.$$

A density point of a set may not belong to this set. Since, by Theorem 5.6.2, almost every point x is a Lebesgue point of the function I_A, we see that almost every point $x \in A$ is a density point of A. In particular, every set of positive measure has density points. If the above limit exists (not necessarily equal to 1), then it is called the density of A at x and we say that the set A has density at the point x. It is clear that if x is a density point of a measurable set A, then the complement of A has zero density at x. Let us give some applications of Lebesgue points.

5.8.4. Theorem. *Let ϱ be an integrable function on $\mathbb{R}^n$ that is bounded on balls and has the integral* 1*, let $\varrho_\varepsilon(y) = \varepsilon^{-n}\varrho(y/\varepsilon)$, $\varepsilon > 0$, and let f be a bounded measurable function on $\mathbb{R}^n$. Suppose that x_0 is a Lebesgue point of the function f. Then*

$$f(x_0) = \lim_{\varepsilon\to 0} f * \varrho_\varepsilon(x_0). \tag{5.8.6}$$

PROOF. Let $|f| \le C$. We may assume that $f(x_0) = 0$. Let $\delta > 0$ be fixed. There exists $R > 0$ such that

$$\int_{|y|\ge R} |\varrho(y)|\, dy \le \frac{\delta}{2C+1}.$$

Let $M = R^n \sup_{|z|\le R} |\varrho(z)|$. Since x_0 is a Lebesgue point and $f(x_0) = 0$, there exists $r_0 > 0$ such that for all $r \in (0, r_0)$ one has

$$\frac{1}{r^n} \int_{|y|\le r} |f(x_0 - y)|\, dy \le \frac{\delta}{2M+1}.$$

Now let $0 < \varepsilon < r_0 R^{-1}$. Then $r := \varepsilon R < r_0$. Hence on account of the estimate $R^n|\varrho(y/\varepsilon)| \le M$ for all $|y| \le r$ we obtain

$$\begin{aligned}\Bigl|\int_{\mathbb{R}^n} f(x_0-y)\varrho_\varepsilon(y)\,dy\Bigr| &\le \int_{\mathbb{R}^n} |f(x_0-y)||\varrho_\varepsilon(y)|\,dy\\ &\le \int_{|y|\le r} |f(x_0-y)||\varrho_\varepsilon(y)|\,dy + \int_{|y|>r} C|\varrho_\varepsilon(y)|\,dy\\ &= \int_{|y|\le r} r^{-n}|f(x_0-y)|R^n|\varrho(y/\varepsilon)|\,dy + C\int_{|y|\ge R} |\varrho(z)|\,dz\\ &\le \frac{\delta M}{2M+1} + \frac{\delta C}{2C+1} < \delta,\end{aligned}$$

which proves our assertion. □

An analogous claim is valid under some other conditions on ϱ, which is discussed in Stein [**905**], [**906**].

5.8.5. Theorem. *Let f be a 2π-periodic function integrable on $[0,2\pi]$. Then, for every Lebesgue point of f, one has*

$$f(x) = \lim_{n\to\infty} \sigma_n(x) = \frac{a_0}{2} + \lim_{r\to 1-} \sum_{n=1}^{\infty} [a_n \cos nx + b_n \sin nx] r^n,$$

where $\sigma_n(x)$ is Fejér's sum (4.3.7) *and a_n and b_n are defined by* (4.3.5).

The proof is left as Exercise 5.8.93.

5.8(iii). Differentiation of measures on $\mathbb{R}^n$

Let μ and ν be two nonnegative Borel measures on $\mathbb{R}^n$ that are finite on all balls. For any $x \in \mathbb{R}^n$ we set

$$\overline{D}_\mu \nu(x) := \limsup_{r\to 0} \frac{\nu\bigl(B(x,r)\bigr)}{\mu\bigl(B(x,r)\bigr)},$$

$$\underline{D}_\mu \nu(x) := \liminf_{r\to 0} \frac{\nu\bigl(B(x,r)\bigr)}{\mu\bigl(B(x,r)\bigr)},$$

where we set $\overline{D}_\mu \nu(x) = \underline{D}_\mu \nu(x) = +\infty$ if $\mu\bigl(B(x,r)\bigr) = 0$ for some $r > 0$.

5.8.6. Definition. *If $\overline{D}_\mu \nu(x) = \underline{D}_\mu \nu(x) < +\infty$, then the number*

$$D_\mu \nu := \overline{D}_\mu \nu(x) = \underline{D}_\mu \nu(x)$$

will be called the derivative of ν with respect to μ at the point x.

5.8.7. Lemma. *Let $0 < c < \infty$. Then*
(i) *If $A \subset \{x\colon\ \underline{D}_\mu \nu(x) \le c\}$, then $\nu^*(A) \le c\mu^*(A)$,*
(ii) *If $A \subset \{x\colon\ \overline{D}_\mu \nu(x) \ge c\}$, then $\nu^*(A) \ge c\mu^*(A)$.*

Proof. (i) By the properties of outer measure it suffices to prove our claim for bounded sets A. Let $A \subset \{x\colon\ \underline{D}_\mu\nu(x) \le c\}$, $\varepsilon > 0$, and let U be an open set containing A. Denote by $\mathcal{F}$ the class of all closed balls $B(a,r) \subset U$ with $r > 0$, $a \in A$ and $\nu\bigl(B(a,r)\bigr) \le (c+\varepsilon)\mu\bigl(B(a,r)\bigr)$. By the definition of $\underline{D}_\mu\nu$ we obtain that $\inf\{r\colon\ B(a,r) \in \mathcal{F}\} = 0$ for all $a \in A$. By Corollary 5.8.2, there exists an at most countable family of disjoint balls $B_j \in \mathcal{F}$ with $\nu\bigl(A \backslash \bigcup_{j=1}^\infty B_j\bigr) = 0$, which yields the estimates

$$\nu^*(A) \le \sum_{j=1}^\infty \nu(B_j) \le (c+\varepsilon)\sum_{j=1}^\infty \mu(B_j) \le (c+\varepsilon)\mu(U).$$

Since $U \supset A$ is arbitrary, we obtain the desired estimate. Assertion (ii) is proved similarly, one has only take for $\mathcal{F}$ the class of balls that satisfy the inequality $\nu\bigl(B(a,r)\bigr) \ge (c-\varepsilon)\mu\bigl(B(a,r)\bigr)$. □

5.8.8. Theorem. *Let μ and ν be two nonnegative Borel measures on $\mathbb{R}^n$ that are finite on all balls. Denote by ν_{ac} the absolutely continuous component of the measure ν with respect to μ (i.e., $\nu = \nu_{ac} + \nu_s$, where $\nu_{ac} \ll \mu$ and $\nu_s \perp \mu$). Then the function $D_\mu\nu$ is defined and finite μ-almost everywhere. In addition, this function is μ-measurable and serves as the Radon–Nikodym density of the measure ν_{ac} with respect to μ.*

Proof. It is clear that the theorem reduces to finite measures. We verify first that $D_\mu\nu$ exists and is finite μ-a.e. Let $S = \{x\colon\ \overline{D}_\mu\nu(x) = +\infty\}$. By Lemma 5.8.7 one has $\mu(S) = 0$. Let $0 < a < b$ and

$$S(a,b) = \bigl\{x\colon\ \underline{D}_\mu\nu(x) < a < b < \overline{D}_\mu\nu(x) < +\infty\bigr\}.$$

By the same lemma

$$b\mu^*\bigl(S(a,b)\bigr) \le \nu^*\bigl(S(a,b)\bigr) \le a\mu^*\bigl(S(a,b)\bigr),$$

whence $\mu^*\bigl(S(a,b)\bigr) = 0$ because $a < b$. Since the union of sets $S(a,b)$ over all positive rational a and b has μ-measure zero, the first claim is proven.

We observe that the functions $x \mapsto \mu\bigl(B(x,r)\bigr)$ and $x \mapsto \nu\bigl(B(x,r)\bigr)$ are Borel (this is seen, for example, from Exercise 5.8.100). Let

$$f_k(x) = \nu\bigl(B(x,1/k)\bigr)/\mu\bigl(B(x,1/k)\bigr)$$

if $\mu\bigl(B(x,1/k)\bigr) > 0$ and $f_k(x) = +\infty$ otherwise. It follows that f_k is finite μ-a.e. and μ-measurable. Hence the function $D_\mu\nu = \lim\limits_{k\to\infty} f_k$ is μ-measurable.

Let us prove the second assertion. Suppose first that $\nu \ll \mu$. It is clear from Lemma 5.8.7 that the set $Z = \{x\colon\ D_\mu\nu(x) = 0\}$ has μ-measure zero. Hence $\nu(Z) = 0$. Let A be a Borel set, let $t > 1$, and let

$$A_m := A \cap \bigl\{x\colon\ t^m \le D_\mu\nu(x) < t^{m+1}\bigr\}, \quad m \in \mathbb{Z}.$$

The sets A_m cover A up to a ν-measure zero set, since ν-a.e. we have the estimate $D_\mu\nu(x) > 0$. Hence on account of the lemma we obtain

$$\begin{aligned}\nu(A) &= \sum_{m=-\infty}^{+\infty} \nu(A_m) \le \sum_{m=-\infty}^{+\infty} t^{m+1}\mu(A_m) \\ &\le t \sum_{m=-\infty}^{+\infty} \int_{A_m} D_\mu\nu\, d\mu = t\int_A D_\mu\nu\, d\mu.\end{aligned}$$

This estimate is true for any $t > 1$. Hence

$$\nu(A) \le \int_A D_\mu\nu\, d\mu.$$

By the estimate $\nu(A_m) \ge t^m\mu(A_m)$ we obtain similarly that

$$\nu(A) \ge \int_A D_\mu\nu\, d\mu.$$

Thus, $D_\mu\nu$ is the Radon–Nikodym density of the measure ν with respect to μ. For completing the proof it remains to verify that $D_\mu\nu_s = 0$ μ-a.e. We take a Borel set B such that $\nu_s(B) = 0$ and $\mu(\mathbb{R}^n\backslash B) = 0$. Let $c > 0$ and $B_c = B \cap \{x\colon\ D_\mu\nu_s(x) \ge c\}$. Then $c\mu(B_c) \le \nu_s(B_c) = 0$, whence $\mu(B_c) = 0$. Therefore, $D_\mu\nu_s = 0$ μ-a.e. on B. $\square$

A multidimensional analog of Theorem 5.1.4 in terms of differentiation of set functions is found in Howard, Pfeffer [**444**].

5.8(iv). The approximate continuity

Let a function f be defined on a measurable set $E \subset \mathbb{R}^n$. We shall say that f is approximately continuous at a point $x \in E$ if there exists a measurable set $E_x \subset E$ such that x is a density point of E_x and $\lim\limits_{y\in E_x, y\to x} f(y) = f(x)$. This property can be reformulated (Exercise 5.8.91) as the following equality: $\operatorname{ap}\lim\limits_{y\to x} f(y) = f(x)$, where the approximate limit $\operatorname{ap}\lim\limits_{y\to x} f(y)$ is defined as a number p such that, for every $\varepsilon > 0$, the set $\{y \in E\colon\ |f(y) - p| < \varepsilon\}$ has x as a density point.

5.8.9. Theorem. *Every finite measurable function on a measurable set E is approximately continuous almost everywhere on E.*

Proof. By Lusin's theorem, for every $\varepsilon > 0$, there exists a continuous function g such that the measure of the set of all points in E where $f \ne g$ is less than ε. Deleting from the set $\{x \in E\colon\ f(x) = g(x)\}$ all points that are not its density points, we obtain the set A of the same measure. Since $f = g$ on A and every point in A is a density point of A, we see from the above-mentioned equivalent description of the approximate continuity that f is approximately continuous at every point in A. Since ε is arbitrary, the theorem is proven. $\square$

An alternative proof is obtained from Exercise 5.8.90. It turns out that this theorem can be inverted. First we establish the following interesting fact.

5.8.10. Lemma. *Let $\{E_\alpha\}$ be an arbitrary family of measurable sets in $\mathbb{R}^n$ such that every point of E_α is its density point. Then their union $E := \bigcup_\alpha E_\alpha$ is measurable.*

In particular, given an arbitrary family $\{E_\alpha\}$ of measurable sets in $\mathbb{R}^n$, let E_α^d denote the set of all density points of E_α. Then the sets $E' := \bigcup_\alpha E_\alpha^d$ and $E'' := \bigcup_\alpha (E_\alpha^d \cap E_\alpha)$ are measurable.

PROOF. We may assume that all sets E_α are contained in a cube, considering their intersections with a fixed open cube. There exist Borel sets $A \subset E$ and $B \supset E$ such that $\lambda_*(E) = \lambda(A)$ and $\lambda^*(E) = \lambda(B)$, where λ is Lebesgue measure. Suppose that E is non-measurable. Then $\lambda(B\backslash A) > 0$ and $\lambda^*(E\backslash A) > 0$. Since almost every point of the set $B\backslash A$ is its density point and $E\backslash A \subset B\backslash A$, it follows that among such points there exists $x \in E\backslash A$ because otherwise we would have $\lambda(E\backslash A) = 0$. Therefore, $x \in E_\alpha$ for some α. Then x is a density point of the set $E_\alpha \cap (B\backslash A) = E_\alpha\backslash A$. This means that $\lambda(E_\alpha\backslash A) > 0$, and we arrive at a contradiction with the equality $\lambda(A) = \lambda_*(E)$. The claim for E' and E'' follows, since every point of E_α^d is its density point. □

Note that every point of the set E' is its density point. This fact enables one to define the so called density topology, in which open sets are the sets of density points of measurable sets (see Exercise 5.8.92).

5.8.11. Theorem. *Suppose that $E \subset \mathbb{R}^n$ is a measurable set and a function $f\colon E \to \mathbb{R}$ is approximately continuous almost everywhere on E. Then f is measurable.*

PROOF. Let $r \in \mathbb{R}$ and $A = \{x \in E\colon\ f(x) < r\}$. Denote by C the set of all points in E at which f is approximately continuous. Let $x \in A \cap C$. By definition, there exists a measurable set $C_x \subset E$ such that the point x belongs to the set C_x and is its density point and the restriction of the function f to C_x is continuous at the point x. Since $f(x) < r$, one can find an open ball U_x centered at x such that $f(y) < r$ for all $y \in U_x \cap C_x$. Let $E_x = U_x \cap C_x$. It is clear that x belongs to the set E_x^d of all density points of the set E_x, hence the set $B := \bigcup_{x\in A\cap C}(E_x^d \cap E_x)$ contains $A\cap C$. Thus, $A\cap C \subset B \subset A$, which gives the equality $A = B\cup(A\backslash C)$. By the above lemma B is measurable. Since $A\backslash C$ has measure zero, the set A is measurable, which means the measurability of the function f. □

5.8(v). Derivates and the approximate differentiability

It is known that there exist nowhere differentiable functions. The following surprising result (its first part is the Denjoy–Young–Saks theorem) shows, in particular, that the set of points of differentiability of an arbitrary function is measurable.

5.8.12. Theorem. *Let f be an arbitrary function on $[a,b]$. Then, for almost every $x \in [a,b]$, one of the following four cases occurs*:

(a) $f'(x)$ *exists and is finite,*

(b) $-\infty < D^+f(x) = D_-f(x) < +\infty$, $D^-f(x) = +\infty$, $D_+f(x) = -\infty$,

(c) $-\infty < D^-f(x) = D_+f(x) < +\infty$, $D^+f(x) = +\infty$, $D_-f(x) = -\infty$,

(d) $D^+f(x) = D^-f(x) = +\infty$, $D_+f(x) = D_-f(x) = -\infty$.

In addition, the upper derivative $\overline{D}f$ and the lower derivative $\underline{D}f$ are measurable as mappings with values in $[-\infty, +\infty]$.

In particular, the set D of all points at which f has a finite derivative is measurable and the function f' on D is measurable. Moreover, D is a Borel set and $f'|_D$ is a Borel function.

PROOF. We verify that one has the equality $D_-f(x) = D^+f(x) < +\infty$ a.e. on the set $E := \{x\colon D_-f(x) > -\infty\}$. Other combinations of derivates are reduced to this one by passing to the functions $-f(x)$, $f(-x)$, $-f(-x)$. The set E is the union of the sets

$$E_{r,n} := \Bigl\{x \in E\colon x > r,\ \frac{f(x) - f(y)}{x - y} > -n,\ \forall\, y \in (r,x)\Bigr\}$$

over all rational $r \in (a,b)$ and all integer $n \geq 0$. Let us verify our claim for a.e. x from each fixed $E_{r,n}$. Passing to the function $f(x-r) + nx$ we reduce the verification to the case of the set $E_{0,0}$. We observe that f on $E_{0,0}$ is monotone, hence can be extended to a monotone function on an interval containing $E_{0,0}$. Since a monotone function is almost everywhere differentiable, the set of points $x \in E_{0,0}$ at which there is no finite limit of the ratio $[f(y) - f(x)]/(y-x)$ as $y \to x$, $y \in E_{0,0}$, has measure zero. Deleting from $E_{0,0}$ this set and the set of all points of E that are not density points of the closure of $E_{0,0}$, we obtain the set E_0 that coincides with $E_{0,0}$ up to a set of measure zero. Let $x \in E_0$. If $y \to x$ and $y \in E_0$, then $[f(y) - f(x)]/(y-x)$ has a finite limit $L(x)$ by our choice of E_0. Then $D_-f(x) \leq L(x) \leq D^+f(x)$. Let $y_n \to x$ and $y_n \notin E_0$. Since x is a density point of E_0, there exist $z_n \in E_0$ with $z_n > y_n$ such that $|z_n - x|/|y_n - x| < (n+1)/n$. Then $f(z_n) \geq f(y_n)$ and hence for all $y_n > x$ we have

$$[f(y_n) - f(x)]/(y_n - x) \leq (n+1)n^{-1}[f(z_n) - f(x)]/(z_n - x),$$

whence one has $D^+f(x) \leq L(x)$. Similarly, we verify that $D_-f(x) \geq L(x)$. Thus, $D^+f(x) = D_-f(x)$ is finite.

Let $A = \{x\colon \overline{D}f(x) > 0\}$. A point x belongs to A if and only if one can find $m \in \mathbb{N}$ and a sequence $h_n \to 0$ such that $h_n \neq 0$ and one has $f(x+h_n) - f(x) \geq m^{-1}|h_n|$. For every pair $k, m \in \mathbb{N}$, we denote by $J_{k,m}$ the union of all intervals $[x, x+h]$ (over all x for which they exist) such that $|h| \leq k^{-1}$ and $f(x+h) - f(x) \geq m^{-1}|h|$. Then $A = \bigcup_{m=1}^\infty \bigcap_{k=1}^\infty J_{k,m}$. This follows by the above characterization of A and the following property:

$$\text{if}\ \ x \in [z, z+h],\ \ 0 < h \leq k^{-1},\ \ f(z+h) - f(z) \geq m^{-1}h,$$

then $|x-z| \le k^{-1}$, $|z+h-x| \le k^{-1}$, and one has at least one of the inequalities

$$f(z+h) - f(x) \ge m^{-1}(z+h-x), \ f(x) - f(z) \ge m^{-1}(x-z).$$

The set $J_{k,m}$ is measurable by Exercise 1.12.87(i). Hence A is measurable, which yields the measurability of $\overline{D}f$, since one can pass to the function $f(x) - cx$. Considering $-f$ we obtain the measurability of $\underline{D}f$. Hence the set D and the function $f'|_D$ are measurable. In fact, they are Borel. Indeed, we observe that the set C of all points of continuity of f (it contains D) is a countable intersection of open sets, since it consists of the points where the oscillation of f vanishes (Exercise 2.12.72), and the set of all points where the oscillation of f is less than $\varepsilon > 0$ is open. Hence, for fixed $m, k \in \mathbb{N}$, the set $C_{m,k}$ of all $x \in C$, such that for some $y \in [a,b]$ with $0 < |x-y| < k^{-1}$ one has $\bigl(f(y) - f(x)\bigr)/(y-x) > m^{-1}$, is Borel (this set is open in C by the continuity of f on C). Then the set $B = \{x \in C\colon\ \overline{D}f(x) > 0\}$ is Borel as well, since $B = \bigcup_{m=1}^\infty \bigcap_{k=1}^\infty C_{m,k}$. Applying this argument to the functions $f(x) - rx$ and $rx - f(x)$, we obtain that the sets $\{x \in C\colon\ \overline{D}f(x) > r\}$ and $\{x \in C\colon\ \underline{D}f(x) < r\}$ are Borel. This yields that D is a Borel set and $f'|_D$ is a Borel function. □

In the Denjoy–Young–Saks theorem one can take any set A in place of an interval and consider the corresponding derivatives along A. On measurability of derivates, see Saks [**840**, §IV.4].

5.8.13. Lemma. *Let f be a function on $[a,b]$ and let E be the set of all points at which f has a nonzero derivative. Then, for every set Z of measure zero, the set $f^{-1}(Z) \cap E$ has measure zero. In other words, $\lambda \circ f^{-1}|_E \ll \lambda|_E$, where λ is Lebesgue measure.*

PROOF. The set E is measurable by Theorem 5.8.12, and the function f is continuous on E, hence is measurable on E. Now it suffices to prove our claim for the sets of the form $E \cap \{f' > n^{-1}\}$ and $E \cap \{f' < -n^{-1}\}$. Hence it suffices to consider the set $A = \{x\colon\ f'(x) > 1\}$ in place of E. Next we reduce everything to the sets

$$A_r = A \cap \Bigl\{x\colon\ \frac{f(x) - f(y)}{x - y} > 1, \forall\, y \in (r, x)\Bigr\}, \quad r \in \mathbb{Q}.$$

Now we may confine ourselves to the set A_0. Deleting from A_0 all points that are not density points, we obtain the set B of the same measure. In addition, the function f on B is increasing. Now we take $\varepsilon > 0$ and find an open set $U \supset Z$ of measure less than ε. Every point $x \in B \cap f^{-1}(Z)$ possesses a sequence of shrinking neighborhoods $U_{x,n} = (x - r_n, x + r_n)$, $r_n = r_n(x)$, such that $x - r_n, x + r_n \in B$, $f(x + r_n) - f(x - r_n) > 2r_n$ and $\bigl(f(x - r_n), f(x + r_n)\bigr) \subset U$. By Vitali's Theorem 5.5.1, the collection of all such neighborhoods contains an at most countable subfamily of disjoint intervals $(x_n - r_n, x_n + r_n)$ that covers $B \cap f^{-1}(Z)$ up to a measure zero set. It remains to observe that the intervals $\bigl(f(x - r_n), f(x + r_n)\bigr)$ are disjoint, since f is increasing on B, and the sum of their lengths is less than ε because they are contained in U. Since $2r_n < f(x+r_n)-f(x-r_n)$, the sum of lengths of the intervals $(x_n - r_n, x_n + r_n)$ is less than ε as well. $\square$

Let $E \subset \mathbb{R}^n$. A mapping $f\colon E \to \mathbb{R}^k$ is called approximately differentiable at a point $x_0 \in E$ if there exists a linear mapping $L\colon \mathbb{R}^n \to \mathbb{R}^k$ such that

$$\operatorname{ap}\lim_{x\to x_0} \frac{|f(x) - f(x_0) - L(x - x_0)|}{|x - x_0|} = 0,$$

where $|v|$ denotes the norm of a vector v. The mapping L (which is obviously uniquely defined) is called the approximate derivative of f at the point x_0 and denoted by $\mathrm{ap}f'(x_0)$. By analogy, one defines the approximate partial derivatives $\mathrm{ap}\partial_{x_i} f(x_0)$. To this end, the function f is considered on the straight lines $\{x_0 + te_i, t \in \mathbb{R}^1\}$. The existence of approximate partial derivatives is considerably weaker than the existence of usual partial derivatives.

Note the following important Whitney theorem [**1013**] (its proof can also be found in Federer [**282**, §3.1]).

5.8.14. Theorem. *Let $f\colon E \to \mathbb{R}^1$ be a measurable function on a measurable set $E \subset \mathbb{R}^n$ equipped with Lebesgue measure λ. Then the following conditions are equivalent*: (i) *f is approximately differentiable a.e. on E,*

(ii) *f has the approximate partial derivatives a.e. on E,*

(iii) *for every $\varepsilon > 0$, there exist a closed set $E_\varepsilon \subset E$ and a function $f_\varepsilon \in C^1(\mathbb{R}^n)$ such that $\lambda(E\backslash E_\varepsilon) < \varepsilon$ and $f|_{E_\varepsilon} = f_\varepsilon|_{E_\varepsilon}$.*

5.8(vi). The class BMO

Let us consider an interesting functional space related to the maximal function. We shall say that a locally integrable function f belongs to the space of functions of bounded mean oscillation $\mathrm{BMO}(\mathbb{R}^n)$ if, for some $A > 0$, for all balls B one has

$$\frac{1}{\lambda(B)} \int_B |f(x) - f_B|\, dx \le A,$$

where

$$f_B := \lambda(B)^{-1} \int_B f(y)\, dy$$

and λ is Lebesgue measure. The smallest possible A is denoted by $\|f\|_{\mathrm{BMO}}$. After factorization by constant functions $\mathrm{BMO}(\mathbb{R}^n)$ with the norm $\|\cdot\|_{\mathrm{BMO}}$ becomes a Banach space. Examples of unbounded functions in $\mathrm{BMO}(\mathbb{R}^n)$ are given in Exercise 5.8.98. For any function $f \in \mathrm{BMO}(\mathbb{R}^n)$, the function $|f(x)|(1+|x|)^{-n-1}$ is integrable. Note the following important John–Nirenberg estimate.

5.8.15. Theorem. *Let $f \in \mathrm{BMO}(\mathbb{R}^n)$. Then, for all $p > 0$, the function $|f|^p$ is locally integrable and, for some constant $c_{n,p}$ independent of f, one has*

$$\frac{1}{\lambda(B)} \int_B |f(x) - f_B|^p\, dx \le c_{n,p} \|f\|^p_{\mathrm{BMO}}$$

for each ball B. In addition, there exist numbers $k_1(n)$ and $k_2(n)$ such that for all $t > 0$ and all balls B, one has

$$\lambda\bigl(x \in B\colon\ |f(x) - f_B| > t\bigr) \le k_1(n)\lambda(B)\exp\bigl(-k_2(n)t/\|f\|_{\mathrm{BMO}}\bigr).$$

The last inequality yields that for all $c < k_2(n)$ one has

$$\int_B \exp\bigl(c|f(x) - f_B|\bigr)\,dx < \infty.$$

Proofs of the stated facts can be found in Stein [**906**].

5.8(vii). Weighted inequalities

Let A_p, $1 \le p < \infty$, be the class of all locally integrable nonnegative functions ω on $\mathbb{R}^n$ such that for some $C > 0$, one has for every ball B

$$\frac{1}{\lambda(B)}\int_B \omega(x)\,dx \le C\Bigl(\frac{1}{\lambda(B)}\int_B \omega(x)^{-p'/p}\,dx\Bigr)^{-p/p'},$$

where $p' = p/(p-1)$ and λ is Lebesgue measure. The membership of ω in A_p is equivalent to the existence of $C' > 0$ such that, for all nonnegative bounded measurable functions f and all balls B, one has

$$(f_B)^p \le \frac{1}{C'}\Bigl(\int_B \omega(x)\,dx\Bigr)^{-1}\int_B f(x)^p\omega(x)\,dx. \tag{5.8.7}$$

The classes A_p have the following relation to the space $\mathrm{BMO}(\mathbb{R}^n)$.

5.8.16. Theorem. (i) *Let $\omega \in A_p$. Then $\ln\omega \in \mathrm{BMO}(\mathbb{R}^n)$.*
(ii) *Let $f \in \mathrm{BMO}(\mathbb{R}^n)$ and $p > 1$. Then $f = c\ln\omega$ for some $c \in \mathbb{R}$ and some $\omega \in A_p$.*

The classes A_p admit yet another description.

5.8.17. Theorem. *Let μ be a nonnegative Borel measure on $\mathbb{R}^n$ of the form $\mu = \omega(x)\,dx$ and let $1 \le p < \infty$. Then $\omega \in A_p$ precisely when there is a number $A > 0$ such that for all $f \in L^p(\mu)$ one has*

$$\mu\bigl(x\colon\ Mf(x) > t\bigr) \le \frac{A}{t^p}\int_{\mathbb{R}^n} |f|^p\,d\mu, \quad \forall\, t > 0.$$

5.8.18. Theorem. *Let $1 < p < \infty$ and $\omega \in A_p$. Then there exists a constant A such that for all $f \in L^p(\omega\,dx)$ one has*

$$\int_{\mathbb{R}^n} |Mf(x)|^p\,\omega(x)\,dx \le A\int_{\mathbb{R}^n} |f(x)|^p\,\omega(x)\,dx.$$

Denote by A_∞ the union of all classes A_p, $p < \infty$. The class A_∞ admits the following characterization.

5.8.19. Theorem. *Let ω be a nonnegative locally integrable function on $\mathbb{R}^n$. The following conditions are equivalent*:
(i) $\omega \in A_\infty$;

(ii) *for every* $\alpha \in (0,1)$*, there exists* $\beta \in (0,1)$ *such that if* B *is a ball and* $E \subset B$ *is a measurable set with* $\lambda(E) \geq \alpha\lambda(B)$*, then*

$$\int_E \omega(x)\,dx \geq \beta \int_B \omega(x)\,dx;$$

(iii) *there exist* $r \in (1,\infty)$ *and* $c > 0$ *such that*

$$\Bigl(\frac{1}{\lambda(B)}\int_B \omega(x)^r\,dx\Bigr)^{1/r} \leq \frac{1}{\lambda(B)}\int_B \omega(x)\,dx$$

for every ball B;

(iv) *there exists* $A > 0$ *such that for every ball* B *one has*

$$\frac{1}{\lambda(B)}\int_B \omega(x)\,dx\ \exp\Bigl(\frac{1}{\lambda(B)}\int_B \ln\frac{1}{\omega(x)}\,dx\Bigr) \leq A.$$

Proofs and additional information related to this subsection can be found in García-Cuerva, Rubio de Francia [**340**], Stein [**906**].

5.8.20. Remark. Let μ and ν be two bounded nonnegative Borel measures on $\mathbb{R}^n$. Let us consider the following relation: $\mu \preceq \nu$ if, for every $\varepsilon > 0$, there exists $\delta > 0$ such that, for every ball B and every Borel $E \subset B$ with $\nu(E)/\nu(B) \leq \delta$, one has $\mu(E)/\mu(B) \leq \varepsilon$. The relation $\preceq$ is an equivalence relation. If ν is Lebesgue measure and $\mu = \omega(x)\,dx$, then condition $\mu \preceq \nu$ is equivalent to $\omega \in A_\infty$. Details can be found in Coifman, Feffermann [**185**].

5.8(viii). Measures with the doubling property

Many results about the maximal function extend to the case when in place of Lebesgue measure one considers a measure μ with the so-called doubling property: for some $c > 0$ one has

$$\mu\bigl(B(x,2r)\bigr) \leq c\mu\bigl(B(x,r)\bigr), \quad \forall\, x, \forall\, r > 0,$$

where $B(x,r)$ is the closed ball of radius r centered at x. Measures with such a property can be considered on general metric spaces, too. It is known that if G is a polynomial of degree d on $\mathbb{R}^n$, then the measure $\mu = |G|^\alpha\,dx$ has the doubling property for all $\alpha > -1/d$. On the other hand, the measure $\mu = \exp|x|\,dx$ does not have this property. There exist singular measures with the doubling property, for example, the measure μ obtained as the weak limit of the sequence of measures $\prod_{k=1}^n [1 + a\cos(3^k 2\pi x)]\,dx$, where $a \in (0,1)$, see, e.g., Stein [**906**, p. 40]. Finally, there exist absolutely continuous measures $\mu = f\,dx$ with the doubling property not equivalent to Lebesgue measure (i.e., f vanishes on a set of positive measure). Exercise 5.8.99 proposes to verify that if $\omega \in A_p$, then the measure $\omega(x)\,dx$ has the doubling property.

Additional information about measures with the doubling property and related references can be found in Heinonen [**418**], Stein [**906**]. When does a positive measure with the doubling property exist on a given space? We shall mention several interesting results in this direction obtained by Volpert and Konyagin [**996**], [**997**].

We shall say that a nonnegative Borel measure μ on a metric space X with the metric ϱ satisfies condition D_γ, where γ is a nonnegative number, if μ is finite on all balls and there exists $C > 0$ such that

$$\mu\bigl(B(x, kR)\bigr) \le Ck^\gamma \mu\bigl(B(x, R)\bigr), \quad \forall\, x \in X, \forall\, R > 0, \forall\, k \in \mathbb{N},$$

where $B(x, r)$ is the closed ball of radius r centered at x. If there exists a positive measure μ on X satisfying condition D_γ, then we say that X belongs to the class Ψ_γ. The existence of a positive measure on X with the doubling property is equivalent to the membership of X in some class Ψ_γ. Let us set $\beta(X) := \inf\{\gamma\colon\ X \in \Psi_\gamma\}$ and $\beta(X) = +\infty$ if there are no such γ.

Now we introduce a metric characteristic of X which is responsible for the existence of measures with the doubling property. We shall say that X belongs to the class Φ_γ, where $\gamma \ge 0$, if there exists a number N such that, for each $x \in X$ and all $R > 0$, $k \in \mathbb{N}$, the ball $B(x, kR)$ contains at most Nk^γ points with the mutual distances at least R. Let $\alpha(X) := \inf\{\gamma\colon\ X \in \Phi_\gamma\}$ and $\alpha(X) = +\infty$ if there are no such γ. It is clear that all these objects depend on the metric ϱ.

5.8.21. Theorem. (i) *If $X \in \Phi_\gamma$, then $X \in \Psi_{\gamma'}$ for all $\gamma' > \gamma$.*

(ii) $\alpha(X) = \beta(X)$.

(iii) *Every nonempty compact set $X \subset \mathbb{R}^n$ with the induced metric belongs to the class Ψ_n and hence is the support of a probability measure with the doubling property.*

In [**997**], an example is constructed showing that assertion (i) may fail for $\gamma' = \gamma$. The following interesting property of covers is deduced in [**997**] from the existence of a measure with the doubling property.

5.8.22. Theorem. *For every $n \in \mathbb{N}$, there exists a number $C(n)$ with the following property: let $B(x_1, R_1),\ldots,B(x_N, R_N)$ be a finite family of closed balls in $\mathbb{R}^n$, let $N_i = \sum_{j=1}^N I_{B(x_j,R_j)}(x_i)$ be the multiplicity of the covering of the point x_i by these balls, and let $N_i' = \sum_{j=1}^N I_{B(x_j,2R_j)}(x_i)$ be the multiplicity of its covering by the balls with the double radii. Then $N_{i_0}' \le C(n)N_{i_0}$ for some $i_0 \in \{1, \ldots, N\}$.*

It is shown in Kaufman, Wu [**498**] that if an atomless Radon probability measure μ on a metric compact K has the doubling property, then there is a Radon probability measure ν on K that has this property as well and is singular with respect to μ. Regarding measures with the doubling property, see also Luukkainen, Saksman [**641**].

5.8(ix). The Sobolev derivative

S.L. Sobolev discovered a new type of derivative, which turned out to be very useful in modern analysis and applications. Sobolev's approach was developed by L. Schwartz, who introduced the concept of generalized derivative not only for functions, but also for more general objects (distributions or generalized functions). Here we briefly explain the principal idea of the theory of generalized derivatives conformably to measures.

5.8.23. Definition. *Let $\Omega \subset \mathbb{R}^n$ be open and let $f \in L^1(\Omega)$. We shall say that a function $g_i \in L^1(\Omega)$ is the generalized partial derivative of f with respect to the variable x_i if, for every smooth function ψ with compact support in Ω, one has*

$$\int_\Omega \partial_{x_i}\psi(x) f(x)\, dx = -\int_\Omega \psi(x) g_i(x)\, dx. \tag{5.8.8}$$

In this case g_i is denoted by $\partial_{x_i} f$.

5.8.24. Definition. *Let μ be a bounded Borel measure on an open set $\Omega \subset \mathbb{R}^n$. We shall say that a bounded measure ν on Ω is the generalized derivative of the measure μ along a vector h if, for every smooth function ψ with compact support in Ω, one has*

$$\int_\Omega \partial_h\psi(x)\, \mu(dx) = -\int_\Omega \psi(x)\, \nu(dx). \tag{5.8.9}$$

Analogous definitions are introduced for locally finite measures. It is clear that if the measure μ is given by a smooth density ϱ with respect to Lebesgue measure, then the measure ν is given by the density $\partial_h\varrho$, i.e., the partial derivative of ϱ along h, provided the latter is integrable. Exercise 5.8.78 proposes to prove that if the measure μ has generalized derivatives along n linearly independent vectors, then it is absolutely continuous with respect to Lebesgue measure. According to Exercise 5.8.79, in the case where $n = 1$ and $\Omega = (a, b)$, the measure μ has the generalized derivative ν along 1 precisely when μ has a density ϱ with respect to Lebesgue measure on (a, b) such that ϱ is equivalent to a function of bounded variation. Thus, for general functions of bounded variation (unlike absolutely continuous functions), their natural derivatives are measures, not functions. Moreover, absolutely continuous functions are specified in the class of functions of bounded variation exactly by that their derivatives are absolutely continuous measures.

The Sobolev space $W^{p,1}(\Omega)$, $p \in [1, \infty)$, is defined as the set of all functions $f \in L^p(\Omega)$ such that their generalized partial derivatives $\partial_{x_i} f$ belong to $L^p(\Omega)$. The mapping

$$\nabla f = (\partial_{x_1} f, \ldots, \partial_{x_n} f)$$

is called the generalized gradient of f. The space $W^{p,1}(\Omega)$ is Banach with respect to the norm

$$\|f\|_{W^{p,1}} := \|f\|_{L^p(\Omega)} + \|\,|\nabla f|\,\|_{L^p(\Omega)}.$$

An equivalent norm is $\|f\|_{L^p(\Omega)} + \sum_{i=1}^n \|\partial_{x_i} f\|_{L^p(\Omega)}$.

In applications, the following Sobolev inequality is useful: there exists a number c_n that depends only on $n > 1$ such that for all $f \in W^{1,1}(\mathbb{R}^n)$ one has

$$\Bigl(\int_{\mathbb{R}^n} |f(x)|^{n/(n-1)}\, dx\Bigr)^{(n-1)/n} \le c_n \int_{\mathbb{R}^n} |\nabla f(x)|\, dx. \tag{5.8.10}$$

One more useful inequality, connecting the integral of a function with the integral of its derivative, is called the Poincaré inequality. We give it in the following formulation.

5.8.25. Theorem. *For every n and every $p \in [1, n)$, there exists a constant $C(n,p)$ such that, for every function $f \in W^{p,1}(\mathbb{R}^n)$ and every ball U, one has*

$$\biggl(\int_U |f - f_U|^{\frac{np}{n-p}}\,dx\biggr)^{\frac{n-p}{np}} \le C(n,p)\biggl(\int_U |\nabla f|^p\,dx\biggr)^{1/p},$$

where

$$f_U := \lambda_n(U)^{-1}\int_U f\,dx$$

and λ_n is Lebesgue measure.

The class $W^{1,1}(\mathbb{R}^1)$ coincides with the space of all integrable absolutely continuous functions whose derivatives are integrable on the whole line.

There is a natural multidimensional analog of functions of bounded variation. Denote by $BV(\Omega)$ the class of all functions f in $L^1(\Omega)$ such that the generalized partial derivatives of the measure $f\,dx$ (in the sense of Definition 5.8.24) are bounded measures on Ω. These measures are denoted by Df_i. Then we obtain a bounded vector-valued measure

$$Df(B) := \bigl(Df_1(B), \ldots, D_nf(B)\bigr).$$

Set

$$\|f\|_{BV} = \|f\|_{L^1(\Omega)} + \|Df\|,$$

where $\|Df\|$ is the variation of the measure Df defined as $\sup_{|e|\le 1}\|(e, Df)\|$, where (e, Df) is the scalar measure obtained by the inner product with the vector e. An equivalent norm: $\|f\|_{BV} = \|f\|_{L^1(\Omega)} + \sum_{i=1}^n \|Df_i\|$.

The following result is due to Krugova [**550**].

5.8.26. Theorem. *Let μ be a convex measure on $\mathbb{R}^n$ with a density ϱ. Then $\varrho \in BV(\mathbb{R}^n)$. If $\varrho(x) > 0$ a.e., then $\varrho \in W^{1,1}(\mathbb{R}^n)$.*

Functions in $BV(\Omega)$ are called functions of bounded variation on Ω. We shall say that a bounded measurable set E has finite perimeter if its indicator function I_E belongs to $BV(\mathbb{R}^n)$. Let

$$\mathrm{P}(E) := \|DI_E\|.$$

A set $E \subset \mathbb{R}^n$ is called a Caccioppolli set if its intersection with each ball has finite perimeter. Sobolev's inequality extends to functions of bounded variation:

$$\biggl(\int_{\mathbb{R}^n} |f(x)|^{n/(n-1)}\,dx\biggr)^{(n-1)/n} \le c_n\|Df\|, \quad \forall\, f \in BV(\mathbb{R}^n). \tag{5.8.11}$$

Inequality (5.8.11) yields the following isoperimetric inequality: for every bounded Caccioppolli set $E \subset \mathbb{R}^n$ one has

$$\lambda_n(E)^{(n-1)/n} \le c_n\mathrm{P}(E).$$

Isoperimetric inequalities are considered in many works, see Burago, Zalgaller [**143**], Chavel [**173**], and Osserman [**732**] for further references.

Let us mention a useful result on the structure of Sobolev functions that resembles Lusin's classical theorem on the structure of measurable functions. A proof and references can be found in Evans, Gariepy [**273**, Ch. 6].

5.8.27. Theorem. *Let $f \in BV(\mathbb{R}^n)$. Then, for every $\varepsilon > 0$, there exists a continuously differentiable function f_ε such that*

$$\lambda_n\bigl(x \in \mathbb{R}^n\colon\ f_\varepsilon(x) \neq f(x)\bigr) \le \varepsilon.$$

If $f \in W^{p,1}(\mathbb{R}^n)$, where $p \in [1, +\infty)$, then f_ε can be chosen such that, in addition, $\|f - f_\varepsilon\|_{W^{p,1}(\mathbb{R}^n)} \le \varepsilon$.

Let $W^{p,1}_{\mathrm{loc}}(\mathbb{R}^n)$ denote the class of all functions f on $\mathbb{R}^n$ such that one has $\zeta f \in W^{p,1}(\mathbb{R}^n)$ for all $\zeta \in C_0^\infty(\mathbb{R}^n)$. Let $W^{p,1}(\mathbb{R}^n, \mathbb{R}^k)$ be the Sobolev class of mappings $f = (f_1, \ldots, f_k)\colon \mathbb{R}^n \to \mathbb{R}^k$ with $f_i \in W^{p,1}(\mathbb{R}^n)$. This class is equipped with the following norm: the sum of the Sobolev norms of the components f_i. By analogy with the case $k = 1$ one defines the class $W^{p,1}_{\mathrm{loc}}(\mathbb{R}^n, \mathbb{R}^k)$.

Regarding Sobolev spaces, see Adams [**2**], Besov, Il'in, Nikol'skiĭ [**86**], Evans, Gariepy [**273**], Goldshtein, Reshetnyak [**371**], Maz'ja [**663**], Stein [**905**], Ziemer [**1051**], and the references therein. Regarding the space BV and Caccioppolli sets, see Ambrosio, Fusco, Pallara [**22**], Federer [**282**], Giusti [**358**], Giaquinta, Modica, Souček [**352**]. Several interesting facts are found in the exercises in this chapter.

5.8(x). The area and coarea formulas and change of variables

Given $f \in W^{p,1}_{\mathrm{loc}}(\mathbb{R}^n, \mathbb{R}^k)$, we denote by $|Jf|$ the absolute value of the k-dimensional Jacobian of f, i.e., the k-dimensional volume of the parallelepiped generated by the vectors ∇f_i, $i = 1, \ldots, k$. In particular, for $n = k$ the number $|Jf(x)|$ equals $|\det(\partial_{x_i} f_j)_{i,j\le n}|$. Let $\mathrm{Card} M$ denote the cardinality of the set M. As above, H^α denotes the Hausdorff measure.

5.8.28. Lemma. *Let $n \le k$ and let a mapping $f\colon \mathbb{R}^n \to \mathbb{R}^k$ be continuous. Denote by E_f the set of all points x at which f is differentiable and the linear mapping $Df(x)$ is injective. Then, for every $\alpha > 1$, the set E_f can be written as the union of a sequence of Borel sets B_j with the following properties: the restrictions $f|_{B_j}$ are injective and there exist invertible linear mappings $G_j\colon \mathbb{R}^n \to \mathbb{R}^k$ such that:*

(i) *the mappings $(f|_{B_j}) \circ G_j^{-1}$ and $G_j \circ (f|_{B_j})^{-1}$ on their natural domains of definition are Lipschitzian with constant α,*

(ii) $\alpha^{-1}|G_j(u)| \le |Df(x)(u)| \le \alpha |G_j(u)|$ *for all $x \in B_j$, $u \in \mathbb{R}^n$,*

(iii) $\alpha^{-n}|\det G_j| \le |Jf(x)| \le \alpha^n |\det G_j|$ *for all $x \in B_j$.*

PROOF. Let us choose $\varepsilon > 0$ such that $\alpha^{-1} + \varepsilon < 1 < \alpha - \varepsilon$. Let us take an everywhere dense countable set $\mathcal{G}$ in the space $\mathrm{GL}(\mathbb{R}^n)$ of all invertible linear operators on $\mathbb{R}^n$ equipped with the operator norm. Let $B(x,r)$ denote the open ball of radius r centered at x. For every $G \in \mathcal{G}$ and $j \in \mathbb{N}$, we consider the set $B(G,j)$ of all points $x \in E_f$ such that

$$(\alpha^{-1}+\varepsilon)|G(u)| \le |Df(x)(u)| \le (\alpha-\varepsilon)|G(u)|, \quad \forall\, u \in \mathbb{R}^n, \tag{5.8.12}$$

$$|f(y)-f(x)-Df(x)(y-x)| \le \varepsilon|G(y-x)|, \quad \forall\, y \in B(x,1/j). \tag{5.8.13}$$

For all $x, y \in B(G,j)$ with $|y-x|<1/j$, we have

$$|f(y)-f(x)| \le |Df(x)(y-x)| + \varepsilon|G(y-x)| \le \alpha|G(y-x)|,$$

$$|f(y)-f(x)| \ge |Df(x)(y-x)| - \varepsilon|G(y-x)| \ge \alpha^{-1}|G(y-x)|.$$

Let us cover $B(G,j)$ by the sets $B\bigl(x,1/(2j)\bigr)\cap B(G,j)$ and choose a countable subcover in this cover. The union of such countable families over all $G \in \mathcal{G}$ and $j \in \mathbb{N}$ gives the required countable family. Indeed, by the obtained estimates we have (i) and (ii). Let us show that every point $x \in E_f$ belongs to some $B(G,j)$. We observe that $Df(x)$ can be written in the form $Df(x) = UT$, where $T \in \mathrm{GL}(\mathbb{R}^n)$ and U is a linear isometry from $\mathbb{R}^n$ to $\mathbb{R}^k$. One has $|Df(x)(u)| = |G(u)|$, $u \in \mathbb{R}^n$. There is $G \in \mathcal{G}$ with $\|TG^{-1}\| < \alpha-\varepsilon$, $\|GT^{-1}\| < (\alpha^{-1}+\varepsilon)^{-1}$. This gives (5.8.12). By the differentiability of f at the point x, there exists $j \in \mathbb{N}$ such that for all $y \in B(x,1/j)$ one has $|f(y)-f(x)-Df(x)(y-x)| \le \varepsilon|y-x|/\|G^{-1}\|$. This gives (5.8.13). Finally, we observe that estimate (iii) in the formulation of the lemma follows by the easily verified fact that the inequality $|T_1(u)| \le |T_2(u)|$ for two linear operators T_1 and T_2 on $\mathbb{R}^n$ implies the inequality $|\det T_1| \le |\det T_2|$. □

The following result contains the so-called area and coarea formulas.

5.8.29. Theorem. *Let $f\colon \mathbb{R}^n \to \mathbb{R}^k$ be a Lipschitzian mapping and let $A \subset \mathbb{R}^n$ and $B \subset \mathbb{R}^k$ be two measurable sets. Then*
(i) *if $n \le k$, then*

$$\int_{A\cap f^{-1}(B)} |Jf(x)|\,dx = \int_B \mathrm{Card}\,\bigl(A\cap f^{-1}(y)\bigr)\,H^n(dy), \tag{5.8.14}$$

(ii) *if $n > k$, then*

$$\int_A |Jf(x)|\,dx = \int_{\mathbb{R}^k} H^{n-k}\bigl(A\cap f^{-1}(y)\bigr)\,dy. \tag{5.8.15}$$

PROOF. (i) Replacing A by $A\cap f^{-1}(B)$, it suffices to consider the case $B=\mathbb{R}^k$. Suppose first that $A \subset E_f$, where E_f is defined in the lemma, fix $\alpha>1$ and take the corresponding partition of E_f into Borel parts B_j. Let $A_j = A\cap B_j$. For every j, we take the operator G_j indicated in the lemma and obtain

$$\alpha^{-n}H^n\bigl(G_j(A_j)\bigr) = \alpha^{-n}|\det G_j|\lambda_n(A_j) \le \int_{A_j} |Jf(x)|\,dx$$
$$\le \alpha^n|\det G_j|\lambda_n(A_j) = \alpha^n H^n\bigl(G_j(A_j)\bigr).$$

In addition, by property (i) of the mappings G_j and Lemma 3.10.12 one has $\alpha^{-n}H^n\bigl(G_j(A_j)\bigr) \le H^n\bigl(f(A_j)\bigr) \le \alpha^n H^n\bigl(G_j(A_j)\bigr)$, whence we have

$$\alpha^{-2n}H^n\bigl(f(A_j)\bigr) \le \int_{A_j} |Jf(x)|\,dx \le \alpha^{2n}H^n\bigl(f(A_j)\bigr).$$

Summing in j and using the equality $\operatorname{Card}\bigl(A_j\cap f^{-1}(y)\bigr) = I_{A_j}(y)$, we obtain

$$\alpha^{-2n}\int_{\mathbb{R}^k} \operatorname{Card}\bigl(A\cap f^{-1}(y)\bigr)\,H^n(dy) \le \int_A |Jf(x)|\,dx \le \alpha^{2n}\int_{\mathbb{R}^k} \operatorname{Card}\bigl(A\cap f^{-1}(y)\bigr)\,H^n(dy).$$

Letting $\alpha \to 1$ we obtain our assertion in the case $A \subset E_f$. Now it suffices to consider the case when A is contained in the set of points at which f has a derivative, but this derivative is not injective (we recall that f is differentiable almost everywhere). Let us fix $\varepsilon > 0$ and write $f = p\circ g$, where p is the projection operator from $\mathbb{R}^k\times\mathbb{R}^n$ onto $\mathbb{R}^k$ and $g\colon \mathbb{R}^n \to \mathbb{R}^k\times\mathbb{R}^n$ is given by $g(x) = \bigl(f(x), \varepsilon x\bigr)$. For all $x \in A$, we have $Dg(x)(u) = \bigl(Df(x)(u), \varepsilon u\bigr)$. It is clear that g and $Dg(x)$ are injective and $\|Dg(x)\| \le C+\varepsilon$. Since $Df(x)$ has a nontrivial kernel, one has $|Jg(x)| \le \varepsilon(L+\varepsilon)^{k-1}$. By the injectivity of Dg we obtain from the first step of the proof

$$H^n\bigl(f(A)\bigr) \le H^n\bigl(g(A)\bigr) = \int_A |Jg(x)|\,dx \le \varepsilon(C+\varepsilon)^{k-1}\lambda_n(A).$$

Letting $\varepsilon \to 0$ we conclude that $H^n\bigl(f(A)\bigr) = 0$, which completes the proof.

Assertion (ii) is proved in a similar manner, see Federer [**282**, §3.2]). □

Letting $n = k$ in the case of a one-to-one mapping f we arrive at the change of variables formula under assumptions much weaker than those in §3.7. By using Theorem 5.8.14 the following more general change of variables formula was proved in Hajłasz [**403**] (earlier this formula had been proved in Kudryavtsev [**551**] under the additional assumption of a.e. differentiability in the usual sense). One needs Lusin's property (N) considered in §3.6.

5.8.30. Theorem. *Let $\Omega \subset \mathbb{R}^n$ be an open set and let $f\colon \mathbb{R}^n \to \mathbb{R}^n$ be a measurable mapping that has the approximate partial derivatives a.e. in Ω. Denote by $|Jf|$ the absolute value of the determinant of the matrix formed by the approximate partial derivatives of the function f. Suppose, in addition, that f has Lusin's property* (N). *Then, for every measurable set $E \subset \Omega$ and every measurable function $u\colon \mathbb{R}^n \to \mathbb{R}^1$, the functions*

$$u\bigl(f(x)\bigr)|Jf(x)|I_E(x), \quad u(y)\operatorname{Card}\bigl(f^{-1}(y)\cap E\bigr)$$

are measurable, where we set $u\bigl(f(x)\bigr)|Jf(x)| = 0$ if the function $u\bigl(f(x)\bigr)$ is not defined. If one of these functions is integrable, then so is the other and

$$\int_E u\bigl(f(x)\bigr)|Jf(x)|\,dx = \int_{\mathbb{R}^n} u(y)\operatorname{Card}\bigl(f^{-1}(y)\cap E\bigr)\,dy.$$

We observe that if a function f has the approximate partial derivatives a.e., then it has a version with Lusin's property (N) (this is clear from Theorem 5.8.14). However, the reader is warned that even when f is continuous, this version may not be continuous. There exist examples of continuous mappings in the class $W^{p,1}_{\text{loc}}$ with $p \le n$ without property (N); see Reshetnyak [**790**], Väisälä [**970**]; for $p < n$ one can even find such homeomorphisms (Ponomarev [**765**]). For such continuous mappings the above formula fails because it implies property (N).

There are many problems in measure theory that are related to Sobolev functions. We mention a result from Aleksandrova, Bogachev, Pilipenko [**9**] on convergence of images of Lebesgue measure under differentiable mappings.

5.8.31. Theorem. (i) *Let $F_j\colon \mathbb{R}^n \to \mathbb{R}^n$ be continuous mappings that converge uniformly on compact sets to a continuous mapping $F\colon \mathbb{R}^n \to \mathbb{R}^n$ and let F_j and F have Lusin's property* (N). *In addition, suppose that almost everywhere there exist the partial derivatives $\partial_{x_i}F_j$ and $\partial_{x_i}F$ such that the mappings $\partial_{x_i}F_j$ converge in measure to $\partial_{x_i}F$ on some set E of finite Lebesgue measure. Finally, suppose that $JF \ne 0$ on E, where JF is the determinant of the matrix formed by the partial derivatives, and that the sequence $\{JF_j\}$ is uniformly integrable on every compact set. Then, the measures $\lambda|_E \circ F_j^{-1}$ converge to the measure $\lambda|_E \circ F^{-1}$ in the variation norm. In addition, if μ is an absolutely continuous probability measure on $\mathbb{R}^n$, then the measures $\mu|_E \circ F_j^{-1}$ converge to the measure $\mu|_E \circ F^{-1}$ in the variation norm.*

(ii) *Let $F_j, F \in W^{p,1}_{\text{loc}}(\mathbb{R}^n, \mathbb{R}^n)$, where $p \ge n$, and let the mappings F_j converge to F in the Sobolev norm $\|\cdot\|_{p,1}$ on every ball. Suppose that E is a measurable set of finite Lebesgue measure and that $JF \ne 0$ on E. Then the measures $\lambda|_E \circ F_j^{-1}$ converge to the measure $\lambda|_E \circ F^{-1}$ in the variation norm.*

J. Moser [**700**] proved the existence of an infinitely differentiable diffeomorphism of a cube in $\mathbb{R}^n$ with any given infinitely differentiable strictly positive Jacobian. Thus, Lebesgue measure on the unit cube can be transformed by a smooth diffeomorphism to any given probability measure with a strictly positive smooth density. See also Rivière, Ye [**812**], where analogous problems are discussed for mappings from Sobolev classes.

Before formulating the following theorem from McCann [**665**] (generalizing a close result from Brenier [**125**]), we recall that every convex function ψ on $\mathbb{R}^n$ is locally Lipschitzian and a.e. differentiable.

5.8.32. Theorem. *Let μ and ν be two Borel probability measures on $\mathbb{R}^n$ such that the measure μ is absolutely continuous. Then, there exists a convex function ψ on $\mathbb{R}^n$ such that $\nu = \mu \circ (\nabla\psi)^{-1}$. In addition, the mapping $\nabla\psi$ is unique μ-a.e. in the class of gradients of convex functions.*

In fact, the requirement on μ is even weaker: it must vanish on all Borel sets of the Hausdorff dimension $n-1$. Results related to this theorem are obtained in Caffarelli [**157**], where one can find applications to integral inequalities.

5.8(xi). Surface measures

A set S in $\mathbb{R}^{n+1}$ will be called an elementary surface if it can be transformed by an orthogonal linear operator to the graph of a Lipschitzian function f restricted to a bounded measurable set $D \subset \mathbb{R}^n$. A set $S \subset \mathbb{R}^{n+1}$ will be called a surface if it is the countable union of elementary surfaces S_j. We shall confine ourselves to considering only elementary surfaces, i.e., graphs of Lipschitzian functions, since the construction of surface measure on more general surfaces reduces to this case.

The surface measure σ_S on the surface $S \subset \mathbb{R}^{n+1}$ is defined as the restriction of the Hausdorff measure H^n to the Borel σ-algebra of S.

It follows by construction that the measure σ_S is σ-finite because it is finite on elementary surfaces. The following result expresses surface measure via Lebesgue measure on $\mathbb{R}^n$.

5.8.33. Proposition. *Suppose that f is a Lipschitzian function on $\mathbb{R}^n$. Let $D \subset \mathbb{R}^n$ be a bounded measurable set and let $S \subset \mathbb{R}^{n+1}$ be the graph of the function f on D. Then*

$$\sigma_S(S) := H^n(S) = \int_D \sqrt{1 + |\nabla f(x)|^2}\, dx. \tag{5.8.16}$$

PROOF. It suffices to consider Borel sets D. Let $F(x) = \bigl(x, f(x)\bigr)$, $x \in D$. By formula (5.8.14) we have

$$\int_D |JF(x)|\, dx = \int_S \operatorname{Card}\bigl(D \cap F^{-1}(y)\bigr)\, H^n(dy) = H^n(S).$$

It remains to observe that $|JF(x)|^2 = 1 + |\nabla f(x)|^2$ by the definition of the absolute value of the Jacobian of the mapping $F\colon \mathbb{R}^n \to \mathbb{R}^{n+1}$. $\square$

If the function f is affine, i.e., $f(x) = (x, h) + c$, where h is a constant vector and c is a constant number, then the n-dimensional measure of the set $F(D)$ (the graph of f on D) is $\sqrt{1 + |h|^2}\lambda_n(D)$. Formula (5.8.16) for smooth functions can be deduced from this. Note also that this formula can be used as a definition of the surface measure for surfaces that are locally representable as graphs of functions (e.g., for elementary surfaces).

Similarly one proves that if a set $S \subset \mathbb{R}^{n+1}$ is given parametrically in the form $S = F(D)$, where $F = (F_1, \ldots, F_{n+1})$ is a Lipschitzian mapping from $\mathbb{R}^n$ to $\mathbb{R}^{n+1}$ and D is a bounded measurable set in $\mathbb{R}^n$, then

$$H^n(S) = \int_D \Bigl|\sum_{k=1}^{n+1} D_k(x)^2\Bigr|^{1/2} dx,$$

where $D_k(x)$ is the absolute value of the Jacobian of the mapping

$$(F_1, \ldots, F_{k-1}, F_{k+1}, \ldots, F_{n+1})\colon\ \mathbb{R}^n \to \mathbb{R}^n.$$

Such a set S may not be an elementary surface, but one can show that up to a set of H^n-measure zero S is an at most countable union of elementary surfaces. When dealing with surfaces it is useful to remember that for any Lipschitzian function f, one can find Borel sets B_j and continuously differentiable functions f_j such that $f|_{B_j} = f_j|_{B_j}$ and the complement to the union of the sets B_j has measure zero.

Hausdorff measures can also be employed for defining length of curves. If a curve $C \subset \mathbb{R}^n$ is defined as the image of the interval $[a, b]$ under a Lipschitzian mapping $f\colon [a, b] \to \mathbb{R}^n$, then

$$H^1(C) = \int_a^b |f'(t)|\, dt,$$

which follows by the area formula.

The following result (its proof is delegated to Exercise 5.8.104) enables one to compute volume integrals by means of surface integrals.

5.8.34. Proposition. *Let f be a Lipschitzian function on $\mathbb{R}^n$ such that $|\nabla f(x)| \geq c > 0$ a.e. If a function g is integrable on $\mathbb{R}^n$, then*

$$\int_{\{f>t\}} g(x)\, dx = \int_t^\infty \int_{\{f=s\}} \frac{g(y)}{|\nabla f(y)|}\, H^{n-1}(dy)\, ds$$

for all $t \in \mathbb{R}$.

The following classical result is also related to surface measures.

5.8.35. Theorem. *Let A be a convex compact set of positive Lebesgue measure λ_n in $\mathbb{R}^n$. If the surface measure of its boundary equals the surface measure of the boundary of a ball B, then*

$$\lambda_n(B) \geq \lambda_n(A).$$

The equality occurs only if A is a ball.

PROOF. We may assume that B is a unit ball centered at the origin. Let $r > 0$. By the Brunn–Minkowski inequality one has

$$\lambda_n(A + rB) \geq \Bigl(\lambda_n(A)^{1/n} + r\lambda_n(B)^{1/n}\Bigr)^n.$$

Taking the expansion of the right-hand side in powers of r, we obtain

$$\lim_{r\to 0} \frac{\lambda_n(A + rB) - \lambda_n(A)}{r} \geq n\lambda_n(A)^{(n-1)/n}\lambda_n(B)^{1/n}.$$

By Exercise 5.8.107, the left-hand side of this inequality equals the surface measure $H^{n-1}(\partial A)$ of the boundary of A. If $A = B$, then this inequality becomes an equality, since $H^{n-1}(\partial B) = n\lambda_n(B)$, which is verified directly. Now the assumption that $H^{n-1}(\partial A) = H^{n-1}(\partial B)$ yields the desired inequality.

Let us consider the case of equality. By Exercise 5.8.107 we obtain that in this case $v_{n-1,1}(A, B) = \lambda_n(A)^{n-1}\lambda_n(B)$ (mixed volumes are defined in §3.10(vii)), which yields the equality in the Minkowski inequality. Therefore, A and B are homothetic, i.e., A is a ball. $\square$

5.8(xii). The Calderón–Zygmund decomposition

5.8.36. Theorem. *Let f be a nonnegative integrable function on $\mathbb{R}^n$. Then, for every number $\alpha > 0$, one can find a sequence of disjoint open cubes Q_k with edges parallel to the coordinate axes such that:*

(i) *for every k one has*

$$\alpha < \frac{1}{\lambda_n(Q_k)} \int_{Q_k} f(x)\,dx \le 2^n \alpha; \tag{5.8.17}$$

(ii) *$f(x) \le \alpha$ for almost all $x \in \mathbb{R}^n \backslash \bigcup_{k=1}^{\infty} Q_k$.*

Proof. We take the cube $Q = [-2^m, 2^m]^n$ with $m \in \mathbb{N}$ such that the integral of f does not exceed $\alpha 2^{(m+1)n}$. The cube Q generates the partition of $\mathbb{R}^n$ into equal closed cubes with the edge length 2^{m+1} and disjoint interiors. Let us take an arbitrary cube Q' in this partition and apply the following operation. We partition Q' into 2^n equal cubes with twice smaller edges. For every cube Q'' in the obtained refinement, two cases are possible:

$$\int_{Q''} f\,dx > \alpha\lambda_n(Q'') \qquad \text{or} \qquad \int_{Q''} f\,dx \le \alpha\lambda_n(Q'').$$

In the first case we declare the interior of Q'' to be one of the required cubes Q_k. We note that (5.8.17) follows from the estimates

$$\alpha < \frac{1}{\lambda_n(Q'')} \int_{Q''} f\,dx \le \frac{2^n}{\lambda_n(Q')} \int_{Q'} f\,dx \le 2^n \alpha. \tag{5.8.18}$$

In the second case, we partition Q'' into 2^n equal cubes with edges half as long. The described operation is applied to all cubes in the first collection and to all cubes of the arising partitions such that whenever the first of the above two possibilities occurs, the interior of the corresponding cube is included in our collection $\{Q_k\}$ and this cube is excluded from further consideration. Estimate (5.8.18) is ensured by the fact that Q' has not been excluded at the previous step, hence

$$\int_{Q'} f\,dx \le \alpha\lambda_n(Q').$$

At the first step this is true due to our choice of m and the equality $\lambda_n(Q') = 2^{(m+1)n}$. Let D be the complement of the obtained sequence of open cubes Q_k. It is clear that D is a closed set. Let us show that $f(x) \le \alpha$ for almost all $x \in D$. Indeed, for almost each $x \in D$, there exists a sequence of closed cubes K_j that contain x, have edges approaching zero and correspond to the second of the above-mentioned cases, i.e., the integral of f over K_j does not exceed $\alpha\lambda_n(K_j)$. By Corollary 5.6.3, for almost all $x \in D$, we have

$$f(x) = \lim_{j\to\infty} \frac{1}{\lambda_n(K_j)} \int_{K_j} f\,dy \le \alpha,$$

which completes the proof. □

Let us observe that $\lambda_n\bigl(\bigcup_{k=1}^\infty Q_k\bigr) \le \alpha^{-1}\|f\|_1$. The Calderón–Zygmund decomposition is connected with the maximal function, see Stein [**905**, Ch. 1].

Exercises

5.8.37.° Prove that if a function f has a finite derivative at every point of the line, then f' has a dense set of continuity points (see, however, Exercise 5.8.119). Hence there exists a closed interval on which the function $|f'|$ is bounded. In particular, f is Lipschitzian on this interval.

HINT: apply Baire's theorem discussed in Exercise 2.12.73 to the functions $k\bigl(f(x+1/k)-f(x)\bigr)$.

5.8.38. Prove that if the derivative of a function f is everywhere finite and equals almost everywhere some continuous function, then it equals that function everywhere and f is continuously differentiable.

HINT: apply Theorem 5.7.7.

5.8.39.° (i) Construct a continuous strictly increasing function f on the real line such that $f'(x)=0$ a.e.

(ii) Show that for such a function one can take

$$f(t) = P\Bigl(\omega\colon\ \sum_{n=1}^\infty \xi_n(\omega)2^{-n} < t\Bigr),$$

where ξ_n are independent random variables (see Chapter 10) on a probability space (Ω, P) such that $P(\xi_n = 1) = p$, $P(\xi_n = 0) = 1-p$, where $p\in(0,1)$ and $p\ne 1/2$.

5.8.40.° (Riesz [**807**]) Prove that a nonnegative function f on $[a,b]$ is Lebesgue integrable precisely when there exists a nondecreasing function F on $[a,b]$ such that $F'(x)=f(x)$ a.e. In addition, the integral of f equals the infimum of the differences $F(b)-F(a)$ over all such functions F.

5.8.41.° Let f be an absolutely continuous function on $[0,1]$. For every $h>0$ let $f_h(x) := h^{-1}\bigl(f(x+h)-f(x)\bigr)$, where $f(x+h)=f(1)$ if $x+h>1$. Show that $\lim\limits_{h\to 0}\|f_h - f'\|_{L^1[0,1]} = 0$.

HINT: the family of functions $f_h, h\in(0,1)$, is uniformly integrable. Indeed, for any fixed $\varepsilon>0$ and $M>0$ and any set E of measure ε one has

$$\begin{aligned}\Bigl|\int_E f_h(x)\,dx\Bigr| &= h^{-1}\Bigl|\int_0^1\int_0^1 f'(t)I_{[x,x+h]}(t)I_E(x)\,dt\,dx\Bigr|\\ &\le Mh^{-1}\int_0^1\int_0^1 I_{[x,x+h]}(t)I_E(x)\,dx\,dt + \int_{\{|f'|>M\}}|f'(t)|\,dt\\ &\le M\varepsilon + \int_{\{|f'|>M\}}|f'(t)|\,dt,\end{aligned}$$

since

$$h^{-1}\int_0^1 I_{[x,x+h]}(t)I_E(x)\,dx \le 1$$

for each fixed t. Taking first a sufficiently large M and then a sufficiently small ε, we make the right-hand side as small as we wish simultaneously for all h.

5.8.42.° Prove Proposition 5.2.8.

HINT: it is clear that f is nondecreasing and hence $f'(x)$ exists a.e. We have $[f_n(x+h)-f_n(x)]/h \geq 0$ if $h>0$, hence $\sum_{n=1}^k f_n'(x) \leq f'(x)$ a.e., whence we obtain convergence of the series $\sum_{n=1}^\infty f_n'(x)$ a.e. to some function g. We may assume that $f_n(0)=0$, passing to $f_n(x)-f_n(a)$. For every k, there exists n_k such that one has $\sum_{n>n_k} f_n(b) < 2^{-k}$, whence by monotonicity we have $\sum_{n>n_k} f_n(x) < 2^{-k}$ for all x. Hence the series of nondecreasing functions $\varphi_k(x) := f(x) - \sum_{n=1}^{n_k} f_n(x)$ converges. According to what we have proved, the series of $\varphi_k'(x)$ converges a.e., hence $\varphi_k'(x) \to 0$ a.e., which yields $f'(x) = g(x)$ a.e. because if $g(x) < f'(x)$, then $\lim\limits_{k\to\infty} \varphi_k'(x) > 0$.

5.8.43. Let $\varrho \in \mathcal{L}(\mathbb{R}^1)$ be absolutely continuous on bounded intervals. Suppose that a function f is either absolutely continuous on bounded intervals or everywhere differentiable. Let $f\varrho'$ and $f'\varrho$ be in $\mathcal{L}^1(\mathbb{R}^1)$. Prove the equality

$$\int_{-\infty}^{+\infty} f'(t)\varrho(t)\,dt = -\int_{-\infty}^{+\infty} f(t)\varrho'(t)\,dt.$$

HINT: suppose first that f is bounded and locally absolutely continuous. Since $\varrho \in \mathcal{L}^1(\mathbb{R}^1)$ by assumption, one can find numbers $a_n \to -\infty$ and $b_n \to +\infty$ such that $|\varrho(a_n)| + |\varrho(b_n)| \to 0$. Then $f(b_n)\varrho(b_n) - f(a_n)\varrho(a_n) \to 0$. By the integration by parts formula for the intervals $[a_n, b_n]$ we arrive at the desired equality. If f is not bounded, we take smooth functions θ_n such that $\theta_n(t) = t$ if $t \in [-n, n]$, $\theta_n(t) = -n-1$ if $t \leq -n-1$, $\theta_n(t) = n+1$ if $t \geq n+1$, and $\sup_{n,t} |\theta_n'(t)| < \infty$. The required equality holds for $\theta_n \circ f$ in place of f, which yields our claim by the Lebesgue dominated convergence theorem. The case where f is everywhere differentiable is less obvious because f may not be absolutely continuous. As above, it suffices to consider the case where f is bounded. We observe that if ϱ does not vanish on a closed interval $[a,b]$, then $f' \in \mathcal{L}^1[a,b]$, hence $f \in AC[a,b]$. It follows by the integration by parts formula that the equality

$$\int_a^b f'(t)\varrho(t)\,dt = -\int_a^b f(t)\varrho'(t)\,dt$$

holds provided that $\varrho(b) = \varrho(a) = 0$ and $\varrho(t) \neq 0$ for all $t \in (a,b)$. It remains to note that the set $U := \{t\colon\ \varrho(t) \neq 0\}$ is a finite or countable union of open intervals (possibly unbounded), and the integrals of $f'\varrho$ and $f\varrho'$ over $\mathbb{R}^1 \backslash U$ vanish, since $\varrho' = 0$ a.e. on $\mathbb{R}^1 \backslash U$.

5.8.44. Let μ be a probability measure on a space X and let f be a nonnegative μ-measurable function. Suppose that φ is a locally absolutely continuous increasing function on $[0, +\infty)$. Prove that

$$\int_X \varphi\bigl(f(x)\bigr)\,\mu(dx) = \varphi(0) + \int_0^\infty \varphi'(t)\mu\bigl(x\colon\ f(x) > t\bigr)\,dt,$$

where both integrals are finite or infinite simultaneously.

HINT: we may assume that $\varphi(0) = 0$ passing to $\varphi(t) - \varphi(0)$. Suppose first that φ is strictly increasing and $f \leq C$. Then the integral of $\varphi \circ f$ with respect to μ equals the Riemann integral of $\mu\bigl(x\colon\ \varphi\circ f(x) > t\bigr)$ over $[0, \varphi(C)]$. Since one has the equality $\mu\bigl(x\colon\ \varphi\circ f(x) > t\bigr) = \mu\bigl(x\colon\ f(x) > \varphi^{-1}(t)\bigr)$, it remains to apply the change of variable formula with $t = \varphi(s)$. The case where φ is not strictly increasing follows by considering the functions $\varphi(t) + t\varepsilon$ with $\varepsilon > 0$ and letting $\varepsilon \to 0$. Let us consider the general case. If $\varphi \circ f \in L^1(\mu)$, we apply the previous case to the functions $\min(f, n)$ in place of f and let $n \to \infty$. Finally, if the right-hand side of the desired equality is finite, then, by the already-proven assertion, we obtain the uniform boundedness of the integrals of $\varphi(\min(f, n))$, which yields the integrability of $\varphi \circ f$ with respect to μ.

5.8.45. Construct a continuous function F on the interval $[0, 1]$ such that, at every point in the interval, it has a finite or infinite derivative f that is almost everywhere finite and integrable, but the function

$$\Phi(x) = F(0) + \int_0^x f(t)\,dt$$

has no finite or infinite derivative at infinitely many points (in particular, Φ does not coincide with F).

HINT: see Lusin [**633**, p. 392].

5.8.46. Show that given $E \subset [0, 1]$ with $\lambda(E) = 0$, there exists a continuous nondecreasing function ψ on $[0, 1]$ with $\psi'(x) = +\infty$ for all $x \in E$.

HINT: there exist open sets $G_n \supset E$ with $\lambda(G_n) < 2^{-n}$. Consider the function $\varphi_n(x) = \lambda(G_n \cap [0, x])$. Then $\varphi_n < 2^{-n}$ and one can set $\psi(x) := \sum_{n=1}^\infty \varphi_n(x)$. If $x_0 \in E$, then, for any fixed n, we have $[x_0, x_0 + h] \subset G_n$ for all sufficiently small $h > 0$, whence $\varphi_n(x_0 + h) = \varphi_n(x_0) + h$. Hence, for every fixed k and all sufficiently small h, we obtain

$$\frac{\psi(x_0 + h) - \psi(x_0)}{h} \geq \sum_{n=1}^k \frac{\varphi_n(x_0 + h) - \varphi_n(x_0)}{h} \geq k.$$

Similarly, one considers $h < 0$. Thus, $\psi'(x_0) = +\infty$.

5.8.47. Construct an example of a continuous function f on $(0, 1)$ that at no point has the usual derivative, but is approximately differentiable almost everywhere.

HINT: see Tolstoff's example in Lusin [**633**, p. 448].

5.8.48. (Lusin [**633**, §46]) (i) Let ψ be a continuous function on $[0, 1]$ such that one has $\psi'(x) = 0$ a.e. Show that there exists a set $E \subset [0, 1]$ of measure 1 such that $\psi(E)$ has measure zero.

(ii) Let ψ be a non-constant continuous function on $[0, 1]$ such that $\psi'(x) = 0$ a.e. Show that there exists a set M of measure zero such that $\psi(M)$ has positive measure.

HINT: use Proposition 5.5.4 to show that $\lambda\bigl(\psi(\{\psi' = 0\})\bigr) = 0$.

5.8.49.° Show that every absolutely continuous function has Lusin's property (N), i.e., takes all measure zero sets to measure zero sets.

5.8.50. Suppose that a function f on $[a, b]$ is differentiable at all points of some set E. Show that $f(E)$ has measure zero precisely when $f'(x) = 0$ a.e. on E.

HINT: use Proposition 5.5.4 and Lemma 5.8.13.

5.8.51. (S. Banach [**50**], M.A. Zareckiĭ) Prove that a function f on $[0, 1]$ is absolutely continuous precisely when it is continuous, is of bounded variation and possesses Lusin's property (N).

HINT: if f is of bounded variation, then f' exists a.e. and is integrable. Let D be the set of all points of differentiability of f. For all $a, b \in (0, 1)$, by the continuity and property (N), Proposition 5.5.4 yields the estimate

$$|f(b)-f(a)| \le \lambda\bigl(f([a,b])\bigr) = \lambda\bigl(f([a,b]\cap D)\bigr) \le \int_a^b |f'(x)|\,dx,$$

ensuring the absolute continuity of f.

5.8.52. Let f be an absolutely continuous function on $[0,1]$ such that for a.e. x one has $f'(x)>0$. Prove that f is strictly increasing and the inverse function is absolutely continuous on $[f(0),f(1)]$.

HINT: the fact that f is strictly increasing follows by the Newton–Leibniz formula. The inverse function is continuous and increasing, so its absolute continuity follows by property (N) verified with the help of Exercise 5.8.50.

5.8.53. Let f be a continuous function on $[0,1]$ and let D be the set of all points of differentiability of f on $(0,1)$. Prove that f is absolutely continuous precisely when f' is integrable on D and $f([0,1]\backslash D)$ has measure zero. In particular, if f is differentiable everywhere in $(0,1)$ and $f'\in L^1[0,1]$, then f is absolutely continuous.

HINT: if f is absolutely continuous, then f' exists a.e. and f has property (N). Conversely, if the above condition is fulfilled, then we can apply the same reasoning as in Exercise 5.8.51.

5.8.54. (M.A. Zareckiĭ) Let f be a continuous strictly increasing function on an interval $[a,b]$. (i) Prove that f is absolutely continuous precisely when f takes the set $\{x\colon\ f'(x)=+\infty\}$ to a measure zero set.

(ii) Let g be the inverse function for f. Prove that g is absolutely continuous precisely when the set $\{x\colon\ f'(x)=0\}$ has measure zero.

HINT: verify that f has property (N) on the set E of all points at which neither finite nor infinite derivative exists; to this end, modify Proposition 5.5.4 for different derivate numbers.

5.8.55. Let f be a continuous function with property (N) on $[a,b]$. Prove that for almost every y the set $f^{-1}(y)$ is at most countable.

HINT: observe that for any compact set $K\subset[a,b]$, there is a measurable set $E\subset K$ such that $f(K)=f(E)$ and the function f is injective on E. To this end, let $g(y)=\min\{x\in f^{-1}(y)\}$, $y\in f(K)$. It is easily verified that g is Borel measurable (see Theorem 6.9.7 in Chapter 6); set $E=g\bigl(f(K)\bigr)$. Let β denote the supremum of numbers α for which there exists a set $E\subset[a,b]$ such that $\lambda(E)\ge\alpha$, $\lambda\bigl(f([a,b]\backslash f(E)\bigr)=0$ and the sets $f^{-1}(y)\cap E$ are at most countable. It is clear that there is a set E_0 with the above properties and $\lambda(E_0)=\beta$. If $\beta=b-a$, then the assertion is proven. Suppose $b-a-\beta>0$. It remains to observe that $f([a,b]\backslash E_0)$ has measure zero. Otherwise by property (N) there is a compact set K in $[a,b]\backslash E_0$ with $\lambda\bigl(f(K)\bigr)>0$. Now the fact established at the first step leads to a contradiction.

5.8.56. Let f be a continuous function on $[a,b]$ with property (N). Let P the set of all points where f has a finite nonnegative derivative and let N be the set of all points where f has a finite nonpositive derivative. Prove that

$$-\lambda\bigl(f(N)\bigr) \le f(b)-f(a) \le \lambda\bigl(f(P)\bigr).$$

Deduce the existence of points of differentiability of f.

HINT: we may assume that $f(a)\le f(b)$ (otherwise consider $-f$). By the previous exercise, for almost every y the compact set $E_y:=f^{-1}(y)$ is at most countable. The set of all such points in $f([a,b])$ is denoted by Y. For every $y \in Y$, there is an isolated point x_y of the set E_y such that $\overline{D}f(x_y) \geq 0$. If E_y consists of a single point, then this point can be taken for x_y. Indeed, the inequality $\overline{D}f(x_y) < 0$ would yield the estimates $f(t) > f(x_y)$ if $t < x_y$ and $f(t) > f(x_y)$ if $t > x_y$ due to the absence of other points x with $f(x) = y$, hence $f(a) > f(b)$. If the compact set E_y is not a singleton, then it contains a pair of isolated points x_1 and x_2 between which there are no other points of E_y (this is easily seen from the fact that any infinite compact set without isolated points is uncountable, but E_y is finite or countable). At least one of these points is a desired one. Let $X = \{x_y\colon\ y \in Y\}$. Denote by X_0 the set of all points in X where f has a finite derivative. We observe that $\lambda(X\backslash X_0) = 0$. Indeed, as x_y is an isolated point in E_y, it is either a strict local extremum (minimum or maximum; the whole set of such points is readily seen to be at most countable) or in some neighborhood of x_y we have $f(t) < f(x_y)$ if $t < x_y$ and $f(t) > f(x_y)$ if $t > x_y$. Theorem 5.8.12 yields that the nondifferentiability points with such a property and $\overline{D}f(x_y) \geq 0$ form a measure zero set. Hence by property (N) we find $\lambda(Y) = \lambda\bigl(f(X)\bigr) = \lambda\bigl(f(X_0)\bigr) \leq \lambda\bigl(f(P)\bigr)$. Finally, $f(b) - f(a) \leq \lambda(Y)$. The second inequality is established in a similar way.

5.8.57. (i) Let f be a continuous function on $[a,b]$ with property (N) and let g be an integrable function on $[a,b]$ such that $f'(x) \leq g(x)$ at almost every point x where $f'(x)$ exists. Show that f is absolutely continuous.

(ii) Show that a continuous function f on $[a,b]$ is absolutely continuous precisely when it has property (N) and the function $f'(x)$ is integrable over the set P of all points at which it exists and is finite and nonnegative. In particular, if f is continuous, has property (N), is a.e. differentiable and f' is integrable, then f is absolutely continuous.

(iii) Show that every continuous function f on $[a,b]$ with property (N) is differentiable on a set of positive measure (but not necessarily a.e.).

HINT: (i) we show that f is of bounded variation; to this end, we observe that by the previous exercise for any $[\alpha,\beta] \subset [a,b]$ one has

$$f(\alpha) - f(\beta) \leq \lambda\bigl(f(P)\bigr) \leq \int_P f'(x)\,dx \leq \int_\alpha^\beta |g(x)|\,dx.$$

This enables us to estimate the total variation of f by $\max f - \min f + 2\|g\|_{L^1}$ because given a finite partition of $[a,b]$ by points x_k, in the finite sum of quantities $|f(x_{k+1}) - f(x_k)|$ the summands with $f(x_{k+1}) - f(x_k) \geq 0$ are estimated by the integrals of $|g|$ over $[x_k, x_{k+1}]$, and the sum of the remaining terms is estimated by the sum of the terms of the first kind and $\max f - \min f$. Assertion (ii) follows from (i) if we set $g(x) = f'(x)$ on P and $g(x) = 0$ outside P. Note that the last claim in (ii) also follows from Proposition 5.5.4 by the same estimate as in Exercise 5.8.51. (iii) If the set of all points of differentiability of f has measure zero, then the set P in (ii) has measure zero as well and hence f is absolutely continuous, which is a contradiction. An example where f is not a.e. differentiable is given in Ruziewicz [**836**].

5.8.58. (Menchoff [**680**]) Let ψ be a continuous function on $[0,1]$ that is not a constant and let $\psi'(x) = 0$ a.e. Then, for every absolutely continuous function φ on $[0,1]$, the function $\psi + \varphi$ has no property (N). In particular, the sum of any absolutely continuous function with the Cantor function has no property (N).

HINT: apply the previous exercise; see also [**680**, p. 645].

5.8.59.° Let f be an absolutely continuous monotone function on an interval $[a,b]$ and let φ be an absolutely continuous function on an interval $[c,d]$ containing $f([a,b])$. Show that $\varphi(f)$ is absolutely continuous on $[a,b]$.

HINT: let f be increasing; given $\varepsilon > 0$ take $\delta > 0$ by the definition of the absolute continuity of φ, and take $\tau > 0$ such that $\sum |f(b_i) - f(a_i)| < \delta$ for every collection of pairwise disjoint intervals (a_i, b_i) with $\sum |b_i - a_i| < \tau$; by the monotonicity of f, if $f(a_i) \neq f(b_i)$ and $f(a_j) \neq f(b_j)$, then the intervals $\bigl(f(a_i), f(b_i)\bigr)$ and $\bigl(f(a_j), f(b_j)\bigr)$ are disjoint.

5.8.60.° Find two absolutely continuous functions $f, g\colon [0,1] \to [0,1]$ such that their composition is not absolutely continuous.

5.8.61.° (Fichtenholz [**292**]) (i) Let a function F on $[a,b]$ be such that the composition $F \circ f$ is absolutely continuous for every absolutely continuous function f with values in $[a,b]$. Prove that F is Lipschitzian.

(ii) Let functions $f\colon [a,b] \to [c,d]$ and $F\colon [c,d] \to \mathbb{R}^1$ be absolutely continuous. Suppose that f satisfies the following Fichtenholz condition: there is a natural number k such that for every y, the set $f^{-1}(y)$ consists of at most k intervals (possibly degenerating to points). Prove that the function $F \circ f$ is absolutely continuous.

(iii) Suppose that a function $f\colon [a,b] \to [c,d]$ is continuous, but does not satisfy the Fichtenholz condition indicated in (ii). Show that there exists an absolutely continuous function F on $[c,d]$ such that the function $F \circ f$ is not absolutely continuous.

5.8.62. (G.M. Fichtenholz [**292**]) Let f be an absolutely continuous function on $[a,b]$ and let φ be an absolutely continuous function on an interval $[c,d]$ containing $f([a,b])$. Show that $\varphi \circ f$ is absolutely continuous on $[a,b]$ precisely when it is of bounded variation.

HINT: the function $\varphi \circ f$ has property (N) and Exercise 5.8.51 applies.

5.8.63. (i) (Lebesgue [**587**]) Show that there exist two functions with property (N) such that their sum does not have this property.

(ii) (Mazurkiewicz [**664**]) There exists a continuous function f with property (N) such that $f(x) + cx$ has no property (N) whenever $c \neq 0$.

(iii) Construct two continuous functions f and g with property (N) on $[0,1]$ such that their product fg has no property (N).

HINT: (i) Let C be the Cantor set of measure zero. It is easily verified that there exists a continuous mapping $\psi = (\psi_1, \psi_2)$ of the set C onto C^2. Let $f(x) = \psi_1(x)$, $g(x) = \psi_2(x)$ if $x \in C$, then extend f and g to continuous functions on $[0,1]$ by the linear interpolation on the intervals adjacent to C. Then $f(C) = g(C) = C$, hence the extensions have property (N). But $f + g$ fails to have this property, since the image of C is an interval due to the fact that $C + C$ is an interval. Passing to $\exp f$ and $\exp g$ we obtain (iii).

5.8.64. (Burenkov [**144**], [**145**]) (i) Construct an absolutely continuous function Φ on the real line and an infinitely differentiable function f such that the function $\Phi\bigl(f(x)\bigr)$ is not absolutely continuous on $[0,1]$.

(ii) Let Φ be a function of bounded variation on $[c,d]$ and let f be a differentiable function on $[a,b]$ such that f' is of bounded variation and $f([a,b]) \subset [c,d]$. Prove that the function $\Phi\bigl(f(x)\bigr)f'(x)$ is of bounded variation on $[a,b]$.

(iii) Let Φ be an absolutely continuous function on $[c,d]$ and let f be a differentiable function on $[a,b]$ such that f' is absolutely continuous and $f([a,b]) \subset [c,d]$.

Prove that the function $\Phi\bigl(f(x)\bigr)f'(x)$ has property (N) and is absolutely continuous on $[a,b]$.

HINT: (i), (ii) see in [**144**]; (iii) the function $(\Phi\circ f)f'$ vanishes on the set $\{f'=0\}$, and every point x in the complement of this set has a neighborhood where the continuously differentiable function f is monotone, hence in this neighborhood the function $\Phi\circ f$ is absolutely continuous (see Exercise 5.8.59). By the absolute continuity of f' we obtain the absolute continuity of the function $(\Phi\circ f)f'$ in the closed subintervals in the considered neighborhood. This gives property (N) and by (ii) implies the absolute continuity of $(\Phi\circ f)f'$ on $[a,b]$. We note that in view of the previous exercise, one cannot refer only to (N)-property of both factors as was done in [**144**], [**145**].

5.8.65. (i) (Banach, Saks [**58**], Bary, Menchoff [**67**]) A continuous function f has the form $f=\varphi\circ\psi$, where φ and ψ are absolutely continuous functions, precisely when f has the following property (S): for every $\varepsilon>0$, there exists $\delta>0$ such that the measure of the set $f(E)$ does not exceed ε whenever the measure of E does not exceed δ.

(ii) (Bary, Menchoff [**67**]) A continuous function f is the composition of two absolutely continuous functions precisely when f takes the set of all points x where there is no finite derivative to a measure zero set.

HINT: see Saks [**840**, Ch. IX, §8].

5.8.66. (i) (Fichtenholz [**293**]) Show that property (S) in the previous exercise does not follow from property (N).

(ii) (Banach [**51**]) Prove that a continuous function f on an interval has property (S) precisely when it has property (N) and assumes almost every value only at finitely many points.

HINT: see Saks [**840**, p. 410].

5.8.67.° (i) Show that if a sequence of increasing functions ψ_n on the real line converges to an increasing function ψ at all points of an everywhere dense set, then it converges to ψ at every point of continuity of ψ.

(ii) Let $\{\psi_n\}$ be a uniformly bounded sequence of increasing functions on $[a,b]$. Show that $\{\psi_n\}$ contains a pointwise convergent subsequence.

HINT: (i) let τ be a point of continuity of ψ and let $\varepsilon>0$. We find an interval $[\alpha,\beta]$ containing τ with the endpoints in the everywhere dense set of convergence such that $|\psi(t)-\psi(s)|<\varepsilon$ whenever $t,s\in[\alpha,\beta]$. There is $m\in\mathbb{N}$ such that $|\psi(\alpha)-\psi_n(\alpha)|<\varepsilon$ and $|\psi(\beta)-\psi_n(\beta)|<\varepsilon$ for all $n\geq m$. Then $|\psi(\tau)-\psi_n(\tau)|<3\varepsilon$, since we have $\psi_n(\alpha)\leq\psi_n(\tau)\leq\psi_n(\beta)$ and $\psi(\alpha)\leq\psi(\tau)\leq\psi(\beta)$.

(ii) By the diagonal method we find a subsequence $\{\psi_{n_k}\}$ convergent at all rational points. The limit function ψ can be extended to an increasing function on $[a,b]$, which has an at most countable set S of discontinuity points. According to (i), outside S one has pointwise convergence of $\{\psi_{n_k}\}$ to ψ. Now it remains to take in $\{\psi_{n_k}\}$ a subsequence convergent at every point of S.

5.8.68.° Let $f_1,\ldots,f_n$ be functions of bounded variation on the interval $[a,b]$ such that $\bigl(f_1(x),\ldots,f_n(x)\bigr)\in U\subset\mathbb{R}^n$ for all $x\in[a,b]$. Suppose that a function $\varphi\colon U\to\mathbb{R}^1$ satisfies the Lipschitz condition. Show that the composition $\varphi(f_1,\ldots,f_n)$ is a function of bounded variation on $[a,b]$.

5.8.69.° Let f and g be functions of bounded variation on $[a, b]$. Show that fg is a function of bounded variation on $[a, b]$, and if $g \geq c > 0$, then so is the function f/g.

HINT: use Exercise 5.8.68 applied to the functions $\varphi(x, y) = xy$ and $\varphi(x, y) = x/y$ on $[a, b] \times [a, b]$.

5.8.70.° Show that the space $BV[a, b]$ of all functions of bounded variation on $[a, b]$ is a Banach space with respect to the norm $\|f\|_{BV} = |f(a)| + V_a^b(f)$.

5.8.71.° (i) Show that every bounded nondecreasing function f on a set $T \subset \mathbb{R}^1$ is of bounded variation and $V(f, T) \leq 2 \sup_{t \in T} |f(t)|$.

(ii) Let f be a function of bounded variation on a set $T \subset \mathbb{R}^1$ and let $V(x) = V\bigl(f, (-\infty, x] \cap T\bigr)$, $x \in T$. Show that V and $V - f$ are nondecreasing functions on T and that the set of points of continuity of V coincides with the set of points of continuity of f.

(iii) Show that if a function f is of bounded variation on a set $T \subset \mathbb{R}^1$, then there exist two nondecreasing functions f_1 and f_2 on the whole real line such that $f = f_1 - f_2$ on T and $V(f, T) = V(f_1, \mathbb{R}^1) + V(f_2, \mathbb{R}^1)$.

5.8.72.° Suppose that a function f on a set $T \subset \mathbb{R}^1$ is of bounded variation. Show that f can be extended to $\mathbb{R}^1$ in such a way that $V(f, \mathbb{R}^1) = V(f, T)$.

HINT: use the previous exercise.

5.8.73.° Let f be a function of bounded variation on $[a, b]$. We redefine f at all discontinuity points, making it left continuous (the discontinuities of f are jumps). Show that the obtained function f_0 is of bounded variation and the following estimate holds: $V(f_0, [a, b]) \leq V(f, [a, b])$.

5.8.74. Let f_n be functions on $[a, b]$ such that $\sup_n V_a^b(f_n) \leq C < \infty$ and $f_n \to f$ in $L^1[a, b]$. Show that f coincides almost everywhere on $[a, b]$ with a function of bounded variation. In this case, we shall say that f is of essentially bounded variation defined by the formula $\|f\|_{BV} := \inf V_a^b(g)$, where inf is taken over all functions g of bounded variation that are equal almost everywhere to f.

HINT: take a subsequence f_{n_k} convergent on a set T of full measure in $[a, b]$, let $g = \lim f_{n_k}$ on T, observe that $V(g, T) \leq C$ and extend g to a function of bounded variation on $[a, b]$ according to Exercise 5.8.72. An alternative proof: Exercise 5.8.79.

5.8.75. Show that a measurable function f on $[a, b]$ is of essentially bounded variation if the following quantity is finite: $\mathrm{ess}V_a^b(f) := \sup\Bigl\{\sum_{i=1}^m |f(t_i) - f(t_{i-1})|\Bigr\}$, where sup is taken over all $m \in \mathbb{N}$ and all points $a < t_0 < t_1 < \ldots < t_m < b$ that are points of the approximate continuity of f.

HINT: approximate f by convolutions with smooth functions and apply the previous exercise.

5.8.76. (i) Show that an integrable function f coincides almost everywhere on $[a, b]$ with some function of bounded variation precisely when

$$\int_a^b |f(x + h) - f(x)| \, dx = O(h) \quad \text{as } h \to 0,$$

where we set $f = 0$ outside $[a, b]$.

(ii) Show that if

$$\int_a^b |f(x+h) - f(x)|\,dx = o(h) \quad \text{as } h \to 0,$$

then f almost everywhere on $[a,b]$ coincides with some constant.

HINT: (i) first verify the necessity of the above condition for nondecreasing functions. For the proof of sufficiency consider the functions $f_h = f * g_h$, where $g_h(x) = h^{-1}g(x/h)$ and g is a smooth probability density with support in $[0,1]$, next apply Exercise 5.8.74. To this end, verify that as $h \to 0$, the functions f_h have uniformly bounded variations on $[a,b]$ and $\|f - f_h\|_{L^1} \to 0$. One can also use Exercise 5.8.79. Another solution is given in Titchmarsh [**947**, Chapter XI, Exercise 10]. (ii) Show that f_h satisfies the same condition, hence $f_h'(x) = 0$; see also Titchmarsh [**947**, Chapter XI, Exercise 4].

5.8.77. Let f be a function of bounded variation on $[a,b]$ and let g be a nonnegative measurable function on the real line with unit integral. Show that the function

$$f * g(x) = \int_{-\infty}^{+\infty} f(x-y)g(y)\,dy,$$

where $f(x) = f(a)$ if $x \le a$ and $f(x) = f(b)$ if $x \ge b$, is of bounded variation and $V(f * g, \mathbb{R}^1) \le V(f, [a,b])$.

5.8.78. (i) Prove that a Borel measure μ on $\mathbb{R}^n$ is absolutely continuous with respect to Lebesgue measure if and only if $\lim\limits_{t\to 0} \|\mu_{th} - \mu\| = 0$ for every $h \in \mathbb{R}^n$, where $\mu_h(B) := \mu(B - h)$.

(ii) Prove that if a Borel measure μ on $\mathbb{R}^n$ is differentiable along n linearly independent vectors, then it is absolutely continuous with respect to Lebesgue measure.

HINT: (i) if the indicated condition holds and $B \in \mathcal{B}(\mathbb{R}^n)$, then the function $h \mapsto \mu(B - h)$ is continuous. Let $\sigma_j = p_j\,dx$, $p_j(x) = j^n p(x/j)$, where p is a smooth probability density. Then $\mu * \sigma_j \ll \lambda_n$ and $\mu * \sigma_j(B) \to \mu(B)$. The converse follows by 4.2.3. (ii) Show that $\|\mu_h - \mu\| \le \|d_h\mu\|$ and apply (i).

5.8.79. Prove that a Borel measure μ on (a,b) has a bounded measure as the generalized derivative along 1 precisely when μ has a density ϱ with respect to Lebesgue measure on (a,b) such that ϱ coincides a.e. with a function of bounded variation. In addition, in this case μ is given by the density $\nu\big((a,x]\big)$, where ν is the generalized derivative of μ.

5.8.80.° Let μ be a Borel measure on the real line (possibly signed) and let $F_\mu(t) = \mu\big((-\infty,t)\big)$. Show that the measure μ is mutually singular with Lebesgue measure if and only if $F_\mu'(t) = 0$ a.e.

HINT: $\mu = \mu_0 + \nu$, where the measure μ_0 is given by an integrable density ϱ and the measure ν is mutually singular with Lebesgue measure. One has $F_\mu'(t) = \varrho(t) + F_\nu'(t) = \varrho(t)$ a.e. by Theorem 5.8.8.

5.8.81.° Let μ be a signed Borel measure on the real line. Show that for all x one has $V\big(F_\mu, (-\infty,x]\big) = V\big(F_\mu, (-\infty,x)\big) = |\mu|\big((-\infty,x)\big)$ and $\|\mu\| = V(F_\mu, \mathbb{R}^1)$.

HINT: the left-hand side is estimated by the right-hand side, since

$$\sum_{i=1}^n \big|\mu\big([a_i,b_i)\big)\big| \le |\mu|\Big(\bigcup_{i=1}^n [a_i,b_i)\Big)$$

for all disjoint finite collections of intervals $[a_i, b_i)$. In order to verify the opposite inequality take $\varepsilon > 0$ and find disjoint compact sets $K_1, K_2 \subset (-\infty, x)$ with $|\mu|\big((-\infty, x)\big) \le \mu(K_1) - \mu(K_2) + \varepsilon$. Next take disjoint intervals $[a_1, b_1), \ldots, [a_n, b_n)$ and $[c_1, d_1), \ldots, [c_k, d_k)$ with

$$\mu(K_1) \le \sum_{i=1}^{n} \mu\big([a_i, b_i)\big) + \varepsilon, \quad |\mu(K_2)| \le \sum_{i=1}^{k} \big|\mu\big([c_i, d_i)\big)\big| + \varepsilon.$$

Then $|\mu|\big((-\infty, x)\big) \le V\big(F_\mu, (-\infty, x)\big) + 3\varepsilon$.

5.8.82.° (i) Let f be a function of bounded variation on $[a, b]$ that vanishes outside of an at most countable set. Show that $V(f, [a, x])' = 0$ a.e.

(ii) Let f be a function of bounded variation on $[a, b]$. Show that

$$V(f, [a, x])' = |f'(x)| \quad \text{a.e.}$$

HINT: (i) let $\{f \ne 0\} = \{x_i\}$, $f_n(x_i) = f(x_i)$ if $i \le n$ and $f_n(x) = 0$ if $x \notin \{x_1, \ldots, x_n\}$. The functions $V(f_n, [a, x])$ increase to $V(f, [a, x])$. By Proposition 5.2.8 one has $f_n' \to f'$ a.e. (ii) If f is left-continuous, then there is a Borel measure μ on $[a, b)$ such that $f(x) - f(a) = F_\mu(x)$ for all $x \in [a, b)$. One has $\mu = \mu_0 + \nu$, where the measure μ_0 is given by the density f', $\nu \perp \lambda$. Since $|\mu| = |\mu_0| + |\nu|$, one has a.e. $V(f, [a, x]) = V\big(f, [a, x)\big) = F_{|\mu|}(x) = F_{|\mu_0|}(x) + F_{|\nu|}(x)$. The measure $|\mu_0|$ is given by the density $|f'|$ and $|\nu| \perp \lambda$, whence $F_{|\nu|}'(x) = 0$ a.e. In the general case we redefine f on the at most countable set of discontinuity points and obtain a left-continuous function g of bounded variation. One has $g'(x) = f'(x)$ a.e. and $V(f, [a, x])' = V(g, [a, x])'$ a.e. by (i).

5.8.83. Let $F = (F_1, \ldots, F_n)\colon \mathbb{R}^n \to \mathbb{R}^n$ be a continuous mapping. Suppose that the functions F_i belong to the Sobolev class $W^{p,r}(\mathbb{R}^n)$, where $pr > n$. Prove that F has Lusin's property (N), i.e., takes all measure zero sets to measure zero sets.

HINT: let $r = 1$; by the Sobolev embedding theorem, for any fixed open cube K, there is a constant C such that, for every cube $Q \subset K$ and all $x, y \in Q$, one has

$$|F(x) - F(y)| \le C\|DF\|_{L^p(Q)}|x - y|^\alpha,$$

where $\alpha = 1 - n/p$. If a set $E \subset K$ has measure zero, then, given $\varepsilon > 0$, it can be covered by a sequence of closed cubes $Q_j \subset K$ with edges r_j and pairwise disjoint interiors such that $\sum_{j=1}^\infty r_j^n < \varepsilon$. The set $F(Q_j)$ is contained in the ball of radius $C\|DF\|_{L^p(Q_j)}\sqrt{n}r_j^\alpha$, whence $\lambda_n^*\big(F(E)\big) \le C^n n^{n/2} \sum_{j=1}^\infty \|DF\|_{L^p(Q_j)}^n r_j^{\alpha n}$, which is estimated by $\mathrm{const}\, \varepsilon^{(p-n)/p}$ by virtue of Hölder's inequality with the exponent p/n. In the case $r > 1$ the reasoning is similar.

5.8.84. (Sierpiński [**878**]) Let $f\colon \mathbb{R} \to \mathbb{R}$ be an arbitrary function and let $\{h_n\}$ be a sequence of nonzero numbers approaching zero. Prove that there exists a function $F\colon \mathbb{R} \to \mathbb{R}$ such that $\lim\limits_{n\to\infty} [F(x + h_n) - F(x)]/h_n = f(x)$ for all x.

HINT: see Bruckner [**135**] or Wise, Hall [**1022**, Example 3.14].

5.8.85. Prove that for every sequence of numbers $h_n > 0$ decreasing to zero, there exists a continuous function $F\colon [0, 1] \to \mathbb{R}$ such that, for every Lebesgue measurable function f on $[0, 1]$, there exists a subsequence $\{h_{n_k}\}$ with

$$\lim_{k\to\infty} \frac{F(x + h_{n_k}) - F(x)}{h_{n_k}} = f(x) \quad \text{a.e. on } [0, 1].$$

HINT: see Bruckner [**135**] or Wise, Hall [**1022**, Example 3.15].

5.8.86. (Fichtenholz [**289**], [**292**]) Let φ be a nondecreasing function on $[c,d]$, let $a=\varphi(c)$, $b=\varphi(d)$, and let a function f be integrable on the interval $[a,b]$. Let

$$F(x)=\int_a^x f(t)\,dt.$$

Suppose that the function $F\circ\varphi$ is absolutely continuous. Prove the equality

$$\int_a^b f(x)\,dx=\int_c^d f\bigl(\varphi(y)\bigr)\varphi'(y)\,dy.$$

HINT: let $E_1=\{x\colon\ \varphi'(x)>0\}$, $E_0=\{x\colon\ \varphi'(x)=0\}$, and let D be the set of all points of differentiability of F at which $F'=f$. Deduce from Lemma 5.8.13 that $\varphi(x)\in D$ for a.e. $x\in E_1$ and hence by the chain rule $(F\circ\varphi)'(x)=f\bigl(\varphi(x)\bigr)\varphi'(x)$. The left-hand side of the formula to be proven equals $F(b)-F(a)$. So, by the Newton–Leibniz formula for $F\circ\varphi$, it suffices to verify that the integral of $(F\circ\varphi)'$ equals the integral of $f(\varphi)\varphi'$, which by the above reduces to the verification of the equality of the integrals of these functions over E_0, i.e., to the equality

$$\int_{E_0}(F\circ\varphi)'\,dy=0.$$

Let $\varepsilon>0$. We take an open set $U\supset E_0$ with

$$\int_{U\backslash E_0}|(F\circ\varphi)'|\,dx<\varepsilon$$

and then take $\delta>0$ corresponding to ε in the definition of the absolute continuity of F; next, by Vitali's theorem, we choose a countable collection of pairwise disjoint intervals $(a_i,b_i)\subset U$ such that E_0 is covered by these intervals up to a set of measure zero and $\varphi(b_i)-\varphi(a_i)\le\delta(b_i-a_i)/(d-c)$ (every point $u\in E_0$ is contained in an arbitrarily small interval $(u-r,u+r)$ with $\varphi(u+r)-\varphi(u-r)\le 2\delta r/(d-c)$, since $\varphi'(a)=0$). Then the sum of lengths of $\bigl(\varphi(b_i),\varphi(a_i)\bigr)$ does not exceed δ and hence $\sum_{i=1}^\infty\bigl|F\bigl(\varphi(b_i)\bigr)-F\bigl(\varphi(a_i)\bigr)\bigr|\le\varepsilon$, whence we obtain

$$\Bigl|\sum_{i=1}^\infty\int_{a_i}^{b_i}(F\circ\varphi)'\,dx\le\varepsilon\Bigr|,$$

which yields

$$\Bigl|\int_{E_0}(F\circ\varphi)'\,dx\Bigr|\le 2\varepsilon.$$

5.8.87. Let f be an increasing absolutely continuous function on $[0,1]$ such that $f(0)=0$, $f(1)=1$. Set

$$D:=\{t\colon\ 0<f'(t)<+\infty\},\ g(t):=\inf\{s\in[0,1]\colon\ f(s)\le t\},\ t\in[0,1].$$

Prove that $f\bigl(g(t)\bigr)=t$, g is strictly increasing and for every bounded Borel function φ, the following equality holds:

$$\int_D\varphi(t)\,dt=\int_0^1\varphi\bigl(g(s)\bigr)g'(s)\,ds.$$

HINT: the equality $f\bigl(g(t)\bigr)=t$ follows by the continuity of f; it gives the injectivity of g. In addition, $f([0,1]\backslash D)\subset[0,1]\backslash f(D)$ because if $f(t)=f(t')$, where $t\in D$, then $t'=t$. The set $f([0,1]\backslash D)$ has measure zero, which follows by property (N) and Proposition 5.5.4 applied to the set $\{t\colon\ f'(t)=0\}$. Hence for a.e. t one has $g(t) \in D$, whence we obtain $f'\bigl(g(t)\bigr)g'(t) = 1$ a.e. If $\varphi(t) = \psi\bigl(f(t)\bigr)f'(t)$, where ψ is a bounded Borel function, then we arrive at the required formula because the integral of $\psi\bigl(f(t)\bigr)f'(t)$ over D equals the integral over $[0,1]$, and the latter equals the integral of ψ by the change of variables formula, but according to what has been said above $\psi(s) = \varphi\bigl(g(s)\bigr)/f'\bigl(g(s)\bigr) = \varphi\bigl(g(s)\bigr)g'(s)$ a.e. Hence our claim is true for $\varphi I_{\{1/n \le f' \le n\}}$, which yields the general case by passing to the limit.

5.8.88. Prove the following generalization of Vitali's Theorem 5.5.2 that was indicated by Lebesgue. Let A be a set in $\mathbb{R}^n$ and let $\mathcal{F}$ be a family of closed sets with the following property: for every $x \in A$, there exist a number $\alpha(x) > 0$, a sequence of sets $F_n(x) \in \mathcal{F}$ and a sequence of cubes $Q_n(x)$ such that $x \in Q_n(x)$, $F_n(x) \subset Q_n(x)$, $\lambda_n\bigl(F_n(x)\bigr) > \alpha(x)\lambda_n\bigl(Q_n(x)\bigr)$ and $\operatorname{diam} Q_n(x) \to 0$. Then $\mathcal{F}$ contains an at most countable subfamily of pairwise disjoint sets F_n whose union covers A up to a set of measure zero.

5.8.89. Let $(X, \mathcal{A}, \mu)$ be a space with a finite nonnegative measure. A family $\mathcal{D} \subset \mathcal{A}$ is called a Vitali system if it satisfies the following conditions: (a) $\varnothing, X \in \mathcal{D}$, all nonempty sets in $\mathcal{D}$ have positive measures, (b) if a set $A \subset X$ is covered by a collection $\mathcal{E} \subset \mathcal{D}$ in such a way that whenever $x \in A$, $B \in \mathcal{D}$ and $x \in B$, there exists $D \in \mathcal{E}$ with $x \in D \subset B$, then one can find an at most countable subcollection of disjoint sets $E_n \in \mathcal{E}$ with $\mu\Bigl(A \backslash \bigcup_{n=1}^\infty E_n\Bigr) = 0$. Suppose that $\mathcal{A}$ contains all singletons and that μ vanishes on them. Suppose we are given a sequence of countable partitions Π_n of the space X into measurable disjoint parts such that Π_{n+1} is a refinement of Π_n. Finally, suppose that the collection $\Pi = \bigcup_{n=1}^\infty \Pi_n$ is dense in the measure algebra $\mathcal{A}/\mu$ and, for every set Z of measure zero and every $\varepsilon > 0$, there exists $E_\varepsilon \in \Pi$ such that $Z \subset E_\varepsilon$ and $\mu(E_\varepsilon) < \varepsilon$. Prove that Π is a Vitali system. Show also that if a measure ν is absolutely continuous with respect to μ, then

$$\frac{d\nu}{d\mu}(x) = \lim_{k\to\infty} \frac{\nu\bigl(B_k(x)\bigr)}{\mu\bigl(B_k(x)\bigr)},$$

where $B_k(x) \in \Pi$ are chosen such that $x \in B_k(x)$, $0 < \mu\bigl(B_k(x)\bigr) < k^{-1}$.

HINT: see Rao [**788**, §5.3], Shilov, Gurevich [**867**, §10].

5.8.90. (i) Let f be a bounded measurable function on a cube in $\mathbb{R}^n$ with Lebesgue measure λ. Prove that the set of points of the approximate continuity of f coincides with the set of its Lebesgue points. In particular, if f is a bounded measurable function on $[0,1]$, then the derivative of the function

$$\int_0^x f(t)\,dt$$

equals $f(x)$ at every point x of the approximate continuity of f.

(ii) Prove that if a function f is integrable on a cube, then every Lebesgue point of f is a point of the approximate continuity, but the converse is not true.

HINT: (i) we may assume that $f(x_0) = 0$. If f is not approximately continuous at x_0, then we can find $\varepsilon > 0$, $q < 1$, and a sequence of balls B_k centered at x_0 with radii decreasing to zero such that $\lambda\bigl(\{|f| < \varepsilon\} \cap B_k\bigr) \le q\lambda(B_k)$. Then the integral of $|f|$ over B_k is not less than $(1-q)\varepsilon\lambda(B_k)$, i.e., x_0 is not a Lebesgue point. If x_0 is a point of the approximate continuity and $|f| \le 1$, then, for every $\varepsilon > 0$, the integral of $|f|$ over B_k does not exceed $\varepsilon\lambda(B_k) + \lambda\bigl(B_k \backslash \{|f| < \varepsilon\}\bigr)$, where the second summand is estimated by $\varepsilon\lambda(B_k)$ for all B_k of sufficiently small radius, i.e., x_0 is a

Lebesgue point. (ii) The first assertion has actually been proved in (i). In order to construct a counter-example to the converse consider the even function f on $[-1,1]$ such that $f(0)=0$, $f=0$ on $[2^{-n-1}, 2^{-n}-8^{-n})$, $f=4^n$ on $(2^{-n}-8^{-n}, 2^{-n})$ for all $n\in\mathbb{N}$.

5.8.91. Prove that the approximate continuity of a function f on $\mathbb{R}^n$ at a point x is equivalent to the equality $\operatorname{ap}\lim\limits_{y\to x} f(y)=f(x)$.

HINT: this equality follows at once from the approximate continuity; to prove the converse, we assume that $x=0$ and $f(x)=0$ and consider the sets $E_k=\{y\colon\ |f(y)|<1/k\}$ and the set $E=\bigcup_{k=1}^\infty \bigl(E_k\backslash[-\varepsilon_k,\varepsilon_k]^n\bigr)$, where $\varepsilon_k>0$ are decreasing to zero sufficiently rapidly.

5.8.92. Let us consider in $[0,1]$ the class Δ of all measurable sets every point of which is a density point, and the empty set.

(i) Prove that Δ is a topology that is strictly stronger than the usual topology of the interval. This topology is called the density topology.

(ii) Show that a function is continuous in the topology Δ precisely when it is Lebesgue measurable.

5.8.93. Prove Theorem 5.8.5.

HINT: prove the first equality by using representation (4.3.7), apply Exercise 4.7.51.

5.8.94. Prove that the spaces $W^{p,1}(\Omega)$ and $BV(\Omega)$ with the indicated norms are Banach spaces.

5.8.95. Prove that $f\in BV(\mathbb{R}^n)$ precisely when there exists a sequence of functions $f_j\in C_0^\infty(\mathbb{R}^n)$ such that $f_j\to f$ in $L^1(\mathbb{R}^n)$ and $\sup_n \|\,|\nabla f_j|\,\|_{L^1(\mathbb{R}^n)}<\infty$.

5.8.96. Prove that $f\in BV(\mathbb{R}^n)$ precisely when for every $i\le n$, the functions

$$\psi_i(x_1,\ldots,x_{n-1})(t)=f(x_1,\ldots,x_{i-1},t,x_i,\ldots,x_{n-1})$$

have bounded essential variations $\|\psi_i(x_1,\ldots,x_{n-1})(\,\cdot\,)\|_{BV}$ for a.e. $(x_1,\ldots,x_{n-1})$ in $\mathbb{R}^{n-1}$ (see Exercise 5.8.74) and

$$\int_{\mathbb{R}^{n-1}} \|\psi_i(x_1,\ldots,x_{n-1})(\,\cdot\,)\|_{BV}\,dx_1\cdots dx_{n-1}<\infty.$$

5.8.97. Suppose that a compact set E has a smooth boundary with the outer normal n. Show that $DI_E=\mathrm{n}\cdot\sigma_{\partial E}$, where $\sigma_{\partial E}$ is the surface measure on the boundary of E. In addition, the perimeter of E equals the surface measure of the boundary of E.

5.8.98. (i) Verify that $\ln|x|\in \mathrm{BMO}(\mathbb{R}^n)$;

(ii) Prove that $\ln|G(x)|\in\mathrm{BMO}(\mathbb{R}^n)$ for every polynomial G on $\mathbb{R}^n$ of degree $d\ge 1$.

(iii) Prove that for every positive number $\alpha<d^{-1}$, there exists a constant $c_{\alpha,d}$ such that for every polynomial G on $\mathbb{R}^n$ of degree $d\ge1$ one has

$$\int_B |G(x)|^{-\alpha}\,dx\le c_{\alpha,d}\Bigl(\int_B |G(x)|\,dx\Bigr)^{-\alpha},$$

where B is the unit ball.

HINT: see references in Stein [**906**, §V.6].

5.8.99. Verify that if $\omega \in A_p$, then the measure $\omega(x)\,dx$ has the doubling property.

HINT: apply (5.8.7) to $B = B(x_0, 2r)$ and $f = I_{B(x_0,r)}$.

5.8.100. Let μ be a nonnegative bounded Borel measure on $\mathbb{R}^n$ and let $B(x,r)$ be the closed ball of radius $r > 0$ centered at x. Prove that

$$\limsup_{y\to x} \mu\bigl(B(y,r)\bigr) \le \mu\bigl(B(x,r)\bigr).$$

HINT: observe that $\mu\bigl(B(x, r+1/k)\bigr) \to \mu\bigl(B(x,r)\bigr)$ and $B(y,r) \subset B(x, r+1/k)$ if $|x-y| < 1/k$.

5.8.101°. Prove Lemma 5.7.4.

HINT: for every n take a function δ_n for $\varepsilon = 2^{-n}$ such that $\delta_{n+1} \le \delta_n$; there is a tagged partition $\mathcal{P}_n$ subordinated to δ_n. The sums $I(f, \mathcal{P}_n)$ have a limit.

5.8.102°. Prove Proposition 5.7.6.

5.8.103°. Prove Lemma 5.7.9.

HINT: let $K_1 = [a_1, b_1], \ldots, K_m = [a_m, b_m]$ be the intervals in $[a,b]$ adjacent to the intervals in $\mathcal{P}_0$. Let $\varepsilon_1 > 0$. For every $j = 1, \ldots, m$, there exists a tagged partition $\mathcal{P}_j$ of the interval K_j such that $\mathcal{P}_j$ is subordinate to δ and one has the estimate $|F(b_j) - F(a_j) - I(f, \mathcal{P}_j)| < \varepsilon_1/m$. Then the collections $\mathcal{P}_j$, $j = 0, \ldots, m$, form a tagged partition $\mathcal{P}$ of the interval $[a,b]$, whence one has

$$\Bigl|I(f,\mathcal{P}_0) - \sum_{i=1}^n [F(d_i) - F(c_i)]\Bigr| \le \Bigl|I(f,\mathcal{P}) - \sum_{i=1}^n [F(d_i) - F(c_i)] - \sum_{j=1}^m [F(b_j) - F(a_j)]\Bigr|$$
$$+ \sum_{j=1}^m |F(b_j) - F(a_j) - I(f,\mathcal{P}_j)| < \varepsilon + \varepsilon_1,$$

which proves the first inequality, since ε_1 is arbitrary. The second inequality is proved similarly by using the first one.

5.8.104. Prove Proposition 5.8.34.

HINT: use (5.8.15).

5.8.105°. Let F be a closed set in $\mathbb{R}^n$. Prove that if x is a density point of F, then one has $\lim\limits_{|y|\to 0} \operatorname{dist}(x+y, F)/|y| = 0$.

5.8.106. Let $Z \subset \mathbb{R}^1$ be a set of measure zero. Show that there exists a measurable set E that has density at no point $z \in Z$, i.e., as $r \to 0$, there is no limit of the ratio $\lambda_n\bigl(E \cap K(z,r)\bigr)/\lambda_n\bigl(K(z,r)\bigr)$, where $K(z,r)$ is the ball of radius r centered at z.

HINT: see Goffman [**366**, p. 175], Kannan, Krueger [**488**, p. 41].

5.8.107. Let A be a convex compact set of positive measure in $\mathbb{R}^n$ and let B be the closed unit ball. (i) Prove that the limit $\lim\limits_{r\to 0} r^{-1}[\lambda_n(A + rB) - \lambda_n(A)]$ exists and equals the surface measure of the boundary of the set A. (ii) Prove that the same limit equals the mixed volume $v_{n-1,1}(A,B)$.

5.8.108. Deduce Theorem 5.8.22 from Theorem 5.8.21.

5.8.109° (Burstin [**152**]) Let f be an a.e. finite Lebesgue measurable function with periods $\pi_n \to 0$ (or, more generally, $f(x+\pi_n) = f(x)$ a.e.). Show that f coincides a.e. with some constant.

HINT: passing to arctg$f$ we may assume that f is bounded; letting

$$f_\varepsilon(x) = \int_{-\infty}^{+\infty} f(x-\varepsilon y)g(y)\,dy,$$

where g is a smooth function with support in $[0,1]$ and the integral 1, we observe that f_ε is a smooth function with periods π_n, whence $f_\varepsilon'(x) = 0$, i.e., f_ε is constant; use that as $\varepsilon \to 0$ the functions f_ε converge to f in measure on every interval.

5.8.110° (Lusin [**633**]) Let E be a measurable set in the unit circle S equipped with the linear Lebesgue measure. Suppose that E has infinitely many centers of symmetry, i.e., points $c \in S$ such that along with every point $e \in E$ the set E contains the point $e' \in S$ symmetric to e with respect to the straight line passing through c and the origin. Prove that the measure of E equals either 0 or 2π.

HINT: observe that the composition of two symmetries of the above type is a rotation (or a shift if we parameterize S by points of the interval), hence the indicator of E has arbitrarily small periods. Another solution is given in Lusin [**633**, p. 195].

5.8.111. Show that a function f on (a,b) is convex precisely when for every $[c,d] \subset (a,b)$, one has

$$f(x) = f(c) + \int_c^x g(t)\,dt, \quad x \in [c,d],$$

where g is a nondecreasing function on $[c,d]$.

HINT: verify first that f is locally Lipschitzian; see Krasnosel'skiĭ, Rutickiĭ [**546**, §1] or Natanson [**707**, Appendix 3].

5.8.112° Let f be an absolutely continuous function on $[a,b]$, let μ be a bounded Borel measure on $[a,b]$, and let $\Phi_\mu(t) = \mu\bigl([a,t)\bigr)$, $\Phi_\mu(a) = 0$.

(i) Prove the equality

$$\int_{[a,b]} f(t)\,\mu(dt) = f(b)\Phi_\mu(b+) - \int_a^b f'(t)\Phi_\mu(t)\,dt.$$

(ii) Let ν be a bounded Borel measure on $[a,b]$. Suppose that the function $\Phi_\nu(t) = \nu\bigl([a,t)\bigr)$ is continuous. Prove the equality

$$\int_{[a,b]} \Phi_\nu(t)\,\mu(dt) = \Phi_\mu(b+)\Phi_\nu(b) - \int_{[a,b]} \Phi_\mu(t)\,\nu(dt).$$

(iii) Let μ be a probability measure on $[0,+\infty)$ and $\Psi_\mu(t) = \mu\bigl([t,+\infty)\bigr)$. Suppose that a function f is absolutely continuous on every closed interval, $f(0) = 0$ and $f' \geq 0$. Prove the equality

$$\int_{[0,+\infty)} f(t)\,\mu(dt) = \int_0^{+\infty} f'(t)\Psi_\mu(t)\,dt.$$

Prove the same equality if the condition $f' \geq 0$ is replaced by the following condition: $f \in L^1(\mu)$ and $f'\Psi_\mu \in L^1(\mathbb{R}^1)$.

HINT: (i) by means of convolution construct a sequence of uniformly bounded measures μ_j with smooth densities p_j vanishing outside $[a-j^{-1}, b+j^{-1}]$ such that $\Phi_{\mu_j}(t) \to \Phi_\mu(t)$ at all points t of continuity of Φ_μ (i.e., everywhere, excepting possibly an at most countable set) and the integrals of every bounded continuous function g against μ_j approach the integral of g against μ. (ii) Use the same measures μ_j and observe that $\Phi_{\mu_j}(t) \to \Phi_\mu(t)$ for ν-a.e. t, since ν has no atoms. Now it suffices to consider the measures μ_j. Hence the required equality follows by (i), where we let $f = \Phi_{\mu_j}$ and take ν in place of μ. (iii) Take $a = 0$. Then, first for bounded f, in (i) pass to the limit as $b \to \infty$ taking into account the equality $\Psi_\mu(t) + \Phi_\mu(t) = 1$.

5.8.113. (i) (Lusin [**633**]) Construct a measurable set $E \subset [0,1]$ such that, letting $f = I_E$, one has

$$\int_0^1 \Bigl|\frac{f(x+t) - f(x-t)}{t}\Bigr|\,dt = \infty \quad \text{for almost all } x \in [0,1].$$

(ii) (Titchmarsh [**946**]) Let $\varphi > 0$ be a continuous function on $(0,1)$ such that the function $1/\varphi$ has an infinite integral. Prove that:

(a) there exists a continuous function f such that

$$\int_0^1 \frac{|f(x+t) - f(x-t)|}{\varphi(t)}\,dt = \infty \quad \text{for a.e. } x;$$

(b) there exists a continuous function g such that the integral of the function $[g(x+t) - g(x)]/\varphi(t)$ in t diverges for a.e. x.

HINT: (i) see Lusin [**633**, p. 464], where it is noted that such examples were constructed by E.M. Landis, V.A. Hodakov and other mathematicians.

5.8.114. Construct a continuous function f such that for every point x in some everywhere dense set of cardinality of the continuum in $[0,1]$, there is no finite limit

$$\lim_{\varepsilon\to 0} \int_\varepsilon^1 \frac{f(x+t) - f(x-t)}{t}\,dt.$$

HINT: see Lusin [**633**, p. 459].

5.8.115. (Rubel [**830**]) Let f be a finite measurable real function on the real line. We consider the following functions with values in $[0,+\infty]$:

$$\varphi(x) = \sup_t |f(x+t) - f(x)|, \quad \varphi^*(x) = \sup_t |f(x+t) - f(x-t)|,$$

$$\Phi(x) = \sup_{t\neq 0}\Bigl|\frac{f(x+t) - f(x)}{t}\Bigr|, \quad \Phi^*(x) = \sup_{t\neq 0}\Bigl|\frac{f(x+t) - f(x-t)}{t}\Bigr|.$$

Show that the functions φ^* and Φ^* may not be measurable, although φ and Φ are always measurable.

HINT: take a set $E \subset [0,1]$ of measure zero such that $E + E$ is non-measurable (Exercise 1.12.67), and let $f(x) = 1$ if $x \in E$, $f(x) = -1$ if $x \in E + 2$, $f(x) = 0$ in all other cases. The measurability of φ follows by the measurability of the function $\sup_t [f(x+t) - f(x)] = \sup_y f(y) - f(x)$. Similarly, one verifies the measurability of Φ.

5.8.116. (N.N. Lusin, D.E. Menchoff) Let $E \subset \mathbb{R}^n$ be a set of finite measure and let $K \subset E$ be a compact set such that E has density 1 at every point of K. Prove that there exists a compact set P without isolated points such that $K \subset P \subset E$ and P has density 1 at every point of K.

HINT: see Bruckner [**135**, pp. 26–28].

5.8.117. Let f be a function of bounded variation on $[a,b]$ such that

$$V(f,[a,b]) = \int_a^b |f'(t)|\,dt.$$

Prove that f is absolutely continuous.

HINT: write f in the form $f = f_1 + f_2$, where f_1 is an absolutely continuous function and $f_2' = 0$ a.e. Verify that $V(f,[a,b]) = V(f_1,[a,b]) + V(f_2,[a,b])$, whence $V(f_2,[a,b]) = 0$ and $f_2 = 0$. An alternative argument: observe that $V(f,[a,c])$ coincides with the integral of $|f'|$ over $[a,c]$ for all $c \in [a,b]$, since $V(f,[c,b])$ is estimated by the integral of $|f'|$ over $[c,b]$. Hence the function $V(f,[a,x])$ is absolutely continuous, which yields the absolute continuity of f.

5.8.118. (i) (Lusin [**632**], [**635**]) Prove that there exists no continuous function f on $[0,1]$ such that $f'(x) = +\infty$ on a set of positive measure.

(ii) Deduce from Theorem 5.8.12 that there is no function with the property mentioned in (i).

5.8.119. (i) (Lusin [**633**]) Prove that there exists a continuous function f on $[0,1]$ such that $f'(x)$ exists a.e. and $f'(x) > 1$ a.e., but in no interval is f increasing.

(ii) (Zahorski [**1045**]) Suppose that a set $E \subset \mathbb{R}^1$ is the countable union of compact sets and every point of E is its density point. Prove that there exists an approximately continuous function φ such that $0 < \varphi(x) \le 1$ if $x \in E$ and $\varphi(x) = 0$ if $x \notin E$.

(iii) Show that there exists an everywhere differentiable function f on the real line such that f' is discontinuous almost everywhere.

(iv) Prove that there exists a differentiable function f on $[0,1]$ with a bounded derivative such that on no interval is f monotone.

HINT: (i) construct a measurable finite function $g > 1$ such that g is not integrable on any interval, then by Theorem 5.1.4 take a continuous function f with $f' = g$ a.e. (ii) Use Exercise 5.8.116; see Zahorski [**1045**, Lemma 11] or Bruckner [**135**, Ch. 2, §6], (iii) Take in (ii) a set E with the everywhere dense complement of measure zero and let

$$f(x) = \int_0^x \varphi(t)\,dt.$$

(iv) See Bruckner [**135**, Ch. 2, §6], Denjoy [**213**], Katznelson, Stromberg [**496**].

5.8.120. (Hahn [**393**], Lusin [**633**, pp. 92–94]) Construct two different continuous functions F and G on $[0,1]$ such that $F(0) = G(0) = 0$ and $F'(x) = G'(x)$ at every point $x \in [0,1]$, where infinite values of the derivatives are allowed.

5.8.121. (Tolstoff [**952**]) Let D be a bounded region in $\mathbb{R}^2$ whose boundary ∂D is a simple piece-wise smooth curve and let φ be a mapping that is continuously differentiable in a neighborhood of the closure of D and maps ∂D one-to-one to a contour Γ bounding a region G. Prove that for every bounded measurable function f on G one has the equality

$$\int_G f\,dx = k \int_D f\bigl(\varphi(y)\bigr) \det \varphi'(y)\,dy,$$

where k is the sign of the integral of $\det \varphi'(y)$ over D (here k is automatically nonzero).

5.8.122.° (i) Suppose that a function f is integrable on $[0,1]$ and a function $\varphi\colon [0,1] \to [0,1]$ is continuously differentiable. Is it true that the function $f\bigl(\varphi(x)\bigr)\varphi'(x)$ is integrable?

(ii) Let a function f be integrable on $[a,b]$, let a function $\varphi\colon [c,d] \to [a,b]$ be absolutely continuous, and let $\varphi([c,d]) \subset [a,b]$. Suppose, in addition, that the function $f\bigl(\varphi(x)\bigr)\varphi'(x)$ is integrable on $[c,d]$. Prove the equality

$$\int_{\varphi(c)}^{\varphi(d)} f(x)\,dx = \int_c^d f\bigl(\varphi(y)\bigr)\varphi'(y)\,dy.$$

HINT: (i) no; consider $f(x) = x^{-1/2}$ and a smooth function φ such that $\varphi(n^{-1}) = 0$, $\varphi(c_n) = n^{-2}$, where c_n is the middle point of $\bigl[(n+1)^{-1}, n^{-1}\bigr]$, and φ is increasing on $\bigl[(n+1)^{-1}, c_n\bigr]$. (ii) If f is continuous, then this equality follows by the Newton–Leibniz formula for the function $F\circ\varphi$, where

$$F(x) = \int_a^x f(t)\,dt.$$

This function is absolutely continuous and $(F\circ\varphi)'(x) = f\bigl(\varphi(x)\bigr)\varphi'(x)$. Hence the image of the measure $\mu := \varphi'\cdot\lambda$ on the interval $[c,d]$ (possibly signed) under the mapping φ coincides with Lebesgue measure on $[a,b]$. This gives the desired equality for any bounded Borel function f, as its right-hand side is the integral of f with respect to the measure $\mu\circ\varphi^{-1}$. By Lemma 5.8.13 our equality extends to bounded Lebesgue measurable functions f. Indeed, we take a bounded Borel function f_0 equivalent to f and set $E := \{x\colon f_0(x) \neq f(x)\}$; then $\varphi(y) \notin E$ for almost every point y at which φ has a finite nonzero derivative, and for almost all other points y one has $f\bigl(\varphi(y)\bigr)\varphi'(y) = f_0\bigl(\varphi(y)\bigr)\varphi'(y) = 0$. Now one can easily pass from bounded f to the general case by using the integrability of f and $f(\varphi)\varphi'$.

5.8.123. (i) Let $E \subset \mathbb{R}^1$ be a set of positive Lebesgue measure. Prove that the set $\mathbb{R}^1\backslash\bigcup_{n=1}^\infty (E + r_n)$, where $\{r_n\} = \mathbb{Q}$, has measure zero.

(ii) Let $A \subset \mathbb{R}^1$ be a set of positive outer measure and let B be an everywhere dense set in $\mathbb{R}^1$. Prove that for every interval I one has $\lambda^*\bigl((A+B)\cap I\bigr) = \lambda(I)$, where λ is Lebesgue measure.

(iii) Suppose we are given two sets $A, B \subset \mathbb{R}^1$ of positive outer measure. Prove that there exists an interval I such that $\lambda^*\bigl((A+B)\cap I\bigr) = \lambda(I)$.

(iv) Construct two sets A and B of positive outer measure on the real line such that $A+B$ contains no open interval.

(v) Suppose we are given two sets A and B of positive outer measure on the real line such that at least one of them is measurable. Prove that $A+B$ contains some open interval.

HINT: (i) is easily deduced from the existence of a density point of E. For other assertions, see Miller [**692**]. We note that (ii) is called Smítal's lemma.

5.8.124. Let $E \subset \mathbb{R}^1$ and $x \in E$. Denote by $\lambda(E, x, x+h)$ the length of the maximal open interval in $(x, x+h)$ that contains no points of E (if $h < 0$, then we consider the interval $(x - |h|, x)$). Let

$$p(E,x) = \limsup_{h\to 0} \lambda(E, x, x+h)/|h|.$$

The set E is said to be porous at x if $p(E,x) < 1$. If E is porous at every point, then we call E a porous set. Finally, a countable union of porous sets is called σ-porous (this concept is due to E.P. Dolzhenko [**230**]).

(i) Prove that every porous set has Lebesgue measure zero and is nowhere dense.

(ii) Construct a compact set of measure zero that is not σ-porous.

(iii) Construct a Borel probability measure on the real line that is singular with respect to Lebesgue measure, but vanishes on every σ-porous set.

(iv) Construct a compact set K on the real line such that every Borel measure on K is concentrated on a σ-porous set.

HINT: see Thomson [**943**, §A11], Tkadlec [**949**], Humke, Preiss [**447**], Zajíček [**1046**].

5.8.125. Prove that every set of positive Lebesgue measure in $\mathbb{R}^2$ contains the vertices of some equilateral triangle.

HINT: we may assume that the origin belongs to the set E and is its density point. We choose a disc U with the center at the origin and radius r such that $\lambda_2(U\cap E) > 10\lambda_2(U)/11$. We may assume that $r=1$. Denote by Φ the set of all points $\varphi\in[0,2\pi)$ such that

$$\int_{E_\varphi} r\,dr \geq 2/5,$$

where E_φ is the set of all points $t\in(0,1]$ such that the point with the polar coordinates (t,φ) belongs to E. Then $10\pi/11 \geq \lambda(\Phi)/2 + 2\bigl(2\pi-\lambda(\Phi)\bigr)/5$, whence one has $\lambda(\Phi)>\pi$; hence there exists an angle $\varphi\in E$ such that $\psi:=\varphi+\pi/6\in\Phi$, where the addition is taken $\operatorname{mod} 2\pi$. We observe that the sets E_φ and E_ψ have a nonempty intersection (since the integral of r over their union does not exceed $1/2$). Let us take a point t in their intersection; we obtain in E two points (t,φ) and (t,ψ) in the polar coordinates that along with the origin form the vertices of an equilateral triangle.

5.8.126°. (Fischer [**299**]) Let a function F be continuous on $[0,1]$ and $F(0)=0$. Prove that F is the indefinite integral of a function in $L^2[0,1]$ precisely when the sequence of functions $n\bigl(F(x+n^{-1})-F(x)\bigr)$ is fundamental in $L^2[0,1]$, where for $x>1$ we set $F(x)=F(1)$.

HINT: if this sequence is fundamental, then it converges in $L^2[0,1]$ to some function f; then

$$\lim_{n\to\infty}\int_0^t n\bigl(F(x+n^{-1})-F(x)\bigr)\,dx = \int_0^t f(x)\,dx$$

for all $t\in[0,1]$. By the continuity of F the left-hand side equals $F(t)-F(0)=F(t)$, as one can see from the equality

$$\int_0^t F(x+1/n)\,dx = \int_{1/n}^{t+1/n} F(y)\,dy.$$

5.8.127. (Denjoy [**214**]) Let a function f be differentiable on $(0,1)$ and let α and β be such that the set $\{x\colon\ \alpha<f'(x)<\beta\}$ is nonempty. Prove that this set has positive measure.

HINT: see Kannan, Krueger [**488**, §5.4].

5.8.128. Give an example of a measurable function on $[0,1]$ that has the Darboux property, i.e., on every interval $[a,b]\subset[0,1]$ it assumes all the values between $f(a)$ and $f(b)$, but does not have the Denjoy property from the previous exercise, i.e., there exist c and d such that the set $\{x\colon\ c<f(x)<d\}$ is nonempty and has measure zero.

HINT: let C be the Cantor set, let $\psi\colon C \to [0,1]$ be a bijection, and let $\{U_n\}$ be complementary open intervals to C in $[0,1]$. Take non-constant affine functions g_n with $g_n([0,1]) \subset U_n$ and set $f(x) = \psi\bigl(g_n^{-1}(x)\bigr)$ if $x \in g_n(C)$, $f(x) = x$ if $x \in C$ and $f(x) = 1$ at all other points. Then $f^{-1}(1/2, 1)$ is nonempty, but has measure zero. If $a < b$, then (a, b) contains some interval U_n in the above-mentioned sequence, hence (a, b) contains the set $g_n(C)$, on which f assumes all values from $[0, 1]$.

5.8.129. (Davies [**207**]) Let a function f on $[0,1]^2$ be approximately continuous in every variable separately. (i) Prove that f is Lebesgue measurable. (ii) Prove that f even belongs to the second Baire class.

5.8.130.° Prove the following Chebyshev inequality for monotone functions: if φ and ψ are nondecreasing finite functions on $[0,1]$ and ϱ is a probability density on $[0, 1]$, then

$$\int_0^1 \varphi(x)\psi(x)\varrho(x)\,dx \geq \int_0^1 \varphi(x)\varrho(x)\,dx \int_0^1 \psi(x)\varrho(x)\,dx.$$

If φ is an increasing function and ψ is decreasing, then the opposite inequality is true.

HINT: subtracting a constant from ψ, we may assume that $\psi\varrho$ has the zero integral. Then, letting

$$\Psi(x) = \int_0^x \psi(t)\varrho(t)\,dt,$$

we have $\Psi(0) = \Psi(1) = 0$, whence $\Psi(x) \leq 0$ by the monotonicity of ψ. If φ is continuously differentiable, then $\varphi' \geq 0$ and hence

$$\int_0^1 \varphi(x)\psi(x)\varrho(x)\,dx = -\int_0^1 \varphi'(x)\Psi(x)\,dx \geq 0.$$

The general case reduces to the considered one because there exists a uniformly bounded sequence of smooth nondecreasing functions φ_n convergent a.e. to φ. For example, one can take $\varphi_n = \varphi * f_n$, where $f_n(t) = nf(t/n)$, f is a smooth probability density with support in $[0, 1]$ and φ is extended by constant values to $[-1, 0]$ and $[1, 2]$. If ψ is decreasing, then we pass to $1 - \psi$ and obtain the opposite inequality.

5.8.131. Let $f\colon [0,\infty) \to [0,\infty)$ be a locally integrable function.

(i) Assume

$$\int_E f(x)\,dx \leq \sqrt{\lambda(E)}$$

for every bounded measurable set E. Prove that

$$\int_0^\infty \frac{f(x)}{1+x}\,dx \leq \int_0^\infty \frac{\sqrt{x}}{(1+x)^2}\,dx < \frac{1}{2}.$$

(ii) Assume that

$$\int_0^T f(x)\,dx \leq T \quad \text{for all } T.$$

Show that the function

$$\frac{f(x)}{1+x^2}$$

is integrable.

HINT: (i) let

$$F(x) = \int_0^x f(y)\,dy,$$

then $F(x) \le \sqrt{x}$ and hence

$$\int_0^t \frac{f(x)}{1+x}\,dx = \frac{F(t)}{1+t} + \int_0^t \frac{F(x)}{(1+x)^2}\,dx \le \frac{F(t)}{1+t} + \int_0^t \frac{\sqrt{x}}{(1+x)^2}\,dx.$$

It remains to let $t \to +\infty$. Assertion (ii) is analogous.

5.8.132. (Gordon [**374**]) In analogy with the definitions in §5.7 we shall consider tagged partitions $P = \{(x_i, E_i)\}$ of the interval $[a,b]$ into finitely many pairwise disjoint measurable sets E_i with $x_i \in E_i$. Such a partition P is said to be subordinate to a positive function δ if $E_i \subset \bigl(x_i - \delta(x_i), x_i + \delta(x_i)\bigr)$ for all i. Prove that:

(i) a function f on $[a,b]$ is Riemann integrable precisely when there exists a number R with the following property: for every $\varepsilon > 0$, there exists a number $\delta > 0$ such that $\Bigl|\sum_{i=1}^n f(x_i)\lambda(E_i) - R\Bigr| < \varepsilon$ for every tagged partition of the interval into measurable sets E_i subordinate to δ;

(ii) a function f on $[a,b]$ is Lebesgue integrable precisely when there exists a number L with the following property: for every $\varepsilon > 0$, there exists a positive function $\delta(\,\cdot\,)$ such that $\Bigl|\sum_{i=1}^n f(x_i)\lambda(E_i) - L\Bigr| < \varepsilon$ for every tagged partition of the interval into measurable sets E_i subordinate to the function δ.

5.8.133. Given a function f on $[a,b]$, its Banach indicatrix $N_f\colon \mathbb{R}^1 \to [0,+\infty]$ is defined as follows: $N_f(y)$ is the cardinality of the set $f^{-1}(y)$.

(i) Prove that the indicatrix of a continuous function is measurable as a mapping with values in $[0,+\infty]$.

(ii) (Banach [**50**]) Prove that a continuous function f is of bounded variation precisely when the function N_f is integrable. In addition, one has

$$\int_{-\infty}^{+\infty} N_f(y)\,dy = V(f,[a,b]). \tag{5.8.19}$$

In particular, $N_f(y) < \infty$ a.e.

(iii) (H. Kestelman) Prove that for a general function f of bounded variation, the difference between the left and right sides of (5.8.19) equals the sum of the absolute values of all jumps of f.

(iv) Prove that if a function f is continuous and $N_f(y) < \infty$ for all y, then f is differentiable almost everywhere.

HINT: (i) let us partition $[a,b]$ into 2^n equal intervals $I_{n,k}$, $k = 1,\ldots,2^n$, such that the first one is closed and the other ones do not contain left ends. Set

$$g_{n,k} = I_{f(I_{n,k})}, \quad g_n = \sum_{k=1}^{2^n} g_{n,k}.$$

It is easily seen that the functions g_n increase pointwise to N_f (see Natanson [**707**, Ch. VIII, §5]). In addition, these functions are Borel measurable. (ii) The integral of $g_{n,k}$ equals the oscillation of f on $I_{n,k}$. One can deduce from this that the integrals of g_n converge to $V(f,[a,b])$, which by the monotone convergence theorem yields the desired equality. (iii) See Kannan, Krueger [**488**, §6.1]. (iv) See van Rooij, Schikhof [**820**, §21]. The Banach indicatrix is also studied in Lozinskiĭ [**624**].

5.8.134. (i) Let f be a continuous function on $[a,b]$ and let E be a Borel set in $[a,b]$. Show that the function $y \mapsto N_f(E,y)$ from $\mathbb{R}^1$ to $[0,+\infty]$ which to every y puts into correspondence the cardinality of the set $E \cap f^{-1}(y)$ is Borel measurable.

(ii) Let f be a continuous function of bounded variation on $[a,b]$ and let $V(x) = V(f,[a,x])$. Prove that for every Borel set B in $[a,b]$ one has

$$\lambda\bigl(V(B)\bigr) = \int_{-\infty}^{+\infty} N_f(B,y)\,dy.$$

Deduce that if $E \subset [a,b]$ is such that $\lambda\bigl(f(E)\bigr) = 0$, then $\lambda\bigl(V(E)\bigr) = 0$.

HINT: (i) if E is a closed interval, then the previous exercise applies. The class of all Borel sets E for which the assertion is true is σ-additive. It remains to apply Theorem 1.9.3(ii) to the class of closed intervals in $[a,b]$, as it generates the σ-algebra $\mathcal{B}([a,b])$. (ii) If $B = [c,d] \subset [a,b]$, then $\lambda\bigl(V([c,d])\bigr) = V(d) - V(c)$ is the integral of $N_f([c,d],y)$ over the real line by the previous exercise. Assertion (i) yields that the right-hand side of the desired equality is a measure as a function of B. The left-hand side is a measure too. Indeed, the function V is increasing, hence for any disjoint sets A and B the set $V(A)\cap V(B)$ is at most countable. The equality of two measures on all closed intervals implies their coincidence on $\mathcal{B}([a,b])$. In order to prove the last assertion take a Borel set $S \supset f(E)$ with $\lambda(S) = 0$ and observe that $\lambda\bigl(V\bigl(f^{-1}(S)\bigr)\bigr) = 0$, since the integral of $N_f\bigl(f^{-1}(S),y\bigr)$ over the real line equals the integral over S.

5.8.135.° Let μ be a measure on a space X and let a function f on $X\times[a,b]$ be such that the functions $x \mapsto f(x,t)$ are integrable and the functions $t \mapsto f(x,t)$ are absolutely continuous. Suppose that the function $\partial f/\partial t$ is integrable with respect to $\mu\otimes\lambda$, where λ is Lebesgue measure. Prove that the function

$$t \mapsto \int_X f(t,x)\,\mu(dx)$$

is absolutely continuous and

$$\frac{d}{dt}\int_X f(t,x)\,\mu(dx) = \int_X \frac{\partial f(x,t)}{\partial t}\,\mu(dx) \quad \text{a.e.}$$

HINT: apply the Newton–Leibniz formula and Fubini's theorem.

5.8.136. (Tolstoff [**951**]) (i) Let φ be a positive monotone function on $(0,1]$ with $\lim\limits_{h\to 0}\varphi(0) = 0$. Prove that for every $\alpha \in (0,1)$, there exists a perfect nowhere dense set $P \subset [0,1]$ of Lebesgue measure α such that for a.e. $x \in P$, there exists a number $\delta(x) > 0$ for which one has $\lambda\bigl((x,x+h)\backslash P\bigr) < \varphi(|h|)|h|$ whenever $|h| < \delta(x)$.

(ii) Prove that for every $\alpha \in (0,1)$, there exists a perfect nowhere dense set $P \subset [0,1]$ of Lebesgue measure α such that for sufficiently small $|h|$ one has

$$\lambda\bigl((x,x+h)\backslash P\bigr) > \varphi(|h|)|h| \quad \text{for all } x \in P.$$

(iii) Let P be a perfect set in $[0,1]$, $[0,1]\backslash P = \bigcup_{n=1}^\infty U_n$, where the U_n's are disjoint intervals. Suppose that $\lambda(U_n) \le q^n$, where $0 < q < 1$. Show that for every $\alpha > 0$, for a.e. x there exists $\delta(x) > 0$ such that $\lambda\bigl((x,x+h)\backslash P\bigr) < |h|^\alpha$ whenever $|h| < \delta(x)$.

(iv) Show that there exist two mutually complementary measurable sets A and B in $[0,1]$ such that $\lambda\bigl(A\cap(x,x+h)\bigr) > \varphi(|h|)|h|$ for a.e. $x \in B$ whenever $|h| < \delta(x)$ and $\lambda\bigl(B\cap(x,x+h)\bigr) > \varphi(|h|)|h|$ for a.e. $x \in A$ whenever $|h| < \delta(x)$.

5.8.137. (Bary [**66**, Appendix, §13]) Let $E \subset [0,1]$ be a measurable set of positive measure and let E_0 be the set of density points of E. Prove that for every $x \in E_0$ and every number α there are numbers λ_n such that $\lambda_n = \alpha n^{-1} + o(1/n)$, $x + \lambda_n \in E_0$, $x - \lambda_n \in E_0$ for all n.

5.8.138. (Brodskiĭ [**130**]) Suppose we are given a continuously differentiable function f on the plane such that its partial derivatives at the point (x_0, y_0) do not vanish. Suppose that x_0 is a density point of a measurable set $A \subset \mathbb{R}^1$ and y_0 is a density point of a measurable set $B \subset \mathbb{R}^1$. Prove that in some neighborhood of $f(x_0, y_0)$, every point has the form $f(x, y)$ with $x \in A$, $y \in B$.

5.8.139. Prove that a necessary and sufficient condition that two sets on the real line are metrically separated in the sense of Exercise 1.12.160 is that at almost all points of one set the density of the other set is zero.

HINT: see Kannan, Krueger [**488**, p. 247].

5.8.140. Let $f = (f_1, \ldots, f_n)\colon \mathbb{R}^n \to \mathbb{R}^n$, where $f_i \in W^{1,1}(\mathbb{R}^n)$. Set

$$\Omega = \bigl\{\det(\partial_{x_j} f_i)_{i,j\le n} \neq 0\bigr\}$$

and denote by $\lambda|_\Omega$ the restriction of Lebesgue measure to Ω. Show that the measure $\lambda|_\Omega \circ f^{-1}$ is absolutely continuous.

HINT: apply Theorem 5.8.27 or the next exercise.

5.8.141. Show that the assertion of the previous exercise remains true for any measurable functions f_i provided that Ω is the set of points where the approximate partial derivatives $\mathrm{ap}\partial_{x_j} f_i$ exist and $\det(\mathrm{ap}\partial_{x_j} f_i)_{i,j\le n} \neq 0$.

HINT: apply Theorem 5.8.14.

5.8.142. (Bogachev, Kolesnikov [**107**]) Let U be an open ball in $\mathbb{R}^d$ and let $F\colon U \to \mathbb{R}^d$ be an integrable mapping such that its derivative DF in the sense of generalized functions is a bounded measure with values in the space of nonnegative symmetric matrices. Let $D_{\mathrm{ac}}F$ be the operator-valued density of the absolutely continuous component of DF and let $\Omega := \{x\colon \det D_{\mathrm{ac}}F(x) > 0\}$. Prove that the measure $\lambda|_\Omega \circ F^{-1}$, where λ is Lebesgue measure, is absolutely continuous.

Bibliographical and Historical Comments

> One gets a strange feeling having seen the same drawings as if drawn by the same hand in the works of four scholars that worked completely independently of each other. An involuntary thought comes that such a striking, mysterious activity of mankind, lasting several thousand years, cannot be occasional and must have a certain goal. Having acknowledged this, we come by necessity to the question: what is this goal?
>
> *I.R. Shafarevich.* On some tendencies of the development of mathematics.

> However, also in my contacts with the American Shakespeare scholars I confined myself to the concrete problems of my research: dating, identification of prototypes, directions of certain allusions. I avoided touching the problem of personality of the Great Bard, the "Shakespeare problem"; neither did I hear those scholars discussing such a problem between themselves.
>
> *I.M. Gililov.* A play about William Shakespeare or the Mystery of the Great Phoenix.

The extensive bibliography in this book covers, however, only a small portion of the existing immense literature on measure theory; in particular, many authors are represented by a minimal number of their most characteristic works. Guided by the proposed brief comments and this incomplete list, the reader, with help of modern electronic data-bases, can considerably enlarge the bibliography. The list of books is more complete (although it cannot pretend to be absolutely complete). For the reader's convenience, the bibliography includes the collected (or selected) works of A.D. Alexandrov [**15**], R. Baire [**47**], S. Banach [**56**], E. Borel [**114**], C. Carathéodory [**166**], A. Denjoy [**215**], M. Fréchet [**321**], G. Fubini [**333**], H. Hahn [**401**], F. Hausdorff [**415**], S. Kakutani [**482**], A.N. Kolmogorov [**535**], Ch.-J. de la Vallée Poussin [**575**], H. Lebesgue [**594**], N.N. Lusin [**637**], E. Marczewski [**652**], J. von Neumann [**711**], J. Radon [**780**], F. Riesz [**808**], V.A. Rohlin [**817**], W. Sierpiński [**881**], L. Tonelli [**956**], G. Vitali [**990**], N. Wiener [**1017**], and G. &W. Young [**1027**], where one can find most of their cited works along with other papers related to measure theory. Many works in the bibliography are only cited in the main text in connection with concrete results (including exercises and hints). Some principal results are accompanied by detailed comments; in many other cases we mention only the final works, which should be consulted concerning the previous publications or the history of the question. Dozens of partial results mentioned in the book have an extremely interesting history, revealed through the reading of old journals, the exposition of which I had to omit with regret.

Most of the works in the bibliography are in English and French; a relatively small part of them (in particular, some old classical works) are in German, Russian, and Italian. For most of the Russian works (excepting a limited number of works from the 1930s–60s), translations are indicated. The reader is warned that in such cases, the titles and author names are given according to the translation even when versions more adequate and closer to the original are possible. Apart from the list of references, I tried to be consistent in the spelling of such names as Prohorov, Rohlin, Skorohod, and Tychonoff, which admit different versions. The letter "h" in such names is responsible for the same sound as in "Hardy" or "Halmos", but in different epochs was transcribed differently, depending on to which foreign language (French, German, or English) the translation was made. Nowadays in official documents it is customary to represent this "h" in the Russian family names as "kh" (although, it seems, just "h" would be enough).

Now several remarks are in order on books on Lebesgue measure and integration. The first systematic account of the theory was given by Lebesgue himself in the first edition of his lectures [**582**] in 1904. In 1907, the first edition of the fundamental textbook by Hobson [**436**] was published, where certain elements of Lebesgue's theory were included (in later editions the corresponding material was considerably reworked and enlarged); next the books by de la Vallée Poussin [**572**] (note that in later editions the Lebesgue integral is not considered) and [**574**] and Carathéodory [**164**] appeared. It is worth noting that customarily the form La Vallée Poussin de is used for the alphabetic ordering; however, in some libraries this author is to be found under "V" or "P", see Burkill [**149**]. These four books are frequently cited in many works of the first half of the 20th century. Let us also mention an extensive treatise Pierpont [**756**]. Some elements of Lebesgue's measure theory were discussed in Hausdorff [**412**] (in later editions this material was excluded). Some background was given in Schönflies [**858**]. Elements of Lebesgue's measure theory were considered in the book Nekrasov [**709**] published in 1907. Early surveys of Lebesgue's theory were La Vallée Poussin [**573**], Bliss [**95**], Hildebrandt [**432**], and a series of articles Borel, Zoretti, Montel, Fréchet [**115**], published in the *Encyclopedie des sciences mathématiques* (the reworked German version was edited by Rosenthal [**823**]). It is worth mentioning that in Lusin's classical monograph [**633**], the first edition of which was published in 1915 and was his magister dissertation (by a special decision of the scientific committee, the degree of Doctor was conferred on Lusin in recognition of the outstanding level of his dissertation), the fundamentals of Lebesgue's theory were assumed to be known (references were given to the books by Lebesgue and de la Vallée Poussin). The subject of Lusin's dissertation was the study of fine properties of the integral (not only the Lebesgue one, but also more general ones), the primitives and trigonometric series. Another very interesting document is the magister dissertation of G.M. Fichtenholz [**288**] (the author of the excellent calculus course [**295**]) completed in February 1918. Unfortunately, due to the well-known circumstances of the time, this remarkable handwritten manuscript was never published and was not available to the broad readership.[1] Fichtenholz's dissertation is a true masterpiece, and many of its results, still not widely known, retain an obvious interest. The manuscript contains 326 pages (the title page is posted on the website of the St.-Petersburg Mathematical Society; the library of the Department of Mechanics and Mathematics of Moscow State University has a copy of the dissertation). The introduction (pp. 1–58) gives a concise course on Lebesgue's integration. The principal original results of G.M. Fichtenholz are concerned with limit theorems for the integral and are commented on in appropriate places below (see also Bogachev [**106**]). The dissertation contains an extensive bibliography (177 titles) and a lot of comments (in addition to historical notes, there are many interesting remarks on mistakes or gaps in many classical works).

In the 1920s the following books appeared: Hahn [**398**], Kamke [**485**], van Os [**731**], Schlesinger, Plessner [**853**], Townsend [**963**]. Vitali's books [**988**], [**989**] also contain large material on Lebesgue's integration. In 1933, the first French edition of the classical book Saks [**840**] was published (the second edition was published in English in 1937); this book still remains one of the most influential reference texts in the subject. The same year was marked by publication of Kolmogorov's celebrated monograph [**532**], which built mathematical probability theory on the basis of abstract measure theory. This short book (of a booklet size), belonging to the most cited scientific works of the 20th century, strongly influenced modern measure theory and became one of the reasons for its growing popularity. Also in the 1930s, the textbooks by Titchmarsh [**947**], Haupt, Aumann [**411**] (the first edition), and Kestelman [**504**] were published. Fundamentals of Lebesgue measure and integration were given in Alexandroff, Kolmogorov [**17**]. The basic results of measure theory were presented in the book Tornier [**961**] on foundations of probability theory, which very closely followed Kolmogorov's approach (a drawback of Tornier's book is a complete omission of indications to the authorship of the presented theorems). In addition, in those years there existed lecture notes published later (e.g., von Neumann [**710**], Vitali, Sansone [**991**]). Note also the book Stone [**914**] containing material on the theory of integration. In 1941 the excellent book Natanson [**706**] was published (I.P. Natanson was Fichtenholz's student and his book was obviously influenced by the aforementioned dissertation of Fichtenholz). In McShane [**668**], the presentation of the

[1] I am most grateful to V.P. Havin, the keeper of the manuscript, for permission to make a copy, and to M.I. Gordin and A.A. Lodkin for their generous help.

theory of the integral is based on the Daniell approach, and then a standard course is given including a chapter on the Lebesgue–Stieltjes integral. Jessen's book [**465**] was composed of a series of journal expositions published in the period 1934–1947. Let us also mention Cramer's book [**190**] on mathematical statistics where a solid exposition of measure and integration was included. It should be noted that Kolmogorov's concept of foundations of probability theory lead to a deep penetration of the apparatus of general measure theory also into mathematical statistics, which is witnessed not only by Cramer's book, but also by many subsequent expositions of the theoretical foundations of mathematical statistics, see Barra [**62**], Lehmann [**600**], Schmetterer [**854**].

After World War II the number of books on measure theory considerably increased because this subject became part of the university curriculum. Below we give a reasonably complete list of such books. A very thorough presentation of measure theory and integration was given in Smirnov [**891**], the first edition of which was published in 1947. In 1950, Natanson's book [**707**] (which was a revised and enlarged version of the already-cited book [**706**]) appeared. This excellent course has become one of the most widely cited textbooks of real analysis. In addition to the standard material it offers a good deal of special topics not found in other sources. Also in 1950, Halmos's classical book [**404**] was published; since then it has become a standard reference in the subject. Three other popular textbooks from the 1950s are Riesz, Sz.-Nagy [**809**], Munroe [**705**], and Kolmogorov, Fomin [**536**]. In my opinion, the book by Kolmogorov and Fomin (it was translated in many languages and had many revised and reprinted editions) is one of the best texts on the theory of functions and functional analysis for university students. It grew from the lecture notes [**533**] on the course "Analysis-III" initiated in 1946 at the Moscow State University by Kolmogorov (he was the first lecturer; among the subsequent lecturers of the course were Fomin, Gelfand, and Shilov). At the time Kolmogorov was planning to write a book on measure theory (the projected book was even mentioned in the bibliography in [**363**], where on p. 19 "the reader is referred to that book for any explanations related to measure theory and the Lebesgue integral"). See also Kolmogorov [**534**]. However, the Halmos book was published, and Kolmogorov gave up his idea, saying, as witnessed by Yu.V. Prohorov, that "there is no desire to write worse than Halmos and no time to write better". By the way, for a similar reason, the book by Marczewski announced in 1947 in *Colloq. Math.*, v. 1, was never completed. Along with these classics of measure theory, one should mention the outstanding treatise of Doob [**231**] on stochastic processes, which became another triumph of applications of general measure theory (it is worth noting that Doob was the scientific advisor of Halmos; see also Bingham [**92**]). Two years later, in 1955, Loève's textbook [**617**] on probability theory was published; this book, a standard reference in probability theory, contains an excellent course on measure and integration. Also in the 1950s, Bourbaki's treatise [**119**] on measure theory appears in several issues. Certainly not suitable as a textbook and, in addition, rather chaotically written, Bourbaki's book offered the reader a lot of useful (and not available from other sources) information in various directions of abstract measure theory. A dozen other books on measure and integration published in the 1950s can be found in the list below. Finally, the famous monograph Dunford, Schwartz [**256**] must be mentioned. Being the most complete encyclopedia of functional analysis, it also presents in depth and detail large portions of measure theory. For the next 50 years the measure-theoretic literature has grown tremendously and it is hardly possible to mention all textbooks and monographs published in many countries and in many languages (e.g., the Russian edition of this book mentions several dozen Russian textbooks). This theory is usually presented in books under the corresponding title as well as under the titles "Real analysis", "Abstract analysis", "Analysis III", as part of functional analysis, probability theory, etc. The following list contains only the books in English, French and German with a few exceptions in Russian, Italian and Spanish (without repeating the already-cited books) that I found in the libraries of several dozen largest universities and mathematical institutes over the world (typically, every particular library possesses considerably less than a half of this list):

Adams, Guillemin [**1**], Akilov, Makarov, Havin [**6**], Aliprantis, Burkinshaw [**18**], Alt [**20**], Amann, Escher [**21**], Anger, Bauer [**25**], Arnaudies [**38**], Artémiadis [**39**], Ash [**41**], [**42**], Asplund, Bungart [**43**], Aumann [**44**], Aumann, Haupt [**45**], Barner, Flohr [**61**], de Barra [**63**], Bartle [**64**], Bass [**68**], Basu [**69**], Bauer [**70**], Bear [**72**], Behrends [**73**], Belkner, Brehmer [**74**], Bellach, Franken, Warmuth, Warmuth [**75**], Benedetto [**76**], Berberian [**78**], [**79**], Berezansky, Sheftel, Us [**80**], Bichteler [**87**], [**88**], Billingsley [**90**], Boccara [**101**], [**102**], Borovkov [**118**], Bouziad, Calbrix [**122**], Brehmer [**124**], Briane, Pagès [**128**], Bruckner, Bruckner, Thomson [**136**], Buchwalter [**139**], Burk [**146**], Burkill [**148**], Burrill [**150**], Burrill, Knudsen [**151**], Cafiero [**158**], Capiński, Kopp [**161**], Carothers [**169**], Chae [**171**], Chandrasekharan [**172**], Cheney [**175**], Choquet [**178**], Chow, Teicher [**179**], Cohn [**184**], Constantinescu, Filter, Weber [**186**], Constantinescu, Weber [**187**], Cotlar, Cignoli [**188**], Courrège [**189**], Craven [**191**], Deheuvels [**209**], DePree, Swartz [**218**], Denkowski, Migórski, Papageorgiou [**216**], Descombes [**219**], DiBenedetto [**221**], Dieudonné [**225**], Dixmier [**229**], Doob [**232**], Dorogovtsev [**234**], Dshalalow [**239**], Dudley [**251**], Durrett [**257**], D'yachenko, Ulyanov [**258**], Edgar [**260**], Eisen [**267**], Elstrodt [**268**], Federer [**282**], Fernandez [**283**], Fichera [**284**], Filter, Weber [**297**], Floret [**301**], Folland [**302**], Fonda [**304**], Foran [**305**], Fremlin [**327**], Fristedt, Gray [**329**], Galambos [**335**], Gänssler, Stute [**337**], Garnir [**344**], Garnir, De Wilde, Schmets [**345**], Gaughan [**347**], Genet [**350**], Gikhman, Skorokhod [**353**] (1st ed.), Gleason [**361**], Goffman [**366**], Goffman, Pedrick [**367**], Goldberg [**370**], Gouyon [**375**], Gramain [**377**], Grauert, Lieb [**378**], Graves [**380**], Günzler [**384**], Gut [**385**], de Guzmán, Rubio [**388**], Haaser, Sullivan [**389**], Hackenbroch [**391**], Hartman, Mikusiński [**410**], Haupt, Aumann, Pauc [**411**], Hennequin, Tortrat [**421**], Henstock

[**422**], [**424**], [**426**], Henze [**427**], Hesse [**429**], Hewitt, Stromberg [**431**], Hildebrandt [**433**], Hinderer [**435**], Hoffman [**438**], Hoffmann, Schäfke [**439**], Hoffmann-Jørgensen [**440**], Hu [**445**], Ingleton [**449**], Jacobs [**452**], Jain, Gupta [**453**], Janssen, van der Steen [**455**], Jean [**457**], Jeffery [**461**], Jiménez Pozo [**468**], Jones [**470**], Kallenberg [**484**], Kamke [**486**], Kantorovitz [**491**], Karr [**494**], Kelley, Srinivasan [**502**], Kingman, Taylor [**518**], Kirillov, Gvishiani [**519**], Klambauer [**521**], Korevaar [**541**], Kovan'ko [**544**], Kowalsky [**545**], Krée [**547**], Krieger [**548**], Kuller [**554**], Kuttler [**561**], Lahiri, Roy [**565**], Lang [**567**], [**568**], Lax [**576**], Leinert [**602**], Letta [**606**], Łojasiewicz [**618**], Lösch [**622**], Lukes, Malý [**630**], Magyar [**643**], Malliavin [**646**], Marle [**656**], Maurin [**660**], Mawhin [**661**], Mayrhofer [**662**], McDonald, Weiss [**666**], McShane [**669**], McShane, Botts [**670**], Medeiros, de Mello [**671**], Métivier [**684**], Michel [**689**], Mikusiński [**691**], Monfort [**695**], Mukherjea, Pothoven [**703**], Neveu [**713**], Nielsen [**714**], Oden, Demkowicz [**728**], Okikiolu [**729**], Pallu de la Barrière [**734**], Panchapagesan [**735**], Parthasarathy [**739**], Pedersen [**742**], Pfeffer [**747**], Phillips [**751**], Picone, Viola [**753**], Pitt [**759**], [**760**], Pollard [**764**], Poroshkin [**766**], Priestley [**770**], Pugachev, Sinitsyn [**773**], Rana [**782**], Randolph [**783**], Rao [**787**], [**788**], Ray [**789**], Revuz [**791**], Richter [**794**], Rosenthal [**825**], Rogosinski [**816**], van Rooij, Schikhof [**820**], Rotar [**827**], Roussas [**828**], Royden [**829**], Ruckle [**832**], Rudin [**835**], Sadovnichiĭ [**838**], Samuélidès, Touzillier [**843**], Sansone, Merli [**844**], Schilling [**852**], Schmitz [**855**], Schmitz, Plachky [**856**], Schwartz [**859**], Segal, Kunze [**862**], Shilov [**865**], Shilov, Gurevich [**867**], Shiryaev [**868**], Sikorski [**883**], Simonnet [**885**], Sion [**886**], Sobolev [**894**], Sohrab [**896**], Spiegel [**900**], Stein, Shakarchi [**907**], Stromberg [**916**], Stroock [**917**], Swartz [**924**], Sz.-Nagy [**926**], Taylor A.E. [**934**], Taylor J.C. [**937**], Taylor S.J. [**938**], Temple [**940**], Thielman [**942**], Tolstow [**953**], Toralballa [**958**], Torchinsky [**960**], Tortrat [**962**], Väth [**973**], Verley [**975**], Vestrup [**976**], Vinti [**977**], Vogel [**994**], Vo-Khac [**995**], Volcic [**998**], Vulikh [**1000**], Wagschal [**1002**], Weir [**1008**], [**1009**], Wheeden, Zygmund [**1012**], Widom [**1014**], Wilcox, Myers [**1019**], Williams [**1020**], Williamson [**1021**], Yeh [**1025**], Zaanen [**1042**], [**1043**], Zamansky [**1048**], Zubieta Russi [**1054**].

Chapters or sections on Lebesgue integration and related concepts (measure, measurable functions) are also found in many calculus (or mathematical analysis) textbooks, e.g., see Amerio [**23**], Beals [**71**], Browder [**133**], Fleming [**300**], Forster [**306**], Godement [**365**], Heuser [**430**], Hille [**434**], Holdgrün [**441**], James [**454**], Jost [**473**], Königsberger [**540**], Lee [**598**], Malik, Arora [**645**], Pugh [**774**], Rudin [**834**], Sprecher [**901**], Tricomi [**964**], Walter [**1004**], Vitali [**988**], or in introductory expositions of the theory of functions, e.g., Bridges [**129**], Brudno [**137**], Kripke [**549**], Lusin [**636**], Oxtoby [**733**], Rey Pastor [**792**], Richard [**793**], Saxe [**846**], Saxena, Shah [**847**]. Various interesting examples related to measure theory are considered in Gelbaum, Olmsted [**349**], Wise, Hall [**1022**]. One could extend this list by adding lecture notes from many university libraries as well as books in all other languages in which mathematical literature is published (e.g., Hungarian, Polish, and other East-European languages, the languages of some former USSR republics, Chinese, Japanese, etc.). Moreover, our list does not include books (of advanced nature) that contain extensive chapters on measure theory (such as Meyer [**686**] and others cited in this text on diverse occasions), but do not offer the background material on integration. See also a series of surveys in Pap [**738**].

The listed books cover (or almost cover) standard graduate courses, but, certainly, considerably differ in many other respects such as depth and completeness and the principles of presentation. Some (e.g., [**251**], [**268**], [**327**], [**431**], [**440**], [**452**], [**788**], [**829**], [**962**], [**1043**]), give a very solid exposition of many themes, others emphasize certain specific directions. I give no classification of the type "textbook or monograph" because in many cases it is difficult to establish a border line, but it is obvious that some of these books cannot be recommended as textbooks for students and some of them have now only a historical interest. On the other hand, even a quick glance at such books is very useful for teaching, since it helps to see the well-known from yet another side, provides new exercises etc. In particular, the acquaintance with those books definitely influenced the exposition in this book.

Many books on the list include extensive collections of exercises, but, in addition, there are books of problems and exercises that are entirely or partly devoted to measure and integration (some of them develop large portions of the theory in form of exercises): Aliprantis, Burkinshaw [**19**], Ansel, Ducel [**27**], Arino, Delode, Genet [**37**], Benoist, Salinier [**77**], Bouyssel [**121**], Capiński, Zastawniak [**162**], Dorogovtsev [**233**], Gelbaum [**348**], George [**351**], Kaczor, Nowak [**475**], Kirillov, Gvishiani [**519**], Kudryavtsev, Kutasov, Chekhov, Shabunin [**553**], Laamri [**562**], Leont'eva, Panferov, Serov [**604**], Letac [**605**], Makarov, Goluzina, Lodkin, Podkorytov [**644**], Ochan [**725**], [**727**], Telyakovskiĭ [**939**], Ulyanov, Bahvalov, D'yachenko, Kazaryan, Cifuentes [**968**], Wagschal [**1003**]. There one can find a lot of simple exercises, which are relatively not so numerous in this book. At present the theory of measure and integration (or parts of this theory) is given in courses on real analysis, measure and integration or is included in courses on functional analysis, abstract analysis, and probability theory. In recent years at the Department of Mechanics and Mathematics of the Lomonosov Moscow University there has been a one-semester course "Real analysis" in the second year of studies (approximately 28 lecture hours and the same amount of time for exercises). The curriculum of the author's course is given in the Appendix below. In addition, several related questions are studied in the course on functional analysis in the third year.

Many books cited above give bibliographical and historical comments; we especially note Anger, Portenier [**26**], Benedetto [**76**], Cafiero [**158**], Chae [**171**], Dudley [**251**], Dunford, Schwartz [**256**], Elstrodt [**268**], Hahn, Rosenthal [**402**], McDonald, Weiss [**666**], Rosenthal [**823**]. Biographies of the best-known mathematicians and recollections about them can be found in their collected works and in journal articles related to memorial dates; see also

Bingham [**91**], Bogoljubov [**109**], Demidov, Levshin [**210**], Menchoff [**681**], Paul [**740**], Phillips [**750**], Polischuk [**763**], Szymanski [**929**], Taylor [**935**], Taylor, Dugac [**936**], Tumakov [**965**], and the book [**683**]. In 1988, 232 letters from Lebesgue to Borel spanning about 20 years were discovered (Borel's part of the correspondence was not found); they are published in [**595**] with detailed comments (this typewritten work is available in the library of Université Paris–VI in Paris; large extracts are published in several issues of the more accessible journal *Revue des mathématiques de l'enseignement supérieur*, and 111 letters are published in [**596**]). Lebesgue's letters, written in a very lively style, reflect many interesting features of the scientific and university life of the time (which will still be familiar to scholars today), the ways of development of analysis of the 20th century, and the personal relations of Lebesgue with other mathematicians.

The history of the development of the theory of measure and integration at the end of the 19th century and the beginning of the 20th is sufficiently well studied. The subsequent period has not yet been adequately analyzed; there are only partial comments and remarks such as given here. Perhaps, an explanation is that an optimal time for the first serious historical analysis of any period in science comes in 50–70 years after the period to be analyzed, when, on the one hand, all available information is sufficiently fresh, and, on the other hand, a new level of knowledge and a retrospective view enable one to give a more objective analysis (in addition, influences of various mafia groups became weaker). If such an assumption is true, then the time for a deeper historical analysis of the development of measure theory up to the middle of the 20th century is coming.

Chapter 1.

§§1.1–1.8. We do not discuss here the works of predecessors of Lebesgue (Borel, Cantor, Darboux, Dini, Hankel, Harnack, Jordan, Peano, Riemann, Stieltjes, Volterra, Weierstrass, and others) that influenced considerably the developments of the theory of measure and integration; concerning this, see Medvedev [**672**]–[**677**], Michel [**688**], Pesin [**743**], [**755**], and the old encyclopedia [**823**]. At the end of the 19th century and the beginning of the 20th widely cited sources in the theory of functions were the books Dini [**228**] and Jordan [**472**].

The principal ideas of measure theory developed in this chapter are due to the French mathematician Henri Lebesgue; for this reason the theory is often called "Lebesgue's measure theory" or "Lebesgue's integration theory". A characteristic fact is that almost all the contents of the modern university course in measure and integration is covered by Lebesgue's lectures [**582**] written on the basis of his doctoral dissertation [**579**] (basic ideas were given in 1901 in [**578**]). A rare example in the history of science! To the foundation stones belong also [**584**], [**587**], [**589**], [**591**], [**593**] (see [**594**]).

As Lebesgue pointed out, his constructions had been influenced by the ideas of Borel [**111**]. Later some polemics between Lebesgue and Borel emerged on priority issues; a sufficiently objective exposition is given in survey articles by Lebesgue himself [**593**] and the historical works [**673**], [**743**], [**965**]. Note also that almost at the same time with Lebesgue, certain important ideas of his theory were developed by Vitali [**979**], [**980**], [**981**] (see also [**990**]) and Young [**1029**] (see also many reprinted papers in [**1027**]; in fact, it is hard to distinguish between the contributions of W.H. Young and those of his wife G.C. Young: see the preface in [**1027**]), but Lebesgue's contribution considerably surpassed the joint contribution of other researchers with regard to the scope and beauty of the whole theory. Lebesgue's theory was quickly and largely recognized; mathematicians in many countries started exploring the new direction and its applications, which led to the creation of big scientific schools. One of the best-known such schools was founded in Russia by N.N. Lusin (whose teacher was another brilliant Russian mathematician D.Th. Egoroff, the author of a theorem now studied in the university courses). In the text of the book and in the comments in relation with concrete results and ideas, we meet the names of many mathematicians that enriched Lebesgue's theory. Among the researchers whose works particularly influenced the theory of measure and integration in the first quarter of the 20th century one should mention G. Vitali, W. Young, J. Radon, C. Carathéodory, F. Riesz, M. Fréchet, N. Lusin, M. Souslin, Ch. de la Vallée Poussin, H. Hahn, F. Hausdorff, P. Daniell, W. Sierpiński, A. Denjoy. In the second quarter of the 20th century the development of measure theory was strongly influenced by Kolmogorov's ideas in this theory as well as in several related fields: probability theory, random processes, dynamical systems, information theory. Among other mathematicians who considerably influenced modern measure theory, essentially formed by the end of the 1950s, one should mention S. Banach, N. Wiener, A. Haar, J. von Neumann, O. Nikodym (a Polish mathematician; after World War II when being in emigration he spelled his name as O.M. Nikodým), S. Saks, A.D. Alexandroff (Aleksandrov), G. Choquet, Yu.V. Prohorov, V.A. Rohlin. In subsequent years, the progress in measure theory was connected with more special directions such as integration on topological spaces (especially infinite-dimensional), geometric measure theory, Sobolev spaces and differentiable measures, as well as with research in related fields: probability theory, dynamical systems, functional analysis, representations of groups, mathematical physics. Fascinating results have been obtained in those directions of measure theory that belong to set theory and mathematical logic. Brief comments on the corresponding results are given below. Additional information can be found in van Dalen, Monna [**196**], Hawkins [**416**], Hochkirchen [**437**], Medvedev [**673**], [**674**], [**675**], Michel [**688**], Pesin [**743**], Pier [**754**], [**755**], Tumakov [**965**].

Shortly before Lebesgue the property of additivity for volumes was studied by Peano, Jordan, Stolz, Harnack, and Cantor (see references in [**672**], [**673**], [**398**], [**755**]). Although the concept of countable additivity was already considered by Borel, the definition of measurability and extension of measure to all measurable sets became an outstanding achievement. We recall that Lebesgue's definition of measurability of a set E in an interval I was given in the form of equality $\lambda^*(E) = \lambda(I) - \lambda^*(I\backslash E)$. Borel used the following procedure: starting from intervals, by taking complements and disjoint countable unions one constructs increasing classes of sets, to which the linear measure extends in a natural way corresponding to the requirement of countable additivity. Note that the actual justification of Borel's construction, i.e., the fact that one obtains a countably additive nonnegative measure on the σ-algebra, was only given via Lebesgue's approach (though, it was shown later that a direct verification is also possible by means of transfinite induction, see, e.g., Areshkin [**30**]). The criterion of measurability of a set A in the form of equality $\lambda^*(A \cup B) = \lambda^*(A) + \lambda^*(B)$ for all B disjoint with A (Exercise 1.12.119), was given by Young [**1029**] who took for his definition a property equivalent to Lebesgue's definition: the existence, for each $\varepsilon > 0$, of an open set U containing the given set A such that the outer measure of $U\backslash A$ is less than ε. Carathéodory [**163**], [**164**] gave the definition of measurability that coincides with Young's criterion and is now called the Carathéodory measurability; he applied his definition to set functions more general than Lebesgue measure, although his first works dealt with sets in $\mathbb{R}^n$. One of early works on the Carathéodory measurability was Rosenthal [**822**]. The definition of measurability adopted in this book arose under the influence of ideas of Nikodym and Fréchet who introduced the metric space of measurable sets with the metric $d(A,B) = \mu(A \triangle B)$, which is equivalent to consideration of the space of indicator functions with the metric from $L^1(\mu)$. The first explicit use of this construction with some applications I found in the work Ważewski [**1006**] of 1923, where the author indicates that the idea is due to Nikodym; this circumstance was also mentioned in Nikodym's paper [**718**]. In Fréchet's papers [**312**], [**315**] of the same years, one finds some remarks concerning the priority issues in this respect with references to Fréchet's earlier papers (in particular, [**310**]), where he considered various metrics on the space of measurable functions, however, he did not explicitly single out the space of measurable sets with the above metric. An interesting application of this space to convergence of set functions was given by Saks [**841**] (see our §4.6). The metric d is sometimes called the Fréchet–Nikodym metric. The aforementioned idea of Nikodym was exploited by himself [**723**], as well as by Kolmogorov (e.g., in [**533**]) for defining measurable sets as we do in this book.

In the early years of development of Lebesgue's theory the subject of studies was Lebesgue measure on the real line and on $\mathbb{R}^n$, as well as more general Borel measures on $\mathbb{R}^n$; in this respect one should mention the works Lebesgue [**591**] and Radon [**778**]. However, yet another advantage of Lebesgue's approach was soon realized: the possibility of extending it to a very abstract framework. One of the first to do this was Fréchet [**308**], [**309**], [**311**], [**313**], [**314**]; it then became commonplace, so that in the 1920–30s the term "measure" applied to abstract set functions, which is clear from the works by Hahn, Nikodym, Banach, Sierpiński, Kolmogorov, and many other researchers of the time. In the same years the problems of probability theory and functional analysis led to measures on infinite-dimensional spaces (Daniell, Wiener, Kolmogorov, Jessen, P. Lévy, Ulam), see Daniell [**198**], [**199**], [**201**], [**202**], Jessen [**463**], Lévy [**610**], Łomnicki, Ulam [**619**], Wiener [**1015**], [**1017**]. A particular role was played by Kolmogorov's works [**528**] (see also [**535**]) and [**532**] laying measure theory in the foundation of probability theory. The total number of works on measures in abstract spaces is extremely large (e.g., Ridder [**795**] published a whole series of papers, only one of which is cited here), and it is not possible to analyze them here. Additional references can be found in Hahn, Rosenthal [**402**] and Medvedev [**673**].

The theorem on extension of a countably additive measure from an algebra to the generated σ-algebra (usually called the Carathéodory theorem) was obtained by Fréchet [**314**] without use of the Carathéodory construction. The fact that the latter provides a short proof of the extension theorem was soon observed; at least, Kolmogorov [**528**], [**532**] mentions it as well-known, and Hahn applies it in [**400**]. A proof by the Carathéodory method was also suggested by Hopf [**442**], [**443**], and became standard. Various questions related to extensions of measures are considered in many works; some of them are cited below in connection with measures on lattices (see also Srinivasan [**903**]). Additional references can be found in those works. In Chapter 7 (Volume 2) we discuss extensions of measures on topological spaces.

The role of the compactness property in measure theory was clear long ago. For example, for general Borel measures on $\mathbb{R}^n$, the existence of approximations by inscribed compacts was observed by Radon [**778**, p. 1309] and Carathéodory [**164**, p. 279]. A convenient and very simple abstract definition in terms of compact classes (discussed in §1.4) was given by Marczewski [**650**] in 1953. Compact classes may not consist of compact sets even in the case where one deals with topological spaces. Such examples are considered in the book, e.g., the classes of cylinders with compact bases. It does not come as a surprise that the concept of compact class entered textbooks. For a discussion of compact classes, see Pfanzagl, Pierlo [**746**].

The first Cantor-type sets were constructed by Smith [**892**] who considered compact sets of measure zero and cardinality of the continuum and compact sets of positive measure without inner points in relation to the Riemann integrability of their indicators. The fact that any open set in $\mathbb{R}^n$ up to a measure zero set is the union of a sequence of open disjoint balls was known long ago, apparently since Vitali's covering theorems (at least, it is mentioned as well-known in Wolff [**1023**]).

The first example of a nonmeasurable set was constructed by Vitali [**983**].

§1.9. Most of the widely used measure-theoretic results on σ-algebras were obtained by W. Sierpiński in the 1920–30s (see Sierpiński [**876**], [**877**], [**881**]), but later some of them were rediscovered by other mathematicians.

Since it would be technically inconvenient to call all such results "Sierpiński theorems", it is reasonable to use terms such as "monotone class theorem". Note that σ-additive classes are also called δ-systems or Dynkin systems. Certainly, whatever our terminology is, the authorship of such theorems is due to Sierpiński.

§1.10. The idea of the A-operation originated in the works of P.S. Alexandroff [**16**] and F. Hausdorff [**413**] in 1916, in which they proved the continuum hypothesis for Borel sets and employed certain representations of Borel sets that contained essential features of this operation. The explicit definition of the A-operation and its investigation was given by M.Ya. Souslin [**899**] under the supervision of N.N. Lusin. The term itself appeared later; Souslin used the term "A-set". A considerable stimulating role was played by Lebesgue's work [**583**], where, on the one hand, a number of important results were obtained, but, on the other hand, a false assertion was given that the projection of any Borel set in the plane is Borel. The analysis of this mistake turned out to be very fruitful. M. Souslin obtained, in particular, the following beautiful results: any Borel set on the real line is Souslin (an A-set in his terminology), there exist non-Borel Souslin sets, and a Souslin set is Borel precisely when its complement is Souslin as well. In addition, the Souslin sets were characterized as the projections of G_δ-sets in the plane. The measurability of Souslin sets was established by Lusin (see [**634**]), and the first published proof was given by Lusin and Sierpiński [**638**]. Szpilrajn-Marczweski [**927**] found a very general result on the stability of some properties such as measurability under the A-operation (see Exercise 6.10.60 in Chapter 6). Concerning the history of discovery of A-sets, see Bogachev, Kolesnikov [**108**], Lorentz [**620**], Tikhomirov [**945**]. W. Sierpiński who was not only an eye-witness of the first steps of this theory, but also one of its active creators, wrote: "Some authors call analytic sets Souslin; it would be more correct to call them Souslin–Lusin sets".

§§1.11, 1.12. General outer measures and the corresponding measurability introduced by Carathéodory [**164**] in the case of $\mathbb{R}^n$ and in exactly the same manner defined in the case of abstract spaces are very efficient tools in measure theory. It should be noted that the definition of outer measure (Maßfunktion) given by Carathéodory included the requirement of additivity for pairs of sets separated by a positive distance ([**164**, p. 239, Property IV]). Such outer measures on metric spaces are now called metric Carathéodory outer measures (see §7.14(x) in our Chapter 7). However, in [**164**, §238] Carathéodory considered the problem of independence of his properties and constructed an example of an outer measure (according to the present terminology) without Property IV; in addition, he constructed an example of an outer measure that is not regular. Outer measures can be generated by general set functions in a slightly different way, described in Exercise 1.12.130 (see, e.g., Poroshkin [**766**], Srinivasan [**902**]). In many textbooks abstract outer measures are introduced from the very beginning, and the measurability is defined according to Carathéodory. It appears that, for a first encounter with the subject, the order of presentation chosen here is preferable. Method I, as one can easily guess, is not a unique method of constructing outer measures. In the literature one encounters finer Methods II, III, and IV (see Munroe [**705**], Bruckner, Bruckner, Thomson [**136**] and §7.14(x)). Rinow [**811**] studied the uniqueness problem for extensions of infinite measures. In connection with outer measures, see also Pesin [**744**].

Theorem 1.12.2 was obtained (in an equivalent formulation) in Sierpiński [**877**], and the included, a slightly shorter, proof was suggested in Jayne [**456**]. Theorem 1.12.9 goes back to S. Saks, although Fréchet [**313**, Theorem 47] had already proved that, for any atomless measure μ and any $\varepsilon > 0$, there exists a finite partition of the space into sets of measure less than ε.

Regarding measure algebras in the context of the theory of Boolean algebras and related problems, see Birkhoff [**93**], Carathéodory [**165**], Dunford, Schwartz [**256**], Kappos [**492**], [**493**], Lacey [**563**], Sikorski [**882**], and Vladimirov [**993**], where there is a discussion of other links to measure theory.

Nikodym [**724**] constructed an example of a nonseparable measure on a σ-algebra in $[0, 1]$. Kodaira, Kakutani [**525**] and Kakutani, Oxtoby [**483**] constructed nonseparable extensions of Lebesgue measure.

Inner measures were considered by Lebesgue and also by Young [**1029**], La Vallée Poussin [**572**], Rosenthal [**822**], Carathéodory [**164**], and then by many other authors, in particular, Hahn [**398**], Hahn, Rosenthal [**402**], Srinivasan [**902**]. More recent works are Fremlin [**327**], Glazkov [**360**], Hoffmann-Jørgensen [**440**], Topsøe [**957**].

Measurable envelopes and measurable kernels were considered in the book Carathéodory [**164**, §§255–257]. By analogy with measurable kernels and measurable envelopes of sets, Blumberg [**96**] considered for an arbitrary function f maximal and minimal (in a certain sense) functions l and u with $l \leq f \leq u$ a.e. The fact that a measure always extends to the σ-algebra obtained by adding a single nonmeasurable set was first published apparently by Nikodym (see [**717**] and Exercise 3.10.37). However, the result had been known to Hausdorff and was contained in his unpublished note "Erweiterung des Systems der messbaren Mengen" dated 1917 (see Hausdorff [**415**, V. 4, p. 324–327]). A detailed study of this question was initiated in Łoś, Marczewski [**621**], and continued in Bierlein [**89**], Ascherl, Lehn [**40**], Lembcke [**603**], and other works.

The Besicovitch and Nikodym sets were constructed in [**83**] and [**715**], respectively; their original constructions have been simplified by many authors, but still remain rather involved. Falconer [**276**] constructed multidimensional analogs of the Nikodym set.

Bernstein's set from Example 1.12.17 is nonmeasurable with respect to every nonzero Borel measure without points of positive measure, which follows by Theorem 1.4.8.

Lemma 1.12.18 is taken from Brzuchowski, Cichoń, Grzegorek, Ryll-Nardzewski [**138**]. Theorem 1.12.19 was proved in Bukovský [**141**] and [**138**].

A number of results and examples connected with measurability are taken from the papers by Sierpiński [**881**]. In [**875**] he constructed an example of a measurable set $A \subset \mathbb{R}$ such that $A - A$ is not measurable. He raised the problem of existence of a Borel set $B \subset \mathbb{R}^1$ such that $B - B$ is not Borel. Lebesgue noted in [**593**] without proof that such a set exists. Later such examples were constructed by several authors (see Exercise 6.10.56 in Chapter 6). Sierpiński [**870**] investigated the measurability of Hamel bases; this question was also considered in Jones [**469**]. In Sierpiński [**874**] a mean value theorem for additive set functions on $\mathbb{R}^n$ was proved. The book Sierpiński [**879**] contains many measure-theoretic assertions that depend on the continuum hypothesis.

Ulam [**966**] constructed an example of an additive but not countably additive set function on the family of all subsets of $\mathbb{N}$, and Tarski [**933**] constructed a nonnegative nonzero additive set function on the family of all subsets of the real line taking values in $\{0, 1\}$ and vanishing on all finite sets.

Hausdorff [**412**, p. 451, 452] constructed an extension of any modular set function on a lattice of sets to the generated algebra. Later this result was rediscovered by several authors in connection with different problems (see, e.g., Smiley [**890**], Pettis [**745**], Kisyński [**520**], Lipecki [**615**]). A thorough discussion of the theory of set functions on lattices of sets, including extension theorems, is given in König [**539**]; see also the books Filter, Weber [**297**], Kelley, Srinivasan [**502**], Rao, Rao [**786**], and the papers Kelley [**501**], Kindler [**515**], [**516**], Rao, Rao [**785**].

Corollary 1.12.41 was proved in Banach, Kuratowski [**57**]; their method was used in Ulam [**967**] (see also comments to Chapter 3).

The problem of possible extensions of Lebesgue measure was discussed very intensively in the 1920–30s. The use of the Hahn–Banach theorem is one of the standard tools in this circle of problems; it was applied, in particular, by Banach himself (see [**49**], [**52**], [**53**]). See also Hulanicki [**446**]. Note that for $n \geq 3$, Lebesgue measure is a unique, up to a constant factor, additive measure on the sphere in $\mathbb{R}^n$ invariant with respect to rotations. The question about this was open for a long time; a positive answer was given in Margulis [**654**], Sullivan [**921**] for $n \geq 5$, and in Drinfeld [**238**] for $n = 3, 4$. On the uniqueness of invariant means, see also Rosenblatt [**821**].

The book Rogers [**813**] contains a discussion of some questions in the discrete geometry related to Lebesgue measure. In relation to Exercise 1.12.94, see also Larman [**570**]. On pavings of the space by smooth bodies, see Gruber [**382**].

In relation to Exercise 1.12.145 we note that a set E is called an Erdős set if there exists a set M of positive Lebesgue measure that has no subsets similar to E (i.e., images of E under nondegenerate affine mappings). The Erdős problem asks whether every infinite set is an Erdős set. This problem is open even for countable sequences decreasing to zero (even for the sequence $\{2^{-n}\}$). A survey on this problem is given in Svetic [**923**].

The theory of set functions was considerably influenced by the extensive treatise of A.D. Alexandroff [**13**]. Additional information about additive set functions is given in Dunford, Schwartz [**256**], Chentsov [**176**], Rao, Rao [**786**]. There are many papers on more general set functions (not necessarily additive), see, e.g., Aleksjuk [**10**], Denneberg [**217**], Drewnowski [**236**], Klimkin [**523**], de Lucia [**626**], Pap [**737**] and the references therein. Natural examples of non-additive set functions are outer measures and capacities; non-additive functions of interval were considered long ago, see Burkill [**147**].

Nonstandard analysis is applied to the theory of integral in Riečan, Neubrunn [**796**]. Measure theory from the point of view of fuzzy sets is considered in Wang, Klir [**1005**]. Ideas of the constructive mathematics applied to measure theory are discussed in Bishop [**94**], Zahn [**1044**]. For applications of measure-theoretic methods to economical models, see Faden [**275**].

There exists an extensive literature on vector measures, which we do not consider (except for the Lyapunov theorem on the range of vector measures proved in Chapter 9 as an application of nonlinear transformations of measures), see, e.g., Bichteler [**87**], Diestel, Uhl [**224**], Dinculeanu [**226**], [**227**], Dunford, Schwartz [**256**], Edwards [**262**], Kluvánek, Knowles [**524**], Kusraev, Malyugin [**560**], Sion [**887**]. Jefferies, Ricker [**460**] consider vector "polymeasures" (e.g., a bi-measure is a function $\mu(A, B)$ that is a measure in every argument).

Chapter 2.

§§2.1.–2.4. The Lebesgue integral belongs among the most important achievements in mathematics of the 20th century. The history of its creation is discussed in van Dalen, Monna [**196**], Hawkins [**416**], Hochkirchen [**437**], Medvedev [**673**], [**674**], [**675**], Michel [**688**], Pesin [**743**], Pier [**754**], [**755**], Tumakov [**965**], and other works cited above in connection with historical comments.

The original Lebesgue definition is described in §2.4 and Exercise 2.12.57. This definition was given in [**578**], and in Lebesgue's dissertation [**579**] it was given as the "analytic definition" after the "geometric definition", according to which the integral of f is the difference of the areas under the graphs of f^+ and f^- (in this spirit one can define the integral with respect to a general Carathéodory measure, see [**788**, §2.2], [**886**]). Finally, the analytic definition is the main one in [**582**]. Later Lebesgue noted other equivalent definitions of his integral. Close, in the sense of ideas, equivalent definitions are given in Exercises 2.12.56, 2.12.57, 2.12.58. The definition of the Lebesgue integral via Lusin's theorem (Exercise 2.12.61) was given, e.g., in Tonelli [**955**], Kovan'ko [**544**] (a close definition with the Riemannian integrability in place of continuity was studied in Hahn [**396**]). The approach based on monotone limits was developed by Young (see [**1028**], [**1030**], [**1031**], [**1033**], [**1036**]), Riesz (see [**803**], [**804**] and Exercise 2.12.60), and Daniell [**198**], [**199**], [**202**], whose method (later generalized by Stone) led to a new view towards the integral. The Daniell–Stone method is discussed in Chapter 7 (Volume 2) because of its connections with integration on topological spaces, although from the point of view of ideas and techniques it could have been placed in Chapter 2. Banach [**54**] considered an axiomatic approach to the integral without using measure theory by postulating the dominated convergence and monotone convergence theorems. In Exercise 2.12.59 one finds a way of introducing the integral without using a.e. convergence, applied in MacNeille [**642**], Mikusiński [**690**], [**691**]. The definition given in the text has been used by many authors; its idea goes back, apparently, to early works of F. Riesz (although Lebesgue's definition by means of his integral sums can be put into the same category). In Riesz [**801**, p. 453] the integral is defined first for a measurable function f with countably many distinct values a_j assumed on sets A_j such that the series $\sum_{j=1}^{\infty} a_j\lambda(A_j)$ converges absolutely, and the sum of the series is taken as the value of the integral. Next the integral extends to the functions that are uniform limits of sequences of functions of the described type. In textbooks, this definition with countably many valued functions was used by Kolmogorov and Fomin [**536**]. It does not involve mean convergence, but from the very beginning infinite series appear in place of finite sums. Simple functions with finitely many values are more convenient in some other respects, in particular, in order to define integral for mappings with values in more general spaces. In Dunford [**252**] such an approach was employed for defining integrals of vector-valued functions, and in Dunford, Schwartz [**256**] the definition with finitely many valued simple functions and approximation in the mean was applied also to scalar functions. The most frequently used in textbooks is the definition given by Theorem 2.5.2, for it opens the shortest way to the monotone convergence theorem and then to other basic theorems on the properties of integral. Yet, the gain is microscopic. Another advantage of such a definition is its constructibility and transparency (the original definition of Lebesgue had these advantages as well); a drawback is the necessity to consider separately nonnegative functions, so that the whole definition is in two steps. A substantial advantage of the definition in the text is its applicability to vector mappings and a clearly expressed idea of completion, its drawback is insufficient constructibility. In order to compensate this drawback we give almost immediately an equivalent definition in the form of Theorem 2.5.2 (in principle, it could have been given right after the main definition, but then the justification of equivalence would be a bit longer). At present, apart from the definitions equivalent to the Lebesgue one, there many wider concepts of integral employed in the most diverse special situations. As yet another equivalent definition, note a construction of the integral by means of the upper and lower generalized Darboux sums (see Exercise 2.12.58). Young [**1031**] defined the integral by means of the lower and upper Darboux sums corresponding to countable partitions into measurable sets. In this work, he derived the following equality for a bounded function f on a measurable set S expressing the Lebesgue integral of f as the Riemann integral of the distribution function. Let $k \le f(x) \le k'$, $I(t) := \lambda(\{f \ge t\})$, $J(t) := \lambda(\{f \le t\})$.

Then the number $\int_k^{k'} I(t)\,dt + k\lambda(S)$ equals the upper integral, and the number $k'\lambda(S) - \int_k^{k'} J(t)\,dt$ equals the lower integral. For measurable functions, both numbers equal the Lebesgue integral.

An important factor favorable for a fast dissemination of the Lebesgue integral was that it enabled one to overcome a number of difficulties that existed in the Riemann theory of integration. For example, Volterra [**999**] constructed an example of an everywhere differentiable function f on $[0,1]$ with a bounded but not Riemann integrable derivative f'. Conditions in limit theorems for the Riemann integrals were rather complicated. Finally, the reduction of multiple Riemann integrals to repeated integrals is not simple at all (see Chapter 3). Gradually, new advantages of the Lebesgue integral have become explicit. They became especially clear when Fréchet [**308**], [**309**] developed Lebesgue's theory for arbitrary general spaces with measures. In particular, this circumstance had a decisive impact on foundations of modern probability theory. An important role was played by the fact that the Stieltjes integral was included in Lebesgue's theory to the same extent as the Riemann integral. Stieltjes invented his integral in [**913**] as a tool for solving certain problems. Then this integral, generalizing the Riemann integral, was also applied by other researchers (see Medvedev [**673**, Ch. VII]), but a possibility of connecting this integral with the Lebesgue one was not immediately observed by Lebesgue. An impetus for finding such a connection was Riesz's work [**800**], where he showed that the general form of a continuous linear function on the space $C[0,1]$ is the Stieltjes integral with respect to a function of bounded variation, i.e., $l(f) = \int f(x)\,d\varphi(x)$. Due to the continuity of f, in the definition of such an integral Riemann-type sums are sufficient, and here there are no problems typical for the Lebesgue integration. However, the indicated integral in general cannot be represented in the form $\int f(x)g(x)\,dx$. For this reason the problem of including the Stieltjes integral in the new theory was not trivial at all. Lebesgue considered this problem in [**592**] and gave a rather artificial solution, which was more precisely described in [**582**, Ch. XI] (2nd ed.) and can be found in Exercise 3.10.111. In the case of multiple integrals, there is no such explicit reduction, although, as we shall see in Chapter 9, here, too, one can separate the atomic part of the measure and transform the continuous part into Lebesgue measure. It is worth noting that shortly after the invention of the Lebesgue integral it was realized (see, e.g., Young [**1031**], Van Vleck [**972**]) that, in turn, it can be expressed by means of the Stieltjes integral or even the Riemann integral (see Theorem 2.9.3), although this is not always convenient. However, further investigations showed that the Stieltjes integral can be naturally included in Lebesgue's theory; it is only necessary to develop the latter for general measures and not only for the classical Lebesgue measure. The reader will find details in Medvedev [**673**, Ch. VII]; here we mention only two works of great importance in this direction: Young [**1038**] and, particularly, Radon [**778**]. Regarding Stieltjes integral, see Carter, van Brunt [**170**], Glivenko [**362**], Gohman [**369**], Gunther [**383**], Hahubia [**505**], Kamke [**486**], Medvedev [**673**], Smirnov [**891**]. The number of articles devoted to modifications or generalizations of the Stieltjes integral is very large; see references in Medvedev [**673**].

Convergence in measure or convergence in probability, called in early works asymptotic convergence, was encountered already in the papers of Borel and Lebesgue, but a systematic treatment was given by Riesz [**799**] and Fréchet [**310**], [**316**], [**317**], and later also by other authors (see, e.g., Slutsky [**889**], Veress [**974**]). Lebesgue [**590**] filled in a gap in his book [**584**] in the justification of the assertion that a.e. convergence implies convergence in measure (the gap was mentioned in the above-cited work of Riesz); Lebesgue adds: "I felicitate myself on the fact that my works are read so thoroughly that one detects even the errors of such a character". The important theorem on a selection of an a.e. convergent subsequence from a sequence convergent in measure was discovered by Riesz [**799**], and in the special case of a sequence convergent in L^2 this theorem was obtained by Weyl [**1011**]. Note that Weyl specified the subclass of "almost uniformly" convergent sequences in the class of all a.e. convergent sequences, but shortly after him Egoroff discovered that Weyl's class coincides with the class of all a.e. convergent sequences. Fréchet and Slutsky showed that if $\xi_n \to \xi$ in measure, then $\varphi(\xi_n) \to \varphi(\xi)$ in measure for any continuous φ; Fréchet established this fact for functions φ of two variables as well. Fréchet (see [**310**], [**312**], [**315**], [**317**], [**319**], [**320**], [**321**]) considered various metrics for convergence in measure, in particular, $\inf\limits_{\varepsilon>0}\bigl\{\mu(|f-g| \geq \varepsilon)+\varepsilon\bigr\}$, and Ky Fan introduced the metric $\inf\limits_{\varepsilon>0}\bigl\{\mu(|f-g| \geq \varepsilon) \leq \varepsilon\bigr\}$. Fréchet [**310**] showed that a.e. convergence cannot be defined by a metric. For infinite measures, one can also consider convergence in measure as convergence in measure on sets of finite measure. It is clear that in the case of a σ-finite measure this convergence is defined by a suitable metric.

Lusin's theorem and Egoroff's theorem were stated without proof by Lebesgue [**580**]. Then the first of them was proved by Vitali in the paper [**982**], which, however, for some time remained unknown to many experts (the paper was in Italian, but most of mathematicians of the time could read Italian; apparently, the point was that in those days the papers of colleagues were read with the same care as now). This theorem was rediscovered by Lusin [**632**], [**631**], after which the result became widely known and very popular (by the way, Vitali in his textbook [**991**] also calls it Lusin's theorem). Before that, Egoroff [**265**] had obtained his remarkable theorem, which is now one of the standard tools in measure theory. Note that Severini [**863**] (see also [**864**]) proved an analogous assertion in some special case, dealing with convergence of orthogonal series in L^2 (almost uniform convergence was established for a subsequence of the partial sums), but he did not state the general result, although his reasoning in fact applies to it; see page 3 of the cited work. In particular, a footnote on that page contains a somewhat vague remark on applicability of the same reasoning under different assumptions: "L'ipotesi che la (5) converga si può sostituire coll'altra che sia in ogni punto di (a, b) determinata: segue infatti dalla (4) che deve allora essere convergente, fatta al più eccezione per i punti di un insieme di misura nulla", i.e., "the hypothesis that (5) converges can be substituted by another one that it be defined at every point of (a,b): in fact, it follows from (4) that it must then converge, with the exception, at most, of points of a set of measure zero". For this reason, we do not call the result the "Egoroff–Severini" theorem as some authors do. The history of discovery of Egoroff's theorem is traced by very interesting letters of Egoroff to Lusin (see Medvedev [**676**]). Let us also note that Borel [**112**] stated without proof several assertions close to the future Lusin theorem, in particular, he noted that if functions f_n on $[0, 1]$ converge pointwise to a function f and for each of them and every $\varepsilon > 0$ there exists a set of measure at least $1 - \varepsilon$ where f_n is continuous, then f has the same property. However, he came to a false conclusion that any measurable function is continuous on a set of full measure. Lebesgue's formulation from the above-cited work [**580**] is this: "Sauf pour les points d'un certain ensemble de mesure nulle, toute fonction mesurable est continue quand on néglige les ensembles de mesure ϵ, ϵ étant aussi petit que l'on veut", i.e. "with the exception of points of some set of measure zero, any measurable function is continuous if one neglects sets of measure ϵ, where ϵ is as small as we wish". In a footnote, Lebesgue mentioned that one cannot let $\epsilon = 0$, thereby correcting an erroneous formulation communicated earlier to Borel (see [**112**]). In order to pass from this a slightly vague formulation to Lusin's theorem proper one should extend a function continuous on a compact to the whole interval. Lebesgue never published a proof of his assertion and later, when Lusin's note was published, he used the term "Lusin's theorem" for this result. The situation with Egoroff's theorem is similar. Lebesgue [**580**] stated the following: "toute série convergente de fonctions mesurables est uniformément convergente quand on néglige certains ensembles de mesure ϵ, ϵ étant aussi petit que l'on veut", i.e., "any convergent series of measurable functions converges uniformly if one neglects certain sets of measure ϵ, however small is ϵ". Taking into account that Lebesgue never left unchallenged any encroachments on his priorities (which is witnessed by a lot of polemical remarks in his papers and a considerable number of special notes serving to clarify such issues), one can suppose that originally he underestimated the utility of his ideas stated in [**580**] and maybe even forgot them, but later did not find it appropriate to refer to an observation that he had not developed himself, since one cannot imagine that Lebesgue was unable to prove such assertions had he been willing do that. Further evidence is a letter of Lebesgue to Borel (see [**595**, p. 299], [**596**, p. 205]), where he writes: "I am very little aware of what, apparently, bothers you to distraction. I know very well that once, in one of December issues, there was a note of yours and a note of mine. But I have never had the texts of those notes, I never returned to that, and all that is very distant. Concerning myself, I must have indicated there a certain property of convergence, I do not know which, but immediate, and which was never useful to me. The only one that I ever *used* indeed is the fact that, given ε, for $n > N$ we have $|R_n| < \varepsilon$ at all points, with the exception of points of some set of measure $\eta(\varepsilon)$ approaching zero together with $\frac{1}{N}$. Obviously, one can transform that in many ways, but I did not do that, I am not concerned with that and saw no interest in that ... Truly, I cannot read anybody and I am not surprised that one cannot read me without being annoyed."

Sierpiński [**869**] observed that a measurable function of a continuous function is not always measurable. In [**871**] he proved the continuity of a measurable function that is convex in the sense of the inequality $f\bigl((x+y)/2\bigr) \leq f(x)/2 + f(y)/2$, which is weaker than the usual convexity.

§§2.5–2.10. The principal results in these sections belong to Lebesgue. Fatou's and B. Levi's theorems are found in [**280**] and [**607**], respectively. In the first edition of Lebesgue's lectures, the integrability of the limit function in the monotone convergence theorem was part of the hypotheses, and B. Levi observed that it follows from the uniform boundedness of the integrals of f_n. The Lebesgue dominated convergence theorem in the general case (with an integrable majorant) was given by him in [**588**]. Young's theorem 2.8.8 was later rediscovered, in particular, it was reproved in Pratt [**768**]. Theorem 2.8.9, usually called the Scheffé theorem, was discovered by Vitali [**985**] who proved that if $f_n \to f$ a.e. and $f_n \geq 0$, then a necessary and sufficient condition for the equality $\lim\limits_{n\to\infty} \int f_n\,dx = \int f\,dx$ is the uniform absolute continuity of the integrals of f_n (which, according to another Vitali theorem discussed in Chapter 4, is equivalent to mean convergence). The fact that a.e. convergence $f_n \to f$ along with convergence of the integrals of $|f_n|$ to the integral of $|f|$ yields the uniform absolute continuity of the integrals of f_n (which is equivalent to mean convergence in the case of a.e. convergence), was also proved by Young, Fichtenholz, and de la Vallée Poussin (see [**1032**], [**1034**], [**287**], [**288**], [**573**]). Hahn [**397**, p. 1774] showed that for any sequence of functions convergent in measure, mean convergence is equivalent to the uniform absolute continuity of integrals. In these works, naturally, Lebesgue measure was considered, but that played no role in the proofs. In Scheffé [**851**], Theorem 2.8.9 was rediscovered and stated for arbitrary probability measures. Such rediscoveries are sometimes useful because very few people read old works. The trivial but very useful inequality that in courses on integration is usually called Chebyshev's inequality is the simplest partial case of a somewhat less obvious inequality for sums of independent random variables that was established in the 19th century first by Bienaymé and later by Chebyshev.

Ter Horst [**941**] discusses an analog of the classical criterion of Riemann–Stieltjes integrability in terms of the discontinuity set of the integrand.

§§2.11–2.12. The Cauchy–Bunyakowsky and Hölder inequalities have a long history. They were first found for the Riemann integrals or even for finite sums. Their extensions to the case of the Lebesgue integral were straightforward and the corresponding "new" inequalities carry the old names. The Cauchy–Bunyakowsky inequality, found by Cauchy in the case of finite sums and by Bunyakowsky (in 1859) for Riemann integrals, is also called the Schwarz inequality, after G. Schwarz who derived it (for double integrals) in 1885. Jensen's inequality was obtained in [**462**]. A classical book on inequalities is Hardy, Littlewood, Polya [**408**]. For an updated survey, see Mitrinović, Pečarić, Fink [**694**]. Inequalities are also considered in §3.10(vi) and §4.7(viii).

Exercise 2.12.115 originates in Kahane [**478**, Ch. III, Theorem 5], where the case $p = 2$ is considered and the functions f_n are independent random variables (which yields a stronger conclusion: the series of f_n diverges a.e.), but the reasoning is the same as in the hint to the exercise.

Chapter 3.

§§3.1–3.2. Decompositions of finitely additive measures into positive and negative parts go back to Jordan. Fréchet [**309**] indicated that a signed countably additive measure on a σ-algebra is bounded and can be decomposed into the difference of two nonnegative measures. For measures on $\mathbb{R}^n$ the result had already been known from Radon [**778**]; the concept of the total variation was also used in Lebesgue [**591**]. Proofs were given in Fréchet [**313**], where the total variation of a signed measure was considered and its countable additivity was established. The decomposition theorem was also obtained by Hahn [**398**]. In some works signed measures are called charges, but here we do not use this term; in many papers it applies not only to countably additive functions, e.g., see Alexandroff [**13**], where this term was introduced.

An important special case of the Radon–Nikodym theorem (the absolute continuity with respect to Lebesgue measure) was found by Lebesgue, the case of Borel measures on $\mathbb{R}^n$ was considered by Radon [**778**] (and later by Daniell [**200**]), and the general result was established by Nikodym [**718**]. We gave a traditional proof of the Radon–Nikodym theorem; the alternative proof from Example 4.3.3 is due to von Neumann.

§§3.3–3.5. The theorem on reduction of multiple integrals to repeated ones for bounded Lebesgue measurable functions was established by Lebesgue himself, and the general theorem is due to Fubini [**331**]. An important complement was given by Tonelli [**954**]. Infinite products of measure spaces were considered by Daniell [**199**] (the countable power of Lebesgue measure on $[0, 1]$ and countable products of arbitrary probability distributions on the real line), Kolmogorov [**532**] (arbitrary products of probability distributions on the real line), and then in the case of a countable product of abstract probability spaces by Hopf [**442**] (who noted that the method of proof in the general case was essentially contained in Kolmogorov's work, although the latter employed compactness arguments), Kakutani [**480**], [**482**] (explicit consideration of arbitrary products of abstract probability spaces and investigation of uncountable products of compact metric spaces with measures), van Kampen [**487**], von Neumann [**710**], and other authors. Several deep results on countable products of measures were obtained by Jessen [**463**] in the case of Lebesgue measure on the unit interval, but he noted that the analogous results were also valid in the general case, and the corresponding formulations were given in Jessen, Wintner [**467**]. The statement on the existence of countable products of arbitrary probability measures is contained in Łomnicki, Ulam [**619**], but the reasoning given there is not sufficient. Uncountable products of abstract probability spaces were already considered by von Neumann in his lectures in the 1930s, but they were published only later in [**710**]. Certainly, implicitly countable products of probability measures arise in many problems of probability theory related to infinite sequences of random variables (see Borel [**113**], Cantelli [**160**]). Explicitly, such constructions in relation to measure theory were considered first in Steinhaus [**911**].

§§3.6–3.7. The change of variables formula for Lebesgue measure in the case of a smooth transformation follows at once from the corresponding theorem for the Riemann integral. More general change of variables formulas are considered in Chapter 5. Comments on Theorem 3.6.9 and its generalizations can be found in the comments to §9.9 in Volume 2.

§§3.8–3.9. Plancherel [**761**], [**762**] obtained a number of important results on the Fourier series and transforms.

An analog of Bochner's theorem for the Fourier series was obtained earlier in Herglotz [**428**], Riesz [**802**]. In addition to the theorem bearing his name, S. Bochner obtained some other results related to the Fourier transforms (see [**103**], [**104**]). F. Riesz [**806**] proved that a positive definite measurable function φ almost everywhere equals some continuous positive definite function ψ, and Crum [**193**] showed that the function $\varphi - \psi$ is positive definite as well. Concerning the Fourier transforms and characteristic functionals, see Bochner [**103**], Kawata [**499**], Lukacs [**628**], [**629**], Okikiolu [**729**], Ramachandran [**781**], Stein, Weiss [**908**], Titchmarsh [**948**], Wiener [**1016**], Wiener, Paley [**1018**].

Convolutions of probability measures are frequently used in probability theory (at least from Chebyshev's works). They are also employed in the integration on groups.

§3.10. We note that Corollary 3.10.3 was not explicitly formulated in the paper Banach, Kuratowski [**57**], where Corollary 1.12.41 was proved, but it was observed later that it follows immediately from the proof (see Banach [**55**]). In Banach's posthumous paper [**55**], the following result was established. Suppose we are given a countable collection of sets $E_n \subset X$. Then, the existence of a probability measure on $\sigma(\{E_n\})$ vanishing on all atoms of $\sigma(\{E_n\})$ (i.e., the sets in $\sigma(\{E_n\})$ that have no nontrivial subsets from $\sigma(\{E_n\})$) is equivalent to the property that the sets of values of the function $\sum_{n=1}^{\infty} I_{E_n} 3^{-n}$ is not a zero set for some Borel probability measure on $[0,1]$ without points of positive measure.

Hausdorff measures were introduced in Hausdorff [**414**]. Federer [**282**] and Rogers [**814**] give a detailed account of this theory. For various generalizations, see Rogers, Sion [**815**], Sion, Willmott [**888**].

Decompositions of additive set functions into countably additive and purely additive components were constructed in Alexandroff [**13**] and Yosida, Hewitt [**1026**]. Our §3.10(iv) describes some later generalizations.

Equimeasurable rearrangements of functions are considered in detail in Chong, Rice [**177**], Lieb, Loss [**612**], and many other books.

An interesting class of measures on $\mathbb{R}^n$ related to symmetries is discussed in the survey Misiewicz, Scheffer [**693**].

In connection with the material in §3.10(vi), see Bobkov [**97**], Bobkov, Götze [**98**], Bobkov, Ledoux [**99**], Borell [**117**], Bogachev [**105**], Brascamp, Lieb [**123**], Buldygin, Kharazishvili [**142**], Burago, Zalgaller [**143**], Dancs, Uhrin [**197**], Hadwiger [**392**], Ledoux [**597**], Leichtweis [**601**], Lieb, Loss [**612**], Pisier [**758**], and Schneider [**857**], where one can find recent results and additional references. Related questions, such as the so-called unimodal measures, are studied in Bertin, Cuculescu, Theodorescu [**82**], Dharmadhikari, Joag-Dev [**220**], Eaton [**259**].

A.D. Alexandroff [**12**] obtained important integral representations of the mixed volumes. They are based on the concept (which is of interest in its own right) of the spherical mapping of a surface defined by means of the unit normal. In addition, A.D. Alexandroff investigated certain curvature measures generated by this mapping.

The Fourier transform takes L^1 to L^∞ and L^2 to L^2. By the interpolation method one proves (see Stein, Weiss [**908**, Ch. V]) that in the case $1 \le p \le 2$ the Fourier transform on $L^1 \cap L^p$ extends to a bounded operator from L^p to L^q, where $q = p/(p-1)$. If $p \neq 2$, then this operator is not surjective, and the extension result fails for $p > 2$ (see Titchmarsh [**948**, Ch. IV]).

Chapter 4.

§§4.1–4.4. The results on the spaces L^2 and L^p traditionally included in courses on measure and integration go back to the works of Riesz [**797**], [**798**], Fréchet [**307**], and Fischer [**298**]. Complete Euclidean spaces are called Hilbert spaces in honor of D. Hilbert who considered concrete spaces of this type in his works on integral equations. First only the spaces l^2 and $L^2[a,b]$ were investigated, later abstract concepts came. Riesz and Fréchet characterized the dual spaces to l^2 or $L^2[a,b]$. The dual spaces to $L^p[a,b]$ with $p > 1$ were described by Riesz [**801**], for general measures on $\mathbb{R}^n$ that was done by Radon [**778**]. The dual to $L^1[a,b]$ was described by Steinhaus [**909**], and the case of an arbitrary bounded measure was considered by Nikodym [**719**] and later by Dunford [**253**].

It is interesting that the first proofs of the Riesz–Fischer theorem had little in common with the ones presented in modern textbooks. F. Riesz considered first the special case where an orthonormal system is the classical system $\sin nx$, $\cos nx$, and then reduced the general case (still for Lebesgue measure) to this special case. E. Fischer deduced the theorem from the completeness of $L^2[a,b]$ that was justified by using indefinite integrals, which also restricted the theorem to Lebesgue measure. It is to be noted that many arguments in the works of that time could now seem a bit strange and not efficient. However, one should not be puzzled: in those days not only were some by now classical theorems unknown, but also many standard methods had not been developed. As an example let us refer the reader to Lebesgue's letters to Fréchet published in Taylor, Dugac [**936**]. In his letters, Lebesgue suggests two different proofs of the fact that, for any Lebesgue measurable function on $[0,1]$, there exists a sequence of polynomials f_n convergent to f almost everywhere. Fréchet had already established the fact for Borel functions and discussed with Lebesgue its extension to general measurable functions. Today even the subject of discussion might seem strange, so customary is the fact that any measurable function almost everywhere equals a Borel function. At that time it was not commonplace, and Lebesgue in four letters presented two different proofs, subsequently correcting defects found in every previous letter. His first proof is this. Let a function f be integrable (e.g., bounded). Then it can be represented as the limit of an almost everywhere convergent sequence of continuous functions, which could be done either by using that $f(x) = \lim\limits_{n\to\infty} n\bigl(F(x+1/n) - F(x)\bigr)$ a.e., where F is the indefinite integral of f, or by approximating f a.e. by the sequence of its trigonometric Fejér sums (see Theorem 5.8.5), whose convergence had been earlier established by Lebesgue (he even proposed the approximation by the usual partial sums of the Fourier series, but then noted that he did not provide any justification of that). Next the general case reduces to this special one by means of the following result of Fréchet (see Exercise 2.12.33): if functions $f_{n,m}$ converge a.e. to f_n as $m \to \infty$, and the functions f_n converge a.e. to f as $n \to \infty$, then one can find subsequences n_k and m_k such that f_{n_k,m_k} converges a.e. to f (Fréchet considered Borel functions, but his proof also worked for Lebesgue measurable ones). By the Weierstrass theorem and the cited result of Fréchet, one obtains polynomial approximations. The second proof by Lebesgue was also based on the above-mentioned result of Fréchet and employed additionally the fact that any measurable function almost everywhere equals a function in the second Baire class (Lebesgue first mistakenly claimed that the first Baire class was enough). When reading Lebesgue's letters one may wonder why he did not apply the result that had already been announced in his paper [**580**] of 1903 and became later known as Lusin's theorem (which has been commented on above). It is very instructive for today's teacher that in the period of formation of measure theory certain elementary things were not obvious even to its creators.

§§4.5–4.6. The principal results about properties of uniformly integrable sequences were obtained by Lebesgue, Vitali, Young, Fichtenholz, de la Vallée Poussin, Hahn, and Nikodym. Formulations in §4.5 give a synthesis of those results.

Theorem 4.6.3, to which Vitali, Lebesgue, Hahn, Nikodym, and Saks contributed, is one of the most important in general measure theory. It is sometimes called the Vitali–Hahn–Saks theorem, which is less precise from the point of view of the history of discovery of this remarkable result. Vitali [**985**] considered the special case where the integrable functions f_n converge almost everywhere and their integrals converge over every measurable set. A very essential step is due to Lebesgue [**589**] who deduced the uniform absolute continuity of the integrals of f_n from convergence of these integrals to zero over every measurable set without assumptions on a.e. convergence. Hahn [**399**] showed that it suffices to require only the existence of a finite limit of integrals over every measurable set. Nikodym [**720**], [**721**], [**722**] proved the uniform boundedness of any sequence of measures bounded on every measurable set and established the countable additivity of the limit in the case of a setwise convergent sequence. The latter assertion was also proved independently by Saks [**841**] who obtained a slightly stronger result by the Baire category method (until then the method of a "glissing hump" was employed). Note that this assertion reduces, by the Radon–Nikodym theorem (already known at the time), to the case of functions considered by Hahn. G.M. Fichtenholz investigated integrals dependent on a parameter and obtained a number of deep results; those results were presented in his magister dissertation defended in 1918 (see his works [**286**], [**285**], [**287**], [**290**], [**294**]). In particular, as early as in 1916 G.M. Fichtenholz proved the surprising result (covering the above-mentioned result of Hahn obtained later) that for setwise convergence of the integrals of functions f_n and their uniform absolute continuity it suffices to have convergence of the integrals over every open set. This result is discussed in Chapter 8. It is mentioned in Fichtenholz's dissertation that the corresponding article was accepted for publication in 1916 (the Proceedings of the Phys. Math. Society at the Kazan University), but, apparently, the publication of scientific journals was already interrupted by World War I and the Russian revolution, and the same material was published later in [**290**]. Some new observations on convergence of measures were made by G.Ya. Areshkin [**28**], [**31**], [**32**], [**33**] and V.M. Dubrovskiĭ [**241**]–[**250**], who investigated certain properties of measures such as the uniform countable additivity and uniform absolute continuity; related properties were also considered by Caccioppoli [**155**], [**156**], and Cafiero [**158**]. The problem of taking limits under the integral sign, very important for applications, and the related properties of sequences of functions or measures were studied in many works; additional references are found in the book Cafiero [**158**]. There are many works on setwise convergence and boundedness of more general set functions, see Aleksjuk [**10**], Areshkin, Aleksjuk, Klimkin [**34**], Drewnowski [**237**], Klimkin [**523**], de Lucia, Pap [**627**]. In most of such works, the method of a "glissing hump" used by Lebesgue and Nikodym turns out to be more efficient.

§4.7. The Banach–Saks property of the spaces L^p, $1 < p < \infty$, was established in Banach, Saks [**59**]. More details are found in the very informative books Diestel [**222**] and Diestel [**223**]. In these books and in Lindenstrauss, Tzafriri [**614**], one can find results on the geometry of L^p.

Theorem 4.7.18 on weak compactness in L^1 took its modern form after the appearance of Eberlein's result on the equivalence of weak compactness and weak sequential compactness in general Banach spaces. The latter result is usually called the Eberlein–Šmulian theorem because one of the implications had been proved earlier by Šmulian, see Dunford, Schwartz [**256**], Diestel [**223**]. The fact that weak sequential compactness in L^1 is equivalent to the uniform integrability can be deduced from the above-mentioned result of Lebesgue [**589**], but explicitly it was stated by Dunford and Pettis (see [**254**], [**255**]). We note that according to the terminology of that time the term "compactness" was used for sequential compactness. Young [**1039**], [**1040**] showed that every uniformly integrable sequence of functions f_n on $[a,b]$ (in fact he required the boundedness of the integrals of $Q(f_n)$, where Q is the indefinite integral of a positive function that monotonically increases to $+\infty$) contains a subsequence of functions such that their indefinite integrals converge pointwise to the indefinite integral of some function f such that the function $Q(f)$ is integrable. We note that the characterization of weak compactness in terms of the uniform integrability can be proved without the Eberlein–Šmulian theorem, although such a proof is considerably longer (see Fremlin [**327**, §247C]). The book Diestel [**223**] gives a concise exposition of the fundamentals of the weak topology in L^1 in relation to the geometry of Banach spaces. The results on the weak compactness in L^1 find many applications outside measure theory as well (see, e.g., Barra [**62**], Lehmann [**600**]). The weak topology in L^∞ is discussed in Alekhno [**7**] and Alekhno, Zabreĭko [**8**].

Corollary 4.7.16 was proved by Radon [**778**, p. 1362, 1363] and rediscovered by Riesz [**805**].

Theorem 4.7.23 was found by V.F. Gaposhkin (see [**338**, Lemma 1.2.4], [**339**, Lemma C]) in the following equivalent formulation: there exist f_{n_k}, $g_k, \psi_k \in L^1(\mu)$ such that the functions g_k converge weakly in $L^1(\mu)$ to some function g and $\sum_{k=1}^\infty \mu(\psi_k \neq 0) < \infty$. It is clear that this implies the assertion in the text if one takes $A_k = \{\psi_k = 0\}$, and the converse follows by letting $\psi_k = I_{D_k}$, $D_k = X\backslash X_{2^{-k}}$. Later a similar result in terms of measures was obtained in Brooks, Chacon [**131**].

Additional remarks on the Komlós theorem are made in Volume 2.

The norm compactness in L^p was investigated by many authors, including Fréchet [**307**], [**318**] (the case $p = 2$), M. Riesz [**810**], Kolmogorov [**530**]; see references in Dunford, Schwartz [**256**] and Sudakov [**919**]. Theorem 4.7.29 is borrowed from Girardi [**356**], [**357**].

In connection with the last assertion of Proposition 4.7.30 obtained in Radon [**778**, p. 1363], we note that for $p = 1$ it was proved in Fichtenholz [**287**] in the following equivalent form: if a sequence of integrable (on an interval) functions f_n converges in measure to an integrable function f, then for convergence of the corresponding integrals over every measurable set it is necessary and sufficient to have the equality $\lim\limits_{n\to\infty} \|f_n\|_{L^1} = \|f\|_{L^1}$.

Hellinger's integral considered in §4.7(viii) was introduced in Hellinger [**420**] (for functions on the real line) and was actively discussed by many authors of the first half of the 20th century (see, in particular, Smirnov [**891**]); Hahn [**394**] clarified its connection to the Lebesgue integral. The assertion in Exercise 4.7.102 is found in Radon [**778**, §VIII], Kudryavtsev [**551**].

Let us mention the very general Kolmogorov integral introduced in the paper [**529**] (see also Kolmogoroff [**526**], [**527**]), which generalized, in particular, Moore, Smith [**696**]. Let $\mathfrak{R}$ be a semiring of subsets in a space X and let φ be a multivalued real function on $\mathfrak{R}$. Let us consider finite partitions $\pi = \{E_k\}$ of the space X into sets $R_k \in \mathfrak{R}$, $k \le n$, and (multivalued) sums $S(\pi) := \sum_{k=1}^{n} \varphi(E_k)$, where the multivaluedness is due to a non-unique choice of $\varphi(E_k)$. The number $I = I(\varphi)$ is called the integral of φ if, for each $\varepsilon > 0$, there exists a finite partition π_ε such that $|I - S(\pi)| < \varepsilon$ for every π that is finer than π_ε and for every possible choice of values of multivalued sums. The principal example: a single-valued set function φ_0, a real function f on X and a multivalued function $\varphi(E) := f(E)\varphi_0(E)$, $f(E) = \{f(x), x \in E\}$. Regarding Kolmogorov's integral, see Goguadze [**368**], Kolmogorov [**535**], Smirnov [**891**].

Integration with respect to additive measures that are not necessarily countably additive started to develop in the 1930s (see, e.g., the classical work Fichtenholz, Kantorovitch [**296**] and references in Dunford, Schwartz [**256**]); although this direction has many links to the usual measure theory, it is not discussed in this book.

Lebesgue [**589**] showed that his integral can be obtained as the limit of certain sums of the Riemann type. Exercise 4.7.101(ii) suggests a simple proof. Jessen [**463**, p. 275] used the martingale convergence theorem to derive a nice result that in the statement of that exercise one can always take $n_m = 2^m$ (see Example 10.3.18 in Chapter 10), and gave a different proof in [**464**]. He also raised the question on the validity of this assertion for the points $x + kn^{-1}$ in place of $x + k2^{-n}$. Marcinkiewicz, Zygmund [**649**] and Ursell [**969**] constructed counter-examples described in Exercise 4.7.101(iii). A more subtle counter-example from Exercise 4.7.101(iv) was constructed by Besicovitch [**84**] who proved that this assertion may fail even for the indicator of an open set. A similar example with a shorter justification was given by Rudin [**833**] who, apparently, was unaware of [**84**]. Close problems are considered in Akcoglu et al. [**3**], Dubins, Pitman [**240**], Fominykh [**303**], Hahn [**395**], Kahane [**477**], Marcinkiewicz, Salem [**648**], Mozzochi [**701**], Pannikov [**736**], Ross, Stromberg [**826**], Ruch, Weber [**831**].

Orlicz spaces defined in Exercise 4.7.126 generalize the spaces L^p; they are discussed in many books, e.g., in Edgar, Sucheston [**261**], Krasnosel'skiĭ, Rutickiĭ [**546**], Rao [**788**].

The theory of L^p-spaces is strongly connected with the theory of interpolation of linear operators, about which see Bergh, Löfström [**81**], Stein, Weiss [**908**].

Chapter 5.

§§5.1–5.4. Functions of bounded variation were considered in the 19th century before the invention of the Lebesgue integral, in particular, by Jordan who introduced them. Absolutely continuous functions were introduced by Vitali. In the first edition of Lebesgue's lectures his theorem on differentiation of the indefinite integral of an integrable function was given without proof in a footnote (in the text only the case of a bounded function was considered). A proof was provided by Vitali and then by Lebesgue.

Lebesgue showed (see [**581**], [**582**], [**585**], [**586**]) that if a continuous function f is of bounded variation and one of its derivates is always finite, then f is absolutely continuous. Lebesgue also proved that if f has a finite derivative at every point such that this derivative is integrable, then f is absolutely continuous (he proved an even more general assertion for one of derivates). The last two works are concerned in fact with filling in the gaps pointed out by Levi [**608**], [**609**] (who also suggested the proofs of the aforementioned facts). Large portions of [**585**], [**586**] are occupied by Lebesgue's polemics with B. Levi with respect to the critical remarks of the latter and the rigor of his arguments. Later Young and Carathéodory showed that if f is continuous and has a finite derivative everywhere with the exception of an at most countable set of points, then f is absolutely continuous provided that f' is integrable; Young [**1037**] proved an analogous assertion for the lower derivative.

Gravé [**379**] constructed examples of continuous strictly increasing functions f such that $f' = 0$ a.e.

A profound discussion of the theory of functions of a real variable is given in Benedetto [**76**], Bruckner [**135**], Bruckner, Bruckner, Thomson [**136**], Carothers [**169**], Ene [**269**], Kannan, Krueger [**488**], Natanson [**707**], van Rooij, Schikhof [**820**], Thomson [**943**].

§§5.5–5.6. Covering theorems, the most important of which was obtained by Vitali [**986**], play an important role in the theory of functions. Generalizations were obtained by Lebesgue [**591**], Besicovitch [**85**], Morse [**699**], and other authors, see the books Guzmán [**386**], Kharazishvili [**509**], Mattila [**658**], Stein [**905**], Stein [**906**], Stein, Weiss [**908**]. In these books as well as in Guzmán [**387**], Okikiolu [**729**], Torchinsky [**959**], one can find some additional information about the maximal function, singular integrals and some other related objects. A classical work on singular integrals is Calderón, Zygmund [**159**]. Interesting results on covering by parallelepipeds can be found in Keleti [**500**].

§5.7. Although we consider only the Lebesgue integral, this section gives a short introduction to the Henstock–Kurzweil integral introduced independently by Kurzweil [**557**] and Henstock [**423**] in the 1950–1960s. It turned out that the Henstock–Kurzweil integral is equivalent to the narrow Denjoy and Perron integrals introduced in 1912 and 1914, respectively. An advantage of the Henstock–Kurzweil definition is that it is entirely elementary. However, no other numerous generalizations of the Lebesgue integral and extensions of the Riemann integral are touched upon here. Among many researchers of generalized integrals one should mention Denjoy (whose work [**211**] gave rise to a flow of publications), Perron, P.S. Alexandroff, Khinchin, Hake, Looman, Burkill, Kolmogorov, Glivenko, Romanovskiĭ, Nemytskiĭ, Tolstoff, McShane, Kurzweil, and Henstock. Several interesting remarks on extensions of the integral are due to Egoroff [**266**]. There is an extensive literature on this subject of scientific or historic character; see Chelidze, Dzhvarsheishvili [**174**], Bartle [**65**], DePree, Swartz [**218**], Goguadze [**368**], Gordon [**373**], Henstock [**422**], [**424**], [**425**] (this paper contains a bibliography with 262 items), [**426**], Kestelman [**504**], Kurtz, Swartz [**556**], Kurzweil [**558**], [**559**], Leader [**577**], Lee, Výborný [**599**], Lusin [**633**], Mawhin [**661**], McLeod [**667**], Medvedev [**673**], Muldowney [**704**], Natanson [**707**], Pesin [**743**], Pfeffer [**749**], Saks [**840**], and Swartz [**925**], where additional references can be found. Romanovski [**818**] developed generalized integrals on abstract sets. Gomes [**372**], Ochan [**726**], Pfeffer [**748**], and Shilov [**866**] give a more detailed account of the Riemann approach (and Jordan's measure) than in standard textbooks of calculus. Certainly, one can study the Henstock–Kurzweil and McShane integrals before the Lebesgue integral, although this creates a perverted impression of the latter (after such courses on integration, students usually do not know any integrals at all). But a brief acquaintance with these integrals after the Lebesgue integral may be rather instructive, in spite of the fact that they are rare in applications. It should be noted that dealing with various generalizations of the Lebesgue integral one should not take too literally the claims that they include the Lebesgue integral: in fact, normally one speaks of constructions generalizing certain special cases of the Lebesgue integral (say, on the real line or on a cube). In addition, every generalization is achieved at the expense of some properties of the Lebesgue integral, but namely the aggregate of all its properties makes the Lebesgue integral so useful in applications.

§5.8. The presented proof of the Besicovitch theorem is borrowed from Evans, Gariepy [**273**]. A number of results in this section (area and coarea formulas, surface measures etc.) are typical representatives of the so-called geometric measure theory, various aspects of which are discussed in many works: David, Semmes [**205**], Edgar [**260**], Evans, Gariepy [**273**], Falconer [**277**], Federer [**282**], Ivanov [**450**], Mattila [**658**], Morgan [**697**], Preiss [**769**], Radó [**776**], Simon [**884**], Vitushkin [**992**]. Theorem 5.8.29 and the corresponding change of variables formula for Lipschitzian mappings were obtained by Federer [**281**]; for everywhere differentiable one-to-one mappings such a formula was obtained in Kudryavtsev, Kaščenko [**552**]. One of the first works in this direction was Schauder [**849**].

The differentiability of measures on $\mathbb{R}^n$ was considered first by Vitali [**986**] (he returned to this subject in [**987**]), Lebesgue [**591**], and Radon [**778**], then these studies were continued by many authors, in particular, Saks [**840**], Buseman, Feller [**153**], Jessen, Marcinkiewicz, Zygmund [**466**]. For abstract theorems on differentiation of measures and covering theorems, see Bruckner, Bruckner, Thomson [**136**], Edgar, Sucheston [**261**], Hayes, Pauc [**417**], Kölzow [**537**], Kenyon, Morse [**503**], Mejlbro, Topsøe [**678**], de Possel [**767**], Saks [**840**], Shilov, Gurevich [**867**], Thomson [**944**], Younovitch [**1041**], Zaanen [**1043**].

Denjoy [**212**], [**213**] and Khintchine [**513**], [**514**] introduced and investigated the approximate continuity and differentiability. Stepanoff [**912**] characterized the measurability as the approximate continuity.

Lusin's property (N) mentioned in this chapter is discussed in a broader context in Chapter 9. Before Lusin, this property was considered by B. Levi in [**608**] in connection with the problem of description of indefinite integrals. It should be noted that B. Levi mistakenly claimed that the sum of two functions with the property (N) has this property as well (Lebesgue constructed the counter-example given in Exercise 5.8.63) and used this claim for the proof of the absolute continuity of any continuous function f such that f possesses the property (N) and f' exists a.e. and is integrable. Later a correct proof was given by Banach and Zareckiĭ (see Exercise 5.8.51).

Appendix

Curriculum of the course “Real Analysis”

1. Rings, algebras and σ-algebras of sets; the existence of the σ-algebra generated by any class of sets. The structure of open sets on the real line. The Borel σ-algebra. §§1.1, 1.2.
2. Functions measurable with respect to a σ-algebra. Basic properties of measurable functions. §2.1.
3. Additive and countably additive measures. The property of countable subadditivity. The criterion of countable additivity. §1.3.
4. Compact classes. The countable additivity of a measure with an approximating compact class. §1.4.
5. Outer measure. The definition of a measurable set. The Lebesgue theorem on the countable additivity of the outer measure on the σ-algebra of measurable sets. The uniqueness of extension. §1.5.
6. Construction of Lebesgue measure on the real line and $\mathbf{R}^n$. Basic properties of Lebesgue measure. §1.7.
7. Almost everywhere convergence. Egoroff's theorem. §2.2.
8. Convergence in measure and its relation to almost everywhere convergence. Fundamental in measure sequences. The Riesz theorem. §2.2.
9. Lusin's theorem. §2.2.
10. The Lebesgue integral for simple functions and its properties. §2.3.
11. The general definition of the Lebesgue integral. §2.4.
12. Basic properties of the Lebesgue integral (linearity, monotonicity). The absolute continuity of the Lebesgue integral. §2.5.
13. Chebyshev's inequality. The criterion of integrability of f in terms of the sets $\{|f| \geq n\}$. §2.9.
14. The dominated convergence theorem. The monotone convergence theorem. Fatou's theorem. §2.8.
15. Connections between the Lebesgue integral and the Riemann integral (proper and improper). §2.10.
16. Hölder's inequality. Minkowski's inequality. §2.11.
17. The spaces $L^p(\mu)$ and their completeness. Connections between different modes of convergence of measurable functions. §2.7, §4.1.
18. The Radon–Nikodym theorem. §3.2.
19. Products of measure spaces. Fubini's theorem. §§3.3, 3.4.
20. Convolution of integrable functions. §3.9.
21. Functions of bounded variation. Absolutely continuous functions. The absolute continuity of the indefinite integral. Connections between absolutely continuous functions and indefinite integrals. The Newton–Leibniz formula and the integration by parts formula for absolutely continuous functions. §§5.1–5.4.

References

[1] Adams M., Guillemin V. Measure theory and probability. Birkhäuser, Boston, 1996; xvi+205 pp. [*413*][1]

[2] Adams R.A. Sobolev spaces. Academic Press, New York, 1975; 268 pp. [*379*]

[3] Akcoglu M., Bellow A., Jones R.L., Losert V., Reinhold-Larsson K., Wierdl M. *The strong sweeping out property for lacunary sequences, Riemann sums, convolution powers, and related matters.* Ergodic Theory Dynam. Systems. 1996. V. 16, №2. P. 207–253. [*435*] [2]

[4] Akhiezer (Achieser) N.I. Theory of approximation. Translated from the Russian. Frederick Ungar Publ., New York, 1956; x+307 pp. (Russian ed.: Moscow, 1947). [*261, 305*]

[5] Akhiezer N.I. The classical moment problem and some related questions in analysis. Translated from the Russian. Hafner Publ., New York, 1965; x+253 pp. (Russian ed.: Moscow, 1961). [*247*]

[6] Akilov G.P., Makarov B.M., Havin V.P. An elementary introduction to the theory of the integral. Leningrad. Gos. Univ., Leningrad, 1969; 348 pp. (in Russian). [*413*]

[7] Alekhno E.A. *Some properties of the weak topology in the space L_∞.* Vests. Nats. Akad. Navuk Belarus. Ser. Fiz.-Mat. Navuk. 2006. №3 (in Russian). [*434*]

[8] Alekhno E.A., Zabreĭko P.P. *On the weak continuity of the superposition operator in the space L_∞.* Vests. Nats. Akad. Navuk Belarus. Ser. Fiz.-Mat. Navuk. 2005. №2. P. 17–23 (in Russian). [*157, 434*]

[9] Aleksandrova D.E., Bogachev V.I., Pilipenko A.Yu. *On the convergence of induced measures in variation.* Mat. Sbornik. 1999. V. 190, N 9. P. 3–20 (in Russian); English transl.: Sb. Math. 1999. V. 190, №9–10. P. 1229–1245. [*382*]

[10] Aleksjuk V.N. Set functions. Leningrad. Gos. Pedagog. Inst., Leningrad, 1982; 78 pp. (in Russian) [*423, 433*]

[11] Alexander R. *Random compact sets related to the Kakeya problem.* Proc. Amer. Math. Soc. 1975. V. 53, №2. P. 415–419. [*66*]

[12] Alexandroff (Aleksandrov) A.D. *On the theory of mixed volumes of convex bodies.* Mat. Sbornik. 1937. V. 2, №5. P. 947–972 (in Russian); English transl. in [**15**]. [*431*]

[13] Alexandroff A.D. *Additive set functions in abstract spaces.* Mat. Sbornik. 1940. V. 8(50). P. 307–348; ibid. 1941. V. 9(51). P. 563–628; ibid. 1943. V. 13(55). P. 169–238. [*422, 429, 431*]

[14] Aleksandrov (Alexandroff) A.D. *Measure, interior and boundary.* Sibirsk. Mat. Zh. 1983. V. 24, №5. P. 12–14 (in Russian); English transl.: Sib. Math. J. 1984. V. 24, №5. P. 657–659. [*237*]

[15] Alexandrov A.D. Selected works. Part 1: Selected scientific papers. Gordon and Breach, Amsterdam, 1996; x+322 pp. [*409*]

[16] Alexandroff P. *Sur la puissance des ensembles mesurables B.* C. R. Acad. Sci. Paris. 1916. T. 162. P. 323–325. [*420*]

[17] Alexandroff P.S., Kolmogorov A.N. Introduction to the theory of functions of a real variable. Moscow, 1932; 270 pp. (3d ed.: 1938; in Russian). [*411*]

---

[1]In square brackets we indicate all page numbers where the work is cited.

[2]The article titles are printed in italics to distinguish them from the book titles.

[18] Aliprantis Ch.D., Burkinshaw O. Principles of real analysis. 3d ed. Academic Press, San Diego, California, 1998; xii+415 pp. [*413*]

[19] Aliprantis Ch.D., Burkinshaw O. Problems in real analysis. A workbook with solutions. 2nd ed. Academic Press, San Diego, California, 1999; viii+403 pp. [*415*]

[20] Alt H.W. Lineare Funktionalanalysis: eine anwendungsorientierte Einführung. 3e Aufl. Springer, Berlin, 1999; xiii+415 S. [*413*]

[21] Amann H., Escher J. Analysis III. Birkhäuser, Basel – Boston – Berlin, 2001; 480 S. [*413*]

[22] Ambrosio L., Fusco N., Pallara D. Functions of bounded variation and free discontinuity problems. Clarendon Press, Oxford University Press, New York, 2000; xviii+434 pp. [*379*]

[23] Amerio L. Analisi matematica con elementi di analisi funzionale. V. III. Methodi matematici e applicazioni. Parte I. Unione Tipografico – Editrice Torinese, Turin, 1981; viii+418 pp. [*414*]

[24] Anderson T.W. *The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities.* Proc. Amer. Math. Soc. 1955. V. 6, №2. P. 170–176. [*225*]

[25] Anger B., Bauer H. Mehrdimensionale Integration. Eine Einführung in die Lebesguesche Theorie. Walter de Gruyter, Berlin – New York, 1976; 188 S. [*413*]

[26] Anger B., Portenier C. Radon integrals. Birkhauser Boston, Boston, 1992; 332 pp. [*415*]

[27] Ansel J.-P., Ducel Y. Exercices corrigés en théorie de la mesure et de l'intégration. Ellipses, Paris, 1995; 126 pp. [*415*]

[28] Areshkin (Areškin) G.Ya. *Completely additive set functions and the Lebesgue–Radon integral.* Trudy Tbilis. Matem. Inst. A.M. Razmadze. 1946. V. 14. P. 173–213 (in Russian). [*433*]

[29] Areshkin G.Ya. *On a problem concerning the possibility of the inversion of limit and total variation in the theory of completely additive set functions.* Uspehi Mat. Nauk. 1949. V. 4, №3. P. 134–135 (in Russian). [*321*]

[30] Areshkin G.Ya. *On measure theory of B-sets.* Soobšč. Akad. Nauk Gruzin. SSR. 1949. V. 7, №8. P. 505–506 (in Russian). [*418*]

[31] Areshkin G.Ya. *On passage to the limit under the Lebesgue–Radon integral sign.* Soobšč. Akad. Nauk Gruzin. SSR. 1949. V. 10, №2. P. 69–76 (in Russian). [*321, 433*]

[32] Areshkin G.Ya.*On convergence in length of curves and on curvilinear Lebesgue integrals.* Dokl. Akad. Nauk SSSR. 1950. V. 72, №5. P. 821–824 (in Russian). [*433*]

[33] Areshkin G.Ya. *On the compactness of a family of completely additive set functions.* Učenie Zap. Leningrad. Gos. Pedagog. Inst. 1962. V. 238. P. 102–118 (in Russian). [*293, 433*]

[34] Areshkin G.Ya., Aleksjuk V.N., Klimkin V.M. *On some properties of vector measures.* Učenie Zap. Leningrad. Gos. Pedagog. Inst. 1971. V. 404. P. 298–321 (in Russian). [*293, 433*]

[35] Areshkin G.Ya., Klimkin V.M. *A generalization of Vitali's theorem on passage to the limit under the integal sign.* Učenie Zap. Leningrad. Gos. Pedagog. Inst. 1969. V. 387. P. 79–91. [*322*]

[36] Arias de Reyna J. Pointwise convergence of Fourier series. Lecture Notes in Math. V. 1785. Springer-Verlag, Berlin, 2002; xviii+175 pp. [*260*]

[37] Arino O., Delode C., Genet J. Mesure et intégration. Exercices et problèmes avec solutions. Maîtrises de mathématiques. Vuibert, Paris, 1976; 239 pp. [*415*]

[38] Arnaudies J.-M. L'intégrale de Lebesgue sur la droite. Vuibert, Paris, 2001; 320 pp. [*413*]

[39] Artémiadis N.K. Real analysis. Southern Illinois University Press, Carbondale; Feffer & Simons, London – Amsterdam, 1976; xii+581 pp. [*413*]

[40] Ascherl A., Lehn J. *Two principles for extending probability measures.* Manuscr. Math. 1977. V. 21. P. 43–50. [*59*]

[41] Ash R.B. Real analysis and probability. Academic Press, New York – London, 1972; xv+476 pp. [*413*]

[42] Ash R.B. Measure, integration, and functional analysis. Academic Press, New York – London, 1972; xiii+284 pp. [*413*]

[43] Asplund E., Bungart L. A first course in integration. Holt, Rinehart and Winston, New York – Toronto – London, 1966; xiii+489 pp. [*413*]

[44] Aumann G. Reelle Funktionen. Springer-Verlag, Berlin – Göttingen – Heidelberg, 1954; 2e Aufl., 1969; 418 S. [*413*]

[45] Aumann G., Haupt O. Einführung in die reelle Analysis. III. 3e Aufl. Walter de Gruyter, Berlin – New York, 1983; 297 S. [*413*]

[46] Baire R. Leçons sur les fonctions discontinues. Gauthier-Villars, Paris, 1904; viii+128 pp. [*148*]

[47] Baire R. Œuvres scientifiques. Gauthier-Villars, Paris, 1990. [*409*]

[48] Ball J.M., Murat F. *Remarks on Chacon's biting lemma.* Proc. Amer. Math. Soc. 1989. V. 107, №3. P. 655–663. [*316*]

[49] Banach S. *Sur le problème de la mesure.* Fund. Math. 1923. V. 4. P. 7–23. [*81, 422*]

[50] Banach S. *Sur les lignes rectifiables et les surfaces dont l'aire est finie.* Fund. Math. 1925. V. 7. P. 225–236. [*388, 406*]

[51] Banach S. *Sur une classe de fonctions continues.* Fund. Math. 1926. V. 8. P. 225–237. [*392*]

[52] Banach S. *Über additive Massfunktionen in abstrakten Mengen.* Fund. Math. 1930. V. 15. P. 97–101. [*422*]

[53] Banach S. Théorie des opérations linéaires. Warszawa, 1932; vii+254 pp. [*422*]

[54] Banach S. *The Lebesgue integral in abstract spaces.* Note II in [**840**] (2nd ed.). P. 320–330. [*424*]

[55] Banach S. *Sur les suites d'ensembles excluant l'existence d'une mesure.* Colloq. Math. 1947–1948. V. 1. P. 103–108. [*430*]

[56] Banach S. Œuvres, V. I, II. PWN-Éditions Scientifiques de Pologne, Warszawa, 1967, 1979. [*409*]

[57] Banach S., Kuratowski K. *Sur une généralisation du problème de la mesure.* Fund. Math. 1929. V. 14. P. 127–131. [*61, 211, 422, 430*]

[58] Banach S., Saks S. *Sur les fonctions absolument continues des fonctions absolument continues.* Fund. Math. 1928. V. 11. P. 113–116. [*392*]

[59] Banach S., Saks S. *Sur la convergence forte dans les champs L^p.* Studia Math. 1930. V. 2. P. 51–57. [*433*]

[60] Banach S., Tarski A. *Sur la décomposition des ensembles de points en parties respectivement congruentes.* Fund. Math. 1924. V. 6. P. 244–277. [*81*]

[61] Barner M., Flohr F. Analysis. II. 2e Aufl., Walter de Gruyter, Berlin, 1989; 449 S. [*413*]

[62] Barra J.-R. Notions fondamentales de statistique mathématique. Dunod, Paris, 1971; xix+258 pp. English transl.: Mathematical basis of statistics. Academic Press, New York – London, 1981; xvi+249 pp. [*412, 434*]

[63] Barra G. de. Measure theory and integration. Ellis Horwood, Chichester; John Wiley & Sons, New York, 1981; 239 pp. (2nd ed.: 2003). [*413*]

[64] Bartle R.G. The elements of integration and Lebesgue measure. John Wiley & Sons, New York, 1995; xii+179 pp. (1st ed.: 1966). [*413*]

[65] Bartle R.G. A modern theory of integration. Amer. Math. Soc., Providence, Rhode Island, 2001; xiv+458 pp. [*437*]

[66] Bary N.K. A treatise on trigonometric series. V. I,II. Translated from the Russian. Pergamon Press Book, Macmillan, New York, 1964; xxiii+553 pp. , xix+508 pp. (Russian ed.: Moscow, 1961). [*85, 261, 407*]

[67] Bary N., Menchoff D. *Sur l'intégrale de Lebesgue–Stieltjes et les fonctions absolument continues de fonctions absolument continues.* Ann. Mat. Pura Appl. (4). 1928. T. 5. P. 19–54. [*392*]

[68] Bass J. Cours de mathématiques. T. III: Topologie, intégration, distributions, équations intégrales, analyse harmonique. Masson et Cie, Paris, 1971; 405 pp. [*413*]

[69] Basu A.K. Measure theory and probability. Prentice-Hall of India, New Delhi, 2003; 236 pp. [*413*]

[70] Bauer H. Measure and integration theory. Walter de Gruyter, Berlin, 2001; xvi+230 pp.; German ed.: Maß- und Integrationstheorie. Walter de Gruyter, Berlin, 1992, 2e Aufl., 260 S.; [*v, 309, 413*]

[71] Beals R. Analysis. An introduction. Cambridge University Press, Cambridge, 2004; x+261 pp. [*414*]

[72] Bear H.S. A primer of Lebesgue integration. Academic Press, San Diego, California, 1995; xii+163 pp. [*413*]

[73] Behrends E. Maß- und Integrationstheorie. Springer-Verlag, Berlin – New York, 1987; 260 S. [*413*]

[74] Belkner H., Brehmer S. Lebesguesche Integrale. VEB Deutscher Verlag der Wissenschaften, Berlin, 1984; 151 S. [*413*]

[75] Bellach J., Franken P., Warmuth E., Warmuth W. Maß, Integral und bedingter Erwartungswert. Vieweg, Braunschweig, 1978; 145 S. [*413*]

[76] Benedetto J.J. Real variable and integration. With historical notes. B.G. Teubner, Stuttgart, 1976; 278 pp. [*160, 413, 415, 436*]

[77] Benoist J., Salinier A. Exercices de calcul intégral: avec rappels de cours. Masson, Paris, 1997; ix+212 pp. [*415*]

[78] Berberian S.K. Measure theory and integration. Macmillan, New York, 1965; xviii+312 pp. [*413*]

[79] Berberian S.K. Fundamentals of real analysis. Springer-Verlag, New York, 1999; xii+479 pp. [*413*]

[80] Berezansky Yu.M., Sheftel Z.G., Us G.F. Functional analysis. V. 1. Translated from the Russian. Birkhäuser, Basel – Boston – Berlin, 1996; xvi+423 pp. (Russian ed.: Kiev, 1990). [*413*]

[81] Bergh J., Löfström J. Interpolation spaces. An introduction. Springer-Verlag, Berlin – New York, 1976; x+207 pp. [*435*]

[82] Bertin E.M.J., Cuculescu I., Theodorescu R. Unimodality of probability measures. Kluwer Academic Publ., Dordrecht, 1997; xiv+251 pp. [*431*]

[83] Besicovitch A.S. *On Kakeya's problem and a similar one*. Math. Z. 1928. B. 27. S. 312–320. [*421*]

[84] Besicovitch A.S. *On a problem concerning Lebesgue integrals*. Math. Ann. 1938. B. 115. S. 613–618. [*314*]

[85] Besicovitch A.S. *A general form of the covering principle and relative differentiation of additive functions*. Proc. Cambridge Philos. Soc. 1945. V. 41. P. 103–110; 1946. V. 42. P. 1–10. [*436*]

[86] Besov O.V., Il'in V.P., Nikolskiĭ S.M. Integral representations of functions and imbedding theorems. V. I,II. Translated from the Russian. Winston & Sons, Washington; Halsted Press, New York – Toronto – London, 1978, 1979; viii+345 pp., viii+311 pp. (Russian ed.: Moscow, 1975). [*379*]

[87] Bichteler K. Integration theory (with special attention to vector measures). Lecture Notes in Math. V. 315. Springer-Verlag, Berlin – New York, 1973; vi+357 pp. [*413, 423*]

[88] Bichteler K. Integration – a functional approach. Birkhäuser, Boston – Berlin, 1998; viii+193 pp. [*413*]

[89] Bierlein D. *Über die Fortsetzung von Wahrscheinlichkeitsfeldern*. Z. Wahrscheinlichkeitstheorie verw. Geb. 1962. B. 1. S. 28–46. [*59, 421*]

[90] Billingsley P. Probability and measure. 3d ed. Wiley, New York, 1995; 593 pp. [*413*]

[91] Bingham N.H. *Studies in the history of probability and statistics XLVI. Measure into probability: from Lebesgue to Kolmogorov*. Biometrika. 2000. V. 87, №1. P. 145–156. [*416*]

[92] Bingham N.H. *Doob: a half-century on.* J. Appl. Probab. 2005. V. 42, №1. P. 257–266. [*412*]

[93] Birkhoff G. Lattice theory. Corrected reprint of the 1967 third edition. Amer. Math. Soc., Providence, Rhode Island, 1979; vi+418 pp. [*421*]

[94] Bishop E. Foundations of constructive analysis. McGraw-Hill Book Co., New York – Toronto, 1967; xiii+370 pp. [*423*]

[95] Bliss G.A. *Integrals of Lebesgue.* Bull. Amer. Math. Soc. 1917. V. 24. P. 1–47. [*410*]

[96] Blumberg H. *The measurable boundaries of an arbitrary function.* Acta Math. 1935. V. 65. P. 263–282. [*421*]

[97] Bobkov S.G. *Isoperimetric and analytic inequalities for log-concave probability measures.* Ann. Probab. 1999. V. 27, №4. P. 1903–1921. [*431*]

[98] Bobkov S.G., Götze F. *Exponential integrability and transportation cost related to logarithmic Sobolev inequalities.* J. Funct. Anal. 1999. V. 163, №1. P. 1–28. [*431*]

[99] Bobkov S.G., Ledoux M. *From Brunn-Minkowski to Brascamp–Lieb and to logarithmic Sobolev inequalities.* Geom. Funct. Anal. 2000. V. 10, №5. P. 1028–1052. [*431*]

[100] Bobynin M.N. *On a theorem of the theory of completely additive set functions.* Uspehi Mat. Nauk. 1952. V. 7, №3. P. 113–120 (in Russian). [*324*]

[101] Boccara N. Functional analysis. An introduction for physicists. Academic Press, Boston, 1990; xiv+327 pp. [*413*]

[102] Boccara N. Intégration. Ellipses, Paris, 1995; 120 pp. [*413*]

[103] Bochner S. Vorlesungen über Fouriersche Integrale. Leipzig, 1932. English transl.: Lectures on Fourier integrals: with an author's supplement on monotonic functions, Stieltjes integrals, and harmonic analysis. Princeton University Press, Princeton, 1959; 333 pp. [*430*]

[104] Bochner S. *Monotone Funktionen, Stieltjessche Integrale und harmonische Analyse.* Math. Ann. 1933. B. 108. S. 378–410 (see [**103**]). [*430*]

[105] Bogachev V.I. Gaussian measures, Amer. Math. Soc., Providence, Rhode Island, 1998; 433 pp. [*198, 431*]

[106] Bogachev V.I. *On G.M. Fichtenholz's works in the theory of the integral.* Istor.–Matem. Issled. Ser. 2. №9(44). 2005 (in Russian). [*411*]

[107] Bogachev V.I., Kolesnikov A.V. *On the Monge–Ampère equation in infinite dimensions.* Infinite Dimen. Anal. Quantum Probab. 2005. V. 8, №4. P. 547–572. [*408*]

[108] Bogachev V.I., Kolesnikov A.V. *Towards the origins of the theory of A-sets: remarks in relation with the Russian translation of Hausdorff's work.* Istor.–Matem. Issled. Ser. 2. №10(45). 2005 (in Russian). [*420*]

[109] Bogoljubov A.N. Mathematicians. Mechanicists. (Biographical reference book). Naukova Dumka, Kiev, 1983 (in Russian). [*416*]

[110] Boman J. *An example of nonuniqueness for a generalized Radon transform.* J. Anal. Math. 1993. V. 61. P. 395–401. [*228*]

[111] Borel E. Leçons sur la théorie des fonctions. Gauthier-Villars, Paris, 1898 (4e éd.: 1950; xii+295 pp.) [*417*]

[112] Borel E. *Un théorème sur les ensembles mesurables.* C. R. Acad. Sci. Paris. 1903. T. 137. P. 966–967. [*427*]

[113] Borel E. *Les probabilités dénombrables et leurs applications arithmétiques.* Rend. Circ. Mat. Palermo. 1909. T. 27. P. 247–271. [*430*]

[114] Borel E. Œuvres de Émile Borel. T. I–IV. Éditions du Centre National de la Recherche Scientifique, Paris, 1972. [*409*]

[115] Borel E., Zoretti L., Montel P., Fréchet M. *Recherches contemporaines sur la théorie des fonctions.* Encyclopédie des sciences mathématiques pures et appliquées. Tome II, V. 1. P. 113–241. Gauthier-Villars, B.G. Teubner, Paris, Leipzig, 1912. [*410*]

[116] Borell C. *Convex measures on locally convex spaces.* Ark. Math. 1974. V. 12, №2. P. 239–252. [*226*]

[117] Borell C. *Convex set functions in d-space.* Period. Math. Hungar. 1975. V. 6, №2. P. 111–136. [*431*]

[118] Borovkov A.A. Probability theory. Translated from the Russian. Gordon and Breach, Amsterdam, 1998; x+474 pp. (Russian ed.: Moscow, 1986). [*413*]

[119] Bourbaki N. Intégration. Ch. I–IV, V, VI, VII–VIII, IX. Hermann et Cie, Paris, 1952, 1956, 1959, 1963, 1969; ii+237+v pp. , ii+131 pp. , 108 pp. , 222 pp. , 133 pp. (English transl. of Ch. I-VI: Springer, 2004). [*412*]

[120] Bourgain J. *An averaging result for l^1-sequences and applications to weakly conditionally compact sets in L^1_X.* Israel J. Math. 1979. V. 32, №4. P. 289–298. [*316*]

[121] Bouyssel M. Intégrale de Lebesgue. Mesure et intégration. Exercices avec solutions et rappels de cours. Cépaduès-éditions, Toulouse, 1997; 383 pp. [*415*]

[122] Bouziad A., Calbrix J. Théorie de la mesure et de l'intégration. Publ. de l'Univ. de Rouen, Mont–Saint-Aignan, 1993; 254 pp. [*413*]

[123] Brascamp H., Lieb E.H. *On extensions of the Brunn–Minkowski and Prékopa–Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation.* J. Funct. Anal. 1976. V. 22. P. 366–389. [*431*]

[124] Brehmer S. Einführung in die Maßtheorie. Akademie-Verlag, Berlin, 1975; 186 S. [*413*]

[125] Brenier Y. *Polar factorization and monotone rearrangement of vector–valued functions.* Comm. Pure Appl. Math. 1991. V. 44. P. 375–417. [*382*]

[126] Brézis H. *How to recognize constant functions. Connections with Sobolev spaces.* Russian Math. Surveys. 2002. V. 54, №4. P. 693–708. [*248*]

[127] Brézis H., Lieb E.H. *A relation between pointwise convergence of functions and convergence of functionals.* Proc. Amer. Math. Soc. 1983. V. 88. P. 486–490. [*298*]

[128] Briane M., Pagès G. Théorie de l'intégration. 2e éd. Vuibert, Paris, 2000; 304 pp. [*413*]

[129] Bridges D.S. Foundations of real and abstract analysis. Springer-Verlag, New York, 1998; xiv+322 pp. [*414*]

[130] Brodskiĭ M.L. *On some properties of sets of positive measure.* Uspehi Matem. Nauk (N.S.). 1949. V. 4, №3. P. 136–138 (in Russian). [*235, 408*]

[131] Brooks J.K., Chacon R.V. *Continuity and compactness of measures.* Adv. Math. 1980. V. 37, №1. P. 16–26. [*434*]

[132] Broughton A., Huff B.W. *A comment on unions of sigma-fields.* Amer. Math. Monthly. 1977. V. 84, №7. P. 553–554. [*84*]

[133] Browder A. Mathematical analysis. An introduction. Springer-Verlag, New York, 1996; xiv+333 pp. [*414*]

[134] Brown A.B., Freilich G. *A condition for existence of a smallest Borel algebra containing a given collection of sets.* Enseign. Math. (2). 1967. V. 13. P. 107–109. [*84*]

[135] Bruckner A.M. Differentiation of real functions. Springer-Verlag, Berlin – Heidelberg, 1978; 246 pp. (2nd ed.: Amer. Math. Soc., Providence, Rhode Island, 1994; 195 pp.) [*332, 395, 401, 402, 436*]

[136] Bruckner A.M., Bruckner J.B., Thomson B.S. Real analysis. Prentice-Hall, 1997; 713 pp. [*210, 413, 421, 436, 438*]

[137] Brudno A.L. The theory of functions of a real variable. Nauka, Moscow, 1971; 120 pp. (in Russian). [*414*]

[138] Brzuchowski J., Cichoń J., Grzegorek E., Ryll-Nardzewski C. *On the existence of nonmeasurable unions.* Bull. Acad. Polon. Sci. 1979. V. 27, №6. P. 447–448. [*421*]

[139] Buchwalter H. Le calcul intégral en licence de mathématiques. Ellipses, Paris, 1991; iv+145 pp. [*413*]

[140] Buczolich Z. *Birkhoff sums and zero-one law.* Real Anal. Exchange. Denton Symp. 2000. P. 13–21. [*172*]

[141] Bukovský L. *Any partition into Lebesgue measure zero sets produces a nonmeasurable set.* Bull. Acad. Polon. Sci. 1979. V. 27, №6. P. 431–435. [*421*]

[142] Buldygin V.V., Kharazishvili A.B. Geometric aspects of probability theory and mathematical statistics. Kluwer Academic Publ., Dordrecht, 2000; x+303 pp. (translated from the 1985 Russian original: The Brunn-Minkowski inequality and its applications, Naukova Dumka, Kiev, 1985, and revised by the authors). [*80, 431*]

[143] Burago D.M., Zalgaller V.A. Geometric inequalities. Translated from the Russian. Springer-Verlag, Berlin, 1988; xiv+331 pp. (Russian ed.: Leningrad, 1980). [*227, 379, 431*]

[144] Burenkov V.I. *On composition of absolutely continuous functions and compositions of functions of bounded variation*. Dokl. Akad. Nauk. SSSR. 1969. V. 189, №2. P. 234–236 (in Russian); English transl.: Soviet Math. Dokl. 1969. V. 10, №2. P. 1379–1381. [*391*]

[145] Burenkov V.I. *On integration by parts and a problem on composition of absolutely continuous functions which arises in this connection*. Trudy Mat. Inst. Steklov. 1975. V. 134. P. 38–46 (in Russian); English transl.: Proc. Steklov Inst. Math. V. 134. P. 45–54. Amer. Math. Soc., Providence. Rhode Island, 1977. [*391*]

[146] Burk F. Lebesgue measure and integration. An introduction. John Wiley & Sons, New York, 1998; xvi+292 pp. [*413*]

[147] Burkill J.C. *Functions of intervals*. Proc. London Math. Soc. 1924. V. 22. P. 275–310. [*423*]

[148] Burkill J.C. The Lebesgue integral. Cambridge University Press, Cambridge, 1951; viii+87 pp. [*413*]

[149] Burkill J.C. *Charles-Joseph de la Vallée Poussin*. J. London Math. Soc. 1964. V. 39. P. 165–175. [*410*]

[150] Burrill C.W. Measure, integration, and probability. McGraw-Hill Book Co., New York, 1972; xiv+464 pp. [*413*]

[151] Burrill C.W., Knudsen J.R. Real variables. Holt, Rinehart and Winston, New York – Montreal – London, 1969; xii+419 pp. [*413*]

[152] Burstin C. *Über eine spezielle Klasse reeller periodischer Funktionen*. Monatsh. Math. Phys. 1915. B. 26. S. 229–262. [*400*]

[153] Buseman H., Feller W. *Zur Differentiation des Lebesgueschen Integrale*. Fund. Math. 1934. V. 22. P. 226–256. [*437*]

[154] Busemann H., Petty C.M. *Problems on convex bodies*. Math. Scand. 1956. V. 4. P. 88–94. [*215*]

[155] Caccioppoli R. *Sull'integrazione delle funzioni discontinue*. Rend. Circ. Mat. Palermo. 1928. T. 52. P. 1–29. [*433*]

[156] Caccioppoli R. *Integrali impropri di Stieltjes. Estensione del teorema di Vitali*. Rend. Acc. Sc. Fis. Mat. Napoli (3). 1929. T. 35. P. 34–40. [*433*]

[157] Caffarelli L. *Monotonicity properties of optimal transportations and the FKG and related inequalities*. Comm. Math. Phys. 2000. V. 214, №3. P. 547–563. [*382*]

[158] Cafiero F. Misura e integrazione. Edizioni Cremonese, Roma, 1959; 451 pp. [*413, 415, 433*]

[159] Calderón A.P., Zygmund A. *On the existence of certain singular integrals*. Acta Math. 1952. V. 88. P. 85–139. [*436*]

[160] Cantelli F.P. *La tendenza ad un limite nel senso del calcolo delle probabilità*. Rend. Circ. Mat. Palermo. 1916. T. 41. P. 191–201. [*430*]

[161] Capiński M., Kopp E. Measure, integral and probability. 2nd ed. Springer-Verlag London, London, 2004; xvi+311 pp. [*413*]

[162] Capiński M., Zastawniak T. Probability through problems. Springer-Verlag, New York, 2001; viii+257 pp. [*415*]

[163] Carathéodory C. *Über das lineare Maß von Punktmengen – eine Verallgemeinerung des Längenbegriffs*. Nachr. Ges. Wiss. Göttingen, Math.-phys. Kl. 1914. S. 404–426. [*418*]

[164] Carathéodory C. Vorlesungen über reelle Funktionen. B.G. Teubner, Leipzig – Berlin, 1918; 704 S. (2e Aufl. 1927; 718 S.) [*100, 410, 418, 419, 420*]

[165] Carathéodory C. Maß und Integral und ihre Algebraisierung. Basel, 1956; English transl.: Algebraic theory of measure and integration. 2nd ed. Chelsea, New York, 1986; 378 pp. [*421*]

[166] Carathéodory C. Gesammelte mathematische Schriften. C.H. Beck, München, 1954. [*409*]

[167] Carlen E., Loss M. *Logarithmic Sobolev inequalities and spectral gaps*. Recent advances in the theory and applications of mass transport, pp. 53–60. Contemp. Math., 353, Amer. Math. Soc., Providence, Rhode Island, 2004. [*325*]

[168] Carlson T. *Extending Lebesgue measure by infinitely many sets*. Pacif. J. Math. 1984. V. 115, №1. P. 33–45. [*61*]

[169] Carothers N.L. Real analysis. Cambridge University Press, Cambridge, 2000; xiv+401 pp. [*413, 436*]

[170] Carter M., Brunt B. van. The Lebesgue–Stieltjes integral. A practical introduction. Springer-Verlag, New York, 2000; x+228 pp. [*425*]

[171] Chae S.B. Lebesgue integration. 2nd ed. Springer-Verlag, New York, 1995; xiv+264 pp. [*413, 415*]

[172] Chandrasekharan K. A course on integration theory. Hindustan Book Agency, New Delhi, 1996; viii+118 pp. [*413*]

[173] Chavel I. Isoperimetric inequalities. Differential geometric and analytic perspectives. Cambridge University Press, Cambridge, 2001; xii+268 pp. [*379*]

[174] Chelidze V.G., Dzhvarsheishvili A.G. The theory of the Denjoy integral and some applications. Translated from the Russian. World Sci., Singapore, 1989; xiii+322 pp. (Russian ed.: Tbilis. Univ., Tbilisi, 1978). [*437*]

[175] Cheney W. Analysis for applied mathematics. Springer-Verlag, New York, 2001; viii+444 pp. [*413*]

[176] Chentsov A.G. Finitely additive measures and relaxation of extremal problems. Translated from the Russian. Consultants Bureau, New York, 1996; xii+244 pp. (Russian ed.: Ekaterinburg, 1993). [*423*]

[177] Chong K.M., Rice N.M. Equimeasurable rearrangements of functions. Queen's Papers in Pure and Applied Mathematics, No. 28. Queen's University, Kingston, Ontario, 1971; vi+177 pp. [*431*]

[178] Choquet G. Lectures on analysis. V. I: Integration and topological vector spaces. W.A. Benjamin, New York – Amsterdam, 1969; xx+360+xxi pp. [*413*]

[179] Chow Y.S., Teicher H. Probability theory. Independence, interchangeability, martingales. 3d ed. Springer-Verlag, New York, 1997; xxii+488 pp. [*413*]

[180] Ciesielski K. *Isometrically invariant extensions of Lebesgue measure*. Proc. Amer. Math. Soc. 1990. V. 110. P. 799–801. [*81*]

[181] Ciesielski K., Fejzić H., Freiling C. *Measure zero sets with non-measurable sum*. Real Anal. Exchange. 2001–2002. V. 27, №2. P. 783–793. [*87*]

[182] Ciesielski K., Pelc A. *Extensions of invariant measures on Euclidean spaces*. Fund. Math. 1985. V. 125. P. 1–10. [*81*]

[183] Clarkson J.A. *Uniformly convex spaces*. Trans. Amer. Math. Soc. 1936. V. 40, №3. P. 396–414. [*325*]

[184] Cohn D.L. Measure theory. Birhäuser, Boston – Basel – Stuttgart, 1980; ix+373 pp. [*413*]

[185] Coifman R.R., Feffermann C. *Weighted norm inequalities for maximal functions and singular integrals*. Studia Math. 1974. V. 51. P. 241–250. [*375*]

[186] Constantinescu C., Filter W., Weber K. Advanced integration theory. With the collaboration of Alexia Sontag. Kluwer Academic Publ., Dordrecht, 1998; x+861 pp. [*413*]

[187] Constantinescu C., Weber K. Integration theory. V. I. Measure and integral. In collaboration with Alexia Sontag. John Wiley & Sons, New York, 1985; xiii+520 pp. [*413*]

[188] Cotlar M., Cignoli R. An introduction to functional analysis. North-Holland, Amsterdam–London; American Elsevier, New York, 1974; xiv+585 pp. [*413*]

[189] Courrège P. Théorie de la mesure. Les cours de Sorbonne. 2e éd. Centre de Documentation Universitaire, Paris, 1965; 233 pp. [*413*]

[190] Cramer H. Mathematical methods of statistics. Princeton University Press, Princeton, 1946; xvi+575 pp. [*412*]

[191] Craven B.D. Lebesgue measure & integral. Pitman, Boston – London, 1982; ix+221 pp. [*413*]

[192] Crittenden R.B., Swanson L.G. *An elementary proof that the unit disc is a Swiss cheese*. Amer. Math. Monthly. 1976. V. 83, №7. P. 552–554. [*91*]

[193] Crum M.M. *On positive definite functions*. Proc. London Math. Soc. (3). 1956. V. 6. P. 548–560. [*430*]

[194] Csiszár I. *Information-type measures of difference of probability distributions and indirect observations*. Studia Sci. Math. Hung. 1967. V. 2. P. 299–318. [*155*]

[195] Csőrnyei M. *How to make Davies' theorem visible*. Bull. London Math. Soc. 2001. V. 33, №1. P. 59–66. [*234*]

[196] Dalen D. van, Monna A.F. Sets and integration: an outline of the development. Wolters – Hoordhoff, Groningen, 1972; 162 pp. [*417, 423*]

[197] Dancs S., Uhrin B. *On a class of integral inequalities and their measure–theoretic consequences*. J. Math. Anal. Appl. 1980. V. 74, №2. P. 388–400. [*431*]

[198] Daniell P.J. *A general form of integral*. Ann. Math. (2). 1917–1918. V. 19. P. 279–294. [*419, 423*]

[199] Daniell P.J. *Integrals in an infinite number of dimensions*. Ann. Math. (2). 1918. V. 20. P. 281–288. [*419, 423, 429*]

[200] Daniell P.J. *Stieltjes derivatives*. Bull. Amer. Math. Soc. 1919. V. 26. P. 444–448. [*429*]

[201] Daniell P.J. *Functions of limited variation in an infinite number of dimensions*. Ann. Math. (2). 1919–1920. V. 21. P. 30–38. [*419*]

[202] Daniell P.J. *Further properties of the general integral*. Ann. Math. (2). 1920. V. 21. P. 203–220. [*419, 423*]

[203] Darji U.B., Evans M.J. *A first return examination of the Lebesgue integral*. Real Anal. Exch. 2001–2002. V. 27, №2. P. 573–582. [*103, 164*]

[204] Darst R.B. *C^∞-functions need not be bimeasurable*. Proc. Amer. Math. Soc. 1971. V. 27. P. 128–132. [*243*]

[205] David G., Semmes S. Analysis of and on uniformly rectifiable sets. Amer. Math. Soc., Providence, Rhode Island, 1993; xii+356 pp. [*437*]

[206] Davies R.O. *On accessibility of plane sets and differentiation of functions of two real variables*. Proc. Cambridge Philos. Soc. 1952. V. 48. P. 215–232. [*234*]

[207] Davies R.O. *Separate approximate continuity implies measurability*. Math. Proc. Camb. Phil. Soc. 1973. V. 73. P. 461–465. [*156, 405*]

[208] Davies R.O. *Two remarks on the measure of product sets*. Classical real analysis, pp. 45–48, Contemp. Math., 42, Amer. Math. Soc., Providence, Rhode Island, 1985. [*235*]

[209] Deheuvels P. L'intégrale. Presses Universitaires de France, Paris, 1980; 229 pp. [*413*]

[210] Demidov S.S., Levshin B.V., eds. The case of Academician Nikolai Nikolaevich Luzin. Russkiĭ Khristianskiĭ Gumanitarnyĭ Institut, St. Petersburg, 1999; 312 pp. (in Russian). [*416*]

[211] Denjoy A. *Une extension de l'intégrale de M. Lebesgue*. C. R. Acad. Sci. Paris. 1912. T. 154. P. 859–862. [*437*]

[212] Denjoy A. *Mémoire sur les nombres derivées des fonctions continues*. J. Math. Pures Appl. (7). 1915. V. 1. P. 105–240. [*438*]

[213] Denjoy A. *Sur les fonctions derivées sommables*. Bull. Soc. Math. France. 1915. V. 43. P. 161–248. [*402, 438*]

[214] Denjoy A. *Sur une propriété des fonctions derivées.* Enseign. Math. 1916. T. 18. P. 320–328. [*404*]

[215] Denjoy A. Articles et mémoires. II. Le champ réel. Notes. II. Le champ réel. Gauthier-Villars, Paris, 1955, 1957. [*409*]

[216] Denkowski Z., Migórski S., Papageorgiou N.S. An introduction to nonlinear analysis: theory. Kluwer Academic Publ., Boston, 2003; xvi+689 pp. [*413*]

[217] Denneberg D. Non-additive measure and integral. Kluwer Academic Publ., Dordrecht, 1994; x+178 pp. [*423*]

[218] DePree J., Swartz C. Introduction to real analysis. John Wiley & Sons, New York, 1988; xii+355 pp. [*413, 437*]

[219] Descombes R. Intégration. Hermann, Paris, 1972; 208 pp. [*413*]

[220] Dharmadhikari S., Joag-Dev K. Unimodality, convexity, and applications. Academic Press, Boston, 1988; xiv+278 pp. [*431*]

[221] DiBenedetto E. Real analysis. Birkhäuser Boston, Boston, 2002; xxiv+485 pp. [*413*]

[222] Diestel J. Geometry of Banach spaces – selected topics. Lecture Notes in Math. V. 485. Springer-Verlag, Berlin – New York, 1975; xi+282 pp. [*282, 285, 319, 433*]

[223] Diestel J. Sequences and series in Banach spaces. Springer, Berlin – New York, 1984; xi+261 pp. [*433, 434*]

[224] Diestel J., Uhl J.J. Vector measures. Amer. Math. Soc., Providence, 1977; xiii+322 pp. [*423*]

[225] Dieudonné J. Éléments d'analyse. Tome II: Chapitres XII à XV. Gauthier-Villars, Paris, 1968; x+408 pp. [*413*]

[226] Dinculeanu N. Vector measures. Pergamon Press, Oxford – New York – Toronto, VEB Deutscher Verlag der Wissenschaften, Berlin, 1967; x+432 pp. [*423*]

[227] Dinculeanu N. Vector integration and stochastic integration in Banach spaces. Wiley–Interscience, New York, 2000; xvi+424 pp. [*423*]

[228] Dini U. Fondamenti per la teorica delle funzioni di variabili reali. Pisa, 1878; 409 pp. German transl.: Grundlagen für eine Theorie der Functionen einer veränderlichen reellen Grösse. Teubner, Leipzig, 1892; xviii+554 S. [*416*]

[229] Dixmier J. L'intégrale de Lebesgue. "Les cours de Sorbonne", Mathématique I. Centre de Documentation Universitaire, Paris, 1964; 93 pp. [*413*]

[230] Dolženko E.P. *Boundary properties of arbitrary functions.* Izvestia AN SSSR. 1967. V. 31, №1. P. 3–14 (in Russian); English transl.: Math. USSR Izv. 1967. V. 1. P. 1–12. [*403*]

[231] Doob J.L. Stochastic processes. John Wiley & Sons, New York, 1953; viii+654 pp. [*412*]

[232] Doob J.L. Measure theory. Springer-Verlag, New York, 1994; xii+210 pp. [*413*]

[233] Dorogovtsev A.Ya. Mathematical analysis. Collection of exercises. Visša Škola, Kiev, 1987; 408 pp. (in Russian). [*415*]

[234] Dorogovtsev A.Ya. Elements of the general theory of measure and integral. Visša Škola, Kiev, 1989; 152 pp. (in Russian). [*413*]

[235] Douglas R.G. *On extremal measures and subspace density.* Michigan Math. J. 1964. V. 11. P. 243–246. [*325*]

[236] Drewnowski L. *Topological rings of sets, continuous set functions. Integration I, II.* Bull. Acad. Polon. Sci. 1972. V. 20. P. 269–286. [*423*]

[237] Drewnowski L. *Equivalence of Brooks–Jewett, Vitali–Hahn–Saks, and Nikodym theorems.* Bull. Acad. Polon. Sci. 1972. V. 20. P. 725–731. [*319, 433*]

[238] Drinfeld V.G. *Finitely-additive measures on S^2 and S^3, invariant with respect to rotations.* Funkt. Anal. i Prilozh. 1984. V. 18. P. 77 (in Russian); English transl.: Funct. Anal. Appl. 1984. V. 18, №3. P. 245–246. [*422*]

[239] Dshalalow J.H. Real analysis. An introduction to the theory of real functions and integration. Chapman & Hall/CRC, Boca Raton, Florida, 2001; xiv+567 pp. [*413*]

[240] Dubins L.E., Pitman J. *A pointwise ergodic theorem for the group of rational rotations.* Trans. Amer. Math. Soc. 1979. V. 251. P. 299–308. [*435*]

[241] Dubrovskiĭ V.M. *On some properties of completely additive set functions and passing to the limit under the integral sign*. Izvest. Akad. Nauk USSR. 1945. V. 9, №4. P. 311–320; Remark: ibid. 1947. V. 11. P. 101–104 (in Russian). [*433*]

[242] Dubrovskiĭ V.M. *On some properties of completely additive set functions and their applications to a generalization of a theorem of H. Lebesgue*. Mat. Sbornik. 1947. V. 20. P. 317–330 (in Russian). [*433*]

[243] Dubrovskiĭ V.M. *On the basis of a family of completely additive functions of set and on the properties of uniform additivity and equi-continuity*. Dokl. Akad. Nauk SSSR. 1947. V. 58. P. 737–740 (in Russian). [*433*]

[244] Dubrovskiĭ V.M. *On properties of absolute continuity and equi-continuity*. Dokl. Akad. Nauk SSSR. 1948. V. 63. P. 483–486 (in Russian). [*433*]

[245] Dubrovskiĭ V.M. *On certain conditions of compactness*. Izv. Akad. Nauk SSSR. 1948. V. 12. P. 397–410 (in Russian). [*433*]

[246] Dubrovskiĭ V.M. *On equi-summable functions and on the properties of uniform additivity and equi-continuity of a family of completely additive set functions*. Izv. Akad. Nauk SSSR. 1949. V. 13. P. 341–356 (in Russian). [*433*]

[247] Dubrovskiĭ V.M. *On the continuity of definite integrals which depend on a parameter*. Dokl. Akad. Nauk SSSR. 1949. V. 66. P. 149–152 (in Russian). [*433*]

[248] Dubrovskiĭ V.M. *On the property of equicontinuity of a family of completely additive set functions with respect to proper and improper bases*. Dokl. Akad. Nauk SSSR. 1951. V. 76. P. 333–336 (in Russian). [*433*]

[249] Dubrovskiĭ V.M. *On a property of a formula of Nikodym*. Dokl. Akad. Nauk SSSR. 1952. V. 85. P. 693–696 (in Russian). [*324, 433*]

[250] Dubrovskiĭ V.M. *On the best majorant of a family of completely additive set functions*. Učen. Zap. Moskovsk. Gos. Univ. 1952. V. 163 (6). P. 89–98 (in Russian). [*433*]

[251] Dudley R.M. Real analysis and probability. Wadsworth & Brooks, Pacific Grove, California, 1989; xii+436 pp. [*62, 228, 413, 415*]

[252] Dunford N. *Integration in general analysis*. Trans. Amer. Math. Soc. 1935. V. 37. P. 441–453. [*424*]

[253] Dunford N. *Uniformity in linear spaces*. Trans. Amer. Math. Soc. 1938. V. 44. P. 305–356. [*431*]

[254] Dunford N. *A mean ergodic theorem*. Duke Math. J. 1939. V. 5. P. 635–646. [*434*]

[255] Dunford N., Pettis B.J. *Linear operations on summable functions*. Trans. Amer. Math. Soc. 1940. V. 47. P. 323–392. [*434*]

[256] Dunford N., Schwartz J.T. Linear operators, I. General Theory. Interscience, New York, 1958; xiv+858 pp. [*240, 282, 283, 321, 413, 415, 421, 423, 424, 434, 435*]

[257] Durrett R. Probability. Theory and Examples. 2nd ed., Duxbury Press, Wadsworth Publ., Belmont, 1996; xiii+503 pp. [*413*]

[258] D'yachenko M.I., Ulyanov P.L. Measure and integral. Faktorial, Moscow, 1998; 160 pp. (in Russian); Spanish transl.: Addison-Wesley Iberoamericana, 2000. [*413*]

[259] Eaton M.L. Lectures on topics in probability inequalities. Centrum voor Wiskunde en Informatica, Amsterdam, 1987; vi+197 pp. [*431*]

[260] Edgar G.A. Integral, probability, and fractal measures. Springer-Verlag, New York, 1998; 286 pp. [*413, 437*]

[261] Edgar G.A., Sucheston L. Stopping times and directed processes. Cambridge University Press, Cambridge, 1992; 428 pp. [*435, 438*]

[262] Edwards R.E. Functional analysis. Theory and applications. Holt, Rinehart and Winston, New York – London, 1965; xiii+781 pp. [*423*]

[263] Edwards R.E. Fourier series. A modern introduction. V. 1,2. 2nd ed. Springer-Verlag, New York – Berlin, 1979, 1982; 224 pp. , xi+369 pp. [*261*]

[264] Eggleston H.G. *Two measure properties of Cartesian product sets*. Quart. J. Math., Oxford, Ser. (2). 1954. V. 5. P. 108–115. [*235*]

[265] Egoroff D.-Th. *Sur les suites de fonctions mesurables*. C. R. Acad. Sci. Paris. 1911. T. 152. P. 244–246. [*426*]
[266] Egoroff D.-Th. *Sur l'intégration des fonctions mesurables*. C. R. Acad. Sci. Paris. 1912. T. 155. P. 1474–1475. [*437*]
[267] Eisen M. Introduction to mathematical probability theory. Prentice Hall, Englewood Cliffs, New Jersey, 1969; xxi+535 pp. [*413*]
[268] Elstrodt J. Mass- und Integrationstheorie. 4e Aufl. Springer-Verlag, Berlin – New York, 2005; xvi+434 S. [*413, 415*]
[269] Ene V. Real functions: current topics. Lecture Notes in Math. V. 1603. Springer-Verlag, Berlin – New York, 1995; xi+310 pp. [*436*]
[270] Erdős P., Kestelman H., Rogers C.A. *An intersection property of sets with positive measure*. Colloq. Math. 1963. V. 11. P. 75–80. [*90*]
[271] Erdős P., Oxtoby J.C. *Partitions of the plane into sets having positive measure in every non-null measurable product set*. Trans. Amer. Math. Soc. 1955. V. 79. P. 91–102. [*235*]
[272] Erdős P., Taylor S.J. *The Hausdorff measure of the intersection of sets of positive Lebesgue measure*. Mathematika. 1963. V. 10. P. 1–9. [*243*]
[273] Evans C., Gariepy R.F. Measure theory and fine properties of functions. CRC Press, Boca Raton – London, 1992; viii+268 pp. [*379, 437*]
[274] Faber V., Mycielski J. *The vanishing of certain integrals*. Houston J. Math. 1990. V. 16, №3. P. 313–316. [*240*]
[275] Faden A.M. Economics of space and time. The measure-theoretic foundations of social science. Iowa State University Press, Ames, Iowa, 1977; xiii+703 pp. [*423*]
[276] Falconer K.J. *Sections of sets of zero Lebesgue measure*. Mathematika. 1980. V. 27, №1. P. 90–96. [*234, 421*]
[277] Falconer K. Fractal geometry. John Wiley & Sons, New York, 1990; xxii+288 pp. [*67, 243, 437*]
[278] Falconer K.J., Mauldin R.D. *Fubini-type theorems for general measure constructions*. Mathematika. 2000. V. 47, №1–2. P. 251–265. [*210*]
[279] Farrell R.H. *Dense algebras of functions in L^p*. Proc. Amer. Math. Soc. 1962. V. 13. P. 324–328. [*308*]
[280] Fatou P. *Séries trigonométriques et séries de Taylor*. Acta Math. 1906. V. 30. P. 335–400. [*428*]
[281] Federer H. *Surface area. II*. Trans. Amer. Math. Soc. 1944. V. 55. P. 438–456. [*437*]
[282] Federer H. Geometric measure theory. Springer-Verlag, New York, 1969; xiv+676 pp. [*79, 243, 312, 373, 379, 381, 413, 430, 437*]
[283] Fernandez P.J. Medida e integracao. Instituto de Matematica Pura e Aplicada, Rio de Janeiro, 1976; viii+198 pp. [*413*]
[284] Fichera G. Lezioni sulle trasformazioni lineari. V. I. Introduzione all'analisi lineare. Istituto Matematico, Università, Trieste, 1954; xvii+502+iv pp. [*413*]
[285] Fichtenholz G. *Un théorème sur l'intégration sous le signe intégrale*. Rend. Circ. Mat. Palermo. 1913. T. 36. P. 111–114. [*234, 433*]
[286] Fichtenholz (Fihtengol'c) G.M. *Definite integrals depending on a parameter*. Mat. Sbornik. 1913. V. 29. P. 53–66 (in Russian). [*433*]
[287] Fichtenholz G. *Notes sur les limites des fonctions représentées par des intégrales définies*. Rend. Circ. Mat. Palermo. 1915. T. 40. P. 153–166. [*428, 433*]
[288] Fichtenholz G.M. Theory of simple integrals dependent on a parameter. Petrograd, 1918; vii+333 pp. (in Russian). [*276, 411, 428*]
[289] Fichtenholz G. *Note sur les fonctions absolument continues*. Bull. Acad. Belgique (5). 1922. V. 8. P. 430–443. [*396*]
[290] Fichtenholz G. *Sur les suites convergentes des intégrales définies*. Bull. International de l'Académie Polonaise des Sciences et des Lettres. Classe des Sciences Mathématiques et Naturelles. Sér. A: Sciences Mathématiques. Année 1923. P. 91–117. Cracovie, 1924. [*276, 433*]

[291] Fichtenholz G. *Sur une fonction de deux variables sans intégrale double*. Fund. Math. 1924. V. 6. P. 30–36. [*233*]

[292] Fichtenholz G.M. *On absolutely continuous functions*. Mat. Sbornik. 1924. V. 31. P. 286–296 (in Russian). [*391, 396*]

[293] Fichtenholz G. *Sur un problème de M. Banach*. Fund. Math. 1927. V. 10. P. 302–304. [*392*]

[294] Fichtenholz G. *Sur l'intégration des suites des fonctions sommables*. C. R. Acad. Sci. Paris. 1927. T. 184. P. 436–438. [*433*]

[295] Fichtenholz G.M. Differential- und Integralrechnung. B. I–III. Translated from the Russian. VEB Deutscher Verlag der Wissenschaften, Berlin, 1986, 1987. [*411*]

[296] Fichtenholz G., Kantorovitch L. *Sur les opérations linéaires dans l'espace de fonctions bornées*. Studia Math. 1936. V. 5. P. 68–98. [*435*]

[297] Filter W., Weber K. Integration theory. Chapman & Hall, London, 1997; xii+294 pp. [*413, 422*]

[298] Fischer E. *Sur la convergence en moyenne*. C. R. Acad. Sci. Paris. 1907. T. 144. P. 1022–1024. [*431*]

[299] Fischer E. *Applications d'un théorème sur la convergence en moyenne*. C. R. Acad. Sci. Paris. 1907. T. 144. P. 1148–1151. [*404*]

[300] Fleming W. Functions of several variables. 2nd ed. Springer-Verlag, Berlin – New York, 1977; xi+411 pp. [*414*]

[301] Floret K. Maß- und Integrationstheorie. B.G. Teubner, Stuttgart, 1981; 360 S. [*413*]

[302] Folland G.B. Real analysis: modern techniques and their applications. 2nd ed. Wiley, New York, 1999; xiv+386 pp. [*413*]

[303] Fominykh M.Yu. *Convergence almost everywhere of Riemann sums*. Vestnik Moskov. Univ. Ser. I Mat. Mekh. 1987. №6. P. 67–70 (in Russian); English transl.: Moscow Univ. Math. Bull. 1987. V. 42, №6. P. 69–72. [*435*]

[304] Fonda A. Lezioni sulla teoria dell'integrale. Edizioni goliardiche, Udine, 2001; 206 pp. [*413*]

[305] Foran J. Fundamentals of real analysis. Marcel Dekker, New York, 1991; xiv+473 pp. [*413*]

[306] Forster O. Analysis. B. 3. Integralrechnung im $\mathbb{R}^n$ mit Anwendungen. Vieweg, Braunschweig, 1999; vii+285 S. [*414*]

[307] Fréchet M. *Sur les ensembles de fonctions et les opérations linéaires*. C. R. Acad. Sci. Paris. 1907. T. 144. P. 1414–1415. [*431, 434*]

[308] Fréchet M. *Définition de l'intégrale sur un ensemble abstrait*. C. R. Acad. Sci. Paris. 1915. T. 161. P. 839–840. [*418, 425*]

[309] Fréchet M. *Sur l'intégrale d'une fonctionnelle étendue a un ensemble abstrait*. Bull. Sci. Math. France. 1915. V. 43. P. 249–265. [*418, 425, 429*]

[310] Fréchet M. *Sur divers modes de convergence d'une suite de fonctions d'une variable*. Bull. Calcutta Math. Soc. 1921. V. 11. P. 187–206. [*418, 426*]

[311] Fréchet M. *Familles additives et fonctions additives d'ensembles abstraits*. Enseignement Math. 1922. V. 22. P. 113–129. [*418*]

[312] Fréchet M. *Sur la distance de deux ensembles*. C. R. Acad. Sci. Paris. 1923. T. 176. P. 1123–1124. [*418, 426*]

[313] Fréchet M. *Des familles et fonctions additives d'ensembles abstraits*. Fund. Math. 1923. V. 4. P. 329–365. [*418, 421, 429*]

[314] Fréchet M. *Des familles et fonctions additives d'ensembles abstraits. Suite*. Fund. Math. 1924. V. 5. P. 206–251. [*419*]

[315] Fréchet M. *Sur la distance de deux ensembles*. Bull. Calcutta Math. Soc. 1924. V. 15. P. 1–8. [*418, 426*]

[316] Fréchet M. *Sur les ensembles compacts de fonctions mesurables*. Fund. Math. 1927. V. 9. P. 25–32. [*321, 426*]

[317] Fréchet M. *Sur la convergence en probabilité*. Metron. 1930. V. 8. P. 3–50. [*426*]

[318] Fréchet M. *Sur les ensembles compacts de fonctions de carrés intégrables.* Acta Sci. Math. (Szeged). 1937. V. 8. P. 116–126. [*434*]

[319] Fréchet M. *Les éléments aléatoires de nature quelconque dans un espace distancié.* Ann. Inst. H. Poincaré. 1948. V. 10. P. 215–310. [*426*]

[320] Fréchet M. Généralités sur les probabilités. Éléments aléatoires. 2d ed. Gauthier-Villars, Paris, 1950; xvi+355 pp. [*171, 426*]

[321] Fréchet M. Pages choisies d'analyse générale. Gauthier-Villars, Paris, 1953; 213 pp. [*409, 426*]

[322] Fremlin D.H. Topological Riesz spaces and measure theory. Cambridge University Press, London, 1974; 266 pp. [*312*]

[323] Fremlin D.H. Consequences of Martin's Axiom. Cambridge University Press, Cambridge, 1985; xii+325 pp. [*78, 80*]

[324] Fremlin D.H. *On the average of inner and outer measures.* Fund. Math. 1991. V. 139, №1. P. 9–15. [*100*]

[325] Fremlin D. *Real-valued-measurable cardinals.* Set theory of the reals (Ramat Gan, 1991), pp. 151–304; Israel Math. Conf. Proc. 6, Bar Ilan Univ., Ramat Gan, 1993. [*80*]

[326] Fremlin D.H. *Weakly α-favourable measure spaces.* Fund. Math. 2000. V. 165, №1. P. 67–94. [*53*]

[327] Fremlin D. Measure theory. V. 1–5. University of Essex, Colchester, 2000–2003. [*74, 98, 235, 237, 325, 413, 421, 434*]

[328] Friedman H. *A consistent Fubini–Tonelli theorem for nonmeasurable functions.* Illinois J. Math. 1980. V. 24. P. 390–395. [*209*]

[329] Fristedt B., Gray L. A modern approach to probability theory. Birkhäuser Boston, Boston, 1997; xx+756 pp. [*413*]

[330] Frumkin P.B. *On a theorem of D.Th. Egoroff on measurable functions.* Dokl. Akad. Nauk SSSR. 1948. V. 60, №6. P. 973–975 (in Russian). [*160*]

[331] Fubini G. *Sugli integrali multipli.* Atti Accad. Lincei Rend., Cl. sci. fis., mat. e natur. Roma. (5). 1907. V. 16, Sem. 1. P. 608–614. [*429*]

[332] Fubini G. *Sulla derivazione per serie.* Atti Accad. Lincei Rend., Cl. sci. fis., mat. e natur. Roma. (5). 1915. V. 24, Sem. 1, f. 3. P. 204–206. [*336*]

[333] Fubini G. Opere scelte. V. 1–3. Edizioni Cremonese, Roma, 1957, 1958, 1962. [*409*]

[334] Fukuda R. *Exponential integrability of sub-Gaussian vectors.* Probab. Theory Relat. Fields. 1990. V. 85. P. 505–521. [*169*]

[335] Galambos J. Advanced probability theory. Marcel Dekker, New York, 1988; viii+405 pp. [*413*]

[336] Galambos J., Simonelli I. Bonferroni-type inequalities with applications. Springer, New York, 1996; 296 pp. [*103*]

[337] Gänssler P., Stute W. Wahrscheinlichkeitstheorie. Springer-Verlag, Berlin – New York, 1977; xii+418 pp. [*413*]

[338] Gaposhkin V.F. *Lacunary series and independent functions.* Uspehi Mat. Nauk. 1966. V. 21, №6. P. 3–83 (in Russian); English transl.: Russian Math. Surveys. 1966. V. 21, №6. P. 1–82. [*317, 434*]

[339] Gaposhkin V.F. *Convergence and limit theorems for sequences of random variables.* Teor. Verojatn. i Primen. 1972. V. 17, №3. P. 401–423 (in Russian); English transl.: Theory Probab. Appl. 1972. V. 17. P. 379–400. [*434*]

[340] García-Cuerva J., Rubio de Francia J.L. Weighted norm inequalities and related topics. North-Holland, Amsterdam, 1985; x+604 p. [*375*]

[341] Gardner R.J. *A positive answer to the Busemann–Petty problem in three dimensions.* Ann. Math. (2). 1994. V. 140, №2. P. 435–447. [*215*]

[342] Gardner R.J. *The Brunn–Minkowski inequality.* Bull. Amer. Math. Soc. 2002. V. 39, №3. P. 355–405. [*226*]

[343] Gardner R.J., Koldobsky A., Schlumprecht T. *An analytic solution to the Busemann–Petty problem on sections of convex bodies.* Ann. Math. 1999. V. 149, №2. P. 691–703. [*215*]

[344] Garnir H.G. Fonctions de variables réelles. T. 2. Librairie Universitaire,Louvain, Gauthier-Villars, Paris, 1965; viii+560 pp. [*413*]

[345] Garnir H.G., De Wilde M., Schmets J. Analyse fonctionnelle. Tome II. Measure et intégration dans l'espace euclidien E_n. Tome III. Espaces fonctionnels usuels. Birkhäuser Verlag, Basel – Stuttgart, 1972, 1973; 287 pp. , 375 pp. [*413*]

[346] Garsia A.M. Topics in almost everywhere convergence. Markham Publ., Chicago, 1970; x+154 pp. [*261*]

[347] Gaughan E. Introduction to analysis. Brooks/Coole, Belmont, 1968; viii+310 pp. [*413*]

[348] Gelbaum B. Problems in real and complex analysis. Springer, New York, 1992; x+488 pp. [*415*]

[349] Gelbaum B.R., Olmsted J.M.H. Counterexamples in analysis. Holden-Day, San Francisco – London – Amsterdam, 1964; xxiv+194 pp. [*414*]

[350] Genet J. Mesure et intégration. Théorie élémentaire. Cours et exercices resolus. Vuibert, Paris, 1976; 322 pp. [*413*]

[351] George C. Exercises in integration. Springer-Verlag, Berlin – New York, 1984; 550 pp. [*87, 91, 173, 307, 415*]

[352] Giaquinta M., Modica G., Souček J. Cartesian currents in the calculus of variations. V. I, II. Springer, Berlin – New York, 1998; xxiv+711 pp. , xxiv+697 pp. [*379*]

[353] Gikhman I.I., Skorokhod A.V. Introduction to the theory of random processes. Translated from the Russian. Saunders, Philadelphia – London – Toronto, 1969; xiii+516 pp. (Russian ed.: Moscow, 1965). [*413*]

[354] Gillis J. *Note on a property of measurable sets.* J. London Math. Soc. 1936. V. 11. P. 139–141. [*90*]

[355] Gillis J. *Some combinatorial properties of measurable sets.* Quart. J. Math. 1936. V. 7. P. 191–108. [*90*]

[356] Girardi M. *Compactness in L_1, Dunford-Pettis operators, geometry of Banach spaces.* Proc. Amer. Math. Soc. 1991. V. 111, №3. P. 767–777. [*434*]

[357] Girardi M. *Weak vs. norm compactness in L_1: the Bocce criterion.* Studia Math. 1991. V. 98. P. 95–97. [*434*]

[358] Giustu E. Minimal surfaces and functions of bounded variation. Birkhäuser Verlag, Basel, 1984; xii+240 pp. [*379*]

[359] Gładysz S., Marczewski E., Ryll-Nardzewski C. *Concerning distances of sets and distances of functions.* Colloq. Math. 1961. V. 8. P. 71–75. [*102*]

[360] Glazkov V.N. *On inner and outer measures.* Sibirsk. Mat. Zh. 1988. V. 29, №1. P. 197–201 (in Russian); English transl.: Sib. Math. J. 1988. V. 29, №1. P. 152–155. [*95, 421*]

[361] Gleason A.M. Fundamentals of abstract analysis. Reading, Addison-Wesley, Massachusetts, 1966; xi+404 pp. [*413*]

[362] Glivenko V.I. The Stieltjes integral. ONTI, Moscow – Leningrad, 1936; 216 pp. (in Russian) [*425*]

[363] Gnedenko B.V., Kolmogorov A.N. Limit distributions for sums of independent random variables. GITTL, Moscow, 1949 (in Russian); English transl.: Addison-Wesley, Cambridge, Mass., 1954; ix+264 pp. [*412*]

[364] Gneiting T. *Curiosities of characteristic functions.* Exposition. Math. 2001. V. 19, №4. P. 359–363. [*246*]

[365] Godement R. Analyse mathématique. IV. Springer-Verlag, Berlin, 2003; xii+599 pp. [*414*]

[366] Goffman C. Real functions. Prindle, Weber & Schmidt, Boston, 1953; xii+263 pp. [*399, 413*]

[367] Goffman C., Pedrick G. First course in functional analysis. Prentice-Hall, Englewood Cliffs, 1965; xi+282 pp. [*413*]
[368] Goguadze D.F. On the Kolmogorov integrals and some their applications. Metsniereba, Tbilisi, 1979; 288 pp. (in Russian). [*435, 437*]
[369] Gohman E.H. The Stieltjes integral and its applications. Fizmatgiz, Moscow, 1958; 192 pp. (in Russian). [*425*]
[370] Goldberg R.R. Methods of real analysis. 2nd ed. John Wiley & Sons, New York – London – Sydney, 1976; x+402 pp. [*413*]
[371] Gol'dshteĭn V.M., Reshetnyak Yu.G. Quasiconformal mappings and Sobolev spaces. Translated and revised from the 1983 Russian original. Kluwer Academic Publ., Dordrecht, 1990; xx+371 pp. (Russian ed.: Introduction to the theory of functions with generalized derivatives and quasiconformal mappings. Moscow, 1983). [*379*]
[372] Gomes R.L. Intégral de Riemann. Junta de Investigąo Matemática, Porto, 1949; viii+309+i pp. [*437*]
[373] Gordon R.A. The integrals of Lebesgue, Denjoy, Perron, and Henstock. Amer. Math. Soc., Providence, Rhode Island, 1994; 395 pp. [*353, 357, 437*]
[374] Gordon R.A. *Some comments on the McShane and Henstock integrals*. Real Anal. Exchange. 1997–1998. V. 23, №1. P. 329–342. [*406*]
[375] Gouyon R. Intégration et distributions. Vuibert, Paris, 1967; 170 pp. [*413*]
[376] Gowurin M.K. *On sequences of indefinite integrals*. Bull. Amer. Math. Soc. 1936. V. 42, №12. P. 930–936. [*276, 322*]
[377] Gramain A. Intégration. Hermann, Paris, 1988; x+229 pp. [*413*]
[378] Grauert H., Lieb I. Differential- und Integralrechnung. III. Integrationstheorie. Kurven- und Flächenintegrale. Vektoranalysis. 2e Aufl. Springer-Verlag, Berlin – New York, 1977; xiv+210 S. [*413*]
[379] Gravé D. *Sur les lignes composées de parties rectilignes*. C. R. Acad. Sci. Paris. 1898. T. 127. P. 1005–1007. [*436*]
[380] Graves L.M. The theory of functions of real variables. McGraw-Hill, New York, 1946 (2nd ed.: 1956; xii+375 pp.). [*413*]
[381] Gromov M. *Monotonicity of the volume of intersection of balls*. Lecture Notes in Math. 1987. V. 1267. P. 1–4. [*246*]
[382] Gruber P.M. *How well can space be packed with smooth bodies? Measure-theoretic results*. J. London Math. Soc. (2). 1995. V. 52, №1. P. 1–14. [*422*]
[383] Gunther N.M. Sur les intégrales de Stieltjes et leur application aux problèmes de la physique mathématique. Trudy Steklov Mat. Inst., Moscow, 1932; 494 pp. (2e éd.: Chelsey, New York, 1949). [*425*]
[384] Günzler H. Integration. Bibliogr. Inst., Mannheim, 1985; 392 S. [*413*]
[385] Gut A. Probability: a graduate course. Springer, New York, 2005; xxiv+603 pp. [*413*]
[386] Guzmán M. de. Differentiation of integrals in $\mathbb{R}^n$. Lecture Notes in Math. V. 481. Springer-Verlag, Berlin – New York, 1975; xii+266 pp. [*67, 346, 353, 436*]
[387] Guzmán M. de. Real variable methods in Fourier analysis. North-Holland, Amsterdam, 1981; 392 pp. [*436*]
[388] Guzmán M. de, Rubio B. Integración: teoría y técnicas. Editorial Alhambra, Madrid, 1979; x+171 pp. [*413*]
[389] Haaser N.B., Sullivan J.A. Real analysis. Dover Publications, New York, 1991; x+341 pp. [*413*]
[390] Hačaturov A.A. *The determination of the value of the measure for a region of n-dimensional Euclidean space from its values for all half-spaces*. Uspehi Matem. Nauk. 1954. V. 9, №3. P. 205–212 (in Russian). [*228*]
[391] Hackenbroch W. Integrationstheorie. Eine Einführung in die Integrationstheorie und ihre Anwendungen. Teubner, Stuttgart, 1987; 144 S. [*413*]
[392] Hadwiger H. Vorlesungen über Inhalt, Oberfläche und Isoperimetrie. Springer-Verlag, Berlin – Heidelberg, 1957; xiii+312 S. [*82, 227, 431*]

[393] Hahn H. *Über den Fundamentalsatz der Integralrechnung*. Monatsh. Math. Phys. 1905. B. 16. S. 161–166. [*402*]

[394] Hahn H. *Über die Integrale des Herrn Hellinger und die Orthogonalinvarianten der quadratischen Formen von unendlich vielen Veränderlichen*. Monatsh. Math. Phys. 1912. B. 23. S. 161–264. [*435*]

[395] Hahn H. *Über Annährung an Lebesgue'sche Integrale durch Riemann'sche Summen*. Sitz. Akad. Wiss. Wien, Math.-naturwiss. Kl. IIa. 1914. B. 123, H. 4. S. 713–743. [*164, 435*]

[396] Hahn H. *Über eine Verallgemeinerung der Riemannschen Integraldefinition*. Monatsh. Math. Phys. 1915. B. 26. S. 3–18. [*423*]

[397] Hahn H. *Einige Anwendungen der Theorie der singulären Integrale*. Sitz. Akad. Wiss. Wien, Math.-naturwiss. Kl. IIa. 1918. B. 127, H. 9. S. 1763–1785. [*428*]

[398] Hahn H. Theorie der reellen Funktionen. B. I. Springer, Berlin, 1921; vii+614 S. [*411, 418, 421, 429*]

[399] Hahn H. *Über Folgen linearer Operationen*. Monatsh. Math. Phys. 1922. B. 32. S. 3–88. [*433*]

[400] Hahn H. *Über die Multiplikation total-additiver Mengenfunktionen*. Annali Scuola Norm. Super. Pisa (2). 1933. V. 2. P. 429–452. [*419*]

[401] Hahn H. Gesammelte Abhandlungen/Collected works. V. 1, 2. With biographical sketches by K. Popper, L. Schmetterer, K. Sigmund, and commentaries on Hahn's work by H. Heuser, H. Sagan, L. Fuchs, W. Frank, D. Preiss, and A. Kluwick. Springer-Verlag, Vienna, 1995, 1996. [*409*]

[402] Hahn H., Rosenthal A. Set functions. Univ. New Mexico Press, Albuquerque, New Mexico, 1948; 324 pp. [*415, 419, 421*]

[403] Hajłasz P. *Change of variables formula under minimal assumptions*. Colloq. Math. 1993. V. 64, №1. P. 93–101. [*381*]

[404] Halmos P. Measure theory. Van Nostrand, New York, 1950; xi+304 pp. [*v, 180, 412*]

[405] Halmos P.R., Savage L.J. *Applications of the Radon–Nikodym theorem to the theory of sufficient statistics*. Ann. Math. Statist. 1949. V. 20. P. 225–241. [*279*]

[406] Hanisch H., Hirsch W.M., Renyi A. *Measures in denumerable spaces*. Amer. Math. Monthly. 1969. V. 76, №5. P. 494–502. [*104*]

[407] Hanner O. *On the uniform convexity of L^p and l^p*. Ark. Math. 1956. B. 3. S. 239–244. [*325*]

[408] Hardy G.H., Littlewood J.E., Pólya G. Inequalities. 2d ed. Cambridge, at the University Press, 1952; xii+324 pp. [*243, 429*]

[409] Hardy G.H., Rogosinski W.W. Fourier series. Cambridge University Press, Cambridge, 1944; 100 pp. [*261*]

[410] Hartman S., Mikusiński J. The theory of Lebesgue measure and integration. Pergamon Press, New York, 1961; 176 pp. [*413*]

[411] Haupt O., Aumann G., Pauc Ch.Y. Differential- und Integralrechnung unter besonderer Berücksichtigung neuerer Ergebnisse. B. III. Integralrechnung. 2e Aufl. Walter de Gruyter, Berlin, 1955; xi+320 S. (1e Aufl.: Haupt O., Aumann G., Berlin, 1938; 183 S.) [*411, 413*]

[412] Hausdorff F. Grundzüge der Mengenlehre. Leipzig, 1914; 476 S. [*410, 422*]

[413] Hausdorff F. *Die Mächtigkeit der Borelschen Mengen*. Math. Ann. 1916. B. 77, №3. S. 430–437. [*420*]

[414] Hausdorff F. *Dimension und äusseres Mass*. Math. Ann. 1919. B. 79. S. 157–179. [*430*]

[415] Hausdorff F. Gesammelte Werke. B. 1–8. Springer, Berlin, 2001–2005. [*409, 421*]

[416] Hawkins T. Lebesgue's theory of integration. Its origins and development. University of Wisconsin Press, Madison – London, 1970; xv+227 pp. [*417, 423*]

[417] Hayes C.A., Pauc C.Y. Derivation and martingales. Springer, Berlin – New York, 1970; vii+203 pp. [*438*]

[418] Heinonen J. Lectures on analysis on metric spaces. Springer, New York, 2001; x+140 pp. [*375*]
[419] Helgason S. The Radon transform. 2nd ed. Birkhäuser, Boston – Basel, 1999; xiv+188 pp. [*227*]
[420] Hellinger E. *Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen.* J. Reine Angew. Math. 1909. B. 136. S. 210–271. [*301, 435*]
[421] Hennequin P.-L., Tortrat A. Théorie des probabilités et quelques applications. Masson et Gie, Paris, 1965; viii+457 pp. [*413*]
[422] Henstock R. Theory of integration. Butterworths, London, 1963; ix+168 pp. [*414, 437*]
[423] Henstock R. *A Riemann–type integral of Lebesgue power.* Canad. J. Math. 1968. V. 20. P. 79–87. [*436*]
[424] Henstock R. Lectures on the theory of integration. World Sci., Singapore, 1988; xiv+206 pp. [*414, 437*]
[425] Henstock R. *A short history of integration theory.* Southeast Asian Bull. Math. 1988. V. 12, №2. P. 75–95. [*437*]
[426] Henstock R. The general theory of integration. Clarendon Press, Oxford University Press, New York, 1991; xii+262 pp. [*414, 437*]
[427] Henze E. Einführung in die Maßtheorie. 2e Aufl. Bibliogr. Inst., Mannheim, 1985; 235 S. [*414*]
[428] Herglotz G. *Über Potenzreichen mit positiven reellen Teil im Einheitskreis.* Leipziger Berichte. 1911. B. 63. S. 501–511. [*430*]
[429] Hesse C. Angewandte Wahrscheinlichkeitstheorie. Ein fundierte Einführung mit über 500 realitätsnahen Beispielen und Aufgaben. Vieweg, Wiesbaden, 2003; x+505 S. [*414*]
[430] Heuser H. Lehrbuch der Analysis. B. 2. 11e Aufl., Teubner, Stuttgart, 2000; 737 S. [*414*]
[431] Hewitt E., Stromberg K. Real and abstract analysis. Englewood Springer, Berlin – New York, 1975; x+476 pp. [*325, 414*]
[432] Hildebrandt T.H. *On integrals related to and extensions of the Lebesgue integrals.* Bull. Amer. Math. Soc. (2). 1917. V. 24. P. 113–144, 177–202. [*410*]
[433] Hildebrandt T.H. Introduction to the theory of integration. Academic Press, New York – London, 1963; ix+385 pp. [*414*]
[434] Hille E. Methods in classical and functional analysis. Addison-Wesley, Reading – London, 1972; ix+486 pp. [*414*]
[435] Hinderer K. Grundbegriffe der Wahrscheinlichkeitstheorie. 3e Aufl. Springer, Berlin, 1985; vii+247 S. [*414*]
[436] Hobson E.W. The theory of functions of a real variable and the theory of Fourier's series. Cambridge, 1907; xv+772 pp. 2d, 3d ed.: V. 1,2. 1926–1927; 780 pp. , 726 pp. [*410*]
[437] Hochkirchen T. *Theory of measure and integration from Riemann to Lebesgue.* In: A history of analysis (H.N. Jahnke ed.), pp. 261–290, Hist. Math., 24, Amer. Math. Soc., Providence, Rhode Island, 2003. [*417, 423*]
[438] Hoffman K. Analysis in Euclidean space. Prentice-Hall, Englewood Cliffs, 1975; 432 pp. [*414*]
[439] Hoffmann D., Schäfke F.W. Integrale. Bibliogr. Inst., Mannheim, 1992; xiv+333 S. [*414*]
[440] Hoffmann-Jørgensen J. Probability with a view toward statistics. V. I, II. Chapman & Hall, New York, 1994; 589 pp. , 533 pp. [*95, 414, 421*]
[441] Holdgrün H.S. Analysis. B. 2. Leins-Verl., Göttingen, 2001; x+485 S. [*414*]
[442] Hopf E. *On causality, statistics and probability.* J. Math. Phys. 1934. V. 13. P. 51–102. [*419, 429*]
[443] Hopf E. Ergodentheorie. Springer, Berlin, 1937; iv+83 S. [*419*]

[444] Howard E.J., Pfeffer W. *Luzin's theorem for charges.* Proc. Amer. Math. Soc. 2004. V. 132, №3. P. 857–863. [*369*]
[445] Hu S. Elements of real analysis. Holden-Day, San Francisco, 1967; xii+365 pp. [*414*]
[446] Hulanicki A. *Invariant extensions of the Lebesgue measure.* Fund. Math. 1962. V. 51. P. 111–115. [*422*]
[447] Humke P.D., Preiss D. *Measures for which σ-porous sets are null.* J. London Math. Soc. (2). 1985. V. 32, №2. P. 236–244. [*404*]
[448] Hunt G.A. Martingales et processus de Markov. Dunod, Paris, 1966; 143 pp. [*309*]
[449] Ingleton A.W. Notes on integration. Mathematical Institute, Oxford University, Oxford, 1972; i+90 pp. [*414*]
[450] Ivanov L.D. Variations of sets and functions. Nauka, Moscow, 1975; 352 pp. (in Russian). [*437*]
[451] Ivanov V.V. *A geometric criterion for Jordan measurability.* Sibirsk. Mat. Zh. 1983. V. 24, №5. P. 48–49 (in Russian); English transl.: Sib. Math. J. 1984. V. 24, №5. P. 685–686. [*237*]
[452] Jacobs K. Measure and integral. Academic Press, New York, 1978; 575 pp. [*414*]
[453] Jain P.K., Gupta V.P. Lebesgue measure and integration. Wiley Eastern, New Delhi, 1986; viii+260 pp. [*414*]
[454] James R.C. Advanced calculus. Wadsworth, Belmont, 1966; x+646 pp. [*414*]
[455] Janssen A.J.E.M., Steen P. van der. Integration theory. Lecture Notes in Math. V. 1078. Springer–Verlag, Berlin – New York, 1984; ii+224 pp. [*414*]
[456] Jayne J. *Generation of σ-algebras, Baire sets, and descriptive Borel sets.* Mathematika. 1977. V. 24, №2. P. 241–256. [*421*]
[457] Jean R. Mesure et intégration. Les Presses de l'Université du Quebec, Montréal, 1975; 305 pp. [*414*]
[458] Jech Th.J. Lectures in set theory, with particular emphasis on the method of forcing. Lecture Notes in Math. V. 217. Springer-Verlag, Berlin – New York, 1971; iv+137 pp. [*78, 80*]
[459] Jech T. Set theory. Academic Press, New York, 1978; xi+621 pp. The third millennium edition, revised and expanded. Springer-Verlag, Berlin, 2003. xiv+769 pp. [*62, 79, 80*]
[460] Jefferies B., Ricker W.J. *Integration with respect to vector valued Radon polymeasures.* J. Austral. Math. Soc. Ser. A. 1994. V. 56, №1. P. 17–40. [*423*]
[461] Jeffery R. The theory of functions of a real variable. 2nd ed., University of Toronto Press, 1953; xi+232 pp. [*414*]
[462] Jensen J.L.W.V. *Sur les fonctions convexes et les inégalités entres les valeurs moyennes.* Acta. Math. 1906. V. 30. P. 175–193. [*429*]
[463] Jessen B. *The theory of integration in a space of an infinite number of dimensions.* Acta Math. 1934. V. 63. P. 249–323. [*419, 429, 435*]
[464] Jessen B. *On the approximation of Lebesgue integrals by Riemann sums.* Ann. Math. (2). 1934. V. 35. P. 248–251. [*435*]
[465] Jessen B. Abstrakt Maal- og Integralkteori. Matematisk Forening i København, København, 1947; iii+108 pp. [*412*]
[466] Jessen B., Marcinkiewicz J., Zygmund A. *Note on the differentiability of multiple integrals.* Fund. Math. 1935. V. 25. P. 217–234. [*437*]
[467] Jessen B., Wintner A. *Distribution functions and the Riemann zeta function.* Trans. Amer. Math. Soc. 1935. V. 38. P. 48–88. [*430*]
[468] Jiménez Pozo M.A. Medida, integración y funcionales. Editorial Pueblo y Educación, Habana, 1989; 504 pp. [*414*]
[469] Jones F.B. *Measure and other properties of a Hamel basis.* Bull. Amer. Math. Soc. 1942. V. 48. P. 472–481. [*86, 422*]
[470] Jones F. Lebesgue integration on Euclidean space. Jones and Bartlett, Boston, 1993; xvi+588 pp. [*414*]

[471] Jørboe O.G., Mejlbro L. The Carleson–Hunt theorem on Fourier series. Lecture Notes in Math. V. 911. Springer-Verlag, Berlin – New York, 1982; 123 pp. [*260*]

[472] Jordan C. Cours d'analyse de l'École Polytechnique. T. 1–3. 2e éd. Paris, 1893–1896; 612 pp. , 627 pp. , 542 pp. [*2, 416*]

[473] Jost J. Postmodern analysis. Springer-Verlag, Berlin, 1998; xviii+353 pp. [*414*]

[474] Kaczmarz S., Nikliborc L. *Sur les suites des fonctions convergentes en moyenne.* Fund. Math. 1928. V. 11. P. 151–168. [*319*]

[475] Kaczor W.J., Nowak M.T. Problems in mathematical analysis III: integration. Amer. Math. Soc., Providence, Rhode Island, 2003. [*415*]

[476] Kadec M.I., Pełczyński A. *Bases, lacunary sequences and complemented subspaces in the spaces L_p.* Studia Math. 1962. T. 21. P. 161–176. [*174*]

[477] Kahane C.S. *Evaluating Lebesgue integrals as limits of Riemann sums.* Math. Japon. 1993. V. 38. P. 1073–1076. [*435*]

[478] Kahane J.-P. Some random series of functions. D.C. Heath, Lexington, Massachusetts, 1968; viii+184 pp.; 2nd ed.: Cambridge University Press, Cambridge, 1985; xiv+305 pp. [*429*]

[479] Kahane J.-P. *Trois notes sur les ensembles parfaits linéaires.* Enseign. Math. (2). 1969. V. 15. P. 185–192. [*66, 103*]

[480] Kakutani S. *Notes on infinite product measures, I, II.* Proc. Imperial Acad. Tokyo. 1943. V. 19. P. 148–151, 184–188. [*429*]

[481] Kakutani S. *Notes on divergent series and integrals.* Proc. Imper. Acad. Tokyo. 1944. V. 20. P. 74–76. [*173*]

[482] Kakutani S. Selected papers. V. 2. Birkhäuser, Boston, 1986. [*409, 429*]

[483] Kakutani S., Oxtoby J.C. *Construction of a nonseparable invariant extension of the Lebesgue measure space.* Ann. Math. (2). 1950. V. 52. P. 580–590. [*81*]

[484] Kallenberg O. Foundations of modern probability. 2nd ed. Springer-Verlag, New York, 2002; xx+638 pp. [*414*]

[485] Kamke E. Das Lebesguesche Integral. Eine Einführung in die neuere Theorie der reellen Funktionen. Teubner, Leipzig, 1925; 151 S. [*411*]

[486] Kamke E. Das Lebesgue–Stieltjes–Integral. B.G. Teubner Verlagsgesellschaft, Leipzig, 1956; vi+226 S. [*414, 426*]

[487] Kampen E.R. van. *Infinite product measures and infinite convolutions.* Amer. J. Math. 1940. V. 62, №2. P. 417–448. [*429*]

[488] Kannan R., Krueger C.K. Advanced analysis on the real line. Springer, New York, 1996; ix+259 p. [*399, 404, 406, 408, 436*]

[489] Kannan R., Lovász L., Simonovits M. *Isoperimetric problems for convex bodies and a localization lemma.* Discrete and Comput. Geom. 1995. V. 13. P. 541–559. [*173*]

[490] Kanovei V.G. The axiom of choice and the axiom of determinacy. Nauka, Moscow, 1984; 65 pp. (in Russian). [*80*]

[491] Kantorovitz S. Introduction to modern analysis. Oxford University Press, New York, 2003; xii+434 pp. [*414*]

[492] Kappos D.A. Strukturtheorie der Wahrscheinlichkeitsfelder und -räume. Springer, Berlin, 1960; 136 S. [*421*]

[493] Kappos D.A. Probability algebras and stochastic spaces. Academic Press, New York, 1969; x+267 pp. [*421*]

[494] Karr A.F. Probability. Springer-Verlag, New York, 1993; xxii+282 pp. [*414*]

[495] Kashin B.S., Saakyan A.A. Orthogonal series. Translated from the Russian. Amer. Math. Soc., Providence, Rhode Island, 1989; xii+451 pp. (Russian ed.: Moscow, 1984). [*261, 306*]

[496] Katznelson Y., Stromberg K. *Everywhere differentiable, nowhere monotone, functions.* Amer. Math. Monthly. 1974. V. 81. P. 349–354. [*402*]

[497] Kaufman R.P., Rickert N.W. *An inequality concerning measures.* Bull. Amer. Math. Soc. 1966. V. 72. P. 672–676. [*244*]

[498] Kaufman R., Wu J.-M. *Two problems on doubling measures.* Rev. Mat. Iberoamer. 1995. V. 11. P. 527–545. [*376*]

[499] Kawata T. Fourier analysis in probability theory. Academic Press, New York, 1972; xii+668 pp. [*430*]

[500] Keleti T. *Density and covering properties of intervals in* $\mathbb{R}^n$. Mathematika. 2000. V. 47. P. 229–242. [*436*]

[501] Kelley J.L. *Measures on Boolean algebras.* Pacif. J. Math. 1959. V. 9, №4. P. 1165–1177. [*422*]

[502] Kelley J.L., Srinivasan T.P. Measure and integral. V. 1. Springer-Verlag, New York – Berlin, 1988; x+150 pp. [*94, 414*]

[503] Kenyon H., Morse A.P. Web derivatives. Mem. Amer. Math. Soc., No. 132. Amer. Math. Soc., Providence, Rhode Island, 1973; xiii+177 pp. [*438*]

[504] Kestelman H. Modern theories of integration. Oxford, 1937. 2nd revised ed. Dover Publications, New York, 1960; x+309 pp. [*411, 437*]

[505] Khakhubia G.P. The generalized Stieltjes–Gunther integral and its applications to some problems of mathematical physics. Gruzin. Politehn. Inst., Tbilisi, 1965; 120 pp. (in Russian). [*425*]

[506] Kharazishvili A.B. Problems of set theory and measure theory. Tbilis. Gos. Univ., Tbilisi, 1978; 178 pp. (in Russian). [*80*]

[507] Kharazishvili A.B. Invariant extensions of the Lebesgue measure. Tbilis. Gos. Univ., Tbilisi, 1983; 204 pp. (in Russian). [*79, 82*]

[508] Kharazishvili A.B. Topological aspects of measure theory. Naukova Dumka, Kiev, 1984; 120 pp. (in Russian). [*80*]

[509] Kharazishvili A.B. Vitali's theorem and its generalizations. Tbilis. Gos. Univ., Tbilisi, 1991; 108 pp. (in Russian). [*436*]

[510] Kharazishvili A.B. Transformation groups and invariant measures. Set-theoretical aspects. World Sci., River Edge, 1998; viii+260 pp. [*82*]

[511] Kharazishvili A.B. Strange functions in real analysis. Marcel Dekker, New York, 2000; viii+297 pp. [*80, 91, 211*]

[512] Kharazishvili A.B. Nonmeasurable sets and functions. Elsevier, Amsterdam, 2004; xi+337 pp. [*82*]

[513] Khintchine A. *Sur la dérivation asymptotique.* C. R. Acad. Sci. Paris. 1917. T. 164. P. 142–144. [*438*]

[514] Khintchine A. *Recherches sur la structure des fonctions mesurables. I,II.* Mat. Sbornik. 1924. V. 31. P. 265–285, 377–433 (in Russian); French transl.: Moscou Recueil Math. 1923. V. 31. P. 265–285; Fund. Math. 1927. V. 9. P. 212–279. [*438*]

[515] Kindler J. *A Mazur–Orlicz type theorem for submodular set functions.* J. Math. Anal. Appl. 1986. V. 120, №2. P. 533–546. [*422*]

[516] Kindler J. *Supermodular and tight set functions.* Math. Nachr. 1987. B. 134. S. 131–147. [*422*]

[517] Kindler J. *Sandwich theorems for set functions.* J. Math. Anal. Appl. 1988. V. 133, №2. P. 529–542. [*100*]

[518] Kingman J.F.C., Taylor S.J. Introduction to measure and probability. Cambridge University Press, London – New York – Ibadan, 1966; x+401 pp. [*414*]

[519] Kirillov A.A., Gvishiani A.D. Theorems and problems in functional analysis. Translated from the Russian. Springer-Verlag, New York – Berlin, 1982; ix+347 pp. (Russian ed.: Moscow, 1979). [*414, 415*]

[520] Kisyński J. *Remark on strongly additive set functions.* Fund. Math. 1968. V. 63. P. 327–332. [*422*]

[521] Klambauer G. Real analysis. American Elsevier, New York, 1973; 436 pp. [*414*]

[522] Klei H.-A., Miyara M. *Une extension du lemme de Fatou.* Bull. Sci. Math. 1991. V. 115, №2. P. 211–221. [*308*]

[523] Klimkin V.M. Introduction to the theory of set functions. Saratov. Gos. Univ., Kuibyshev, 1989; 210 pp. (in Russian). [*423, 433*]

[524] Kluvánek I., Knowles G. Vector measures and control systems. North-Holland, Amsterdam – Oxford; American Elsevier, New York, 1976; ix+180 pp. [*423*]

[525] Kodaira S., Kakutani S. *A nonseparable translation-invariant extension of the Lebesgue measure space.* Ann. Math. (2). 1950. V. 52. P. 574–579. [*81*]

[526] Kolmogoroff (Kolmogorov) A. *La définition axiomatique de l'intégrale.* C. R. Acad. Sci. Paris. 1925. T. 180. P. 110–111 (English transl.: *Axiomatic definition of the integral*, [**535**, V. 1, p. 13–14]). [*435*]

[527] Kolmogoroff A. *Sur les bornes de la généralisation de l'intégrale.* Manuscript dated by 1925, English transl.: *On the limits of the generalization of the integral*, [**535**, V. 1, p. 15–32]). [*435*]

[528] Kolmogorov A.N. *General measure theory and probability calculus.* Trudy Komm. Akad. Matem. 1929. T. 1. P. 8–21 (in Russian); English transl.: [**535**, V. 2, p. 48–59]. [*419*]

[529] Kolmogoroff A. *Untersuchungen über den Integralbegriff.* Math. Ann. 1930. B. 103. S. 654–696; English transl.: *Studies on the concept of integral*, [**535**, V. 1, p. 100–143]. [*435*]

[530] Kolmogoroff A. *Ueber Kompaktheit der Funktionenmengen bei der Konvergenz im Mittel.* Nachr. Ges. Wiss. Göttingen. 1931. B. 9. S. 60–63; English transl.: *On the compactness of sets of functions in the case of convergence in mean*, [**535**, V. 1, p. 147–150]. [*434*]

[531] Kolmogoroff A. *Beiträge zur Maßtheorie.* Math. Ann. 1933. B. 107. S. 351–366; English transl.: *On measure theory*, [**535**, V. 1, p. 161–180]. [*248*]

[532] Kolmogoroff A. Grundbegriffe der Wahrscheinlichkeitsrechnung. Springer, Berlin, 1933; iv+62 S. English transl.: Foundations of the Theory of Probability. Chelsea, New York, 1950; viii+71 pp. [*411, 419, 429*]

[533] Kolmogorov A.N. Lectures on a course of Analysis–III. Moscow Univ., Moscow, 1946–1947 (in Russian). [*412, 418*]

[534] Kolmogorov A.N. *Towards exposition of the foundations of Lebesgue's measure theory.* Uspehi Mat. Nauk. 1950. V. 5, №1. P. 211–213 (in Russian). [*412*]

[535] Kolmogorov A.N. Selected works. V. I,II. Kluwer Academic Publ., Dordrecht, 1991, 1992. [*409, 419, 435*]

[536] Kolmogorov A.N., Fomin S.V. Introductory real analysis. V. 1. Metric and normed spaces. V. 2. Measure. The Lebesgue integral. Hilbert space. Transl. from the 2nd Russian ed. Corr. repr. Dover, New York, 1975; xii+403 pp.; German ed.: Reelle Funktionen und Funktionalanalysis. VEB Deutscher Verlag Wiss., Berlin, 1975 (Russian ed.: Elements of the theory of functions and functional analysis, Moscow, 1954, 1960). [*vi, 62, 65, 67, 412, 424*]

[537] Kölzow D. Differentiation von Maßen. Lecture Notes in Math. V. 65. Springer-Verlag, Berlin – New York, 1968; 102 S. [*438*]

[538] Komlós J. *A generalization of a problem of Steinhaus.* Acta Math. Acad. Sci. Hungar. 1967. V. 18. P. 217–229. [*290*]

[539] König H. Measure and integration. An advanced course in basic procedures and applications. Springer–Verlag, Berlin, 1997; xxii+260 pp. [*422*]

[540] Königsberger K. Analysis 2. 4e Aufl. Springer, Berlin, 2002; xii+459 S. [*414*]

[541] Korevaar J. Mathematical methods. V. I: linear algebra, normed spaces, distributions, integration. Academic Press, London, 1968; x+505 pp. [*414*]

[542] Körner T.W. *Besicovitch via Baire.* Studia Math. 2003. V. 158, №1. P. 65–78. [*66*]

[543] Kostelyanec P.O., Rešetnyak Yu.G. *The determination of a completely additive function by its values on half-spaces.* Uspehi Matem. Nauk. 1954. V. 9, №3. P. 135–140 (in Russian). [*228*]

[544] Kovan'ko A.S. The Lebesgue integral. Lvov. Univ., Lvov, 1951; 204 pp. (in Russian). [*414, 423*]

[545] Kowalsky H.-J. Vektoranalysis. II. Walter de Gruyter, Berlin – New York, 1976; 251 S. [*414*]

[546] Krasnosel'skiĭ M.A., Rutickiĭ Ja.B. Convex functions and Orlicz spaces. Translated from the first Russian ed. Noordhoff, Groningen, 1961; xi+249 pp. (Russian ed.: Moscow, 1958). [*320, 400, 435*]

[547] Krée P. Intégration et théorie de la mesure. Une approche géométrique. Cours et exercices corrigés. Ellipses, Paris, 1997; x+211 pp. [*414*]

[548] Krieger H.A. Measure-theoretic probability. University Press of America, Lanham, 1980; xi+382 pp. [*414*]

[549] Kripke B. Introduction to analysis. W.H. Freeman, San Francisco – London, 1968; viii+274 pp. [*414*]

[550] Krugova E.P. *On differentiability of convex measures.* Matem. Zametki. 1995. V. 57, №6. P. 51–61 (in Russian); English transl.: Math. Notes. 1995. V. 58, №6. P. 1294–1301. [*378*]

[551] Kudryavtsev (Kudryavcev) L.D. *On p-variation of mappings and summability of powers of the Radon–Nikodym derivative.* Uspehi Mat. Nauk. 1955. V. 10, №2. P. 167–174 (in Russian). [*381, 435*]

[552] Kudryavtsev L.D., Kaščenko Yu.D. *On change of variable in the Lebesgue integral.* Dokl. Akad. Nauk SSSR. 1952. V. 84, №5. P. 869–871 (in Russian). [*437*]

[553] Kudryavtsev L.D., Kutasov A.D., Chehlov V.I., Shabunin M.I. A collection of problems in mathematical analysis. Functions of several variables. Nauka, Fizmatlit, Moscow, 1995; 496 pp. (in Russian). [*415*]

[554] Kuller R.G. Topics in modern analysis. Prentice-Hall, Englewood Cliffs, 1969; vii+296 p. [*414*]

[555] Kuratowski K., Mostowski A. Set theory. With an introduction to descriptive set theory. 2nd ed. North-Holland, Amsterdam – New York – Oxford; Polish Sci. Publ., Warsaw, 1976; xiv+514 pp. [*78, 79*]

[556] Kurtz D.S., Swartz C.W. Theories of integration. The integrals of Riemann, Lebesgue, Henstock–Kurzweil, and McShane. World Sci., River Edge, 2004; xiii+268. [*437*]

[557] Kurzweil J. *Generalized ordinary differential equations and continuous dependence on a parameter.* Czech. Math. J. 1957. V. 7(82). P. 418–446. [*436*]

[558] Kurzweil J. Nichtabsolut konvergente Integrale. B.G. Teubner Verlagsgesellschaft, Leipzig, 1980; 184 S. [*437*]

[559] Kurzweil J. Integration between the Lebesgue integral and the Henstock–Kurzweil integral. Imperial College Press, London, 2002; 140 pp. [*437*]

[560] Kusraev A.G., Malyugin S.A. Some problems in the theory of vector measures. Akad. Nauk SSSR Sibirsk. Otdel., Inst. Mat., Novosibirsk, 1988; 182 pp. (in Russian). [*423*]

[561] Kuttler K. Basic analysis. Rinton Press, Princeton, 2001; xiv+541 pp. [*414*]

[562] Laamri I.H. Mesures, intégration, convolution et transformée de Fourier des fonctions. Dunod, Paris, 2001; xii+338 pp. [*415*]

[563] Lacey H.E. The isometric theory of classical Banach spaces. Springer-Verlag, Berlin – New York, 1974; x+270 pp. [*421*]

[564] Lacey M.T. *Carleson's theorem: proof, complements, variations.* Publ. Mat. 2004. V. 48, №2. P. 251–307. [*260*]

[565] Lahiri B.K., Roy K.C. Real analysis. World Press, Calcutta, 1988; xvi+399 pp. [*414*]

[566] Lamperti J.W. Probability. A survey of the mathematical theory. 2nd ed. John Wiley & Sons, New York, 1996; xii+189 pp. [*vii*]

[567] Lang S. Real analysis. 2nd ed. Addison-Wesley, Reading, Massachusetts, 1983; xv+533 pp. [*414*]

[568] Lang S. Real and functional analysis. 3d ed. Springer, New York, 1993; xiv+580 pp. [*414*]

[569] Larman D.G. *On the exponent of convergence of a packing of spheres.* Mathematika. 1966. V. 13. P. 57–59. [*91*]

[570] Larman D.G. *On packings of unequal spheres in* R^n. Canad. J. Math. 1968. V. 20. P. 967–969. [*422*]

[571] Larman D.G., Rogers C.A. *The existence of a centrally symmetric convex body with central sections that are unexpectedly small*. Mathematika. 1975. V. 22, №2. P. 164–175. [*215*]

[572] la Vallée Poussin Ch.J. de. Cours d'analyse infinitésimale. T. 1, 2. 2e éd. Paris, 1909, 1912; xi+423 pp. , xi+464 pp. (3me éd.: Louvain – Paris, 1914). [*410, 421*]

[573] la Vallée Poussin Ch.J. de. *Sur l'intégrale de Lebesgue*. Trans. Amer. Math. Soc. 1915. V. 16. P. 435–501. [*410, 428*]

[574] la Vallée Poussin Ch.J. de. Intégrales de Lebesgue, fonctions d'ensemble, classes de Baire. Gauthier-Villars, Paris, 1916; 154 pp. (2e éd., 1934, 193 pp.). [*410*]

[575] la Vallée Poussin Ch.J. de. Œuvres scientifiques. V. I–III. Académie Royale de Belgique, Brussels; Circolo Matematico di Palermo, Palermo, 2000–2004. [*409*]

[576] Lax P. Theory of functions of a real variable. Courant Inst. Math. Sci., New York Univ., New York, 1959; xi+154 pp. [*414*]

[577] Leader S. The Kurzweil–Henstock integral and its differentials. Marcel Dekker, New York – Basel, 2001; 355 pp. [*437*]

[578] Lebesgue H. *Sur une généralisation de l'intégrale définie*. C. R. Acad. Sci. Paris. 1901. T. 132. P. 1025–1028. [*416, 423*]

[579] Lebesgue H. *Intégrale, longueur, aire*. Ann. Mat. Pura Appl. (3). 1902. T. 7. P. 231–359. [*416, 423*]

[580] Lebesgue H. *Sur une propriété des fonctions*. C. R. Acad. Sci. 1903. T. 137, №26. P. 1228–1230. [*426, 432*]

[581] Lebesgue H. *Sur l'existence des dérivées*. C. R. Acad. Sci. Paris. 1903. T. 136. P. 659–661. [*436*]

[582] Lebesgue H. Leçons sur l'intégration et la recherche des fonctions primitives. Paris, Gauthier-Villars, 1904; 2e éd., Paris, 1928; xii+340 pp. [*410, 416, 423, 436*]

[583] Lebesgue H. *Sur les fonctions représentables analytiquement*. J. Math. Pures Appl. (6). 1905. V. 1. P. 139–216. [*420*]

[584] Lebesgue H. Leçons sur les séries trigonométriques. Paris, 1906; 138 pp. [*416, 426*]

[585] Lebesgue H. *Sur les fonctions dérivées*. Atti Accad. Lincei Rend. Cl. sci. fis., mat. e natur. Roma. (5). 1906. V. 15, Sem. 2, f. 1. P. 3–18. [*436*]

[586] Lebesgue H. *Encore une observation sur les fonctions dérivées*. Atti Accad. Lincei Rend. Cl. sci. fis., mat. e natur. Roma. (5). 1907. V. 16, Sem. 1. P. 92–100. [*436*]

[587] Lebesgue H. *Sur la recherche des fonctions primitives par l'intégration*. Atti Accad. Lincei Rend. Cl. sci. fis., mat. e natur. Roma. (5). 1907. V. 16. P. 283–290. [*391, 416*]

[588] Lebesgue H. *Sur la méthode de M. Goursat pour la résolution de l'équation de Fredholm*. Bull. Math. Soc. France. 1908. V. 36. P. 3–19. [*428*]

[589] Lebesgue H. *Sur les intégrales singulières*. Ann. Fac. Sci. Univ. Toulouse (3). 1909. V. 1. P. 25–117. [*416, 433, 434*]

[590] Lebesgue H. *Sur les suites de fonctions mesurables*. C. R. Acad. Sci. Paris. 1909. T. 149. P. 102–103. [*426*]

[591] Lebesgue H. *Sur l'intégration des fonctions discontinues*. Ann. Sci. Ecole Norm. Sup. 1910. V. 27. P. 361–450. [*416, 418, 429, 436, 437*]

[592] Lebesgue H. *Sur l'intégrale de Stieltjes et sur les opérations fonctionnelles linéaires*. C. R. Acad. Sci. Paris. 1910. T. 150. P. 86–88. [*425*]

[593] Lebesgue H. *Remarques sur les théories de la mesure et de l'intégration*. Ann. Sci. Ecole Norm. Sup. 1918. V. 35. P. 191–250. [*416, 422*]

[594] Lebesgue H. Œuvres scientifiques. V. 1–5. L'Enseignement Mathématique, Université de Genève, Genève, 1972. [*409, 416*]

[595] Lebesgue H. Lettres d'Henri Lebesgue à Émile Borel. With an afterword by Bernard Bru and Pierre Dugac. Cahiers du Séminaire d'Histoire des Mathématiques, 12, P. 1–511, Univ. Paris VI, Paris, 1991. [*416, 427*]

[596] Lebesgue H. Les lendemains de l'intégrale. Lettres à Émile Borel. Éd. par B. Bru B. et P. Dugac. Vuibert, Paris, 2004; 240 pp. [*416, 427*]

[597] Ledoux M. The concentration of measure phenomenon. Amer. Math. Soc., Providence, Rhode Island, 2001; x+181 pp. [*431*]

[598] Lee J.R. Advanced calculus with linear analysis. Academic Press, New York – London, 1972; xiv+218 pp. [*414*]

[599] Lee P.Y., Výborný R. Integral: an easy approach after Kurzweil and Henstock. Cambridge University Press, Cambridge, 2000; xii+311 pp. [*437*]

[600] Lehmann E.L. Testing statistical hypotheses. John Wiley & Sons, New York, Chapman & Hall, London, 1959; xiii+369 pp. [*412, 434*]

[601] Leichtweiss K. Konvexe Mengen. Springer-Verlag, Berlin – New York, 1980; 330 S. [*431*]

[602] Leinert M. Integration und Maß. Friedr. Vieweg & Sohn, Braunschweig, 1995; xii+208 S. [*414*]

[603] Lembcke J. *On a measure extension theorem of Bierlein.* Lecture Notes in Math. 1980. V. 794. P. 45–48. [*421*]

[604] Leont'eva T.A., Panferov V.S., Serov V.S. Problems in the theory of functions of a real variable. Moskov. Gos. Uinv., Moscow, 1997; 208 pp. (in Russian). [*415*]

[605] Letac G. Exercises and solutions manual for "Integration and probability" by Paul Malliavin. Springer-Verlag, New York, 1995; viii+142 pp. [*415*]

[606] Letta G. Teoria elementare dell'integrazione. Editrice Tecnico Scientifica, Pisa, 1973; vii+204 pp. (rev. ed.: Boringhieri, Torino, 1976; 289 pp.) [*414*]

[607] Levi B. *Sopra l'integrazione delle serie.* Rend. Ist. Lombardo (2). 1906. T. 39. P. 775–780. [*428*]

[608] Levi B. *Ricerche sopra le funzioni derivata.* Atti Accad. Lincei Rend. Cl. sci. fis., mat. e natur. Roma. (5). 1906. V. 15, Sem. 1, f. 12. P. 674–684. [*436, 438*]

[609] Levi B. *Ancora alcune osservazioni sulle funzioni derivate.* Atti Accad. Lincei Rend. Cl. sci. fis., mat. e natur. Roma. (5). 1906. V. 15, Sem. 2, f. 6. P. 358–368. [*436*]

[610] Lévy P. Théorie de l'addition des variables aléatoires. Gautier-Villars, Paris, 1937 (2e éd., 1954; 385 pp.). [*419*]

[611] Lichtenstein L. *Ueber die Integration eines bestimmten Integrals in Bezug auf einen Parameter.* Nachr. Ges. Wiss. Göttingen, Math.-phys. Kl. 1910. S. 468–475. [*234*]

[612] Lieb E.H., Loss M. Analysis. 2nd ed. Amer. Math. Soc., Providence, Rhode Island, 2001; 346 pp. [*214, 325, 431*]

[613] Liese F., Vajda I. Convex statistical distances. BSB B.G. Teubner Verlag, Leipzig, 1987; 224 pp. [*154*]

[614] Lindenstrauss J., Tzafriri L. Classical Banach spaces. V. I,II. Springer, Berlin – New York, 1977, 1979; xiii+190 pp. , x+243 pp. [*433*]

[615] Lipecki Z. *On strongly additive set functions.* Colloq. Math. 1971. V. 22. P. 255–256. [*422*]

[616] Lipecki Z. *Unique extensions of positive set functions.* Arch. Math. 1983. B. 41. S. 71–79. [*61*]

[617] Loève M. Probability theory. Van Nostrand, Toronto – New York – London, 1955; xv+515 pp.; (4th ed.: v. 1,2. Springer, New York, 1977, 1978; xvii+425 pp. , xvi+413 pp.) [*vi, 412*]

[618] Łojasiewicz S. An introduction to the theory of real functions. Wiley, New York, 1988; ix+230 pp. [*414*]

[619] Łomnicki Z., Ulam S. *Sur la théorie de la mesure dans les espaces combinatoires et son application au calcul des probabilités I. Variables indépendentes.* Fund. Math. 1934. V. 23. P. 237–278. [*419, 430*]

[620] Lorentz G.G. *Who discovered analytic sets?* Math. Intelligencer. 2001. V. 23, №4. P. 28–32. [*420*]

[621] Łoś J., Marczewski E. *Extensions of measure.* Fund. Math. 1949. V. 36. P. 267–276. [*421*]

[622] Lösch F. Höhere Mathematik: eine Einführung für Studierende und zum Selbststudium (von H. Mangoldt, K. Knopp). B. 4: Mengenlehre, Lebesguesches Maß und

Integral, Topologische Räume, Vektorräume, Funktionalanalysis, Integralgleichungen. 4e Aufl. S. Hirzel, Stuttgart, 1990; xv+612 S. [*414*]
[623] Lovász L., Simonovits M. *Random walks in a convex body and an improved volume algorithm.* Random Structures Algorithms. 1993. V. 4, №4. P. 359–412. [*173*]
[624] Lozinskiĭ S.M. *On the indicatrix of Banach.* Vestnik Leningrad. Univ. 1958. V. 13, №7. P. 70–87 (in Russian). [*406*]
[625] Lubotzky A. Discrete groups, expanding graphs and invariant measures. Birkhäuser, Basel – Boston, 1994; 195 pp. [*82*]
[626] Lucia P. de. Funzioni finitamente additive a valori in un gruppo topologico. Quaderni dell'Unione mat. ital., 29. Pitagora Ed., Bologna, 1985; viii+187 pp. [*423*]
[627] Lucia P. de, Pap E. *Convergence theorems for set functions.* P. 125–178 in [**738**]. [*433*]
[628] Lukacs E. Characteristic functions. 2nd ed. Hafner Publ., New York, 1970; x+350 pp. [*241, 430*]
[629] Lukacs E. Developments in characteristic function theory. Griffin, London, 1983; viii+182 pp. [*430*]
[630] Lukes J., Malý J. Measure and integral. Matfyzpress, Prague, 1995; 178 pp. [*414*]
[631] Lusin N. *Sur les propriétés des fonctions mesurables.* C. R. Acad. Sci. Paris. 1912. T. 154. P. 1688–1690. [*426*]
[632] Lusin (Luzin) N.N. *On the fundamental theorem of the integral calculus.* Mat. Sbornik. 1912. V. 28. P. 266–294 (in Russian). [*332, 402, 426*]
[633] Lusin N.N. Integral and trigonometric series. Moscow, 1915 (2nd ed.: Moscow, 1951; 551 pp.; published also in Mat. Sbornik. 1916. V. 30, №1. P. 1–242; in Russian). [*332, 388, 400, 402, 410, 437*]
[634] Lusin N. *Sur la classification de M. Baire.* C. R. Acad. Sci. Paris. 1917. T. 164. P. 91–94. [*420*]
[635] Lusin N. *Sur la notion de l'intégrale.* Ann. Mat. Pura Appl. (3). 1917. V. 26. P. 77–129. [*332, 402*]
[636] Lusin N.N. The theory of functions of a real variable. 2nd ed. Učpedgiz, Moscow, 1948; 319 pp. (in Russian). [*414*]
[637] Lusin N.N. Collected works. V. 1–3. Fizmatgiz, Moscow, 1953, 1958, 1959 (in Russian). [*409*]
[638] Lusin N., Sierpiński W. *Sur quelques propriétés des ensembles (A).* Bull. Acad. Sci. Cracovie, Sér. A. 1918. №4–5a. P. 35–48. [*420*]
[639] Luther N.Y. *Unique extension and product measure.* Canad. J. Math. 1967. V. 19. P. 757–763. [*99, 236*]
[640] Luther N.Y. *A decomposition of measures.* Canad. J. Math. 1968. V. 20. P. 953–959. [*99*]
[641] Luukkainen J., Saksman E. *Every complete doubling metric space carries a doubling measure.* Proc. Amer. Math. Soc. 1998. V. 126, №2. P. 531–534. [*376*]
[642] MacNeille H.M. *A unified theory of integration.* Proc. Nat. Acad. Sci. U.S.A. 1941. V. 27. P. 71–76. [*162, 424*]
[643] Magyar Z. The Lebesgue integral. Akadémiai Kiadó, Budapest, 1997; x+106 pp. [*414*]
[644] Makarov B.M., Goluzina M.G., Lodkin A.A., Podkorytov A.N. Selected problems in real analysis. Translated from the Russian. Amer. Math. Soc., Rhode Island, Providence, 1992; 370 pp. (Russian ed.: Moscow, 1992). [*415*]
[645] Malik S.C., Arora S. Mathematical analysis. 2nd ed. Wiley, New York, 1992; xii+903 pp. [*414*]
[646] Malliavin P., Airault H., Kay L., Letac G. Integration and probability. Springer-Verlag, New York – Berlin, 1995; 322 pp. [*414*]
[647] Mallory D. *Extension of set functions to measures and applications to inverse limit measures.* Canad. Math. Bull. 1975. V. 18, №4. P. 547–553. [*52*]

[648] Marcinkiewicz J., Salem R. *Sur les sommes riemanniennes.* Compositio Math. 1940. T. 7. P. 376–389. [*435*]
[649] Marcinkiewicz J., Zygmund A. *Mean values of trigonometrical polynomials.* Fund. Math. 1937. V. 28. P. 131–166. [*435*]
[650] Marczewski E. *On compact measures.* Fund. Math. 1953. V. 40. P. 113–124. [*419*]
[651] Marczewski E. *Remarks on the convergence of measurable sets and measurable functions.* Colloq. Math. 1955. V. 3. P. 118–124. [*100, 165*]
[652] Marczewski E. Collected mathematical papers. PAN, Warszawa, 1996. [*409*]
[653] Marczewski E., Steinhaus H. *On a certain distance of sets and the corresponding distance of functions.* Colloq. Math. 1958. V. 6. P. 319–327. [*102*]
[654] Margulis G.A. *Some remarks on invariant measures.* Monatsh. Math. 1980. B. 90, №3. S. 233–235. [*422*]
[655] Margulis G.A. *Finitely additive invariant measures on Euclidean spaces.* Ergodic Theory Dynam. Syst. 1982. V. 2. P. 383–396. [*81*]
[656] Marle C.-M. Mesures et probabilités. Hermann, Paris, 1974; 474 pp. [*414*]
[657] Martin D.A. *A simple proof that determinancy implies Lebesgue measurability.* Rend. Sem. Mat. Univ. Politec. Torino. 2003. V. 61, №4. P. 393–397. [*80*]
[658] Mattila P. Geometry of sets and measures in Euclidean spaces. Cambridge University Press, Cambridge, 1995; 343 pp. [*436, 437*]
[659] Mauldin R.D. *Countably generated families.* Proc. Amer. Math. Soc. 1976. V. 54. P. 291–297. [*61*]
[660] Maurin K. Analysis. Part II. Integration, distributions, holomorphic functions, tensor and harmonic analysis. Reidel, Dordrecht, 1980; xvii+829 pp. [*414*]
[661] Mawhin J. Analyse. Fondements, techniques, évolution. De Boeck Université, Brussels, 1992; iv+808 pp. [*414, 437*]
[662] Mayrhofer K. Inhalt und Maß. Wien, 1952; vii+265 S. [*414*]
[663] Maz'ja V.G. Sobolev spaces. Translated from the Russian. Springer-Verlag, Berlin, 1985; xix+486 pp. (Russian ed.: Leningrad, 1985). [*379*]
[664] Mazurkiewicz S. *Sur les fonctions qui satisfont à la condition (N).* Fund. Math. 1930. V. 16. P. 348–352. [*391*]
[665] McCann R.J. *Existence and uniqueness of monotone measure-preserving maps.* Duke Math. J. 1995. V. 80. P. 309–323. [*382*]
[666] McDonald J.N., Weiss N.A. A course in real analysis. Biographies by C.A. Weiss. Academic Press, San Diego, California, 1999; xx+745 pp. [*414, 415*]
[667] McLeod R.M. The generalized Riemann integral. Mathematical Association of America, Washington, D.C., 1980; xiii+275 pp. [*437*]
[668] McShane E.J. Integration. Princeton, 1944; viii+394 pp. [*411*]
[669] McShane E.J. Unified integration. Academic Press, New York – London, 1983; xiii+607 pp. [*414*]
[670] McShane E.J., Botts T.A. Real analysis. Van Nostrand, Princeton, 1959; ix+272 pp. [*414*]
[671] Medeiros L.A., Mello E.A. de. A integral de Lebesgue. Inst. matem., Univ. Rio de Janeiro, Rio de Janeiro, 1985; 157 pp. [*414*]
[672] Medvedev F.A. Development of set theory in the nineteenth century. Nauka, Moscow, 1965; 232 pp. (in Russian). [*416*]
[673] Medvedev F.A. Development of the concept of the integral. Nauka, Moscow, 1974; 423 pp. (in Russian). [*416, 417, 419, 423, 425, 437*]
[674] Medvedev F.A. Scenes from the history of real functions. Translated from the Russian. Birkhäuser Verlag, Basel, 1991; 265 pp. (Russian ed.: Moscow, 1975). [*416, 417, 423*]
[675] Medvedev F.A. The French school of the theory of functions and sets on the border of the XIX–XX centuries. Nauka, Moscow, 1976; 232 pp. (in Russian). [*416, 417, 423*]

[676] Medvedev F.A. *Letters of D.Th. Egoroff to N.N. Lusin.* Istor.–Matem. Issled. 1980. V. 25. P. 335–361 (in Russian). [*427*]
[677] Medvedev F.A. The early history of the axiom of choice. Nauka, Moscow, 1982; 304 pp. (in Russian). [*416*]
[678] Mejlbro L., Topsøe F. *Vitali systems in* $\mathbf{R}^n$ *with irregular sets.* Math. Scand. 1996. V. 79, №1. P. 86–100. [*438*]
[679] Melnikov M.S. *Dependence of volume and diameter of sets in an n-dimensional Banach space.* Uspehi Mat. Nauk. 1963. V. 18, №4. P. 165–170 (in Russian). [*214*]
[680] Menchoff D. *Sur la représentation conforme des domaines plans.* Math. Ann. 1926. B. 95, №5. S. 641–670. [*390*]
[681] Menshov D.E. *Reminiscences of the early years and the beginning of the Moscow school of the theory of functions.* Istor.-Matem. Issled. 1983. №27. P. 312–333 (in Russian). [*416*]
[682] Mergelyan S.N. *Uniform approximations of functions of a complex variable.* 1952. V. 7, №2. P. 31–122 (in Russian); English transl.: Amer. Math. Soc. Translations №101, 1954. [*91*]
[683] Message d'un mathématicien: Henri Lebesgue. Pour le centenaire de sa naissance. Introductions et extraits choisis par Lucienne Félix. Librairie Scientifique et Technique Albert Blanchard, Paris, 1974; ix+259 pp. [*416*]
[684] Métivier M. Notions fondamentales de la théorie des probabilités: maîtrises de mathématiques. 2e éd., Dunod, Paris, 1972; xi+334 pp. [*414*]
[685] Meyer M., Reisner S., Schmuckenschläger M. *The volume of the intersection of a convex body with its translates.* Mathematika. 1993. V. 40, №2. P. 278–289. [*246*]
[686] Meyer P.-A. Probability and potentials. Blaisdell, Waltham – Toronto – London, 1966; xiii+266 pp. [*415*]
[687] Miamee A.G. *The inclusion* $L^p(\mu) \subseteq L^q(\nu)$. Amer. Math. Monthly. 1991. V. 98, №4. P. 342–345. [*310*]
[688] Michel A. Constitution de la théorie moderne de l'intégration. Mathesis. Librairie Philosophique J. Vrin, Paris, 1992; 338 pp. [*416, 417, 423*]
[689] Michel H. Maß- und Integrationstheorie. I. VEB Deutscher Verlag der Wissenschaften, Berlin, 1978; 204 S. [*414*]
[690] Mikusiński J. *Sur une définition de l'intégrale de Lebesgue.* Bull. Acad. Polon. Sci. 1964. V. 12. P. 203–204. [*162, 424*]
[691] Mikusiński J. The Bochner integral. Birkhäuser Verlag, Basel – Stuttgart, 1978; xii+233 pp. [*414, 424*]
[692] Miller H.I. *A theorem connected with results of Steinhaus and Smítal.* J. Math. Anal. Appl. 1987. V. 124. P. 27–32. [*403*]
[693] Misiewicz J.K., Scheffer C.L. *Pseudo isotropic measures.* Nieuw Arch. Wisk. (4). 1990. V. 8, №2. P. 111–152. [*431*]
[694] Mitrinović D.S., Pečarić J.E., Fink A.M. Classical and new inequalities in analysis. Kluwer Academic Publ., Dordrecht, 1993; xviii+740 pp. [*429*]
[695] Monfort A. Cours de probabilités. Economica, Paris, 1980; 210 pp. [*414*]
[696] Moore E.H., Smith H.L. *A general theory of limits.* Amer. J. Math. 1922. V. 44. P. 102–121. [*435*]
[697] Morgan F. Geometric measure theory. A beginner's guide. 3d ed. Academic Press, San Diego, California, 2000; 226 pp. [*437*]
[698] Morse A.P. *Convergence in variation and related topics.* Trans. Amer. Math. Soc. 1937. V. 41. P. 48–83. [*344*]
[699] Morse A.P. *Perfect blankets.* Trans. Amer. Math. Soc. 1947. V. 61. P. 418–442. [*436*]
[700] Moser J. *On the volume elements on a manifold.* Trans. Amer. Math. Soc. 1965. V. 120. P. 286–294. [*382*]
[701] Mozzochi C.J. *On a Riemann sum construction of Rudin.* Proc. Amer. Math. Soc. 1969. V. 22. P. 718. [*435*]

[702] Mozzochi C.J. On the pointwise convergence of Fourier series. Lecture Notes in Math. V. 199. Springer-Verlag, Berlin – New York, 1971; vii+87 pp. [*260*]
[703] Mukherjea A., Pothoven K. Real and functional analysis. Plenum Press, New York, 1978; x+529 pp. [*414*]
[704] Muldowney P. A general theory of integration in function spaces, including Wiener and Feynman integration. Longman Scientific, Wiley, New York, 1987; viii+115 pp. [*437*]
[705] Munroe M.E. Introduction to measure and integration. Addison–Wesley, Cambridge, Massachusetts, 1953; x+310 pp. (2nd ed.: Measure and integration. Addison–Wesley, Reading, Massachusetts – London – Don Mills, 1971; xii+290 pp.) [*412, 421*]
[706] Natanson I.P. Foundations of the theory of functions of a real variable. Leningrad. Univ., Leningrad, 1941; 296 pp. (in Russian). [*411, 412*]
[707] Natanson I.P. Theory of functions of a real variable. GTTI, Moscow – Leningrad, 1950; 399 pp. (in Russian); English transl.: V. 1,2. Frederick Ungar Publ., New York, 1955, 1961; 277 pp., 265 pp. [*vi, 62, 149, 400, 406, 412, 436, 437*]
[708] Natterer F. The mathematics of computerized tomography. B.G. Teubner, Stuttgart; John Wiley & Sons, New York, 1986; x+222 pp. [*227*]
[709] Nekrasov V.L. Structure and measure of linear point domains. Tomsk Technol. Inst., Tomsk, 1907; viii+254 pp. (in Russian). [*410*]
[710] Neumann J. von. Functional Operators. I. Measures and Integrals. Annals of Mathematics Studies, no. 21. Princeton University Press, Princeton, 1950; vii+261 pp. [*411, 429*]
[711] Neumann J. von. Collected works. V. 1–4. Pergamon Press, New York, 1961. [*409*]
[712] Neumann J. von. Invariant measures. Amer. Math. Soc., Providence, Rhode Island, 1999; xvi+134 pp. [*82*]
[713] Neveu J. Bases mathématiques du calcul des probabilités. Masson et Cie, Paris, 1964; xiii+203 pp. English transl.: Mathematical foundations of the calculus of probability. Holden-Day, San Francisco, 1965; 231 pp. [*vi, 414*]
[714] Nielsen O.A. An introduction to integration and measure theory. John Wiley & Sons, New York, 1997; xvi+473 pp. [*320, 414*]
[715] Nikodym O.[2] *Sur la mesure des ensembles plans dont tous les points sont rectilinéairement accessibles.* Fund. Math. 1927. V. 10. P. 116–168. [*421*]
[716] Nikodym O. *Sur une propriété de la mesure généralisée des ensembles.* Prace Matem.–Fiz. 1928–1929. T. 36. P. 65–71. [*89*]
[717] Nikodym O. *Sur les fonctions d'ensemble.* Comptes Rendus du I Congrès des Math. des Pays Slaves, Varsovie, 1929 (1930), pp. 304–313. [*229, 421*]
[718] Nikodym O. *Sur une généralisation des intégrales de M. J. Radon.* Fund. Math. 1930. V. 15. P. 131–179. [*418, 429*]
[719] Nikodym O. *Contribution à la théorie des fonctionnelles linéaires en connexion avec la théorie de la mesure des ensembles abstraits.* Mathematica (Cluj). 1931. V. 5. P. 130–141. [*306, 431*]
[720] Nikodym O. *Sur les suites de fonctions parfaitement additives d'ensembles abstraits.* C. R. Acad. Sci. Paris. 1931. T. 192. P. 727. [*433*]
[721] Nikodym O. *Sur les familles bornées de fonctions parfaitement additives d'ensemble abstrait.* Monatsh. Math. Phys. 1933. B. 40. S. 418–426. [*433*]
[722] Nikodym O. *Sur les suites convergentes de fonctions parfaitement additives d'ensemble abstrait.* Monatsh. Math. Phys. 1933. B. 40. S. 427–432. [*433*]
[723] Nikodym O. *Sur la mesure vectorielle parfaitement additive dans un corps abstrait de Boole.* Mem. Acad. Roy. Belg. 1938. T. 17, F. 7. P. 1–40. [*418*]
[724] Nikodym O. *Sur l'existence d'une mesure parfaitement additive et non séparable.* Mem. Acad. Roy. Belg. 1938. T. 17, F. 8. P. 1–29. [*421*]

[2] Another (later) spelling: Nikodým O.M.

[725] Ochan Yu.S. Collection of problems and theorems in the theory of functions of a real variable. Prosveščenie, Moscow, 1965; 231 pp. (in Russian). [*415*]

[726] Ochan Yu.S. The integral. Moskov. Pedagog. Inst., Moscow, 1973; 313 pp. (in Russian). [*437*]

[727] Ochan Yu.S. Collection of problems in mathematical analysis. Prosveščenie, Moscow, 1981; 271 pp. (in Russian). [*415*]

[728] Oden J.T., Demkowicz L.F. Applied functional analyis. CRC Press, Boca Raton, 1996; xiv+653 pp. [*414*]

[729] Okikiolu G.O. Aspects of the theory of bounded integral operators in L^p-spaces. Academic Press, London – New York, 1971; ix+522 pp. [*414, 430, 436*]

[730] Olevskiĭ A.M. Fourier series with respect to general orthogonal systems. Springer-Verlag, New York – Heidelberg, 1975; viii+136 pp. [*261*]

[731] Os C.H. van. Moderne integraalrekening. Noordhoff, Groningen, 1925; 204 pp. [*411*]

[732] Osserman R. *The isoperimetric inequality.* Bull. Amer. Math. Soc. 1978. V. 84, №6. P. 1182–1238. [*379*]

[733] Oxtoby J.C. Measure and category. A survey of the analogies between topological and measure spaces. Springer-Verlag, New York – Berlin, 1971; viii+95 pp. [*93, 414*]

[734] Pallu de la Barrière R. Intégration. Ellipses, Paris, 1997; 208 pp. [*414*]

[735] Panchapagesan T.V. Medida e integración. Parte I. Teoría de la medida. V. I, II. Universidad de los Andes, Facultad de Ciencias, Departamento de Matemáticas, Mérida, 1991; xviii+344 pp., 345–708 pp. [*414*]

[736] Pannikov B.V. *An example of the divergence of Riemann sums.* Sibirsk. Mat. Zh. 1988. V. 29, №3. P. 208–210 (in Russian); English transl.: Sib. Math. J. 1989. V. 29, №3. P. 503–506. [*435*]

[737] Pap E. Null-additive set functions. Kluwer Academic Publ., Dordrecht, 1995; xii+315 pp. [*423*]

[738] Pap E. (ed.) Handbook of measure theory. V. 1,2. North-Holland, Amsterdam, 2002. [*415*]

[739] Parthasarathy K.R. Introduction to probability and measure. Springer-Verlag, New York, Macmillan India, 1978; xii+312 pp. (2nd rev. ed.: Hindustan Book Agency, New Delhi, 2005; 338 pp.) [*vi, 414*]

[740] Paul S. Die Moskauer mathematische Schule um N.N. Lusin: Entstehungs- und Entwicklungsgeschichte, Arbeitsprinzipien, Zerfall; unter besonderer Berücksichtigung der Kampagne gegen Lusin im Sommer 1936. Kleine, Bielefeld, 1997; 238 S. [*416*]

[741] Peano G. Applicazioni geometriche del calcolo infinitesimale. Torino, 1887; xii+334 pp. [*2*]

[742] Pedersen G.K. Analysis now. Springer-Verlag, New York, 1989; xiv+277 pp. [*414*]

[743] Pesin I.N. Classical and modern integration theories. Translated from the Russian. Academic Press, New York — London, 1970; xviii+195 pp. (Russian ed.: Moscow, 1966). [*416, 417, 423, 437*]

[744] Pesin Y.B. Dimension theory in dynamical systems. Contemporary views and applications. University of Chicago Press, Chicago, 1997; xii+304 pp. [*421*]

[745] Pettis J. *On the extension of measures.* Ann. Math. 1951. V. 54. P. 186–197. [*422*]

[746] Pfanzagl J., Pierlo W. Compact systems of sets. Lecture Notes in Math. V. 16. Springer-Verlag, Berlin – New York, 1966; ii+48 pp. [*419*]

[747] Pfeffer W.F. Integrals and measures. Marcel Dekker, New York, 1977; ix+259 pp. [*414*]

[748] Pfeffer W.F. The Riemann approach to integration. Local geometric theory. Cambridge University Press, Cambridge, 1993; xvi+302 pp. [*437*]

[749] Pfeffer W.F. Derivation and integration. Cambridge University Press, Cambridge, 2001; xvi+266 pp. [*437*]

[750] Phillips E.R. *Nicolai Nicolaevich Luzin and the Moscow school of the theory of functions.* Hist. Math. 1978. V. 5. P. 275–305. [*416*]

[751] Phillips E.R. An introduction to analysis and integration theory. Corr. repr. of the 1971 ed. Dover Publications, New York, 1984; xxviii+452 pp. [*414*]

[752] Phillips R.S. *On linear transformations.* Trans. Amer. Math. Soc. 1940. V. 48. P. 516–541. [*303*]

[753] Picone M., Viola T. Lezione sulla teoria moderna dell'integrazione. Einaudi, Torino, 1952; 404 pp. [*414*]

[754] Pier J.-P. *Intégration et mesure 1900–1950.* Development of mathematics 1900–1950. Ed. by J.-P. Pier. Birkhäuser, Boston – Basel, 1994, pp. 517–564. [*417, 423*]

[755] Pier J.-P. Histoire de l'intégration. Vingt-cinq siècles de mathématiques. Culture Scientifique. Masson, Paris, 1996; xii+307 pp. [*416, 417, 423*]

[756] Pierpont J. Lectures on the theory of functions of real variables. V. 1,2. Ginn & Co., Boston, 1905, 1912; xvi+250 pp., xiii+645 pp. [*410*]

[757] Pinsker M.S. Information and information stability of random variables and processes. Translated from the Russian. Holden-Day, San Francisco – London – Amsterdam, 1964; xii+243 pp. (Russian ed.: Moscow, 1960). [*155*]

[758] Pisier G. The volume of convex bodies and Banach space geometry. Cambridge University Press, Cambridge, 1989; xvi+250 pp. [*431*]

[759] Pitt H.R. Integration, measure and probability. Oliver & Boyd, Edinburgh – London, 1963; viii+110 pp. [*414*]

[760] Pitt H.R. Measure and integration for use. Oxford University Press, London – New York, 1985; xii+143 pp. [*414*]

[761] Plancherel M. *Contribution à l'étude de la répresentation d'une fonction arbitraire par des intégrales définies.* Rend. Circ. Mat. Palermo. 1910. T. 30. P. 289–335. [*430*]

[762] Plancherel M. *Démonstration du théorème de Riesz–Fischer et du théorème de Weyl sur les suites convergentes en moyenne.* Bull. Sci. Math (2). 1923. V. 47. P. 195–204. [*430*]

[763] Polischuk E.M. Émile Borel (1871–1956). Nauka, Leningrad, 1980; 168 pp. (in Russian). [*416*]

[764] Pollard D. A user's guide to measure theoretic probability. Cambridge University Press, Cambridge, 2002; xiv+351 p. [*414*]

[765] Ponomarev S.P. *On the N-property of homeomorphisms of the class W^1_p.* Sibirsk. Mat. Zh. 1987. V. 28, №2. P. 140–148 (in Russian); English transl.: Sib. Math. J. 1987. V. 28, №2. P. 291–298. [*382*]

[766] Poroshkin A.G. The theory of measure and integral. Syktyvkar. Gos. Univ., Syktyvkar, 1996; 172 pp. (in Russian). [*414, 420*]

[767] Possel R. de. *Sur la dérivation abstraite des fonctions d'ensembles.* J. Math. Pures Appl. 1936. V. 15. P. 391–409. [*438*]

[768] Pratt J.W. *On interchanging limits and integrals.* Ann. Math. Statist. 1960. V. 31. P. 74–77. [*428*]

[769] Preiss D. *Geometry of measures in R^n: distribution, rectifiability, and densities.* Ann. Math. (2). 1987. V. 125, №3. P. 537–643. [*437*]

[770] Priestley H.A. Introduction to integration. Clarendon Press, Oxford University Press, New York, 1997; xii+306 pp. [*414*]

[771] Pták P., Tkadlec J. *A note on determinacy of measures.* Časopis Pěst. Mat. 1988. V. 113, №4. P. 435–436. [*244*]

[772] Pták V. *A combinatorial lemma on the existence of convex means and its application to weak compactness.* Proc. Sympos. Pure Math., V. VII, pp. 437–450. Amer. Math. Soc., Providence, Rhode Island, 1963. [*90*]

[773] Pugachev V.S., Sinitsyn I.N. Lectures on functional analysis and applications. Translated from the Russian. World Sci., River Edge, 1999; xx+730 pp. (Russian ed.: Moscow, 1996). [*414*]

[774] Pugh C.C. Real mathematical analysis. Springer-Verlag, New York, 2002; xii+437 pp. [*414*]

[775] Rademacher H. *Über eine Eigenschaft von meßbaren Mengen positiven Maßes*. Jahr. Deutsch. Math. Verein. 1921. B. 30. S. 130–132. [*85*]

[776] Radó T. Length and area. Amer. Math. Soc., New York, 1948; v+572 pp. [*437*]

[777] Radó T., Reichelderfer P. *A theory of absolutely continuous transformations in the plane*. Trans. Amer. Math. Soc. 1941. V. 49. P. 258–307. [*102*]

[778] Radon J. *Theorie und Anwendungen der absolut additiven Mengenfunktionen*. Sitz. Akad. Wiss. Wien, Math.-naturwiss. Kl. IIa. 1913. B. 122. S. 1295–1438. [*418, 419, 425, 429, 431, 434, 437*]

[779] Radon J. *Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten*. Ber. Verh. Königl. Sächs. Ges. Wiss. Leipzig. 1917. B. 69. S. 262–277. [*227*]

[780] Radon J. Gesammelte Abhandlungen, V. 1,2. Birkhäuser, Basel – Boston, 1987. [*409*]

[781] Ramachandran B. Advanced theory of characteristic functions. Statist. Publ. Soc., Calcutta, 1967; vii+208 pp. [*430*]

[782] Rana I.K. An introduction to measure and integration. 2nd ed. Amer. Math. Soc., Rhode Island, Providence, 2002; 459 pp. [*414*]

[783] Randolph J.F. Basic real and abstract analysis. Academic Press, New York, 1968; 504 pp. [*414*]

[784] Rao B.V. *On discrete Borel spaces and projective sets*. Bull. Amer. Math. Soc. 1969. V. 75. P. 614–617. [*211*]

[785] Rao K.P.S. Bhaskara, Rao B.V. *On extensions of measures*. Colloq. Math. 1987. V. 53, №1. P. 43–47. [*422*]

[786] Rao K.P.S. Bhaskara, Rao M. Bhaskara. Theory of charges. A study of finitely additive measures. Academic Press, New York – London, 1983; x+315 pp. [*99, 422, 423*]

[787] Rao M.M. Probability theory with applications. Academic Press, New York, 1984; 495 pp. [*414*]

[788] Rao M.M. Measure theory and integration. John Wiley & Sons, New York, 1987; xiv+540 pp. (2nd ed.: Marcel Dekker, New York, 2004; xx+761 pp.). [*242, 312, 320, 397, 414, 423*]

[789] Ray W.O. Real analysis. Prentice Hall, Englewood Cliffs, New Jersey, 1988; xii+307 pp. [*414*]

[790] Reshetnyak Yu.G. *Property N for the space mappings of class* $W^1_{n,\mathrm{loc}}$. Sibirsk. Mat. Zh. 1987. V. 28, №5. P. 149–153 (in Russian); English transl.: Sib. Math. J. 1987. V. 28, №5. P. 810–813. [*382*]

[791] Revuz D. Mesure et intégration. Hermann, Paris, 1997. [*414*]

[792] Rey Pastor J. Elementos de la teoria de funciones. 3 ed. Ibero-Americana, Madrid – Buenos Aires, 1953; 560 pp. [*414*]

[793] Richard U. Teoria delle funzioni. Ist. Matematico Del Politecnico di Torino, Torino, 1951; viii+367 pp. [*414*]

[794] Richter H. Wahrscheinlichkeitstheorie. 2e Aufl. Springer, Berlin, 1966; xii+462 S. [*414*]

[795] Ridder J. *Integration in abstrakten Räumen*. Fund. Math. 1935. V. 24. P. 72–117. [*419*]

[796] Riečan B., Neubrunn T. Integral, measure, and ordering. Kluwer Academic Publ., Dordrecht; Ister Science, Bratislava, 1997; xiv+378 pp. [*423*]

[797] Riesz F. *Sur les systèmes orthogonaux de fonctions*. C. R. Acad. Sci. Paris. 1907. T. 144. P. 615–619. [*431*]

[798] Riesz F. *Sur une espèce de géométrie analytique des systèmes de fonctions sommables*. C. R. Acad. Sci. Paris. 1907. T. 144. P. 1409–1411. [*431*]

[799] Riesz F. *Sur les suites de fonctions mesurables*. C. R. Acad. Sci. Paris. 1909. T. 148. P. 1303–1305. [*426*]

[800] Riesz F. *Sur les opérations fonctionnelles linéaires.* C. R. Acad. Sci. Paris. 1909. T. 149. P. 974–977. [*425*]

[801] Riesz F. *Untersuchungen über Systeme integrierbarer Funktionen.* Math. Ann. 1910. B. 69. S. 449–497. [*424, 431*]

[802] Riesz F. *Sur certains systèmes singuliers d'équations intégrales.* Ann. École Norm. Sup. (3). 1911. V. 28. P. 33–62. [*430*]

[803] Riesz F. *Sur quelques points de la théorie des fonctions sommables.* C. R. Acad. Sci. Paris. 1912. T. 154. P. 641–643. [*423*]

[804] Riesz F. *Sur intégrale de Lebesgue.* Acta Math. 1920. V. 42. P. 191–205. [*423*]

[805] Riesz F. *Sur la convergence en moyenne, I,II.* Acta Sci. Math. 1928–29. T. 4. P. 58–64, 182–185. [*434*]

[806] Riesz F. *Über Sätze von Stone und Bochner.* Acta Univ. Szeged. 1933. V. 6. P. 184–198. [*430*]

[807] Riesz F. *Sur l'intégrale de Lebesgue comme l'opération inverse de la dérivation.* Ann. Scuola Norm. Super. Pisa (2). 1936. V. 5. P. 191–212. [*386*]

[808] Riesz F. Œuvres complètes, V. I,II. Acad. Sci., Budapest, 1960. [*409*]

[809] Riesz F., Sz.-Nagy B. Leçons d'analyse fonctionnelle. Académie des Sciences de Hongrie, Akadémiai Kiadó, Budapest, 1952; viii+449 pp. English transl.: Functional analysis. Dover, New York, 1990; xii+504 pp. [*163, 412*]

[810] Riesz M. *Sur les ensembles compacts de fonctions sommables.* Acta Sci. Math. Szeged. 1933. V. 6. P. 136–142. [*295, 434*]

[811] Rinow W. *Über die Eindeutigkeit der Caratheodoryschen Erweiterung eines Maßes.* Math. Nachr. 1969. B. 42. S. 289–294. [*421*]

[812] Rivière T., Ye D. *Resolutions of the prescribed volume from equation.* Nonlinear Diff. Eq. Appl. 1996. V. 3. P. 323–369. [*382*]

[813] Rogers C. Packing and covering. Cambridge University Press, New York, 1964; viii+111 pp. [*422*]

[814] Rogers C.A. Hausdorff measures. Cambridge University Press, London, 1970; viii+179 pp. [*430*]

[815] Rogers C.A., Sion M. *On Hausdorff measures in topological spaces.* Monatsh. Math. 1963. B. 67. S. 269–278. [*430*]

[816] Rogosinski W.W. Volume and integral. Oliver and Boyd, Edinburgh and London; Interscience Publishers, New York, 1952; x+160 pp. [*414*]

[817] Rohlin V.A. Selected works. Moscow Centre of Continuous Mathematical Education, Moscow, 1999 (in Russian). [*409*]

[818] Romanovski P. *Intégrale de Denjoy dans les espaces abstraits.* Mat. Sbornik USSR. 1941. T. 9 (51). №1. P. 67–120. [*437*]

[819] Romero J.L. *When is $L^p(\mu)$ contained in $L^q(\mu)$?* Amer. Math. Monthly. 1983. V. 90. P. 203–206. [*310*]

[820] Rooij A.C.M. van, Schikhof W.H. A second course on real functions. Cambridge University Press, Cambridge, 1982; xiii+200 pp. [*406, 414, 436*]

[821] Rosenblatt J. *Uniqueness of invariant means for measure-preserving transformations.* Trans. Amer. Math. Soc. 1981. V. 265, №2. P. 623–636. [*422*]

[822] Rosenthal A. *Beiträge zu Carathéodory's Messbarkeitstheorie.* Nachr. Ges. Wiss. Göttingen. 1916. S. 305–321. [*418, 421*]

[823] Rosenthal A. *Neuere Untersuchungen über Funktionen reeller Veränderlichen.* Nach den Referaten von L. Zoretti, P. Montel und M. Fréchet bearbeiten von A. Rosenthal. Encyklopädie der Mathematischen Wissenschaften. T. III, H. 2. S. 851–1187. Verlag, B.G. Teubner, Leipzig, 1923–1927. [*410, 415, 416*]

[824] Rosenthal H.P. *On relatively disjoint families of measures, with some applications to Banach space theory.* Studia Math. 1970. V. 37. P. 13–36. [*303*]

[825] Rosenthal J.S. A first look at rigorous probability theory. World Sci., Singapore, 2000; xiv+177 pp. [*414*]

[826] Ross K.A., Stromberg K. *Jessen's theorem on Riemann sums for locally compact groups*. Pacific J. Math. 1967. V. 20. P. 135–147. [*435*]
[827] Rotar V.I. Probability theory. Translated from the Russian. World Sci., River Edge, 1997; xviii+414 pp. (Russian ed.: Moscow, 1992). [*414*]
[828] Roussas G.G. An introduction to measure-theoretic probability. Amsterdam, Elsevier, Academic Press, New York, 2004; xviii+443 pp. [*414*]
[829] Royden H.L. Real analysis. 3d ed., Prentice Hall, Englewood Cliffs, New Jersey, 1988; 444 pp. (1st ed.: Macmillan, 1963). [*vi, 414*]
[830] Rubel L.A. *A pathological Lebesgue-measurable function*. J. London Math. Soc. 1963. V. 38. P. 1–4. [*401*]
[831] Ruch J.-J., Weber M. *Quelques résultats dans la théorie des sommes riemanniennes*. Exposition. Math. 1997. V. 15, №3. P. 279–288. [*435*]
[832] Ruckle W.H. Modern analysis. Measure theory and functional analysis with applications. PWS-Kent Publ., Boston, 1991; xv+265 pp. [*414*]
[833] Rudin W. *An arithmetic property of Riemann sums*. Proc. Amer. Math. Soc. 1964. V. 15. P. 321–324. [*314*]
[834] Rudin W. Principles of mathematical analysis. 2nd ed. McGraw-Hill, New York, 1964; ix+270 pp. [*138, 414*]
[835] Rudin W. Real and complex analysis. 2nd ed., McGraw-Hill, New York, 1974; xii+452 pp. [*414*]
[836] Ruziewicz S. *Remarque à la note de M. Banach "Sur une classe de fonctions continues"*. Fund. Math. 1926. V. 8. P. 173–174. [*390*]
[837] Saadoune M., Valadier M. *Extraction of a "good" subsequence from a bounded sequence of integrable functions*. J. Convex Anal. 1995. V. 2, №1-2. P. 345–357. [*299*]
[838] Sadovnichiĭ V.A. Theory of operators. Translated from the second Russian edition. Consultants Bureau, New York, 1991; xii+396 pp. (Russian ed.: Moscow, 1986). [*414*]
[839] Sadovnichiĭ V.A., Grigor'yan A.A., Konyagin S.V. Problems from students mathematical competitions. Moskov. Gos. Univ., Moscow, 1987; 312 pp. (in Russian). [*172*]
[840] Saks S. Théorie de l'intégrale. Warszawa, 1933; vii+290 pp. 2nd ed. (cor. and complem.): Theory of the integral. Warszawa, 1937; xv+343 pp. [*332, 372, 392, 411, 437*]
[841] Saks S. *On some functionals*. Trans. Amer. Math. Soc. 1933. V. 35, №2. P. 549–556; Addition: ibid., №4. P. 965–970. [*276, 433*]
[842] Salem R., Zygmund A. *Some properties of trigonometric series whose terms have random signs*. Acta Math. 1954. V. 91. P. 245–301. [*142*]
[843] Samuélidès M., Touzillier L. Analyse fonctionnelle. Cépaduès Éditions, Toulouse, 1989; iv+289 pp. [*414*]
[844] Sansone G., Merli L. Lezioni sulla teoria delle funzioni di variabili reali. Cedam, Padova, 1964; 236 pp. [*414*]
[845] Sarason D. *Functions of vanishing mean oscillation*. Trans. Amer. Math. Soc. 1975. V. 207. P. 391–405. [*174*]
[846] Saxe K. Beginning functional analysis. Springer-Verlag, New York, 2002; xii+197 pp. [*414*]
[847] Saxena S.Ch., Shah S.M. Introduction to real variable theory. 2nd ed. Prentice-Hall of India, New Delhi, 1980; xv+334 pp. [*414*]
[848] Schaefer H.H. Topological vector spaces. Springer-Verlag, Berlin – New York, 1971; xi+294 pp. [*281*]
[849] Schauder J.P. *The theory of surface measure*. Fund. Math. 1926. V. 8. P. 1–48. [*437*]
[850] Schechtman G., Schlumprecht Th., Zinn J. *On the Gaussian measure of the intersection of symmetric, convex sets*. Ann. Probab. 1998. V. 26. P. 346–357. [*239*]
[851] Scheffé H. *A useful convergence theorem for probability distributions*. Ann. Math. Statist. 1947. V. 18. P. 434–438. [*428*]

[852] Schilling R. Measures, integrals and martingales. Cambridge University Press, Cambridge, 2005; 380 pp. [*414*]

[853] Schlesinger L., Plessner A. Lebesguesche Integrale und Fouriersche Reihen. De Gruyter, Berlin – Leipzig, 1926; 229 S. [*411*]

[854] Schmetterer L. Einführung in die mathematische Statistik. 2e Aufl. Springer-Verlag, Wien – New York, 1966; ix+597 S. [*412*]

[855] Schmitz N. Einführung in die Maß- und Integrationstheorie. Gesellschaft zur Förderung der Mathematischen Statistik, Münster, 1987; 144 S. [*414*]

[856] Schmitz N., Plachky D. Vorlesungen über Wahrscheinlichkeitstheorie. I. Verlag Anton Hain, Meisenheim am Glan, 1976; vi+216 S. [*414*]

[857] Schneider R. Convex bodies: the Brunn–Minkowski theory. Cambridge University Press, 1993; xiii+490 pp. [*431*]

[858] Schönflies A. Entwicklung der Mengenlehre und ihrer Anwendungen. Leipzig, Berlin, 1913; 388 S. [*410*]

[859] Schwartz L. Analyse mathématique. V. I,II. Hermann, Paris, 1967; xxix+830 pp., xxiv+571 pp. [*414*]

[860] Segal I.E. *The span of the translations of a function in a Lebesgue space*. Proc. Nat. Acad. Sci. U.S.A. 1944. V. 30. P. 165–169. [*327*]

[861] Segal I.E. *Equivalences of measure spaces*. Amer. J. Math. 1951. V. 73. P. 275–313. [*312*]

[862] Segal I., Kunze R.A. Integrals and operators. Springer, Berlin – New York, 1978; xiv+371 pp. [*414*]

[863] Severini C. *Sopra gli sviluppi in serie di funzioni ortogonali*. Atti Acc. Gioenia di Catania (5). 1910. T. 3, Mem. XI. P. 1–7. [*426*]

[864] Severini C. *Sulle successioni di funzioni ortogonali*. Atti Acc. Gioenia di Catania (5). 1910. T. 3, Mem. XIII. P. 1–10. [*426*]

[865] Shilov G.Ye. Mathematical analysis. A special course. Translated from the Russian. Pergamon Press, Oxford – New York – Paris, 1965; xii+481 pp. (Russian ed.: Moscow, 1961). [*414*]

[866] Shilov G.E. Mathematical analysis. Functions of several real variables. Nauka, Moscow, 1972; 624 pp. (in Russian). [*437*]

[867] Shilov G.E., Gurevich B.L. Integral, measure and derivative: a unified approach. Translated from the Russian. Dover Publ., New York, 1977; xiv+233 pp. (2nd Russian ed.: Moscow, 1967). [*397, 414, 438*]

[868] Shiryaev A.N. Probability. Translated from the Russian. Springer-Verlag, New York, 1996; xvi+623 pp. (3nd Russian ed.: Moscow, 2004). [*vi, 414*]

[869] Sierpiński W. *Sur une propriété générale des ensembles de points*. C. R. Acad. Sci. Paris. 1916. T. 162. P. 716–717. [*428*]

[870] Sierpiński W. *Sur la question de la mesurabilité de la base de M. Hamel*. Fund. Math. 1920. V. 1. P. 105–111. [*422*]

[871] Sierpiński W. *Sur les fonctions convexes mesurables*. Fund. Math. 1920. V. 1. P. 125–129. [*428*]

[872] Sierpiński W. *Sur un problème concernant les ensembles mesurables superficiellement*. Fund. Math. 1920. V. 1. P. 112–115. [*232*]

[873] Sierpiński W. *Sur les rapports entre l'existence des intégrales* $\int_0^1 f(x,y)dx$, $\int_0^1 f(x,y)dy$ *et* $\int_0^1 dx \int_0^1 f(x,y)dy$. Fund. Math. 1920. V. 1. P. 142–147. [*232*]

[874] Sierpiński W. *Sur les fonctions d'ensembles additives et continues*. Fund. Math. 1922. V. 3. P. 240–246. [*422*]

[875] Sierpiński W. *Sur l'ensemble de distances entre les points d'un ensemble*. Fund. Math. 1925. V. 7. P. 144–148. [*422*]

[876] Sierpiński W. *Sur les ensembles boreliens abstraits*. Ann. Soc. Polon. Math. 1927. T. 6. P. 50–53. [*419*]

[877] Sierpiński W. *Un théorème général sur les familles des ensembles*. Fund. Math. 1928. V. 12. P. 206–210. [*419*]

[878] Sierpiński W. *Sur une propriété de fonctions quelconques d'une variable réelle*. Fund. Math. 1935. V. 25. P. 1–4. [*395*]
[879] Sierpiński W. Hypothèse du continu. 2nd ed. Chelsea, New York, 1956; xvii+274 pp. [*78, 422*]
[880] Sierpiński W. On the congruence of sets and their equivalence by finite decomposition. New York, 1967; 115 pp. [*82*]
[881] Sierpiński W. Œuvres choisies. T. I–III. PWN-éditions Scientifiques de Pologne, Warsaw, 1974, 1975, 1976. [*409, 419, 422*]
[882] Sikorski R. Boolean algebras. Springer-Verlag, Berlin – Heidelberg, 1960; ix+176 pp. [*421*]
[883] Sikorski R. Advanced calculus: functions of several variables. Polish Sci. Publ., Warszawa, 1969; 460 pp. [*414*]
[884] Simon L. Lectures on geometric measure theory. Centre for Mathematical Analysis, Australian National University, Canberra, 1983; vii+272 pp. [*437*]
[885] Simonnet M. Measures and probabilities. Springer-Verlag, New York, 1996; xiv+510 pp. [*414*]
[886] Sion M. Introduction to the methods of real analysis. Holt, Rinehart and Winston, New York – Montreal – London, 1968; x+134 pp. [*414, 423*]
[887] Sion M. A theory of semigroup valued measures. Lecture Notes in Math. V. 355. Springer-Verlag, Berlin – New York, 1973; iv+140 pp. [*423*]
[888] Sion M., Willmott R.C. *Hausdorff measures on abstract spaces*. Trans. Amer. Math. Soc. 1966. V. 123. P. 275–309. [*430*]
[889] Slutsky E. *Über stochastische Asymptoten und Grenzwerte*. Metron. 1925. V. 5. P. 3–89. [*171, 426*]
[890] Smiley M.F. *An extension of metric distributive lattices with an application in general analysis*. Trans. Amer. Math. Soc. 1944. V. 56, №3. P. 435–447. [*422*]
[891] Smirnov V.I. A course of higher mathematics. V. 5. [Integration and functional analysis]. Translated from the Russian. Pergamon Press, Oxford – New York; Addison-Wesley, Reading, Mass.– London, 1964; xiv+635 pp.; Smirnow W.I. Lehrgang der höheren Mathematik. Teil V. Zweite, berichtigte Auflage. VEB Deutscher Verlag der Wissenschaften, Berlin, 1967; xiii+570 S. (1st Russian ed.: Moscow, 1947; 2nd ed.: Moscow, 1959). [*412, 426, 435*]
[892] Smith H.J.S. *On the integration of discontinuous functions*. Proc. London Math. Soc. 1875. V. 6. P. 140–153. [*419*]
[893] Sobolev S.L. Introduction to the theory of cubature formulae. Nauka, Moscow, 1974; 808 pp. (in Russian). [*325*]
[894] Sobolev V.I. Lectures on supplementary chapters of mathematical analysis. Nauka, Moscow, 1968; 288 pp. (in Russian). [*414*]
[895] Sodnomov B. *On a property of sets of positive measure*. Fund. Math. 1967. V. 60. P. 187–190. [*87*]
[896] Sohrab H.H. Basic real analysis. Birkhäuser Boston, Boston, 2003; xiv+559 pp. [*414*]
[897] Solovay R. *A model of set theory in which every set of reals is Lebesgue measurable*. Ann. Math. 1970. V. 92. P. 1–56. [*80*]
[898] Solovay R. *Real-valued measurable cardinals*. In: Axiomatic Set Theory (Proc. Sympos. Pure Math., V. XIII, Part I), pp. 397–428. Amer. Math. Soc., Providence, Rhode Island, 1971. [*80*]
[899] Souslin M. *Sur une définition des ensembles mesurables B sans nombres transfinis*. C. R. Acad. Sci. Paris. 1917. T. 164, №2. P. 89–91. [*420*]
[900] Spiegel M.R. Theory and problems of real variables: Lebesgue measure and integration with applications to Fourier series. MacGraw-Hill, New York, 1964; 194 pp. [*414*]
[901] Sprecher D.A. Elements of real analysis. 2nd ed. Dover Publications, New York, 1987; viii+343 pp. [*414*]

[902] Srinivasan T.P. *On measurable sets.* J. Indian Math. Soc. (N.S.) 1954. V. 18. P. 1–8. [*420, 421*]

[903] Srinivasan T.P. *On extensions of measures.* J. Indian Math. Soc. (N.S.) 1955. V. 19. P. 31–60. [*419*]

[904] Stampacchia G. *Sulla successioni di funzioni continue rispetto ad una variabile e misurabili rispetto ad un'altra.* Atti Accad. Lincei. Rend. Cl. sci. fis., mat. e natur. Roma. (8). 1949. V. 6. P. 198–201. [*160*]

[905] Stein E. Singular integrals and differentiability properties of functions. Princeton University Press, Princeton, 1970; xiv+290 pp. [*367, 379, 386, 436*]

[906] Stein E.M. Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals. Princeton University Press, Princeton, 1993; xiii+695 pp. [*65, 66, 353, 367, 374, 375, 398, 436*]

[907] Stein E.M., Shakarchi R. Real analysis. Measure theory, integration, and Hilbert spaces. Princeton University Press, Princeton, 2005; xx+402 pp. [*414*]

[908] Stein E., Weiss G. Introduction to Fourier analysis on Euclidean spaces. Princeton University Press, Princeton, 1971; x+297 pp. [*238, 320, 430, 431, 435, 436*]

[909] Steinhaus H. *Additive und stetige Funktionaloperationen.* Math. Z. 1919. B. 5. S. 186–221. [*431*]

[910] Steinhaus H. *Sur les distances des points des ensembles de mesure positive.* Fund. Math. 1920. V. 1. P. 93–104. [*85, 100*]

[911] Steinhaus H. *Les probabilités dénombrables et leur rapport à la théorie de la mesure.* Fund. Math. 1923. V. 4. P. 286–310. [*430*]

[912] Stepanoff W. *Sur une propriété caractéristique des fonctions mesurables.* Mat. Sbornik. 1924. V. 31. P. 487–489. [*438*]

[913] Stieltjes T.J. *Recherches sur les fractions continues.* Ann. Fac. Sci. Toulouse (1). 1894. V. 8. P. J.1–J.122. [*425*]

[914] Stone M.H. Linear transformations in Hilbert space and their applications to analysis. Amer. Math. Soc. Colloq. Publ., V. 15, New York, 1932; viii+622 pp. [*411*]

[915] Stromberg K. *The Banach–Tarskii paradox.* Amer. Math. Monthly. 1979. V. 86 (March). P. 151–161. [*81*]

[916] Stromberg K. An introduction to classical real analysis. Wadsworth, Belmont, 1981; ix+575 pp. [*414*]

[917] Stroock D.W. A concise introduction to the theory of integration. 3d ed. Birkhäuser Boston, Boston, 1999; xiv+253 pp. [*414*]

[918] Subramanian B. *On the inclusion* $L^p(\mu) \subset L^q(\mu)$. Amer. Math. Monthly. 1978. V. 85, №6. P. 479–481. [*310*]

[919] Sudakov V.N. *Criteria of compactness in function spaces.* Uspehi Mat. Nauk. 1957. V. 12, №3. P. 221–224 (in Russian). [*318, 434*]

[920] Suetin P.K. Classical orthogonal polynoms. Nauka, Moscow, 1976; 328 pp. (in Russian). [*261*]

[921] Sullivan D. *For* $n > 3$ *there is only one finitely additive rotationally invariant measure on the* n*-sphere defined on all Lebesgue measurable sets.* Bull. Amer. Math. Soc. 1981. V. 4. P. 121–123. [*422*]

[922] Sun Y. *Isomorphisms for convergence structures.* Adv. Math. 1995. V. 116, №2. P. 322–355. [*237*]

[923] Svetic R.E. *The Erdos similarity problem: a survey.* Real Anal. Exchange. 2000–2001. V. 26 (2). P. 525–539. [*422*]

[924] Swartz Ch. Measure, integration and function spaces. World Sci., River Edge, 1994; xii+277 pp. [*319, 414*]

[925] Swartz Ch. Introduction to gauge integrals. World Sci., Singapore, 2001; x+157 pp. [*353, 437*]

[926] Sz.-Nagy B. Introduction to real functions and orthogonal expansions. Oxford University Press, New York, 1965; xi+447 pp. [*414*]

[927] Szpilrajn E. *Sur certains invariants de l'opération (A)*. Fund. Math. 1933. V. 21. P. 229–235. [*420*]
[928] Szpilrajn E. *Sur l'extension de la mesure lebesguienne*. Fund. Math. 1935. V. 25. P. 551–558. [*80*]
[929] Szymanski W. *Who was Otto Nikodym?* Math. Intell. 1990. V. 12, №2. P. 27–31. [*416*]
[930] Tagamlickiĭ Ya.A. *On the integration of sequences of functions*. Dokl. Akad. Nauk SSSR. 1947. V. 57, №1. P. 17–19 (in Russian). [*321*]
[931] Talagrand M. Pettis integral and measure theory. Mem. Amer. Math. Soc. 1984. V. 51, №307. P. 1–224. [*235*]
[932] Talagrand M. *Maharam's problem*. http://arxiv.org: math.FA/0601689. P. 1–30. [*75*]
[933] Tarski A. *Une contribution à la théorie de la mesure*. Fund. Math. 1930. V. 15. P. 42–50. [*422*]
[934] Taylor A.E. General theory of functions and integration. 2d ed. Dover Publ., New York, 1985; xvi+437 pp. (1st ed. Blaisdell, 1965). [*414*]
[935] Taylor A.E. *A study of Maurice Fréchet. I–III.* Arch. Hist. Exact Sci. 1982. V. 27, №3. P. 233–295; ibid., 1985. V. 34, №4. P. 279–380; ibid., 1987. V. 37, №1. P. 25–76. [*416*]
[936] Taylor A.E., Dugac P. *Quatre lettres de Lebesgue à Fréchet*. Rev. Histoire Sci. Appl. 1981. V. 34, №2. P. 149–169. [*416, 432*]
[937] Taylor J.C. An introduction to measure and probability. Springer-Verlag, New York, 1997; xviii+299 pp. [*414*]
[938] Taylor S.J. Introduction to measure and integration. Cambridge University Press, London, 1973; 266 pp. [*414*]
[939] Telyakovskiĭ S.A. A collection of problems in the theory of functions of a real variable. Nauka, Moscow, 1980; 112 pp. (in Russian). [*415*]
[940] Temple G. The structure of Lebesgue integration theory. Clarendon Press, Oxford, 1971; 184 pp. [*414*]
[941] Ter Horst H.J. *Riemann–Stieltjes and Lebesgue–Stieltjes integrability*. Amer. Math. Monthly. 1984. V. 91. P. 551–559. [*428*]
[942] Thielman H. Theory of functions of a real variable. Prentice-Hall, Englewood Cliffs, New Jersey, 1953; 209 pp. [*414*]
[943] Thomson B.S. Real functions. Lecture Notes in Math. V. 1170. Springer-Verlag, Berlin – New York, 1985; vii+229 pp. [*404, 436*]
[944] Thomson B.S. *Differentiation*. P. 179–247 in [**738**]. [*438*]
[945] Tikhomirov V.M. *The discovery of A-sets*. Istor.-Matem. Issled. 1993. №34. P. 129–139 (in Russian). [*420*]
[946] Titchmarsh E.C. *The convergence of certain integrals*. Proc. London Math. Soc. (2). 1926. V. 24. P. 347–358. [*401*]
[947] Titchmarsh E. The theory of functions. Clarendon Press, Oxford, 1932; x+454 pp. [*308, 394, 411*]
[948] Titchmarsh E.C. Introduction to the theory of Fourier integrals. 3d ed. Chelsea, New York, 1986; x+394 pp. [*430, 431*]
[949] Tkadlec J. *Construction of a finite Borel measure with σ-porous sets as null sets*. Real Anal. Exchange. 1986–1987. №1. P. 349–353. [*404*]
[950] Tolstoff (Tolstov) G.P. *Une remarque sur le théorème de D.Th. Egoroff*. Dokl. Akad. Nauk SSSR. 1939. V. 22, №6. P. 309–311 (in Russian); French transl.: C. R. (Dokl.) Acad. Sci. URSS. 1939. V 22. P. 305–307. [*159*]
[951] Tolstoff G. *Sur les points de densité des ensembles linéaires mesurables*. Rec. Math. [Mat. Sbornik] N.S. 1942. V. 10. P. 249–264 (in Russian; French summary). [*407*]
[952] Tolstov G.P. *On change of variables in multiple integrals*. Uspehi Mat. Nauk. 1950. V. 5, №4. P. 162–169 (in Russian). [*402*]
[953] Tolstow (Tolstov) G.P. Maß und Integral. Translated from the Russian. Akademie-Verlag, Berlin, 1981; 296 S. (Russian ed.: Moscow, 1976). [*414*]

[954] Tonelli L. *Sull'integrazione per parti.* Rend. Accad. Lincei Rend. Cl. sci. fis., mat. e natur. Roma. (5). 1909. V. 18. P. 246–253. [*429*]

[955] Tonelli L. *Sulla nozione di integrale.* Ann. Mat. Pura Appl. (4). 1923. T. 1. P. 105–145. [*423*]

[956] Tonelli L. Opere scelte. V. 1. Edizioni Cremonese, Roma, 1960. [*409*]

[957] Topsøe F. *On construction of measures.* Proc. Conf. Topology and Measure. I (Zinnowitz, 1974), part 2, pp. 343–381, Ernst-Moritz-Arndt-Universität, Greifswald, 1978. [*421*]

[958] Toralballa L.V. Theory of functions. Charles Merill Books, Ohio, 1963; 671 pp. [*414*]

[959] Torchinsky A. Real variable methods in harmonic analysis. Academic Press, New York, 1986; 462 pp. [*436*]

[960] Torchinsky A. Real variables. Addison-Wesley, New York, 1988; 403 pp. [*414*]

[961] Tornier E. Wahrscheinlichkeitsrechnung und allgemeine Integrationstheorie. Teubner, Leipzig, 1936; 160 S. [*411*]

[962] Tortrat A. Calcul des probabilités et introduction aux processus aléatoires. Masson, Paris, 1971; xiv+303 pp. [*414*]

[963] Townsend E.J. Functions of real variables. Henry Holt, New York, 1928; ix+405 pp. [*411*]

[964] Tricomi F.G. Istituzioni di analisi superiore (metodi matematici della fisica). 2 ed. CEDAM, Padua, 1970; ix+472 pp. [*414*]

[965] Tumakov I.M. Henri Leon Lebesgue (1875–1941). Nauka, Moscow, 1975; 119 pp. (in Russian). [*416, 417, 423*]

[966] Ulam S. *Concerning functions of sets.* Fund. Math. 1929. V. 14. P. 231–233. [*422*]

[967] Ulam S. *Zur Masstheorie in der allgemeinen Mengenlehre.* Fund. Math. 1930. V. 16. P. 140–150. [*77, 422*]

[968] Ulyanov P.L., Bahvalov A.N., D'yachenko M.I., Kazaryan K.S., Cifuentes P. Real analysis in exercises. Fizmatlit, Moscow, 2005; 416 pp. (in Russian). [*415*]

[969] Ursell H.D. *On the behaviour of a certain sequence of functions derived from a given one.* J. London Math. Soc. 1937. V. 12. P. 229–232. [*435*]

[970] Väisälä J. *Quasiconformal maps and positive boundary measure.* Analysis. 1989. V. 9. P. 205–216. [*382*]

[971] Vakhania N.N., Kvaratskhelia V.V. *Unconditional convergence of weakly subgaussian series in Banach spaces.* Theory Probab. Appl. 2006. [*169*]

Vallée Poussin Ch.J. de la: see la Vallée Poussin Ch.J. de [**572**]–[**575**]

[972] Van Vleck E.B. *Haskin's momental theorem and its connection with Stieltjes problem of moments.* Trans. Amer. Math. Soc. 1917. V. 18. P. 326–330. [*425*]

[973] Väth M. Integration theory. A second course. World Sci., River Edge, 2002; viii+277 pp. [*414*]

[974] Veress P. *Über Funktionenmengen.* Acta Sci. Math. Szeged. 1927. V. 3. P. 177–192. [*321, 426*]

[975] Verley J.-L. Théorie élémentaire de l'intégration. "Les cours de Sorbonne". 2e éd. Centre de docum. Université Paris-V, Paris, 1968; 185 pp. [*414*]

[976] Vestrup E.M. The theory of measures and integration. Wiley-Interscience, Hoboken, New Jersey, 2003; xviii+594 pp. [*103, 229, 414*]

[977] Vinti C. Lezioni sulla teoria dell'integrazione. Galeno, Perugia, 1988; 206 pp. [*414*]

[978] Visintin A. *Strong convergence results related to strict convexity.* Comm. Partial Differential Equations. 1984. V. 9, №5. P. 439–466. [*299*]

[979] Vitali G. *Sulle funzioni integrali.* Atti. R. Accad. sci. Torino. 1904–1905. T. 40. P. 1021–1034. [*417*]

[980] Vitali G. *Sull'integrabilità delle funzioni.* Rend. Ist. Lombardo (2). 1904. T. 37. P. 69–73. [*417*]

[981] Vitali G. *Sui gruppi di punti.* Rend. Circ. Mat. Palermo. 1904. T. 18. P. 116–126. [*417*]

[982] Vitali G. *Una proprietà delle funzioni misurabili*. Rend. R. Ist. Lombardo (2). 1905. T. 38. P. 600–603. [*426*]

[983] Vitali G. *Sul problema della misura dei gruppi di punti di una retta*. Bologna, Tipogr. Gamberini e Parmeggiani, 1905. [*419*]

[984] Vitali G. *Sulle funzioni ad integrale nullo*. Rend. Circ. Mat. Palermo. 1905. T. 20. P. 136–141. [*149*]

[985] Vitali G. *Sull'integrazione per serie*. Rend. Circ. Mat. Palermo. 1907. T. 23. P. 137–155. [*428, 433*]

[986] Vitali G. *Sui gruppi di punti e sulle funzioni di variabili reali*. Atti. R. Accad. sci. Torino. 1908. T. 43. P. 75–92. [*436, 437*]

[987] Vitali G. *Sulle funzioni continue*. Fund. Math. 1926. V. 8. P. 175–188. [*437*]

[988] Vitali G. Lezioni di analisi infinitesimale, con intercalate alcune nozioni sulla teoria della misura degli aggregati di punti e sulla teoria degli integrali di Lebesgue. La Litotipo, Padova, 1927; v+495 pp. [*411, 414*]

[989] Vitali G. Geometria nello spazio Hilbertiano. N. Zanichelli, Bologna, 1929; 283 pp. [*411*]

[990] Vitali G. Opere sull'analisi reale e complessa. Carteggio. Edizioni Cremonese, Bologna, 1984. [*409, 417*]

[991] Vitali G., Sansone G. Moderna teoria delle funzioni di variabile reale. Parte I,II (Parte I: G. Vitali, Parte II: G. Sansone). 3a ed. Zanichelli, Bologna, 1951, 1952; 222 pp., 614 pp. (1a ed.: Bologna, 1935) [*411, 426*]

[992] Vitushkin A.G. On multidimensional variations. GITTL, Moscow, 1955; 220 pp. (in Russian). [*437*]

[993] Vladimirov D.A. Boolesche Algebren. Translated from the Russian. Akademie-Verlag, Berlin, 1978; viii+245 S. (Russian ed.: Moscow 1969). Revised: Boolean algebras in analysis. Kluwer Academic Publ., Dordrecht, 2002; xxii+604 pp. [*421*]

[994] Vogel W. Wahrscheinlichkeitstheorie. Vandenhoeck & Ruprecht, Göttingen, 1970; 385 S. [*414*]

[995] Vo-Khac Kh. Mesure, intégration, convolution, & analyse de Fourier. Interprétation dans le langage des probabilités. Ellipses, Paris, 1984; 256 pp. [*414*]

[996] Vol'berg A.L., Konyagin S.V. *A homogeneous measure exists on any compactum in* $\mathbb{R}^n$. Dokl. Akad. Nauk SSSR. 1984. V. 278, №4. P. 783–786 (in Russian); English transl.: Soviet Math. Dokl. 1984. V. 30, №2. P. 453–456. [*375*]

[997] Vol'berg A.L., Konyagin S.V. *On measures with the doubling condition*. Izv. Akad. Nauk SSSR. 1987. V. 51, №3. P. 666–675 (in Russian); English transl.: Math. USSR Izv. 1988. V. 30, №3. P. 629–638. [*375*]

[998] Volcic A. Lezioni sull'integrale. 2a ed. Edizioni goliardiche, Udine, 1997; 156 pp. [*414*]

[999] Volterra V. *Sui principii del calcolo integrale*. G. Battaglini. 1881. T. 19. P. 333–372. [*425*]

[1000] Vulikh B.Z. A brief course in the theory of functions of a real variable. Nauka, Moscow, 1973; 352 pp. (in Russian); English transl.: Mir, Moscow, 1976; 356 pp. [*104, 414*]

[1001] Wagon S. The Banach–Tarski paradox. Cambridge University Press, Cambridge, 1993; xviii+253 pp. [*81, 82*]

[1002] Wagschal C. Dérivation, intégration. Hermann, Paris, 1999; viii+472 pp. [*414*]

[1003] Wagschal C. Intégration. Exercices et problèmes corrigés. Hermann, Paris, 1999; vi+130 pp. [*415*]

[1004] Walter W. Analysis 2. 5e Aufl. Springer, Berlin, 2002, xiv+408 S. [*414*]

[1005] Wang Z.Y., Klir G.J. Fuzzy measure theory. Plenum Press, New York, 1992; x+354 pp. [*423*]

[1006] Ważewski T. *Sur les ensembles mesurables*. C. R. Acad. Sci. Paris. 1923. T. 176. P. 69–70. [*418*]

[1007] Weber H. *Ein Fortsetzungssatz für gruppenwertige Maße*. Arch. Math. 1980. B. 34. S. 157–159. [*61*]

[1008] Weir A.J. Lebesgue integration and measure. Cambridge University Press, London – New York, 1973; xii+281 pp. [*414*]

[1009] Weir A.J. General integration and measure. Cambridge University Press, London – New York, 1974; xi+298 pp. [*414*]

[1010] Wesler O. *An infinite packing theorem for spheres*. Proc. Amer. Math. Soc. 1960. V. 11. P. 324–326. [*91*]

[1011] Weyl H. *Über die Konvergenz von Reihen, die nach Orthogonalfunktionen fortschreiten*. Math. Ann. 1909. B. 67. S. 225–245. [*426*]

[1012] Wheeden R.L., Zygmund A. Measure and integral. Marcel Dekker, New York – Basel, 1977; 274 pp. [*414*]

[1013] Whitney H. *Totally differentiable and smooth functions*. Pacif. J. Math. 1951. V. 1. P. 143–159. [*373*]

[1014] Widom H. Lectures on measure and integration. Van Nostrand Reinhold, New York, 1969; viii+166 pp. [*414*]

[1015] Wiener N. *Differential space*. J. Math. Phys. 1923. V. 2. P. 131–174. [*419*]

[1016] Wiener N. The Fourier integral and certain of its applications. Dover, New York, 1958; 201 pp. [*430*]

[1017] Wiener N. Collected works. V. 1–3, The MIT Press, Cambridge, 1976–1981. [*409, 419*]

[1018] Wiener N., Paley R. Fourier transforms in the complex domain. Amer. Math. Soc., New York, 1934; viii+184 pp. [*430*]

[1019] Wilcox H.J., Myers D.L. An introduction to Lebesgue integration and Fourier series. Dover Publications, New York, 1994; viii+159 pp. [*414*]

[1020] Williams D. Probability with martingales. Cambridge University Press, Cambridge, 1994; xv+251 pp. [*414*]

[1021] Williamson J.H. Lebesgue integration. Holt, Rinehart and Winston, New York, 1962; viii+117 pp. [*414*]

[1022] Wise G.L., Hall E.B. Counterexamples in probability and real analysis. Oxford University Press, New York – Oxford, 1994; xii+211 pp. [*81, 228, 395, 414*]

[1023] Wolff J. *Sur les séries* $\sum \frac{A_k}{z-\alpha_k}$. C. R. Acad. Sci. Paris. 1921. T. 173. P. 1056–1057. [*419*]

[1024] Wolff T. *Recent work connected with the Kakeya problem*. Prospects in Mathematics, pp. 129–162, Amer. Math. Soc., Providence, Rhode Island, 1999. [*66*]

[1025] Yeh J. Lectures on real analysis. World Sci., River Edge, 2000; xvi+548 pp. [*414*]

[1026] Yosida K., Hewitt E. *Finitely additive measures*. Trans. Amer. Math. Soc. 1952. V. 72. P. 46–66. [*431*]

[1027] Young G.C., Young W.H. Selected papers. Presses Polytechniques et Universitaires Romandes, Lausanne, 2000; x+870 pp. [*409, 417*]

[1028] Young W.H. *The general theory of integration*. Proc. Royal Soc. London. 1904. V. 73. P. 445–449. [*423*]

[1029] Young W.H. *Open sets and the theory of content*. Proc. London Math. Soc. (2). 1905. V. 2. P. 16–51. [*93, 417, 418, 421*]

[1030] Young W.H. *Upper and lower integration*. Proc. London Math. Soc. (2). 1905. V. 2. P. 52–66. [*423*]

[1031] Young W.H. *On the general theory of integration*. Phil. Trans. Royal Soc. London (A). 1905. V. 204. P. 221–252. [*423, 424, 425*]

[1032] Young W.H. *On parametric integration*. Monatsh. Math. Phys. 1910. B. 21. S. 125–149. [*428*]

[1033] Young W.H. *On a new method in the theory of integration*. Proc. London Math. Soc. (2). 1911. V. 9. P. 15–50. [*423*]

[1034] Young W.H. *On semi-integrals and oscillating successions of functions*. Proc. London Math. Soc. (2). 1911. V. 9. P. 286–324. [*134, 428*]

[1035] Young W.H. *On successions of integrals and Fourier series*. Proc. London Math. Soc. 1913. V. 11. P. 43–95. [*316*]

[1036] Young W.H. *On the new theory of integration*. Proc. Royal Soc., Ser. A. 1913. V. 88. P. 170–178. [*423*]

[1037] Young W.H. *On derivatives and their primitive functions*. Proc. London Math. Soc. 1913. V. 12. P. 207–217. [*436*]

[1038] Young W.H. *On integration with respect to a function of bounded variation*. Proc. London Math. Soc. 1914. V. 13. P. 109–150. [*425*]

[1039] Young W.H. *Note on the existence of converging sequences in certain oscillating successions of functions*. Proc. Royal Soc. (A). 1915. V. 92. P. 353–356. [*434*]

[1040] Young W.H. *On successions with subsequences converging to an integral*. Proc. London Math. Soc. (2). 1926. V. 24. P. 1–20. [*434*]

[1041] Younovitch B. *Sur la derivation des fonctions absolument additives d'ensemble*. C. R. (Dokl.) Acad. Sci. URSS. 1941. V. 30. P. 112–114. [*438*]

[1042] Zaanen A.C. Linear analysis. Measure and integral, Banach and Hilbert space, linear integral equations. Interscience, New York; North-Holland, Amsterdam; P. Noordhoff N.V., Groningen, 1953; vii+601 pp. [*414*]

[1043] Zaanen A.C. Integration. North-Holland, Amsterdam, 1967; xiii+604 pp. [*310, 312, 320, 414, 438*]

[1044] Zahn P. Ein konstruktiver Weg zur Masstheorie und Funktionalanalysis. Wissenschaftliche Buchgesellschaft, Darmstadt, 1978; 350 S. [*423*]

[1045] Zahorski Z. *Sur la première dérivée*. Trans. Amer. Math. Soc. 1950. V. 69. P. 1–54. [*402*]

[1046] Zajíček L. *Sets of σ-porosity and sets of σ-porosity (q)*. Časopis Pěst. Mat. 1976. V. 101. P. 350–359. [*404*]

[1047] Zalcman L. *Uniqueness and nonuniqueness for the Radon transform*. Bull. London Math. Soc. 1982. V. 14, №3. P. 241–245. [*228*]

[1048] Zamansky M. Introduction à l'algèbre et l'analyse modernes. 3me éd. Dunod, Paris, 1967; xviii+435 pp. [*414*]

[1049] Zhang G.Y. *Intersection bodies and the Busemann–Petty inequalities in R^4*. Ann. Math. (2). 1994. V. 140, №2. P. 331–346. [*215*]

[1050] Zhang G. *A positive solution to the Busemann-Petty problem in $\mathbf{R}^4$*. Ann. Math. 1999. V. 149, №2. P. 535–543. [*215*]

[1051] Ziemer W. Weakly differentiable functions. Springer-Verlag, New York – Berlin, 1989; xvi+308 pp. [*379*]

[1052] Zink R.E. *On the structure of measure spaces*. Acta Math. 1962. V. 107. P. 53–71. [*93*]

[1053] Zorich V.A. Mathematical analysis. V. I,II. Translated from the Russian. Springer, Berlin – New York, 2004; xviii+574 pp., xv+681 pp. (Russian ed.: Moscow, 1981). [*158, 234, 260*]

[1054] Zubieta Russi G. Teoría de la integral. Editorial Trillas, Mexico City, 1970; 222 pp. [*414*]

[1055] Zygmund A. Trigonometric series. V. I,II. 3d ed. Cambridge University Press, Cambridge, 2002; xiv+383 pp., viii+364 pp. [*261*]

Author Index

Subject Index

Notation: