V. I. Bogachev

1

Measure Theory

Springer

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 ° 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., Σ, Λ, Γ, Ξ), 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. $\square$

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 $B = \varnothing$. □

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. □

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 $f^{-1}(\mathcal{F})\subset\sigma\bigl(f^{-1}(\mathcal{F})\bigr)$. □

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. □

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 $n \le N$. 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 $n \le N$. Then $x \in B_n$ for every $n \le N$, whence we obtain $x \in \bigcap_{n=1}^{N} B_n$, which is a contradiction. The above proved equality yields the estimate

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

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 $k \to \infty$. □

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

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

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*

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

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). $\square$

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*

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

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

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

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

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 $(X\backslash A) \triangle (X\backslash A_\varepsilon) = A \triangle A_\varepsilon$. Next, the union of two measurable sets A and B is measurable. Indeed, let $\varepsilon > 0$ and let $A_\varepsilon, B_\varepsilon \in \mathcal{A}$ be such that $\mu^*(A \triangle A_\varepsilon) < \varepsilon/2$ and $\mu^*(B \triangle B_\varepsilon) < \varepsilon/2$. Since

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

one has

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

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

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

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

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

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

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

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

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

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

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

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

$$|\nu(A) - \mu^*(A)| = |\nu(A) - \nu(B) + \mu^*(B) - \mu^*(A)|$$
$$\le |\nu(A) - \nu(B)| + |\mu^*(B) - \mu^*(A)| \le 2\varepsilon.$$

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 $A'', A' \in \sigma(\mathcal{A}) \subset \mathcal{A}_\mu$. $\square$

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. □

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. $\square$

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 $I_k = \{|x_i| \leq k, i = 1, \ldots, n\}$ 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 $Q_j = c_j I + h_j$, $c_j > 0$, $h_j \in W$, 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 $\lambda_n^*(A \bigtriangleup A_\varepsilon) < \varepsilon$. Then

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

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. $\square$

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 semialgebra 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 $A_{n_1,\dots,n_k} = A_{n_1}$, and in the second, $A_{n_1,\dots,n_k} = A_k$. □

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

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

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

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

We need the following technical assertion. Let $\bigl(\mathbb{N}^\infty\bigr)^\infty$ denote the space of all sequences $\eta = (\eta^1, \eta^2, \dots)$ 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 $\sigma_1, \dots, \sigma_m$ and $\tau^m_1, \dots, \tau^m_n$ 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

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

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 $A^{\nu_1,\ldots,\nu_m}_{n_1,\ldots,n_k} \in \mathcal{E}$ and let

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

Keeping the notation of the above lemma, for any natural numbers $\eta_1,\ldots,\eta_l$ we find $\sigma \in \mathbb{N}^\infty$ and $\tau = (\tau^m) \in \bigl(\mathbb{N}^\infty\bigr)^\infty$ 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

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

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

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

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

(ii) Set

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

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 $E_{n_1,\dots,n_k} \in \mathcal{E}$. Let $\varepsilon > 0$. For every collection $m_1, \dots, m_k$ of natural numbers, denote by $D_{m_1,\dots,m_k}$ the union of the sets $E_{n_1,\dots,n_k}$ over all $n_1 \le m_1,\dots,n_k \le m_k$. Let

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

It is clear that as $m \to \infty$, the sets M_m monotonically increase to A, and the sets $M_{m_1,\dots,m_k,m}$ with fixed $m_1, \dots, m_k$ monotonically increase to $M_{m_1,\dots,m_k}$. 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

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

Therefore, for all k one has

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

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

It remains to prove that $E \subset A$. Let $x \in E$. Then, for all k we have $x \in D_{m_1,\dots,m_k}$. Hence $x \in E_{n_1,\dots,n_k}$ for some collection $n_1, \dots, n_k$ such that $n_1 \le m_1,\dots,n_k \le m_k$. Such collections will be called admissible. Our task is to construct an infinite sequence $n_1, n_2, \dots$ such that all its initial intervals $n_1, \dots, n_k$ are admissible. In this case $x \in \bigcap_{k=1}^\infty E_{n_1,\dots,n_k} \subset A$. 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 $n_1, \dots, n_k$ is called extendible if, for every $l \ge k$, there exists an admissible collection $p_1, \dots, p_l$ with $p_1 = n_1, \dots, p_k = n_k$. 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 $n \le m_1$ 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 $B_{n_1,\ldots,n_k} = f(A_{n_1,\ldots,n_k})$ are compact by the continuity of f and that

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

Indeed, the left-hand side of this equality is contained in the right-hand side for any sets and mappings. Let $y \in \bigcap_{k=1}^{\infty} f(A_{n_1,\ldots,n_k})$. Then, for every k, there exists $x_k \in A_{n_1,\ldots,n_k}$ with $f(x_k) = y$. If for infinitely many indices k the points x_k coincide with one and the same point x, then $x \in \bigcap_{k=1}^{\infty} A_{n_1,\ldots,n_k}$ by the monotonicity of $A_{n_1,\ldots,n_k}$. 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 $x \in A_{n_1,\ldots,n_k}$ for all k, since $x_j \in A_{n_1,\ldots,n_k}$ for all $j \geq k$ and $A_{n_1,\ldots,n_k}$ is a closed set. Thus, $x \in \bigcap_{k=1}^{\infty} A_{n_1,\ldots,n_k}$. By the continuity of f we obtain $f(x) = y$. $\square$

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 $(t_1, \ldots, t_k) \in \mathbb{R}^k$ 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 $(a_1, \ldots, a_k, t_1, \ldots, t_k) \mapsto \sum_{i=1}^{k} t_i a_i$ 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

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

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

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

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

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

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

$$\mathfrak{m}(E) \geq \mathfrak{m}(E \cap A) + \mathfrak{m}\big(E \cap (X \backslash A)\big).$$

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

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

Since

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

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

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

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

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

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 $C_n := \{x\colon\ \operatorname{dist}(x, C) \le n^{-1}\}$ 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 $D_k := \bigl\{x \in A\colon\ (k+1)^{-1} < \operatorname{dist}(x, C) \le k^{-1}\bigr\}$. 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. $\square$

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

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

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. $\square$

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

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

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 $\mu(A\triangle C) \le \mu(A\triangle B)+\mu(B\triangle C)$. 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 $|\mu(A) - \mu(B)| \le \mu(A \bigtriangleup B) = d(A,B)$.

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 non-negative 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

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

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

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

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 $A-A=\bigl\{a_1-a_2:\ a_1,a_2\in A\bigr\}$ 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 $\mu_*(S)=\sup\bigl\{\mu(B):\ B\subset S, B\in\mathcal{B}\bigr\}$. 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

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

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 $i \leq n$, 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 $t \leq s$ satisfying the conditions: 1) $t \leq t$, 2) if $t \leq s$ and $s \leq u$, then $t \leq u$ 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 $t \leq s$ and $s \leq t$, then $t = s$. But we do not require this. We write $t < s$ if $t \leq s$ and $t \neq s$. The set T is called linearly ordered if all its elements are pairwise comparable and, in addition, if $t \leq s$ and $s \leq t$, 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

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

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

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

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 $0 \le x, y \le 1$, with the following properties: $\lambda_2(K_j) \le 1/j$ 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. $\square$

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

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

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

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

Note that τ_* can also be defined by the formula

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

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 $\mu_*|_{\mathcal{A}} = \mu$. $\square$

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 $\mu_*(\varnothing) = 0$. □

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. $\square$

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 $\mathfrak{m}(A) \le \mathfrak{m}(B)$ 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

$$\beta(R_1 \cup R_2) + \beta(R_1 \cap R_2) = \beta(R_1) + \beta(R_2), \quad \forall\, R_1,\, R_2 \in \mathfrak{R}. \qquad (1.12.10)$$

If in (1.12.10) we replace the equality sign by "$\le$", then we obtain the definition of a submodular function, and the change of "$=$" to "$\ge$" 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 $\beta(R_1) \le \beta(R_2)$ 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. $\square$

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. $\square$

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 $T(\beta_1)$,..., $T(\beta_n)$, 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 $\mu(A_n) \ge \varepsilon > 0$ for all n, then $\mu(B) \ge \varepsilon$.

(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 $\mu(B_k) \le \sum_{n=k}^\infty \mu(A_n)$, $\mu(B_k) \ge \mu(A_k)$. 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 $\mu(A_n) \ge \varepsilon > 0$ 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 $\lambda(A_n) \ge \varepsilon > 0$ 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 $\lambda(E_q) \le 2\alpha(q)$. 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 $\lambda(E_k) \ge \alpha$ 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 $\mu(X' \bigtriangleup X_{t_n}) < \mu(X')2^{-n}$.

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 $\lambda_n\big(E \cap B(x,r)\big) = 0$ for some $r > 0$, and let E_1 be the set of all x such that $\lambda_n\big(E \cap B(x,r)\big) = \omega_n r^n$ 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

$$\tau_*(F) \le \sum_{n=1}^{\infty} \tau_*(F_n) \quad \text{whenever } F, F_n \in \mathcal{A}_{\mathcal{F}} \text{ and } F \subset \textstyle\bigcup_{n=1}^{\infty} F_n,$$

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) $\mu(E) = \sum_\alpha \mu(E \cap X_\alpha)$ for each set $E \in \mathcal{A}$, where convergence of the series $\sum_\alpha c_\alpha$, $c_\alpha \ge 0$, 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 $\mu\big((\Psi'\backslash\Psi) \cap X_\alpha\big) = 0$, 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 $\mu\big((X \times y) \cap E_\alpha\big) = 0$ 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 $\widetilde{\mu}(A \triangle A') = 0$.

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

$$\widetilde{\mathcal{A}} := \{A \triangle Z, A \in \mathcal{A}, Z \in \mathcal{Z}\}.$$

It is easily seen that $\widetilde{\mathcal{A}}$ is a σ-algebra and contains $\mathcal{A}$ and $\mathcal{S}$. Set $\widetilde{\mu}(A \triangle Z) := \mu(A)$ for $A \in \mathcal{A}$ and $Z \in \mathcal{Z}$. The definition is unambiguous because if $A \triangle Z = A' \triangle Z'$, $A, A' \in \mathcal{A}$, $Z, Z' \in \mathcal{Z}$, then $A \triangle A' = Z \triangle Z'$, whence $\mu(A \triangle A') = \mu_*(Z \triangle Z') = 0$, since $Z \triangle Z' \in \mathcal{Z}$. 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

$\sigma_\mu(A, B) = \mu(A \triangle B)/\mu(A \cup B)$ 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 $\mu(A \triangle B) = 0$.

(b) Show that if $A_n, A \in M_0$ and $\sigma_\mu(A_n, A) \to 0$, then $\mu(A_n \triangle A) \to 0$.

(c) Show that if $\mu(A_n \triangle A) \to 0$ 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 $x=3\sum_{n=1}^\infty\varepsilon_n 4^{-n}$, $\varepsilon_n\in\{0,1\}$. Show that $E+\dfrac{1}{2}E=[0,3/2]$, 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) $F[x,y)\ge 0$ whenever $x<y$ coordinatewise, 2) $F(x^j)\to F(x)$ whenever the vectors x^j increase to x, 3) $F(x)\to 0$ as $\max_k x_k\to-\infty$ and $F(x)\to 1$ 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 $f\colon X \to \mathbb{R}^1$ is called measurable with respect to $\mathcal{A}$ (or $\mathcal{A}$-measurable) if $\{x\colon f(x) < c\} \in \mathcal{A}$ for every $c \in \mathbb{R}^1$.*

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 $B \in \mathcal{B}(\mathbb{R}^1)$.*

Proof. Let f be $\mathcal{A}$-measurable. Denote by $\mathcal{E}$ the collection of all sets $B \in \mathcal{B}(\mathbb{R}^1)$ 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)

$$f^{-1}\Bigl(\bigcup_{n=1}^{\infty} B_n\Bigr) = \bigcup_{n=1}^{\infty} f^{-1}(B_n) \in \mathcal{A}, \quad f^{-1}(\mathbb{R}^1 \backslash B_n) = X \backslash f^{-1}(B_n) \in \mathcal{A}.$$

Since $\mathcal{E}$ contains the rays $(-\infty, c)$, we obtain that $\mathcal{B}(\mathbb{R}^1) \subset \mathcal{E}$, i.e., $\mathcal{B}(\mathbb{R}^1) = \mathcal{E}$. 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

$$\bigl\{x\colon\ f(x)+g(x)<c\bigr\} = \bigl\{x\colon\ f(x)<c-g(x)\bigr\}$$
$$= \bigcup_{r_n}\Bigl(\bigl\{x\colon\ f(x)<r_n\bigr\}\cap\bigl\{x\colon\ r_n<c-g(x)\bigr\}\Bigr),$$

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. □

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. □

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 $\lim\limits_{n\to\infty} f_n(x) = f(x)$ for all $x \in X$. □

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

$$X_n^m := \bigcap_{i\geq n} \Bigl\{x\colon\ |f_i(x) - f(x)| < \frac{1}{m}\Bigr\} \in \mathcal{A}_\mu.$$

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

$$\mu(X\backslash X_\varepsilon) = \mu\Bigl(\bigcup_{m=1}^\infty (X\backslash X_{k(m)}^m)\Bigr) \leq \sum_{m=1}^\infty \mu(X\backslash X_{k(m)}^m) \leq \varepsilon \sum_{m=1}^\infty 2^{-m} = \varepsilon.$$

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

$$|f_{n_i}(x) - f_{n_j}(x)| \leq |f_{n_i}(x) - f_{n_{i-1}}(x)| + |f_{n_{i-1}}(x) - f_{n_{i-2}}(x)| + \dots$$
$$+ |f_{n_{j+1}}(x) - f_{n_j}(x)| \leq \sum_{l=j}^{\infty} 2^{-l} \leq 2^{-j+1} \leq 2^{-m},$$

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 $\bigl(f(x), g(x)\bigr) \in Y$ and $\bigl(f_n(x), g_n(x)\bigr) \in Y$ 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

$$\mu\Bigl(x\colon\ \bigl|\Psi\bigl(f(x), g(x)\bigr) - \Psi\bigl(f_{j_n}(x), g_{j_n}(x)\bigr)\bigr| > c\Bigr) > c \qquad (2.2.1)$$

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

$$\Psi\bigl(f_{i_n}(x), g_{i_n}(x)\bigr) \to \Psi\bigl(f(x), g(x)\bigr) \quad \text{a.e.},$$

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

$$\mu\Bigl(x\colon\ |f_{1,k_1}(x) - f_1(x)| \geq 2^{-1}\Bigr) \leq 2^{-1}.$$

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, $i = 1, \dots, n$, 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*

$$\int_X |f_i(x) - f_j(x)|\,\mu(dx) < \varepsilon \quad \text{for all } i,\, j \geq n.$$

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 $C = \max_{x\in X,\, i\le N} |f_i(x)| + 1$ and $\delta = \varepsilon(2C)^{-1}$. If $\mu(D) < \delta$ and $n \ge N$, 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

$$\sum_{k=-\infty}^{+\infty} \varepsilon k\, \mu\big(x\colon\ \varepsilon k \le f(x) < \varepsilon(k+1)\big)$$

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

$$\sum_{k=-\infty}^{+\infty} \frac{k}{n}\, \mu\Big(x\colon\ \frac{k}{n} \le f(x) < \frac{k+1}{n}\Big).$$

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

$$\Big|\int_D f_n(x)\,\mu(dx)\Big| + \Big|\int_D g_n(x)\,\mu(dx)\Big| \le \varepsilon, \quad \forall\, n \in \mathbb{N}. \tag{2.4.1}$$

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 $n \ge N$

$$\begin{aligned}
&\Big|\int_X f_n(x)\,\mu(dx) - \int_X g_n(x)\,\mu(dx)\Big| \\
&\le \Big|\int_{X_\delta} \big[f_n(x) - g_n(x)\big]\,\mu(dx) + \int_{X\backslash X_\delta} f_n(x)\,\mu(dx) - \int_{X\backslash X_\delta} g_n(x)\,\mu(dx)\Big| \\
&\quad \le \varepsilon\mu(X) + \Big|\int_{X\backslash X_\delta} f_n(x)\,\mu(dx)\Big| + \Big|\int_{X\backslash X_\delta} g_n(x)\,\mu(dx)\Big| \le \varepsilon\big(\mu(X)+1\big),
\end{aligned}$$

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. □

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*:

$$I(f) := \sup\biggl\{\int_X \varphi\, d\mu\colon\ \varphi \le f \text{ a.e., } \varphi \text{ is simple}\biggr\}.$$

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. □

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*

$$\mu\bigl(x\colon\ |f(x)| \geq R\bigr) \leq \frac{1}{R}\int_X |f(x)|\,\mu(dx). \tag{2.5.1}$$

PROOF. Set $A_R = \bigl\{x\colon\ |f(x)| \geq R\bigr\}$. It is clear that $R \cdot I_{A_R}(x) \leq |f(x)|$ 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*

$$\sup_{n\geq 1}\int_X \min(f,n)\,d\mu < \infty.$$

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 $|f_n(x) - g_n(x)| \leq n^{-1}$ for all x. Since $f_n(x) \to f(x)$, one has $g_n(x) \to f(x)$. Whenever $n \geq k$, we have $|f_n - f_k| = f_n - f_k$, hence

$$\begin{aligned}\int |g_n - g_k|\,d\mu &= \int |g_n - f_n + f_n - f_k + f_k - g_k|\,d\mu\\ &\leq \int |g_n - f_n|\,d\mu + \int |f_n - f_k|\,d\mu + \int |f_k - g_k|\,d\mu\\ &\leq \frac{1}{n}\mu(X) + \int f_n\,d\mu - \int f_k\,d\mu + \frac{1}{k}\mu(X).\end{aligned}$$

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 $|f(x)| \leq g(x)$ 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

$$\min(f^+,n) \leq \min(g,n) \quad\text{and}\quad \min(f^-,n) \leq \min(g,n).$$

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. $\square$

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. $\square$

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 $f = \lim\limits_{n\to\infty} f_n$ 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 $L^1(\mu)$. $\square$

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

$$\int_D \big[|\varphi(x)| + |\varphi_n(x)|\big]\,\mu(dx) < \varepsilon$$

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

$$\int_X |\varphi(x) - \varphi_n(x)|\,\mu(dx)$$
$$\leq \int_{X_\delta} |\varphi(x) - \varphi_n(x)|\,\mu(dx) + \int_{X\backslash X_\delta} |\varphi(x) - \varphi_n(x)|\,\mu(dx) \leq \varepsilon\mu(X) + \varepsilon,$$

which proves (2.7.1). $\square$

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

$$\begin{aligned}\int_X &|g_n(x) - g_k(x)|\,\mu(dx)\\ &\le \int_X \Bigl[|g_n(x) - f_n(x)| + |f_n(x) - f_k(x)| + |f_k(x) - g_k(x)|\Bigr]\,\mu(dx)\\ &\qquad\qquad \le \frac{1}{n} + \frac{1}{k} + \int_X |f_n(x) - f_k(x)|\,\mu(dx).\end{aligned}$$

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*

$$|f_n(x)| \leq \Phi(x) \quad \textit{a.e. for every } n,$$

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 $|f| \leq \Phi$ 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

$$|f_n(x) - f(x)| \leq \frac{\varepsilon}{2\mu(X) + 1}$$

for all $n \geq N$. Therefore, for $n \geq N$ we have

$$\begin{aligned}
&\int_X |f(x) - f_n(x)|\,\mu(dx) \\
&\qquad \leq \int_{X\backslash X_\delta} |f(x) - f_n(x)|\,\mu(dx) + \int_{X_\delta} |f(x) - f_n(x)|\,\mu(dx) \\
&\qquad\qquad \leq 2\int_{X\backslash X_\delta} \Phi(x)\,\mu(dx) + \frac{\varepsilon}{2\mu(X)+1}\mu(X_\delta) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

The theorem is proven. □

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 $f_n(x) \leq f_{n+1}(x)$ a.e. for each $n \in \mathbb{N}$. Suppose that*

$$\sup_n \int_X f_n(x)\,\mu(dx) < \infty.$$

Then, the function $f(x) = \lim_{n\to\infty} f_n(x)$ *is almost everywhere finite and integrable. In addition, equality* (2.8.1) *holds true.*

Proof. For $n \leq m$ 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 $|f_n(x)| \leq |f(x)| + |f_1(x)|$ 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*

$$\sup_n \int_X f_n(x)\,\mu(dx) \leq K < \infty.$$

Then, the function f *is* μ*-integrable and*

$$\int_X f(x)\,\mu(dx) \leq K.$$

Moreover,

$$\int_X f(x)\,\mu(dx) \leq \liminf_{n\to\infty} \int_X f_n(x)\,\mu(dx).$$

Proof. Set $g_n(x) = \inf_{k\geq n} f_k(x)$. Then

$$0 \leq g_n \leq f_n, \quad g_n \leq g_{n+1}.$$

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). $\square$

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 $L^p\bigl([a, +\infty)\bigr)$ 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*

$$\int_X |fg| \, d\mu \le \Bigl(\int_X |f|^p \, d\mu\Bigr)^{1/p} \Bigl(\int_X |g|^q \, d\mu\Bigr)^{1/q}. \qquad (2.11.1)$$

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). $\square$

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 $g=I_{\{f\ne 0\}}$ we arrive at the following estimate.

2.11.4. Corollary. *Suppose that $f\in\mathcal{L}^p(\mu)$ and $\mu\bigl(x\colon\ f(x)\ne 0\bigr)<\infty$. Then*

$$\int_X |f|\,d\mu \le \mu\bigl(x\colon\ f(x)\ne 0\bigr)^{1/q}\Bigl(\int_X |f|^p\,d\mu\Bigr)^{1/p},\quad q=p(p-1)^{-1}.$$

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 $f_1\in\mathcal{L}^{p_1}(\mu)$,..., $f_n\in\mathcal{L}^{p_n}(\mu)$. 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 $|f+g|^p\le2^p(|f|^p+|g|^p)$, 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 $\|f+g\|_p \le \|f\|_p + \|g\|_p$. □

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*

$$E_{(f_i),B} = \bigl\{x\colon\ \bigl(f_1(x),\dots,f_n(x),\dots\bigr) \in B\bigr\}, \quad f_i \in \mathcal{F},\ B \in \mathcal{B}(\mathbb{R}^\infty). \quad (2.12.1)$$

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 $f_1,\dots,f_n,\dots$ 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 $\mathcal{E} = \sigma(\mathcal{F})$. □

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*

$$g(x) = \psi\bigl(f_1(x),\dots,f_n(x),\dots\bigr), \quad (2.12.2)$$

where $f_i \in \mathcal{F}$ and ψ is a Borel function on $\mathbb{R}^\infty$. If $\mathcal{F}$ is a finite family $\{f_1,\dots,f_n\}$, 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 $E_1,\dots,E_k$ with coefficients $c_1,\dots,c_k$, 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 $f_i^{(j)}$, $j = 1,\dots,k$, there will correspond the subsequences $J_i^{(j)}$. Set $\varphi_j(x_1,x_2,\dots) := \psi_j(x_{J_1^{(j)}}, x_{J_2^{(j)}},\dots)$. It is clear that φ_j is a Borel function on $\mathbb{R}^\infty$. Then g can be written in the form

$$g = c_1\psi_1(f_{J_1^1}, f_{J_2^1},\dots) + \cdots + c_k\psi_k(f_{J_1^k}, f_{J_2^k},\dots) = \sum_{j=1}^k c_j\varphi_j(f_1,f_2,\dots).$$

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 $g_k = \varphi_k(f_1,f_2,\dots)$, 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 $\mathbb{R}^n$. $\square$

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. $\square$

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. □

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 $\varphi(t_0,t_0)=0$. □

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 $f_{n,k} = \max(f_n, \dots, f_{n+k})$ 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*

$$\int_{-\infty}^{\infty} g(t)\,dF(t) := \int_{\mathbb{R}} g(t)\,\mu(dt) \qquad (2.12.7)$$

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*

$$\begin{aligned}\int_X f\ln f\,d\mu - \int_X f\,d\mu\Bigl(\ln\int_X f\,d\mu\Bigr) \\ \ge \int_X f\ln g\,d\mu - \int_X f\,d\mu\Bigl(\ln\int_X g\,d\mu\Bigr),\end{aligned} \tag{2.12.13}$$

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*

$$\|\mu-\nu\|^2 := \left(\int_X |f-1|\,d\mu\right)^2 \le 2\int_X f\ln f\,d\mu,$$

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_nf_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 $\sup\limits_{x,y\in B} |f(x) - f(y)| \le \varepsilon$ (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 $\sup\limits_{x,y\in E} |f(x) - f(y)| \le \varepsilon$. 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 $\mu(E) \le \delta_k < \delta$. It remains to choose a point x_n in every set B_n and put $g|_{B_n} = f(x_n)$. Then $|g(x) - f(x)| \le \varepsilon$ 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 $|f_{n_k}(x)| \ge \varepsilon$ for all $x \in A_k$ and all k.

Hint: let $g_m(x) := \sup_{n\ge m} |f_n(x)|$. 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 $E := \bigcap_{m\ge 1}\{x\colon\, g_m(x) > \varepsilon\}$ 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 $\liminf\limits_{k\to\infty} |t_{n_k}| > 0$.

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

$$\lim_{n\to\infty} \mu\Bigl(\bigcup_{m\ge n} \bigl\{x\colon\ |f_m(x) - f_n(x)| \ge \varepsilon\bigr\}\Bigr) = 0.$$

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 $\lim\limits_{y\to y_0} f(x,y) = f(x,y_0)$ 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 $\lim\limits_{y\to y_0} f(x,y) = f(x,y_0)$ 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 $\lim\limits_{y\to 0} f(x,y) = 0$ 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.° 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 $\bigl\{x\colon\ \omega_f(x) \geq j^{-1}\bigr\}$.

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 $0 \leq f \leq 1$ 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 $\mu(f^{-1}(I_n)) \geq 2^{-n}$; there exists $y \in \bigcap_{n=1}^\infty I_n$; then $\mu\bigl(x\colon\ |f(x) - y|^{-1} \geq 2^n\bigr) \geq 2^{-n}$.

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.° 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.° 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.° 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.° 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 $\liminf_{n\to\infty} \bigl|f_n(x)/\|f_n\|_1^{1/p} - C_n\bigr| < \infty$ a.e., whence $\limsup_{n\to\infty} |f_n(x)| = \infty$ 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

$$\alpha = p(q-s)\bigl(q(p-s)\bigr)^{-1}, \quad \beta = s(p-q)\bigl(q(p-s)\bigr)^{-1}.$$

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 $\Lambda \le p$ and verify the required properties by using that $p(f) \le 0$ whenever $f \le 0$ 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 $\alpha_i \ge 0$ 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 $\mu(f<c) \le 1/2$ for all $c<m$ and $\mu(f<c) \ge 1/2$ 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

$$M_\varepsilon^p := \Bigl\{f\in\mathcal{L}^p(\mu)\colon\ \mu\bigl(x\colon\ |f(x)|\geq\varepsilon\|f\|_{L^p(\mu)}\bigr)\geq\varepsilon\Bigr\}.$$

(i) Show that $\mathcal{L}^p(\mu)=\bigcup_{\varepsilon>0}M_\varepsilon^p$.

(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 $f\in M_\varepsilon^p$ 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 $f\notin M_\varepsilon^p$, 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. $\square$

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

$$\|\mu\| \le 2\sup\bigl\{|\mu(A)|:\ A \in \mathcal{A}\bigr\} \le 2\|\mu\|. \tag{3.1.4}$$

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 $A \in \mathcal{A}$

$$\begin{aligned}\int_A h \, d\mu &= \int_A f \, d\mu + n^{-1}\mu(A \cap X_n^+) \le \int_A f \, d\mu + \eta(A \cap X_n^+) \\ &= \int_{A\setminus X_n^+} f \, d\mu + \nu(A \cap X_n^+) \le \nu(A\setminus X_n^+) + \nu(A \cap X_n^+) = \nu(A).\end{aligned}$$

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 $\mu_0(A) := \nu\big(A \cap (X\backslash Y)\big)$. 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. $\square$

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 $\mathcal{A}_1\otimes\mathcal{A}_2$. $\square$

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. $\square$

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. $\square$

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. $\square$

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 $\lambda(A_x) = f(x)$. $\square$

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

$$\int_{X\times Y} f_n \, d(\mu\otimes\nu) = \int_Y \biggl(\int_X f_n \, d\mu\biggr) d\nu \le \int_Y \biggl(\int_X f \, d\mu\biggr) d\nu,$$

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

$$\int_X |f|^p\, d\mu = p\int_0^\infty t^{p-1}\mu\bigl(x\colon\ |f(x)| > t\bigr)\, dt. \tag{3.4.4}$$

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 $\{f \ne 0\}$. 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 μ_{n+1}, μ_{n+2},... 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. $\square$

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. $\square$

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 $S = \{\alpha_i\}$, $Q_\alpha = \bigcap_{i:\ \alpha_i=\alpha} K^{(i)}_{\alpha_i}$. 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

$$\mu\Bigl(\Bigl(\prod_{i=1}^n A_i \backslash \prod_{i=1}^n K_i\Bigr) \times \prod_{\alpha\notin\{\alpha_1,\ldots,\alpha_n\}} X_\alpha\Bigr)$$
$$\leq \sum_{i=1}^n \bigotimes_{i=1}^n \mu_{\alpha_i}\Bigl((A_i\backslash K_i) \times \prod_{j\leq n, j\neq i} X_j\Bigr) = \sum_{i=1}^n \mu_i(A_i\backslash K_i) < \varepsilon.$$

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. $\square$

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. □

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 $[0,1]$. □

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 $x/2 + \mathrm{const}$ (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 $f(E) = D$. □

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

$$\lambda_n\big(F(E)\big) = \int_E |JF(x)|\,dx \tag{3.7.2}$$

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 $\Delta_j := \lambda_n\big(F(E_j)\big) - \lambda_n\big(L_j(E_j)\big)$, and in this notation one has

$$\begin{aligned}\lambda_n\big(F(E)\big) &= \sum_{j=1}^{m^n} \lambda_n\big(F(E_j)\big) = \sum_{j=1}^{m^n}\Big[\lambda_n\big(L_j(E_j)\big) + \Delta_j\Big] \\ &= \sum_{j=1}^{m^n} |\det F'(x_j)|\lambda_n(E_j) + \sum_{j=1}^{m^n}\Delta_j.\end{aligned}$$

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,

$$\Delta_j = \lambda_n\big(F(E_j)\big) - \lambda_n\big(L_j(E_j)\big) \le 2C_n\varepsilon\lambda_n(E_j),$$

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

$$\lambda_n\big(L_j(E_j)\big) - \lambda_n\big(F(E_j)\big) \le K_n\sqrt{\varepsilon}\lambda_n(E_j), \quad j = 1, \ldots, m^n. \tag{3.7.4}$$

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 $\bigl\|\bigl(I - G_jF'(x)\bigr)h\bigr\| \le \|h\|/2$ 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

$$\begin{aligned}\Psi(x) &= x - G_jF(x) + y - x_j + G_jF(x_j)\\ &= y + (x - x_j) + G_j\bigl(F(x_j) - F(x)\bigr) = y + G_j\bigl(r(x, x_j)\bigr).\end{aligned}$$

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

$$\begin{aligned}\lambda_n\bigl(L_j(E_j)\bigr) - \lambda_n\bigl(F(E_j)\bigr) &\le \lambda_n\bigl(L_j(E_j)\bigr) - \lambda_n\bigl(L_j(Q_j)\bigr)\\ &= |\det F'(x_j)|\bigl[\lambda_n(E_j) - \lambda_n(Q_j)\bigr]\\ &= |\det F'(x_j)|\bigl(1 - (1-\sqrt{\varepsilon})^n\bigr)\lambda_n(E_j),\end{aligned}$$

which yields (3.7.4). Thus, formula (3.7.1) is established for cubes. The general case easily follows from this. $\square$

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 $\bigl(F(a), F(b)\bigr)$, one has*

$$\int_a^b g\bigl(F(t)\bigr)F'(t)\,dt = \int_{F(a)}^{F(b)} g(s)\,ds. \tag{3.7.6}$$

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

$$\frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \exp[-i(y,x)] \exp[-\alpha|x|^2]\,dx = \frac{1}{(2\alpha)^{n/2}} \exp\Big[-\frac{1}{4\alpha}|y|^2\Big].$$

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

$$p_\sigma(x) = \frac{1}{(2\pi\sigma)^{n/2}} \exp\Big(-\frac{|x|^2}{2\sigma}\Big)$$

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*

$$\widehat{\partial_{x_j} f}(y) = i y_j \widehat{f}(y).$$

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

$$\Bigl| \int_{\mathbb{R}^n} f \, d\mu \Bigr| \le \varepsilon + \Bigl| \int_{\mathbb{R}^n} f_0 \, d\mu \Bigr| \le \varepsilon + \Bigl| \int_{\mathbb{R}^n} [f_0 - g] \, d\mu \Bigr|$$
$$\le 2\varepsilon + \int_{\mathbb{R}^n \backslash K} |g| \, d|\mu| \le 4\varepsilon.$$

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 $R > R_1$

$$\begin{aligned} |J_R - f(x)| &\le \Bigl|J_R - \frac{f(x)}{\pi}\int_{-T_2}^{T_2}\frac{\sin(Rt)}{t}\,dt\Bigr| \\ &+ \Bigl|\frac{f(x)}{\pi}\int_{-T_2}^{T_2}\frac{\sin(Rt)}{t}\,dt - f(x)\Bigr| \le \Bigl|J_R - \frac{f(x)}{\pi}\int_{-T_2}^{T_2}\frac{\sin(Rt)}{t}\,dt\Bigr| + \frac{\varepsilon}{3} \\ &= \frac{1}{\pi}\Bigl|\int_{-\infty}^{+\infty} f(t+x)\frac{\sin(Rt)}{t}\,dt - \int_{-T_2}^{T_2} f(x)\frac{\sin(Rt)}{t}\,dt\Bigr| + \frac{\varepsilon}{3} \\ &\le \frac{1}{\pi}\Bigl|\int_{-T_2}^{T_2}[f(t+x)-f(x)]\frac{\sin(Rt)}{t}\,dt\Bigr| \\ &+ \frac{1}{\pi}\Bigl|\int_{\{|t|\ge T_2\}} f(t+x)\frac{\sin(Rt)}{t}\,dt\Bigr| + \frac{\varepsilon}{3} < \frac{\varepsilon}{\pi} + \frac{\varepsilon}{\pi} + \frac{\varepsilon}{3} < \varepsilon. \end{aligned}$$

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

$$f(x_1, x_2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ix_1y_1} g_1(y_1, x_2)\, dy_1,$$

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

$$\begin{aligned} g_1(y_1, x_2) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ix_2y_2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-iy_2z_2} g_1(y_1, z_2)\, dz_2\, dy_2 \\ &= \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{+\infty} e^{ix_2y_2} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-iy_1z_1 - iy_2z_2} f(z_1, z_2)\, dz_1\, dz_2\, dy_2, \end{aligned}$$

which yields (3.8.2). □

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. □

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 $\psi = \check{f}$. □

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*

$$\mu\big(x\colon\ |x| \geq t\big) \leq \frac{\sqrt{e}}{\sqrt{e}-1}\int_{\mathbb{R}^n}\big[1-\widetilde{\mu}(y/t)\big]\,\gamma(dy), \quad \forall\, t>0, \tag{3.8.6}$$

where γ is the standard Gaussian measure on $\mathbb{R}^n$.

Proof. We know that $\widetilde{\gamma}(x) = \exp\big(-|x|^2/2\big)$. Let γ_t be the image of γ under the mapping $x \mapsto x/t$. Then $\widetilde{\gamma_t}(x) = \exp\big(-t^{-2}|x|^2/2\big)$. Therefore, by (3.8.5), we obtain

$$\mu\big(x\colon\ |x| \geq t\big) = \mu\big(x\colon\ \widetilde{\gamma_t}(x) \leq e^{-1/2}\big) \leq \frac{1}{1-e^{-1/2}}\int_{\mathbb{R}^n}[1-\widetilde{\mu}(y)]\,\gamma_t(dy).$$

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*

$$\mu\big(x\colon\ |x| \geq r^{-2}\big) \leq 6nr^2 + 3\sup_{|z|\leq r}|1-\widetilde{\mu}(z)|. \tag{3.8.7}$$

Proof. The left-hand side of (3.8.7) is majorized by the integral of the function $3|1-\widetilde{\mu}(r^2 y)|$ against the measure γ, since $\sqrt{e}(\sqrt{e}-1)^{-1} < 3$. The integral over the ball of radius r^{-1} is majorized by $3\sup_{|z|\leq r}|1-\widetilde{\mu}(z)|$, as $|r^2 y| \leq r$ if $|y| \leq r^{-1}$. By Chebyshev's inequality one has

$$\gamma\big(y\colon\ |y| > r^{-1}\big) \leq r^2\int_{\mathbb{R}^n}|y|^2\,\gamma(dy) = nr^2.$$

It remains to observe that $|1-\widetilde{\mu}| \leq 2$. □

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 $g(x,y) = f_0\big(F(x,y)\big)$, 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,

$$\|f*g\|_{L^1(\mathbb{R}^n)} \leq \|f\|_{L^1(\mathbb{R}^n)}\|g\|_{L^1(\mathbb{R}^n)}. \tag{3.9.2}$$

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

$$\mu * \nu(B) = \int_{\mathbb{R}^n \times \mathbb{R}^n} I_B(x+y)\mu(dx)\,\nu(dy) \qquad (3.9.5)$$
$$= \int_{\mathbb{R}^n} \mu(B-y)\,\nu(dy) = \int_{\mathbb{R}^n} \nu(B-x)\,\mu(dx).$$

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

$$\int_Z \varphi g\, d\mu_{\sigma,T} = \int_Z \varphi\, d\nu_{\sigma,T} = \int_{X\times Y} \varphi\circ T\, d(\nu\otimes\sigma) = \int_{X\times Y} \varphi\circ T f\, d(\mu\otimes\sigma)$$
$$\le \|f\|_{L^p(\mu)} \Bigl(\int_Y \int_X \bigl|\varphi\bigl(T(x,y)\bigr)\bigr|^q \,\mu(dx)\,\sigma(dy)\Bigr)^{1/q} = \|f\|_{L^p(\mu)} \|\varphi\|_{L^q(\mu_{\sigma,T})},$$

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

$$|\operatorname{diam} S_a(A) - \varepsilon|^2 \leq |x-y|^2 = |b-c|^2 + |(x-y,a)|^2$$
$$\leq |b-c|^2 + |M_c - m_b|^2 = |(b+m_b a) - (c+M_c a)|^2 \leq (\operatorname{diam} A)^2$$

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 $n \leq 3$, but for $n \geq 4$ 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*

$$H^s_\delta(A) = \inf\biggl\{\sum_{i=1}^\infty \alpha(s)\Bigl(\frac{\operatorname{diam} C_j}{2}\Bigr)^s:\ A \subset \bigcup_{j=1}^\infty C_j,\ \operatorname{diam} C_j \leq \delta\biggr\},$$

$$H^s(A) = \lim_{\delta\to 0} H^s_\delta(A) = \sup_{\delta>0} H^s_\delta(A).$$

We note that the second equality in the definition of H^s is fulfilled, since $H^s_\delta \geq H^s_{\delta'}$ 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

$\left(\bigcup_{j=1}^{\infty} C_j\right) \bigcap \left(\bigcup_{j=1}^{\infty} C'_j\right) = \varnothing$. Hence

$$H^s_\delta(A) + H^s_\delta(B) \le \sum_{j=1}^{\infty} \alpha(s)2^{-s}(\operatorname{diam} C_j)^s + \sum_{j=1}^{\infty} \alpha(s)2^{-s}(\operatorname{diam} C'_j)^s,$$

whence we obtain that $H^s_\delta(A)+H^s_\delta(B) \le H^s_\delta(A\cup B)$, 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^k_j with diameters at most k^{-1} and

$$\sum_{j=1}^{\infty} \alpha(s)2^{-s}(\operatorname{diam} C^k_j)^s \le H^s_{1/k}(A) + k^{-1}.$$

The set $B = \bigcap_{k=1}^{\infty} \bigcup_{j=1}^{\infty} C^k_j$ is Borel and

$$H^s_{1/k}(B) \le \sum_{j=1}^{\infty} \alpha(s)2^{-s}(\operatorname{diam} C^k_j)^s \le H^s_{1/k}(A) + k^{-1},$$

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 $H^s_\delta(A) = 0$ 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^1_δ 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 $\lambda_1^*(A) \le H^1_\delta(A)$. 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 $\lambda_1^*(A) \ge H^1_\delta(A)$. □

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 $c \le 1$. 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

$$H^n_\delta\Bigl(U\backslash\bigcup_{j=1}^\infty U_j\Bigr) \le H^n\Bigl(U\backslash\bigcup_{j=1}^\infty U_j\Bigr) = 0.$$

Hence

$$H^n_\delta(U) = \sum_{j=1}^\infty H^n_\delta(U_j) \le \sum_{j=1}^\infty \lambda_n(U_j) = \lambda_n(U).$$

This contradiction shows that $c \le 1$. 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

$$\lambda_n(U) \le \sum_{j=1}^\infty \lambda_n(C_j) \le \sum_{j=1}^\infty \alpha(n)2^{-n}(\operatorname{diam} C_j)^n$$

and hence $\lambda_n(U) \le H^n_\delta(U) \le H^n(U)$. □

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., $|f(x)-f(y)| \le \Lambda|x-y|$ for all $x,y \in \mathbb{R}^n$. Then, for every $s \ge 0$ and every $A \subset \mathbb{R}^n$, we have $H^s\bigl(f(A)\bigr) \le \Lambda^s H^s(A)$.*

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 $\operatorname{diam} f(C_j) \le \Lambda \operatorname{diam} C_j \le \Lambda\delta$ and the sets $f(C_j)$ cover $f(A)$. Hence

$$H^s_{\Lambda\delta}\bigl(f(A)\bigr) \le \Lambda^s \sum_{j=1}^\infty \alpha(s)2^{-s}(\operatorname{diam} C_j)^s,$$

so $H^s_{\Lambda\delta}\bigl(f(A)\bigr) \le \Lambda^s H^s_\delta(A)$. 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:*

$$m(A_1 \cup \cdots \cup A_n) \geq m(A_1) + \cdots + m(A_n)$$

for all disjoint $A_1, \ldots, A_n \in \mathcal{R}$.

(i) *For all $A \in \mathcal{R}$ we set*

$$m_{\text{add}}(A) := \inf\Bigl\{\sum_{j=1}^{n} m(A_j)\colon \quad A = \bigcup_{j=1}^{n} A_j, A_j \in \mathcal{R},\ A_j \text{ are disjoint}\Bigr\}.$$

Then m_{add} is an additive set function, $m_{\text{add}} \leq m$, and $m_{\text{add}} \geq \nu$ for each additive set function $\nu\colon \mathcal{R} \to [0, +\infty]$ such that $\nu \leq m$.

(ii) *Set*

$$m_\sigma(A) := \inf\Bigl\{\sum_{j=1}^{\infty} m(A_j)\colon \quad A = \bigcup_{j=1}^{\infty} A_j,\ A_j \in \mathcal{R},\ A_j \text{ are disjoint}\Bigr\},\ A \in \mathcal{R}.$$

Then m_σ is a countably additive set function, $m_\sigma \leq m$, and $m_\sigma \geq \nu$ for each countably additive set function $\nu\colon \mathcal{R} \to [0, +\infty]$ such that $\nu \leq m$.

Proof. (i) Let $E_1, E_2 \in \mathcal{R}$, $E_1 \cap E_2 = \varnothing$. We show that $m_{\text{add}}(E_1 \cup E_2) \leq m_{\text{add}}(E_1) + m_{\text{add}}(E_2)$. 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

$$\sum_{i=1}^{k} m(E_1^i) < m_{\text{add}}(E_1) + \varepsilon,\ \sum_{j=1}^{n} m(E_2^j) < m_{\text{add}}(E_2) + \varepsilon.$$

Then E_1^i and E_2^j are disjoint, hence

$$m_{\text{add}}(E_1 \cup E_2) \leq \sum_{i=1}^{k} m(E_1^i) + \sum_{j=1}^{n} m(E_2^j) < m_{\text{add}}(E_1) + m_{\text{add}}(E_2) + 2\varepsilon.$$

It remains to use that ε is arbitrary. Let us establish the opposite inequality. Now we may assume that $m_{\text{add}}(E_1 \cup E_2) < \infty$. 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 $\mathcal{R}$. $\square$

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 $m_0' = m_0$, $m_1' = m_1$, $m_2 = m_\sigma$. $\square$

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. □

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 $\widehat{f * f} = (2\pi)^{n/2}(\widehat{f})^2$. □

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

$$\begin{aligned} p_t * f(x) &= \int_{\mathbb{R}^n} f(y)p_t(x-y)\,dy \\ &= (2\pi)^{-n/2}\int_{\mathbb{R}^n} \varphi(z)\exp[-i(z,x)]\exp[-t|z|^2/2]\,dz \geq 0. \end{aligned}$$

By the continuity of f we have $f * p_{1/k}(x) \to f(x)$. Hence $f \geq 0$. 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

$$\int_{\mathbb{R}^n} f(x)\psi_k(x)\,dx = \pi^{n/2}\int_{\mathbb{R}^n} \varphi(x)p_{1/k}(x)\,dx \leq \pi^{n/2}\varphi(0)$$

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 $x \neq 0$, 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*

$$w\bigl(tx+(1-t)y\bigr) \ge u(x)^t v(y)^{1-t}, \quad \forall\, x, y \in \mathbb{R}^n. \tag{3.10.5}$$

Then

$$\int_{\mathbb{R}^n} w(x)\,dx \ge \biggl(\int_{\mathbb{R}^n} u(x)\,dx\biggr)^t \biggl(\int_{\mathbb{R}^n} v(y)\,dy\biggr)^{1-t}. \tag{3.10.6}$$

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,

$$\int_{-\infty}^{+\infty} w\bigl(tx'+(1-t)y', x_n\bigr)\,dx_n$$
$$\ge \biggl(\int_{-\infty}^{+\infty} u(x', x_n)\,dx_n\biggr)^t \biggl(\int_{-\infty}^{+\infty} v(y', y_n)\,dy_n\biggr)^{1-t}$$

by the one-dimensional case, since for fixed $x', y' \in \mathbb{R}^{n-1}$ we have

$$w\bigl(tx'+(1-t)y', tx_n+(1-t)y_n\bigr) \ge u(x', x_n)^t v(y', y_n)^{1-t}.$$

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

$$A(s) := \bigl\{x\colon\ u(x) \ge s\bigr\},\ B(s) := \bigl\{x\colon\ v(x) \ge s\bigr\},\ C(s) := \bigl\{x\colon\ w(x) \ge s\bigr\}.$$

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

$$t\lambda_1\bigl(A(s)\bigr) + (1-t)\lambda_1\bigl(B(s)\bigr) \le \lambda_1\bigl(C(s)\bigr). \tag{3.10.7}$$

Indeed, it suffices to verify that, for arbitrary compact sets $K \subset tA(s)$ and $K' \subset (1-t)B(s)$, we have $\lambda_1(K)+\lambda_1(K') \le \lambda_1(K+K')$. 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). □

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 $\gamma_1 \leq \gamma \leq \gamma_2$. 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 $x \leq y$. (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 $T_n(y) = \sum_{i=1}^{n} f(x_i, y)(a_{i+1}-a_i)$ 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

$$|Lf(s)|^2\le\int_0^\infty e^{-st}|f(t)|^2t^{1/2}\,dt\int_0^\infty e^{-st}t^{-1/2}\,dt=\sqrt{\pi}s^{-1/2}\int_0^\infty e^{-st}|f(t)|^2t^{1/2}\,dt.$$

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 $f_1, \ldots, f_n$ 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 $F^{-1}(y) = \{x_1, \ldots, x_k\}$, where $JF(x_i) \neq 0$, 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

$$\left|\int_{-\infty}^{+\infty} f(x)\, dx\right| = \lim_{T \to +\infty} \int_{-\infty}^{+\infty} \left|(2T)^{-1} \int_{-T}^{T} f(x + t)\, dt\right| dx,$$

$$\int_0^1 \Bigl|\sum_{n=-\infty}^{\infty} f(x + n)\Bigr|\, dx = \lim_{N \to \infty} \int_{-\infty}^{+\infty} \Bigl|(2N + 1)^{-1} \sum_{n=-N}^{N} f(x + n)\Bigr|\, dx.$$

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^s_δ and H^s from §3.10(iii). Verify that if $s < t$ and $H^s(A) < \infty$, then $H^t(A) = 0$, and if $H^s_\delta(A) = 0$ 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 $H^s_\delta(B) > C$; in B we find a bounded set B' with $H^s_\delta(B') > C$, next in B' we find a compact set K with $H^s_\delta(K) > C$, 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

$$|\mu(E)| \geq \Gamma(n/2)\Bigl(2\sqrt{\pi}\Gamma\bigl((n+1)/2\bigr)\Bigr)^{-1}.$$

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 $\mu \circ f^{-1}\bigl(f(U)\bigr) = \nu \circ f^{-1}\bigl(f(U)\bigr)$ 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

$$\int_0^1 f(x) \, d\Phi(x) = \int_{\Phi(0)}^{\Phi(1)} f\bigl(\Phi^{-1}(y)\bigr) \, dy$$

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 nonmeasurable 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 $L^p\bigl([a,+\infty)\bigr)$ 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

$$\|f\|_\infty := \operatorname{esssup}_{x\in X}|f(x)| := \inf_{\Omega:\ \mu(X\backslash\Omega)=0}\ \sup_{x\in\Omega}|f(x)|, \quad f\in\mathcal{L}^\infty(\mu).$$

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. □

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 $\varepsilon \to 0$. □

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 $|f| \le 1$. 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 $f_j = fI_{B_j}$. We find m such that the integral of ϱ over $\mathbb{R}^n \backslash B_m$ is less than $\delta/4$. For $j \ge m+k$ 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

$$\|f - f * \varrho_\varepsilon\|_{L^1(B)} = \|f_j - f * \varrho_\varepsilon\|_{L^1(B)}$$
$$\le \|f_j - f_j * \varrho_\varepsilon\|_{L^1(B)} + \|(f_j - f) * \varrho_\varepsilon\|_{L^1(B)} \le \|f_j - f_j * \varrho_\varepsilon\|_{L^1(B)} + \frac{\delta}{2}.$$

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 $0 < \alpha < \beta \le 1$ 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 $|fg| \le f^2 + g^2$. 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) $(x,x)\geq 0$ 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:

$$|(x,y)| \leq \|x\|\,\|y\|. \tag{4.3.1}$$

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 $\|h-x_n\|^2\leq d^2+n^{-1}$. We show that the sequence $\{x_n\}$ is fundamental. To this end, it suffices to observe that

$$\|x_n-x_k\| \leq \frac{1}{\sqrt{n}}+\frac{1}{\sqrt{k}}.$$

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

$$\|h-p\|\leq \|h-x_n\|,\quad \|h-p\|\leq\|h-x_k\|.$$

It remains to apply the estimate

$$\|x_n-x_k\|\leq \|x_n-p\|+\|x_k-p\|\leq \frac{1}{\sqrt{n}}+\frac{1}{\sqrt{k}},$$

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

$$|L(\varphi)| \leq \int_X |\varphi|\,d\lambda \leq \|1\|_{L^2(\lambda)}\|\varphi\|_{L^2(\lambda)}.$$

The estimate $|L(\varphi_1-\varphi_2)| \leq \|1\|_{L^2(\lambda)}\|\varphi_1-\varphi_2\|_{L^2(\lambda)}$ 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 $\Omega = \{x\colon\ \psi(x) \leq 0\}$. Then Ω belongs to $\mathcal{A}$. Substituting in (4.3.2) the function $\varphi = I_\Omega$, we obtain

$$\mu(\Omega) = \int_\Omega \psi\,d\lambda \leq 0,$$

whence $\mu(\Omega) = 0$. Let $\Omega_1 = \{x\colon\ \psi(x) > 1\}$. By using that $\mu(\Omega_1) \leq \lambda(\Omega_1)$, 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 $f_n = fI_{\{\psi\geq 1/n\}}$ are bounded and increase pointwise to f such that

$$\int_X f_n\,d\mu = \int_X I_{\{\psi\geq 1/n\}}(1-\psi)\,d\lambda = \int_X I_{\{\psi\geq 1/n\}}\,d\nu \leq \nu(X).$$

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 $I_A I_{\{\psi\geq 1/n\}} \to I_A$ μ-a.e., hence ν-a.e. (here we use the absolute continuity of ν with respect to μ). Hence

$$\nu(A) = \lim_{n\to\infty}\int_X I_A I_{\{\psi\geq 1/n\}}\,d\nu = \lim_{n\to\infty}\int_X I_A I_{\{\psi\geq 1/n\}} f\,d\mu = \int_A f\,d\mu$$

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). □

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 $\left|\sum_{n=1}^N I_{A_n}(x)\right| \le |I_A(x)|$. 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. $\square$

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. □

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 $f \in L^p(\mu)$. □

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

$$\Bigl|\int_{A_{p+j}} f_i\,d\mu\Bigr| \le \int_{X_{p+j}} |f_i|\,d\mu \le C.$$

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

$$\begin{aligned}\Bigl|\int_{A_j} f_i\,d\mu\Bigr| &= \Bigl|\int_{B\cup A_j} f_i\,d\mu - \int_{B\backslash A_j} f_i\,d\mu\Bigr| \\ &\le \Bigl|\int_{B\cup A_j} f_i\,d\mu\Bigr| + \Bigl|\int_{B\backslash A_j} f_i\,d\mu\Bigr| \le 2m.\end{aligned}$$

Thus,

$$\Bigl|\int_{A} f_i\,d\mu\Bigr| \le 2mp + C(k-p) \quad \text{for all } i \text{ and } A \in \mathcal{A}.$$

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*

$$\int_A |f|\, d\mu < \varepsilon \quad \textit{for all } f \in \mathcal{F} \textit{ if } \mu(A) < \delta.$$

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 $f \in \mathcal{F}$. $\square$

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. $\square$

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 $X_0 = \bigcup_{n=1}^\infty \bigl\{x\colon |f_n(x)| \neq 0\bigr\}$. 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

$$M_{k,m} = \biggl\{ A \in \mathcal{A}\colon \ \biggl| \int_A (f_k - f_m)\, d\mu \biggr| \le \varepsilon \biggr\}, \quad k, m \in \mathbb{N}.$$

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

$$\biggl| \int_A (f_k - f_m)\, d\mu \biggr| \le \varepsilon \quad \text{if } \mu(A \,\triangle\, B) \le r. \tag{4.5.2}$$

Let us take a positive number $\delta < r$ such that, whenever $\mu(A) \le \delta$, one has

$$\biggl| \int_A f_j\, d\mu \biggr| \le \varepsilon, \quad j = 1, \ldots, n.$$

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

$$\biggl| \int_A f_j\, d\mu \biggr| \le 3\varepsilon$$

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

$$\left|\int_B f_n\,d\mu\right| \le \varepsilon \quad \text{for all } n \text{ if } \mu(B) < \delta.$$

Next we pick N such that $\mu(A_n) < \delta$ for all $n \ge N$. Then, for all $n \ge N$, we obtain $|\nu(A_n)| \le \varepsilon$, 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 $f \in L^1(\mu)$. $\square$

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. $\square$

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. $\square$

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

$$\begin{aligned}&\int G\bigl(|f(x)|\bigr)\,\mu(dx)\\ &\quad\le \alpha_1\mu\bigl(x\colon\ 1<|f(x)|\le 2\bigr)+(\alpha_1+\alpha_2)\mu\bigl(x\colon\ 2<|f(x)|\le 3\bigr)+\ldots\\ &\qquad\qquad = \sum_{n=1}^{\infty}\alpha_n\mu_n(f)=\sum_{n=1}^{\infty}\sum_{k=C_n}^{\infty}\mu_k(f).\end{aligned}$$

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 $x \ne 0$, there exists $l \in F$ with $l(x) \ne 0$). 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

$$\varepsilon^p(|a|^p+|b|^p)\le 4|a-b|^p \Rightarrow \Bigl|\frac{a+b}{2}\Bigr|^p \le (1-\delta)\frac{|a|^p+|b|^p}{2}. \tag{4.7.1}$$

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

$$\Bigl|\frac{f(x)+g(x)}{2}\Bigr|^p\le(1-\delta)\frac{|f(x)|^p+|g(x)|^p}{2},\quad \forall\, x\in\Omega. \tag{4.7.2}$$

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 $n \ge 1$. 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*

$$\lim_{n\to\infty} \int f_{n_k} g \, d\mu = \int f g \, d\mu, \quad \forall\, g \in L^\infty(\mu).$$

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 $f_{n,k} := f_n I_{\{|f_n| \le k\}}$, $n, k \in \mathbb{N}$. 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

$$\begin{aligned}\|g_k - g_l\|_{L^1(\mu)} &= \int (g_k - g_l)\mathrm{sign}(g_k - g_l)\, d\mu \\ &= \lim_{j\to\infty} \int (f_{n_j,k} - f_{n_j,l})\mathrm{sign}(g_k - g_l)\, d\mu \\ &\le \liminf_{j\to\infty} \|f_{n_j,k} - f_{n_j,l}\|_{L^1(\mu)} \to 0\end{aligned}$$

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

$$\int_{X \setminus X_n} |f_n| \, d\mu \geq \varepsilon.$$

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*

$$|\mu(A)| \leq \varepsilon \quad \textit{for all } \mu \in M \textit{ whenever } A \in \mathcal{A} \textit{ and } \nu(A) \leq \delta.$$

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 $\mu_1, \dots, \mu_n \in M$ 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 $\mu_1^n, \dots, \mu_{k_n}^n$ 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 $j = 1, \dots, k_n$, 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 $|\mu_n|(S_n) \geq \varepsilon > 0$. According to Exercise 4.7.137(ii), there exist sets $R_n \in \mathfrak{R}$ such that

$$|\mu_k|(S_n \triangle R_n) < \varepsilon 2^{-n}/4, \quad k \in \mathbb{N}. \tag{4.7.3}$$

Then $|\mu_n|(R_n) \geq 3\varepsilon/4$. 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

$$R_k \cap R_j \subset (S_k \triangle R_k) \cup (S_j \triangle R_j),$$

whence

$$|\mu_n|(R_k \cap R_j) \leq |\mu_n|(S_k \triangle R_k) + |\mu_n|(S_j \triangle R_j).$$

Hence $|\mu_n|\big(R_k \backslash (R_1 \cup \cdots \cup R_{k-1})\big) < \varepsilon/2$, whence $|\mu_n|(E_k \triangle R_k) < \varepsilon/2$. Thus, $|\mu_n|(E_n) \geq \varepsilon/4$, which leads to a contradiction with (ii) for $\mathfrak{R}$. □

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 $\pi_1 \leq \pi_2$ 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 $\pi_1 \leq \pi_3$, $\pi_2 \leq \pi_3$, 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

$$\overline{\rm osc} f|_A := \mu(A)^{-1}\int_A\biggl|f(x)-\mu(A)^{-1}\int_A f(y)\,\mu(dy)\biggr|\,\mu(dx)$$

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

$$\begin{aligned}
&\int_{C_m} \biggl|(f+h) - \int_{C_m} (f+h)\,d\mu\biggr|\,d\mu \\
&\le \int_{C_m} \biggl|(f+g) - \int_{C_m} (f+g)\,d\mu\biggr|\,d\mu + \int_{C_m} \biggl|(h-g) - \int_{C_m} (h-g)\,d\mu\biggr|\,d\mu \\
&\qquad \le \int_{C_m} \biggl|f - \int_{C_m} f\,d\mu\biggr|\,d\mu + 2\mu(C_m) \sup_{x\in C_m} |h(x) - g(x)| < \varepsilon\mu(C_m).
\end{aligned}$$

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

$$\biggl|\int_{A_j} f_N\,d\mu\biggr| < \varepsilon\mu(A_j), \quad \biggl|\int_{A_j} f\,d\mu - \int_{A_j} |f_N|\,d\mu\biggr| < \varepsilon\mu(A_j), \quad j = 1, \ldots, k.$$

Then, for every A_j, we obtain

$$\begin{aligned}
&\mu(A_j)^{-1} \int_{A_j} \biggl|f_N - \mu(A_j)^{-1} \int_{A_j} f_N\,d\mu\biggr|\,d\mu \\
&\qquad\ge \mu(A_j)^{-1} \int_{A_j} |f_N|\,d\mu - \mu(A_j)^{-1} \biggl|\int_{A_j} f_N\,d\mu\biggr| \\
&\qquad\qquad\ge \mu(A_j)^{-1} \int_{A_j} f\,d\mu - \varepsilon - \varepsilon \ge \varepsilon,
\end{aligned}$$

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 $j \ge k+1$. $\square$

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 $\lim_{n\to\infty} \|f - f_n\|_{L^1(\mu)} = 0$.*

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:

$$\begin{aligned}\int_X |f - g|\, d\lambda &= \int_X |\sqrt{f} - \sqrt{g}|(\sqrt{f} + \sqrt{g})\, d\lambda \\ &\le \left(2 - 2\int_X \sqrt{fg}\, d\lambda\right)^{1/2} \left(2 + 2\int_X \sqrt{fg}\, d\lambda\right)^{1/2}.\end{aligned}$$

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

$$\int_X f^\alpha g^{1-\alpha}\, d\lambda \le \left(\int_X f^{1/2} g^{1/2}\, d\lambda\right)^{1/p},$$

whence

$$1 - \int_X f^\alpha g^{1-\alpha}\, d\lambda \ge k_\alpha\left(1 - \int_X f^{1/2} g^{1/2}\, d\lambda\right).$$

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.$^\circ$ Let μ be a probability measure and let f be a measurable function such that $\sup_{p\geq 1} \|f\|_{L^p(\mu)} < \infty$. Prove that $f \in L^\infty(\mu)$.

Hint: use Chebyshev's inequality.

4.7.46.$^\circ$ 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.$^\circ$ Prove the formula for the Legendre polynomials in Example 4.3.7.

4.7.48.$^\circ$ 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.$^\circ$ 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.$^\circ$ (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 $\{f_n g_k\}$ 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 $f_n g_n - I_A I_B = (f_n - I_A)g_n + I_A(g_n - I_B)$ and Fubini's theorem that $f_n g_n \to I_A I_B$ in $L^p(\mu\otimes\nu)$. Applying this assertion in the case $p = 2$ and noting that the elements $f_n g_k$ have unit norms and are mutually orthogonal, we obtain the second claim. The reasoning in (ii) is much the same.

4.7.51.$^\circ$ 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, $n \geq 2$, 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 $\theta\colon\ r \mapsto \ln \|f\|^r_{L^r(\mu)}$ 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 $g \in \bigcup_{q\geq 1} L^q(\mu)$.

(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 $\{f \neq 0\}$ and $fg \in L^1(\mu)$ for every function $g \in L^q(\mu)$, where $1 < q \leq \infty$, 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 $1 \leq p < \infty$. 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 $\mu(|f| \geq c) < \infty$ 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 $gI_{\{|f|\geq c\}}$ 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

$$L(f) = \int f g_L \, d\mu \quad \text{for all } f \in L^1(\mu).$$

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

$$\sup\biggl\{\biggl|\int_X fg\,d\mu\biggr|:\ g \in \mathcal{P}, \|g\|_{L^1(\mu)} = 1\biggr\} < \infty.$$

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

$$\sum_{k=1}^{n} \mu(A_k)^{1-p} \left| \int_{A_k} f \, d\mu \right|^p \le C$$

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

$$\left| \int_0^1 fg \, dx \right| \le C^{1/p} \|g\|_q$$

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

$$\int_0^1 f_0 g \, dx = \int_0^1 fg \, dx$$

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 $L^p(\mathrm{IR}^n)$.

(i) (A.N. Kolmogorov; for $p = 1$, A.N. Tulaikov) Prove that the closure of $\mathcal{K}$ in $L^p(\mathrm{IR}^n)$ 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 $\sup_{f\in\mathcal{K}} \|f - S_r f\|_p \leq \varepsilon$, where $S_r f$ is Steklov's function defined by the equality

$$S_r f(x) := \lambda_n\bigl(B(x,r)\bigr)^{-1} \int_{B(x,r)} f(y)\,dy,$$

$B(x,r)$ 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

$$\sup_{f\in\mathcal{K}} \lim_{h\to 0} \int_{\mathrm{IR}^n} |f(x+h) - f(x)|^p\,dx = 0.$$

(iii) (V.N. Sudakov) Show that conditions (a) and (b) (or (a) and (b')) yield the boundedness of $\mathcal{K}$ in $L^p(\mathrm{IR}^n)$, 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 $g_\delta \in C_0^\infty(\mathrm{IR}^n)$ 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 $\|f - Sf\| \le 1$, $f \in \mathcal{K}$, 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 $\varphi(t+s) \leq N\varphi(t) + N\varphi(s)$. 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 $\varphi(2x) \leq C\varphi(x)$. 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 $\varphi(x+y) \leq C_1[\varphi(x) + \varphi(y)]$, $C_1 = \max(C/2, 1)$; then $\varphi \circ |f_n - f| \leq C_1[\varphi \circ |f_n| + \varphi \circ |f|]$. 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

$$\|f\|_\varphi := \inf\Bigl\{\alpha > 0\colon \ \int \varphi(|f|/\alpha)\,d\mu \leq 1\Bigr\}$$

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

$$f^*(t) = \inf\bigl\{s \geq 0\colon \ \mu\bigl(x\colon \ |f(x)| > s\bigr) \leq t\bigr\}$$

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 $C_k = \left(\bigcup_{j=1}^{p_k} A_j^k\right)\backslash\left(\bigcup_{j=1}^{p_k} B_j^k\right)$. Then

$$C_k \in \mathfrak{R}, \quad C_k \bigtriangleup E_k \subset \Bigl(\bigcup_{j=p_k+1}^{\infty} A_j^k\Bigr)\backslash\Bigl(\bigcup_{j=p_k+1}^{\infty} B_j^k\Bigr),$$

whence $|\mu_n|(C_k \bigtriangleup E_k) < \varepsilon 2^{-k}/4$ for all n and k. It is clear by the definition of C_k that $C_i \cap C_j \subset (C_i \bigtriangleup E_i) \cup (C_j \bigtriangleup E_j)$. 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

$$C_n \bigtriangleup D_n = \bigcup_{j=1}^{n-1} (C_n \cap C_j)$$

and by (4.7.22) we obtain $|\mu_n|(C_n \bigtriangleup D_n) < \varepsilon/2$. Hence $|\mu_n|(E_n \bigtriangleup D_n) < 3\varepsilon/4$. 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

$$\nu_n(S \bigtriangleup E_n) < \varepsilon 2^{-n}/4, \quad n \in \mathbb{N}.$$

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 $|\mu_n|(S \bigtriangleup D_p) < \varepsilon/4$. If $n > p$, then we obtain

$$|\mu_n|(S \bigtriangleup D_p) \leq |\mu_n|(S \bigtriangleup D_n) + |\mu_n|(D_n \bigtriangleup D_p) < \varepsilon.$$

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 $\widehat{f}\widehat{g}$ 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 $|f_n(x)| \leq g_n(x)$ for all n. Show that $|f(x)| \leq g(x)$ a.e. Construct an example demonstrating that the estimates $|f_n(x)| \leq |g_n(x)|$ do not imply that $|f(x)| \leq |g(x)|$ 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 $L^p(\mu)$, $1 \leq p < \infty$.

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 $0 \leq f \leq 1$, $f|_Q = 1$, $f = 0$ 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 $\varrho \geq 0$ 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 $t_1 \leq t_2 \leq \cdots \leq t_{n+1}$ 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

$$\sup_{t\in T} |f(t)| \leq |f(t_0)| + V(f,T).$$

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,

$$V_a^b(\alpha f + \beta g) \leq |\alpha| V_a^b(f) + |\beta| V_a^b(g) \tag{5.2.1}$$

for any two functions f and g of bounded variation and arbitrary real numbers α and β. This is obvious from the estimate

$$\bigl|\alpha f(t_{i+1}) + \beta g(t_{i+1}) - \alpha f(t_i) - \beta g(t_i)\bigr| \\ \leq |\alpha|\,|f(t_{i+1}) - f(t_i)| + |\beta|\,|g(t_{i+1}) - g(t_i)|.$$

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 $x \geq y$ we have

$$V(x) - V(y) = V_y^x(f) \geq |f(x) - f(y)| \geq f(x) - f(y).$$

Then $|V(x) - V(y)| \geq |f(x) - f(y)|$, 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 $k \leq n$ 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

$$\lambda^*\bigl(E \bigcap \textstyle\bigcup_{j=1}^k I_j\bigr) \geq \lambda^*(E_n) - \lambda^*\bigl(E_n \backslash \textstyle\bigcup_{j=1}^k I_j\bigr) > \lambda^*(E) - \varepsilon.$$

Finally, one has $\lambda\bigl(\bigcup_{j=1}^k I_j\bigr) \leq \lambda(G) < \lambda^*(E) + \varepsilon$. □

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 $\mu(E) = 0$. □

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

$$V(f,[a,b]) \le \int_a^b |g(x)|\,dx.$$

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

$$\Omega_+ = \{x\colon\ g(x)\ge 0\},\quad \Omega_- = \{x\colon\ g(x)<0\}.$$

Then we find finitely many pairwise disjoint intervals (a_1,b_1), ...,(a_n,b_n) in $[a,b]$ such that

$$\lambda\Bigl(\Omega_+ \triangle \bigcup_{i=1}^n (a_i,b_i)\Bigr) < \delta. \tag{5.3.3}$$

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

$$\lambda\Bigl(\Omega_- \triangle \bigcup_{i=1}^k (c_i,d_i)\Bigr) < \delta.$$

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

$$\sum_{i=1}^n |f(b_i)-f(a_i)| \ge \sum_{i=1}^n \int_{a_i}^{b_i} |g(x)|\,dx - \frac{1}{4}\varepsilon \ge \int_{\Omega_+} |g(x)|\,dx - \frac{1}{2}\varepsilon.$$

Similarly, we obtain

$$\sum_{i=1}^k |f(d_i)-f(c_i)| \ge \int_{\Omega_-} |g(x)|\,dx - \frac{1}{2}\varepsilon.$$

Thus,

$$V(f,[a,b]) \ge \sum_{i=1}^n |f(b_i)-f(a_i)| + \sum_{i=1}^k |f(d_i)-f(c_i)| \ge \int_a^b |g(x)|\,dx - \varepsilon,$$

which completes the proof. □

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 $\lambda(\Omega \bigtriangleup A) < \delta$. Then

$$\int_\Omega f(x)\,dx \le \int_A f(x)\,dx + \int_{\Omega\bigtriangleup A} |f(x)|\,dx \le \int_A f(x)\,dx + \varepsilon = \varepsilon.$$

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

$$\begin{aligned}\int_a^x \frac{F(y+h_n)-F(y)}{h_n}\,dy &= \frac{1}{h_n}\int_{a+h_n}^{x+h_n} F(y)\,dy - \frac{1}{h_n}\int_a^x F(y)\,dy \\ &= \frac{1}{h_n}\int_x^{x+h_n} F(y)\,dy - \frac{1}{h_n}\int_a^{a+h_n} F(y)\,dy,\end{aligned}$$

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. □

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

$$\lambda(B_k) > \frac{1}{2}\sup\bigl\{\lambda(B)\colon\ B \in S_{k-1}\bigr\}, \quad k > 1, \tag{5.5.1}$$
$$S_{k-1} := \bigl\{B \in S\colon\ B \cap (B_1 \cup \cdots \cup B_{k-1}) = \varnothing\bigr\}.$$

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 $\lambda(B_{m+1})\ge\frac{1}{2}\lambda(B)$. 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 $\lambda(B_l)\ge\frac{1}{2}\lambda(B)$. 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 $B \subset V_j$. $\square$

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*

$$\lambda\bigl(x\colon\ (Mf)(x) > t\bigr) \le \frac{C_n}{t} \int_{\mathbb{R}^n} |f(y)|\, dy, \qquad (5.6.1)$$

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 $E_t = \{x\colon\ (Mf)(x) > t\}$. 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

$$E_t := \{x\colon\ (Mf)(x) > t\} \subset \{x\colon\ (Mg)(x) > t/2\},$$

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

$$2C_n p|f(y)|\int_0^{2|f(y)|} t^{p-2}\,dt = 2C_n p|f(y)|\frac{|2f(y)|^{p-1}}{p-1} = 2^p C_n\frac{p}{p-1}|f(y)|^p.$$

Thus, the indicated double integral equals $2^pC_n\frac{p}{p-1}\|f\|_{L^p}^p$. 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*

$$\lim_{r\to 0} \frac{1}{\lambda\bigl(B(x,r)\bigr)} \int_{B(x,r)} |f(y) - f(x)|\, dy = 0. \qquad (5.6.3)$$

For every point x with such a property, called a Lebesgue point of f, one has

$$\lim_{r\to 0} \frac{1}{\lambda\bigl(B(x,r)\bigr)} \int_{B(x,r)} f(y)\, dy = f(x). \qquad (5.6.4)$$

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

$$\limsup_{k\to\infty} \frac{1}{\lambda\bigl(B\bigl(x,r_k(x)\bigr)\bigr)} \int_{B(x,r_k(x))} f(y)\, dy \ge f(x) + \frac{1}{m} \qquad (5.6.5)$$

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 $E_0 = \bigl\{x \in U\colon\ (Mf_1)(x) \ge f_1(x) + m^{-1}\bigr\}$. The measure of E_0 is estimated by means of Chebyshev's inequality and (5.6.1) as follows:

$$\begin{aligned}
\lambda(E_0) &\le \lambda\Bigl(x \in E_0\colon\ f_1(x) \le -\frac{1}{2m}\Bigr) + \lambda\Bigl(x \in E_0\colon\ f_1(x) \ge -\frac{1}{2m}\Bigr)\\
&\le 2m\|f_1\|_{L^1} + \lambda\Bigl(x \in E_0\colon\ (Mf_1)(x) \ge \frac{1}{2m}\Bigr)\\
&\le 2m\|f_1\|_{L^1} + 2mC_n\|f_1\|_{L^1} \le 2m\varepsilon(C_n+1).
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we obtain $\lambda(E) = 0$. Replacing f by $-f$ we obtain that the set of all points x where

$$\liminf_{k\to\infty} \frac{1}{\lambda\bigl(B\bigl(x,r_k(x)\bigr)\bigr)} \int_{B(x,r_k(x))} f(y)\, dy \le f(x) - \frac{1}{m}$$

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. □

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 $\mathcal{K}$. □

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*

$$\limsup_{R_0\in\mathcal{R}_0,\,\mathrm{diam}(R_0)\to 0} \frac{1}{\lambda(R_0)}\int_{R_0} f(x-y)\,dy = +\infty \quad \textit{for a.e. } x.$$

(ii) *There is a compact set* $K \subset \mathbb{R}^2$ *of positive measure such that*

$$\liminf_{R\in\mathcal{R},\,\mathrm{diam}(R)\to 0} \frac{1}{\lambda(R)}\int_{R} I_K(x-y)\,dy = 0 \quad \textit{for all } x.$$

(iii) *If* $f \in \mathcal{L}^p(\mathbb{R}^n)$, *where* $p>1$, *then*

$$\lim_{R_0\in\mathcal{R}_0,\,\mathrm{diam}(R_0)\to 0} \frac{1}{\lambda(R_0)}\int_{R_0} f(x-y)\,dy = f(x) \quad \textit{for a.e. } x.$$

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 $\delta := \max\limits_i |x_i - x_{i-1}|$ 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* $[c,d]\subset\big(x-\delta(x), x+\delta(x)\big)$. *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. $\square$

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

$$|f(u)-f(x)-f'(x)(u-x)|\le\varepsilon|u-x|, \qquad \forall\, u\in\bigl(x-\delta(x),x+\delta(x)\bigr)\cap[a,b]. \tag{5.7.2}$$

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

$$\bigl|I(f', \mathcal{P}) - [f(b) - f(a)]\bigr| \le \Bigl|I(f', \mathcal{P}_1) - \sum_{i \in J_1} [f(b_i) - f(a_i)]\Bigr| \\ + \sum_{i \in J_0} |f(b_i) - f(a_i)| \le \varepsilon(b-a) + 2\varepsilon,$$

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 $\lambda(A_n) = 0$. $\square$

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. □

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. □

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,

$$\begin{aligned}\Big|I(|f|, \widehat{\mathcal{P}}_1) - I(|f|, \widehat{\mathcal{P}}_2)\Big| &= \Big|I(|f|, \widehat{\mathcal{P}}_1') - I(|f|, \widehat{\mathcal{P}}_2')\Big| \\ &\leq \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\big||f(x_i)| - |f(y_j)|\big|\lambda(I_i \cap K_j) \\ &\leq \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}|f(x_i) - f(y_j)|\lambda(I_i \cap K_j).\end{aligned}$$

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

$$\delta(x) = \mathrm{dist}(x, [0, 1] \backslash U_n).$$

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 $\varepsilon\lambda(A_i)/4$:

$$\left|I(f,\widehat{\mathcal{P}}) - \int_0^1 f(x)\,dx\right| = \left|\sum_{i=1}^k \int_{a_i}^{b_i} [f(x_i) - f(x)]\,dx\right|$$
$$\leq \sum_{i=1}^k \int_{A_i} |f(x_i) - f(x)|\,dx + \sum_{i=1}^k |f(x_i)|\lambda(B_i) + \sum_{i=1}^k \int_{B_i} |f(x)|\,dx < \varepsilon.$$

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. □

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

$$I_k := \{j\colon\ j < k, B_j \cap B_k \neq \varnothing\}, \quad M_k := I_k \cap \{j\colon\ r_j \le 3r_k\}.$$

Let us show that

$$\mathrm{Card}(M_k) \le 20^n. \tag{5.8.1}$$

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

$$\begin{aligned}\lambda_n\bigl(B(a_k, 5r_k)\bigr) &\ge \sum_{j\in M_k} \lambda_n\bigl(B(a_j, r_j/3)\bigr) = \sum_{j\in M_k} C_n r_j^n 3^{-n} \\ &\ge \sum_{j\in M_k} C_n r_k^n 4^{-n} = \mathrm{Card}(M_k) C_n r_k^n 4^{-n},\end{aligned}$$

where $\lambda_n\bigl(B(a, r)\bigr) = C_n r^n$. Hence the obtained estimates yield the inequality $5^n \ge \mathrm{Card}(M_k) 4^{-n}$.

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 $1 \le i, j < k$, $B_i \cap B_k \neq \varnothing$, $B_j \cap B_k \neq \varnothing$, $r_i > 3r_k$, $r_j > 3r_k$. 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

$$\mathfrak{m}(A \cap U) \le \sum_{j=1}^{N_n} \mathfrak{m}\biggl((A \cap U) \cap \Bigl(\bigcup_{B \in \mathcal{F}^1_j} B\Bigr)\biggr).$$

So, there exists $j \in \{1, \ldots, N_n\}$ with

$$\mathfrak{m}\biggl((A \cap U) \cap \Bigl(\bigcup_{B \in \mathcal{F}^1_j} B\Bigr)\biggr) \ge \frac{1}{N_n}\mathfrak{m}(A \cap U).$$

Therefore, there exists a finite collection $B_1, \ldots, B_{k_1} \in \mathcal{F}^1_j$ such that

$$\mathfrak{m}\biggl((A \cap U) \cap \Bigl(\bigcup_{i=1}^{k_1} B_i\Bigr)\biggr) \ge (1-\alpha)\mathfrak{m}(A \cap U),$$

which yields (5.8.5), since

$$\mathfrak{m}(A \cap U) = \mathfrak{m}\biggl((A \cap U) \cap \Bigl(\bigcup_{i=1}^{k_1} B_i\Bigr)\biggr) + \mathfrak{m}\biggl((A \cap U)\backslash\Bigl(\bigcup_{i=1}^{k_1} B_i\Bigr)\biggr)$$

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. $\square$

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

$$\begin{aligned}\left|\int_{\mathbb{R}^n} f(x_0-y)\varrho_\varepsilon(y)\,dy\right| &\le \int_{\mathbb{R}^n} |f(x_0-y)||\varrho_\varepsilon(y)|\,dy\\ &\le \int_{|y|\le r} |f(x_0-y)||\varrho_\varepsilon(y)|\,dy + \int_{|y|>r} C|\varrho_\varepsilon(y)|\,dy\\ &= \int_{|y|\le r} r^{-n}|f(x_0-y)|R^n|\varrho(y/\varepsilon)|\,dy + C\int_{|y|\ge R}|\varrho(z)|\,dz\\ &\le \frac{\delta M}{2M+1} + \frac{\delta C}{2C+1} < \delta,\end{aligned}$$

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. $\square$

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. □

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*

$$\Bigl(\int_U |f - f_U|^{\frac{np}{n-p}}\, dx\Bigr)^{\frac{n-p}{np}} \le C(n,p)\Bigl(\int_U |\nabla f|^p\, dx\Bigr)^{1/p},$$

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

$$Df(B) := \bigl(Df_1(B), \ldots, D_n f(B)\bigr).$$

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:

$$\Bigl(\int_{\mathbb{R}^n} |f(x)|^{n/(n-1)}\, dx\Bigr)^{(n-1)/n} \le c_n\|Df\|, \quad \forall\, f \in BV(\mathbb{R}^n). \tag{5.8.11}$$

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*

$$\lambda_n\bigl(x \in \mathbb{R}^n\colon\ f_\varepsilon(x) \neq f(x)\bigr) \leq \varepsilon.$$

If $f \in W^{p,1}(\mathbb{R}^n)$, where $p \in [1, +\infty)$, then f_ε can be chosen such that, in addition, $\|f - f_\varepsilon\|_{W^{p,1}(\mathbb{R}^n)} \leq \varepsilon$.

Let $W^{p,1}_{\text{loc}}(\mathbb{R}^n)$ 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 $W^{p,1}_{\text{loc}}(\mathbb{R}^n, \mathbb{R}^k)$.

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 $f \in W^{p,1}_{\text{loc}}(\mathbb{R}^n, \mathbb{R}^k)$, 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 $|\det(\partial_{x_i} f_j)_{i,j\leq n}|$. Let $\mathrm{Card} M$ denote the cardinality of the set M. As above, H^α denotes the Hausdorff measure.

5.8.28. Lemma. *Let $n \leq k$ 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) $\alpha^{-1}|G_j(u)| \leq |Df(x)(u)| \leq \alpha |G_j(u)|$ *for all $x \in B_j$, $u \in \mathbb{R}^n$,*

(iii) $\alpha^{-n}|\det G_j| \leq |Jf(x)| \leq \alpha^n |\det G_j|$ *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

$$(\alpha^{-1}+\varepsilon)|G(u)| \le |Df(x)(u)| \le (\alpha-\varepsilon)|G(u)|, \quad \forall\, u \in \mathbb{R}^n, \qquad (5.8.12)$$

$$|f(y)-f(x)-Df(x)(y-x)| \le \varepsilon|G(y-x)|, \quad \forall\, y \in B(x,1/j). \qquad (5.8.13)$$

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 $|\det T_1| \le |\det T_2|$. $\square$

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*

$$\int_{A\cap f^{-1}(B)} |Jf(x)|\,dx = \int_B \mathrm{Card}\,\bigl(A\cap f^{-1}(y)\bigr)\,H^n(dy), \qquad (5.8.14)$$

(ii) *if $n > k$, then*

$$\int_A |Jf(x)|\,dx = \int_{\mathbb{R}^k} H^{n-k}\bigl(A\cap f^{-1}(y)\bigr)\,dy. \qquad (5.8.15)$$

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 $\mathrm{Card}\bigl(A_j \cap f^{-1}(y)\bigr) = I_{A_j}(y)$, we obtain

$$\alpha^{-2n}\int_{\mathbb{R}^k} \mathrm{Card}\bigl(A \cap f^{-1}(y)\bigr)\,H^n(dy) \le \int_A |Jf(x)|\,dx \\ \le \alpha^{2n}\int_{\mathbb{R}^k} \mathrm{Card}\bigl(A \cap f^{-1}(y)\bigr)\,H^n(dy).$$

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*

$$u\bigl(f(x)\bigr)|Jf(x)|I_E(x), \quad u(y)\mathrm{Card}\bigl(f^{-1}(y) \cap E\bigr)$$

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

$$\int_E u\bigl(f(x)\bigr)|Jf(x)|\,dx = \int_{\mathbb{R}^n} u(y)\mathrm{Card}\bigl(f^{-1}(y) \cap E\bigr)\,dy.$$

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 $W^{p,1}_{\rm loc}$ 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 $JF \neq 0$ 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 $F_j, F \in W^{p,1}_{\rm loc}(\mathbb{R}^n, \mathbb{R}^n)$, 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 $JF \neq 0$ 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*

$$\sigma_S(S) := H^n(S) = \int_D \sqrt{1 + |\nabla f(x)|^2}\, dx. \qquad (5.8.16)$$

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*

$$\alpha < \frac{1}{\lambda_n(Q_k)} \int_{Q_k} f(x)\,dx \leq 2^n \alpha; \tag{5.8.17}$$

(ii) *$f(x) \leq \alpha$ 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:

$$\int_{Q''} f\,dx > \alpha\lambda_n(Q'') \qquad \text{or} \qquad \int_{Q''} f\,dx \leq \alpha\lambda_n(Q'').$$

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

$$\alpha < \frac{1}{\lambda_n(Q'')} \int_{Q''} f\,dx \leq \frac{2^n}{\lambda_n(Q')} \int_{Q'} f\,dx \leq 2^n \alpha. \tag{5.8.18}$$

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

$$\int_{Q'} f\,dx \leq \alpha\lambda_n(Q').$$

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 $f(x) \leq \alpha$ 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

$$f(x) = \lim_{j\to\infty} \frac{1}{\lambda_n(K_j)} \int_{K_j} f\,dy \leq \alpha,$$

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

$$|f(b)-f(a)| \leq \lambda\bigl(f([a,b])\bigr) = \lambda\bigl(f([a,b]\cap D)\bigr) \leq \int_a^b |f'(x)|\,dx,$$

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 $\lambda(E)\geq\alpha$, $\lambda\bigl(f([a,b]\backslash f(E)\bigr)=0$ 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

$$-\lambda\bigl(f(N)\bigr) \leq f(b)-f(a) \leq \lambda\bigl(f(P)\bigr).$$

Deduce the existence of points of differentiability of f.

HINT: we may assume that $f(a)\leq f(b)$ (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 $\nu\bigl((a, x]\bigr)$, 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 $F_\mu(t) = \mu\bigl((-\infty, t)\bigr)$. 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 $V\bigl(F_\mu, (-\infty, x]\bigr) = V\bigl(F_\mu, (-\infty, x)\bigr) = |\mu|\bigl((-\infty, x)\bigr)$ and $\|\mu\| = V(F_\mu, \mathbb{R}^1)$.

HINT: the left-hand side is estimated by the right-hand side, since

$$\sum_{i=1}^n \bigl|\mu\bigl([a_i, b_i)\bigr)\bigr| \le |\mu|\Bigl(\bigcup_{i=1}^n [a_i, b_i)\Bigr)$$

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 $|\mu|\bigl((-\infty, x)\bigr) \le \mu(K_1) - \mu(K_2) + \varepsilon$. Next take disjoint intervals $[a_1, b_1), \ldots, [a_n, b_n)$ and $[c_1, d_1), \ldots, [c_k, d_k)$ with

$$\mu(K_1) \le \sum_{i=1}^{n} \mu\bigl([a_i, b_i)\bigr) + \varepsilon, \quad |\mu(K_2)| \le \sum_{i=1}^{k} \bigl|\mu\bigl([c_i, d_i)\bigr)\bigr| + \varepsilon.$$

Then $|\mu|\bigl((-\infty, x)\bigr) \le V\bigl(F_\mu, (-\infty, x)\bigr) + 3\varepsilon$.

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. $V(f, [a, x]) = V\bigl(f, [a, x)\bigr) = F_{|\mu|}(x) = F_{|\mu_0|}(x) + F_{|\nu|}(x)$. 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 $\lambda_n^*\bigl(F(E)\bigr) \le C^n n^{n/2} \sum_{j=1}^\infty \|DF\|_{L^p(Q_j)}^n r_j^{\alpha n}$, 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 $\mathrm{diam}Q_n(x) \to 0$. 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 $\mathrm{ap}\lim\limits_{y\to x} f(y)=f(x)$.

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 $0 < \varphi(x) \leq 1$ 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 $\bmod 2\pi$. 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, $\inf_{\varepsilon>0}\bigl\{\mu(|f-g| \geq \varepsilon)+\varepsilon\bigr\}$, and Ky Fan introduced the metric $\inf_{\varepsilon>0}\bigl\{\mu(|f-g| \geq \varepsilon) \leq \varepsilon\bigr\}$. 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 $f\bigl((x+y)/2\bigr) \le f(x)/2 + f(y)/2$, 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 $f_n \ge 0$, 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 $R_k \in \mathfrak{R}$, $k \leq n$, 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

Measure Theory

Volume II

V.I. Bogachev

Measure Theory

Volume II

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 to Volume 2

Introductory notes on Volume 2 appear in the general introduction in Volume 1, so we confine ourselves to several remarks of a more technical nature. Chapter 6 has partly an auxiliary character; yet, I hope, the reader will find a lot of interesting and useful things also in this chapter. It contains a brief exposition of the basic facts about Borel and Baire sets and Souslin spaces, including several measurable selection theorems. Chapter 7 is devoted to measures on topological spaces. Among the diverse classes of measures discussed here, Radon measures are the most important. Along with the properties of measures, we study the properties of the corresponding functionals on spaces of functions, in particular, the Riesz theorem and its generalizations. In spite of the considerable length of this chapter (the longest in the book), the subsequent chapters use a relatively small number of its results and constructions. Chapter 8 gives a modern presentation of the theory of weak convergence of measures. In particular, we consider metrics and topologies on spaces of measures and weak compactness. Chapter 9 is concerned with nonlinear transformations of measures and isomorphisms of measure spaces, including the theory of Lebesgue–Rohlin spaces. Finally, Chapter 10 is devoted to conditional measures and conditional expectations. In addition to the classical results and various subtleties related to these objects, we give a brief introduction to the theory of martingales (at a level meeting the basic needs of measure theory) and present a number of results from ergodic theory that are directly linked to measure theory and illustrate its ideas and methods. All these chapters are almost independent in the technical sense (so that they can be read selectively with minimal reference to the previous material or can be used for preparing various special courses), but, as one can easily observe, in the sense of ideas they are all strongly connected and altogether form the foundations of modern measure theory. The study of various transformations of measures is the leitmotiv of this volume.

The numeration of chapters continues the numeration of Volume 1. The references to assertions, remarks, and exercises comprise the chapter number, section number, and assertion number. For example, Definition 1.1.1 is found in §1 of Chapter 1 (i.e., in Volume 1), and within each section all the assertions are numbered consecutively independently of their type. The numeration of formulas is organized similarly, but the formula numbers are given in brackets.

The bibliographical and historical comments on this volume concern only the chapters in this volume, but on several occasions they interrelate with the comments in Volume 1. It is reasonable to consider all the comments as one essay presented in two parts. At the end of this volume the reader will find the cumulative bibliography for both volumes, in which the works cited only in Volume 1 are marked by the asterisk (without indication of pages where they are cited), and in the works cited in both volumes, the page numbers referring to Volume 1 and Volume 2 are preceded by **I** and **II**, respectively; the absence of such indicators means that the work is cited only in the present volume.

The book is completed by the cumulative author and subject indices to both volumes, where the page numbers referring to Volume 1 and Volume 2 are preceded by **I** and **II**, respectively.

Finally, knowledge of all the material of Volume 1 is not assumed in this volume. For most of this volume it is enough to be acquainted with the basic course from Volume 1; however, it is necessary to be familiar with the standard university course of functional analysis including elements of general topology. In those cases where we have to resort to the results in the complementary material of Volume 1, the exact references are provided. Some additional necessary facts are presented in the appropriate places.

Comments and remarks can be sent to vibogach@mail.ru.

Moscow, August 2006 Vladimir Bogachev

Contents

Contents of Volume 1

CHAPTER 6

Borel, Baire and Souslin sets

Now we have already not a single mathematical space, but infinitely many of them, and it is unknown which one is the most adequate model of the space of the physical reality. So one has to construct samples of different spaces in an analytical way.

A.N. Kolmogorov. Modern controversies on the nature of mathematics.

6.1. Metric and topological spaces

In this section, we recall the basic concepts related to topological spaces and prove several facts necessary for the sequel. In addition, we give some examples of topological spaces interesting from the point of view of measure theory. Our presentation is oriented towards a reader acquainted with metric spaces, but without topological background. The information given here is sufficient for understanding the main part of the text (it is most important to be familiar with the concepts of compactness and continuity). However, the reader is warned that for mastering a number of more special examples and many complementary results in §6.10 and the concluding sections in other chapters, it is necessary to have at least minimal topological background (in spite of the fact that formally all the necessary concepts are introduced). More details can be found in Kuratowski [**1082**], Engelking [**532**]. The term "a topological space (X,τ)" denotes a set X with a family τ of its subsets containing X and the empty set and closed with respect to finite intersections and arbitrary unions. The sets in the family τ are called open. Actually a shorter term "a topological space X" is used, which means, of course, that the family of open sets τ (called a topology in X) is fixed. One has to indicate a topology explicitly when on the same set X several different topologies are introduced. Such a situation will be encountered below. A base of the topology is a family of open sets such that every nonempty open set is a union of some sets in this family.

A neighborhood of a point in a topological space is any open set containing this point. A point x in a set A is called isolated if it has a neighborhood not containing other points from A. A point a is called a limit point of the set A if every neighborhood of a contains a point $b \neq a$ from A.

A set in a topological space is called closed if its complement is open. The closure of a set A in a topological space X is defined as the intersection of all closed sets containing A (i.e., the smallest closed set containing A).

Every subset X_0 of a topological space X is a topological space with the *induced topology* that consists of all sets $U \cap X_0$, where U is open in X.

An important subclass of topological spaces is the class of metric spaces. We recall that a metric space (X, ϱ) is a set X endowed with a function $\varrho\colon X^2 \to [0, +\infty)$ (called a metric) possessing the following properties:

(1) $\varrho(x, y) = 0$ precisely when $x = y$;

(2) $\varrho(x, y) = \varrho(y, x)$ for all $x, y \in X$;

(3) $\varrho(x, z) \le \varrho(x, y) + \varrho(y, z)$ for all $x, y, z \in X$.

Let a be a point in a metric space (X, ϱ) and $r > 0$. The sets

$$\{x \in X\colon\ \varrho(x, a) < r\} \quad \text{and} \quad \{x \in X\colon\ \varrho(x, a) \le r\}$$

are called the open and closed balls, respectively, with the center a and radius r.

It is readily verified by using property (3) that the family of all open sets in a metric space X (i.e., sets in which every point is contained with some ball of a positive radius centered at that point) satisfies the above axioms of a topological space. Below we encounter many important examples of topological spaces whose topology is not generated by a metric.

A topological space (X, τ) is called metrizable if there is a metric on X such that the collection of all open sets for this metric is precisely τ.

It is worth noting that essentially different metrics may generate the same topology. For example, the usual metric on $\mathbb{R}^1$ generates the same topology as the bounded metric $\varrho(x, y) = |x - y|/(1 + |x - y|)$.

A locally convex space is a linear space X equipped with a family of seminorms p_α, $\alpha \in A$, such that for every $x \ne 0$ there is p_α with $p_\alpha(x) > 0$. Such a family generates a topology on X whose base consists of the sets

$$U_{x_0, \alpha_1, \ldots, \alpha_n, \varepsilon} := \bigl\{x \in X\colon\ p_{\alpha_i}(x - x_0) < \varepsilon,\, i = 1, \ldots, n\bigr\},\ \varepsilon > 0.$$

Some special cases have already been considered in Chapter 4. Complete metrizable locally convex spaces are called Fréchet spaces.

A mapping f from a topological space X to a topological space Y is called continuous at a point x if, for every nonempty open set W containing $f(x)$, there exists a nonempty open set U containing x such that $f(U) \subset W$. A mapping is called continuous if it is continuous at every point. It is left to the reader to verify that the continuity of a mapping $f\colon X \to Y$ is equivalent to the following: for every open set $W \subset Y$, the set $f^{-1}(W)$ is open in X, or, equivalently, for every closed set $Z \subset Y$, the set $f^{-1}(Z)$ is closed in X. Note, however, that the image of an open set may not be open. A mapping is called open if it takes every open set to an open one. The class of all continuous mappings from X to Y is denoted by $C(X, Y)$; if $Y = \mathbb{R}^1$, then this class is denoted by $C(X)$. The set of all bounded functions in $C(X)$ is denoted by $C_b(X)$. It is easily verified that $C_b(X)$ is a Banach space with the norm

$\|f\| = \sup_{x\in X} |f(x)|$. A family F of functions on a topological space X is said to be equicontinuous at a point x if for every $\varepsilon > 0$, there is a neighborhood U of x such that $|f(x) - f(y)| < \varepsilon$ for all $y \in U$ and all $f \in F$. A family F of functions on a locally convex space X is said to be uniformly equicontinuous if for every $\varepsilon > 0$, there is a neighborhood of zero U in X such that $|f(x) - f(y)| < \varepsilon$ for all $f \in F$ if $x - y \in U$. Both notions are similarly defined for mappings with values in metric or locally convex spaces or mappings on metric spaces.

In the study of topological spaces, the concept of a net is very useful. This concept generalizes that of a sequence to the case of an uncountable index set.

A nonempty set T is called directed if it is equipped with a partial order (see §1.12(vi)) satisfying the following condition: for each $t, s \in T$, there exists $u \in T$ with $t \le u$ and $s \le u$.

A directed set may contain elements that are not comparable. For example, $\mathbb{R}^2$ can be equipped with the partial order $(x, y) \le (u, v)$ defined by $x \le u$, $y \le v$. Clearly, not all elements are comparable, but every two are majorized by a certain third element.

A net in X is a family of elements $\{x_t\}_{t\in T}$ in X indexed by a directed set T. Similarly, we define a net of sets $\{U_t\}_{t\in T}$ in X. A net $\{x_t\}_{t\in T}$ is called a subnet of a net $\{y_s\}_{s\in S}$ if there is a mapping $\pi\colon T \to S$ such that $x_t = y_{\pi(t)}$ and, for each $s_0 \in S$, there exists $t_0 \in T$ with $\pi(t) \ge s_0$ for all $t \ge t_0$. A net of sets $\{A_t\}_{t\in T}$ in a space X is called decreasing if $A_t \subset A_s$ whenever $s \le t$. Such a net is called decreasing to the set $\bigcap_{t\in T} A_t$. A net of real functions $\{f_t\}_{t\in T}$ on a space X is called decreasing if $f_t \le f_s$ whenever $s \le t$. Similarly, one defines increasing nets of sets and functions. In the case of an increasing net of sets A_t one says that it increases to the set $\bigcup_{t\in T} A_t$. The corresponding notation: $A_t \downarrow A$, $A_t \uparrow A$, $f_t \downarrow f$, $f_t \uparrow f$. A net $\{x_t\}_{t\in T}$ in a topological space X converges to an element x if, for every nonempty open set U containing x, there exists an index t_0 such that $x_t \in U$ for all $t \in T$ with $t_0 \le t$. Notation: $\lim\limits_t x_t = x$. It is worth noting that convergence of a countable net is not the same as convergence of a sequence. For example, let $T = \mathbb{Z}$ be equipped with the usual ordering and let $x_n = n^{-1}$ if $n \ge 0$ and $x_n = n$ otherwise. Then the countable net $\{x_n\}$ converges to zero, but is not even bounded. The following simple fact is left as Exercise 6.10.20.

6.1.1. Lemma. *Let X and Y be two topological spaces. A mapping $f\colon X \to Y$ is continuous at a point x precisely when for every net x_α convergent to x, the net $f(x_\alpha)$ converges to $f(x)$.*

The reasoning analogous to the proof of this lemma shows that every point x in the closure of a set A in a topological space X is either an isolated point of A or the limit of some net of points in A (such points are called limit points or cluster points of A).

A mapping $f\colon X \to Y$ between topological spaces is called *a homeomorphism* if it maps X one-to-one on Y and both mappings f and f^{-1} are continuous. Topological spaces between which there is a homeomorphism are called *homeomorphic*.

An important role in the theory of topological spaces is played by diverse separation axioms. We need only the few simplest ones listed below.

6.1.2. Definition. *Let X be a topological space.* (i) *X is called Hausdorff if every two distinct points in X possess disjoint neighborhoods.*

(ii) *A Hausdorff space X is called regular if, for every point $x \in X$ and every closed set Z in X not containing x, there exist disjoint open sets U and V such that $x \in U$, $Z \subset V$.*

(iii) *A Hausdorff space X is called completely regular if, for every point $x \in X$ and every closed set Z in X not containing x, there exists a continuous function $f\colon X \to [0,1]$ such that $f(x) = 1$ and $f(z) = 0$ for all $z \in Z$.*

(iv) *A Hausdorff space X is called normal if, for all disjoint closed sets Z_1 and Z_2 in X, there exist disjoint open sets U and V such that $Z_1 \subset U$ and $Z_2 \subset V$.*

(v) *A Hausdorff space is called perfectly normal if every closed set $Z \subset X$ has the form $Z = f^{-1}(0)$ for some continuous function f on X.*

Sets of the form indicated in (v) are called *functionally closed*.

It is clear that any metric space satisfies all conditions (i)–(v). For example, for f in (v) one can take $f(x) = \operatorname{dist}(x, Z)$, where the distance $\operatorname{dist}(x, Z)$ from the point x to the set Z is defined as the infimum of distances from x to points in Z. Throughout we consider only Hausdorff spaces.

6.1.3. Lemma. *For any nonempty disjoint closed sets Z_1 and Z_2 in a metric space, there exists a continuous function f such that $Z_1 = f^{-1}(0)$ and $Z_2 = f^{-1}(1)$.*

PROOF. Let $f_i(x) = \operatorname{dist}(x, Z_i)$ and $f = f_1/(f_1 + f_2)$. □

In addition to the regularity properties, topological spaces may differ in the following properties related to covers. An open cover of a set is a collection of open sets the union of which contains this set.

6.1.4. Definition. (i) *A Hausdorff space X is called compact if every open cover of X contains a finite subcover. If this is true for countable covers, then X is called countably compact. A countable union of compact sets is called a σ-compact space.*

(ii) *A Hausdorff space X is called Lindelöf if every open cover of X contains an at most countable subcover.*

(iii) *A Hausdorff space X is called paracompact if in every open cover $\{U_\alpha\}$ of X one can inscribe a locally finite open cover $\{W_\beta\}$, i.e., every point has a neighborhood that meets only finitely many sets W_β. A space X is called countably paracompact if the indicated property is fulfilled for all at most countable open covers $\{U_\alpha\}$.*

(iv) *A Hausdorff space X is called sequentially compact if every infinite sequence in X has a convergent subsequence.*

Sets with compact closure (i.e., subsets of compact sets) are called relatively compact. In metrizable spaces (unlike general spaces), the compactness and countable compactness of a set K are equivalent and are also equivalent to the sequential compactness of K.

Note that sometimes in the definition of Lindelöf spaces one includes that the space must be regular. The properties to be Lindelöf or paracompact are not inherited by subsets. If in a space X every subset possesses one of the listed properties, then that property is called hereditary. For example, X is hereditary Lindelöf provided that every collection of open sets in X contains an at most countable subcollection with the same union. Among the listed classes of topological spaces the most important for applications in measure theory are compact and completely regular spaces. Also frequent are *locally compact* spaces, i.e., spaces in which every point has a neighborhood with compact closure.

6.1.5. Lemma. *Let K be a nonempty compact set in a completely regular space X and let U be an open set containing K. Then, there exists a continuous function $f\colon X \to [0,1]$ such that $f|_K = 1$ and $f|_{X\setminus U} = 0$.*

The proof is delegated to Exercise 6.10.21.

Let X be a completely regular space. Then there exists (and is unique) a compact space βX called the *Stone–Čech compactification* of the space X such that X is homeomorphically embedded into βX as a dense subset and every bounded continuous function on X extends to a continuous function on βX (see Engelking [**532**, §3.6]). A completely regular space is called *Čech complete* if it is a G_δ-set (i.e., a countable intersection of open sets) in βX. Polish spaces (see below) and locally compact spaces are Čech complete.

Let X_t be a family of nonempty topological spaces parameterized by indices t from some nonempty set T. The product $X = \prod_{t\in T} X_t$ of the spaces X_t has a natural topology (called the product topology) consisting of all possible unions of the products of the form $U_{t_1} \times \ldots U_{t_n} \times \prod_{t\neq t_i} X_t$, where U_{t_i} is an open set in X_{t_i}.

If $X_t = X$ for all $t \in T$, then the product of the spaces X_t is denoted by X^T. This space is naturally identified with the space of all mappings $x\colon T \to X$. Under this identification, the product topology becomes the topology of pointwise convergence. If $T = \mathbb{N}$, then the corresponding product is denoted by X^∞. An important example is the space $\mathbb{R}^\infty$ of all real sequences $x = (x_n)$. The countable product of metric spaces X_n with metrics ϱ_n is metrizable by the metric

$$\varrho(x,y) = \sum_{n=1}^{\infty} 2^{-n} \frac{\varrho_n(x_n,y_n)}{\varrho_n(x_n,y_n)+1}.$$

It is readily verified that if all X_n are complete separable metric spaces, then so is their product with the above metric. For example, $\mathbb{R}^\infty$ is a complete separable metric space.

One of the simplest examples of infinite products (but very important for measure theory) is the countable power $\mathbb{N}^\infty$ of the set of natural numbers, i.e., the set of all infinite sequences $\nu = (\nu_i)$ of natural numbers. Convergence in $\mathbb{N}^\infty$ is just coordinate-wise convergence. As above, we equip $\mathbb{N}^\infty$ with the metric

$$\varrho(\nu, \mu) = \sum_{j=1}^{\infty} 2^{-j} \frac{|n_j - m_j|}{|n_j - m_j| + 1}, \quad \nu = (n_j), \mu = (m_j). \tag{6.1.1}$$

6.1.6. Theorem. (R. Baire) *The space $\mathbb{N}^\infty$ with the product topology is homeomorphic to the space $\mathcal{R}$ of all irrational numbers in $(0,1)$ (or in $\mathbb{R}^1$) with its usual topology.*

PROOF. For every $\nu = (n_i) \in \mathbb{N}^\infty$, let $h(\nu) := \sum_{k=1}^{\infty} 2^{-n_1-\cdots-n_k}$. It is readily seen that h is a homeomorphism between $\mathbb{N}^\infty$ and the complement of the countable set M of binary rational numbers in $[0,1]$. It remains to observe that there is homeomorphism h_0 of $[0,1]$ such that $h_2(M) = \mathbb{Q} \cap [0,1]$; see Engelking [**532**, 4.3H, p. 279]. □

6.1.7. Corollary. *The space $\mathbb{N}^\infty$ contains a closed subspace that can be continuously mapped one-to-one onto $\mathbb{R}^1$.*

PROOF. The space $\mathbb{N}^\infty$ is homeomorphic to $\mathbb{N}^\infty \times \mathbb{N}$, and the closed subspace $\mathbb{N}^\infty \times \{1\}$ of $\mathbb{N}^\infty \times \mathbb{N}$ is homeomorphic to the space of irrational numbers. We add to $\mathbb{N}^\infty \times \{1\}$ the set of all points of the form $(n, 1, 1, \ldots) \times \{2\}$, which is closed, countable, and disjoint with $\mathbb{N}^\infty \times \{1\}$. This additional set can be continuously mapped one-to-one onto the space of rational numbers. □

Another useful example for measure theory is the countable power of the two-point set.

6.1.8. Example. The Cantor set C is homeomorphic to $\{0,1\}^\infty$.

A justification is left as Exercise 6.10.25. Uncountable products are nonmetrizable, excepting the case where at most countably many factors are singletons (see Exercise 6.10.23). The following important result is called Tychonoff's theorem; see [**532**, Theorem 3.2.4].

6.1.9. Theorem. *If nonempty spaces X_t are compact, then their product is compact as well.*

Now we introduce a class of spaces that is very important for measure theory.

6.1.10. Definition. *A topological space homeomorphic to a complete separable metric space is called Polish. The empty set is also included in the class of Polish spaces.*

6.1.11. Example. Every open or closed subset of a Polish space is Polish.

PROOF. We have to show that every set Y that is either open or closed in a complete separable metric space X can be equipped with a metric generating the original topology and making Y a complete space (clearly, it remains separable). In the case of a closed set the metric of X works, and if Y is open, then we take the metric

$$\varrho_0(x,y) = \varrho(x,y) + \frac{|\mathrm{dist}(x, X\backslash Y) - \mathrm{dist}(y, X\backslash Y)|}{|\mathrm{dist}(x, X\backslash Y) - \mathrm{dist}(y, X\backslash Y)| + 1}.$$

The verification of the fact that we obtain the required metric is left as a simple exercise. □

We recall that countable intersections of open sets are called G_δ-sets or sets of the type G_δ. Countable unions of closed sets are called F_σ-sets. The above example is a special case of a general result (see Engelking [**532**, Theorem 4.3.23, Theorem 4.3.24]), according to which any G_δ-set in a complete metric space is metrizable by a complete metric and, conversely, if a subspace of a metric space is metrizable by a complete metric, then this subspace is a G_δ-set. Polish spaces have the following characterizations (proofs can be found in Engelking [**532**, Theorem 4.2.10, Theorem 4.3.24, Corollary 4.3.25]).

6.1.12. Theorem. (i) *Polish spaces are precisely the spaces that are homeomorphic to closed subspaces in* $\mathbb{R}^\infty$.

(ii) *Every separable metric space* X *is homeomorphic to a subset of* $[0,1]^\infty$, *and if* X *is complete, then this subset is a* G_δ*-set.*

6.1.13. Theorem. *Every nonempty complete separable metric space is the image of* $\mathbb{N}^\infty$ *under a continuous mapping.*

PROOF. Let us equip $\mathbb{N}^\infty$ with the metric (6.1.1). We represent the given space X in the form $X = \bigcup_{j=1}^\infty E(j)$, where the sets $E(j)$ are closed (not necessarily disjoint) of diameter less than 2^{-3}. By induction, for every k, we find a closed set $E(n_1,\ldots,n_k)$ of diameter less than 2^{-k-2} with

$$E(n_1,\ldots,n_k) = \bigcup_{j=1}^\infty E(n_1,\ldots,n_k,j).$$

For every $\nu = (n_i) \in \mathbb{N}^\infty$, the closed sets $E(n_1,\ldots,n_k)$ are decreasing and have diameters less than 2^{-k-2}. Hence they shrink to a single point denoted by $f(\nu)$. Note that $f(\mathbb{N}^\infty) = X$. Indeed, every point x belongs to some set $E(n_1)$, then to $E(n_1,n_2)$ and so on, which yields an element ν such that $f(\nu) = x$. In addition, f is locally Lipschitzian. Indeed, let ϱ be the metric in X. If $\varrho(\nu,\mu) < 1/4$, then there exists k with $2^{-k-2} \le \varrho(\nu,\mu) < 2^{-k-1}$. Then $\nu_i = \mu_i$ if $i \le k$. Hence $f(\mu)$ and $f(\nu)$ belong to $E(n_1,\ldots,n_k)$, whence we obtain $\varrho\bigl(f(\mu), f(\nu)\bigr) < 2^{-k-2} \le \varrho(\mu,\nu)$. Thus, f is continuous. □

6.1.14. Corollary. *Every nonempty Polish space is the image of* $\mathbb{N}^\infty$ *under a continuous mapping.*

Certainly, such a mapping may not be injective (e.g., in the case of a finite space). But every Polish space without isolated points can be represented as the image of $\mathbb{N}^\infty$ under an injective continuous mapping (see Rogers, Jayne [**1589**, §2.4]). For injective mappings we have the following.

6.1.15. Theorem. *For any Polish space X, one can find a closed set $Z \subset \mathbb{N}^\infty$ and a one-to-one continuous mapping f of the set Z onto X.*

PROOF. By Theorem 6.1.12, we may assume that X is a closed subspace in $\mathbb{R}^\infty$. By Corollary 6.1.7, there exists a closed set $E \subset \mathbb{N}^\infty$ that can be mapped continuously and one-to-one onto $\mathbb{R}^1$. Then E^∞ is closed in the countable power of $\mathbb{N}^\infty$ and admits a continuous one-to-one mapping onto $\mathbb{R}^\infty$. Since the countable power of $\mathbb{N}^\infty$ is homeomorphic to $\mathbb{N}^\infty$ and the preimage of a closed set under a continuous mapping is closed, we obtain the required representation. □

A mapping mentioned in the above theorem may not be a homeomorphism (i.e., the inverse mapping may be discontinuous). For example, the set $\mathcal{R}$ of irrational numbers (we recall that $\mathcal{R}$ is homeomorphic to $\mathbb{N}^\infty$) contains a closed set that can be mapped continuously and one-to-one onto $[0,1]$, but such a mapping cannot be a homeomorphism because $\mathcal{R}$ contains no intervals.

By a modification of the proof of Theorem 6.1.13 one establishes the following lemma (see Kuratowski [**1082**, §36] and Exercise 6.10.33).

6.1.16. Lemma. *Every nonempty complete metric space without isolated points contains a subset homeomorphic to $\mathbb{N}^\infty$.*

Nonempty closed sets without isolated points are called perfect.

6.1.17. Proposition. *Any two bounded perfect nowhere dense sets on the real line are homeomorphic. In particular, every set of this type is homeomorphic to the Cantor set and has cardinality of the continuum.*

PROOF. Let $E \subset [0,1]$ be a set of this type and let $0, 1 \in E$. We construct a homeomorphism h of $[0,1]$ that maps the Cantor set C onto E. To this end, we enumerate the countable family $\mathcal{U}$ of disjoint open intervals complementary to E in $[0,1]$ as follows. Let $U_{1,1} \in \mathcal{U}$ be an interval of the maximal length. Next we pick an interval $U_{2,1} \in \mathcal{U}$ of the maximal length on the left from $U_{1,1}$ and an interval $U_{2,2} \in \mathcal{U}$ of the maximal length on the right from $U_{1,1}$. We proceed by induction and, for every n, obtain 2^{n-1} open intervals $U_{n,k}$ that have the same mutual disposition as the intervals $J_{n,k}$ which appear in the construction of the Cantor set. Clearly, this process exhausts all intervals in $\mathcal{U}$. Let h be an affine homeomorphism between $J_{n,k}$ and $U_{n,k}$ for all n, k, so h is an increasing function that maps $[0,1]\backslash C$ homeomorphically onto $[0,1]\backslash E$. It is readily seen that h extends uniquely to a homeomorphism of $[0,1]$ by the formula $h(t) = \inf\{h(u)\colon\ u \notin C, u > t\}$. □

In measure theory, the following representation of metrizable compacts, obtained by P.S. Alexandroff, is useful. A simple proof is found in many books; see Engelking [**532**, 4.5.9, p. 291].

6.1.18. Proposition. *Any nonempty metric compact K is a continuous image of some compact set K_0 in $[0,1]$. Moreover, one can take for K_0 the Cantor set C.*

Let us give examples of more exotic topological spaces useful for constructing various counter-examples in measure theory.

6.1.19. Example. The Sorgenfrey line Z is defined as the real line with the topology whose base consists of all intervals $[x, r)$, where x is a real number, r is a rational number and $x < r$. The Sorgenfrey interval $[0, 1)$ is equipped with the same topology. Similarly, the Sorgenfrey plane Z^2 is the plane with the topology generated by the rectangles $[a, b)\times[c, d)$. Usual open sets on the real line (or in the plane) are open in the Sorgenfrey topology, since every interval (a, b) is the union of the sets $[a + 1/n, b)$.

The Sorgenfrey line has the following properties, see Arkhangel'skiĭ, Ponomarev [**68**], Engelking [**532**], Steen, Seebach [**1774**]:

(1) the space Z is not metrizable, but it is Lindelöf, paracompact and perfectly normal, and every point has a countable base of neighborhoods;

(2) every compact subset of Z is at most countable.

The set D of all points of the form $(x, -x)$ in the Sorgenfrey plane is closed and is discrete in the induced topology, i.e., every point is open in the induced topology. This follows by the equality $(x, -x) = D \cap [x, x + 1)\times[-x, -x + 1)$.

6.1.20. Example. Let $X = C_0 \bigcup C_1 \subset \mathbb{R}^2$, where

$$C_0 = \{(x, 0)\colon\ 0 < x \le 1\} \quad \text{and} \quad C_1 = \{(x, 1)\colon\ 0 \le x < 1\}.$$

Let us equip X with the topology generated by the base consisting of all sets of one of the following two types:

$$\bigl\{(x, i) \in X\colon\ x_0 - 1/k < x < x_0,\ i = 0, 1\bigr\} \cup \{(x_0, 0)\},$$

where $0 < x_0 \le 1$, $k \in \mathbb{N}$, and

$$\bigl\{(x, i) \in X\colon\ x_0 < x < x_0 + 1/k,\ i = 0, 1\bigr\} \cup \{(x_0, 1)\},$$

where $0 \le x_0 < 1$, $k \in \mathbb{N}$. The space X is called "two arrows of P.S. Alexandroff" (or "two arrows", "double arrow") and has the following properties:

(i) X is a compact space;

(ii) X is perfectly normal and hereditary Lindelöf;

(iii) X is a non-metrizable separable space, in which every point has a countable base of neighborhoods. Every metrizable subset of X is at most countable;

(iv) the natural projection of X onto $[0, 1]$ (with the usual topology) is continuous.

See Arkhangel'skiĭ, Ponomarev [**68**, p. 146] or Engelking [**532**, 3.10C] for a proof and Exercise 6.10.87 for an alternative description of this topology.

6.1.21. Example. Let Ω be an ordinal number. The set of all ordinals α with $\alpha \le \Omega$ is denoted by $[0, \Omega]$. It is equipped with the *order topology*, the base of which consists of all sets of the form $\{x < \alpha\}$, $\{\alpha < x < \beta\}$, $\{x > \alpha\}$, where $\alpha, \beta \le \Omega$. The space $[0, \Omega)$ with the deleted point Ω is equipped with the induced topology. The space $[0, \Omega]$ is compact. Indeed, given its open cover $\{U_t\}_{t\in T}$, we consider the set M of all $x \le \Omega$ such that the closed interval $[0, x]$ is not covered by finitely many elements of the given cover. Since $[0, \Omega]$ is well-ordered, M contains the smallest element x_0. There exists t_0 with $x_0 \in U_{t_0}$. It is easy to see that there exists an element $y \in [0, x_0) \cap U_{t_0}$ (if x_0 is the minimal element in U_{t_0}, then x_0 has an immediate predecessor, which leads to a contradiction). Then $y \notin M$ and there exist $t_1, \ldots, t_n \in T$ such that $[0, y] \subset \bigcup_{i=1}^n U_{t_i}$. Hence $[0, x_0] \subset \bigcup_{i=0}^n U_{t_i}$, i.e., $x_0 \notin M$.

6.2. Borel sets

One of the most frequently used σ-algebras on a topological space X is the *Borel σ-algebra* generated by all open sets; it is denoted by the symbol $\mathcal{B}(X)$. It is clear that $\mathcal{B}(X)$ is generated by all closed sets, too. The sets in $\mathcal{B}(X)$ are called the *Borel sets* in the space X. The property of a set to be Borel depends on the space in which it is considered. For example, one always has $X \in \mathcal{B}(X)$.

Borel sets owe the name to the classical works of E. Borel [**230**], [**234**].

6.2.1. Definition. *Let X and Y be topological spaces. A mapping $f\colon X \to Y$ is called Borel (or Borel measurable) if $f^{-1}(B) \in \mathcal{B}(X)$ for all sets $B \in \mathcal{B}(Y)$.*

6.2.2. Lemma. *Every continuous mapping between topological spaces is Borel measurable.*

Proof. Let X, Y be topological spaces and let $f\colon X \to Y$ be a continuous mapping. Denote by $\mathcal{E}$ the class of all sets $B \in \mathcal{B}(Y)$ such that $f^{-1}(B) \in \mathcal{B}(X)$. Obviously, the class $\mathcal{E}$ is a σ-algebra and by the continuity of f it contains all open sets (we recall that the preimage of any open set under a continuous mapping is open). Therefore, $\mathcal{E} = \mathcal{B}(Y)$. $\square$

6.2.3. Lemma. *Let $(X, \mathcal{A})$ be a measurable space, let E be a separable metric space, and let $f\colon X \to E$ be measurable, i.e., $f^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}(E)$. Then, there exists a sequence of measurable mappings f_n with an at most countable range uniformly convergent to f.*

Proof. For every n we cover E by a finite or countable collection of balls of diameter less than $1/n$. From this collection we construct a cover of E by disjoint Borel sets $B_{n,k}$, $k \in \mathbb{N}$, of diameter less than $1/n$. Next we choose in every $B_{n,k}$ a point c_k and let $f_n(x) = c_k$ if $x \in f^{-1}(B_{n,k})$. Then the distance between $f_n(x)$ and $f(x)$ does not exceed $1/n$ for all x. $\square$

6.2.4. Lemma. *Let X be a topological space and let Y be a subset of X with the induced topology. Then $\mathcal{B}(Y) = \{B \cap Y\colon\ B \in \mathcal{B}(X)\}$.*

In particular, for all $Y \in \mathcal{B}(X)$ we have $\mathcal{B}(Y) = \{B \in \mathcal{B}(X)\colon\ B \subset Y\}$.

Proof. Let $\mathcal{E} := \{E \subset Y\colon\ E = B \cap Y,\ B \in \mathcal{B}(X)\}$. It is easy to see that $\mathcal{E}$ is a σ-algebra. By the definition of the induced topology, all open sets in the space Y belong to $\mathcal{E}$ because they are intersections of Y with open sets in X. Hence $\mathcal{B}(Y) \subset \mathcal{E}$. On the other hand, the class $\mathcal{E}_0$ of all sets $B \in \mathcal{B}(X)$ such that $B \cap Y \in \mathcal{B}(Y)$, is a σ-algebra too and contains all open sets in X. So $\mathcal{B}(X) \subset \mathcal{E}_0$, which completes the proof. The last claim is obvious. □

Let us consider certain elementary properties of Borel mappings.

6.2.5. Lemma. *Let $(\Omega, \mathcal{A})$ be a measurable space and let T be a metric space (or, more generally, a perfectly normal space, i.e., a space in which every closed set is the set of zeros of a continuous function). A mapping $f\colon \Omega \to T$ is measurable with respect to the σ-algebras $\mathcal{A}$ and $\mathcal{B}(T)$ precisely when for every continuous real function ψ on T, the function $\psi \circ f$ is measurable with respect to $\mathcal{A}$.*

Proof. The necessity of the above condition is obvious, since $\psi^{-1}(U)$ is open in T for any open $U \subset \mathbb{R}^1$. For the proof of the converse we verify that $f^{-1}(Z) \in \mathcal{A}$ for every closed set $Z \subset T$. We observe that Z has the form $Z = \psi^{-1}(0)$ for some continuous function ψ (if T is perfectly normal, then this is true by definition, in the case of a metric space one can take $\psi(x) = \operatorname{dist}(x, Z)$). Now we obtain $f^{-1}(Z) = (\psi \circ f)^{-1}(0) \in \mathcal{A}$. □

6.2.6. Corollary. *Suppose that in the situation of Lemma* 6.2.5 *the mapping $f\colon \Omega \to T$ is the pointwise limit of a sequence of measurable mappings $f_n\colon (\Omega, \mathcal{A}) \to \bigl(T, \mathcal{B}(T)\bigr)$. Then f is measurable with respect to $\mathcal{A}$ and $\mathcal{B}(T)$.*

6.2.7. Corollary. *The statement of the previous corollary remains valid if Ω is a topological space with the Borel σ-algebra and the mappings f_n are continuous.*

The last corollary may fail for arbitrary completely regular spaces T. Let us consider the following example of R.M. Dudley.

6.2.8. Example. Let T be the space of all functions f from $[0,1]$ to $[0,1]$ equipped with the topology of pointwise convergence. According to Tychonoff's theorem, T is compact. Let us take for Ω the interval $[0,1]$ with the Borel σ-algebra. Let $f_n\colon \Omega \to T$ be defined by the formula

$$f_n(\omega)(s) = \max(1 - n \mid \omega - s \mid, 0),\ \omega \in \Omega, s \in [0,1].$$

The mappings f_n converge pointwise to the mapping $f\colon \omega \mapsto I_{\{\omega\}}$, i.e., $f(\omega)(s) = 1$ if $s = \omega$ and $f(\omega)(s) = 0$ if $s \neq \omega$. Each mapping f_n is continuous, hence measurable if T is equipped with the Borel σ-algebra, but f is not measurable. Indeed, the set $U_C = \bigcup_{\omega \in C} \{x \in T\colon\ x(\omega) > 0\}$ is open in T for every subset $C \subset \Omega$, and $f^{-1}(U_C) = C$. Let C be a non-Borel set. Then the preimage of U_C is not measurable.

6.2.9. Proposition. *Let X be a metric space and let $\mathcal{E}$ be some class of subsets of X containing all open sets and closed with respect to countable* *unions of pairwise disjoint sets and countable intersections. Then $\mathcal{E}$ contains all Borel sets.*

PROOF. Follows by Theorem 1.12.2. $\square$

6.2.10. Definition. (i) *An isomorphism of two measurable spaces $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ is a one-to-one mapping $j\colon X \to Y$ such that $j(\mathcal{A}) = \mathcal{B}$ and $j^{-1}(\mathcal{B}) = \mathcal{A}$.*

(ii) *A measurable space $(S, \mathcal{B})$ is called standard if it is isomorphic to the space $\bigl(M, \mathcal{B}(M)\bigr)$ for some Borel set M in a Polish space.*

Sometimes standard measurable spaces are called standard Borel spaces. We shall see below that there are only two non-isomorphic classes of standard measurable spaces of infinite cardinality: countable and uncountable.

Let us prove the following interesting result of Kuratowski on extensions of isomorphisms (see also Kuratowski [**1082**, §35, VII]).

6.2.11. Theorem. *Let X and Y be Polish spaces, $A \subset X$, $B \subset Y$, and let $f\colon A \to B$ be a Borel isomorphism, i.e., a one-to-one Borel mapping such that f^{-1} is Borel measurable provided that A and B are equipped with the induced Borel σ-algebras. Then, one can find two sets $A^* \in \mathcal{B}(X)$ and $B^* \in \mathcal{B}(Y)$ and a Borel isomorphism $f^*\colon A^* \to B^*$ such that $A \subset A^*$, $B \subset B^*$ and $f^*|_A = f$.*

PROOF. Let $g := f^{-1}\colon B \to A$. Clearly, one can find Borel mappings $f^*\colon X \to Y$ and $g^*\colon Y \to X$ such that $f^*|_A = f$ and $g^*|_B = g$. Let us set $A^* := \bigl\{x \in X\colon\ g^*\bigl(f^*(x)\bigr) = x\bigr\}$, $B^* := \bigl\{y \in Y\colon\ f^*\bigl(g^*(y)\bigr) = y\bigr\}$. It is readily seen that A^* and B^* are Borel sets and f^* is a Borel isomorphism between them. $\square$

6.3. Baire sets

Another important σ-algebra on a topological space X is generated by all sets of the form

$$\{x \in X\colon\ f(x) > 0\},$$

where f is a continuous function on X. This σ-algebra is called the *Baire σ-algebra* and is denoted by $\mathcal{B}a(X)$. It is clear that this is the smallest σ-algebra with respect to which all continuous functions on X are measurable. The same σ-algebra is generated by the class of all bounded continuous functions. The sets in $\mathcal{B}a(X)$ are called the *Baire sets* in the space X. Baire sets owe the name to the classical works of R. Baire [**93**], [**94**] on the theory of functions.

The sets of the form $\{x \in X\colon\ f(x) > 0\}$, where $f \in C(X)$, are called *functionally open* and their complements are called *functionally closed*.

In a metric space, any closed set is the set of zeros of a continuous function. Hence the Borel and Baire σ-algebras of a metric space coincide. Below we discuss other cases of coincidence and give examples of non-coincidence. The following lemma is obvious from the fact that every closed set on the real line has the form $f^{-1}(0)$, $f \in C(\mathbb{R}^1)$.

6.3.1. Lemma. *A set U is functionally open precisely when it has the form $U = \varphi^{-1}(W)$, where $\varphi \in C(X)$ and $W \subset \mathbb{R}^1$ is open.*

A set Z is functionally closed precisely when it has the form $Z = \psi^{-1}(0)$, where $\psi \in C(X)$.

6.3.2. Lemma. *Let Z_1 and Z_2 be disjoint functionally closed sets in a topological space X. Then, there exists a function $f \in C_b(X)$ with values in $[0,1]$ such that $Z_1 = f^{-1}(0)$, $Z_2 = f^{-1}(1)$.*

PROOF. The sets Z_i have the form $Z_i = \psi_i^{-1}(0)$, where $\psi_i \in C_b(X)$ and $0 \le \psi_i \le 1$. One can take $f = \psi_1/(\psi_1 + \psi_2)$. $\square$

6.3.3. Lemma. *Every Baire set is determined by some countable family of functions, i.e., has the form*

$$\bigl\{x\colon\ \bigl(f_1(x), f_2(x), \ldots, f_n(x), \ldots\bigr) \in B\bigr\},\ \ f_i \in C(X),\ \ B \in \mathcal{B}(\mathbb{R}^\infty). \qquad (6.3.1)$$

Moreover, every set of this form is Baire, and we can take $f_i \in C_b(X)$.

PROOF. We prove first that every set of the form (6.3.1) is Baire. This is true if B is closed, since it has the form $B = \psi^{-1}(0)$ for some continuous function ψ on $\mathbb{R}^\infty$ and the function $x \mapsto \psi\bigl(\bigl(f_n(x)\bigr)_{n=1}^\infty\bigr)$ is continuous. It is easily verified that for any fixed sequence $\{f_n\}$, the class $\mathcal{B}_0$ of all sets $B \in \mathcal{B}(\mathbb{R}^\infty)$ such that

$$\bigl\{x\colon\ \bigl(f_1(x), f_2(x), \ldots, f_n(x), \ldots\bigr) \in B\bigr\} \in \mathcal{B}a(X)$$

is a σ-algebra. Hence it contains $\mathcal{B}(\mathbb{R}^\infty)$ and thus coincides with $\mathcal{B}(\mathbb{R}^\infty)$. On the other hand, the class $\mathcal{E}$ of all Baire sets E representable in the form (6.3.1) with $f_i \in C_b(X)$, contains all sets of the form $\{f > 0\}$, $f \in C(X)$. In addition, this class is a σ-algebra. Indeed, the complement of any set $E \in \mathcal{E}$ has the form (6.3.1) with the same f_i and the set $\mathbb{R}^\infty \backslash B$ in place of B. If $E_j \in \mathcal{E}$ are represented by means of the sets $B_j \in \mathcal{B}(\mathbb{R}^\infty)$ and functions $f_{j,n}$, then $E = \bigcap_{j=1}^\infty E_j$ can be written in the form (6.3.1) as well. To this end, we write the space $\mathbb{R}^\infty$ as its countable power and take $B = \prod_{j=1}^\infty B_j$. $\square$

The following result follows immediately from the definitions. Nevertheless, it is useful in applications because perfectly normal spaces constitute a sufficiently large class. Some examples are given below.

6.3.4. Proposition. *Let X be a perfectly normal space. Then we have $\mathcal{B}(X) = \mathcal{B}a(X)$.*

6.3.5. Corollary. *The equality $\mathcal{B}(X) = \mathcal{B}a(X)$ is true in any of the following cases*:

(i) *X is a metric space,*

(ii) *X is a regular space such that every family of its open subsets contains a countable subfamily with the same union* (*i.e., X is hereditary Lindelöf*).

PROOF. Both conditions imply that X is perfectly normal (see Section 6.1 or Engelking [**532**, §3.8]). $\square$

The following lemma shows that if in Lemma 6.2.5 one deals with Baire sets in place of Borel sets, then no restrictions on the space are needed.

6.3.6. Lemma. *Let $(\Omega, \mathcal{A})$ be a measurable space and let T be a topological space. A mapping $f\colon \Omega \to T$ is measurable with respect to the σ-algebras $\mathcal{A}$ and $\mathcal{B}a(T)$ precisely when for every continuous real function ψ on T, the function $\psi \circ f$ is measurable with respect to $\mathcal{A}$.*

PROOF. The necessity of this condition is obvious, and its sufficiency is verified in the same manner as in Lemma 6.2.5: the class $\mathcal{E}$ of all sets $B \in \mathcal{B}a(T)$ with $f^{-1}(B) \in \mathcal{A}$ is a σ-algebra and contains all sets $\psi^{-1}(0)$, where $\psi \in C(T)$. □

6.3.7. Corollary. *Let $(\Omega, \mathcal{A})$ be a measurable space, let T be a topological space, and let a mapping $f\colon \Omega \to T$ be the pointwise limit of a sequence of measurable mappings $f_n\colon (\Omega, \mathcal{A}) \to \bigl(T, \mathcal{B}a(T)\bigr)$. Then f is measurable with respect to $\mathcal{A}$ and $\mathcal{B}a(T)$.*

6.4. Products of topological spaces

Let T be a nonempty index set and let X_t, $t \in T$, be a family of nonempty spaces equipped with σ-algebras $\mathcal{A}_t$. We recall that the product of the family $\{X_t\}_{t\in T}$ is the set of all collections of the form $x = \{x_t,\, t \in T\}$, where $x_t \in X_t$ for every $t \in T$. This product is denoted by $\prod_{t\in T} X_t$. In Chapter 3, we have already discussed the σ-algebra $\mathcal{A} = \bigotimes_{t\in T} \mathcal{A}_t$ generated by all finite products of sets from $\mathcal{A}_t$. This section is concerned with the situation where all X_t are topological spaces and $\mathcal{A}_t$ are the Borel or Baire σ-algebras. The space $X = \prod_{t\in T} X_t$ is equipped with the product topology, i.e., open sets are unions of basic open sets of the form $U_{t_1,\ldots,t_n} = \{x \in X\colon\, x_{t_i} \in U_{t_i}, i = 1, \ldots, n\}$, where U_{t_i} is an open set in X_{t_i}. The principal question concerns the relations between $\bigotimes_{t\in T} \mathcal{B}(X_t)$ and $\mathcal{B}(X)$ and between $\bigotimes_{t\in T} \mathcal{B}a(X_t)$ and $\mathcal{B}a(X)$.

6.4.1. Lemma. *Let B_n be Borel sets in spaces X_n, where $n \in \mathbb{N}$. Then $B = \prod_{n=1}^\infty B_n$ is a Borel set in $X = \prod_{n=1}^\infty X_n$ with the product topology. In addition, one has $\bigotimes_{n=1}^\infty \mathcal{B}(X_n) \subset \mathcal{B}(X)$.*

PROOF. Since $B = \bigcap_{k=1}^\infty \bigl(B_k \times \prod_{n\neq k} X_n\bigr)$, it suffices to verify that for any $B \in \mathcal{B}(X_1)$ one has $B \times \prod_{n=2}^\infty X_n \in \mathcal{B}(X)$. This is true for open B. Since the class $\mathcal{E}$ of all $B \in \mathcal{B}(X_1)$ such that $B \times \prod_{n=2}^\infty X_n \in \mathcal{B}(X)$ is a σ-algebra, this class coincides with $\mathcal{B}(X_1)$. The second claim follows from the first one. □

6.4.2. Lemma. (i) *Let X, Y be Hausdorff spaces and let Y have a countable base* (*e.g., let Y be a separable metric space*). *Then we have the equality $\mathcal{B}(X \times Y) = \mathcal{B}(X) \otimes \mathcal{B}(Y)$.*

(ii) *Let X_n, where $n \in \mathbb{N}$, be nonempty Hausdorff spaces such that $\prod_{n=1}^\infty X_n$ is hereditary Lindelöf* (*e.g., let all X_n have countable bases*). *Then we have the equality $\mathcal{B}\Bigl(\prod_{n=1}^\infty X_n\Bigr) = \bigotimes_{n=1}^\infty \mathcal{B}(X_n)$.*

(iii) *If every X_n is compact, then $\mathcal{B}a\bigl(\prod_{n=1}^\infty X_n\bigr) = \bigotimes_{n=1}^\infty \mathcal{B}a(X_n)$.*

Proof. (i) According to the previous lemma, it suffices to show that every open set U in $X\times Y$ belongs to $\mathcal{B}(X)\otimes\mathcal{B}(Y)$. Let $\{V_n\}$ be a countable base of Y. Then U can be represented as the union of sets $U_\alpha\times V_n$, where the sets U_α are open in X. For fixed n, let W_n be the union of all sets U_α with $U_\alpha\times V_n\subset U$. Then $U=\bigcup_{n=1}^\infty (W_n\times V_n)\in\mathcal{B}(X)\otimes\mathcal{B}(Y)$.

(ii) By the Lindelöf property, every open set in $\prod_{n=1}^\infty X_n$ can be represented as a countable union of finite products of open sets in the spaces X_n, since it is a certain union of elements of the standard base.

(iii) Follows by the Weierstrass theorem, according to which the set of finite sums of products of functions from $C(X_n)$ is dense in $C\bigl(\prod_{n=1}^\infty X_n\bigr)$. □

This lemma does not extend to arbitrary spaces (even metric ones).

6.4.3. Example. Let X be a Hausdorff space of cardinality greater than that of the continuum. Then $\mathcal{B}(X\times X)\neq\mathcal{B}(X)\otimes\mathcal{B}(X)$.

Proof. We show that the diagonal

$$\Delta := \bigl\{(x,x)\colon\ x\in X\bigr\},$$

which is closed in $X\times X$, does not belong to $\mathcal{B}(X)\otimes\mathcal{B}(X)$. To this end, let $\mathcal{E}$ denote the class of all sets $E\subset X\times X$ such that E and its complement are representable as unions of the continuum (or fewer) of rectangles $A\times B$, $A,B\subset X$. By definition, $\mathcal{E}$ contains all rectangles. In addition, $\mathcal{E}$ is a σ-algebra. Indeed, the class $\mathcal{E}$ is closed with respect to complementation. It admits countable unions. Indeed, if $E_n\in\mathcal{E}$, then the complement to $\bigcup_{n=1}^\infty E_n$ can be written in the form of a union of the continuum of rectangles. To see this, we observe that if $(X\times X)\backslash E_n=\bigcup_\alpha E_{n,\alpha}$, where $E_{n,\alpha}$ are rectangles and α belongs to some set of indices A of cardinality of the continuum, then the complement to $\bigcup_{n=1}^\infty E_n$ is $\bigcap_{n=1}^\infty\bigcup_\alpha E_{n,\alpha}=\bigcup_{(\alpha_n)\in A^\infty} D_{(\alpha_n)}$, where $D_{(\alpha_n)}=\bigcap_{n=1}^\infty E_{n,\alpha_n}$ are rectangles, and the set A^∞ is of cardinality of the continuum or less. Therefore, $\mathcal{E}$ contains the σ-algebra generated by rectangles. It is clear that Δ does not belong to $\mathcal{E}$. □

We recall that the graph of a mapping $f\colon\ X\to Y$ is the subset of $X\times Y$ defined as $\Gamma_f := \bigl\{\bigl(x,f(x)\bigr)\colon\ x\in X\bigr\}$.

6.4.4. Theorem. *Let $(X,\mathcal{A})$, $(Y,\mathcal{B})$ and $(Z,\mathcal{E})$ be measurable spaces and let $f\colon\ (X,\mathcal{A})\to(Z,\mathcal{E})$ and $g\colon\ (Y,\mathcal{B})\to(Z,\mathcal{E})$ be measurable mappings. Suppose that*

$$\Delta_Z := \{(z,z)\colon\ z\in Z\}\in\mathcal{E}\otimes\mathcal{E}.$$

Then

$$\bigl\{(x,y)\in X\times Y\colon\ f(x)=g(y)\bigr\}\in\mathcal{A}\otimes\mathcal{B}.$$

In particular, the graph of the mapping f belongs to $\mathcal{A}\otimes\mathcal{E}$.

Proof. The mapping $(f,g)\colon\ X\times Y\to Z\times Z$ is measurable with respect to the pair of σ-algebras $\mathcal{A}\otimes\mathcal{B}$ and $\mathcal{E}\otimes\mathcal{E}$. By hypothesis, $\mathcal{A}\otimes\mathcal{B}$ contains the preimage of Δ_Z under this mapping, which yields the first claim. The second claim follows by the first one if we set $(Y,\mathcal{B})=(Z,\mathcal{E})$ and $g(y)=y$. □

6.4.5. Corollary. *Let X and Y be Hausdorff spaces such that*

$$\Delta_Y := \bigl\{(y,y)\colon\ y\in Y\bigr\} \in \mathcal{B}(Y)\otimes\mathcal{B}(Y).$$

Then, the graph of every Borel mapping $f\colon\ X\to Y$ belongs to $\mathcal{B}(X)\otimes\mathcal{B}(Y)$. In particular, this is the case if $Y\times Y$ is hereditary Lindelöf.

PROOF. The first claim follows by the above theorem. The second one is seen from the fact that the complement to the diagonal of Y^2 is open and can be written as a union of open rectangles $U\times V$, so it remains to choose among these rectangles a finite or countable collection with the same union, which yields that $\Delta_Y\in\mathcal{B}(Y)\otimes\mathcal{B}(Y)$. □

6.4.6. Lemma. *Suppose that $(X,\mathcal{B})$ is a measurable space and a function $f\colon\ X\times\mathbb{R}^1\to\mathbb{R}^1$ satisfies the following conditions: for every fixed $t\in\mathbb{R}^1$, the function $x\mapsto f(x,t)$ is $\mathcal{B}$-measurable, and for every fixed $x\in X$, the function $t\mapsto f(x,t)$ is right-continuous. Then, the function f is measurable with respect to $\mathcal{B}\otimes\mathcal{B}(\mathbb{R}^1)$. The same is true in the case of the left continuity. Moreover, f may be a mapping with values in a separable metric space.*

PROOF. We may assume that $0\le f\le 1$. For every natural n, we partition the interval $[0,1]$ into 2^n equal intervals by the points $k2^{-n}$. Let

$$f_n(x,t) = f\bigl(x, m+(k+1)2^{-n}\bigr) \quad \text{if } t\in \bigl[m+k2^{-n}, m+(k+1)2^{-n}\bigr),$$

where $m\in\mathbb{Z}$, $k=0,\ldots,2^n-1$. Note that $\lim\limits_{n\to\infty} f_n(x,t)=f(x,t)$ for all (x,t). Indeed, given $\varepsilon>0$, by hypothesis, there exists $\delta>0$ such that for all $s\in[t,t+\delta)$ one has $|f(x,t)-f(x,s)|<\varepsilon$. Let $2^{-n}<\delta$. Then we can find k such that $k2^{-n}\le t<(k+1)2^{-n}<t+\delta$. Hence $|f(x,t)-f_n(x,t)|<\varepsilon$. It remains to observe that the functions f_n are measurable with respect to $\mathcal{B}\otimes\mathcal{B}(\mathbb{R}^1)$ by the measurability of f in x. In the case of the left continuity the reasoning is similar. With an obvious modification the proof remains valid for mappings to separable metric spaces (see Corollary 6.2.6). □

It is worth noting that a function of two variables that is Borel in every variable separately may not be Borel in two variables (see Exercise 6.10.43).

Some additional information is given in §6.10(i).

6.5. Countably generated σ-algebras

We say that a family $\mathcal{S}$ of subsets of a space X separates the points in X if for every two distinct points x and y, there is a set $S\in\mathcal{S}$ such that either $x\in S$ and $y\notin S$ or $y\in S$ and $x\notin S$. A family $\mathcal{F}$ of functions on X is said to separate the points of X if for every two distinct points x and y, there is $f\in\mathcal{F}$ such that $f(x)\ne f(y)$.

6.5.1. Definition. *Let $\mathcal{E}$ be a σ-algebra of subsets of a space X.*

(i) *$\mathcal{E}$ is called countably generated or separable if it is generated by an at most countable family of sets E_n, i.e., $\mathcal{E}=\sigma(\{E_n\})$.*

(ii) *$\mathcal{E}$ is called countably separated if there exists an at most countable collection of sets $E_n\in\mathcal{E}$ separating the points.*

6.5.2. Example. The Borel σ-algebra of a separable metric space is separable and countably separated. Indeed, a countable base of open sets generates the Borel σ-algebra and separates the points.

It is clear that the σ-algebra $\sigma(\{f_n\})$ generated by a countable family of real functions f_n on a space X is countably generated because it is generated by the sets $\{f_n < r_k\}$ where $\{r_k\}$ are all rational numbers.

6.5.3. Lemma. *Let Γ be a family of functions on a space X. The generated σ-algebra $\sigma(\Gamma)$ separates the points in X precisely when Γ separates the points in X.*

PROOF. If Γ separates the points in X, then the sets from $\sigma(\Gamma)$ of the form $f^{-1}(a,b)$, $f \in \Gamma$, a, $b \in \mathbb{R}^1$, separate them too. Suppose now that $\sigma(\Gamma)$ separates the points in X, but Γ does not, i.e., there exist two distinct points x and y with $f(x) = f(y)$ for all $f \in \Gamma$. Let us consider the class $\mathcal{E}$ of all sets $E \subset X$ such that either $\{x,y\} \subset E$ or $\{x,y\} \subset X\backslash E$. It is readily verified that $\mathcal{E}$ is a σ-algebra. By our assumption $\mathcal{E}$ contains all sets $\{f < c\}$, $f \in \Gamma$, $c \in \mathbb{R}^1$, hence $\mathcal{E}$ contains the generated σ-algebra. This leads to a contradiction, since $\sigma(\Gamma)$ separates the points x and y. $\square$

6.5.4. Proposition. *Let $\mathcal{F}$ be a family of continuous real functions separating the points of a topological space X such that $X \times X$ is hereditary Lindelöf. Then $\mathcal{F}$ contains a finite or countable subfamily separating the points in X. In particular, this is true if X is a separable metric space.*

PROOF. For every $f \in \mathcal{F}$, let $U(f) = \{(x,y) \in X \times X\colon\ f(x) \neq f(y)\}$. Denote by C the complement of the diagonal in the space $X \times X$. The sets $U(f)$ form an open cover of C. By our assumption on $X \times X$, one can find a finite or countable subfamily of sets $U(f_n)$ covering C. It is clear that the family of functions f_n separates the points in X. In fact, we only need that C be Lindelöf. $\square$

6.5.5. Theorem. *Let $(E,\mathcal{E})$ be a measurable space. Then $\mathcal{E}$ is countably generated if and only if there exists an $\mathcal{E}$-measurable function $f\colon\ E \to [0,1]$ such that $\mathcal{E} = \{f^{-1}(B)\colon\ B \in \mathcal{B}([0,1])\}$.*

PROOF. For any function $f\colon\ E \to [0,1]$, the collection of sets $f^{-1}(B)$, where $B \in \mathcal{B}([0,1])$, is a countably generated σ-algebra. For a countable collection of generating sets one can take $f^{-1}([0,r_n])$, where $\{r_n\}$ are all rational numbers in $[0,1]$. Conversely, if $\mathcal{E} = \sigma(\{A_n\})$, then let

$$f = \sum_{n=1}^{\infty} 3^{-n} I_{A_n}.$$

The measurability of f is obvious. Since the preimages of Borel sets form a σ-algebra $\mathcal{A}$, for the proof of the equality $\mathcal{A} = \mathcal{E}$ it is sufficient to verify that $\mathcal{A}$ contains all sets A_n. The latter is easily seen from the equalities $A_1 = f^{-1}([1/3,2/3])$, $A_2 = f^{-1}([1/9,2/9] \cup [1/3+1/9, 1/3+2/9])$, and so on. The theorem is proven. $\square$

6.5.6. Corollary. *Let $(X,\mathcal{A},\mu)$ be a measure space with a finite measure μ, let $(E,\mathcal{E})$ be a space with a countably generated σ-algebra $\mathcal{E}$, and let $F\colon X\to E$ be a μ-measurable mapping, i.e., $F^{-1}(\mathcal{E}) := \{F^{-1}(B)\colon B\in\mathcal{E}\}$ is contained in $\mathcal{A}_\mu$. Then, there exists a mapping $F_0\colon X\to E$ such that $F_0(x)=F(x)$ for μ-a.e. x and $F_0^{-1}(\mathcal{E})\subset\mathcal{A}$, i.e., F_0 is $(\mathcal{A},\mathcal{E})$-measurable.*

PROOF. By the above theorem, $\mathcal{E} = f^{-1}\bigl(\mathcal{B}(\mathbb{R}^1)\bigr)$ for some function f on E. The function $f\circ F$ is measurable with respect to μ and hence has an $\mathcal{A}$-measurable modification g. There is a set $Z\in\mathcal{A}$ of zero μ-measure outside of which g coincides with $f\circ F$. Let $F_0(x)=F(x)$ if $x\notin Z$ and $F_0(x)=e$ if $x\in Z$, where e is an arbitrary fixed element of E. It is clear that $F_0=F$ μ-a.e. Let $E\in\mathcal{E}$. Then $E=f^{-1}(B)$, where $B\in\mathcal{B}(\mathbb{R}^1)$. Since $F_0|_{X\backslash Z}=F|_{X\backslash Z}$ and $F_0|_Z=e$, we obtain

$$\begin{aligned}F_0^{-1}(E) &= \bigl(F_0^{-1}(E)\cap Z\bigr)\cup\bigl(F_0^{-1}(E)\cap(X\backslash Z)\bigr)\\ &= \bigl(F_0^{-1}(E\cap\{e\})\cap Z\bigr)\cup\bigl(g^{-1}(B)\cap(X\backslash Z)\bigr).\end{aligned}$$

Finally, $\bigl(F_0^{-1}(E\cap\{e\})\cap Z\bigr)$ is either empty or coincides with Z. □

6.5.7. Theorem. *Let $(E,\mathcal{E})$ be a measurable space. The following conditions are equivalent*:

(i) *$\mathcal{E}$ is a countably separated σ-algebra*;

(ii) *there exists an injective $\mathcal{E}$-measurable function $f\colon E\to[0,1]$*;

(iii) $\Delta_E := \{(x,x)\colon\ x\in E\}\in\mathcal{E}\otimes\mathcal{E}$;

(iv) *there exists a separable σ-algebra $\mathcal{E}_0\subset\mathcal{E}$ such that all singletons belong to $\mathcal{E}_0$.*

PROOF. If (i) is fulfilled and $\{E_n\}\subset\mathcal{E}$ is a countable family separating the points in E, then the function $f=\sum_{n=1}^\infty 3^{-n}I_{E_n}$ is $\mathcal{E}$-measurable and injective, as is easily seen. In order to derive property (iii) from property (ii) we observe that

$$\Delta_E = \{(x,y)\in E\times E\colon\ f(x)=f(y)\} = g^{-1}(\Delta_{[0,1]}),$$

where $g(x,y)=\bigl(f(x),f(y)\bigr)$, $g\colon E^2\to[0,1]^2$. Since the mapping g is measurable with respect to $\mathcal{E}\otimes\mathcal{E}$ and $\mathcal{B}([0,1]^2)$ and the diagonal is a Borel set, one has $\Delta_E\in\mathcal{E}\otimes\mathcal{E}$. Now let (iii) be fulfilled. We observe that every set $A\in\mathcal{E}\otimes\mathcal{E}$ is contained in the σ-algebra generated by sets $A_n\times A_k$ for some finite or countable collection of sets $A_n\in\mathcal{E}$ (Exercise 1.12.54). We take such a collection $\{A_n\}$ for $A=\Delta_E$. It remains to observe that for every $x\in E$, we have $\{x\}=\Delta_E\cap\{x\}\times E\in\sigma(\{A_n\})$. Indeed, the class of all sets $B\in\mathcal{E}\otimes\mathcal{E}$ with the property that $B\cap\{x\}\times E\in\sigma(\{A_n\})$, is a σ-algebra. In addition, this class contains all sets $A_n\times A_k$, since the section of $A_n\times A_k$ at the point x either is empty or coincides with A_k. Thus, all sets in $\sigma(\{A_n\times A_k\})$ enjoy the above-mentioned property, hence Δ_E has this property as well. Finally, (iv) yields (i): according to Lemma 6.5.3, any countable family of sets generating $\mathcal{E}_0$ must separate the points in E. □

The next theorem characterizes the class of measurable spaces that possess both countability properties considered above.

6.5.8. Theorem. *Let $(E, \mathcal{E})$ be a measurable space. Then $\mathcal{E}$ is countably generated and countably separated precisely when the space $(E, \mathcal{E})$ is isomorphic to some subset M in $[0,1]$ with the induced Borel σ-algebra, i.e., there exists an $\bigl(\mathcal{E}, \mathcal{B}(M)\bigr)$-measurable one-to-one mapping $f\colon E \to M$ such that*

$$\mathcal{E} = \{f^{-1}(B)\colon B \in \mathcal{B}(M)\}.$$

PROOF. By Example 6.5.2, the indicated condition is sufficient. Suppose that $\mathcal{E}$ is countably generated and countably separated. Let us take a countable collection of sets A_n separating the points in E and generating $\mathcal{E}$. Then the function f considered in the proof of Theorem 6.5.5 is injective. Let $M = f(E)$. It is clear that f is an isomorphism of the measurable spaces $(E, \mathcal{E})$ and $\bigl(M, \mathcal{B}(M)\bigr)$. $\square$

Of course, a separable σ-algebra $\mathcal{E}$ may not separate the points in the space, but if it does separate, then by Lemma 6.5.3 it is countably separated. On the other hand, a countably separated σ-algebra may not be countably generated. Let us consider a non-trivial example of this sort.

6.5.9. Example. Let $\mathcal{E}$ be some σ-algebra of subsets of $[0,1]$ containing all Souslin sets and belonging to the σ-algebra $\mathcal{L}$ of all Lebesgue measurable sets (for example, one can take $\mathcal{E} = \mathcal{L}$). Then $\mathcal{E}$ is not countably generated, although it contains all Borel sets; in particular, it is countably separated.

PROOF. Assume the contrary. As shown above, there exists an $\mathcal{E}$-measurable function $f\colon E \to [0,1]$ such that $\mathcal{E} = \{f^{-1}(B)\colon B \in \mathcal{B}([0,1])\}$. Since $\mathcal{E} \subset \mathcal{L}$, the function f is Lebesgue measurable. By Lusin's theorem, there is a compact set $K \subset [0,1]$ of positive Lebesgue measure such that the restriction of f to K is continuous. Then every set $E \subset K$ belonging to $\mathcal{E}$ is Borel, since $E = f^{-1}\bigl(B \cap f(K)\bigr)$ for some Borel set $B \subset [0,1]$ and $f(K)$ is compact. This leads to a contradiction, since we show in §6.7 that every compact set of positive Lebesgue measure contains non-Borel Souslin subsets (see Corollaries 6.7.11 and 6.7.13). $\square$

6.6. Souslin sets and their separation

In this section, we begin the study of Souslin sets in topological spaces. We discuss some basic properties of Souslin sets; then in the next section we concentrate on the case where the whole space is Souslin (for example, is complete separable metric), and finally return to general spaces.

6.6.1. Definition. *A set in a Hausdorff space is called Souslin if it is the image of a complete separable metric space under a continuous mapping. A Souslin space is a Hausdorff space that is a Souslin set. The empty set is Souslin as well.*

Souslin sets are also called analytic sets. The complement of a Souslin set in a Souslin space is called co-Souslin or coanalytic.

Note also that the images of Polish spaces under continuous one-to-one mappings to Hausdorff spaces are called *Lusin spaces*. It will be clear from the discussion below that not every Souslin space is Lusin.

Theorem 6.1.13 yields the following characterization.

6.6.2. Lemma. *A nonempty set in a Hausdorff space is Souslin precisely when it can be represented as the image of the space* $\mathbb{N}^\infty$ *under a continuous mapping.*

6.6.3. Proposition. *Every nonempty Souslin set is the image of the space* $\mathcal{R}$ *of irrational numbers of the interval* $(0,1)$ *under some continuous mapping and also is the image of* $(0,1)$ *under some Borel mapping.*

Proof. The first claim follows at once from Theorem 6.1.6. The second claim is an obvious corollary of the first one, since $\mathcal{R}$ can be represented as the image of $(0,1)$ under the Borel mapping that is identical on $\mathcal{R}$ and takes all rational numbers to $\sqrt{1/2}$. □

6.6.4. Lemma. *Every Souslin space is hereditary Lindelöf.*

Proof. Let X be a Souslin space. Then X is the image of a separable metric space M under a continuous mapping F. For any open sets $U_\alpha \subset X$, the sets $F^{-1}(U_\alpha)$ are open in M and cover the set $F^{-1}(\bigcup_\alpha U_\alpha)$. Hence one can choose a finite or countable subcover, which yields a countable subcover in $\{U_\alpha\}$. Thus, X is hereditary Lindelöf. □

6.6.5. Lemma. (i) *The image of a Souslin set under a continuous mapping to a Hausdorff space is a Souslin set.*

(ii) *Every open or closed subset of a Souslin space is Souslin.*

(iii) *If* A_n *is a Souslin set in a space* X_n *for every* $n \in \mathbb{N}$, *then* $\prod_{n=1}^\infty A_n$ *is a Souslin set in the space* $\prod_{n=1}^\infty X_n$.

Proof. Claim (i) is obvious.

(ii) Let $X = f(E)$, where $f\colon\, E \to X$ is a continuous mapping and E is a complete separable metric space. If $A \subset X$ is a closed set, then $E_0 = f^{-1}(A)$ is a closed subspace in E and hence is a complete separable metric space. If A is open, then $E_0 = f^{-1}(A)$ is an open set. According to Exercise 6.1.11 the space E_0 is homeomorphic to a complete separable metric space E_1, i.e., A is a continuous image of E_1.

(iii) If $A_n = f_n(E_n)$, where E_n is a complete separable metric space and $f_n\colon\, E_n \to X_n$ is a continuous mapping, then $E = \prod_{n=1}^\infty E_n$ is a complete separable metric space and $f = (f_1, f_2, \ldots)\colon\, E \to \prod_{n=1}^\infty X_n$ is a continuous mapping. □

Now we prove that the class of Souslin sets is closed under the A-operation; in particular, it admits countable unions and countable intersections. However, as will be shown below, the complement of a Souslin set even in the interval $[0,1]$ may not be Souslin.

Let $\mathcal{S}_X$ denote the class of all Souslin sets in a topological space X.

6.6.6. Theorem. *The class $\mathcal{S}_X$ in a Hausdorff space X is closed with respect to the A-operation. In particular, if sets A_n are Souslin, then so are $\bigcap_{n=1}^{\infty} A_n$ and $\bigcup_{n=1}^{\infty} A_n$.*

PROOF. (1) First we show that countable unions and countable intersections of Souslin sets are Souslin. Suppose that A_n is a Souslin set in X. Then there exist a separable metric space E_n and a continuous mapping $f_n\colon E_n \to X$ with $A_n = f_n(E_n)$. The union E of the spaces E_n becomes a complete separable metric space if the distances between the points of different spaces E_n and E_m are defined to be 1, and the distances between the points in every E_n are unchanged. We define the mapping $f\colon E \to X$ as follows: $f|_{E_n} = f_n$. Then f is continuous and $f(E) = \bigcup_{n=1}^{\infty} A_n$. According to what we have proved earlier, the set $A = \prod_{n=1}^{\infty} A_n$ is Souslin in the space X^{∞}. Let

$$D = \bigl\{(x_n) \in A\colon\ x_n = x_1,\ \forall\, n \geq 1\bigr\}.$$

Then D is closed, hence is a Souslin set in A. Set $g\bigl((x_n)\bigr) = x_1$ if $(x_n) \in D$. Then g is continuous and $g(D) = \bigcap_{n=1}^{\infty} A_n$.

(2) Let $\mathbb{A} = \bigl(A(n_1, \ldots, n_k)\bigr)$ be a table of Souslin sets. Let $N(n_1, \ldots, n_k)$ denote the set in $\mathbb{N}^{\infty}$ consisting of all $\nu = (\nu_i)$ such that $\nu_1 = n_1, \ldots, \nu_k = n_k$. Note that one has

$$\begin{aligned} C := & \bigcup_{(n_i)\in\mathbb{N}^{\infty}} \bigcap_{k=1}^{\infty} A(n_1, \ldots, n_k) \times N(n_1, \ldots, n_k) \\ = & \bigcap_{k=1}^{\infty} \bigcup_{(n_1,\ldots,n_k)\in\mathbb{N}^k} A(n_1, \ldots, n_k) \times N(n_1, \ldots, n_k). \end{aligned}$$

Indeed, a point (x, ν) belongs to the left-hand side precisely when

$$(x, \nu) \in \bigcap_{k=1}^{\infty} A(\nu_1, \ldots, \nu_k) \times N(\nu_1, \ldots, \nu_k).$$

Hence it belongs to the right-hand side. Conversely, if it belongs to the right-hand side, then we have $x \in A(\nu_1, \ldots, \nu_k)$ for every k, whence we obtain

$$(x, \nu) \in \bigcap_{k=1}^{\infty} A(\nu_1, \ldots, \nu_k) \times N(\nu_1, \ldots, \nu_k).$$

As shown in (1), the set C is Souslin in the space $X \times \mathbb{N}^{\infty}$. Let us consider the natural projection $\pi_X\colon X \times \mathbb{N}^{\infty} \to X$. It remains to verify that

$$\pi_X(C) = S(\mathbb{A}) = \bigcup_{(n_i)\in\mathbb{N}^{\infty}} \bigcap_{k=1}^{\infty} A(n_1, \ldots, n_k).$$

Indeed, it suffices to show that

$$\pi_X\Bigl(\bigcap_{k=1}^{\infty} A(n_1,\ldots,n_k)\times N(n_1,\ldots,n_k)\Bigr) = \bigcap_{k=1}^{\infty} A(n_1,\ldots,n_k).$$

The left-hand side of this equality obviously belongs to the right-hand side. If x belongs to the right-hand side, then for every k, the point x is the projection of some pair (x,ν^k) from $A(n_1,\ldots,n_k)\times N(n_1,\ldots,n_k)$. This means that $\nu_i^k = n_i$ if $i\le k$. Then the point x is the projection of the pair (x,ν), where $\nu=(n_1,n_2,\ldots)$. The proof is complete. □

6.6.7. Corollary. *Every Borel subset of a Souslin space is a Souslin space.*

PROOF. Denote by $\mathcal{E}$ the class of all Borel sets B in a Souslin space X such that B and $X\backslash B$ are Souslin sets. We know that the class $\mathcal{E}$ contains all closed sets. By construction it is closed with respect to complementation. Finally, the above theorem yields that this class admits countable intersections. Hence $\mathcal{E}$ is a σ-algebra containing all closed sets, i.e., one has $\mathcal{E}=\mathcal{B}(X)$. □

6.6.8. Theorem. *Every Souslin set in a Hausdorff space can be obtained from closed sets by means of the A-operation.*

PROOF. Let a set A be the image of the space $\mathbb{N}^\infty$ under a continuous mapping f. For every finite sequence $n_1,\ldots,n_k$ we denote by $F_{n_1,\ldots,n_k}$ the closure of $f(C_{n_1,\ldots,n_k})$, where

$$C_{n_1,\ldots,n_k} = \bigl\{(m_i)\in\mathbb{N}^\infty\colon\ (m_1,\ldots,m_k)=(n_1,\ldots,n_k)\bigr\}.$$

Let us show that $A=\bigcup_{(n_i)\in\mathbb{N}^\infty}\bigcap_{k=1}^\infty F_{n_1,\ldots,n_k}$. It suffices to prove that $f\bigl((n_i)\bigr)=\bigcap_{k=1}^\infty F_{n_1,\ldots,n_k}$ for all $(n_i)\in\mathbb{N}^\infty$. Suppose that this is not true for some element $(n_i)\in\mathbb{N}^\infty$. Then there exists a point $x\in\bigcap_{k=1}^\infty F_{n_1,\ldots,n_k}$ that differs from $f\bigl((n_i)\bigr)$. Since X is Hausdorff, the points x and $f\bigl((n_i)\bigr)$ have disjoint neighborhoods. Hence there exists an open set U such that $f\bigl((n_i)\bigr)\in U\subset\overline{U}$ and $x\notin\overline{U}$. By the continuity of f for all sufficiently large k we have $f(C_{n_1,\ldots,n_k})\subset U$, whence $x\in\overline{f(C_{n_1,\ldots,n_k})}\subset\overline{U}$, which is a contradiction. □

The following separation theorem is very important in the theory of Souslin sets.

6.6.9. Theorem. *Let A_i, $i\in\mathbb{N}$, be pairwise disjoint Souslin sets in a Hausdorff space X. Then, there exist pairwise disjoint Borel sets B_i such that $A_i\subset B_i$ for all $i\in\mathbb{N}$.*

PROOF. (1) First we make several general remarks. We shall say that disjoint sets M_i are Borel separated if there exist disjoint Borel sets B_i with $M_i\subset B_i$. If for every $i\in\mathbb{N}$, disjoint sets M and M_i are Borel separated, then so are the sets M and $\bigcup_{i=1}^\infty M_i$. Indeed, if $B_i,C_i\in\mathcal{B}(X)$, $M\subset B_i$, $M_i\subset C_i$, $C_i\cap B_i=\varnothing$, then $B:=\bigcap_{i=1}^\infty B_i$ and $C=\bigcup_{i=1}^\infty C_i$ are disjoint Borel sets and $M \subset B$, $\bigcup_{i=1}^{\infty} M_i \subset C$. Similarly, one verifies that if for every $i, j \in \mathbb{N}$, we have disjoint Borel separated sets M_i and P_j, then the sets $\bigcup_{i=1}^{\infty} M_i$ and $\bigcup_{j=1}^{\infty} P_j$ are Borel separated. In addition, $\bigcap_{i=1}^{\infty} M_i$ and $\bigcup_{i=1}^{\infty} P_i$ are Borel separated.

(2) Now we consider the case where we have only two disjoint Souslin sets. It is clear from step (1) of the proof of Theorem 6.6.6 that this reduces to the following situation: we have closed sets C and D in a complete separable metric space E and a continuous mapping $f\colon E \to X$ with $f(C) \cap f(D) = \varnothing$. Suppose that $f(C)$ and $f(D)$ cannot be separated by disjoint Borel sets. We represent E in the form $E = \bigcup_{i=1}^{\infty} E(i)$, where $E(i)$ are closed sets of diameter less than 1. According to the above observations, for some $n_1, m_1 \in \mathbb{N}$ the sets $f\bigl(C \cap E(n_1)\bigr)$ and $f\bigl(D \cap E(m_1)\bigr)$ are not Borel separated. By induction, for every k we construct closed sets $E(n_1, \ldots, n_k)$ and $E(m_1, \ldots, m_k)$ of diameter less than $1/k$ in E such that the sets $f\bigl(C \cap E(n_1, \ldots, n_k)\bigr)$ and $f\bigl(D \cap E(m_1, \ldots, m_k)\bigr)$ are not Borel separated and $E(p_1, \ldots, p_k) = \bigcup_{j=1}^{\infty} E(p_1, \ldots, p_k, j)$, where for all p_i and j the sets $E(p_1, \ldots, p_k, j)$ are closed and have diameter less than $(k+1)^{-1}$. By the completeness of E, there exist points $a, b \in E$ such that given $\varepsilon > 0$, for all sufficiently large k the sets $C \cap E(n_1, \ldots, n_k)$ and $D \cap E(m_1, \ldots, m_k)$ belong to the ε-neighborhoods of the points a and b, respectively. Note that $a \in C$, $b \in D$, since C and D are closed. Then, by the continuity of f, for all sufficiently large k the sets $f\bigl(C \cap E(n_1, \ldots, n_k)\bigr)$ and $f\bigl(D \cap E(m_1, \ldots, m_k)\bigr)$ belong to disjoint open neighborhoods of the points $f(a)$ and $f(b)$ (which are distinct, since $f(C) \cap f(D) = \varnothing$), i.e., are Borel separated. This contradiction proves the theorem in the considered partial case.

(3) Let us consider the general case of a countable family of disjoint Souslin sets A_i. As we proved, there exist disjoint Borel sets B_1 and C_1 with $A_1 \subset B_1$, $\bigcup_{i=2}^{\infty} A_i \subset C_1$. Further, there exist disjoint Borel sets $\widehat{B}_2$ and $\widehat{C}_2$ with $A_2 \subset \widehat{B}_2$ and $\bigcup_{i=3}^{\infty} A_i \subset \widehat{C}_2$. We set $B_2 = \widehat{B}_2 \cap C_1$ and $C_2 = \widehat{C}_2 \cap C_1$. Continuing this process by induction, we obtain the required sets B_i. □

6.6.10. Corollary. *Suppose that the complement of a Souslin set A in a Hausdorff space X is Souslin. Then A is a Borel set.*

Proof. There exist $B, C \in \mathcal{B}(X)$ such that $B \cap C = \varnothing$, $A \subset B$ and $X \backslash A \subset C$. Then $A = B$ and $X \backslash A = C$. □

The proof of the following result of P.S. Novikoff can be found in Dellacherie [**425**, p. 251], Rogers, Jayne [**1589**, p. 58].

6.6.11. Theorem. *Let A_n, $n \in \mathbb{N}$, be Souslin sets in a Hausdorff space such that $\bigcap_{n=1}^{\infty} A_n$ is a Borel set. Then there exist Borel sets B_n such that $A_n \subset B_n$ and $\bigcap_{n=1}^{\infty} A_n = \bigcap_{n=1}^{\infty} B_n$.*

Let us also mention Lusin's theorem on separation by coanalytic set (see Dellacherie [**425**, p. 247], Hoffmann-Jørgensen [**841**, p. 80], or Lusin [**1209**, Ch. III] for a proof).

6.6.12. Theorem. *Let A and B be Souslin sets in a Polish space X. Then there exist coanalytic sets C and D such that*

$$A\backslash B \subset C,\ B\backslash A \subset D,\ C \cap D = \varnothing,\ C \cup D = X\backslash(A \cap B).$$

6.7. Sets in Souslin spaces

In this section, we discuss Souslin sets in Souslin spaces. In particular, everything said below applies to complete separable metric spaces and their Borel subsets. In addition to several general results, we shall obtain an example of a non-Borel Souslin set. As above, Γ_f denotes the graph of a mapping f.

6.7.1. Lemma. *Let X and Y be Souslin spaces. Then the graph Γ_f of any Borel mapping $f\colon X \to Y$ is a Borel, hence Souslin, subset in the Souslin space $X\times Y$. Conversely, if $f\colon X \to Y$ has a Souslin graph, then f is Borel measurable.*

PROOF. The first assertion follows from Corollary 6.4.5 and Lemma 6.6.4. In order to prove the converse, we observe that for any $B \in \mathcal{B}(Y)$, the sets $f^{-1}(B)$ and $f^{-1}(Y\backslash B)$ are Souslin as the projections of $\Gamma_f \cap (X\times B)$ and $\Gamma_f \cap \bigl(X\times(Y\backslash B)\bigr)$, respectively. By Corollary 6.6.10, we obtain the inclusion $f^{-1}(B) \in \mathcal{B}(X)$. □

6.7.2. Theorem. *Let X be a Souslin space* (*e.g., a complete separable metric space*) *and let A be its subset. The following are equivalent*:

(i) *A is a Souslin set*;

(ii) *A can be obtained by the A-operation on closed sets in X*;

(iii) *A is the projection of a closed set in the space $X\times\mathbb{N}^\infty$*;

(iv) *A is the projection of a Borel set in $X\times\mathbb{R}^1$.*

PROOF. The equivalence of (i) and (ii) follows by Theorem 6.6.8 and Theorem 6.6.6 taking into account that all closed sets in a Souslin space are Souslin. Since the spaces $X\times\mathbb{N}^\infty$ and $X\times\mathbb{R}^1$ are Souslin, all Borel sets in them are Souslin by Corollary 6.6.7. Hence (iii) and (iv) imply (i). In order to deduce (iii) from (i), we observe that the set A is the image of $\mathbb{N}^\infty$ under some continuous mapping $f\colon \mathbb{N}^\infty \to X$, hence coincides with the projection of Γ_f on X. Note that Γ_f is closed in the Souslin space $\mathbb{N}^\infty\times X$. Finally, we verify that (i) yields (iv). To this end, we represent A as the image of $\mathbb{R}^1$ under a Borel mapping f. This can be done by using Proposition 6.6.3. It remains to observe that the graph of f is a Borel subset of $\mathbb{R}^1\times X$, and A is its projection on X. □

6.7.3. Theorem. *Let X and Y be Souslin spaces and let $f\colon X \to Y$ be a Borel mapping. Then, for all Souslin sets $A \subset X$ and $C \subset Y$, the sets $f(A)$ and $f^{-1}(C)$ are Souslin. In particular, this is true if f is continuous.*

If f is injective, then the mapping $f^{-1}\colon f(X) \to X$ is Borel.

PROOF. By Lemma 6.7.1, the graph of the mapping $f|_A$ is a Souslin set in the Souslin space $A{\times}Y$. Hence its projection on Y, equal to $f(A)$, is a Souslin set. Similarly, $f^{-1}(C)$ is the projection on X of the set $D = \Gamma_f \cap (X\times C)$. It remains to observe that D is a Souslin set, since so are Γ_f and $X\times C$. If f is injective, then $f(B) \in \mathcal{B}\bigl(f(X)\bigr)$ for any $B \in \mathcal{B}(X)$ by Corollary 6.6.10, since $f(B)$ and $f(X)\backslash f(B) = f(X\backslash B)$ are Souslin sets in $f(X)$. □

Even for continuous injective f the set $f(B)$ with $B \in \mathcal{B}(X)$ need not belong to $\mathcal{B}(Y)$: take a non-Borel Souslin set $X \subset [0,1]$ (see below) and its identical embedding into $Y = [0,1]$. However, see Theorem 6.8.6.

6.7.4. Theorem. *Let X be a Souslin space. Then, there exist a Souslin subset S in the interval $[0,1]$ and a one-to-one Borel mapping h from the space X onto S such that h is an isomorphism of the measurable spaces $\bigl(X, \mathcal{B}(X)\bigr)$ and $\bigl(S, \mathcal{B}(S)\bigr)$.*

PROOF. As we know, the space $X\times X$ is Souslin. By Lemma 6.6.4 it is hereditary Lindelöf. According to Corollary 6.4.5, the diagonal in $X\times X$ belongs to $\mathcal{B}(X)\otimes\mathcal{B}(X)$, whence by Theorem 6.5.7 we obtain the existence of an injective Borel function $h\colon\ X \to [0,1]$. Set $S = f(X)$. By Theorem 6.7.3 the set S is Souslin and $h\colon\ X \to S$ is a Borel isomorphism. □

6.7.5. Corollary. *The Borel σ-algebra of a Souslin space is countably generated and countably separated.*

6.7.6. Corollary. *Let μ be a finite measure on a measurable space $(X, \mathcal{A})$, let Y be a Souslin space, and let $F\colon\ X \to Y$ be a μ-measurable mapping, i.e., $F^{-1}(B) \in \mathcal{A}_\mu$ for all $B \in \mathcal{B}(Y)$. Then, there exists a mapping $G\colon\ X \to Y$ such that $F = G$ μ-a.e. and $G^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}(Y)$.*

PROOF. One can apply Corollary 6.5.6. □

6.7.7. Theorem. *Let X be a completely regular Souslin space. Then*

(i) *X is perfectly normal; in particular, the Borel and Baire σ-algebras in X coincide;*

(ii) *there exists a countable family of continuous functions on X separating the points in X.*

PROOF. Let U be open in X. By the complete regularity, for every point $x \in U$, there exists a continuous function $f_x\colon\ X \to [0,1]$ such that $f_x(x) = 1$ and $f_x = 0$ outside U. The open sets $U_x = \bigl\{z\colon\ f_x(z) > 0\bigr\}$ cover U. By Lemma 6.6.4, there is an at most countable subcover $\{U_{x_n}\}$ of the set U. It remains to observe that $U = \{f > 0\}$, where the function $f = \sum_{n=1}^\infty 2^{-n} f_{x_n}$ is continuous. Indeed, $f = 0$ outside U. For every $y \in U$, there exists n with $y \in U_{x_n}$, i.e., $f_{x_n}(y) > 0$. Thus, X is a perfectly normal space.

The space $X\times X$ is Souslin as well. By Lemma 6.6.4 it is hereditary Lindelöf. Hence (ii) follows by Proposition 6.5.4. □

We note that even a countable Souslin space may not be completely regular (Exercise 6.10.78).

6.7.8. Corollary. *Every compact subset in a Souslin space is metrizable.*

PROOF. Since all closed subsets of Souslin spaces are Souslin, it suffices to establish the metrizability of every compact Souslin space K. In turn, it suffices to show the existence of a countable family of continuous functions separating the points in K (Exercise 6.10.24). Since every compact space is completely regular, assertion (ii) applies. $\square$

Let us show that there exist non-Borel Souslin sets. First we prove an interesting auxiliary result.

6.7.9. Proposition. *Suppose that we are given a complete separable metric space X. Then:*

(i) *there exists a closed set $Z \subset X \times \mathbb{N}^\infty$ such that every closed set in X coincides with one of the sections $Z_\nu := \{x \in X\colon\ (x,\nu) \in Z\}$, $\nu \in \mathbb{N}^\infty$;*

(ii) *there exists a Souslin set $A \subset X \times \mathbb{N}^\infty$ such that every Souslin set in X coincides with one of the sections $A_\nu := \{x \in X\colon\ (x,\nu) \in A\}$, $\nu \in \mathbb{N}^\infty$.*

PROOF. (i) Let $\{U_n\}$ be a countable base of the topology in X. Set

$$Z = \Bigl\{(x,\nu) \in X \times \mathbb{N}^\infty\colon\ \nu = (n_i),\ x \notin \bigcup_{i=1}^\infty U_{n_i}\Bigr\}.$$

Every closed set in X is the complement of some union of the sets U_n, hence coincides with one of the sections Z_ν, $\nu \in \mathbb{N}^\infty$. Note that Z is closed, since its complement is open. Indeed, let x and $\nu = (\nu_i)$ be such that x belongs to U_{ν_i} for some i. Then for all (x',η) sufficiently close to (x,ν), we have $\eta_i = \nu_i$ and $x' \in U_{\eta_i} = U_{\nu_i}$.

(ii) Let us apply (i) to the space $X \times \mathbb{N}^\infty$ and take a corresponding closed set $Z \subset X \times \mathbb{N}^\infty \times \mathbb{N}^\infty$. Let

$$A = \bigl\{(x,\nu) \in X \times \mathbb{N}^\infty\colon\ (x,\eta,\nu) \in Z \text{ for some } \eta \in \mathbb{N}^\infty\bigr\}.$$

The set A is Souslin, since it can be represented as the projection of a closed set in $X \times \mathbb{N}^\infty \times \mathbb{N}^\infty$. Every Souslin set E in the space X is the projection of some closed set in $X \times \mathbb{N}^\infty$, i.e., the projection of some section Z_ν. Therefore, we have $E = A_\nu$. $\square$

6.7.10. Theorem. *The space $\mathbb{N}^\infty$ contains a Souslin set that is not Borel.*

PROOF. Let us apply assertion (ii) of the above proposition to $X = \mathbb{N}^\infty$ and take a corresponding Souslin set $A \subset \mathbb{N}^\infty \times \mathbb{N}^\infty$. The set

$$S = \bigl\{\nu \in \mathbb{N}^\infty\colon\ (\nu,\nu) \in A\bigr\}$$

is Souslin in $\mathbb{N}^\infty$ as the projection of the intersection of A with the diagonal. Its complement

$$\mathbb{N}^\infty \backslash S = \bigl\{\nu \in \mathbb{N}^\infty\colon\ \nu \notin A_\nu\bigr\}$$

is not Souslin, since otherwise due to our choice of A, we would have for some ν the equality

$$\mathbb{N}^\infty \backslash S = A_\nu,$$

which yields simultaneously $\nu \notin A_\nu$ and $\nu \in A_\nu$ by the construction of S. Therefore, S is not Borel. □

6.7.11. Corollary. *A non-Borel Souslin set exists in every space that contains a subset homeomorphic to the space $\mathbb{N}^\infty$, in particular, in every nonempty complete metric space without isolated points.*

PROOF. If a set X_0 in a space X is homeomorphic to $\mathbb{N}^\infty$ and A is a non-Borel Souslin set in X_0, then A is Souslin and non-Borel in the space X. The second claim of the corollary follows by Lemma 6.1.16. □

6.7.12. Theorem. *If f is a continuous mapping of a complete separable metric space X onto an uncountable Hausdorff space Y, then X contains a set E that is homeomorphic to the Cantor set C such that f maps E homeomorphically onto $f(E)$.*

PROOF. In every set $f^{-1}(y)$, $y \in Y$, we choose a point and obtain an uncountable set $X_0 \subset X$, on which f is injective. We shall consider X_0 as a metric space and take the set X_1 of all points $x \in X_0$ every neighborhood of which contains uncountably many points in X_0. It is easily verified that the metric space X_1 is uncountable and has no isolated points. One can find a Souslin scheme $\mathbb{A}$ in X indexed by finite sequences $(n_1, \ldots, n_k)$ of 0 and 1 such that every set $A(n_1, \ldots, n_k)$ is open, meets X_1, has diameter at most $1/k$, the closure of $A(n_1, \ldots, n_k, n_{k+1})$ is contained in $A(n_1, \ldots, n_k)$, and the closure of $f\bigl(A(n_1, \ldots, n_k)\bigr)$ does not meet $f\bigl(A(m_1, \ldots, m_k)\bigr)$ if $(m_1, \ldots, m_k)$ does not coincide with $(n_1, \ldots, n_k)$. The required scheme is constructed inductively. First we take balls $A(0)$ and $A(1)$ of radius less than 1 with the centers $a_1 \in X_1$ and $a_2 \in X_1$ such that the closures of their images under f do not meet. Then in $A(0)$ we find balls $A(0,0)$ and $A(0,1)$ of radius less than $1/2$ such that their closures lie in $A(0)$ and the closures of their images do not meet. We do the same with $A(1)$. The process continues inductively. This scheme defines a homeomorphism $g\colon \{0,1\}^\infty \to X$, $g\bigl((n_i)\bigr) = \bigcap_{i=1}^\infty A(n_1, \ldots, n_i)$. One can also define a homeomorphism $h\colon C \to X$ by the formula $h(c) = A(c_1) \cap A(c_1, c_2) \cap \cdots$, where $c = 2c_1/3 + 2c_2/9 + \cdots$, $c_i \in \{0,1\}$. It is verified that f is injective on the set $E = g(\{0,1\}^\infty) = h(C)$, which by the compactness of this set means that $f|_E$ is a homeomorphism. An analogous reasoning is presented in more detail in Kuratowski [**1082**, §36.V, p. 455], Hoffmann-Jørgensen [**841**, §1.5.H]. □

6.7.13. Corollary. *Every uncountable Souslin space contains a set that is homeomorphic to the Cantor set and has cardinality of the continuum.*

It follows by the above that the classes of Souslin and Borel subsets in a Souslin space X have cardinality at most of the continuum, and if X is uncountable, then their cardinality is precisely $\mathfrak{c}$.

6.7.14. Remark. We know that all Borel sets on the real line are obtained by means of the Souslin operation on closed sets, which, however, produces non-Borel sets as well. Hausdorff raised the question on the existence of an operation that produces exactly the Borel sets. The precise formulation is this. Let $\mathfrak{M}$ be some family of sets. Denote by $\mathfrak{B}(\mathfrak{M})$ the smallest class of sets that contains $\mathfrak{M}$ and is closed with respect to countable unions and countable intersections. For example, if $\mathfrak{M}$ is the class of all open sets on the real line, then $\mathfrak{B}(\mathfrak{M}) = \mathcal{B}(\mathbb{R}^1)$. Hausdorff asked: does there exist a set $N \subset \mathbb{N}^\infty$ such that for every family of sets $\mathfrak{M}$, one has the equality $\mathfrak{B}(\mathfrak{M}) = \bigcup_{(n_i)\in N} \bigcap_{i=1}^\infty M_{n_i}$, where $M_{n_i} \in \mathfrak{M}$? Sierpiński [**1714**] proved that there are no such sets N.

6.8. Mappings of Souslin spaces

Let X and Y be Souslin spaces and let $f\colon X \to Y$ be a Borel mapping. In this section, we discuss descriptive properties of the sets of points $y \in Y$ such that the equation $f(x) = y$ has a unique solution, n solutions or infinitely many solutions. We consider a somewhat more general problem concerning the analogous properties of the sections $A_y := \{x \in X\colon (x, y) \in A\}$ of sets $A \in X\times Y$. The former problem is a partial case of this more general one if we take for A the graph of f.

Let $\operatorname{Card} M$ denote the cardinality of a set M and let $\aleph_0$ denote the cardinality of $\mathbb{N}$.

We recall that by Theorem 6.7.3 the images of Souslin sets under Borel mappings between Souslin spaces are Souslin. However, it is important here that the range space is Souslin.

6.8.1. Example. The identity mapping from $X = \mathbb{R}^1$ with the usual topology onto the Sorgenfrey line Z (see Example 6.1.19) is Borel, since any open set in Z is an at most countable union of semiclosed intervals. But Z is not Souslin by Corollary 6.7.13, since it contains no uncountable compact sets.

6.8.2. Theorem. *Let A be a Souslin set in $X\times Y$. Then, for any $n \in \mathbb{N}$, the sets $\{y \in Y\colon \operatorname{Card} A_y \geq \aleph_0\}$ and $\{y \in Y\colon \operatorname{Card} A_y \geq n\}$ are Souslin. The set $\{y \in Y\colon \operatorname{Card} A_y = 1\}$ is the difference of two Souslin sets.*

Proof. We take a countable algebra $\mathcal{E} \subset \mathcal{B}(X)$ separating the points in X. Then the condition $\operatorname{Card} A_y \geq n$, which means that there exist n distinct points $x_1, \ldots, x_n$ in A_y, is equivalent to the existence of pairwise disjoint sets $E_1, \ldots, E_n$ in $\mathcal{E}$ with $E_j \cap A_y \neq \varnothing$ for all $j \leq n$. Letting π_Y be the projection operator from $X\times Y$ on Y, the latter can be written as follows:

$$y \in \bigcap_{j=1}^{n} \pi_Y\bigl((E_j \times Y) \cap A\bigr).$$

Let $\mathcal{E}_n$ be the family of all collections $\{E_1, \ldots, E_n\}$ consisting of n pairwise disjoint sets $E_j \in \mathcal{E}$. The cardinality of $\mathcal{E}_n$ is at most countable and

$$\{y \in Y\colon \operatorname{Card} A_y \geq n\} = \bigcup_{\sigma\in\mathcal{E}_n} \bigcap_{E\in\sigma} \pi_Y\bigl((E\times Y) \cap A\bigr).$$

Since the set $\pi_Y\bigl((E\times Y)\cap A\bigr)$ is Souslin, the set $\{y\in Y\colon \operatorname{Card} A_y\ge n\}$ is Souslin as well. This yields that

$$\{y\in Y\colon \operatorname{Card} A_y\ge \aleph_0\}=\bigcap_{n=1}^{\infty}\{y\in Y\colon \operatorname{Card} A_y\ge n\}$$

is a Souslin set. The last claim follows trivially by the first one. □

We observe that although X and Y are Souslin spaces throughout this section, in the above theorem we need not assume this because the case of general spaces reduces to the considered one due to the fact that the projections of A to X and Y are Souslin sets. See also Theorem 6.10.18 below.

6.8.3. Corollary. *Let X and Y be Souslin spaces and let $f\colon X\to Y$ be a Borel mapping. Then the sets*

$$\bigl\{y\in Y\colon \operatorname{Card} f^{-1}(y)\ge n\bigr\}\quad\textit{and}\quad \bigl\{y\in Y\colon \operatorname{Card} f^{-1}(y)\ge \aleph_0\bigr\}$$

are Souslin. The set $\{y\in Y\colon \operatorname{Card} f^{-1}(y)=1\}$ is the difference of two Souslin sets.

We note that the difference of two Souslin sets can be a set of a more complex nature: it may be neither Souslin nor co-Souslin. But if the set A is closed, then the set $\{y\in Y\colon \operatorname{Card} A_y=1\}$ turns out to be the complement of a Souslin set. In particular, if f in the above corollary is continuous, then $\{y\in Y\colon \operatorname{Card} f^{-1}(y)=1\}$ is the complement of a Souslin set (the proof can be found in Hoffmann-Jørgensen [**841**], where there are some more general results).

We now discuss the properties of injective Borel mappings. In particular, we shall characterize the Borel sets in complete separable metric spaces as the injective continuous images of closed subsets in the space $\mathbb{N}^\infty$ (or, which amounts to the same thing, in the space of irrational numbers).

6.8.4. Lemma. *Let X be a complete separable metric space. Then, every Borel set in X is the injective image of some closed set in $X\times\mathbb{N}^\infty$ under the natural projection $X\times\mathbb{N}^\infty\to X$.*

Proof. We show that the class $\mathcal{E}$ of all Borel sets with the indicated property contains all open sets and is closed with respect to formation of countable unions of disjoint sets and countable intersections. Then a reference to Proposition 6.2.9 completes the proof. Let G be open in X. Then the set

$$E=\bigl\{(x,t)\in X\times(0,+\infty)\colon \operatorname{dist}(x,X\backslash G)=t^{-1}\bigr\}$$

is closed in $X\times(0,+\infty)$, G coincides with its projection on X, and the projection operator is injective on E. However, this is not yet what we wanted because a set from $X\times\mathbb{N}^\infty$ is required. By Corollary 6.1.7, there exist a closed set $D\subset\mathbb{N}^\infty$ and a continuous one-to-one mapping f of the set D onto $(0,+\infty)$. Then the set

$$Z=\bigl\{(x,n)\in X\times\mathbb{N}^\infty\colon \bigl(x,f(n)\bigr)\in E\bigr\}$$

has the required properties. Thus, all open sets in X belong to the class $\mathcal{E}$.

Suppose now that sets $A_j \in \mathcal{E}$ are pairwise disjoint. Let us take closed sets Z_j in $X\times\mathbb{N}^\infty\times\{j\}$ that are projected injectively onto A_j. It is easily seen that the set $Z=\bigcup_{j=1}^\infty Z_j$ is closed in $X\times\mathbb{N}^\infty\times\mathbb{N}$ and is projected one-to-one onto $\bigcup_{j=1}^\infty A_j$. Since the space $\mathbb{N}^\infty\times\mathbb{N}$ is homeomorphic to $\mathbb{N}^\infty$ by means of the homeomorphism

$$h\colon\ (\eta,k)\mapsto(k,\eta_1,\eta_2,\ldots),\quad \eta=(\eta_i),$$

the set $C=\bigl\{(x,\eta)\colon\ \bigl(x,h^{-1}(\eta)\bigr)\in Z\bigr\}$ is closed in $X\times\mathbb{N}^\infty$ and is projected one-to-one onto $\bigcup_{j=1}^\infty A_j$.

Finally, for arbitrary $A_j\in\mathcal{E}$, we choose closed sets $C_j\subset X\times\mathbb{N}^\infty$ that are projected one-to-one onto A_j. Let us consider the set

$$Z=\bigl\{(x,\eta^1,\eta^2,\ldots)\colon\ x\in X,\ \eta^j\in\mathbb{N}^\infty,\ (x,\eta^j)\in C_j,\ j\in\mathbb{N}\bigr\}\subset X\times\bigl(\mathbb{N}^\infty\bigr)^\infty.$$

It is clear that the set Z is closed in $X\times\bigl(\mathbb{N}^\infty\bigr)^\infty$ and is projected one-to-one onto $\bigcap_{j=1}^\infty A_j$. Similarly to the previous step, it remains to observe that the space $\bigl(\mathbb{N}^\infty\bigr)^\infty$ is homeomorphic to $\mathbb{N}^\infty$. □

6.8.5. Corollary. *Every Borel set in a Polish space is the image of some closed set in $\mathbb{N}^\infty$ under a continuous one-to-one mapping.*

PROOF. Follows by the lemma and Theorem 6.1.15. □

6.8.6. Theorem. *Let B be a Borel set in a complete separable metric space X, let Y be a Souslin space, and let $f\colon B\to Y$ be an injective Borel mapping. Then $f(B)$ is a Borel set in Y.*

PROOF. By Theorem 6.7.4 it suffices to prove our claim for mappings to $[0,1]$. The graph of f is a Borel set in $X\times[0,1]$, and its projecting to $[0,1]$ is injective due to the injectivity of f. Hence the assertion reduces to the case of continuous f. Now we assume that $Y=[0,1]$ and f is continuous. In addition, by Lemma 6.8.4 we can assume that B is a closed subset in $X\times\mathbb{N}^\infty$, i.e., is a complete separable metric space. As in the proof of Theorem 6.1.13, to every finite sequence of natural numbers $n_1,\ldots,n_k$, we associate a nonempty closed set $E(n_1,\ldots,n_k)\subset B$ of diameter less than 2^{-k-2} in such a way that

$$B=\bigcup_{j=1}^\infty E(j),\quad E(n_1,\ldots,n_k)=\bigcup_{j=1}^\infty E(n_1,\ldots,n_k,j).$$

Let $A(n)=E(n)\backslash\bigcup_{j=1}^{n-1}E(j)$, and for $k>1$ let

$$A(n_1,\ldots,n_k)=A(n_1,\ldots,n_{k-1})\cap E(n_1,\ldots,n_k)\backslash\bigcup_{j<n_k}E(n_1,\ldots,n_{k-1},j).$$

If $k\in\mathbb{N}$ is fixed, the Borel sets $A(n_1,\ldots,n_k)$ are disjoint and their union over all $n_1,\ldots,n_k$ is B. By the injectivity of f the Souslin sets $f\bigl(A(n_1,\ldots,n_k)\bigr)$

are pairwise disjoint. According to Theorem 6.6.9, there exist disjoint Borel sets $B(n_1,\ldots,n_k)$ in Y such that

$$f\bigl(A(n_1,\ldots,n_k)\bigr)\subset B(n_1,\ldots,n_k).$$

We can have the inclusion $B(n_1,\ldots,n_k)\subset\overline{f\bigl(A(n_1,\ldots,n_k)\bigr)}$ by passing to the Borel sets $B(n_1,\ldots,n_k)\cap\overline{f\bigl(A(n_1,\ldots,n_k)\bigr)}$. Let us show that

$$f(B)=\bigcap_{k=1}^{\infty}\ \bigcup_{(n_1,\ldots,n_k)\in\mathbb{N}^k}B(n_1,\ldots,n_k), \tag{6.8.1}$$

whence our assertion follows in an obvious way. To this end, we first observe that

$$\bigcap_{k=1}^{\infty}\ \bigcup_{(n_1,\ldots,n_k)\in\mathbb{N}^k}B(n_1,\ldots,n_k)=\bigcup_{(n_i)\in\mathbb{N}^\infty}\ \bigcap_{k=1}^{\infty}B(n_1,\ldots,n_k). \tag{6.8.2}$$

Indeed, the right-hand side of (6.8.2) belongs to the left-hand side in an obvious way. Conversely, if a point y belongs to the left-hand side of (6.8.2), then for every k, this point is contained in exactly one of the sets $B(n_1,\ldots,n_k)$ due to their disjointness. The corresponding indices are denoted by $n_1(k),\ldots,$ $n_k(k)$. One has $n_i(k+1)=n_i(k)$ whenever $i\le k$, since $y\notin B(m_1,\ldots,m_k)$ if $(m_1,\ldots,m_k)\neq(n_1,\ldots,n_k)$. Thus, $y\in\bigcap_{k=1}^{\infty}B\bigl(n_1(1),n_2(2),\ldots,n_k(k)\bigr)$, and (6.8.2) is established. The set defined by equality (6.8.2) will be denoted by D. Then one has

$$f(B)\subset\bigcup_{(n_i)\in\mathbb{N}^\infty}\ \bigcap_{k=1}^{\infty}f\bigl(A(n_1,\ldots,n_k)\bigr)\subset D.$$

On the other hand,

$$D\subset\bigcup_{(n_i)\in\mathbb{N}^\infty}\ \bigcap_{k=1}^{\infty}\overline{f\bigl(E(n_1,\ldots,n_k)\bigr)}\subset f(B).$$

Indeed, if $y\in\bigcap\limits_{k=1}^{\infty}\overline{f\bigl(E(n_1,\ldots,n_k)\bigr)}$, the set $E(n_1,\ldots,n_k)$ contains points $x_{n_1,\ldots,n_k}$ such that $f(x_{n_1,\ldots,n_k})\to y$. The sequence $\{x_{n_1,\ldots,n_k}\}$ is fundamental. Since B is complete, this sequence converges to some $x\in B$, whence one has $y=f(x)$ by the continuity of f. Therefore, we obtain (6.8.1). $\square$

It is worth noting that the image of a Souslin space under an injective continuous mapping may not be Borel: it suffices to take a non-Borel Souslin set in $[0,1]$ and consider its embedding in $[0,1]$. However, the above theorem obviously remains valid for all Souslin spaces that are injective images of Polish spaces (the so-called Lusin spaces).

6.8.7. Corollary. *A set in a Polish space is Borel precisely when it is the image of a closed subset of* $\mathbb{N}^\infty$ *under a continuous injective mapping.*

6.8.8. Corollary. *Any two uncountable Borel sets in Polish spaces are Borel isomorphic.*

PROOF. By the previous corollary and the above-established fact that all continuous injective mappings of Polish spaces take Borel sets to Borel ones, we obtain that it suffices to prove the existence of a Borel isomorphism between $\mathbb{R}^1$ and any uncountable closed subset in $\mathbb{N}^\infty$, or, equivalently, in the space $\mathcal{R}$ of irrational numbers in $(0,1)$. We observe that if two uncountable Borel spaces A and B are Borel isomorphic, then for any at most countable subset $C \subset A$, the spaces $A\backslash C$ and B are Borel isomorphic as well. If C is infinite, then it suffices to take in B a part C' corresponding to the set C with the added countable subset $D \subset A\backslash C$, to establish a one-to-one correspondence between D and C', and keep the initial isomorphism between $A\backslash(C\cup D)$ and $B\backslash C'$. The case of finite C can be reduced to the considered one. Thus, we may neglect countable subsets. If now M is a closed subset in the space $\mathcal{R}$ of irrational numbers in the interval $(0,1)$, then it coincides up to a countable set with some closed subset A in the closed interval. If the interior of A is not empty, it is Borel isomorphic to $(0,+\infty)$. So we may assume that A has no interior points. The set of all points $x \in A$ that possess a neighborhood meeting A in an at most countable set, is at most countable. Hence we can assume by the above observation that A is perfect. By Proposition 6.1.17 it remains to consider the case when A is the Cantor set (the interior of A, if it is nonempty, is obviously Borel isomorphic to $(0,1)$). In that case, the existence of a Borel isomorphism is verified directly (for example, by using the ternary expansion for the Cantor set and the binary expansion for the interval). $\square$

It is clear from the above that there are only two classes of pairwise isomorphic infinite standard measurable spaces in the sense of Definition 6.2.10: countable and of cardinality of the continuum.

6.8.9. Theorem. *Let $\{f_n\}$ be a sequence of Borel functions on a Souslin space X separating the points in X. Then $\{f_n\}$ generates the Borel σ-algebra of X.*

PROOF. It follows by our hypothesis that the countable family of Borel sets B_n of the form $f_k^{-1}\bigl((r_i,r_j)\bigr)$, where $\{r_j\}$ are all rational numbers, separates the points in X. It was shown in the proof of Theorems 6.7.4, 6.5.5 and 6.5.7 that the function $h=\sum_{n=1}^\infty 3^{-n}I_{B_n}$ maps X one-to-one onto the Souslin set $S:=f(X)$ in $[0,1]$ and for every $B\in\mathcal{B}(X)$, we have $B=h^{-1}\bigl(h(B)\bigr)$, where $h(B)\in\mathcal{B}(S)$. This means that there exists a set $C\in\mathcal{B}(\mathbb{R}^1)$ such that $h(B)=C\cap S$ and $B=h^{-1}(C)$. Thus, the function h generates $\mathcal{B}(X)$. Hence $\mathcal{B}(X)=\sigma(\{f_n\})$. $\square$

6.8.10. Example. Let K be a compact metric space and let a sequence $\{x_n\}$ be dense in K. Then the Borel σ-algebra of the separable Banach space $C(K)$ is generated by the functions $\varphi\mapsto\varphi(x_n)$ on $C(K)$, since they separate the points of $C(K)$.

The following deep and important result is due to P.S. Novikoff [**1383**].

6.8.11. Theorem. *There exist two disjoint coanalytic sets in* $\{0,1\}^\infty$ *that cannot be separated by Borel sets. The same is true for any uncountable Polish space.*

PROOF. Let us show that there exist Souslin sets A_0 and A_1 in $\{0,1\}^\infty$ such that the sets $A_0\backslash A_1$ and $A_1\backslash A_0$ cannot be separated by Borel sets. We know that there is a Souslin set $S\subset\{0,1\}\times\{0,1\}^\infty\times\{0,1\}^\infty$ that is universal for the Souslin sets in $\{0,1\}\times\{0,1\}^\infty$. We have $S=\{0\}\times S_0\cup\{1\}\times S_1$, where S_0 and S_1 are Souslin sets in $\{0,1\}^\infty\times\{0,1\}^\infty$. We show that $S_0\backslash S_1$ and $S_1\backslash S_0$ cannot be separated by Borel sets. Suppose that B_0 and B_1 are disjoint Borel sets with $S_0\backslash S_1\subset B_0$, $S_1\backslash S_0\subset B_1$. Clearly, we may assume that B_0 is the complement of B_1. In Exercise 6.10.31, the Borel classes $\mathcal{B}_\alpha$ corresponding to at most countable ordinals α are introduced such that their union is the class of all Borel sets of a given space. So one has $B_0\in\mathcal{B}_\tau$ for some ordinal τ with $0\le\tau<\omega_1$. According to that exercise, there is a Borel set C_0 in $\{0,1\}^\infty$ that does not belong to $\mathcal{B}_\tau$. Let C_1 be its complement. The set $C:=\{0\}\times C_0\cup\{1\}\times C_1$ is Souslin in $\{0,1\}\times\{0,1\}^\infty$. As S is universal, there is a point $x\in\{0,1\}^\infty$ with $C=S_x$, hence $C_0=(S_0)_x$, $C_1=(S_1)_x$. Since $C_0\cap C_1=\varnothing$, we obtain $C_0\subset(S_0)_x\backslash(S_1)_x\subset(B_0)_x$, $C_1\subset(S_1)_x\backslash(S_0)_x\subset(B_1)_x$. This yields that $(B_0)_x=C_0$ and $(B_1)_x=C_1$, since $C_0\cup C_1=\{0,1\}^\infty$ and $(B_0)_x\cap(B_1)_x=\varnothing$. The set $(B_0)_x$ is of class $\mathcal{B}_\tau$ in $\{0,1\}^\infty$, which contradicts our choice of C_0. Thus, we obtain two Souslin sets S_0 and S_1 in $\{0,1\}^\infty\times\{0,1\}^\infty$ such that $S_0\backslash S_1$ and $S_1\backslash S_0$ cannot be separated by Borel sets. By Theorem 6.6.12 there are coanalytic sets C_0 and C_1 such that $C_0\cap C_1=\varnothing$,

$$C_0\cup C_1=\{0,1\}^\infty\backslash(A_0\cap A_1),\ A_0\backslash A_1\subset C_0,\ A_1\backslash A_0\subset C_1.$$

As $A_0\backslash A_1$ and $A_1\backslash A_0$ cannot be separated by Borel sets, the same is true for the sets C_0 and C_1. Taking into account Theorem 6.7.3 and Theorem 6.8.6, we see that the last assertion of the theorem follows by the fact that any uncountable Polish space is Borel isomorphic to $\{0,1\}^\infty$. □

6.9. Measurable choice theorems

Let $F\colon X\to Y$ be some mapping. For every point $y\in F(X)$, we can pick an element $x=G(y)\in F^{-1}(y)$. Thus, we obtain a mapping G such that $F\circ G$ is the identity mapping on the range of F. The mapping G is called a selection or section of the mapping F or, alternatively, an inverse or implicit function $x=G(y)$ defined from the equation $y=F(x)$. However, in applications it is important to have a mapping G with certain additional properties. For instance, if F is continuous or Borel, it would be nice to preserve these properties for G. It is easy to give examples showing that even for one-to-one continuous mappings F the inverse may be discontinuous. We shall see below that for a Borel mapping F, one cannot always find a Borel mapping G. But it is remarkable that one can always take for G a mapping with nice measurability properties (a measurable selection). This is the content of the following Jankoff theorem, which belongs to the so-called measurable selection (or choice) theorems.

6.9.1. Theorem. *Let X and Y be Souslin spaces and let $F\colon X \to Y$ be a Borel mapping such that $F(X) = Y$. Then, one can find a mapping $G\colon Y \to X$ such that $F\bigl(G(y)\bigr) = y$ for all $y \in Y$ and G is measurable with respect to the σ-algebra generated by all Souslin subsets in Y. In addition, the set $G(Y)$ belongs to the σ-algebra $\sigma(\mathcal{S}_X)$ generated by Souslin sets in X.*

PROOF. Suppose first that F is continuous. Since X is the image of the space $\mathbb{N}^\infty$ under a continuous mapping p, it suffices to prove our claim for $\mathbb{N}^\infty$ and take for the required mapping the composition of p with the mapping obtained for $\mathbb{N}^\infty$. Thus, we may assume that $X = \mathbb{N}^\infty$. The set $\mathbb{N}^\infty$ is equipped with the lexicographic order: $(n_i) < (k_i)$ if either $n_1 < k_1$, or $n_1 = k_1, \ldots, n_m = k_m$ and $n_{m+1} < k_{m+1}$ for some $m \geq 1$. Let $x \leq z$ if $x < z$ or $x = z$. For every $y \in Y$, we take for $G(y)$ the smallest in the sense of the lexicographic order element of the set $F^{-1}(y)$ (which is nonempty by hypothesis and is closed by the continuity of F). Note that such an element exists. Indeed, let $F^{-1}(y)$ be denoted by Z. We take any element $x^1 = (x^1_i) \in Z$ such that $x^1_1 \leq z_1$ for all $z = (z_i) \in Z$. Next we find an element $x^2 = (x^2_i) \in Z$ such that $x^2_2 = x^1_2$ and $x^2_2 \leq z_2$ for all $z = (z_i) \in Z$ such that $z_1 = x^1_1$. Then we find an element $x^3 \in Z$ with $x^3_1 = x^2_1$, $x^3_2 = x^2_2$ and $x^3_3 \leq z_3$ for all $z = (z_i) \in Z$ with $z_i = x^2_i$ for $i = 1, 2$. By induction, we find elements $x^k \in Z$ with the following properties: $x^{k+1}_i = x^k_i$ if $i \leq k$ and $x^{k+1}_{k+1} \leq z_{k+1}$ for all $z = (z_i) \in Z$ such that $z_i = x^k_i$ for all $i \leq k$. Let us consider the element $x = (x^i_i)$. The sequence of elements x^k converges to x in $\mathbb{N}^\infty$. Since Z is closed, we have $x \in Z$. In addition, $x \leq z$ for all $z \in Z$. Indeed, otherwise for some k we would have $x_1 = z_1, \ldots, x_k = z_k$, $z_{k+1} < x_{k+1} = x^{k+1}_{k+1}$. Then $z_i = x^i_i$ if $i \leq k$, which leads to a contradiction with our choice of x_{k+1}.

By construction $F\bigl(G(y)\bigr) = y$. We verify that for any Borel set $B \subset \mathbb{N}^\infty$, the set $G^{-1}(B)$ is contained in the σ-algebra $\mathcal{A}$ generated by all Souslin subsets in Y. Since the family of all Borel sets B with this property is a σ-algebra, it suffices to consider closed sets of the form

$$B = \bigl\{(n_i) \in \mathbb{N}^\infty\colon\ (n_i) \leq (b_i)\bigr\},$$

where $b_i \in \mathbb{N}$ are fixed. It is easy to see that these sets generate $\mathcal{B}(\mathbb{N}^\infty)$. It is clear that $G^{-1}(B) = F(B)$. Indeed, if $G(y) \in B$, then $y \in F(B)$. If $y = F(\eta)$ with $\eta \in B$, then $G(y) \leq \eta$, hence $G(y) \in B$, i.e., one has $y \in G^{-1}(B)$. Since $F(B)$ is Souslin, the set $G^{-1}(B)$ belongs to the σ-algebra $\mathcal{A}$.

Let us consider the general case. Then the graph of the mapping F, i.e., the set $\Gamma := \bigl\{\bigl(x, F(x)\bigr),\, x \in X\bigr\}$ is a Souslin subset of the space $X \times Y$ (see Lemma 6.7.1). The projection $\pi_Y\colon\ \Gamma \to Y$ is continuous. By the above, there exists a measurable mapping $\Psi\colon\ (Y, \mathcal{A}) \to \ \bigl(\Gamma, \mathcal{B}(\Gamma)\bigr)$ with $\pi_Y \circ \Psi(y) = y$ for all $y \in Y$. Let $\pi_X\colon \Gamma \to X$ be the natural projection. Set $G = \pi_X \circ \Psi$. Then

$$F\bigl(G(y)\bigr) = F\Bigl(\pi_X\bigl(\Psi(y)\bigr)\Bigr) = \pi_Y\bigl(\Psi(y)\bigr) = y, \quad \forall\, y \in Y,$$

since $\Psi(y) = \bigl(x, F(x)\bigr)$, where $x = \pi_X\bigl(\Psi(y)\bigr)$ and $F(x) = \pi_Y\bigl(\Psi(y)\bigr)$. By the continuity of π_X and measurability of Ψ with respect to $\mathcal{A}$ we obtain that G is $\mathcal{A}$-measurable.

Let us show that $G(Y) \in \sigma(\mathcal{S}_X)$. Let $T(x) = G\bigl(F(x)\bigr)$. We have $G(Y) = \{x \in X\colon T(x) = x\}$. The set on the right is the intersection of the sets $\{f_n = f_n \circ T\}$, where $\{f_n\}$ is a countable family of Borel functions on X separating points. It remains to observe that the function $f_n \circ T$ is measurable with respect to the σ-algebra $\sigma(\mathcal{S}_X)$. This follows by the $\bigl(\sigma(\mathcal{S}_Y), \mathcal{B}(X)\bigr)$-measurability of G and the $\bigl(\sigma(\mathcal{S}_X), \sigma(\mathcal{S}_Y)\bigr)$-measurability of F (the latter is a consequence of the Borel measurability of F, see Theorem 6.7.3). □

Let us observe that the mapping G constructed in the proof in the case where $X = \mathbb{N}^\infty$ and F is continuous has the following property: the set $G(Y)$ is coanalytic, i.e., its complement is Souslin. Indeed, since $G(y)$ is the minimal element in $F^{-1}(y)$, the set $\mathbb{N}^\infty \backslash G(Y)$ is the projection of the set

$$B = \bigl\{(x, z) \in \mathbb{N}^\infty \times \mathbb{N}^\infty\colon\ F(x) = F(z), z < x\bigr\},$$

where the relation $z < x$ is understood in the sense of the lexicographic order. It is readily seen that B is a Borel set. We observe that $G(Y)$ is Souslin only if it is Borel. This is impossible for a non-Borel Souslin set $Y \subset [0, 1]$, since F is injective on $G(Y)$ and $Y = F\bigl(G(Y)\bigr)$. Hence our method may produce non-Souslin sets $G(Y)$. Below we give an example where there is no selection G at all such that $G(Y)$ is Souslin.

The given proof applies to a more general problem of selecting a single-valued branch of a multivalued mapping, which we now discuss.

Let X be some space and let $(\Omega, \mathcal{B})$ be a measurable space. Suppose $\Psi\colon \Omega \to 2^X$ is a mapping with values in the set of all nonempty subsets of X, i.e., $\Psi(\omega) \subset X$ and $\Psi(\omega) \neq \varnothing$ for all $\omega \in \Omega$. The graph of the multivalued mapping Ψ is the set $\Gamma_\Psi := \bigl\{(\omega, u) \in \Omega \times X\colon\ \omega \in \Omega,\ u \in \Psi(\omega)\bigr\}$.

Let $\pi_\Omega\colon \Omega \times X \to \Omega$ and $\pi_X\colon \Omega \times X \to X$ denote the natural projections. The graphs of multivalued mappings are precisely the sets $\Gamma \subset \Omega \times X$ with $\pi_\Omega(\Gamma) = \Omega$.

A selection of Ψ is a mapping $\zeta\colon \Omega \to X$ such that $\zeta(\omega)$ belongs to $\Psi(\omega)$ for all $\omega \in \Omega$.

A typical example of a multivalued mapping is the inverse to a mapping $F\colon X \to \Omega$, i.e., $\Psi(\omega) = F^{-1}(\omega)$. Certainly, in order that Ψ be everywhere defined, the equality $F(X) = \Omega$ is required. The method of proof of the previous theorem yields the following assertion (we do not explain the necessary changes in the reasoning because in Theorem 6.9.5 below we prove a more general fact).

6.9.2. Theorem. *Let Ω and X be Souslin spaces and let the graph of a mapping Ψ from Ω to the set of nonempty subsets of X be a Souslin* *(for example, Borel) set. Then, there exists a mapping* $f\colon \Omega \to X$ *that is measurable with respect to the* σ*-algebra* $\sigma(\mathcal{S}_\Omega)$ *generated by all Souslin sets in* Ω *and satisfies the relation* $f(\omega) \in \Psi(\omega)$ *for all* $\omega \in \Omega$*.*

Let us give a sufficient condition in order to have a Borel selection.

6.9.3. Theorem. *Let* X *be a complete separable metric space and let* Ψ *be a mapping on* $(\Omega, \mathcal{B})$ *with values in the set of nonempty closed subsets of* X*. Suppose that for every open set* $U \subset X$*, we have*

$$\widehat{\Psi}(U) := \bigl\{\omega\colon\ \Psi(\omega) \cap U \neq \varnothing\bigr\} \in \mathcal{B}.$$

Then Ψ *has a selection* ζ *that is measurable with respect to the pair of* σ*-algebras* $\mathcal{B}$ *and* $\mathcal{B}(X)$*.*

Proof. Let $\{x_n\}$ be any countable everywhere dense set in X. We define a mapping $\zeta_0\colon\ \Omega \to X$ as follows: $\zeta_0(\omega) = x_n$ if n is the smallest number with $\Psi(\omega) \cap B(x_n, 1) \neq \varnothing$, where $B(x, r)$ is the open ball of radius r with the center at x. It is clear that ζ_0 assumes countably many values and is $\mathcal{B}$-measurable, since

$$\zeta^{-1}(x_n) = \widehat{\Psi}\bigl(B(x_n, 1)\bigr) \backslash \bigcup_{m=1}^{n-1} \widehat{\Psi}\bigl(B(x_m, 1)\bigr).$$

Now we construct inductively $\mathcal{B}$-measurable mappings ζ_k with countably many values $\{x_n\}$ such that for all ω one has

$$\operatorname{dist}\bigl(\zeta_k(\omega), \zeta_{k+1}(\omega)\bigr) < 2^{-k+1}, \quad \operatorname{dist}\bigl(\zeta_k(\omega), \Psi(\omega)\bigr) < 2^{-k},$$

where dist denotes the distance in X. Suppose that ζ_k is already constructed. Let $\Omega_i = \zeta_k^{-1}(x_i)$. If $\omega \in \Omega_i$, then we have $\Psi(\omega) \cap B(x_i, 2^{-k}) \neq \varnothing$. Now we define ζ_{k+1} on Ω_i as follows: $\zeta_{k+1}(\omega) = x_n$ if n is the smallest number with

$$\Psi(\omega) \cap B(x_i, 2^{-k}) \cap B(x_n, 2^{-k-1}) \neq \varnothing.$$

As above, the mapping ζ_{k+1} is $\mathcal{B}$-measurable. In addition, we have the estimates $\operatorname{dist}\bigl(\zeta_{k+1}(\omega), \Psi(\omega)\bigr) < 2^{-k-1}$ and

$$\operatorname{dist}\bigl(\zeta_{k+1}(\omega), \zeta_k(\omega)\bigr) < 2^{-k} + 2^{-k-1} < 2^{-k+1}.$$

In particular, $\{\zeta_k(\omega)\}$ is a fundamental sequence; its limit we denote by $\zeta(\omega)$. It is clear that $\zeta(\omega) \in \Psi(\omega)$. Taking into account the $\mathcal{B}$-measurability of ζ, we see that ζ is as required. $\square$

It is clear that in this theorem the completeness of X can be replaced with the completeness of $\Psi(\omega)$. In fact, this reduces to the considered case if we take the completion of X.

6.9.4. Corollary. *In the situation of the above theorem, one can find a sequence of* $\mathcal{B}$*-measurable selections* ζ_n *such that for every* ω *the sequence* $\{\zeta_n(\omega)\}$ *is dense in* $\Psi(\omega)$*.*

PROOF. Let $\{x_n\}$ be an everywhere dense sequence in X. For every pair $(n,i) \in \mathbb{N}^2$, we set

$$\Psi_{ni}(\omega) = \Psi(\omega) \cap B(x_n, 2^{-i}) \quad \text{if } \omega \in \widehat{\Psi}\bigl(B(x_n, 2^{-i})\bigr)$$

and $\Psi_{ni}(\omega) = \Psi(\omega)$ otherwise. The multivalued mapping $\overline{\Psi_{ni}}$ associating to a point ω the closure of the set $\Psi_{ni}(\omega)$, takes the values in the family of complete subsets of X. For any open set $U \subset X$, we have

$$\begin{aligned}&\bigl\{\omega\colon\ \overline{\Psi_{ni}}(\omega) \cap U \neq \varnothing\bigr\} = \bigl\{\omega\colon\ \Psi_{ni}(\omega) \cap U \neq \varnothing\bigr\}\\ &= \widehat{\Psi}\Bigl(B(x_n, 2^{-i}) \cap U\Bigr) \bigcup \Bigl[\Bigl(\Omega \backslash \widehat{\Psi}\bigl(B(x_n, 2^{-i})\bigr)\Bigr) \cap \widehat{\Psi}(U)\Bigr] \in \mathcal{B}.\end{aligned}$$

By the above theorem, $\overline{\Psi_{ni}}$ has a $\mathcal{B}$-measurable selection ζ_{ni}. We verify that the closure of $\{\zeta_{ni}(\omega)\}$ is $\Psi(\omega)$. Let $x \in \Psi(\omega)$ and $\varepsilon > 0$. We pick i and n such that $2^{1-i} < \varepsilon$ and $\operatorname{dist}(x_n, x) < 2^{-i}$. Then $\omega \in \widehat{\Psi}\bigl(B(x_n, 2^{-i})\bigr)$ and $\zeta_{ni}(\omega)$ belongs to the closure of $B(x_n, 2^{-i})$. Hence $\operatorname{dist}\bigl(x, \zeta_{ni}(\omega)\bigr) \le 2^{1-i} < \varepsilon$. $\square$

6.9.5. Theorem. *Suppose that Ω and X are Souslin spaces and $\sigma(\mathcal{S}_\Omega)$ is the σ-algebra generated by all Souslin sets in Ω. Let the graph of a multivalued mapping Ψ from Ω to the set of nonempty subsets of X be a Souslin set in $\Omega \times X$. Then, there exists a sequence of selections ζ_n that are measurable as mappings from $\bigl(\Omega, \sigma(\mathcal{S}_\Omega)\bigr)$ to $\bigl(X, \mathcal{B}(X)\bigr)$, such that for every ω, the sequence $\{\zeta_n(\omega)\}$ is dense in the set $\Psi(\omega)$.*

PROOF. Denote by Γ the graph of Ψ. There exists a continuous mapping h from a complete separable metric space Z onto Γ. Denote by π the projection $\Gamma \to \Omega$, $(\omega, x) \mapsto \omega$. By the continuity of $\pi \circ h$, the multivalued mapping $\Phi = (\pi \circ h)^{-1}$ on Ω takes values in the set of nonempty closed subsets of Z. We observe that Φ has the closed graph Γ_Φ in $\Omega \times Z$ by the continuity of $\pi \circ h$. Therefore, for any open set $U \subset Z$, the set $\widehat{\Phi}(U)$ is Souslin in Ω since it coincides with the projection of $\Gamma_\Phi \cap (\Omega \times U)$ on Ω. Let us apply Corollary 6.9.4 to $\mathcal{B} = \sigma(\mathcal{S}_\Omega)$ and Φ and Z in place of Ψ and X. We obtain $\mathcal{B}$-measurable sections η_n of the mapping Φ such that the sequences $\{\eta_n(\omega)\}$ are dense in the sets $\Phi(\omega)$. For any ω, the point $h\bigl(\eta_n(\omega)\bigr)$ has the form $\bigl(\omega, \zeta_n(\omega)\bigr)$. The mappings ζ_n are as required. Indeed, the inclusion $\eta_n(\omega) \in (\pi \circ h)^{-1}(\omega)$ yields the equality $\omega = \pi \circ h \circ \eta_n(\omega)$, whence we obtain $h\bigl(\eta_n(\omega)\bigr) \in \{\omega\} \times \Psi(\omega)$, hence $\zeta_n(\omega) \in \Psi(\omega)$. The measurability of ζ_n with respect to $\sigma(\mathcal{S}_\Omega)$ is seen from the formula $\zeta_n = \pi_X \circ h \circ \eta_n$, where π_X is the projection to X. $\square$

6.9.6. Theorem. *Let X and Y be Polish spaces and let $\Gamma \in \mathcal{B}(X \times Y)$. Suppose, additionally, that the set $\Gamma_x := \{y \in Y\colon\ (x, y) \in \Gamma\}$ is nonempty and σ-compact for all $x \in X$. Then Γ contains the graph of some Borel mapping $f\colon\ X \to Y$.*

For a proof, see Kechris [**968**, §35] (see also Arsenin, Lyapunov [**72**, §15]). Interesting generalizations are obtained in Levin [**1165**]. Other sufficient conditions are given in Burgess [**282**].

An important partial case when there exists a Borel inverse mapping is that of a continuous mapping of a metrizable compact space. This follows by Theorem 6.9.6 or by Theorem 6.9.3, but we give a direct justification.

6.9.7. Theorem. *Let X be a compact metric space, let Y be a Hausdorff topological space, and let $f\colon X \to Y$ be a continuous mapping. Then, there exists a Borel set $B \subset X$ such that $f(B) = f(X)$ and f injective on B. In addition, the mapping $f^{-1}\colon f(X) \to B$ is Borel.*

PROOF. The set $f(X)$ is compact metrizable. Hence we may further assume that Y coincides with the metrizable compact $f(X)$. Suppose first that $X \subset [0,1]$. Set $g(y) = \inf\{x\colon f(x) = y\}$, $y \in f(X)$. The function g is Borel, since for every $c \in \mathbb{R}^1$, the set $\{y\colon g(y) \le c\}$ is closed. Indeed, let $g(y_n) \le c$ and let y be the limit of $\{y_n\}$. One can find $x_n \in X$ such that $f(x_n) = y_n$ and $x_n \le c + 1/n$. Passing to a subsequence we may assume that $\{x_n\}$ converges to some $x \in X$. Then $f(x) = y$ and $x \le c$, whence $g(y) \le c$. It is clear that $f\bigl(g(y)\bigr) = y$, hence the function g is injective, the set $B := g(Y)$ is Borel and $f(B) = Y$. Alternatively, one could observe that $B = X \backslash \bigcup_{n=1}^{\infty} B_n$, where

$$B_n := \{x \in X\colon \exists\, t \in X, f(t) = f(x), x - t \ge 1/n\},$$

and the sets B_n are closed by the continuity of f and the compactness of X. The mapping f on B is injective. In the general case, by Proposition 6.1.18, there exists a compact set $K \subset [0,1]$ such that $X = \varphi(K)$ for some continuous mapping φ. Let us apply the already proven assertion to the mapping $f \circ \varphi$ and find a Borel set $B_0 \subset [0,1]$ such that the mapping $f \circ \varphi$ is injective and $f\bigl(\varphi(B_0)\bigr) = f\bigl(\varphi(K)\bigr) = f(X)$. Then φ is injective on B_0 and hence the set $B := \varphi(B_0)$ is Borel in X. It is clear that f is injective on B. $\square$

The metrizability of X is essential even if $Y = [0,1]$: it suffices to consider the projection of the space "two arrows" (see Exercise 6.10.36). Certainly, in this theorem neither the compactness of X nor the continuity of f can be omitted. For example, if f is a continuous function on $[0,1]$ such that for some Borel set X, the set $f(X)$ is not Borel, then $f(X)$ cannot be the injective continuous image of a Borel set B. A similar example is constructed with a Borel function f on $[0,1]$ with non-Borel $f([0,1])$. P.S. Novikoff [**1383**] discovered that there might be no Borel selection even in the case where f is a Borel function such that $f([0,1]) = [0,1]$. A classical example (with the plane in place of the interval) can be found, e.g., in the book Lusin [**1209**, Ch. III, p. 220], and the next theorem contains its modification suggested by J. Saint Raymond.

6.9.8. Theorem. *There exists a continuous mapping F of $\mathbb{N}^\infty \times \{0,1\}$ on $\mathbb{N}^\infty$ such that no Souslin set is injectively mapped by F onto $\mathbb{N}^\infty$. In particular, there is no selection G with Souslin $G(\mathbb{N}^\infty)$, hence there is no Borel selection.*

Proof. By Theorem 6.8.11 there exist two disjoint sets C_0 and C_1 in $\mathbb{N}^\infty$ with Souslin complements A_0 and A_1 such that there is no Borel set separating C_0 and C_1. One can find continuous surjections $F_0\colon \mathbb{N}^\infty \to A_0$, $F_1\colon \mathbb{N}^\infty \to A_1$. Let $F\colon \mathbb{N}^\infty \times \{0,1\} \to \mathbb{N}^\infty$ be defined by $F(\nu,0) = F_0(\nu)$, $F(\nu,1) = F_1(\nu)$. We obtain a continuous surjection, since $A_0 \cup A_1 = \mathbb{N}^\infty$. Suppose there is a Souslin set $S \subset \mathbb{N}^\infty \times \{0,1\}$ on which F is injective and $F(S) = \mathbb{N}^\infty$. Let $S_i := \{\nu \in \mathbb{N}^\infty\colon (\nu,i) \in S\}$, $i = 0,1$. We observe that the sets $B_0 := F_0(S_0) = G^{-1}(\mathbb{N}^\infty \times \{0\})$ and $B_1 := F_0(S_1) = G^{-1}(\mathbb{N}^\infty \times \{1\})$ are Souslin and disjoint and their union is $\mathbb{N}^\infty$. Hence both sets are Borel. One has $B_i \subset A_i$. Hence $C_0 \subset B_1$, $C_1 \subset B_0$, which contradicts the fact that C_0 and C_1 cannot be separated by Borel sets. Since the image of $\mathbb{N}^\infty$ under an injective Borel mapping is a Borel set, there is no Borel selection. □

6.9.9. Corollary. *There exists a Borel function $f\colon [0,1] \to [0,1]$ with $f([0,1]) = [0,1]$ such that there is no Borel function $g\colon [0,1] \to [0,1]$ with $f\bigl(g(y)\bigr) = y$ for all $y \in [0,1]$. In particular, there is no Borel set in $[0,1]$ that would be injectively mapped by f onto $[0,1]$.*

6.9.10. Corollary. *There exists a continuous mapping $g\colon \mathbb{N}^\infty \to [0,1]$ with $g(\mathbb{N}^\infty) = [0,1]$ that has no Borel selections.*

Proof. Indeed, let Γ be the graph of the function f from Novikoff's example and let π be the projection operator of Γ to the axis of ordinates. Then Γ is a Borel set in $[0,1]^2$ and there exists a continuous mapping h from the space $\mathbb{N}^\infty$ onto Γ. The mapping $g := \pi \circ h$ is the required one. Indeed, if there exists a Borel set $B \subset \mathbb{N}^\infty$ that is injectively mapped by g onto $[0,1]$, then $B_0 := h(B)$ is Borel in Γ. The projection of B_0 on the axis of abscissas, denoted by B_1, is a Borel set as well (by the injectivity of the projection operator on Γ) and $f(B_1) = [0,1]$. The function f is injective on B_1 by the injectivity of π on B_0, which follows by the injectivity of g on B. □

The proof of the next measurable choice result can be found in Castaing, Valadier [**319**].

6.9.11. Theorem. *Let X be a complete separable metric space. Suppose that the graph of a mapping Ψ with values in the set of nonempty closed subsets of X belongs to $\mathcal{B}\otimes\mathcal{B}(X)$. Denote by $\widehat{\mathcal{B}}$ the intersection of the Lebesgue completions of $\mathcal{B}$ over all probability measures on $\mathcal{B}$. Then, there exists a sequence of selections ζ_n that are measurable as mappings from $(\Omega,\widehat{\mathcal{B}})$ to $\bigl(X,\mathcal{B}(X)\bigr)$, and for every ω, the sequence $\{\zeta_n(\omega)\}$ is dense in the set $\Psi(\omega)$.*

We now prove a useful result from Leese [**1143**].

6.9.12. Theorem. *Let $(\Omega,\mathcal{B})$ be a measurable space and let X be a Souslin space. Suppose that $A \in S\bigl(\mathcal{B}\otimes\mathcal{B}(X)\bigr)$. Then $\pi_\Omega(A) \in S(\mathcal{B})$ and there is a $\bigl(\sigma\bigl(S(\mathcal{B})\bigr),\mathcal{B}(X)\bigr)$-measurable mapping $\xi\colon \pi_\Omega(A) \to X$ whose graph is contained in A.*

PROOF. We have $\pi_\Omega(A)\in S(\mathcal{B})$ by Corollary 6.10.10 proven below, hence we may assume that $\pi_\Omega(A)=\Omega$. Let $\mathcal{J}:=\mathbb{N}^\infty$. The set A admits a Souslin representation $A=\bigcup_{\eta\in\mathcal{J}}\bigcap_{n=1}^\infty A_{\eta_1,\ldots,\eta_n}\times B_{\eta_1,\ldots,\eta_n}$, where $A_{\eta_1,\ldots,\eta_n}\in\mathcal{B}$ and $B_{\eta_1,\ldots,\eta_n}$ are closed in X (this follows by Exercise 6.10.69). Suppose first that $X=\mathcal{J}$. Let $A_\eta=\bigcap_{n=1}^\infty A_{\eta_1,\ldots,\eta_n}$, $B_\eta=\bigcap_{n=1}^\infty B_{\eta_1,\ldots,\eta_n}$. It is readily seen that A is the projection on $\Omega\times X$ of the set

$$E:=\bigcup_{\eta\in\mathcal{J}}\bigcap_{n=1}^\infty A_{\eta_1,\ldots,\eta_n}\times B_{\eta_1,\ldots,\eta_n}\times N_{\eta_1,\ldots,\eta_n}=\bigcup_{\eta\in\mathcal{J}}A_\eta\times B_\eta\times\{\eta\},$$

where $N_{\eta_1,\ldots,\eta_n}:=\{\nu\in\mathcal{J}\colon\ \nu_1=\eta_1,\ldots,\nu_n=\eta_n\}$, and $E\in S\bigl(\mathcal{B}\times\mathcal{B}(X\times\mathcal{J})\bigr)$. The sections E_ω, where $\omega\in\Omega$, are closed. Indeed, if $(x,\nu)\notin E_\omega$, then $(\omega,x)\notin A_\nu\times B_\nu$. Hence, for some n, either $\omega\notin A_{\nu_1,\ldots,\nu_n}$ or $x\notin B_{\nu_1,\ldots,\nu_n}$. In the first case $X\times N_{\nu_1,\ldots,\nu_n}$ is a neighborhood of (x,ν) disjoint with E_ω. In the second case x has a neighborhood U disjoint with $B_{\nu_1,\ldots,\nu_n}$, so $U\times N_{\nu_1,\ldots,\nu_n}$ is a neighborhood of (x,ν) disjoint with E_ω. Let $\Psi(\omega):=E_\omega$. For any open set U in $X\times\mathcal{J}$, we have $\widehat{\Psi}(U)=\pi_\Omega\bigl(E\cap(\Omega\times U)\bigr)$. Hence $\widehat{\Psi}(U)\in S(\mathcal{B})$. By Theorem 6.9.3 there is a $\bigl(\sigma\bigl(S(\mathcal{B})\bigr),\mathcal{B}(X\times\mathcal{J})\bigr)$-measurable mapping $\zeta\colon\ \Omega\to X\times\mathcal{J}$ whose graph belongs to E. It remains to set $\xi:=\pi_{\mathcal{J}}\circ\zeta$.

In the general case, there is a continuous surjection f from $\mathcal{J}$ onto X. Now we set $E:=\bigcup_{\eta\in\mathcal{J}}\bigl(A_\eta\times f^{-1}(B_\eta)\times\{\eta\}\bigr)$. It is clear that E belongs to $\mathcal{S}\bigl(\mathcal{B}\times\mathcal{B}(\mathcal{J}\times\mathcal{J})\bigr)$. Note that $\pi_\Omega(E)=\Omega$ as $\pi_\Omega(A)=\Omega$. By the first step we find a $\bigl(\sigma\bigl(S(\mathcal{B})\bigr),\mathcal{B}(\mathcal{J}\times\mathcal{J})\bigr)$-measurable mapping $\zeta=(\zeta_1,\zeta_2)\colon\ \Omega\to\mathcal{J}\times\mathcal{J}$ whose graph is contained in E. Finally, the mapping $\xi:=f\circ\zeta_1$ has the required properties. □

The next theorem from Aumann [**80**] and Sainte-Beuve [**1636**] gives measurable selections on measure spaces (it has a modification applicable to certain complete σ-algebras rather than measures; see the cited papers). Although this theorem follows directly from Theorem 6.9.12 and the relations $S(\mathcal{A})\subset\mathcal{A}_\mu=\mathcal{A}$, we give an independent proof.

6.9.13. Theorem. *Let $(\Omega,\mathcal{A},\mu)$ be a complete probability space, let X be a Souslin space, and let Ψ be a multivalued mapping from Ω to the set of nonempty subsets of X such that its graph Γ_Ψ belongs to $\mathcal{A}\otimes\mathcal{B}(X)$. Then, there exists an $\bigl(\mathcal{A},\mathcal{B}(X)\bigr)$-measurable mapping $f\colon\ \Omega\to X$ such that $f(\omega)\in\Psi(\omega)$ for all $\omega\in\Omega$.*

PROOF. Let us recall that there exist two sequences $\{A_n\}\subset\mathcal{A}$ and $\{B_n\}\subset\mathcal{B}(X)$ such that Γ_Ψ belongs to $\sigma(\{A_n\times B_n\})$, in particular, it belongs to $\mathcal{A}_0\otimes\mathcal{B}(X)$, where $\mathcal{A}_0$ is the σ-algebra generated by $\{A_n\}$. We know that there exists an $\mathcal{A}_0$-measurable function $h\colon\ \Omega\to[0,1]$ such that $\mathcal{A}_0=\bigl\{h^{-1}(B)\colon\ B\in\mathcal{B}([0,1])\bigr\}$. Thus, h gives a one-to-one mapping from $\mathcal{A}_0$ onto $\mathcal{B}(E)$, where $E:=h(\Omega)$. Hence the mapping $g\colon\ (\omega,x)\mapsto\bigl(h(\omega),x\bigr)$, $\Omega\times X\to E\times X$, takes $\mathcal{A}_0\otimes\mathcal{B}(X)$ to $\mathcal{B}(E)\otimes\mathcal{B}(X)$. In particular, we have $g(\Gamma_\Psi)\in\mathcal{B}(E)\otimes\mathcal{B}(X)$. The set $g(\Gamma_\Psi)$ is the graph of the multivalued mapping $\Phi\colon\ y\mapsto\bigcup_{\omega\in h^{-1}(y)}\Psi(\omega)$. Now it suffices to prove our claim for Φ and the probability space $\big(E, \mathcal{B}(E)_\nu, \nu\big)$, where $\nu := \mu\circ h^{-1}$. Indeed, if we have a ν-measurable mapping $f_1\colon E\to X$ with $f_1(y)\in\Phi(y)$, then there exists a set $B\in\mathcal{B}(E)$ with $\nu(B)=1$ on which f_1 is Borel. Then $h^{-1}(B)\in\mathcal{A}_0$, $\mu\big(h^{-1}(B)\big)=1$, and we can set $f(\omega) := f_1\big(h(\omega)\big)$ for all $\omega\in h^{-1}(B)$, and for all other points ω we can pick $f(\omega)\in\Psi(\omega)$ in an arbitrary way. Let us observe that $\Psi(\omega)=\Psi(\omega')$ if $h(\omega)=h(\omega')$ since $I_{\Gamma_\Psi}(\omega,x)=\varphi\big(h(\omega),x\big)$, where φ is a Borel function on $[0,1]\times X$. Hence $f(\omega)\in\Psi(\omega)$ for all $\omega\in\Omega$.

Finally, the claim for E follows by the already known results for Souslin spaces, since the graph of Φ is the intersection of $E\times X$ with some Borel set D in $[0,1]\times X$. The projection S of the set D on $[0,1]$ is a Souslin set and contains E. Hence it remains to extend ν to a Borel measure on S and take the multivalued mapping on S with the Souslin graph $(S\times X)\cap D$. □

Evstigneev [**545**] and Graf [**718**] obtained an analogous result in the case where X is compact and the graph of Ψ belongs to $S\big(\mathcal{A}\otimes\mathcal{B}a(X)\big)$. Another related result is given in Exercise 6.10.77.

We now discuss yet another aspect of measurable selections. Let $(E,\mathcal{E})$ be a measurable space and let R be an equivalence relation on E, i.e., R is a subset of E^2 that contains the diagonal, $(y,x)\in R$ whenever $(x,y)\in R$, and if $(x,y),(y,z)\in R$, then $(x,z)\in R$. A set S is called a section or selection of R if S meets every equivalence class in exactly one point. If the equivalence classes have a reasonable descriptive structure, one might ask whether there is a nice selection. However, the classical Vitali example, where the equivalence on $[0,1]$ is defined by setting $x\sim y$ if $x-y\in\mathbb{Q}$, shows that there might be no measurable section even if each equivalence class is countable. It turns out that the measurable structure of the factor-space E/R must be taken into account.

The following very general result is due to Hoffmann-Jørgensen [**841**].

Let $R(x)$ denote the equivalence class of x. For every $A\subset E$, let

$$R(A) := \{y\in E\colon\ \exists x\in A \text{ with } (x,y)\in R\}.$$

6.9.14. Theorem. *Let $\mathcal{E}^*$ be a class of subsets of E that contains $\mathcal{E}$ and is closed under countable unions and countable intersections. Suppose there is a Souslin scheme $\{A_{n_1,\dots,n_k}\}$ with values in $\mathcal{E}$ such that:*

(i) $E=\bigcup_{n=1}^\infty A_n$, $A_{n_1,\dots,n_k}=\bigcup_{n=1}^\infty A_{n_1,\dots,n_k,n}$,

(ii) *for every $x\in E$ and every $(n_i)\in\mathbb{N}^\infty$, the intersection of the sets $R(x)\cap A_{n_1,\dots,n_k}$ is a single point, provided that these sets are not empty,*

(iii) $R(A_{n_1,\dots,n_k})\in\mathcal{E}^*$.

Then R has a section S such that $E\backslash S\in\mathcal{E}^*$.

Proof. Let us define a Souslin scheme $\{H_{n_1,\dots,n_k}\}$ by induction as follows: $H_n=A_n\backslash\bigcup_{k=1}^{n-1}R(A_k)$ and

$$H_{n_1,\dots,n_k,n_{k+1}} = (A_{n_1,\dots,n_k,n_{k+1}}\cap H_{n_1,\dots,n_k})\backslash \bigcup_{j=1}^{n_{k+1}-1} R(A_{n_1,\dots,n_k,j}).$$

Now we set

$$S_k = \bigcup_{(n_1,\dots,n_k)\in\mathbb{N}^k} H_{n_1,\dots,n_k}, \quad S = \bigcap_{k=1}^{\infty} S_k$$

and show that S has the required properties. To see that $E\backslash S \in \mathcal{E}^*$, it suffices to show that $E\backslash H_{n_1,\dots,n_k} \in \mathcal{E}^*$. This is easily verified by induction due to the inclusion $E\backslash(A\backslash B) = (X\backslash A) \cup (A \cap B) \in \mathcal{E}^*$ for all $A \in \mathcal{E}$ and $B \in \mathcal{E}^*$, which holds, since $\mathcal{E} \subset \mathcal{E}^*$, $\mathcal{E}$ is stable under complementation and $\mathcal{E}^*$ is stable under finite intersections and unions. Let us show that S is a section. Let $x \in E$. There exists $m_1 := \min\{n\colon R(x) \cap A_n \neq \varnothing\}$. Then $R(x) \cap A_k = \varnothing$ if $k < m_1$. Hence $R(x) \cap R(A_k) = \varnothing$ for all $k < m_1$ and $R(x) \cap A_{m_1} = R(x) \cap H_{m_1} \neq \varnothing$. Therefore, $R(x) \subset R(A_{m_1})$ and $R(x) \cap H_n$ for all $n \neq m_1$. By using that $A_{m_1} = \bigcup_{n=1}^{\infty} A_{m_1,n}$, we find a number m_2 such that $R(x) \cap A_{m_1,m_2} = R(x) \cap H_{m_1,m_2} \neq \varnothing$ and $R(x) \cap H_{m_1,n} = \varnothing$ for all $n \neq m_2$. By induction we obtain a sequence $\{m_k\}$ such that

$$R(x) \cap A_{m_1,\dots,m_k} = R(x) \cap H_{m_1,\dots,m_k} \neq \varnothing \quad \text{and} \quad R(x) \cap H_{m_1,\dots,m_k,n} = \varnothing$$

whenever $n \neq m_{k+1}$. As $H_{n_1,\dots,n_{k+1}} \subset H_{n_1,\dots,n_k}$, one has $R(x) \cap H_{n_1,\dots,n_k} = \varnothing$ if $(n_1,\dots,n_k) \neq (m_1,\dots,m_k)$. On account of these relations we have the equality $S \cap R(x) = \bigcap_{k=1}^{\infty} A_{m_1,\dots,m_k} \cap R(X)$, which by property (ii) of the scheme $\{A_{n_1,\dots,n_k}\}$ yields that $S \cap R(x)$ consists of a single point. □

6.9.15. Example. Let E be a complete separable metric space and let $\mathcal{E}^* = \mathcal{S}_E$ be the class of all Souslin sets in E. One can find a Souslin scheme $\{A_{n_1,\dots,n_k}\}$ that consists of closed sets $A_{n_1,\dots,n_k}$ of diameter at most $1/k$ such that condition (i) in the theorem is fulfilled. Then condition (ii) is fulfilled too for any equivalence relation with closed equivalence classes. Hence in order to obtain a coanalytic section one has only to ensure condition (iii).

6.9.16. Corollary. *Let E be a regular Souslin space and let $\mathcal{E}^*$ be a class of subsets of E that contains all Souslin sets and is closed under countable unions and countable intersections. Suppose R is an equivalence relation on E such that each equivalence class is closed and $R(A) \in \mathcal{E}^*$ for each closed set A. Then R has a coanalytic section S.*

Proof. There is a continuous surjection $f\colon \mathbb{N}^\infty \to X$. One can find a Souslin scheme $\{Z_{n_1,\dots,n_k}\}$ in $\mathbb{N}^\infty$ that consists of closed sets $Z_{n_1,\dots,n_k}$ of diameter at most $1/k$ such that condition (i) in the theorem is fulfilled. Let $A_{n_1,\dots,n_k} := \overline{f(Z_{n_1,\dots,n_k})}$. The Souslin scheme $\{A_{n_1,\dots,n_k}\}$ satisfies condition (i) in the theorem. By our assumption, $R(A_{n_1,\dots,n_k}) \in \mathcal{E}^*$. Let us verify condition (ii). Let $x \in X$ and $(n_i) \in \mathbb{N}^\infty$ be such that the sets $R(x) \cap A_{n_1,\dots,n_k}$ are not empty. Hence $Z_{n_1,\dots,n_k} \neq \varnothing$ and there is a unique element ν in $\bigcap_{k=1}^{\infty} Z_{n_1,\dots,n_k}$. We show that $f(\nu) \in R(x)$. Suppose not. Since X is regular, one can find disjoint open sets V and W such that $f(\nu) \in V$ and $R(x) \subset W$. By the continuity of f one has an open ball U containing ν with $f(U) \subset V$. There is a sufficiently large number k such that $f(Z_{n_1,\dots,n_k}) \subset V$, hence $A_{n_1,\dots,n_k}$ is contained in the complement of W and does not meet $R(x)$, a contradiction. It is seen from the same reasoning that $f(\nu)$ is a unique element of $\bigcap_{k=1}^{\infty} A_{n_1,\ldots,n_k}$. Indeed, if y is another element of this set, we find open sets V and W such that $f(\nu) \in V$, $y \in W$, $\overline{V} \cap \overline{W} = \varnothing$, which leads to a contradiction by the above reasoning. $\square$

It is worth noting that if we omit the regularity assumption on X, but require that the sets $R(A)$ be Souslin for all Souslin sets $A \subset X$, the above proof shows that there is a selection S that belongs to the σ-algebra $\sigma(\mathcal{S}_X)$. Indeed, it suffices to take $\mathcal{E}^* = \sigma(\mathcal{S}_X)$ and $A_{n_1,\ldots,n_k} = f(Z_{n_1,\ldots,n_k}) \in \mathcal{S}_X$.

6.9.17. Corollary. *Let X be a regular Souslin space, let Y be a Hausdorff space, and let $F\colon X \to Y$ be a continuous surjection. Then there exists a coanalytic set $S \subset X$ that is mapped by F one-to-one onto Y. If X is not regular, then S can be found in $\sigma(\mathcal{S}_X)$.*

PROOF. Let $(x, y) \in R$ if $F(x) = F(y)$. Then the equivalence classes are closed. In addition, $R(A) = F^{-1}\bigl(F(A)\bigr)$ is a Souslin set for every Souslin set $A \subset X$. Hence the previous corollary applies. If X is not regular, then we use the observation made above. $\square$

Under stronger assumptions one can find a Borel section.

6.9.18. Corollary. *Let R be an equivalence relation on a topological space X with closed equivalence classes. Then R admits a Borel section under any of the following conditions*:

(i) *the space X is Polish and $R(U) \in \mathcal{B}(X)$ for every open set U* (*or $R(Z) \in \mathcal{B}(X)$ for every closed set Z*);

(ii) *the space X is Lusin and $R(B) \in \mathcal{B}(X)$ for every Borel set B.*

PROOF. (i) We may assume that X is a complete separable metric space and apply the theorem to $\mathcal{E} = \mathcal{E}^* = \mathcal{B}(X)$ and the same Souslin scheme as in Example 6.9.15. (ii) By hypothesis, there is a one-to-one continuous mapping f of a complete separable metric space E onto X. Let us set $\mathcal{E} = \mathcal{E}^* = \mathcal{B}(X)$ and apply the theorem to the Souslin scheme $\{f(A_{n_1,\ldots,n_k})\}$ with $A_{n_1,\ldots,n_k}$ from Example 6.9.15. $\square$

Additional information can be found in Burgess [**280**], [**281**].

6.10. Supplements and exercises

6.10(i). Borel and Baire sets

We note that apart from the σ-algebra $\sigma(\mathcal{F})$ generated by a class of sets $\mathcal{F}$ in a space X, one can consider the smallest class of sets that contains $\mathcal{F}$ and is closed with respect to countable unions and countable intersections (but may not be closed with respect to complementation). This class is denoted by $\mathbb{B}(\mathcal{F})$. The class $\mathbb{B}(\mathcal{F})$ can be smaller than $\sigma(\mathcal{F})$: for example, the class of all Souslin subsets of the interval is closed with respect to countable unions and countable intersections, but is not closed with respect to complementation; the same is true for the class of at most countable subsets of the interval. Certain sufficient conditions for the equality $\mathbb{B}(\mathcal{F}) = \sigma(\mathcal{F})$ can be found in Exercise 6.10.32 and Jayne [**887**].

We know that the Borel σ-algebra of any subspace consists of the intersections of that subspace with Borel sets of the whole space. The situation with the Baire structure is different.

6.10.1. Example. There exist a completely regular space X, its closed Baire subset X_0, and a Baire subset B of X_0 (with the induced topology) such that B cannot be the intersection of a Baire set in X with X_0. Moreover, one can take for X_0 a functionally closed set in X.

PROOF. Let X be the Sorgenfrey plane (see Example 6.1.19) and let X_0 be the straight line in the plane given by the equation $x + y = 0$. Obviously, X_0 is a functionally closed subset of X, since the function $(x, y) \mapsto x + y$ is continuous on X. For any real number x, the open set $[x, x+1)\times[-x, -x+1)$ meets X_0 precisely at the point $(x, -x) \in X_0$. Thus, every point in X_0 is open in the induced topology, hence so is every subset of X_0. Therefore, all subsets of X_0 are Baire from the point of view of this subspace. It remains to observe that X is separable, hence has only the continuum of Baire sets (any continuous function is uniquely determined by its values on a countable everywhere dense set), whence we obtain the existence of a subset B in X_0 that is not Baire in X. In Exercise 6.10.81 it is proposed to verify that the intersections of X_0 with Baire subsets of X are Borel sets with respect to the usual topology of the plane. □

The following result is partially inverse to Proposition 6.3.4 (see Halmos [**779**], Ross, Stromberg [**1612**] for a proof).

6.10.2. Theorem. *If X is compact, then $\mathcal{B}(X) = \mathcal{B}a(X)$ precisely when X is perfectly normal.*

Recall that βX is the Stone–Čech compactification of a completely regular space X.

6.10.3. Theorem. (i) *Let X be completely regular and $X \in \mathcal{B}a(\beta X)$. Then, every closed Baire set in X is functionally closed.* (ii) *Any compact Baire set in a completely regular space is functionally closed.* (iii) *Let X be compact and let $B \in \mathcal{B}a(X)$. If $A \subset B$ and $A \in \mathcal{B}a(B)$, then $A \in \mathcal{B}a(X)$.*

For proofs and references, see Comfort, Negrepontis [**365**]. In applications, one also encounters spaces with distinct families of Borel and Baire sets.

6.10.4. Example. Suppose that X is any of the following spaces:

(i) an uncountable product of compact intervals (which is a compact space),

(ii) the space of all functions on an interval with the topology of pointwise convergence (i.e., the product $\mathbb{R}^{\mathfrak{c}}$ of the continuum of real lines),

(iii) the subspace in $\mathbb{R}^{\mathfrak{c}}$ consisting of all bounded functions. Then $\mathcal{B}a(X)$ is strictly smaller than $\mathcal{B}(X)$.

For the proof it suffices to use the following important result (going back to M.F. Bokshtein, see Engelking [**532**, 2.7.12(c)]) that describes the structure of Baire sets in product spaces.

6.10.5. Theorem. *Suppose that $(X_t)_{t\in T}$ is a family of separable spaces and Y is a separable metric space. Then, for every continuous mapping $F\colon \prod_{t\in T} X_t \to Y$, there exist a finite or countable set $S\subset T$ and a continuous mapping $F_0\colon \prod_{s\in S} X_s \to Y$ such that $F = F_0\circ\pi_S$, where π_S denotes the natural projection from $\prod_{t\in T} X_t$ to $\prod_{s\in S} X_s$. In particular, $\mathcal{B}a(\prod_{t\in T} X_t)$ is generated by the coordinate mappings to the spaces $\bigl(X_t, \mathcal{B}a(X_t)\bigr)$.*

The Baire σ-algebra can be generated by a family of functions that is much smaller than the whole class $C(X)$. We have already seen this in Proposition 6.5.4. The following result (which also follows from Bokshtein's theorem) was obtained in Edgar [**513**], [**514**]. Its proof can be found in Exercise 6.10.67. The definition of the weak topology is given in §4.7(ii).

6.10.6. Theorem. *Let X be a locally convex space equipped with the weak topology $\sigma(X,X^*)$. Then the corresponding Baire σ-algebra coincides with the σ-algebra $\sigma(X^*)$ generated by X^*. In particular, the Baire σ-algebra of any product of real lines $\mathbb{R}^\Lambda$ coincides with the σ-algebra generated by the coordinate functions.*

The following result from Kellerer [**974**] gives some information on the behavior of the Borel and Baire structures under multiplication of topological spaces.

6.10.7. Proposition. *Let (X_α), $\alpha\in A$, be a family of nonempty spaces, $X=\prod_\alpha X_\alpha$. The equality $\mathcal{B}a(X) = \bigotimes_\alpha \mathcal{B}a(X_\alpha)$ holds in any of the following cases*:

(a) *every finite subproduct of the spaces X_α is Lindelöf* (*for example, every X_α is either compact or separable metric*);

(b) *$A=\{1,2\}$ and at least one of the spaces X_1 and X_2 is separable metric*;

(c) *$A=\{1,2\}$, the space X_1 is locally compact and σ-compact and X_2 is separable.*

On the other hand, there exist a discrete space X_1 and a separable compact space X_2 such that $\mathcal{B}a(X_1\times X_2)\neq \mathcal{B}a(X_1)\bigotimes \mathcal{B}a(X_2)$.

It is unknown whether the equality $\mathcal{B}a(X\times Y) = \mathcal{B}a(X)\otimes\mathcal{B}a(Y)$ is true for all separable spaces.

Now we prove a useful result due to V.V. Sazonov.

6.10.8. Proposition. *Let X be a σ-compact topological space and let Γ be a family of continuous functions separating the points in X. Then the equality $\mathcal{B}a(X) = \sigma(\Gamma)$ holds.*

PROOF. We verify that $\mathcal{B}a(X) \subset \sigma(\Gamma)$. One can assume that Γ is an algebra of functions, passing to the algebra generated by the family Γ. Let $f \in C(X)$. It is easy to see that by the σ-compactness of X and the Weierstrass theorem, there exists a sequence of functions $f_n \in \Gamma$ such that $f(x) = \lim\limits_{n\to\infty} f_n(x)$ for every $x \in X$. Thus, the function f is measurable with respect to $\sigma(\Gamma)$. □

In diverse problems, some other σ-algebras of subsets in a topological space X may be useful. Let us mention some of them: the σ-algebra $\sigma_{\mathcal{K}}(X)$ generated by all compact subsets of X, the σ-algebra $\sigma_{G_\delta}(X)$ generated by all closed G_δ-sets in X, the σ-algebra $\sigma_B(X)$ generated by all balls in a metric space X. A simple example of a metric space X with distinct σ-algebras $\mathcal{B}(X)$ and $\sigma_B(X)$ is any uncountable discrete space in which the balls are singletons and the whole space (e.g., let all nonzero mutual distances equal 1). Then $\sigma_B(X)$ coincides with the σ-algebra of all sets that are either at most countable or have at most countable complements.

There exists a Banach space X with $\mathcal{B}(X) \neq \sigma_B(X)$ (see Fremlin [**624**]). On the other hand, there exists a nonseparable metric space for which one has $\mathcal{B}(X) = \sigma_B(X)$ (see Exercise 6.10.44).

Some additional information is given in Hoffmann-Jørgensen [**841**], [**845**], [**847**], Jayne [**887**], Kharazishvili [**988**], Mauldin [**1274**], [**1277**].

6.10(ii). Souslin sets as projections

The following theorem shows how to define Souslin sets without the Souslin operation. We recall that the symbols $\mathcal{E}_\sigma$, $\mathcal{E}_\delta$, $\mathcal{E}_{\sigma\delta}$ denote, respectively, the classes of countable unions, countable intersections, and countable intersections of countable unions of elements in the class $\mathcal{E}$. Let $\mathcal{N}$ denote the class of all cylinders in $\mathbb{N}^\infty$, i.e., the class of all sets of the form

$$C(p_1, \dots, p_k) = \{(n_i) \in \mathbb{N}^\infty \colon\ n_1 = p_1, \dots, n_k = p_k\}.$$

Given two classes of sets $\mathcal{E}$ and $\mathcal{F}$ in spaces X and Y, let

$$\mathcal{E}\times\mathcal{F} := \{E\times F \subset X\times Y \colon\ E \in \mathcal{E}, F \in \mathcal{F}\}.$$

Let $S(\mathcal{E})$ denote the class of all sets obtained by the Souslin operation on sets in $\mathcal{E}$.

6.10.9. Theorem. *Suppose that a class $\mathcal{E}$ of subsets of a nonempty set X contains the empty set. Then, the following conditions for a set $A \subset X$ are equivalent:*

(i) $A \in S(\mathcal{E})$;

(ii) *A is the projection on X of an $(\mathcal{E}\times\mathcal{N})_{\sigma\delta}$-set in the space $X\times\mathbb{N}^\infty$;*

(iii) *there exists a space* Y *with a compact class of subsets* $\mathcal{K}$ *such that* A *is the projection on* X *of an* $(\mathcal{E}\times\mathcal{K})_{\sigma\delta}$*-set in* $X\times Y$;

(iv) *there exists a space* Y *with a compact class of subsets* $\mathcal{K}$ *such that* A *is the projection on* X *of a set in* $X\times Y$ *belonging to* $S(\mathcal{E}\times\mathcal{K})$.

(v) *there exists a Souslin space* Y *such that* A *is the projection on* X *of a set in* $X\times Y$ *belonging to the class* $S(\mathcal{E}\times\mathcal{S}_Y)$, *where* $\mathcal{S}_Y$ *is the class of all Souslin sets in* Y.

PROOF. Let (i) be fulfilled. There exist $A(n_1,\ldots,n_k)\in\mathcal{E}$ such that

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

Let us consider the set

$$C=\bigcap_{k=1}^\infty\bigcup_{(n_1,\ldots,n_k)\in\mathbb{N}^k}A(n_1,\ldots,n_k)\times C(n_1,\ldots,n_k).$$

It is clear that $C\in(\mathcal{E}\times\mathcal{N})_{\sigma\delta}$. We show that A is the projection of C on X. Indeed, x belongs to the projection of C precisely when there exists $\eta=(\eta_j)$ in $\mathbb{N}^\infty$ with $(x,\eta)\in C$, i.e., when for every k, there exists $\sigma^k=(n^k_j)\in\mathbb{N}^\infty$ such that $x\in A(n^k_1,\ldots,n^k_k)$ and $\eta_j=n^k_j$ for all $j=1,\ldots,k$. The latter is equivalent to that $x\in A(\eta_1,\ldots,\eta_k)$ for all k, which proves our claim about the projection of C. Hence (i) yields (ii).

We recall that $\mathcal{N}$ is a compact class (see Lemma 3.5.3). Hence (ii) implies (iii), whence condition (iv) follows at once because $(\mathcal{E}\times\mathcal{K})_{\sigma\delta}\subset S(\mathcal{E}\times\mathcal{K})$.

Let (iv) be fulfilled. Suppose first that A is the projection of some set B in $(\mathcal{E}\times\mathcal{K})_{\sigma\delta}$, i.e., we derive (i) from (iii). We have

$$B=\bigcap_{k=1}^\infty\bigcup_{n=1}^\infty A_{kn}\times B_{kn},\quad A_{kn}\in\mathcal{E},\ B_{kn}\in\mathcal{K}.$$

Set $A(n_1,\ldots,n_k)=\bigcap_{j=1}^k A_{jn_j}$, $B(n_1,\ldots,n_k)=\bigcap_{j=1}^k B_{jn_j}$. Then a standard argument shows that

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

Let us introduce the table of sets $A'(n_1,\ldots,n_k)$ that coincide with the sets $A(n_1,\ldots,n_k)$ if $B(n_1,\ldots,n_k)\neq\varnothing$ and are empty otherwise. This is possible since the empty set belongs to $\mathcal{E}$ and the class $S(\mathcal{E})$ admits finite intersections, so that $A'(n_1,\ldots,n_k)$ belongs to $S(\mathcal{E})$. For completing the proof in the case under consideration it remains to verify that

$$A=\pi_X(B)\in S\bigl(\{A'(n_1,\ldots,n_k)\}\bigr).\tag{6.10.1}$$

The first equality is the definition of A. For the proof of the second one we have to show that for every fixed sequence (n_i) we have the equalities

$$\pi_X\Bigl(\bigcap_{k=1}^{\infty} A(n_1,\ldots,n_k)\times B(n_1,\ldots,n_k)\Bigr) \tag{6.10.2}$$
$$= \bigcap_{k=1}^{\infty} \pi_X\bigl(A(n_1,\ldots,n_k)\times B(n_1,\ldots,n_k)\bigr) = \bigcap_{k=1}^{\infty} A'(n_1,\ldots,n_k).$$

The second equality in (6.10.2) is obvious. The left-hand side of (6.10.2) belongs to the right-hand side. Suppose that a point x belongs to the projection of every set $A(n_1,\ldots,n_k)\times B(n_1,\ldots,n_k)$. Then the sets

$$\bigl(\{x\}\times Y\bigr)\cap\Bigl(\bigcap_{k=1}^{m} A(n_1,\ldots,n_k)\times B(n_1,\ldots,n_k)\Bigr)$$

are nonempty. Since the classes $\mathcal{K}$ and $\mathcal{N}$ are compact, it follows by Proposition 1.12.4 that $\bigl(\{x\}\times Y\bigr)\cap\bigl(\bigcap_{k=1}^{\infty} A(n_1,\ldots,n_k)\times B(n_1,\ldots,n_k)\bigr)\neq\varnothing$. It is clear that the projection of any element in this set is x. Thus, we have proved (6.10.2), hence (6.10.1).

Now let A be the projection of $B\in S(\mathcal{E}\times\mathcal{K})$. According to what has already been proved, B is the projection on $X\times Y$ of some $(\mathcal{E}\times\mathcal{K}\times\mathcal{N})_{\sigma\delta}$-set $C\subset X\times Y\times\mathbb{N}^\infty$. The class $\mathcal{H}:=\mathcal{K}\times\mathcal{N}$ is compact by Lemma 3.5.3. Therefore, A is the projection of an $(\mathcal{E}\times\mathcal{H})_{\sigma\delta}$-set in the space $X\times(Y\times\mathbb{N}^\infty)$ and by the above we have $A\in S(\mathcal{E})$. Thus, (iv) implies (i), hence (i)–(iv) are equivalent.

It is clear that (v) follows from (ii). Finally, let (v) be fulfilled. According to Theorem 6.7.4, the space Y is Borel isomorphic to a Souslin subset of the interval $[0,1]$. This isomorphism also identifies the classes of Souslin sets. For this reason, we may assume from the very beginning that Y is a Souslin set in $[0,1]$. Then

$$\mathcal{E}\times\mathcal{S}_Y\subset\mathcal{E}\times\mathcal{S}_{[0,1]}\subset\mathcal{E}\times S(\mathcal{K})\subset S(\mathcal{E}\times\mathcal{K}),$$

where $\mathcal{K}$ is the class of all compact sets in $[0,1]$. Hence (iv) is fulfilled. □

6.10.10. Corollary. *Let $\mathcal{E}$ be a σ-algebra of subsets of a space X and let Y be a Souslin space. Then the projection on X of any set $M\in S\bigl(\mathcal{E}\otimes\mathcal{B}(Y)\bigr)$ belongs to $S(\mathcal{E})$. If the graph of $f\colon X\to Y$ belongs to $S\bigl(\mathcal{E}\otimes\mathcal{B}(Y)\bigr)$, then f is measurable with respect to $\bigl(\sigma\bigl(S(\mathcal{E})\bigr),\mathcal{B}(Y)\bigr)$, in particular, f is measurable with respect to every measure on $\mathcal{E}$.*

PROOF. We have $S\bigl(\mathcal{E}\otimes\mathcal{B}(Y)\bigr)=S\bigl(\mathcal{E}\times\mathcal{B}(Y)\bigr)$ by Exercise 6.10.69. If $B\in\mathcal{B}(Y)$, then $f^{-1}(B)=\pi_X\bigl(\Gamma_f\cap(X\times B)\bigr)\in S(\mathcal{E})$. □

Let us consider an application to hitting times of random processes.

6.10.11. Example. Suppose that $(\Omega,\mathcal{F},P)$ is a probability space. Let us set $T=[0,+\infty)$ and let $\mathcal{B}=\mathcal{B}(T)$. Given any set $A\in T\times\Omega$, let

$$h_A(\omega)=\inf\bigl\{t\geq 0\colon\ (t,\omega)\in A\bigr\},$$

where $h(\omega) = +\infty$ if $(t,\omega) \in A$ for no t. If $A \in \mathcal{S}(\mathcal{B}\otimes\mathcal{F})$, then h_A is $\sigma\bigl(\mathcal{S}(\mathcal{F})\bigr)$-measurable, hence is P-measurable. Indeed, for every $c > 0$, the set $\{h_A < c\}$ is the projection of the set $\bigl([0,c)\times\Omega\bigr) \cap A \in \mathcal{S}(\mathcal{B}\otimes\mathcal{F})$.

In particular, if a mapping ξ from $T\times\Omega$ to a measurable space $(E,\mathcal{E})$ is $(\mathcal{B}\otimes\mathcal{F},\mathcal{E})$-measurable, then, for every set $A \in E$, the mapping h defined by $h(\omega) = \inf\{t \geq 0\colon\ \xi(t,\omega) \in A\}$ is P-measurable.

6.10(iii). $\mathcal{K}$-analytic and $\mathcal{F}$-analytic sets

We recall that a multivalued mapping Ψ from a topological space X to the set of nonempty subsets of a topological space Y is called upper semicontinuous if for every $x \in X$ and every open set V in Y containing the set $\Psi(x)$, there exists a neighborhood U of the point x such that $\Psi(U) := \bigcup_{u\in U} \Psi(u) \subset V$.

6.10.12. Definition. *Let X be a Hausdorff space.* (i) *A set $A \subset X$ is called $\mathcal{K}$-analytic if there exists an upper semicontinuous mapping Ψ on $\mathbb{N}^\infty$ with values in the set of nonempty compact sets in X such that the equality $A = \bigcup_{\sigma\in\mathbb{N}^\infty} \Psi(\sigma)$ holds.*

(ii) *A set $A \subset X$ is called $\mathcal{F}$-analytic or $\mathcal{F}$-Souslin if it is obtained by means of the Souslin operation on closed sets in X.*

Jayne [**886**] proved (the proof can also be read in Rogers, Jayne [**1589**, §2.8]) that for a Hausdorff space X, the following conditions are equivalent:

(a) X is $\mathcal{K}$-analytic,

(b) X is a continuous image of a $F_{\sigma\delta}$-set in some compact space,

(c) X is a continuous image of a $K_{\sigma\delta}$-set (a countable intersection of countable unions of compact sets) in some Hausdorff space,

(d) X is a continuous image of a Lindelöf G_δ-set in some compact space.

The most important properties of $\mathcal{K}$-analytic spaces are listed in the following theorem. For a proof, see Rogers, Jayne [**1589**].

6.10.13. Theorem. (i) *Every $\mathcal{K}$-analytic set is $\mathcal{F}$-analytic and Lindelöf.*

(ii) *The class of all $\mathcal{K}$-analytic sets in a given space is closed with respect to the Souslin operation.*

(iii) *The image of any $\mathcal{K}$-analytic set under any upper semicontinuous multivalued mapping with values in the nonempty compact sets in a Hausdorff space is $\mathcal{K}$-analytic.*

(iv) *A set A in a Hausdorff space X is $\mathcal{K}$-analytic precisely when it is the projection of a closed $\mathcal{K}$-analytic set in $X\times\mathbb{N}^\infty$.*

(v) *In any Souslin space X, the classes of $\mathcal{K}$-analytic sets, $\mathcal{F}$-analytic sets, and Souslin sets coincide.*

It follows from (iii) that every Souslin set is $\mathcal{K}$-analytic. The class of $\mathcal{K}$-analytic sets is larger: for instance, any compact K is $\mathcal{K}$-analytic (as the image of $\mathbb{N}^\infty$ under the constant multivalued mapping $\Psi(\sigma) \equiv K$), but a nonmetrizable compact space is not Souslin. Although $\mathcal{K}$-analytic sets form a broader class than Souslin sets, they possess many nice properties of the latter. In particular, any finite Borel measure on such a space is tight (Exercise 7.14.125).

Let us observe that all equivalent descriptions of Souslin sets encountered in this book fall into the following two categories: (1) representations by means of the A-operation on certain classes of sets (intervals, closed sets, open sets, etc.) and (2) representations by means of images of nice spaces under certain classes of mappings, where one can vary source spaces (Polish spaces, the space of irrational numbers, subsets in certain product spaces, etc.) as well as the classes of mappings (continuous, Borel measurable, projections, etc.), in particular, such mappings can be single-valued or multivalued as in this subsection. Obviously, one can hardly list all possible alternate equivalent options. However, there is yet another approach not discussed in this book and going back to Lusin: (3) scrible representations. This approach is discussed in Kuratowski, Mostowski [**1083**], Lusin [**1209**].

6.10(iv). Blackwell spaces

6.10.14. Definition. *A measurable space $(X, \mathcal{A})$ is called a Blackwell space if the σ-algebra $\mathcal{A}$ is countably generated and contains all singletons and, in addition, has no proper sub-σ-algebras with these two properties.*

This interesting class of spaces was introduced in Blackwell [**180**] (without the requirement of separation of points, which is now usually included). Such spaces admit the following description (the proof is left as Exercise 6.10.64).

6.10.15. Theorem. *Let $(X, \mathcal{A})$ be a measurable space such that $\mathcal{A}$ is countably generated and contains all one-point sets. Then the following conditions are equivalent*:

(i) *$(X, \mathcal{A})$ is a Blackwell space*;

(ii) *every one-to-one $\mathcal{A}$-measurable mapping from X onto a measurable space $(Y, \mathcal{B})$, where the σ-algebra $\mathcal{B}$ is countably generated and contains all one-point sets, is an isomorphism*;

(iii) *every injective $\mathcal{A}$-measurable mapping f from X to a Polish space Y is an isomorphism between $(X, \mathcal{A})$ and $\bigl(f(X), \mathcal{B}\bigl(f(X)\bigr)\bigr)$.*

Some authors (see, e.g., Meyer [**1311**]) use another terminology, according to which the Blackwell spaces are isomorphic to Souslin subspaces of the real line (a different characterization of this class is given in Exercise 6.10.64). It is clear from Theorem 6.8.9 that such spaces are Blackwell in the sense of the above definition. However, the converse is false (see Orkin [**1403**], Rao, Rao [**1532**]). Thus, Blackwell spaces up to isomorphisms form some class of subspaces of the real line with the induced Borel σ-algebras and this class strictly contains the class of Souslin subspaces. It should be noted that a non-Souslin set complementary to a Souslin one may not be Blackwell (Exercise 6.10.65). It is consistent with the standard axioms that the non-Borel coanalytic sets are not Blackwell spaces (see Orkin [**1403**], Rao, Rao [**1532**]). About Blackwell spaces, see also Shortt, Rao [**1704**].

Let us say that a measurable space $(X, \mathcal{A})$ has the Doob property if for every pair of measurable spaces $(E, \mathcal{E})$ and $(F, \mathcal{F})$ and every mapping f from E to F such that $\mathcal{E} = \{f^{-1}(B)\colon\ B \in \mathcal{F}\}$, every $(\mathcal{E}, \mathcal{A})$-measurable mapping from E to X has the form $h \circ f$, where $h\colon\ F \to X$ is measurable with respect to the pair $(\mathcal{F}, \mathcal{A})$. The space $\mathbb{R}^1$ with its Borel σ-algebra has the Doob property by Theorem 2.12.3 because $\mathcal{E} = \sigma(I_B \circ f\colon\ B \in \mathcal{F})$. Spaces with the Doob property are investigated in Pintacuda [**1459**], Pratelli [**1484**]. An example of a nonseparable space with this property is constructed in [**1484**]. However, if $\mathcal{A}$ is countably generated and the measurable space $(X, \mathcal{A})$ has the Doob property, then it is standard Borel and is Borel isomorphic either to $\mathbb{R}^1$ or to a set in $\mathbb{N}$.

6.10(v). Mappings of Souslin spaces

6.10.16. Lemma. *Let X be a Polish space, let Y be a metric space, and let $f\colon\ X \to Y$ be a Borel mapping. Then the set $f(X)$ is separable.*

Proof. Suppose that the set $f(X)$ is nonseparable. Then, there exists an uncountable set $S \subset f(X)$ all points of which have mutual distances greater than some $\varepsilon > 0$. If we show that S has cardinality of the continuum, then we obtain a contradiction with the fact noted in §6.7 that $\mathcal{B}(X)$ has cardinality at most of the continuum. Indeed, the cardinality of the set of all subsets of S is greater than that of the continuum. Then the same is true for the set of all sets $f^{-1}(E)$, $E \subset S$. All such sets belong to $\mathcal{B}(X)$, since every subset of S is closed. Now we show that S has cardinality of the continuum (it is clear that the cardinality of S is not greater than that of the continuum). To this end, we consider disjoint Borel sets $f^{-1}(s)$, $s \in S$, pick in each of them an arbitrary element z_s and define the mapping $g\colon\ X \to X$ as follows: $g(x) = z_s$ if $x \in f^{-1}(s)$, $g(x) = z$ if $x \notin f^{-1}(S)$, where $z \notin f^{-1}(S)$ is an arbitrary fixed element. Then g is a Borel mapping. Indeed, g is constant on the Borel set $X \backslash f^{-1}(S)$, and for any Borel set $B \subset f^{-1}(S)$, we have $g^{-1}(B) = f^{-1}(A)$, where $A = \{s \in S\colon\ z_s \in B\}$. Since A is closed (as is every set in S), one has $f^{-1}(A) \in \mathcal{B}(X)$. According to Corollary 6.7.13, the uncountable set $g(X)$ has cardinality of the continuum. Then S also does. □

6.10.17. Corollary. *Let f be a Borel mapping from a Souslin space X to a metric space Y. Then the set $f(X)$ is separable.*

Now we prove the following important result due to Lusin.

6.10.18. Theorem. *Suppose that X and Y are Souslin spaces and A is a Souslin set in $X \times Y$. Then the set $\{y \in Y\colon\ \mathrm{Card}\, A_y > \aleph_0\}$, where $A_y := \{x\colon\ (x, y) \in A\}$, is Souslin. In particular, if $f\colon\ X \to Y$ is a Borel mapping, then the set $\{y \in Y\colon\ \mathrm{Card}\, f^{-1}(y) > \aleph_0\}$ is Souslin.*

Proof. There exist a complete separable metric space M and a continuous mapping $\varphi = (\varphi_1, \varphi_2)$ from M onto A. For every $y \in Y$, the set

$$M(y) := \bigl\{z \in M\colon\ \varphi_2(z) = y\bigr\} \subset \varphi_1^{-1}(A_y)$$

is closed in M, hence is a complete separable metric space. Denote by D the subset in M^∞ consisting of all sequences without isolated points. According to Exercise 6.10.74, the set D is G_δ in M and hence is a Polish space. Note that the set A_y is uncountable precisely when there exists a sequence $\{x_k\} \in D$ with the following property: $\varphi_2(x_k) = y$ for all k and $\varphi_1(x_k) \neq \varphi_1(x_n)$ for all distinct k and n. Indeed, if such a sequence exists, then its closure is uncountable and belongs to A_y by the continuity of φ_2. Conversely, if A_y is uncountable, then by means of the axiom of choice we pick in M an uncountable set P that is mapped by φ one-to-one onto A_y. Let us delete from P all points each of which has a neighborhood meeting P at an at most countable set. We obtain an uncountable set $P_0 \subset P$ that contains a countable everywhere dense sequence $\{x_k\}$. It is clear that $\{x_k\}$ has no isolated points. Let us set

$$S = \bigcap_{k=1}^{\infty} \bigcup_{m=k+1}^{\infty} \Bigl\{ (\{x_i\}, y) \in D \times Y \colon\ \varphi_2(x_k) = y,\ \varphi_1(x_k) \neq \varphi_1(x_m) \Bigr\}.$$

It is readily seen that the set S is Borel in $D \times Y$ (all the intersected sets are Borel), hence is Souslin. Denote by π_Y the projection operator from $D \times Y$ to Y. Then by the above-mentioned characterization of uncountable A_y we obtain the equality $\{y \in Y \colon \operatorname{Card} A_y > \aleph_0\} = \pi_Y(S)$, which completes the proof. □

This theorem should be compared with Theorem 6.8.2 proved above.

6.10(vi). Measurability in normed spaces

There are many works devoted to the study of measurability in Banach spaces with the norm topology or with the weak topology. We recall that the weak topology of an infinite-dimensional Banach space X is not metrizable. Even a ball in a separable space may not be metrizable in the weak topology. For example, this is the case for balls in the space l^1 (Exercise 6.10.35). If X is separable and reflexive, then the closed balls in the weak topology are metrizable compact (the converse is true as well). If X is separable, then $\mathcal{B}(X)$ is generated by the half-spaces of the form $\bigl\{x \in X \colon\ l(x) < c\bigr\}$, $l \in X^*$, $c \in \mathbb{R}^1$. In the general case, this is not true. If X is nonseparable, then the operation of addition $X \times X \to X$ may fail to be measurable with respect to $\mathcal{B}(X) \otimes \mathcal{B}(X)$ and $\mathcal{B}(X)$. Talagrand [**1828**] proved that X is a measurable vector space, i.e., the operation $(t, x, y) \mapsto tx + y$, $\mathbb{R}^1 \times X \times X \to X$ is measurable with respect to $\mathcal{B}(\mathbb{R}^1) \otimes \mathcal{B}(X) \otimes \mathcal{B}(X)$ and $\mathcal{B}(X)$ precisely when $\mathcal{B}(X) \otimes \mathcal{B}(X) = \mathcal{B}(X \times X)$. In the same work, there is an example of a nonseparable Banach space X such that this equality is fulfilled. In addition, it is shown that the continuum hypothesis implies the measurability of the space l^∞ in the above sense. It is proved in Talagrand [**1827**] that in the space l^∞, the Borel σ-algebras corresponding to the weak topology and norm topology do not coincide. On measurability in Banach spaces, see Edgar [**513**], [**514**], Talagrand [**1834**].

6.10(vii). The Skorohod space

We consider an interesting class of spaces introduced by Skorohod [**1739**] and frequently used in the theory of random processes. Let E be a metric space with a metric ϱ. The Skorohod space $D_1(E)$ is the space of mappings $x\colon [0,1] \to E$ that are right continuous and have left limits for all $t > 0$, equipped with the metric

$$d(x,y) = \inf\Bigl\{\varepsilon > 0 \mid \ \exists\, h \in \Lambda[0,1]\colon \ |t - h(t)| \le \varepsilon, \ \varrho\bigl(x(t), y(h(t))\bigr) \le \varepsilon\Bigr\},$$

where $\Lambda[0,1]$ is the set of homeomorphisms h of the interval $[0,1]$ such that $h(0) = 0$, $h(1) = 1$. Similarly, one defines the Skorohod space of mappings with values in completely regular spaces (see Jakubowski [**878**]). If the space E is Polish, then so is $D_1(E)$ (the proof for $E = \mathbb{R}^1$ can be found in Billingsley [**169**]; in the general case the reasoning is similar). In the case of complete E, the space $D_1(E)$ is not always complete with respect to the metric d, but is complete with respect to the following metric that defines the same topology:

$$d_0(x,y) = \inf\Bigl\{\varepsilon > 0 \mid \ \exists\, h \in \Lambda[0,1]\colon \\ \sup_{t>s}\Bigl|\log \frac{h(t) - h(s)}{t - s}\Bigr| \le \varepsilon, \varrho\bigl(x(t), y(h(t))\bigr) \le \varepsilon\Bigr\}.$$

Similarly, one defines the Skorohod space $D(E)$ of mappings on the half-line. In the case $E = \mathbb{R}^1$, a detailed discussion of the Skorohod space can be found in Billingsley [**169**]. It is readily verified that for any separable metric space E, the Borel σ-algebra of $D_1(E)$ is generated by the mappings $x \mapsto x(t)$, $t \in [0,1]$. The analogous question for more general spaces is considered in Jakubowski [**878**] and Bogachev [**207**]. To these works and also to Lebedev [**1117**], Mitoma [**1322**], we refer for additional information on Skorohod spaces. The descriptive properties of Skorohod spaces turn out to be a subtle matter. We mention a result of Kolesnikov [**1017**].

6.10.19. Theorem. *Let E be a coanalytic set in a Polish space M. Then $D_1(E)$ is a coanalytic set in $D_1(M)$.*

As observed by Kolesnikov [**1017**], the space $D_1(\mathbb{Q})$ is not Souslin. In addition, he proved in the same work that under the assumption of the existence of nonmeasurable projections of coanalytic sets (which is consistent with the usual axioms), there exists a Souslin subset E of the interval such that the space $D_1(E)$ is not universally measurable in $D_1([0,1])$. The Skorohod space can be equipped with some other natural topologies different from those mentioned above. The role of Skorohod spaces in the theory of random processes is explained by the fact that many important random processes possess sample paths belonging to such spaces, so the distributions of these processes are naturally defined on Skorohod spaces.

Exercises

6.10.20.° Prove Lemma 6.1.1.

HINT: Let f be continuous at x, $x = \lim_\alpha x_\alpha$, and let W be a neighborhood of $f(x)$. We find a neighborhood U of x such that $f(U) \subset W$ and take α_0 such that $x_\alpha \in U$ for all α with $\alpha_0 \le \alpha$. Then $f(x_\alpha) \in W$. Conversely, if we have the indicated condition for nets and W is a neighborhood of the point $f(x)$, then we take for T the set of all neighborhoods of the point x equipped with the following order: $U \le V$ if $V \subset U$. Since the intersection of two neighborhoods is a neighborhood, we obtain a directed set. If we suppose that every neighborhood U of the point x contains a point x_U with $f(x_U) \notin W$, then we obtain the net $\{x_U\}_{U\in T}$ convergent to x, which contradicts our condition, since the net $f(x_U)$ does not converge to $f(x)$.

6.10.21.° Prove Lemma 6.1.5.

HINT: For every $x \in K$, there is a continuous function $f_x\colon X \to [0,1]$ with $f_x(x) = 1$ vanishing outside U. The open sets $\{y\colon f_x(y) > 1/2\}$ cover K, and one can find a finite subcover corresponding to some points $x_1, \ldots, x_n$. The function $g = (f_{x_1} + \cdots + f_{x_n})/n\colon X \to [0,1]$ vanishes outside U and is greater than $(2n)^{-1}$ on K. Now let $f = \psi \circ g$, where the function $\psi\colon [0,1] \to [0,1]$ is continuous, equals 1 on $[1/(2n), 1]$ and $\psi(0) = 0$.

6.10.22.° Let K be a compact set in a completely regular space X. (i) Prove that every continuous function f on K extends to a continuous function on X with the same maximum of the absolute value. (ii) Let f be a continuous mapping from K to a Fréchet space Y. Show that f extends to a continuous mapping on all of the space X with values in the closed convex envelope of $f(K)$.

HINT: (i) the set $\mathcal{F}$ of all continuous functions on K possessing bounded continuous extensions to X is a subalgebra in $C(K)$ and contains constants. This subalgebra separates the points of K by the complete regularity of X. By the Stone–Weierstrass theorem, there exists a sequence of functions $f_n \in \mathcal{F}$ uniformly convergent to f on K. We may assume that $|f_n(x) - f_{n+1}(x)| < 2^{-n}$ for all $x \in K$. By induction we find continuous functions g_n on X such that $|g_n(x) - g_{n+1}(x)| < 2^{-n}$ for all $x \in X$ and $g_n|_K = f_n|_K$. Since X is completely regular, there exists a continuous function $\zeta_1\colon X \to [0,1]$ equal to 1 on K and 0 outside the open set $V_1 := \{|f_1 - f_2| < 1/2\}$. Letting $g_1 := \zeta_1 f_1$, $g_2 := \zeta_1 f_2$, $f_n' := \zeta_1 f_n$, $n \ge 3$, we continue this process applied to the functions f_n'. The sequence $\{g_n\}$ converges uniformly on X and its limit on K is f. Thus, we obtain an extension of f to a bounded continuous function g on X. Now we obtain the equality $\max_X |g(x)| = \max_K |f(x)|$ by passing to the function $\theta \circ g$, where $\theta(t) = t$ if $t \in [-M, M]$, $\theta(t) = M$ if $t > M$, $\theta(t) = -M$ if $t < -M$. (ii) Since $f(K)$ is compact, its closed convex envelope V is compact as well. There is a sequence of functionals $l_n \in Y^*$ separating the points in V. The mapping $h = (l_n)\colon Y \to \mathbb{R}^\infty$ takes V to the convex compact set Q and is a homeomorphism on V. Hence it suffices to prove our assertion for $h \circ f$. Let us extend all functions $l_n \circ f$ to bounded continuous functions ψ_n on X and apply Dugundji's theorem, according to which there is a continuous mapping $g\colon \mathbb{R}^\infty \to Q$ that is identical on Q (see Engelking [**532**, 4.5.19]).

6.10.23.° Let X_t, $t \in T$, be an uncountable collection of metric spaces containing more than one point. Show that the topological product of X_t is not metrizable.

HINT: in a metrizable space, every point has a countable base of neighborhoods.

6.10.24.° Prove that a compact space X is metrizable precisely when there is a countable family of continuous functions f_n separating the points in X.

HINT: the necessity of this condition is obvious; for the proof of sufficiency we embed X into $\mathbb{R}^\infty$ by a continuous mapping $x \mapsto \bigl(f_n(x)\bigr)$, which gives a compact set in $\mathbb{R}^\infty$; then we verify that on this set the original topology coincides with the topology induced from $\mathbb{R}^\infty$.

6.10.25. Show that the Cantor set C is homeomorphic to $\{0,1\}^\infty$.

HINT: consider the mapping $h(x) = \sum_{n=1}^\infty 2x_n 3^{-n}$, $x = (x_n)$, $x_n \in \{0,1\}$; see Engelking [**532**, 3.1.28].

6.10.26. Let K be a nonempty compact set without isolated points. Prove that K can be continuously mapped onto $[0,1]$.

HINT: suppose not. Then no compact subset of K can be mapped continuously onto $[0,1]$, since otherwise such a mapping could be extended to all of K. There exists a nonconstant continuous function φ_1 on K with values in $[0,1]$. By our assumption, there exists $c_1 < c_2$ such that $K_{1,1} := \{\varphi_1 \le c_1\}$ and $K_{1,2} := \{\varphi_1 \ge c_2\}$ are nonempty and $[c_1, c_2]$ does not meet $\varphi_1(K)$. This enables us to find a continuous function f_1 with $f|_{K_{1,1}} = 0$, $f|_{K_{1,2}} = 1/2$. Applying this reasoning to $K_{1,1}$ and $K_{1,2}$ we obtain $K_{1,1} = K_{2,1} \cup K_{2,2}$, $K_2 = K_{2,3} \cup K_{2,4}$ with disjoint compact sets $K_{i,j}$. Take a continuous function f_2 assuming the values 0, 1/4, 1/2, 3/4 on $K_{2,1}$, $K_{2,2}$, $K_{2,3}$, $K_{2,4}$. By induction, we continue this process and find disjoint compact sets $K_{n,m}$, $m = 1, \ldots, 2^n$, such that $K_{n-1,1} = K_{n,1} \cup K_{n,2}$, $K_{n-1,2} = K_{n,3} \cup K_{n,4}$ and so on. Then we find a continuous function f_n that assumes the values $0, 2^{-n}, \ldots, 1-2^{-n}$ on the 2^n disjoint compact sets of the nth step. The obtained continuous functions converge uniformly to a function f whose range is $[0,1]$.

6.10.27. The space $\mathcal{D}(\mathbb{R}^1)$ is the set of all infinitely differentiable functions with compact support equipped with the locally convex topology τ generated by all norms of the form $p_{\{a_k\}}(\varphi) = \sum_{k=-\infty}^\infty a_k \max\bigl\{|\varphi^{(m)}(x)|\colon\ x \in [k, k+1], m \le a_k\bigr\}$, where one takes for $\{a_k\}$ all two-sided sequences of natural numbers. A sequence φ_j converges to φ in this topology if and only if the functions φ_j vanish outside some common interval and all the derivatives of φ_j converge uniformly to the corresponding derivatives of φ. This topology τ is the topology of the locally convex inductive limit of the sequence of spaces $\mathcal{D}_n$ consisting of smooth functions with support in $[-n, n]$ and equipped with the sequence of norms $\max|\varphi^{(m)}(t)|$. The space of all linear functions on $\mathcal{D}(\mathbb{R}^1)$ continuous in the topology τ is denoted by $\mathcal{D}'(\mathbb{R}^1)$ and is called the space of distributions (generalized functions). Similarly, one defines $\mathcal{D}(\mathbb{R}^d)$ and $\mathcal{D}'(\mathbb{R}^d)$.

(i) Prove that the topology τ is strictly weaker than the topology τ_1 on $\mathcal{D}(\mathbb{R}^1)$ in which the open sets are all those sets that give open intersections with all $\mathcal{D}_n$ (where $\mathcal{D}_n$ is given the above-mentioned topology generated by countably many norms). To this end, show that the quadratic form $F(\varphi) = \sum_{n=1}^\infty \varphi(n)\varphi^{(n)}(0)$ is discontinuous in the topology τ, but is continuous in τ_1.

(ii) Prove that the topology τ is strictly stronger than the topology τ_2 on $\mathcal{D}(\mathbb{R}^1)$ generated by the norms $p_\psi(\varphi) = \sup|\psi(x)\varphi^{(m)}(x)|$, where one takes all nonnegative integers m and positive locally bounded functions ψ. To this end, verify that the linear function $F(\varphi) = \sum_{n=1}^\infty \varphi^{(n)}(n)$ is continuous in the topology τ, but is discontinuous in the topology τ_2.

(iii) Prove that the space $\mathcal{D}(\mathbb{R}^1)$ with the topology τ of the inductive limit of the spaces $\mathcal{D}_n$ is not a k_R-space (X is called a k_R-space if for the continuity of a function on X, its continuity on all compact sets is sufficient).

It should be noted that in some textbooks of functional analysis the topologies τ_1 or τ_2 are mistakenly introduced as equal to τ. Fortunately, convergence of countable sequences in all the three topologies is the same.

6.10.28. Let X be a separable metric space and let $\mathcal{F}$ be some collection of Borel sets in X. Suppose that $r_n > 0$ are numbers decreasing to zero and that for every $x \in X$ and every $n \in \mathbb{N}$, there exists a set E in the σ-algebra generated by $\mathcal{F}$ such that $B(x, r_{n+1}) \subset E \subset B(x, r_n)$, where $B(x, r)$ is the open ball of radius r centered at x. Show that $\sigma(\mathcal{F}) = \mathcal{B}(X)$.

HINT: see Hoffmann-Jørgensen [**848**, 1.9].

6.10.29.° Show that the σ-algebra $\mathcal{E}$ generated by all one-point subsets of $\mathbb{R}$ is not countably generated. Deduce that not every sub-σ-algebra of $\mathcal{B}(\mathbb{R})$ is countably generated.

HINT: use that every set in $\mathcal{E}$ is either at most countable or its complement is at most countable.

6.10.30. Let $\mathcal{E}$ be the algebra of all finite unions of intervals (open, closed or semiclosed) in $[0, 1]$. By induction, we define classes of sets B_n, $n \in \mathbb{N}$, as follows: B_n is the collection of all countable intersections and countable unions of sets in B_{n-1}, $B_0 = \mathcal{E}$. Prove that $\bigcup_{n=0}^{\infty} B_n$ is not a σ-algebra, in particular, does not coincide with the Borel σ-algebra.

HINT: see assertion (vi) in the next exercise or Kuratowski [**1082**, §30, XIV], Rogers, Jayne [**1589**, §4.3].

6.10.31. Let $\mathcal{E}$ be a class of subsets of a space X with $\varnothing \in \mathcal{E}$. (i) Let Ω be the set of all finite or countable ordinal numbers. The classes $\mathcal{E}_\alpha$, $\alpha \in \Omega$, are defined by means of transfinite induction as follows: $\mathcal{E}_0 = \mathcal{E}$ and $\mathcal{E}_\alpha$ consists of all sets of the form $\bigcup_{n=1}^{\infty} A_n$, where $A_n \in \mathcal{E}_{\beta_n}$ with $\beta_n < \alpha$, and $X \backslash A$, where $A \in \mathcal{E}_\beta$ with $\beta < \alpha$. Show that $\sigma(\mathcal{E}) = \bigcup_{\alpha \in \Omega} \mathcal{E}_\alpha$.

(ii) Let $\mathcal{E}$ be an algebra of sets. For all $\alpha \in \Omega$ we define the classes $\mathcal{B}_\alpha$ as follows: $\mathcal{B}_\alpha$ consists of all countable unions and countable intersections of sets in $\mathcal{B}_\beta$ with $\beta < \alpha$, $\mathcal{B}_0 = \mathcal{E}$. Show that $\sigma(\mathcal{E}) = \bigcup_{\alpha \in \Omega} \mathcal{B}_\alpha$. Show that this may be false if $\mathcal{E}$ is not an algebra.

(iii) Prove that if the class $\mathcal{E}$ is infinite and its cardinality is not greater than that of the continuum, then the cardinality of $\sigma(\mathcal{E})$ equals the cardinality of the continuum.

There is another hierarchy of Borel classes $\widehat{\mathcal{B}}_\alpha$, $0 \le \alpha < \omega_1$, defined as follows. Given a topological space X, let $\mathcal{B}_0$ be the class of all open sets in X. If the ordinal α is even (limit ordinals count as even), let $\widehat{\mathcal{B}}_{\alpha+1}$ be the family of complements of sets in $\widehat{\mathcal{B}}_\alpha$. If α is odd, let $\widehat{\mathcal{B}}_{\alpha+1}$ be the family of countable unions of sets in $\widehat{\mathcal{B}}_\alpha$. If α is a limit ordinal, let $\widehat{\mathcal{B}}_\alpha$ be the family of countable unions of sets chosen from the families $\widehat{\mathcal{B}}_\beta$ with $\beta < \alpha$.

It is clear that $\widehat{\mathcal{B}}_\alpha \subset \mathcal{B}_\alpha$ and that the union of all $\widehat{\mathcal{B}}_\alpha$ is $\mathcal{B}(X)$. The classes $\widehat{\mathcal{B}}_\alpha$ are easier to deal with in some transfinite induction constructions because at every step only one type of operation (complementation or sum) is involved.

(iv) Suppose that the open sets in X are $\mathcal{F}_\sigma$-sets. Show that $\widehat{\mathcal{B}}_\alpha \subset \widehat{\mathcal{B}}_{\alpha+2}$ and $\widehat{\mathcal{B}}_\alpha \subset \widehat{\mathcal{B}}_{\gamma+1}$, provided that γ is a limit ordinal and $\alpha < \gamma$.

(v) Let X have a countable topology base. Prove that for every class $\widehat{\mathcal{B}}_\alpha$ in X, $0 \le \alpha < \omega_1$, there is a set $E \subset X \times \mathbb{N}^\infty$ of class $\widehat{\mathcal{B}}_\alpha$ that is universal in the sense that: (a) every section $E_y = \{x \in X\colon (x,y) \in E\}$, $y \in \mathbb{N}^\infty$, is of class $\widehat{\mathcal{B}}_\alpha$ in X; (b) for every set A of class $\widehat{\mathcal{B}}_\alpha$ in X, there is $y \in \mathbb{N}^\infty$ such that $A = E_y$.

(vi) Prove that for each α with $1 \le \alpha < \omega_1$, the space $\mathbb{N}^\infty$ contains a set of class $\widehat{\mathcal{B}}_\alpha$ that belongs to no $\widehat{\mathcal{B}}_\beta$ with $0 \le \beta < \alpha$. The same is true for any uncountable Polish space.

Hint: in (i) and (ii) use that every countable family of indices α_n is majorized by some β. (iii) The fact that the cardinality of $\sigma(\mathcal{E})$ does not exceed $\mathfrak{c}$ follows by (i). Let $\mathcal{E}$ be a countable family $\{E_n\}$ and let $f = \sum_{n=1}^\infty 3^{-n} I_{E_n}$. If f assumes infinitely many values, then $\sigma(\mathcal{E})$ contains infinitely many disjoint sets, whence it follows that the cardinality of $\sigma(\mathcal{E})$ is at least $\mathfrak{c}$. But if f has only finitely many values, then $\sigma(\{E_n\})$ is finite, hence so is $\{E_n\}$. In (iv) and (v) use transfinite induction. (vi) Let $\alpha \ge 2$. There is a set E of class $\widehat{\mathcal{B}}_\alpha$ in $\mathbb{N}^\infty \times \mathbb{N}^\infty$ that is universal for the $\widehat{\mathcal{B}}_\alpha$-sets in $\mathbb{N}^\infty$. Let $A = \Delta \cap E$, where Δ is the diagonal in $\mathbb{N}^\infty \times \mathbb{N}^\infty$. Show that A is of class $\widehat{\mathcal{B}}_\alpha$ in $\mathbb{N}^\infty \times \mathbb{N}^\infty$. Take the set $B = \Delta \backslash A$ and show that B is in $\widehat{\mathcal{B}}_{\alpha+1}$, but belongs to no $\widehat{\mathcal{B}}_\beta$ with $0 \le \beta < \alpha + 1$.

6.10.32. (Sierpiński [**1715**]) Let $\mathcal{F}$ be a family of subsets in a set X and let $\mathbb{B}(\mathcal{F})$ be the class of all sets that can be obtained from $\mathcal{F}$ by means of finite or countable intersections and unions in an arbitrary order. Prove that $\mathbb{B}(\mathcal{F})$ coincides with the σ-algebra $\sigma(\mathcal{F})$ generated by $\mathcal{F}$ if and only if $E_1 \backslash E_2 \in \mathbb{B}(\mathcal{F})$ for all sets $E_1, E_2 \in \mathcal{F}$.

6.10.33. Show that every complete nonempty metric space without isolated points contains a Borel set that is homeomorphic to $\mathbb{N}^\infty$.

Hint: modify the proof of Theorem 6.1.13.

6.10.34.$^\circ$ Let X be a locally convex space and let X_0 be its linear subspace equipped with the induced topology. Show that the σ-algebra in X_0 generated by the dual space X_0^* coincides with the intersection of X_0 with the σ-algebra in X generated by X^*.

6.10.35.$^\circ$ Show that the closed unit ball in l^1 is not metrizable in the weak topology.

Hint: weak and strong convergences are equivalent for countable sequences in l^1, hence the metrizability of the ball in the weak topology would imply the coincidence of the weak and strong topologies on the ball, which is impossible, since every nonempty weakly open set contains a straight line and meets the sphere.

6.10.36. Let X be the space "two arrows" from Example 6.1.20. Prove that $\mathcal{B}(X)$ is the class of all sets B for which there exists a set $E \in \mathcal{B}[0,1]$ such that $B \triangle \pi^{-1}(E)$ is at most countable, where $\pi\colon X \to [0,1]$ is the natural projection. Hence $\mathcal{B}(X) \subset \mathcal{B}(\mathbb{R}^2)$ and $\mathcal{B}(X)$ is generated by a countable family and singletons (but is not countably generated). In addition, every measure on $\mathcal{B}(X)$ is separable. Finally, if $B \in \mathcal{B}(X)$ is uncountable, then $\pi|_B$ is not injective.

Hint: the class $\mathcal{B}$ of all sets B with the indicated property is a σ-algebra. All one-point sets are closed in X and hence belong to $\mathcal{B}(X)$, i.e., $\mathcal{B}(X)$ contains all countable sets. By the continuity of π, we have $\pi^{-1}(E) \in \mathcal{B}(X)$ for all $E \in \mathcal{B}[0,1]$, whence one has $\mathcal{B} \subset \mathcal{B}(X)$. Since X is hereditary Lindelöf, every open set is an at most countable union of elements of the considered topology base. The elements of the base differ only in one point sets from the preimages of intervals under π, whence it follows that $\mathcal{B}(X) \subset \mathcal{B}$. The elements of the base with rational endpoints along with singletons generate $\mathcal{B}(X)$, since every element of the base is a countable union of elements with rational endpoints with a possible added point. Finally, a countable family of sets in $\mathcal{B}(X)$ cannot separate the points in X, hence cannot generate $\mathcal{B}(X)$.

6.10.37. Construct an example of two countably generated σ-algebras $\mathcal{B}_1$ and $\mathcal{B}_2$ such that $\mathcal{B}_1 \cap \mathcal{B}_2$ is not countably generated.

HINT: see Rao, Rao [**1532**]; one can take the sub-σ-algebras $\mathcal{B}_1$ and $\mathcal{B}_2$ in $\mathcal{B}(\mathbb{R}^1)$ consisting of the sets invariant with respect to translations to 1 and π, respectively.

6.10.38. Let X be a space with a countably generated σ-algebra $\mathcal{A}$ and let $X_0 \subset X$. Show that the σ-algebra of subsets of X_0 that have the form $X_0 \cap A$ with $A \in \mathcal{A}$, is countably generated as well.

HINT: apply Theorem 6.5.5.

6.10.39. Let X be a normal space and let X_0 be closed in X. Prove that $\mathcal{B}a(X_0) = \{B \cap X_0\colon\ B \in \mathcal{B}a(X)\}$.

HINT: any function $f \in C(X_0)$ extends to a continuous function on X, see Engelking [**532**, Theorem 2.1.8].

6.10.40°. Let $(X, \mathcal{B})$ be a measurable space and let Y and S be separable metric spaces. Suppose that a mapping $F\colon X \times Y \to S$ is continuous in y for every fixed $x \in X$ and, for every fixed $y \in Y$, the mapping $x \mapsto F(x, y)$ is measurable with respect to $\mathcal{B}$ and $\mathcal{B}(S)$. Prove that the mapping F is measurable with respect to $\mathcal{B}\otimes\mathcal{B}(Y)$ and $\mathcal{B}(S)$.

HINT: for every n take a countable partition of Y into Borel sets $B_{n,j}$ of diameter at most 2^{-n}, pick in $B_{n,j}$ a point $b_{n,j}$ and set $f_n(x, y) = f(x, b_{n,j})$ whenever $y \in B_{n,j}$; the obtained mappings are measurable with respect to $\mathcal{B}\otimes\mathcal{B}(Y)$ and $\mathcal{B}(S)$ and converge pointwise to f.

6.10.41. (Rudin [**1625**]) Let X be a metric space, let Y be a topological space, let E be a locally convex space, and let a mapping $f\colon X \times Y \to E$ be continuous in every argument separately. Prove that f is a pointwise limit of a sequence of continuous mappings. In particular, if E is metrizable, then f is Borel measurable.

HINT: use the following consequence of paracompactness: for every n one can find continuous functions $\varphi_{\alpha,n}\colon X \to [0, 1]$ with the following properties: one has $\sum_\alpha \varphi_{\alpha,n}(x) = 1$ for all x, every point has a neighborhood in which all functions $\varphi_{\alpha,n}$, with the exception of finitely many of them, vanish and the support of every function $\varphi_{\alpha,n}$ has diameter at most $1/n$. Choose $x_{\alpha,n}$ such that $\varphi_{\alpha,n}(x_{\alpha,n}) > 0$ and set $f_n(x, y) = \sum_\alpha \varphi_{\alpha,n}(x) f(x_{\alpha,n}, y)$.

6.10.42. (i) Let X and Y be Souslin spaces, let $A \subset X \times Y$ be a Souslin set, let $\pi_X(A)$ be the projection of A on X, and let f be a bounded Borel function on A (or, more generally, let the sets $\{f < r\}$ be Souslin). Show that the sets

$$\bigl\{x \in \pi_X(A)\colon\ \inf_y f(x, y) < r\bigr\} \quad \text{and} \quad \{x \in \pi_X(A)\colon\ \inf_y f(x, y) \le r\}$$

are Souslin. Prove an analogous assertion for the sets

$$\bigl\{x \in \pi_X(A)\colon\ \sup_y f(x, y) > r\bigr\} \quad \text{and} \quad \bigl\{x \in \pi_X(A)\colon\ \sup_y f(x, y) \ge r\bigr\}.$$

Show that if $(E, \mathcal{E})$ is a measurable space, $A \in \mathcal{E} \otimes \mathcal{B}(Y)$, and f is a bounded $\mathcal{E} \otimes \mathcal{B}(Y)$-measurable function on A, then the sets

$$\left\{x \in \pi_E(A)\colon \inf_y f(x,y) < r\right\} \quad \text{and} \quad \left\{x \in \pi_E(A)\colon \inf_y f(x,y) \leq r\right\}$$

belong to $\mathcal{S}(\mathcal{E})$.

(ii) Show that there exists a bounded Borel function f on the plane such that the function $g(x) = \sup_y f(x,y)$ is not Borel.

HINT: in (i) represent the indicated sets as projections; in (ii) consider the indicator of a Borel set whose projection is not Borel measurable.

6.10.43.° Prove that there exists a non-Borel (even nonmeasurable) function in the plane that is Borel in every variable separately.

HINT: see Exercise 3.10.49.

6.10.44. (Talagrand [**1826**]) Show that there is a nonseparable metric space whose Borel σ-algebra is generated by balls.

6.10.45.° Give an example of a compact space whose Borel σ-algebra is not generated by closed G_δ-sets.

HINT: consider the product of the continuum of compact intervals.

6.10.46. Give an example of a Polish space whose Borel σ-algebra is not generated by compact sets.

HINT: consider any infinite-dimensional separable Banach space X; observe that σ-algebra generated by compact sets in X is contained in the σ-algebra of all sets A such that either A or $X \backslash A$ is a first category set.

6.10.47. (Bourbaki [**242**, Ch. V, §8, n 5], Chentsov [**335**]) For every $x \in \mathbb{R}$, let I_x be a copy of $[0,1]$ and let U_x be a copy of $(0,1]$. Prove that $\prod_{x \in \mathbb{R}} U_x$ is not Borel in the compact space $\prod_{x \in \mathbb{R}} I_x$.

HINT: see a more general fact in Exercises 7.14.157 and 7.14.158, and also Wise, Hall [**1993**, Example 6.24].

6.10.48.° Let X be the space "two arrows" from Example 6.1.20. Prove that the mappings $f_1\colon (0,1) \to X$, $f_1(x) = (x,1)$, $f_2\colon (0,1) \to X$, $f_2(x) = (x,0)$, are Borel measurable, but $f = (f_1, f_2)\colon (0,1) \to X \times X$ is not Borel measurable.

HINT: the induced topology of the diagonal of $X \times X$ is discrete, hence every subset of it is Borel in the induced topology.

6.10.49. Suppose that sets $E(n_1, \ldots, n_k)$ form a monotone table and satisfy the following condition: if $E(n_1, \ldots, n_k) \cap E(m_1, \ldots, m_p)$ is nonempty for some $k \leq p$, then $n_1 = m_1, \ldots, n_k = m_k$. Prove that

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

HINT: the left-hand side always belongs to the right-hand side; verify the inverse inclusion by using that if x belongs to the set $\bigcup_{(n_i) \in \mathbb{N}^\infty} E(n_1, \ldots, n_k)$ for all indices $k = 1, \ldots, n$, then there exist $m_1, \ldots, m_n$ such that $x \in E(m_1, \ldots, m_n)$; this gives a sequence (m_n) with $x \in \bigcap_{k=1}^{\infty} E(m_1, \ldots, m_k)$.

6.10.50. Let $(X, \mathcal{A})$ be a measurable space, let $S \subset [0, \infty)$ be a countable set, and let $\{A_s\}_{s \in S} \subset \mathcal{A}$ be a cover of X such that $A_s \subset A_t$ whenever $s < t$. Set

$$f(x) = \inf\{s \in S\colon x \in A_s\}.$$

Show that the function f is measurable with respect to $\mathcal{A}$, $f(x) \leq s$ if $x \in A_s$, $f(x) \geq s$ if $x \notin A_s$.

6.10.51. Under the assumption of Martin's axiom prove that there exists an injective function $f\colon \mathbb{R}^1 \to \mathbb{R}^1$ that is nonmeasurable with respect to every probability measure whose domain of definition is a σ-algebra and contains all singletons.

HINT: see Kharazishvili [**992**, Theorem 6, p. 173].

6.10.52. (Sierpiński [**1720**]) Construct a sequence of continuous functions f_n on $[0,1]$ that has cluster points in the topology of pointwise convergence, but all such cluster points are nonmeasurable functions.

6.10.53.° Let f be a surjective Borel mapping of a Souslin space X onto a Souslin space Y and let a set $E \subset Y$ be such that $f^{-1}(E)$ is a Borel set in X. Prove that E is Borel as well.

HINT: the sets $E = f\bigl(f^{-1}(E)\bigr)$ and $Y\backslash E = f\bigl(X\backslash f^{-1}(E)\bigr)$ are disjoint Souslin.

6.10.54. (i) (Purves [**1505**], the implication (b)⇒(a) was obtained by Lusin [**1209**]) Prove that for a Borel mapping F from a Borel subset X of a Polish space to a Polish space Y, the following conditions are equivalent:

(a) $F(B)$ is Borel in Y for every Borel set $B \subset X$;

(b) the set of all values y such that $F^{-1}(y)$ is uncountable, is at most countable.

(ii) (Maitra [**1236**]) Prove that the equivalent conditions (a) and (b) are also equivalent to the following condition: $F^{-1}\bigl(F(B)\bigr)$ is Borel in X for every Borel set $B \subset X$.

6.10.55. (i) Let $(X, \mathcal{B})$ be a measurable space, let $Y \subset X$, and let us set $\mathcal{B}_Y = \{Y \cap B, B \in \mathcal{B}\}$. Prove that every $\mathcal{B}_Y$-measurable function on Y is the restriction of some $\mathcal{B}$-measurable function on all of X.

(ii) (Shortt [**1702**]) Let $\mathcal{B}$ be a σ-algebra of subsets of a space X. Suppose that $\mathcal{B}$ is countably generated and countably separated. Prove that $(X, \mathcal{B})$ is a standard measurable space precisely when for every measurable space $(\Omega, \mathcal{F})$ and every set $\Omega' \subset \Omega$, every mapping $f\colon \Omega' \to (X, \mathcal{B})$ that is measurable with respect to $\mathcal{F} \cap \Omega'$, extends to a measurable mapping $(\Omega, \mathcal{F}) \to (X, \mathcal{B})$.

HINT: (i) it suffices to consider bounded functions passing to $\mathrm{arctg} f$; observe that if sets $B_i \cap Y$, where $i = 1, \ldots, k$ and $B_i \in \mathcal{B}$, are pairwise disjoint, then one can find pairwise disjoint sets $B_i' \in \mathcal{B}$ with $B_i' \cap Y = B_i \cap Y$. Assuming that $0 < f < 1$, consider the sets $A_{i,n} = \{(i-1)2^{-n} < f \leq i2^{-n}\}$, $i = 1, \ldots, 2^n$. Let $f_n = i2^{-n}$ if $x \in A_{i,n}$. Then $|f - f_n| \leq 2^{-n}$. By using the above observation, one can find $\mathcal{B}$-measurable functions g_n such that $g_n|_Y = f_n$ and $\max|g_n - g_{n-1}| = \max|f_n - f_{n-1}| \leq 2^{2-n}$. The required extension can be defined by $g = \lim\limits_{n\to\infty} g_n$.

6.10.56. (i) (Sodnomov [**1759**], [**1760**], Erdős, Stone [**535**], Rogers [**1588**]) Construct two Borel sets A and B on the real line such that the set $A + B$ is not Borel. Show that this is possible even if A is compact and B is a G_δ-set. Construct also a Borel set B on the real line such that $B - B$ is not Borel.

(i) (Rao [**1531**]) Show that there is no countably generated σ-algebra $\mathcal{E}$ in $\mathbb{R}^1$ that is contained in the σ-algebra of Lebesgue measurable sets and has the property that $A + B \in \mathcal{E}$ for all Borel sets A, B.

6.10.57. (Sierpiński [**1713**]) Show that every Souslin set $E \subset [0,1]$ can be represented as $E = f\bigl([0,1)\bigr)$ with some left continuous function f.

6.10.58.° Show that there exists a countable family of intervals on the real line such that it generates the Borel σ-algebra, but every proper subfamily does not.

HINT: consider the intervals $\bigl((k-1)2^{-n}, k2^{-n}\bigr)$ with integer n and k and verify that they separate points, but if one deletes one interval, then this property is lost; say, if one deletes an interval with $k=2l$, then its left endpoint and the middle point are not separated by the remaining intervals; see also Elstrodt [**530**, p. 109], Rao, Shortt [**1535**].

6.10.59. (Jackson, Mauldin ([**874**]) Let $\mathbb{R}^d$ be equipped with some norm and let $\mathcal{L}_0$ be the smallest class of sets containing all open balls for this norm and closed with respect to the operations of complementation and countable union of disjoint sets. Prove that $\mathcal{L}_0 = \mathcal{B}(\mathbb{R}^d)$ (it is shown in Keleti, Preiss [**970**] that the analogous assertion fails for infinite-dimensional separable Banach spaces).

6.10.60. (Szpilrajn [**1812**]) Let $\mathcal{N}$ be some class of subsets of a set X with the following properties: if $N_1 \in \mathcal{N}$ and $N_2 \subset N_1$, then $N_2 \in \mathcal{N}$, and if $N_i \in \mathcal{N}$, then $\bigcup_{i=1}^\infty N_i \in \mathcal{N}$ (such a class is sometimes called a zero class). Suppose that we are given a class $\mathcal{M}$ of subsets of X satisfying the following conditions: (a) $\mathcal{M}$ is closed with respect to countable unions and countable intersections, (b) $\mathcal{M}$ contains the complements of all sets in $\mathcal{N}$, (c) for every set S, there exists a set $\widetilde{S} \in \mathcal{M}$ such that $S \subset \widetilde{S}$ and if $M \in \mathcal{M}$ is such that $S \subset M \subset \widetilde{S}$, then $\widetilde{S}\backslash M \in \mathcal{N}$. Prove that the class $\mathcal{M}$ is closed with respect to the A-operation and derive from this that the class of measurable sets is closed under the A-operation.

HINT: see, e.g., Rogers, Jayne [**1589**, Theorem 2.9.2]. For applications to measurable sets, take for $\mathcal{N}$ the class of all measure zero sets and for $\widetilde{S}$ a measurable envelope of S.

6.10.61. (Mazurkiewicz [**1283**]) Let Z be a closed subset of $\mathbb{N}^\infty$ and $f\colon\ Z \to Y$ a continuous mapping with values in a Souslin space Y. Show that there exists a coanalytic set $E \subset Z$ such that $f(E) = f(Z)$ and f is injective on E. In particular, every Souslin set is the continuous and one-to-one image of some coanalytic set.

HINT: see Theorem 6.9.1 or Kuratowski [**1082**, §39, p. 491].

6.10.62.° Let X be a Souslin space and let f be a Borel function on X. Prove that there is a stronger topology on X generating the initial Borel structure such that X remains Souslin and f becomes continuous.

HINT: observe that the graph of f is a Souslin space that is Borel isomorphic to X (the natural projection operator is a Borel isomorphism), and f is continuous in the topology on X imported from this graph.

6.10.63.° Let X be a Borel set in a Polish space and let $\mathcal{A}$ be a countably generated sub-σ-algebra in $\mathcal{B}(X)$. Prove that there exist a Souslin set $E \in \mathbb{R}^1$ and a Borel function f on X with $f(X) = E$ such that $\mathcal{A} = \{f^{-1}(B), B \in \mathcal{B}(E)\}$.

6.10.64. (i) Prove Theorem 6.10.15. (ii) Let $\mathcal{A}$ be a countably generated σ-algebra in a space X and let $\mathcal{A}$ contain all singletons. Prove that $(X, \mathcal{A})$ is isomorphic to a Souslin subspace of the real line with the induced Borel σ-algebra precisely when for every $\mathcal{A}$-measurable function f, the set $f(X)$ is Souslin.

HINT: (i) use the existence of an injective $\mathcal{A}$-measurable function f generating $\mathcal{A}$. (ii) Take the same function as in (i) and prove that all sets $f(A)$, $A \in \mathcal{A}$, are Souslin, hence Borel in $f(X)$ by the separation theorem, since $f(A) \cap f(X\backslash A) = \varnothing$.

6.10.65. (Maitra [**1235**]) (i) Let A be a Blackwell coanalytic set in a Polish space. Show that for every injective Borel mapping $f\colon\ X\to Y$, where Y is a Polish space, the set $f(A)$ is coanalytic. (ii) Construct an example of a coanalytic set in $[0,1]$ that is not Blackwell.

HINT: (i) apply Theorem 6.2.11; (ii) take a non-Borel Souslin set $E\subset[0,1]$ and a continuous mapping f from the space $\mathcal{R}$ of irrational numbers in $(0,1)$ onto E; use Exercise 6.10.61 to obtain a coanalytic set $A\subset\mathcal{R}$ such that $f(A)=E$ and f is injective on A. The set A is a required one.

6.10.66. Let $\mathcal{K}$ be a class of subsets of a set X such that every collection of sets in $\mathcal{K}$ with the empty intersection has a finite subcollection with the empty intersection. Suppose that for every pair of distinct points x and y, there exist sets $K_x, K_y\in\mathcal{K}$ such that $x\notin K_x$ and $y\notin K_y$. Show that X can be equipped with a Hausdorff topology such that X and all sets in $\mathcal{K}$ are compact.

HINT: consider the topology generated by all sets $X\backslash K$, where $K\in\mathcal{K}$.

6.10.67. Prove Theorem 6.10.6.

HINT: it is clear that $\sigma(X^*)$ is contained in the Baire σ-algebra of the space X with the weak topology. In order to verify the inverse inclusion it suffices to show that for every weakly continuous function F on X, the set $\{x\in X\colon\ F(x)>0\}$ belongs to $\sigma(X^*)$. One can assume that X is embedded as an everywhere dense linear subspace in $\mathbb{R}^T$, where $T=X^*$. Then the weak topology of X coincides with the one induced from $\mathbb{R}^T$. For any rational r, let $U_r=\{x\in X\colon\ F(x)>r\}$, $V_r=\{x\in X\colon\ F(x)<r\}$. There exist open sets $\widetilde{U}_r$ and $\widetilde{V}_r$ in $\mathbb{R}^T$ such that $\widetilde{U}_r\cap X=U_r$, $\widetilde{V}_r\cap X=V_r$. Note that $\widetilde{U}_r\cap\widetilde{V}_r=\varnothing$, since X is dense in $\mathbb{R}^T$. Now we can use Bokstein's theorem (see [**532**, 2.7.12(c)]), according to which there exist a countable set S and open sets U'_r, V'_r in $\mathbb{R}^S$ such that $U'_r\cap V'_r=\varnothing$, $\widetilde{U}_r\subset\pi_S^{-1}(U'_r)$, $\widetilde{V}_r\subset\pi_S^{-1}(V'_r)$. The open sets U'_r, V'_r in the metrizable space $\mathbb{R}^S$ are Baire, hence $X\cap\pi_S^{-1}(U'_r)$ and $X\cap\pi_S^{-1}(V'_r)$ are contained in $\sigma(X^*)$. It remains to observe that $\{x\in X\colon\ F(x)>0\}$ coincides with the union of the sets $X\cap\pi_S^{-1}(U'_r)$ over all rational $r>0$, which is verified directly.

6.10.68.° Let $\mathcal{A}$ and $\mathcal{B}$ be two σ-algebras and let $E\in S(\mathcal{A}\otimes\mathcal{B})$. Show that there exists two sequences $\{A_n\}\subset\mathcal{A}$ and $\{B_n\}\subset\mathcal{B}$ such that $E\in S(\{A_n\}\times\{B_n\})$.

HINT: every $\mathcal{A}\otimes\mathcal{B}$-Souslin set is generated by a countable table of sets in $\mathcal{A}\otimes\mathcal{B}$, hence it remains to apply Exercise 1.12.54.

6.10.69.° Let $\mathcal{E}$ be a σ-algebra in a space X, let Y be a Souslin space, and let Z be a Souslin set in Y. Show that $S\bigl(\mathcal{E}\otimes\mathcal{B}(Z)\bigr)\subset S\bigl(\mathcal{E}\times\mathcal{B}(Y)\bigr)$.

HINT: it suffices to verify that $\mathcal{E}\otimes\mathcal{B}(Z)\subset S\bigl(\mathcal{E}\times\mathcal{B}(Y)\bigr)$; since $\mathcal{E}\times\mathcal{B}(Z)$ is a semialgebra and $S\bigl(\mathcal{E}\times\mathcal{B}(Y)\bigr)$ is a monotone class, it remains to observe that $E\times B\in S\bigl(\mathcal{E}\times\mathcal{B}(Y)\bigr)$ for all $E\in\mathcal{E}$ and $B\in\mathcal{B}(Z)\subset S\bigl(\mathcal{B}(Y)\bigr)$.

6.10.70. (Jayne [**887**]) Let X be a topological space. Show that $\mathcal{B}a(X)$ is the smallest class of sets that contains all functionally closed sets and admits countable unions of disjoint sets and arbitrary countable intersections.

6.10.71.° Let X be a topological space and let $\mathcal{F}_0$ be the class of all functionally closed sets in X. Show that $S\bigl(\mathcal{B}a(X)\bigr)=S(\mathcal{F}_0)$. In particular, in every metric space, all Borel sets are $\mathcal{F}$-analytic.

HINT: observe that if $F\in\mathcal{F}_0$, then $X\backslash F=\bigcup_{n=1}^\infty F_n$, where $F_n\in\mathcal{F}_0$; consider the class $\mathcal{E}:=\{B\in\mathcal{B}a(X)\colon\ B, X\backslash B\in S(\mathcal{F}_0)\}$ and verify that $\mathcal{E}=\mathcal{B}a(X)$.

6.10.72. Let $(X, \mathcal{A}, \mu)$ be a probability space, $\mathcal{A}$ a countably generated σ-algebra, $(T, \mathcal{B})$ a measurable space, and let μ_t, where $t \in T$, be a family of bounded measures on $\mathcal{A}$ absolutely continuous with respect to μ such that for every $A \in \mathcal{A}$, the function $t \mapsto \mu_t(A)$ is measurable with respect to $\mathcal{B}$. Prove that one can find an $\mathcal{A}\otimes\mathcal{B}$-measurable function f on $X \times T$ such that for every $t \in T$, the function $x \mapsto f(x,t)$ is the Radon–Nikodym density of the measure μ_t with respect to μ.

HINT: if $X = [0,1]$ and $\mathcal{A} = \mathcal{B}([0,1])$, then, by Theorem 5.8.8, for every t, the Radon–Nikodym density of μ_t with respect to μ is given by the equality

$$f(x,t) = \lim_{n\to\infty} \mu_t([x-\varepsilon_n, x+\varepsilon_n])/\mu([x-\varepsilon_n, x+\varepsilon_n]),$$

where $\varepsilon_n = n^{-1}$, $f(x,t) = 0$ if $\mu([x-\varepsilon_n, x+\varepsilon_n]) = 0$ for some n. One can assume that the measure μ has no atoms, since for its purely atomic part the claim is obvious. It is readily seen that the functions $\mu_t([x-\varepsilon_n, x+\varepsilon_n])/\mu([x-\varepsilon_n, x+\varepsilon_n])$ are measurable with respect to $\mathcal{B}([0,1])\otimes\mathcal{B}$, since the numerator and denominator are continuous in x due to the absence of atoms and are $\mathcal{B}$-measurable in t. The above limit exists for a.e. x if t is fixed, for all other x we set $f(x,t) = 0$. In the general case, according to Theorem 6.5.5, there exists an $\mathcal{A}$-measurable function $\xi\colon X \to [0,1]$ such that $\mathcal{A} = \bigl\{\xi^{-1}(B), B \in \mathcal{B}([0,1])\bigr\}$. Set $\nu = \mu\circ\xi^{-1}$, $\nu_t = \mu_t\circ\xi^{-1}$. Then $\nu_t \ll \nu$ and by the above there exists a $\mathcal{B}([0,1])\otimes\mathcal{B}$-measurable version $(x,t) \mapsto \varrho(x,t)$ of the Radon–Nikodym densities of the measures ν_t with respect to ν. Set $f(x,t) := \varrho\bigl(\xi(x), t\bigr)$. The function f is measurable with respect to $\mathcal{A}\otimes\mathcal{B}$. Let t be fixed. Given a set $A \in \mathcal{A}$, we can find a set $B \in \mathcal{B}([0,1])$ with $A = \xi^{-1}(B)$. Since $I_B\bigl(\xi(x)\bigr) = I_A(x)$, we obtain

$$\mu_t(A) = \nu_t(B) = \int_B \varrho(y,t)\,\nu(dy) = \int_X I_B\bigl(\xi(x)\bigr) f(x,t)\,\mu(dx) = \int_A f(x,t)\,\mu(dx).$$

6.10.73. (i) (C. Doléans-Dade) Let $(X, \mathcal{A}, \mu)$ be a probability space, $(T, \mathcal{B})$ a measurable space, and let $f_n(x,t)$ be a sequence of $\mathcal{A}\otimes\mathcal{B}$-measurable functions on $X \times T$ such that for every fixed t, the sequence of functions $x \mapsto f_n(x,t)$ is fundamental in measure μ. Show that there exists an $\mathcal{A}\otimes\mathcal{B}$-measurable function f such that $f_n(\,\cdot\,,t) \to f(\,\cdot\,,t)$ in measure μ for every t.

(ii) (Stricker, Yor [**1793**]) Let $(X, \mathcal{A}, \mu)$ be a probability space with a separable measure μ, $(T, \mathcal{B})$ a measurable space, and let $f_n(x,t)$ be a sequence of $\mathcal{A}\otimes\mathcal{B}$-measurable functions on $X\times T$ such that for every fixed t, the functions $x \mapsto f_n(x,t)$ are integrable against the measure μ and converge weakly in $L^1(\mu)$. Show that there exists an $\mathcal{A}\otimes\mathcal{B}$-measurable function f integrable in x such that $f_n(\,\cdot\,,t) \to f(\,\cdot\,,t)$ weakly in $L^1(\mu)$ for every t.

HINT: (i) one can assume that the functions f_n are uniformly bounded, passing to $\operatorname{arctg} f_n$. Then, for every t, the sequence $f_n(\,\cdot\,,t)$ is fundamental in $L^2(\mu)$. The functions

$$g_{n,k}(t) = \int_X |f_n(x,t) - f_k(x,t)|^2\,\mu(dx)$$

are measurable with respect to $\mathcal{B}$. For every $p \in \mathbb{N}$, let $m_p(t)$ be the smallest m such that $g_{n,k}(t) \le 8^{-p}$ for all $n,k \ge m$. It is easy to see from the proof of the Riesz theorem that for every t, the sequence $f_{m_p(t)}(x,t)$ converges μ-a.e. In addition, it is readily verified that the functions $m_p(t)$ are $\mathcal{B}$-measurable. Hence the function $f_{m_p(t)}(x,t)$ is measurable with respect to $\mathcal{A}\otimes\mathcal{B}$. The desired function is defined as follows: $f(x,t) = \lim\limits_{p\to\infty} f_{m_p(t)}(x,t)$ if this limit exists and $f(x,t) = 0$ otherwise. Assertion (ii) follows by Exercise 6.10.72.

6.10.74. Let (M,d) be a separable metric space and let $D \subset M^\infty$ consist of all sequences without isolated points. Show that D is a G_δ-set.

HINT: let $\{a_m\}$ be a countable everywhere dense set in M. Verify that for every fixed k, m, and n, the set of all sequences $\{x_j\} \in M^\infty$ such that $d(x_k, a_m) \leq 2^{-n-1}$ and $d(x_j, a_m) \geq 2^{-n}$ for all $j \neq k$, is closed.

6.10.75. Let τ be an uncountable ordinal. Show that for every continuous function f on the space $[0,\tau)$ with the order topology, there exists $\tau_0 < \tau$ such that f is constant on $[\tau_0, \tau)$.

HINT: for every k, there exists $\alpha_k < \tau$ such that $|f(\alpha) - f(\beta)| < 1/k$ whenever $\alpha, \beta > \alpha_k$. Indeed, otherwise one could construct by induction an increasing sequence α_{k_n} with $|f(\alpha_{k_{n+1}}) - f(\alpha_{k_n})| \geq 1/k$. This contradicts the continuity of f since such a sequence converges to $\sup \alpha_{k_n}$. There exists $\tau_0 < \tau$ such that $\alpha_k < \tau_0$ for all k. It is clear that τ_0 is the required ordinal.

6.10.76.° A set E in a topological space X is said to have the Baire property if there exists an open set U such that $E \triangle U$ is a first category set. Show that the class $\mathcal{BP}(X)$ of all sets in X with the Baire property is a σ-algebra containing $\mathcal{B}(X)$.

HINT: if F is closed, then $F \in \mathcal{BP}(X)$, since one can take for U the interior of F. If $A \in \mathcal{BP}(X)$ and $A \triangle B$ is a first category set, then it is easy to see that $B \in \mathcal{BP}(X)$. This yields that if $E \in \mathcal{BP}(X)$, then $X\backslash E \in \mathcal{BP}(X)$, since $(X\backslash E) \triangle (X\backslash U) = E \triangle U$, where U is open. Finally, it is readily verified that $\mathcal{BP}(X)$ admits countable unions. All open sets belong to $\mathcal{BP}(X)$ by definition.

6.10.77. Suppose that X and Y are compact spaces, Y is metrizable, μ is a probability measure on $\mathcal{B}(Y)$, and $f\colon X \to Y$ is continuous. Prove that there exists a $\bigl(\mathcal{B}(Y)_\mu, \mathcal{B}(X)\bigr)$-measurable mapping $g\colon Y \to X$ with $g(y) \in f^{-1}(y)$ for all $y \in f(X)$.

HINT: see Graf [**718**].

6.10.78. (i) Let $X = \mathbb{Q}$ be equipped with the topology which is obtained by reinforcing the usual induced topology with the complement of the sequence $\{1/n\}$. Show that we obtain a countable Hausdorff space (in particular, a Souslin space) that has a countable base but is not regular. (ii) Construct a countable Hausdorff space with a countable base such that some point in this space is not a Baire set.

HINT: (i) see Arkhangel'skiĭ, Ponomarev [**68**, Ch. II, Problem 103]; (ii) see Steen, Seebach [**1774**, p. 98, Counterexample 80].

6.10.79.° (A.D. Alexandroff [**30**]) Let Z_n be disjoint functionally closed sets in a topological space X.

(i) Let Z_n have pairwise disjoint functionally open neighborhoods U_n such that $Z := \bigcup_{n=1}^\infty Z_n$ is closed. Prove that Z is functionally closed.

(ii) Suppose that every union of sets Z_n is functionally closed. Show that the sets Z_n possess pairwise disjoint functionally open neighborhoods U_n.

(iii) Show that if the space X is normal, then the assumption that all unions of Z_n are closed yields that they are functionally closed.

HINT: (i) there are continuous functions $f_n\colon X \to [0, 3^{-n}]$ such that $f_n = 0$ outside U_n and $Z_n = \{f_n = 3^{-n}\}$. The function $f = \sum_{n=1}^\infty f_n$ is continuous. Note that Z coincides with $f^{-1}(S)$, where S is the closed countable set consisting of the numbers $s_n := \sum_{k=1}^n 3^{-k}$ and their limit $1/2$. Indeed, $f|_{Z_n} = s_n$ since supports of f_n are disjoint. Hence $Z_n \subset f^{-1}(S)$, i.e., $Z \subset f^{-1}(S)$. If $x \in f^{-1}(S)$ and $f(x) = s_n$, then $x \in Z_n$, since $\sum_{j=n+1}^\infty 3^{-j} < 3^{-n}$. If we had $f(x) = 1/2$, then by the above the point x would be a limit point of Z, hence it would belong to one of the sets Z_n because Z is closed. However, this is impossible. (ii) Sets U_n are constructed by induction. We find disjoint functionally open neighborhoods U_1 and V_1 of the functionally closed sets Z_1 and $\bigcup_{n=2}^{\infty} Z_n$. Next, in V_1 we find disjoint functionally open neighborhoods of the sets Z_2 and $\bigcup_{n=3}^{\infty} Z_n$ and so on. (iii) For any normal space X the reasoning in (ii) is applicable to arbitrary closed sets, hence the assertion follows by (i).

6.10.80°. (A.D. Alexandroff [**30**]) (i) Let F_n be functionally closed sets in a topological space X and let $F_{n+1} \subset F_n$ for all $n \in \mathbb{N}$. Prove that there exist functionally open sets G_n such that $F_n \subset G_n$ and $G_{n+1} \subset G_n$ for all n and $\bigcap_{n=1}^{\infty} G_n = \bigcap_{n=1}^{\infty} F_n$.

(ii) Suppose that in (i) one has the equality $\bigcap_{n=1}^{\infty} F_n = \varnothing$. Let $Z_n := F_n \backslash G_{n+1}$. Show that the sets Z_n are disjoint, and for every sequence $\{n_k\}$ of natural numbers, the set $\bigcup_{k=1}^{\infty} Z_{n_k}$ is functionally closed.

HINT: (i) there exist $f_n \in C(X)$ with $0 \le f_n \le 1$, $F_n = f_n^{-1}(0)$. Let us set $h_n := f_1 + \cdots + f_n$ and $G_n := \{x\colon\ h_n(x) < 1/n\}$. Then $F_n \subset G_n$, $G_{n+1} \subset G_n$. If $x \notin \bigcap_{n=1}^{\infty} F_n$, then there exists n such that $h_n(x) > 0$. Hence there exists $m > n$ such that $h_m(x) > 1/m$, i.e., $x \notin \bigcap_{n=1}^{\infty} G_n$.

(ii) It is obvious that the sets Z_n are disjoint, since $Z_n \cap F_{n+1} = \varnothing$. By induction we find two sequences of functionally open sets U_n and V_n with the following properties: $Z_n \subset U_n \subset G_n$, $U_n \cap V_n = \varnothing$, $F_{n+1} \subset V_n$, $U_{n+1} \subset V_n$. To this end, we include Z_1 and F_2 into disjoint functionally open sets U_1 and V_1 contained in G_1. Next we consider the functionally open set $G_2 \cap V_1$ containing disjoint functionally closed sets Z_2 and F_3 and so on. The sets U_n are disjoint. Every set $\bigcup_{k=1}^{\infty} Z_{n_k}$ is functionally closed. Indeed, suppose x is not in this set. We find k such that $x \notin G_{n_k}$. Since $X \backslash G_{n_k}$ and F_{n_k} are disjoint functionally closed sets and $Z_j \subset F_{n_k}$ for all $j \ge n_k$, the point x has a functionally open neighborhood W not meeting the sets Z_j, $j \ge n_k$. Since x does not belong to $Z_{n_1}, \ldots, Z_{n_{k-1}}$, there exists a functionally open neighborhood of x not meeting $\bigcup_{k=1}^{\infty} Z_{n_k}$. By Exercise 6.10.79 the set $\bigcup_{k=1}^{\infty} Z_{n_k}$ is functionally closed.

6.10.81. Show that the set $D := \{(s, -s) \in Z^2\}$ in the Sorgenfrey plane Z^2 (see Example 6.1.19) is Baire and that for every Baire set $B \in Z^2$, the intersection $B \cap D$ is Borel with respect to the usual topology of the plane.

HINT: the function $(x, y) \mapsto x + y$ is continuous on Z^2, hence D is functionally closed. Therefore, it suffices to verify that every functionally closed set $F \subset D$ belongs to $\mathcal{B}(\mathbb{R}^2)$. Let $F = f^{-1}(0)$, where $f \in C(Z^2)$. Let us write $x = (t, s)$ and set $B_k(x) := [t, t+1/k) \times [s, s+1/k)$,

$$U_n := \bigl\{x\colon\ |f(x)| < (2n)^{-1}\bigr\},\ U_{n,k} := \bigl\{x \in F\colon\ B_k(x) \subset U_n\bigr\}.$$

Let $W_{n,k}$ be the closure of $U_{n,k}$ in the usual topology and let $B := \bigcap_{n=1}^{\infty} \bigcup_{k=1}^{\infty} W_{n,k}$. Then $B \in \mathcal{B}(\mathbb{R}^2)$ and it suffices to show that $F = B$. Since for every $n \in \mathbb{N}$, by the continuity of f we have $F = \bigcup_{k=1}^{\infty} U_{n,k}$, one obtains $F \subset B$. Let $x \notin F$. Then, for some n, we have $|f(x)| \ge 1/n$. There is $k \in \mathbb{N}$ such that $|f(y)| > (2n)^{-1}$ for all $y \in B_k(x)$. This yields that x belongs to no $W_{n,k}$, since the sets $B_k(x)$ and $B_k(z)$ meet if $z \in D$ and $|x - z| < (2k)^{-1}$. Thus, $x \notin B$, i.e., $B \subset F$.

6.10.82. Let X be a topological space such that there exists a continuous injective mapping h from X to some metric space. Let $A \subset X$. Suppose that every infinite sequence in A has a limit point in X. Show that the closure of A is metrizable and compact.

HINT: observe that $h(\overline{A}) = \overline{h(A)}$. Indeed, $h(\overline{A}) \subset \overline{h(A)}$ by the continuity of h. If $y \in \overline{h(A)}$, then $y = \lim\limits_{n\to\infty} h(x_n)$, where $x_n \in A$. Hence either $y \in h(A)$, or we may assume that $\{x_n\}$ is infinite and then $y = h(x)$, where x is a limit point of $\{x_n\}$. One can also conclude that the set $\overline{h(A)}$ in a metric space is compact. The same is true for every subset of A, whence it follows that $h^{-1}\colon \overline{h(A)} \to \overline{A}$ is continuous, since the preimages of all closed sets in $\overline{A}$ are compact. Thus, h is a homeomorphism between $\overline{A}$ and $\overline{h(A)}$.

6.10.83. Prove that the σ-algebra generated by Souslin sets in $[0,1]$ is strictly smaller than the σ-algebra of all Lebesgue measurable sets.

HINT: the first σ-algebra has the cardinality of the continuum $\mathfrak{c}$ (see Exercise 6.10.31), and the cardinality of the second one is $2^{\mathfrak{c}}$. A much deeper fact is contained in the next result.

6.10.84. (Kunugui [**1079**]) Prove that the σ-algebra generated by Souslin sets in $[0,1]$ is not closed with respect to the A-operation.

6.10.85. Let $(X,\mathcal{A})$ be a measurable space and let a function $f\colon [0,1]\times X \to \mathbb{R}^1$ be such that, for every $t \in [0,1]$, the function $x \mapsto f(t,x)$ is $\mathcal{A}$-measurable, and, for every $x \in X$, the function $t \mapsto f(t,x)$ is increasing. Suppose that $f(1,x) \geq 0$. Show that the function $g(x) := \inf\bigl\{t \in [0,1]\colon f(t,x) \geq 0\bigr\}$ is $\mathcal{A}$-measurable.

HINT: one has $g(x) = \inf\bigl\{t \in [0,1] \cap \mathbb{Q}\colon f(t,x) \geq 0\bigr\}$, since for every $t \in [0,1]$ and $\varepsilon > 0$, there is a rational number $s \in (t, t+\varepsilon)$ and $f(s,x) \geq f(t,x)$. Let $\{t_n\}$ be the set of all rational numbers in $[0,1]$ and let $g_n(x)$ be the minimal number in the finite set $\{t_1,\ldots,t_n,1\}$ such that $f(t_i,x) \geq 0$ (such a number exists since $f(1,x) \geq 0$). It is readily seen that the function g_n is $\mathcal{A}$-measurable. Hence so is $g(x) = \lim\limits_{n\to\infty} g_n(x)$.

6.10.86. Let $(X,\mathcal{A})$ be a measurable space and let Y be a metrizable Souslin space. For any $A \subset X\times Y$ let

$$A^{cl} := \bigl\{(x,y) \in X\times Y\colon\ y \in \overline{A_x}\bigr\}, \quad A^{int} := \bigl\{(x,y) \in X\times Y\colon\ y \in \operatorname{Int} A_x\bigr\},$$

and $A^{bd} = A^{cl}\backslash A^{int}$, where $A_x := \{y \in Y\colon (x,y) \in A\}$, $\overline{M}$ is the closure of a set M and Int$M$ is the interior of M. Prove that:

(i) if Y is metrizable by a metric d, then for all r one has

$$\bigl\{(x,y) \in X\times Y\colon\ d(y,A_x) \leq r\bigr\} \in \mathcal{S}\bigl(\mathcal{A}\otimes\mathcal{B}(Y)\bigr), \quad \text{where } d(y,\varnothing) := +\infty,$$

(ii) if $A \in \mathcal{S}\bigl(\mathcal{A}\otimes\mathcal{B}(Y)\bigr)$, then $A^{cl} \in \mathcal{S}\bigl(\mathcal{A}\otimes\mathcal{B}(Y)\bigr)$,

(iii) if $X\times Y\backslash A \in \mathcal{S}\bigl(\mathcal{A}\otimes\mathcal{B}(Y)\bigr)$, then $X\times Y\backslash A^{cl} \in \mathcal{S}\bigl(\mathcal{A}\otimes\mathcal{B}(Y)\bigr)$,

(iv) if $A \in \mathcal{A}\otimes\mathcal{B}(Y)$, then $A^{bd} \in \mathcal{S}\bigl(\mathcal{A}\otimes\mathcal{B}(Y)\bigr)$.

HINT: let $E = \{(x,y,z) \in X\times Y^2\colon (x,z) \in A\}$; apply Exercise 6.10.42 and the equality $d(y,A_x) = \inf\{d(y,z)\colon z \in E_{(x,y)}\}$. Now (ii) follows from (i), since one has $A^{cl} = \{(x,y)\colon d(y,A_x) = 0\}$. Finally, (iii) and (iv) follow from (ii).

6.10.87.$^\circ$ (i) Let us equip the set $X = [0,1]^2$ with the order topology with respect to the lexicographic ordering, i.e., $(x_1,y_1) < (x_2,y_2)$ if $x_1 < x_2$ and if $x_1 = x_2$ and $y_1 < y_2$. Show that X is compact and the natural projection $f\colon X \to [0,1]$ is continuous. (ii) Show that the space "two arrows", denoted by X_0, is closed in X.

(iii) Show that the sets $\{x\}\times(0,1)$ are open in X, hence one can find an open set in X whose projection is not Lebesgue measurable.

HINT: (i) a neighborhood of a point $(x,0)$ contains a strip; (ii) and (iii) are straightforward.

CHAPTER 7

Measures on topological spaces

> As soon as we establish what is required from a naval architect *in his speciality*, then immediately the corresponding volume of knowledge from calculus and mechanics is set up. But here one must be very careful not to introduce superfluous requirements; for the fact that the upper deck is covered with wood does not necessitate the study of botany, or that a sofa in the ward-room is upholstered with leather does not force one to study zoology; the same is here: if a consideration of some particular question involves a certain formula, then it is much better to present it without proof rather than introduce in the course a whole branch of mathematics in order to give a full derivation of that single formula.
>
> *A.N. Krylov.* My recollections.

7.1. Borel, Baire and Radon measures

In classical measure theory, it is customary to fix some domain of definition of a measure (say, the σ-algebra of all measurable sets). This domain is either given in advance or is obtained as a result of some extension procedure (for example, the Lebesgue–Carathéodory extension). However, in many applications, as we shall see below, the choice of domain of measure turns out to be a very delicate question, and the problem of extension to a larger domain is not always solved by completing. Typical examples of such a situation are related to measures on topological spaces or spaces equipped with filtrations. Such problems occur in the study of the distributions of random processes in functional spaces. This chapter is devoted to a broad circle of problems related to regularity and domains of definition of measures. We discuss Borel and Baire measures and their regularity properties such as tightness, τ-additivity etc. We shall see that any Baire measure is regular. On the other hand, we shall encounter examples of Borel measures that are neither regular nor tight, and examples of Borel measures on compact spaces that are not Radon (although are tight). It will be shown that there exist Baire measures without countably additive extensions to the Borel σ-algebra. This picture will be complemented by the theorem that every tight Baire measure can be extended to a Borel measure and has a unique extension to a Radon measure. In particular, any Baire measure on a compact space X can be (uniquely) extended to a Radon measure on X (although non-Radon extensions to $\mathcal{B}(X)$ may exist as well). Radon measures are most frequently encountered in real applications, so they are given particular attention. Throughout we consider measures of *bounded variation* unless the opposite is explicitly said (regarding infinite measures, see §7.11 and §7.14(xviii)). In addition, we consider only Hausdorff spaces (although not everywhere is this essential).

7.1.1. Definition. *Let X be a topological space.*

(i) *A countably additive measure on the Borel σ-algebra $\mathcal{B}(X)$ is called a Borel measure on X.*

(ii) *A countably additive measure on the Baire σ-algebra $\mathcal{B}a(X)$ is called a Baire measure on X.*

(iii) *A Borel measure μ on X is called a Radon measure if for every B in $\mathcal{B}(X)$ and $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset B$ such that $|\mu|(B\backslash K_\varepsilon) < \varepsilon$.*

A set in a topological space X is called *universally measurable* if it belongs to the Lebesgue completion of $\mathcal{B}(X)$ with respect to every Borel measure on X. A set measurable with respect to every Radon measure on X is called *universally Radon measurable.* A mapping F from X to a topological space Y is called *universally measurable* if so are the sets $F^{-1}(B)$ for all $B \in \mathcal{B}(Y)$.

The following lemma shows that Borel measures are uniquely determined by their values on open sets.

7.1.2. Lemma. *If two Borel measures on a topological space coincide on all open sets, then they coincide on all Borel sets.*

Proof. It suffices to verify that a Borel measure μ vanishing on all open sets is identically zero. The measures μ^+ and μ^- are nonnegative and coincide on all open sets. Then $\mu^+ = \mu^-$ by Lemma 1.9.4 because the class of all open sets admits finite intersections. Since $\mu^+ \perp \mu^-$, one has $\mu^+ = \mu^- = 0$. $\square$

We observe that, by definition, a measure μ is Radon if and only if the measure $|\mu|$ is Radon. This is also equivalent to that both measures μ^+ and μ^- are Radon.

Radon measures constitute the most important class of measures for applications. As we shall see later, on many spaces (including complete separable metric spaces) all Borel measures are Radon. However, first we consider an example due to Dieudonné [**445**], which shows that even on a compact space a Borel measure may fail to be Radon.

7.1.3. Example. There exists a compact topological space X with a Borel measure μ such that μ assumes only two values 1 and 0, but is not Radon.

Proof. We take for X the set of all ordinals not exceeding the first uncountable ordinal ω_1. Then X is an uncountable well-ordered set with the maximal element ω_1, and for any $\alpha \neq \omega_1$ the set $\{x\colon\, x \leq \alpha\}$ is at most countable. We equip X with the order topology (§6.1); in this topology X is compact. Let $X_0 = X\backslash\{\omega_1\}$. Denote by $\mathcal{F}_0$ the class of all uncountable closed subsets in the space X_0 equipped with the induced topology. The measure

μ on $\mathcal{B}(X)$ is defined as follows: $\mu(B) = 1$ if B contains a set from $\mathcal{F}_0$ and $\mu(B) = 0$ otherwise. Let us show that μ is countably additive. To this end, let us introduce the class $\mathcal{E}$ of all sets $E \subset X$ such that either E or $X \backslash E$ contains an element from $\mathcal{F}_0$. The class $\mathcal{E}$ is a σ-algebra: it is closed under complementation and countable intersections, since $F := \bigcap_{n=1}^{\infty} F_n \in \mathcal{F}_0$ if $F_n \in \mathcal{F}_0$. Indeed, if F is countable, there is $\alpha < \omega_1$ such that $F \subset [0, \alpha]$. By induction one can easily find a strictly increasing sequence of ordinals $\alpha_j \in (\alpha, \omega_1)$ that contains infinitely many elements from every F_n (because F_n is uncountable and $[0, \alpha_j]$ is countable). Then $\{\alpha_j\}$ has a limit $\alpha' \in F$ and $\alpha' > \alpha$. In addition, $\mathcal{B}(X) \subset \mathcal{E}$. Indeed, if A is closed and uncountable, then $A \cap X_0 \in \mathcal{F}_0$. If A is at most countable, then its complement contains an element from $\mathcal{F}_0$ since $A \subset [0, \alpha]$ for some $\alpha < \omega_1$. Suppose now that $\{B_n\} \subset \mathcal{B}(X)$ is a sequence of disjoint sets. As shown above, at most one of them contains an element from $\mathcal{F}_0$, and if there is no such B_n, every $X \backslash B_n$ contains a set $F_n \in \mathcal{F}_0$, hence $\bigcap_{n=1}^{\infty} F_n \in \mathcal{F}_0$, so $\bigcup_{n=1}^{\infty} B_n$ has no subsets from $\mathcal{F}_0$. Therefore, μ is countably additive. Every point $x \neq \omega_1$ has a neighborhood of measure zero, hence $\mu(K) = 0$ for every compact set $K \subset X_0$. Since $\mu(\{\omega_1\}) = 0$, μ is not Radon (moreover, it even has no support, i.e., the smallest closed set of full measure because ω_1 belongs to every closed set of full measure; see below about supports of measures). $\square$

The measure μ constructed in this example is called the Dieudonné measure.

Thus, in order to ensure the Radon property of a measure, it is not enough to be able to approximate its value on the whole space by the values on compact sets. The latter property has a special name.

7.1.4. Definition. *A nonnegative set function μ defined on some system $\mathcal{A}$ of subsets of a topological space X is called tight on $\mathcal{A}$ if for every $\varepsilon > 0$, there exists a compact set K_ε in X such that $\mu(A) < \varepsilon$ for every element A in $\mathcal{A}$ that does not meet K_ε. An additive set function μ of bounded variation on an algebra* (*or a ring*) *is called tight if its total variation $|\mu|$ is tight.*

A Borel measure μ is tight if and only if for every $\varepsilon > 0$ there is a compact set K_ε such that $|\mu|(X \backslash K_\varepsilon) < \varepsilon$. However, already in the case of a general Baire measure one has to formulate this property in the way indicated in the foregoing definition because nonempty compact sets may not belong to the domain of such a measure.

It is clear that any measure on a compact space is tight. What is missing for a tight Borel measure to be Radon?

7.1.5. Definition. *A nonnegative set function μ defined on some system $\mathcal{A}$ of subsets of a topological space is called regular if for every A in $\mathcal{A}$ and every $\varepsilon > 0$, there exists a closed set F_ε such that $F_\varepsilon \subset A$, $A \backslash F_\varepsilon \in \mathcal{A}$ and $\mu(A \backslash F_\varepsilon) < \varepsilon$.*

An additive set function μ of bounded variation on an algebra (*or a ring*) *is called regular if its total variation $|\mu|$ is regular.*

By definition, every Radon measure on a Hausdorff space is regular and tight. It is clear that if a Borel measure is regular and tight, then it is Radon, since the intersection of a compact set and a closed set is compact. However, a regular Borel measure may fail to be tight. Let us consider an example.

7.1.6. Example. Let M be a nonmeasurable subset of the interval $[0,1]$ with zero inner measure and unit outer measure (see Chapter 1). We consider M with the usual metric as a metric space. Then every Borel subset of this space has the form $M \cap B$, where B is a Borel subset in $[0,1]$. We define a measure on M by the formula $\mu(M \cap B) = \lambda(B)$, where λ is Lebesgue measure, i.e., μ is the restriction of λ to M in the sense of Definition 1.12.11. Since Lebesgue measure is regular (see, for example, Theorem 1.4.8), the measure μ is regular as well (we recall that the closed sets in M are the intersections of M with closed subsets of $[0,1]$). But it is not tight, since every compact set K in the space M is also compact in $[0,1]$, hence, by construction, has Lebesgue measure zero, whence we obtain $\mu(K) = 0$.

The above example of a non-tight measure on a separable metric space might seem artificial because of a rather exotic choice of the space M, and one might be tempted to choose for M a more constructive space. In the subsequent sections we shall see that exotic spaces are inevitable in such examples and that this circumstance has deep set-theoretic reasons. The following theorem shows that one cannot take for M a Borel set in $[0,1]$. This is one of the most important theorems in measure theory and is often used in applications.

7.1.7. Theorem. *Let X be a metric space. Then every Borel measure μ on X is regular. If X is complete and separable, then the measure μ is Radon.*

PROOF. We can assume that $\mu \geq 0$. The regularity of μ has actually been proven in Theorem 1.4.8 (no specific features of $\mathbb{R}^n$ have been used). Let us suppose that X is complete and separable and show that the measure μ is tight. Let $\varepsilon > 0$. By the separability of X, for every natural n, one can cover X by a finite or countable family of open balls U_n^j of radius $\varepsilon 2^{-n}$. By using the countable additivity of μ, one can find a finite union $W_n = \bigcup_{j=1}^{m_n} U_n^j$ such that $\mu(X \backslash W_n) < \varepsilon 2^{-n}$. The set $W = \bigcap_{n=1}^{\infty} W_n$ is completely bounded, since for every $\delta > 0$, it can be covered by finitely many balls of radius δ. In addition, $\mu(X \backslash W) \leq \sum_{n=1}^{\infty} \mu(X \backslash W_n) < \varepsilon$. It remains to observe that the closure K of the set W is compact by the completeness of X. The tightness and regularity yield that our measure is Radon. $\square$

7.1.8. Corollary. *Every Baire measure μ on a topological space X is regular. Moreover, for every Baire set E and every $\varepsilon > 0$, there exists a continuous function f on X such that $f^{-1}(0) \subset E$ and $|\mu|\bigl(E \backslash f^{-1}(0)\bigr) < \varepsilon$.*

More generally, for any family Γ of continuous functions on X, every measure μ on the σ-algebra $\sigma(\Gamma)$ generated by Γ is regular.

Proof. It suffices to consider nonnegative measures. We recall that the set E has the form

$$E = \{x\colon\ (f_1(x), \ldots, f_n(x), \ldots) \in B\},$$

where $B \in \mathcal{B}(\mathbb{R}^\infty)$ and $f_n \in C(X)$ (in the second case, $f_n \in \Gamma$). Let μ_0 be the image of the measure μ under the mapping $x \mapsto (f_1(x), \ldots, f_n(x), \ldots)$ from X to $\mathbb{R}^\infty$. This mapping is continuous. Since $\mathbb{R}^\infty$ is a metric space, by the above theorem, there exists a continuous function g on $\mathbb{R}^\infty$ such that $g^{-1}(0) \subset B$ and

$$\mu_0\big(B\backslash g^{-1}(0)\big) < \varepsilon.$$

It remains to observe that the function $f(x) = g\big(f_1(x), \ldots, f_n(x), \ldots\big)$ is continuous on X and by the definition of the image measure, we have the equality $\mu\big(E\backslash f^{-1}(0)\big) = \mu_0\big(B\backslash g^{-1}(0)\big)$. □

We recall that a topological space X is called perfectly normal if every closed set in X has the form $f^{-1}(0)$, where $f \in C(X)$. It is clear that in this case the Borel σ-algebra coincides with the Baire one. So the following assertion follows from the definition and the previous corollary.

7.1.9. Corollary. *Every Borel measure on a perfectly normal space is regular.*

7.1.10. Lemma. *Let μ be a Baire measure on a topological space X. Then, for every $B \in \mathcal{B}a(X)$ and $\varepsilon > 0$, there exists a continuous function $\psi\colon X \to [0,1]$ such that*

$$\left|\int_X \psi\, d\mu - \mu(B)\right| < \varepsilon.$$

In addition, there exists a continuous function $\zeta\colon X \to [-1,1]$ such that

$$\left|\int_X \zeta\, d\mu - |\mu|(B)\right| < \varepsilon.$$

Proof. As in Corollary 7.1.8, it suffices to prove both assertions in the case $X = \mathbb{R}^\infty$. In this special case, one can find a closed set $Z \subset B$ and an open set $U \supset B$ with $|\mu|(U\backslash Z) < \varepsilon/2$. It remains to take a continuous function $\psi\colon X \to [0,1]$ that equals 1 on Z and 0 outside U (clearly, this is possible since $\mathbb{R}^\infty$ is a metrizable space). It is easy to see that ψ is a required function.

For the proof of the second assertion we take the Hahn decomposition $\mu = \mu^+ - \mu^-$ and find disjoint closed sets Z_1 and Z_2 such that $Z_1 \cup Z_2 \subset B$, $\mu^-(Z_1) = 0$, $\mu^+(Z_2) = 0$ and $\mu^+(B\backslash Z_1) + \mu^-(B\backslash Z_2) < \varepsilon/4$. In addition, we can find disjoint open sets $U_1 \supset Z_1$ and $U_2 \supset Z_2$ for which the inequality $|\mu|(U_1\backslash Z_1) + |\mu|(U_2\backslash Z_2) < \varepsilon/4$ holds. Finally, let us take a continuous function ζ equal to 1 on Z_1, -1 on Z_2 and 0 outside $U_1 \cup U_2$. Then

$$\left|\int_X \zeta\, d\mu - |\mu|(B)\right| \le |\mu|(U_1\backslash Z_1) + |\mu|(U_2\backslash Z_2) + |\mu|\big(B\backslash(Z_1 \cup Z_2)\big) < \varepsilon.$$

The lemma is proven. □

7.1.11. Lemma. *If a Borel or Baire measure μ is tight (or Radon), then every measure absolutely continuous with respect to μ is tight (respectively, Radon).*

PROOF. Let μ be a tight Borel or Baire measure and let $\nu = f \cdot \mu$, where $f \in \mathcal{L}^1(\mu)$. Then the measure ν is tight by the absolute continuity of the Lebesgue integral. Similarly, one proves that ν is Radon for a Radon measure μ. $\square$

In analogy with the case of scalar functions we shall say that a mapping of a measure space $(X, \mathcal{A}, \mu)$ to a topological space Y is μ-measurable if it is $\bigl(\mathcal{A}_\mu, \mathcal{B}(Y)\bigr)$-measurable. For example, if μ is a Borel measure, then $\mathcal{A}_\mu$ is the completion of $\mathcal{B}(X)$ with respect to μ.

We shall need the following modification of Egoroff's theorem.

7.1.12. Theorem. *Let Y be a separable metric space, $(X, \mathcal{A}, \mu)$ a space with a finite measure, and let $f_n\colon X \to Y$ be a sequence of mappings measurable with respect to the pair of the σ-algebras $\mathcal{A}$ and $\mathcal{B}(Y)$ and convergent μ-a.e. to a mapping f. Then, for every $\varepsilon > 0$, there exists a set $X_\varepsilon \in \mathcal{A}$ such that $|\mu|(X\backslash X_\varepsilon) < \varepsilon$ and the restrictions of the mappings f_n to the set X_ε converge uniformly to the restriction of f.*

PROOF. The arguments employed in the proof of Egoroff's theorem for real functions remain valid if we observe that

$$\bigl\{x\colon\ \varrho_Y\bigl(f_n(x), f_k(x)\bigr) \le r\bigr\} \in \mathcal{A}$$

for all $r \ge 0$, $n, k \in \mathbb{N}$, where ϱ_Y is the metric of Y. This follows by the fact that the mappings $x \mapsto \bigl(f_n(x), f_k(x)\bigr)$, $(X, \mathcal{A}) \to \bigl(Y \times Y, \mathcal{B}(Y) \otimes \mathcal{B}(Y)\bigr)$, are measurable and the function $(x, y) \mapsto \varrho_Y(x, y)$ is continuous, hence measurable with respect to the σ-algebra $\mathcal{B}(Y \times Y)$, which coincides with the σ-algebra $\mathcal{B}(Y) \otimes \mathcal{B}(Y)$ by the separability of Y. $\square$

Now we give a generalization of Lusin's classical theorem.

7.1.13. Theorem. *Let X be a topological space with a Radon measure μ, let Y be a complete separable metric space, and let $f\colon X \to Y$ be a μ-measurable mapping (i.e., $f^{-1}(B) \in \mathcal{B}_\mu(X)$ for all $B \in \mathcal{B}(Y)$). Then, for every $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset X$ such that $|\mu|(X\backslash K_\varepsilon) < \varepsilon$ and $f|_{K_\varepsilon}$ is continuous.*

If X is completely regular and Y is a Fréchet space, then there exists a continuous mapping $f_\varepsilon\colon X \to Y$ such that $|\mu|\bigl(x\colon f(x) \ne f_\varepsilon(x)\bigr) < \varepsilon$.

PROOF. We observe that if our claim is true for μ-measurable mappings f_n convergent to f a.e., then it is true for f as well. Indeed, for each n, we find a compact set K_n on which f_n is continuous with $|\mu|(X\backslash K_n) < \varepsilon 4^{-n}$ and use Egoroff's theorem to obtain a compact set K_0 with $|\mu|(X\backslash K_0) < \varepsilon/4$ on which convergence is uniform. Then we set $K_\varepsilon := \bigcap_{n \ge 0} K_n$. Now it suffices to prove our claim for mappings with countably many values because f can be uniformly approximated by such mappings. To this end, given

$k \in \mathbb{N}$, we partition Y into disjoint Borel parts B_j of diameter less than $1/j$, choose arbitrary elements $y_j \in B_j$ and define f_k as follows: $f_k = y_j$ on $f^{-1}(B_j)$. Every f_k is the pointwise limit of mappings with finitely many values, so it remains to note that if f assumes finitely many distinct values $c_1, \ldots, c_n$, then our assertion is true. Indeed, every set $A_j := f^{-1}(c_j)$ contains a compact set K_j with $|\mu|(A_j \backslash K_j) < \varepsilon/n$. The mapping f is continuous on $K_1 \cup \cdots \cup K_n$, since the sets K_j are disjoint and every point in K_j has a neighborhood that does not meet other sets K_i. The last assertion follows by Exercise 6.10.22, which enables us to extend continuous mappings from compact sets in completely regular spaces.

In the case where X is a metric space, the following alternative proof of Lusin's theorem was given in Dellacherie [**426**]. One may assume that X is compact and μ is a probability measure. The mapping $g\colon x \mapsto \bigl(x, f(x)\bigr)$ from X to $X \times Y$ is measurable with respect to μ. Hence there exists a compact set $S \subset X \times Y$ such that $\mu \circ g^{-1}(S) > 1 - \varepsilon$. Let K denote the projection of S on X. It is clear that K is compact and $\mu(K) = \mu \circ g^{-1}(S) > 1 - \varepsilon$. The mapping g on K takes values in the compact projection of S on Y, whence we obtain the continuity of f on K. Indeed, suppose a sequence of points $x_n \in K$ converges to a point $x_0 \in K$. The sequence $\bigl(x_n, f(x_n)\bigr) \in S$ contains a subsequence convergent to a point in S. This point can be only $\bigl(x_0, f(x_0)\bigr)$. Hence $\{f(x_n)\}$ converges to $f(x_0)$. $\square$

If we only require that K_ε be closed, then the first assertion of the theorem (with a similar proof) is valid for regular Borel measures. The second assertion (for Radon measures) extends to arbitrary separable metric spaces in the following weaker form: the mapping f_ε takes values in some separable Banach space, in which Y is isometrically embedded (for Y itself, there might be no such a mapping: it suffices to take $Y = \{0, 1\}$, $X = [0, 1]$, $f = I_{[0,1/2]}$). The case where Y is a Souslin space is considered in Corollary 7.4.4. A non-trivial generalization of this theorem is given in §7.14(ix).

7.2. τ-additive measures

There is one more important regularity property that is intermediate between the usual regularity and the Radon property.

7.2.1. Definition. *A Borel measure μ on a topological space X is called τ-additive (or τ-regular, τ-smooth) if for every increasing net of open sets $(U_\lambda)_{\lambda \in \Lambda}$ in X, one has the equality*

$$|\mu|\Bigl(\bigcup_{\lambda \in \Lambda} U_\lambda\Bigr) = \lim_{\lambda} |\mu|(U_\lambda). \tag{7.2.1}$$

If (7.2.1) is fulfilled for all nets with $\bigcup_\lambda U_\lambda = X$, then μ is called τ_0-additive (or weakly τ-additive).

It is clear from the definition that a measure μ is τ-additive precisely when its total variation $|\mu|$ is τ-additive (the same is true for the τ_0-additivity).

One can verify that any regular τ_0-additive Borel measure is τ-additive (see Exercise 7.14.66). On the other hand, there exist τ_0-additive measures that are not τ-additive.

7.2.2. Proposition. (i) *Every Radon measure is τ-additive.*

(ii) *Every τ-additive measure on a regular space is regular. In particular, every τ-additive measure on a compact space is Radon.*

(iii) *Every tight τ-additive measure is Radon.*

(iv) *Every Borel measure on a separable metric space X is τ-additive. Moreover, this is true if X is hereditary Lindelöf.*

PROOF. (i) Suppose we are given an increasing net of open sets U_λ, a Radon measure μ and $\varepsilon > 0$. We find a compact set $K \subset \bigcup_{\lambda\in\Lambda} U_\lambda$ with $|\mu|\bigl(\bigcup_{\lambda\in\Lambda} U_\lambda \backslash K\bigr) < \varepsilon$. It remains to take a finite subcover of K by sets U_λ. (ii) Suppose we are given a τ-additive measure μ on a regular space X. Denote by $\mathcal{E}$ the class of all Borel sets E in X such that

$$|\mu|(E) = \sup\bigl\{|\mu|(Z)\colon\ Z \subset E \text{ is closed}\bigr\} = \inf\bigl\{|\mu|(U)\colon\ U \supset E \text{ is open}\bigr\}.$$

We know that $\mathcal{E}$ is a σ-algebra (see the proof of Theorem 1.4.8). Hence it suffices to show that every open set U belongs to $\mathcal{E}$. By the regularity of X the set U can be represented in the form of the union of a family of open sets V such that $\overline{V} \subset U$. Therefore, U is covered by the directed family of open subsets of U consisting of finite unions of sets V of the above type, partially ordered by inclusion. Let $\varepsilon > 0$. Then, by the τ-additivity of μ, there exists a finite family of open sets $V_i \subset \overline{V_i} \subset U$, $i = 1, \ldots, n$, such that letting $W = \bigcup_{i=1}^n V_i$, we have $|\mu|(U\backslash W) < \varepsilon$. Then $|\mu|(U\backslash \overline{W}) < \varepsilon$. If X is compact, then by the regularity of μ we obtain the Radon property. (iii) The restrictions of a τ-additive measure to all compact subspaces are Radon, which by virtue of tightness yields the Radon property on the whole space. (iv) It suffices to use the countable additivity of our measure and the property that every open cover of any subset of X contains an at most countable subcover. □

7.2.3. Corollary. *Let two τ-additive measures μ and ν on a space X coincide on all sets from some class $\mathcal{U}$ that contains a base of the topology in X and is closed with respect to finite intersections. Then $\mu = \nu$.*

PROOF. Every open set U in X can be represented in the form of the union of a net of increasing open sets U_α that are finite unions of sets in $\mathcal{U}$. It is easily seen that $\mu(U_\alpha) = \nu(U_\alpha)$. By the τ-additivity we obtain $\mu(U) = \nu(U)$. By Lemma 7.1.2 both measures coincide on all Borel sets. □

We note that Example 7.1.6 gives a τ-additive measure that is not Radon. Let us consider another interesting example.

7.2.4. Example. Let $X = [0, 1)$ be the Sorgenfrey interval (with its topology generated by all semiclosed intervals $[a, b) \subset X$). Then X is hereditary Lindelöf and all Borel sets in X are the same as in the usual topology of the interval (since every open set in the Sorgenfrey topology is an at most countable union of intervals $[a, b)$). The usual Lebesgue measure on this space is regular and τ-additive, but is not Radon, since compact subsets in X are at most countable.

7.2.5. Proposition. *Let μ be a regular Borel measure. Then the following conditions are equivalent*:

(i) *the measure μ is τ-additive*;

(ii) *for every increasing net $\{U_\alpha\}$ of open sets with union U one has the equality*

$$\mu(U) = \lim_\alpha \mu(U_\alpha); \tag{7.2.2}$$

(iii) *for every decreasing net $\{Z_\alpha\}$ of closed sets with intersection Z one has the equality*

$$\mu(Z) = \lim_\alpha \mu(Z_\alpha); \tag{7.2.3}$$

(iv) *for every decreasing net $\{Z_\alpha\}$ of closed sets with $\bigcap_\alpha Z_\alpha = Z = \varnothing$, one has equality* (7.2.3).

Proof. Relations (7.2.2) and (7.2.3) are equivalent for any measure and are fulfilled for τ-additive measures. It follows by (7.2.3) and the regularity of μ that the measures μ^+ and μ^- satisfy (7.2.3), hence satisfy (7.2.2). Therefore, (7.2.1) is fulfilled, i.e., μ is τ-additive. Thus, (i)–(iii) are equivalent. Finally, let (iv) be fulfilled and let $\{Z_\alpha\}$ be a decreasing net of closed sets. Let us fix $\varepsilon > 0$ and take an open set U such that $Z = \cap_\alpha Z_\alpha \subset U$ and $|\mu|(U\backslash Z) < \varepsilon$. Then the closed sets $Z_\alpha\backslash U$ decrease to the empty set, so $\lim\limits_\alpha \mu(Z_\alpha\backslash U) = 0$. It remains to observe that we have the inequalities $\mu(Z_\alpha) = \mu(Z_\alpha\backslash Z) + \mu(Z)$ and $|\mu(Z_\alpha\backslash Z) - \mu(Z_\alpha\backslash U)| \le |\mu|(U\backslash Z) < \varepsilon$. $\square$

We recall that a function f on a topological space X is called lower semicontinuous if for all $c \in \mathbb{R}^1$, the sets $\{x\colon\ f(x) > c\}$ are open (see Engelking [**532**, 1.7.14]). It is clear that such functions are Borel. Note that the pointwise limit of an increasing net of lower semicontinuous functions is lower semicontinuous as well. A function f is called upper semicontinuous if all sets $\{f < c\}$ are open, i.e., the function $-f$ is lower semicontinuous.

7.2.6. Lemma. *Let μ be a regular τ-additive* (*for example, Radon*) *measure on a topological space X and let $\{f_\alpha\}$ be an increasing net of lower semicontinuous nonnegative functions such that the function $f = \lim\limits_\alpha f_\alpha$ is bounded. Then*

$$\lim_\alpha \int_X f_\alpha(x)\,\mu(dx) = \int_X f(x)\,\mu(dx).$$

Proof. One can assume that μ is nonnegative; the general case is obtained from the Jordan–Hahn decomposition. In addition, one can assume that $f < 1$. Set

$$f_{\alpha,n} = \frac{1}{n}\sum_{k=1}^n I_{\{f_\alpha > (k-1)/n\}}, \quad f_n = \frac{1}{n}\sum_{k=1}^n I_{\{f > (k-1)/n\}}.$$

By the lower semicontinuity of the functions f_α, the function f is lower semicontinuous as well. Thus, the sets $\{f_\alpha > (k-1)/n\}$ are open and for any fixed n and k, they form a net increasing to the open set $\{f > (k-1)/n\}$. By the τ-additivity we have $\lim_\alpha \mu\bigl(f_\alpha > (k-1)/n\bigr) = \mu\bigl(f > (k-1)/n\bigr)$. Hence, for every n, we have

$$\lim_\alpha \int_X f_{\alpha,n}\, d\mu = \int_X f_n\, d\mu.$$

In view of the estimates $|f_{\alpha,n} - f_\alpha| \le 1/n$, $|f_n - f| \le 1/n$ this completes the proof. □

7.2.7. Corollary. *If μ is a regular τ-additive measure on a topological space X and $\{f_\alpha\} \subset C_b(X)$ is a net decreasing to zero, then*

$$\lim_\alpha \int_X f_\alpha(x)\, \mu(dx) = 0.$$

PROOF. Let α_0 be any fixed element. We observe that the net $f_{\alpha_0} - f_\alpha$, $\alpha \ge \alpha_0$, increases to f_{α_0} and consists of nonnegative functions. It remains to apply the above lemma and the additivity of integral. □

Lemma 10.5.5 in Chapter 10 contains a close result for not necessarily lower semicontinuous functions contained in the image of a lifting of an arbitrary measure μ.

7.2.8. Lemma. *Let X be a completely regular space and let μ be a τ-additive measure on X. Then, for every $B \in \mathcal{B}(X)$ and $\varepsilon > 0$, there exists a continuous function $\psi\colon X \to [0,1]$ such that*

$$\left| \int_X \psi\, d\mu - \mu(B) \right| < \varepsilon.$$

In addition, there exists a continuous function $f\colon X \to [-1,1]$ such that

$$\left| \int_X f\, d\mu - |\mu|(B) \right| < \varepsilon.$$

In particular, this is true for any Radon measure.

PROOF. By Lemma 7.1.10, it suffices to show that for every $\delta > 0$, there exists a set $B_\delta \in \mathcal{B}a(X)$ such that $|\mu|(B_\delta \triangle B) < \delta$. Since the measure μ is regular by the complete regularity of X and τ-additivity, one can find an open set $G \supset B$ with $|\mu|(G\backslash B) < \delta/2$. Due to the complete regularity of X, the set G is a union of an increasing net of functionally open sets. By using the τ-additivity of μ once again, we find a functionally open set $B_\delta \subset G$ with $|\mu|(G\backslash B_\delta) < \delta/2$. □

Let us introduce the following notation.

Notation. Given a topological space X, we shall use throughout the following symbols:

$\mathcal{M}_\sigma(X)$ is the set of all Baire measures,

$\mathcal{M}_{\mathcal{B}}(X)$ is the set of all Borel measures,

$\mathcal{M}_r(X)$ is the set of all Radon measures,
$\mathcal{M}_t(X)$ is the set of all tight Baire measures,
$\mathcal{M}_\tau(X)$ is the set of all τ-additive Borel measures.

The symbols $\mathcal{M}^+_\sigma(X)$, $\mathcal{M}^+_\mathcal{B}(X)$, $\mathcal{M}^+_r(X)$, $\mathcal{M}^+_t(X)$, $\mathcal{M}^+_\tau(X)$ stand for the corresponding classes of nonnegative measures. Finally, the symbol $\mathcal{P}$ will denote the subclass of probability measures in the respective classes.

An important application of the property of τ-additivity concerns the concept of support of a Borel measure. For every Borel measure μ, one can consider the closed set S_μ that is the intersection of all closed sets of full μ-measure (i.e., the complements of sets of $|\mu|$-measure zero). If this set also has full measure, then it is called the *support* of μ and is denoted by $\operatorname{supp}\mu$ (in this case we say that the measure μ has support). The measure μ on a compact space constructed in Example 7.1.3 has no support (one has $S_\mu = \{\omega_1\}$).

7.2.9. Proposition. *Every τ-additive measure has support. In particular, every Radon measure has support and every Borel measure on a separable metric space has support.*

Proof. By the τ-additivity, the union of any family of open sets of measure zero has measure zero. $\square$

We recall that the weight of a metric space (X, d) is the minimal cardinality of a topology base in X (and also the minimal cardinality $\mathfrak{m}$ with the property that every set $S \subset X$ with $\inf_{x,y\in S, x\neq y} d(x, y) > 0$ is of cardinality at most $\mathfrak{m}$; see Engelking [**532**, Theorem 4.1.15]).

7.2.10. Proposition. *The weight of a metric space (X, d) is nonmeasurable in the sense of* §1.12(x) *precisely when every Borel measure on X is τ-additive* (*and then is Radon if X is complete*). *An equivalent condition*: *every Borel measure on X has support.*

Proof. Let the weight $\mathfrak{m}$ of X be measurable. Then X contains a set S of measurable cardinality $\mathfrak{m}' \le \mathfrak{m}$ such that $d(x, y) \ge r > 0$ for all $x \neq y$ in S. Indeed, by Zorn's lemma, for every $n \in \mathbb{N}$, there is a maximal family M_n of points such that $d(x, y) \ge n^{-1}$ whenever $x, y \in M_n$. The cardinality of some family M_n must be measurable, since otherwise the cardinality of their union would be nonmeasurable, which contradicts the above-cited theorem. There is a probability measure on the class of all subsets in S vanishing on all singletons. Its extension to $\mathcal{B}(X)$ has no support, since the sets $S\backslash\{x\}$ are closed and have measure 1. Conversely, suppose that there is a Borel probability measure μ on X that is not τ-additive. Then μ has no support, for its support would be separable. Indeed, any nonseparable metric space contains an uncountable collection of disjoint balls, which cannot all be of positive measure. Therefore, we obtain a family Γ of open sets of μ-measure zero such that their union has a positive μ-measure. According to Stone's theorem (see Engelking [**532**, Theorem 4.4.1]), there is a sequence Γ_n of collections of open subsets of sets in Γ such that for every fixed n, the sets in Γ_n are pairwise disjoint, and the union of sets in all collections Γ_n coincides with the union of all sets in Γ. Hence there is n such that the union of sets in Γ_n has a positive measure. Thus, since the sets in Γ_n also have measure zero (as subsets of elements of Γ), we may assume that Γ consists of disjoint sets. On the set of all subsets of Γ we obtain a nonzero measure ν by setting $\nu(E) = \mu\bigl(\bigcup_{K\in E} K\bigr)$, $E \subset \Gamma$. This measure is well-defined due to the disjointness of sets in Γ. All one-element subsets in Γ have ν-measure zero. This shows that the cardinality of X is measurable. Finally, if a Borel measure μ on X has support, then, as noted above, this support is separable, hence μ is τ-additive (and is Radon if X is complete). $\square$

7.3. Extensions of measures

In this section, we discuss several important questions related to extensions of measures to larger σ-algebras. In particular, we shall see that every tight Baire measure can be extended to a Radon measure. Such constructions are efficient in the study of measures on large functional spaces such as the space of all functions on an interval. Before proving theorems on extensions of tight measures, let us consider the following simple example of a tight Baire measure that has a Radon extension to the Borel σ-algebra, but this extension cannot be obtained by means of Lebesgue's completion of $\mathcal{B}a(X)$.

7.3.1. Example. Let $X = \mathbb{R}^T$, where T is an uncountable set (for example, an interval of the real line), let x_0 be any element in X (for example, the identically zero function), and let ν be the measure on the σ-algebra $\mathcal{B}a(X)$ defined by the formula: $\nu(B) = 1$ if $x_0 \in B$ and $\nu(B) = 0$ otherwise (i.e., ν is Dirac's measure at x_0). It is clear that this measure is tight and by the same formula can be extended to $\mathcal{B}(X)$. However, the one-point set x_0 is nonmeasurable with respect to Lebesgue's completion of the measure ν on $\mathcal{B}a(X)$. Indeed, otherwise this set would be a union of a set in $\mathcal{B}a(X)$ and a set of outer measure zero with respect to ν on $\mathcal{B}a(X)$, which is impossible, since no singleton is Baire in our space, whereas the point x_0 has outer measure 1.

The next theorem and its corollary are very useful in applications. The proof employs the inner measure μ_* generated by a nonnegative additive set function μ on an algebra $\mathcal{A}$ by the formula

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

in accordance with the general construction from §1.12(viii).

7.3.2. Theorem. *Suppose an algebra $\mathcal{A}$ of subsets of a Hausdorff space X contains a base of the topology. Let μ be a regular additive set function of bounded variation on $\mathcal{A}$.*

(i) *Suppose that μ is tight. Then it admits a unique extension to a Radon measure on X.*

(ii) *Suppose that X is regular and that for every increasing net $\{U_\alpha\}$ of open sets in $\mathcal{A}$ with $X = \bigcup_\alpha U_\alpha$, we have $|\mu|(X) = \lim\limits_\alpha |\mu|(U_\alpha)$. Then μ admits a unique extension to a τ-additive measure on $\mathcal{B}(X)$.*

If μ is nonnegative, then in both cases the corresponding extensions for all $B \in \mathcal{B}(X)$ are given by the formula

$$\widehat{\mu}(B) = \inf\bigl\{\mu_*(U)\colon\ U \textit{ is open in } X \textit{ and } B \subset U\bigr\}. \tag{7.3.1}$$

PROOF. It suffices to prove the theorem for nonnegative measures, since the positive and negative parts of any set function μ with the properties from (i) or (ii) possess those properties as well. First we verify claim (ii), which is more difficult, and then explain the changes to be made for the proof of (i). Let us show that

$$\lim_\alpha \mu(U_\alpha) = \mu_*(U) \tag{7.3.2}$$

for every net of increasing open sets $U_\alpha \in \mathcal{A}$ with $\bigcup\limits_\alpha U_\alpha = U$. Indeed, otherwise

$$\mu_*(U) - \lim_\alpha \mu(U_\alpha) \geq \varepsilon > 0.$$

By the regularity of μ on the algebra $\mathcal{A}$, there exists a closed set $Z \subset U$ from $\mathcal{A}$ with $\mu(Z) > \mu_*(U) - \varepsilon/2$. Let $W = X\backslash Z$. Then

$$\begin{aligned} \lim_\alpha \mu(U_\alpha \cup W) &\leq \lim_\alpha \mu(U_\alpha) + \mu(X) - \mu(Z) \\ &\leq \lim_\alpha \mu(U_\alpha) + \mu(X) - \mu_*(U) + \varepsilon/2 \leq \mu(X) - \varepsilon/2 < \mu(X), \end{aligned}$$

which contradicts the equality $\mu(X) = \lim\limits_\alpha \mu(U_\alpha \cup W)$ that follows by the equality $X = \bigcup\limits_\alpha (U_\alpha \cup W)$ due to the τ_0-additivity of μ.

Now we show that

$$\lim_\alpha \mu_*(U_\alpha) = \mu_*(U) \tag{7.3.3}$$

for every net of arbitrary open sets U_α increasing to U. We verify first that

$$\lim_\alpha \mu_*(U_\alpha) \geq \mu(V) \tag{7.3.4}$$

for any open set $V \subset U$ in $\mathcal{A}$. To this end, we denote by $\mathcal{W}$ the class of all open sets W in $\mathcal{A}$ such that $W \subset U_\alpha$ for some α. It is clear that $\mathcal{W}$ is a directed (by increasing) family of sets with union U. According to (7.3.2) we have

$$\mu_*(U) = \sup\{\mu(W),\ W \in \mathcal{W}\}.$$

Since $V = \bigcup\limits_{W\in\mathcal{W}} (V \cap W)$, we obtain similarly

$$\mu(V) = \mu_*(V) = \sup\{\mu(V \cap W),\ W \in \mathcal{W}\}. \tag{7.3.5}$$

By the definition of $\mathcal{W}$ we have $V \cap W \subset U_\alpha$ for some α, whence $\mu(V \cap W) \leq \mu(W) \leq \mu_*(U_\alpha)$. Therefore, $\mu(V \cap W) \leq \lim\limits_\alpha \mu_*(U_\alpha)$. By (7.3.5) we arrive at (7.3.4). Taking the supremum over all open sets $V \subset U$ in $\mathcal{A}$, we obtain from (7.3.2) that $\lim\limits_\alpha \mu_*(U_\alpha) \geq \mu_*(U)$. Since $\mu_*(U_\alpha) \leq \mu_*(U)$, we arrive at (7.3.3).

Let us verify two other properties of μ_*: if U_1 and U_2 are open, then

$$\mu_*(U_1 \cup U_2) \le \mu_*(U_1) + \mu_*(U_2), \tag{7.3.6}$$

and if $U_1 \cap U_2 = \varnothing$, then

$$\mu_*(U_1 \cup U_2) = \mu_*(U_1) + \mu_*(U_2). \tag{7.3.7}$$

Indeed, by the hypothesis of the theorem, there exist two nets of increasing open sets W_α^1 and W_β^2 from $\mathcal{A}$ such that $U_1 = \bigcup_\alpha W_\alpha^1$ and $U_2 = \bigcup_\beta W_\beta^2$. Then

$$\mu(W_\alpha^1 \cup W_\beta^2) \le \mu(W_\alpha^1) + \mu(W_\beta^2) \le \mu_*(U_1) + \mu_*(U_2).$$

By using (7.3.2) we obtain $\mu_*(W_\alpha^1 \cup U_2) \le \mu_*(U_1) + \mu_*(U_2)$ for every fixed α. Now (7.3.6) follows from (7.3.3). Similarly, we verify (7.3.7).

Let us now consider the set function

$$\nu(A) = \inf\bigl\{\mu_*(U)\colon\ U \text{ is open and } A \subset U\bigr\}, \quad A \subset X.$$

It follows from (7.3.3) and (7.3.6) that ν is a Carathéodory outer measure (see Chapter 1). Therefore,

$$\mathcal{A}_\nu = \bigl\{A\colon\ \nu(A\cap B) + \nu\bigl((X\backslash A)\cap B\bigr) = \nu(B),\ \forall\, B \subset X\bigr\}$$

is a σ-algebra, on which the set function ν is countably additive. We show that $\mathcal{B}(X) \subset \mathcal{A}_\nu$. It suffices to verify that every open set U belongs to $\mathcal{A}_\nu$. To this end, it suffices to establish the estimate

$$\nu(U\cap B) + \nu\bigl((X\backslash U)\cap B\bigr) \le \nu(B) \tag{7.3.8}$$

for every $B \subset X$ (the reverse inequality follows by (7.3.6)). Suppose that B is open. Then (7.3.8) is written in the form

$$\mu_*(U\cap B) + \nu\bigl((X\backslash U)\cap B\bigr) \le \mu_*(B). \tag{7.3.9}$$

By the regularity of X there exists a net of increasing open sets U_α with $U = \bigcup_\alpha U_\alpha$ and $Z_\alpha := \overline{U_\alpha} \subset U$ for all α, where $\overline{U_\alpha}$ denotes the closure of U_α. Then

$$B = (B\cap U_\alpha) \cup \bigl(B\cap (X\backslash U_\alpha)\bigr) \supset (B\cap U_\alpha) \cup \bigl(B\cap (X\backslash Z_\alpha)\bigr).$$

Since the set $B\cap (X\backslash Z_\alpha)$ is open, we obtain from (7.3.7) that

$$\mu_*(B) \ge \mu_*(B\cap U_\alpha) + \mu_*\bigl(B\cap (X\backslash Z_\alpha)\bigr).$$

We observe that $\mu_*\bigl(B\cap (X\backslash Z_\alpha)\bigr) \ge \nu\bigl(B\cap (X\backslash U)\bigr)$, since $B\cap (X\backslash U)$ belongs to $B\cap (X\backslash Z_\alpha)$ and the latter set is open. Thus,

$$\mu_*(B) \ge \mu_*(B\cap U_\alpha) + \nu\bigl(B\cap (X\backslash U)\bigr).$$

By using (7.3.2), we obtain (7.3.9). Now let B be arbitrary and let $W \supset B$ be open. Then $\mu_*(W) \ge \nu(B\cap U) + \nu\bigl(B\cap (X\backslash U)\bigr)$. Therefore, we have (7.3.8). Thus, $U \in \mathcal{A}_\nu$ and hence $\mathcal{B}(X) \subset \mathcal{A}_\nu$. It remains to take for the desired extension $\widehat{\mu}$ the restriction of ν to $\mathcal{B}(X)$. The measure $\widehat{\mu}$ is τ-additive by (7.3.3), since $\nu(U) = \mu_*(U)$ for every open U. If U is open and belongs to $\mathcal{A}$, we have $\nu(U) = \mu(U)$. Let $A \in \mathcal{A}$. Given $\varepsilon > 0$, by the regularity of μ we find an open set $U \in \mathcal{A}$ such that $A \subset U$ and $\mu(A) > \mu(U) - \varepsilon$, i.e.,

$\mu(A) > \nu(U) - \varepsilon \geq \nu(A) - \varepsilon$. Hence $\mu(A) \geq \nu(A)$. Then $\mu(X\backslash A) \geq \nu(X\backslash A)$, as $X\backslash A \in \mathcal{A}$. Therefore, $\mu(A) = \nu(A)$. Note that $\mathcal{A}$ may not belong to $\mathcal{B}(X)$, but is contained in the completion of $\mathcal{A} \cap \mathcal{B}(X)$. The uniqueness of extension follows by Corollary 7.2.3.

We now proceed to assertion (i). In the case of a regular space it follows by the already proven assertion. In the general case, the above reasoning can be slightly modified. We observe that in the proof of existence, the regularity of X was only used in order to verify that $\mathcal{B}(X) \subset \mathcal{A}_\nu$. Hence, by taking into account that our tight measure satisfies the condition indicated in (ii), we conclude that the reasoning preceding the above-mentioned verification remains valid. In order to show that also the inclusion $\mathcal{B}(X) \subset \mathcal{A}_\nu$ is still true, we observe that $\mathcal{A}_\nu$ contains all open sets U such that $X\backslash U$ is compact. Then $U = \bigcup_\alpha U_\alpha$ for some net of increasing open sets U_α with $Z_\alpha = \overline{U_\alpha} \subset U$. This follows from the fact that every point in U and the compact complement to U have disjoint neighborhoods. Thus, the subsequent reasoning of the previous step remains valid and $U \in \mathcal{A}_\nu$. Hence all compact sets are in $\mathcal{A}_\nu$. It remains to show that every closed set Z is contained in $\mathcal{A}_\nu$. Let us take a sequence of compact sets K_n such that $\mu^*(K_n) > \mu(X) - 1/n$. Let $K = \bigcup_{n=1}^\infty K_n$. We observe that $\nu(K_n) = \mu^*(K_n)$. Indeed, every open set V containing K_n contains an open set $W \in \mathcal{A}$ that contains K_n, since every point $x \in K_n$ has a neighborhood $W_x \subset V$ from $\mathcal{A}$, and the cover obtained in this way has a finite subcover. Therefore, $\mu^*(K_n) \leq \mu(W) \leq \mu_*(V)$, whence $\mu^*(K_n) \leq \nu(K_n)$. On the other hand, by the regularity of μ, one has $\mu^*(K_n) = \inf \mu(W)$, where inf is taken over all open $W \supset K_n$ in $\mathcal{A}$. Since $\mu(W) \geq \nu(K_n)$, this yields the estimate $\mu^*(K_n) \geq \nu(K_n)$. It follows by the above on account of completeness of the σ-algebra $\mathcal{A}_\nu$ that $\nu(X\backslash K) = 0$. It remains to observe that Z coincides up to a ν-measure zero set with the set $\bigcup_{n=1}^\infty (Z \cap K_n)$, which belongs to $\mathcal{A}_\nu$ by the above, since the sets $Z \cap K_n$ are compact. The uniqueness of extension follows from the fact that every two extensions coincide on all finite unions of elements of a base from $\mathcal{A}$, hence coincide on all compact sets because every open neighborhood of a compact set contains a neighborhood that is a finite union of elements of the base. $\square$

7.3.3. Corollary. *Let X be a completely regular space. Then*:

(i) *every tight Baire measure μ on X admits a unique extension to a Radon measure*;

(ii) *every Baire measure μ on X that is τ_0-additive on $\mathcal{B}a(X)$ in the sense that $|\mu|(X) = \sup_\alpha |\mu|(U_\alpha)$ for all increasing nets of functionally open sets U_α such that $X = \bigcup_\alpha U_\alpha$, admits a unique extension to a τ-additive Borel measure.*

Proof. According to Corollary 7.1.8, every Baire is regular. Since X is completely regular, functionally open sets form a base of the topology. $\square$

7.3.4. Corollary. *Let X be a σ-compact completely regular space. Then every Baire measure on X has a unique extension to a Radon measure.*

PROOF. It suffices to observe that on a σ-compact space, every Baire measure is tight. □

Now we are able to reinforce Corollary 7.3.3.

7.3.5. Corollary. *Let X be a completely regular space and let Γ be a family of continuous functions on X separating the points in X. Then, every tight measure μ on the σ-algebra $\sigma(\Gamma)$ generated by Γ admits a unique extension to a Radon measure on X. Moreover, the same is true if μ is a regular and tight additive set function of bounded variation on the algebra $\mathfrak{A}(\Gamma)$ generated by Γ.*

PROOF. As above, it is sufficient to consider nonnegative measures, passing to the Jordan decomposition (in the case of the algebra $\mathfrak{A}(\Gamma)$ this is possible due to our assumption of boundedness of variation). This corollary differs from the main theorem in that $\mathfrak{A}(\Gamma)$ may not contain a base of the topology (for example, this is the case if $X = l^2$ with the usual Hilbert norm and $\Gamma = (l^2)^*$). It is clear that the main theorem applies to the space X with the topology τ generated by Γ (i.e., the weakest topology with respect to which all functions in Γ are continuous). Note that all compact sets in the initial topology are compact in the topology τ. Let μ_τ denote the unique Radon extension of μ to (X, τ). We take a set K_n with $\mu^*(K_n) > \mu(X) - 1/n$ that is compact in the initial topology. By the above theorem we obtain $\mu_\tau(K_n) = \mu^*(K_n) > \mu(X) - 1/n$. Hence the measure μ_τ is concentrated on the set $X_0 = \bigcup_{n=1}^\infty K_n$. We shall consider X_0 with the initial topology, in which it is σ-compact. According to Proposition 6.10.8, every Baire set B in the space X_0 has the form $B = X_0 \cap E$, $E \in \sigma(\Gamma)$. Therefore, the restriction μ_0 of the measure μ_τ to $\mathcal{B}a(X_0)$ is well-defined. By using the previous corollary we extend μ_0 to a Radon measure $\widehat{\mu}$ on X_0 and then to a Radon measure on all of X by setting $\widehat{\mu}(X \backslash X_0) = 0$. It is clear that $\widehat{\mu}$ is the required extension. Let us verify its uniqueness. Let μ_1 and μ_2 be two Radon measures that coincide on the algebra $\mathfrak{A}(U)$ generated by Γ. Then these measures coincide on all compact sets in the topology τ, hence on all compact sets in the initial topology, whence we have $\mu_1 = \mu_2$. □

7.3.6. Corollary. *Let X be a locally convex space with the σ-algebra $\sigma(X^*)$ and let μ be a tight measure on $\sigma(X^*)$. Then μ has a unique extension to a Radon measure on X. The same is true for every tight regular additive set function of bounded variation on the algebra generated by X^* (or by any subspace in X^* separating the points in X).*

PROOF. The set X^* separates the points in X. □

7.3.7. Example. Let X be a normed space and let μ be a measure on the σ-algebra $\mathcal{E}$ in the space X^* generated by the elements of X. Then μ has a unique extension to a Radon measure on X^* with the weak* topology.

PROOF. By the Banach–Alaoglu theorem the balls in X^* are compact in the weak* topology. Hence the measure μ is tight. □

7.3.8. Example. Let X be the product of the continuum of copies of $[0,1]$. Then Dirac's measure δ at zero considered on the Baire σ-algebra of X has a Borel extension that is not Radon.

PROOF. It is known that the ordinal interval $(0,\omega_1)$ is homeomorphic to a subset of X (see Engelking [**532**, Theorem 2.3.23]). By using this homeomorphism we transport the Dieudonné measure μ to X and obtain a non-regular Borel measure μ on X that assumes only the values 0 and 1. Its Baire restriction has a unique Radon extension μ_0. Then μ_0 must be Dirac's measure at some point x_0. Clearly, μ is a non-regular Borel extension of δ_{x_0}. As shown by Keller [**971**], there is a homeomorphism h of X such that $h(x_0)=0$. In fact, Keller proved the result for the countable power, but the uncountable case follows at once by splitting $[0,1]^{\mathfrak{c}}$ into a product of countable powers of $[0,1]$. Now $\mu\circ h^{-1}$ is a non-regular Borel extension of δ_0. □

The results obtained above enable us to identify tight Baire measures on a completely regular space X with their (unique) Radon extensions.

We recall that Lebesgue's extension may not be sufficient for obtaining the extension guaranteed by Theorem 7.3.2 (see Example 7.3.1).

Finally, there exist Baire measures without Borel extensions at all.

7.3.9. Example. Let $I=[0,1)$ be the Sorgenfrey interval with the topology from Example 6.1.19 and let $X=I^2$ be equipped with the product topology (i.e., X is the set $[0,1)^2$ in the Sorgenfrey plane). According to Exercise 6.10.81, the set $T:=\{(t,s)\in X\colon\ t+s=1\}$ is Baire and for any $B\in\mathcal{B}a(X)$, the intersection $B\cap T$ is Borel with respect to the usual topology of the plane. Hence the formula

$$\mu(B):=\lambda\bigl(t\in[0,1)\colon\ (t,1-t)\in B\bigr),$$

where λ is Lebesgue measure on $[0,1)$, defines a Baire probability measure on X. Every point $x=(u,1-u)\in T$ is measurable with respect to μ and has measure zero because x belongs to the Baire set

$$E(x):=X\cap[u,u+1)\times[1-u,2-u),$$

for which we have $\mu(E)=0$, since if $t\in[0,1)$ and $(t,1-t)\in E$, then $t=u$. If the measure μ could be extended to a countably additive measure on $\mathcal{B}(X)$, then all subsets of T would be measurable with respect to the extension, which along with the equality $\mu(\{x\})=0$, $x\in T$, would give a probability measure on the set of all subsets of $[0,1)$ vanishing on all singletons. According to Corollary 1.12.41, this is impossible under the continuum hypothesis. Exercise 7.14.69 proposes to construct analogous examples without use of the continuum hypothesis; moreover, one can even take a locally compact space for X.

Regarding extensions of Baire measures, see also §7.14(iii). The following non-trivial reinforcement of assertion (i) of Corollary 7.3.3 is easily deduced from a deep result presented in Exercise 7.14.84. It enables us to drop the complete regularity assumption on X.

7.3.10. Theorem. *Any tight Baire measure on a Hausdorff space has a Radon extension.*

One more construction of Radon extensions was given in Henry [**812**].

7.3.11. Theorem. *Let X be a Hausdorff space, let $\mathcal{A}$ be a subalgebra in $\mathcal{B}(X)$, and let μ be a nonnegative additive set function on $\mathcal{A}$ satisfying the following condition: for every $A \in \mathcal{A}$ and every $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset A$ such that $\mu^*(A\backslash K_\varepsilon) < \varepsilon$. Then μ extends to a Radon measure on X.*

PROOF. Let us consider the set of all pairs $(\mathcal{E}, \eta)$, where $\mathcal{E}$ is a subalgebra in $\mathcal{B}(X)$ containing $\mathcal{A}$ and η is a nonnegative additive function on $\mathcal{E}$ that extends μ and possesses on $\mathcal{E}$ the same property of inner compact regularity as μ has on $\mathcal{A}$. Such pairs are partially ordered by the following relation: $(\mathcal{E}_1, \eta_1) \leq (\mathcal{E}_2, \eta_2)$ if $\mathcal{E}_1 \subset \mathcal{E}_2$ and $\eta_2|_{\mathcal{E}_1} = \eta_1$. It is clear that every linearly ordered part $(\mathcal{E}_\alpha, \eta_\alpha)$ of this set has an upper bound. Indeed, the union $\mathcal{E}$ of all algebras $\mathcal{E}_\alpha$ is an algebra (because for any two such algebras, one of them is contained in the other), and $\mathcal{A} \subset \mathcal{E}$. The function η on $\mathcal{E}$ defined by the equality $\eta(E) = \eta_\alpha(E)$ if $E \in \mathcal{E}_\alpha$ is well-defined for the same reason, is additive and extends μ. Finally, it is clear that η has the required approximation property. By Zorn's lemma, there is a maximal element $(\mathcal{B}, \nu)$. We show that $\mathcal{B} = \mathcal{B}(X)$ and that ν is a Radon measure. Let us observe that ν is countably additive on $\mathcal{B}$ due to the existence of an approximating compact class. Therefore, one can extend ν to the σ-algebra $\sigma(\mathcal{B})$, and the extension is inner compact regular as well, which is seen from the proof of assertion (iii) in Proposition 1.12.4. By the maximality of $\mathcal{B}$ this shows that $\mathcal{B}$ itself is a σ-algebra and ν is a measure. Suppose that there is a closed set Z not belonging to $\mathcal{B}$. We shall obtain a contradiction if we prove the existence of a measure $\widetilde{\nu}$ that extends ν to $\mathcal{B}_0 := \sigma(\mathcal{B} \cup \{Z\})$ and is inner compact regular in the same sense as μ. Set

$$\widetilde{\nu}(C) = \nu^*(C \cap Z) + \nu_*\bigl(C \cap (X\backslash Z)\bigr).$$

According to the proof of Theorem 1.12.14, $\mathcal{B}_0$ is the class of all sets of the form $C = (A \cap Z) \cup \bigl(B \cap (X\backslash Z)\bigr)$, where $A, B \in \mathcal{B}$, and $\widetilde{\nu}$ is a measure on $\mathcal{B}_0$ extending ν. We verify the inner compact regularity of $\widetilde{\nu}$. Given $A \in \mathcal{B}_0$ and $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset A$ with $\nu^*(A\backslash K_\varepsilon) < \varepsilon$. Then

$$\widetilde{\nu}^*\bigl((A \cap Z)\backslash(K_\varepsilon \cap Z)\bigr) \leq \widetilde{\nu}^*(A\backslash K_\varepsilon) \leq \nu^*(A\backslash K_\varepsilon) \leq \varepsilon.$$

Let $\widetilde{Z} \in \mathcal{B}$ be a measurable envelope of Z with respect to the measure ν (see §1.12(iv)). Then $\widetilde{\nu}(\widetilde{Z}) = \widetilde{\nu}(Z)$, since $\nu_*(\widetilde{Z}\backslash Z) = 0$ by the definition of a measurable envelope. Since $B \cap (X\backslash\widetilde{Z}) \in \mathcal{B}$, there exists a compact set $S_\varepsilon \subset B \cap (X\backslash\widetilde{Z})$ with $\nu^*\bigl(\bigl(B \cap (X\backslash\widetilde{Z})\bigr)\backslash S_\varepsilon\bigr) < \varepsilon$. Since $Z \subset \widetilde{Z}$, one has $S_\varepsilon \subset B \cap (X\backslash Z)$. On account of the equality $\widetilde{\nu}(\widetilde{Z}) = \widetilde{\nu}(Z)$ we obtain

$$\begin{aligned}\widetilde{\nu}\bigl(B \cap (X\backslash Z)\bigr) &= \widetilde{\nu}\bigl(B \cap (X\backslash\widetilde{Z})\bigr) = \nu\bigl(B \cap (X\backslash\widetilde{Z})\bigr)\\ &= \nu^*\bigl(\bigl(B \cap (X\backslash\widetilde{Z})\bigr)\backslash S_\varepsilon\bigr) + \nu_*(S_\varepsilon) < \nu_*(S_\varepsilon) + \varepsilon.\end{aligned}$$

Therefore, $(\widetilde{\nu})^*\bigl((B\cap(X\backslash Z))\backslash S_\varepsilon\bigr) < \varepsilon$. Finally, $K_\varepsilon\cap Z$ is compact, $(K_\varepsilon\cap Z)\cup S_\varepsilon$ is compact as well, $(K_\varepsilon \cap Z) \cup S_\varepsilon \subset C$, and $(\widetilde{\nu})^*\bigl(C\backslash((K_\varepsilon \cap Z) \cup S_\varepsilon)\bigr) < 2\varepsilon$ as required. □

The difference between this theorem and the previous results is that the algebra $\mathcal{A}$ may be very small, but in place of tightness a stronger assumption is imposed. An analogous theorem holds for infinite measures as well (see [**812**]). Extensions of measures are also discussed in §9.8.

7.4. Measures on Souslin spaces

7.4.1. Theorem. *Let μ be a Borel measure on a Hausdorff space X. Then every Souslin set in X is measurable with respect to μ, i.e., belongs to $\mathcal{B}(X)_\mu$.*

PROOF. We know that any Souslin set is representable as the result of the Souslin operation on closed sets in X. It remains to use that the Souslin operation preserves the measurability according to Theorem 1.10.5. □

7.4.2. Example. Let X and Y be Souslin spaces and let f be a Borel function on $X\times Y$ that is bounded from below. Set

$$g(x) = \inf_{y\in Y} f(x,y).$$

Then the function g is measurable with respect to every Borel measure on X. If the function f is bounded above, then the function

$$h(x) = \sup_{y\in Y} f(x,y)$$

is measurable with respect to every Borel measure on X.

PROOF. We observe that the set $\{x\colon\ g(x) < c\}$ for any c is the projection on X of the Borel set $\bigl\{(x,y) \in X\times Y\colon\ f(x,y) < c\bigr\}$, i.e., is Souslin. In the case of the function h we consider the set $\{x\colon\ h(x) > c\}$. □

A slightly more general fact is contained in Exercise 6.10.42.

7.4.3. Theorem. *If X is a Souslin space, then every Borel measure μ on X is Radon and is concentrated on a countable union of metrizable compact sets. In addition, for every B in $\mathcal{B}(X)$ and every $\varepsilon > 0$, there exists a metrizable compact set $K_\varepsilon \subset B$ such that $|\mu|(B\backslash K_\varepsilon) < \varepsilon$.*

PROOF. It suffices to show that for every $\varepsilon > 0$, there is a compact set K_ε such that $|\mu|(X\backslash K_\varepsilon) < \varepsilon$. Then it will follow that μ is Radon. Indeed, compact subsets of Souslin spaces are metrizable by Corollary 6.7.8, and on metrizable compact sets all Borel measures are Radon. The tightness can be verified in two ways. The first possibility is to take a continuous mapping f from $\mathbb{N}^\infty$ onto X and apply Theorem 6.9.1. Hence we obtain a mapping $g\colon X \to \mathbb{N}^\infty$ such that $f\bigl(g(x)\bigr) = x$ for all $x \in X$ and, in addition, for every $B \in \mathcal{B}(\mathbb{N}^\infty)$, the set $g^{-1}(B)$ belongs to the σ-algebra generated by all Souslin sets. As shown above, g is measurable with respect to μ. It remains to observe that $\mu = (\mu\circ g^{-1})\circ f^{-1}$ and $\mu\circ g^{-1}$ is a Borel, hence Radon, measure on $\mathbb{N}^\infty$. By the continuity of f the measure μ is Radon as well. The second possibility is to apply Theorem 7.14.34. To this end, one has to verify that the set function $B\mapsto |\mu|^*\bigl(f(B)\bigr)$ is a Choquet capacity. This possibility is left as Exercise 7.14.89. $\square$

7.4.4. Corollary. *Let ν be a Radon measure on a topological space T, let X be a Souslin space, and let $f\colon T\to X$ be measurable with respect to $\bigl(\mathcal{B}(T)_\mu, \mathcal{B}(X)\bigr)$. Then, for every $\varepsilon>0$, there is a compact set $S_\varepsilon\subset T$ such that $|\nu|(T\backslash S_\varepsilon)<\varepsilon$ and $f|_{S_\varepsilon}$ is continuous.*

PROOF. We find a compact set $K\subset X$ with $|\nu|\circ f^{-1}(X\backslash K)<\varepsilon$ and then apply Theorem 7.1.13 to the mapping $f\colon f^{-1}(K)\to K$. $\square$

7.5. Perfect measures

In this section, we discuss an interesting class of measures important for applications: *perfect measures*. For notational simplicity we consider here only finite nonnegative measures.

7.5.1. Definition. *Let $(X,\mathcal{S})$ be a measurable space. A nonnegative measure μ on $\mathcal{S}$ is called perfect if for every $\mathcal{S}$-measurable real function f and every set $E\subset\mathbb{R}$ with $f^{-1}(E)\in\mathcal{S}$, there exists a Borel set B such that $B\subset E$ and $\mu\bigl(f^{-1}(B)\bigr)=\mu\bigl(f^{-1}(E)\bigr)$.*

It terms of $\mu\circ f^{-1}$ perfectness means that the completion of $\mathcal{B}(\mathbb{R}^1)$ with respect to $\mu\circ f^{-1}$ contains all sets E such that $f^{-1}(E)\in\mathcal{S}$. Indeed, for the set $D=\mathbb{R}^1\backslash E$ we also have $f^{-1}(D)=X\backslash f^{-1}(E)\in\mathcal{S}$, hence there is a Borel set $B'\subset D$ with $\mu\bigl(f^{-1}(B')\bigr)=\mu\bigl(f^{-1}(D)\bigr)$. Then for the Borel sets B and $B''=\mathbb{R}^1\backslash B'$ we have $B\subset E\subset B''$ and $\mu\circ f^{-1}(B)=\mu\circ f^{-1}(B'')$, since

$$\begin{aligned}\mu\circ f^{-1}(B'') &= \mu\bigl(f^{-1}(\mathbb{R}^1\backslash B')\bigr)=\mu(X)-\mu\bigl(f^{-1}(B')\bigr)\\ &= \mu(X)-\mu\bigl(f^{-1}(D)\bigr)=\mu(X)-\mu\bigl(X\backslash f^{-1}(E)\bigr)=\mu\bigl(f^{-1}(E)\bigr).\end{aligned}$$

In particular, we have $f(X)\in\mathcal{B}(\mathbb{R}^1)_{\mu\circ f^{-1}}$. However, the set $f(A)$ may fail to be $\mu\circ f^{-1}$-measurable for a set $A\in\mathcal{S}$, although the set $f(A)$ is always $\mu_A\circ f^{1}$-measurable. For example, the identity mapping on the interval $[0,1]$ with Lebesgue measure (which is perfect, as we shall see) can be redefined on a measure zero set Z in such a way that the image of Z will be nonmeasurable with respect to Lebesgue measure (note that Lebesgue measure is transformed into itself).

It is clear from the definition that a perfect measure μ is perfect on every σ-algebra $\mathcal{S}_1\subset\mathcal{S}$.

7.5.2. Proposition. *A measure μ on $(X,\mathcal{S})$ is perfect if and only if for every $\mathcal{S}$-measurable real function f, there exists a Borel set $B\subset\mathbb{R}$ such that*

$$B\subset f(X)\quad\textit{and}\quad \mu\bigl(f^{-1}(B)\bigr)=\mu(X).$$

Proof. The above condition is obviously fulfilled for any perfect measure. Suppose now that it is fulfilled for some measure μ on $\mathcal{S}$. Let f be an $\mathcal{S}$-measurable function, $E \subset \mathbb{R}^1$ and $f^{-1}(E) \in \mathcal{S}$. Let us take an arbitrary point $c \in E$ and consider the following function: $f_0(x) = f(x)$ if $x \in f^{-1}(E)$, $f_0(x) = c$ if $x \notin f^{-1}(E)$. It is clear that f_0 is an $\mathcal{S}$-measurable function and $f_0(X) = E$. Hence there is a Borel set $B \subset E$ with $\mu\bigl(f_0^{-1}(B)\bigr) = \mu(X)$. If $c \notin B$, then $f_0^{-1}(B) = f^{-1}(B)$, whence $\mu\bigl(f^{-1}(B)\bigr) = \mu(X)$. Therefore, $\mu\bigl(f^{-1}(E)\bigr) = \mu(X)$. If $c \in B$, then $f_0^{-1}(B) = f^{-1}(B) \cup \bigl(X\backslash f^{-1}(E)\bigr)$, whence one has

$$\mu\bigl(f^{-1}(B)\bigr) + \mu\bigl(X\backslash f^{-1}(E)\bigr) = \mu(X).$$

Therefore, $\mu\bigl(f^{-1}(B)\bigr) - \mu\bigl(f^{-1}(E)\bigr) = 0$. □

7.5.3. Example. Let $X \subset [0,1]$, $\lambda^*(X) = 1$, $\lambda_*(X) = 0$, where λ is Lebesgue measure, and let μ be the restriction of λ to $\mathcal{B}(X)$, i.e., one has $\mu(B \cap X) = \lambda(B)$, $B \in \mathcal{B}([0,1])$. Then μ is not perfect (it suffices to take the function $f\colon X \to [0,1]$, $f(x) = x$).

Let us mention some elementary properties of perfect measures. These almost immediate properties are often useful in applications.

7.5.4. Proposition. (i) *A measure μ on a σ-algebra $\mathcal{S}$ is perfect precisely when its completion is perfect on $\mathcal{S}_\mu$.*

(ii) *If a measure μ on a σ-algebra $\mathcal{S}$ is perfect, then its restriction to any set $E \in \mathcal{S}_\mu$ equipped with the trace of an arbitrary sub-σ-algebra in $\mathcal{S}_\mu$ is a perfect measure.*

(iii) *Let a measure μ on $(X, \mathcal{S})$ be perfect, let $(Y, \mathcal{A})$ be a measurable space, and let $F\colon X \to Y$ be an $(\mathcal{S}, \mathcal{A})$-measurable mapping. Then, the induced measure $\mu \circ F^{-1}$ on $\mathcal{A}$ is perfect.*

Proof. (i) Let a measure μ on $\mathcal{S}$ be perfect and let f be an $\mathcal{S}_\mu$-measurable function. We shall assume that the set $f(X)$ is uncountable, since otherwise it can be taken as a required Borel set. We pick a point $c \in f(X)$ with $\mu\bigl(f^{-1}(c)\bigr) = 0$. There exist an $\mathcal{S}$-measurable function f_0 and a set $X_0 \in \mathcal{S}$ such that $\mu(X_0) = \mu(X)$ and $f_0 = f$ on X_0. The function f_0 can be redefined in such a way that $f_0(x) = c$ if $x \in X\backslash X_0$. We take a Borel set $B \subset f_0(X)$ with $\mu\bigl(f_0^{-1}(B)\bigr) = \mu(X)$. It is clear that $B \subset f(X)$ and $\mu\bigl(f^{-1}(B)\bigr) = \mu(X)$. Claim (ii) follows by (i).

(iii) If a function f is measurable with respect to $\mathcal{A}$, then the function $f \circ F$ is measurable with respect to $\mathcal{S}$. Hence there exists a Borel set $B \subset f\bigl(F(X)\bigr)$ with $\mu\bigl(F^{-1}\bigl(f^{-1}(B)\bigr)\bigr) = \mu(X) = \mu \circ F^{-1}(Y)$. By Proposition 7.5.2, the measure $\mu \circ F^{-1}$ is perfect. □

As explained above, the image of a space X with a perfect measure μ under a μ-measurable real function f is measurable with respect to the image measure $\mu \circ f^{-1}$ (but it may not be measurable with respect to other Borel measures, for example, with respect to Lebesgue measure). The same is true for mappings f with values in a measurable space $(E, \mathcal{E})$ if $\mathcal{E}$ is countably generated and countably separated because $(E, \mathcal{E})$ is isomorphic to a subset of $\mathbb{R}^1$ with the Borel σ-algebra. However, in general, the image of a space with a complete perfect measure under a measurable mapping to a space with a complete perfect measure may not be measurable.

7.5.5. Example. Let $X = \{0\}$ be equipped with Dirac's measure δ and let Y be the product of the continuum of intervals. We equip Y with Dirac's measure δ at zero considered on the δ-completion of $\mathcal{B}a(Y)$. Then in both cases the measure δ is perfect, the natural embedding $X \to Y$ is measurable, but the point zero is not in $\mathcal{B}a(Y)_\delta$.

The previous example also shows that the restriction of a perfect measure to a nonmeasurable set of full outer measure may be a perfect measure.

The next result shows that the class of perfect measures is very large. Most measures actually encountered are perfect. The same result describes close connections between perfect and compact measures.

7.5.6. Theorem. (i) *Every measure possessing an approximating compact class is perfect.*

(ii) *A measure μ on a σ-algebra $\mathcal{S}$ is perfect if and only if it possesses an approximating compact class on every countably generated sub-σ-algebra $\mathcal{S}_1 \subset \mathcal{S}$.*

(iii) *A measure μ on $(X, \mathcal{S})$ is perfect if and only if it is quasi-compact in the following sense: for every sequence $\{A_i\} \subset \mathcal{S}$ and every $\varepsilon > 0$, there exists a set $A \in \mathcal{S}$ such that $\mu(A) > \mu(X) - \varepsilon$ and the sequence $\{A \cap A_i\}$ is a compact class.*

(iv) *A measure on a countably separated σ-algebra is perfect if and only if it has a compact approximating class.*

PROOF. (i) We show that any measure μ with an approximating compact class $\mathcal{K}$ is quasi-compact. As explained in §1.12(ii), we may assume that the class $\mathcal{K}$ belongs to $\mathcal{S}$ and admits finite unions and countable intersections. Given $\varepsilon > 0$ and sets $A_n \in \mathcal{S}$, we find $C_n \subset A_n$ and $B_n \subset X \backslash A_n$ such that

$$C_n, B_n \in \mathcal{K},\ \mu(A_n \backslash C_n) < \varepsilon 2^{-n-1},\ \mu\big((X \backslash A_n) \backslash B_n\big) < \varepsilon 2^{-n-1}.$$

Let $A = \bigcap_{n=1}^\infty (C_n \cup B_n)$. Then $C_n \cap A \in \mathcal{K}$. It is easy to see that we have $A_n \cap A = C_n \cap A$, which proves the compactness of the class $\{A_n \cap A\}$. In addition, $\mu(A) > \mu(X) - \varepsilon$.

We now prove that any quasi-compact measure μ is perfect. Let f be an $\mathcal{S}$-measurable function. Let $\{I_n\}$ be the countable set of all open intervals with rational endpoints. Let $A_n = f^{-1}(I_n)$. For every $\varepsilon_k = 2^{-k}$, we take a set E_k with $\mu(E_k) > \mu(X) - 2^{-k}$ such that the class $\{E_k \cap A_n\}$ is compact. Set $E = \bigcup_{k=1}^\infty E_k$. It is clear that $\mu(E) = \mu(X)$. It remains to observe that the sets $f(E_k)$ are closed. Indeed, let k be fixed and let t be a limit point of the set $f(E_k)$. Then there exist numbers n_j such that the intervals I_{n_j} are decreasing and $t = \bigcap_{j=1}^\infty I_{n_j}$. It is clear that they all meet $f(E_k)$ because $t \in \overline{f(E_k)}$. Hence the sets $E_k \cap A_{n_j}$ are nonempty. By the definition of a compact class, there exists a point x in their intersection. Then $f(x) = t$ since $f(x) \in f(E_k \cap A_{n_j}) \subset I_{n_j}$. Hence $f(E_k)$ is closed. By Proposition 7.5.2 the measure μ is perfect.

(ii) Let the measure μ be perfect and let a σ-algebra $\mathcal{S}_1$ be generated by a countable family of sets $A_i \in \mathcal{S}$. As shown in Theorem 6.5.5, one has $\mathcal{S}_1 = f^{-1}\bigl(\mathcal{B}(\mathbb{R}^1)\bigr)$, where $f = \sum_{n=1}^{\infty} 3^{-n} I_{A_n}$. Due to our assumption the set $f(X)$ is $\mu \circ f^{-1}$-measurable. Hence the class $\mathcal{E}$ of its compact subsets is approximating for the measure $\mu \circ f^{-1}$. Then the class of sets $f^{-1}(E)$, where $E \in \mathcal{E}$, is compact and approximating for μ on $\mathcal{S}_1$.

If μ has an approximating compact class on every countably generated σ-algebra in $\mathcal{S}$, then the reasoning in (i) yields that μ is quasi-compact on $\mathcal{S}$, hence is perfect as shown above. Claim (iii) follows by the already proven assertions.

(iv) If a measure μ on a countably separated σ-algebra $\mathcal{S}$ in X is perfect, then we take an injective $\mathcal{S}$-measurable real function f on X and denote by $\mathcal{K}$ the class of all sets of the form $f^{-1}(E)$, where E is a compact subset in $f(X)$. If we are given a family of sets $K_\alpha = f^{-1}(E_\alpha) \in \mathcal{K}$ such that every finite subfamily has a nonempty intersection, then all finite families of compact sets E_α have nonempty intersections. Hence $\bigcap_\alpha E_\alpha \neq \varnothing$. By the injectivity of f we obtain $\bigcap_\alpha K_\alpha \neq \varnothing$. Thus, the class $\mathcal{K}$ is compact (even $\aleph$-compact, see below). Furthermore, $\mathcal{K}$ approximates μ, as for every $A \in \mathcal{S}$ the measure $\mu|_A$ is perfect, which gives compact sets $E_n \subset f(A)$ with $\mu\bigl(f^{-1}(E_n)\bigr) \to \mu(A)$. $\square$

Vinokurov, Mahkamov [**1930**] and Musiał [**1346**] give examples of spaces with perfect, but not compact measures. Since their constructions are rather involved, we do not reproduce them here.

Certainly, it can happen that on a given σ-algebra there are perfect and non-perfect measures. The following result deals with the situation where all measures on a given σ-algebra are perfect.

7.5.7. Theorem. (i) *Let $X \subset \mathbb{R}$. Every Borel measure on $\mathcal{B}(X)$ is perfect if and only if X is universally measurable, i.e., is measurable with respect to the completion of every Borel measure on* $\mathbb{R}$.

(ii) *Let $(X, \mathcal{S})$ be a measurable space. If for every $\mathcal{S}$-measurable function f, the set $f(X) \subset \mathbb{R}$ is universally measurable, then every measure on every sub-σ-algebra $\mathcal{S}_1 \subset \mathcal{S}$ is perfect. Conversely, if every measure on every countably generated sub-σ-algebra $\mathcal{S}_1 \subset \mathcal{S}$ is perfect, then for every $\mathcal{S}$-measurable function f, the set $f(X) \subset \mathbb{R}$ is universally measurable.*

(iii) *Let $\mathcal{S}$ be a countably generated σ-algebra in a space X. Every probability measure on $\mathcal{S}$ is perfect if and only if for some* (*and then for every*) *sequence of sets A_n generating $\mathcal{S}$, the set of values of the function $h := \sum_{n=1}^{\infty} 3^{-n} I_{A_n}$ is universally measurable on the real line.*

Proof. (i) If X is measurable with respect to a Borel measure μ on the real line, then μ is Radon on X, hence perfect. The converse follows by Proposition 7.5.2.

(ii) If we are given a measure μ on $\mathcal{S}$ and an $\mathcal{S}$-measurable function f, then the measurability of $f(X)$ with respect to $\mu\circ f^{-1}$ gives a Borel set $B\subset f(X)$ of full measure with respect to $\mu\circ f^{-1}$. By Proposition 7.5.2 the measure μ is perfect on $\mathcal{S}$. The same is true for any sub-σ-algebra in $\mathcal{S}$.

Suppose that every measure μ on every countably generated sub-σ-algebra in $\mathcal{S}$ is perfect. If we are given an $\mathcal{S}$-measurable function f and a measure ν on $\mathcal{B}\bigl(f(X)\bigr)$, then we can consider the measure $\mu\colon f^{-1}(E)\mapsto \nu(E)$ on the countably generated σ-algebra of sets $f^{-1}(E)$, $E\in\mathcal{B}\bigl(f(X)\bigr)$. Since by hypothesis the measure μ is perfect, its image ν is perfect as well. By (i) the set $f(X)$ is universally measurable.

(iii) If every measure on $\mathcal{S}$ is perfect, then $h(X)$ is universally measurable (for any sequence $\{A_n\}\subset\mathcal{S}$) according to (ii). Conversely, suppose that for some sequence of sets A_n generating $\mathcal{S}$ the set $h(X)$ is universally measurable on the real line. Every $\mathcal{S}$-measurable function f has the form $g\circ h$, where g is a Borel function on the real line. According to (i) every Borel measure on $h(X)$ is perfect. By (ii) the set $g\bigl(h(X)\bigr)$ is universally measurable. □

7.5.8. Example. If X is a Souslin space (for example, a Borel set in a Polish space), then every measure μ on an arbitrary sub-σ-algebra $\mathcal{S}_1$ in $\mathcal{S}:=\mathcal{B}(X)$ is perfect. This is clear from assertion (ii) in the above theorem and the fact that the image of a Souslin space under a Borel function is universally measurable. However, μ may not be extendible to all of $\mathcal{B}(X)$ and not approximated from within by compact sets (see Example 9.8.1).

7.5.9. Example. (Sazonov [**1656**]) Under the continuum hypothesis, there exists a measurable space $(X,\mathcal{S})$ such that every measure on $\mathcal{S}$ is perfect, but there is a sub-σ-algebra $\mathcal{S}_1\subset\mathcal{S}$ on which there are non-perfect measures. Indeed, we take for X the interval $[0,1]$ with the σ-algebra $\mathcal{S}$ of all subsets. We know (see §1.12(x)) that under the continuum hypothesis, every measure on $\mathcal{S}$ is concentrated on a countable set, hence is perfect. On the other hand, there are non-perfect measures on $[0,1]$, as we have seen in Example 7.5.3.

So far in our discussion of perfect measures no topological concepts have been involved. It is time to do this.

7.5.10. Theorem. (i) *Every Radon measure on a topological space is perfect. Hence every tight Baire measure is perfect.*

(ii) *A Borel measure on a separable metric space is perfect if and only if it is Radon.*

(iii) *A Borel measure on a metric space is Radon if and only if it is perfect and τ-additive.*

Proof. The first claim in (i) follows from Theorem 7.5.6 and Theorem 7.3.10. The second claim follows from the first one and Proposition 7.5.4. For the proof of assertion (ii) we suppose that a measure μ on a separable metric space X is perfect and take a countable family of open balls U_n with all possible rational radii and centers at the points of a countable everywhere dense set. The function

$$\xi = \sum_{n=1}^{\infty} 3^{-n} I_{U_n}$$

maps X one-to-one onto the set $\xi(X) \subset \mathbb{R}^1$. If we equip this set with the usual topology, the mapping $\xi^{-1}\colon \xi(X) \to X$ becomes continuous. Indeed, let $t \in \xi(X)$ and $\varepsilon > 0$. In the ε-neighborhood of the point $x = \xi^{-1}(t)$ we pick a ball U_{n_0} containing this point. Let $s \in \xi(X)$ and $|t - s| < 3^{-n_0-1}$. Then the point $y = \xi^{-1}(s)$ belongs to U_{n_0}, since otherwise $I_{U_{n_0}}(y) = 0$ and $|t - s| = |\xi(x) - \xi(y)| \geq 3^{-n_0}/2$. Hence ξ^{-1} is continuous. By hypothesis, there exists a Borel set $B \subset \xi(X)$ such that $\mu(X) = \mu\bigl(\xi^{-1}(B)\bigr)$. Since the measure $\mu \circ \xi^{-1}$ on the real line is Radon, for every $\varepsilon > 0$, one can find a compact set $C_\varepsilon \subset B$ with $\mu\bigl(\xi^{-1}(C_\varepsilon)\bigr) > \mu(X) - \varepsilon$. It remains to observe that $K_\varepsilon = \xi^{-1}(C_\varepsilon)$ is compact by the continuity of ξ^{-1}. Claim (iii) follows from (ii), since any τ-additive measure on a metric space has separable support because any nonseparable metric space contains an uncountable collection of disjoint balls. □

7.5.11. Example. (i) There exists a τ-additive Borel measure on a separable metric space that is not perfect.

(ii) There exists a perfect measure on a locally compact space possessing an approximating compact class, but not τ-additive.

(iii) There exists a perfect τ-additive Borel measure (which even has an approximating compact class) that is not tight.

PROOF. For the proof of (i) we take the measure from Example 7.5.3. In order to construct an example in (ii), we take for X the space X_0 from Example 7.1.3 (the space of countable ordinals), and consider the measure μ that equals 0 on all countable sets and 1 on their complements (such sets exhaust all Borel sets in X_0). One can verify that μ is not τ-additive, but possesses an approximating compact class (namely, consisting of the empty set and all sets of measure 1). Finally, Lebesgue measure on the Sorgenfrey interval from Example 7.2.4 can be taken in (iii). This measure is perfect, since the Borel σ-algebra corresponding to the Sorgenfrey topology is the usual Borel σ-algebra of the interval. By Theorem 7.5.6(ii) it has an approximating compact class. However, this measure vanishes on all compact sets in the Sorgenfrey interval, since they are finite. □

Some authors call a measure μ on a σ-algebra $\mathcal{A}$ in a space X compact if it has an approximating class $\mathcal{K} \subset \mathcal{A}$ that is compact in the following stronger sense: every collection of sets in $\mathcal{K}$ that has an empty intersection possesses a finite subcollection whose intersection is empty. In this terminology, measures (and classes) compact in our sense are called countably compact, semicompact or $\aleph_0$-compact. For the above-mentioned stronger property we shall use the term $\aleph$-*compactness*. It is clear that any Radon measure possesses this stronger property. However, not every compact (in our sense) measure is $\aleph$-compact (see Exercise 1.12.105).

7.6. Products of measures

In this section, we discuss regularity properties of product measures on topological spaces. First of all, the kind of problems we have as compared to the already discussed product measures must be explained. The point is that the product of Baire or Borel σ-algebras may be strictly smaller than the Baire and Borel σ-algebra of the product space. There are no problems if we deal with countable products of Borel probability measures on separable metric spaces (or on Souslin spaces).

7.6.1. Example. Let μ_n be Borel probability measures on separable metric spaces X_n. Then $\mu = \bigotimes_{n=1}^{\infty} \mu_n$ is a Borel probability measure on the separable metric space $X = \prod_{n=1}^{\infty} X_n$.

Proof. The measure μ is defined on the σ-algebra $\mathcal{E}$ generated by finite products of Borel sets in X_n. But $\mathcal{E} = \mathcal{B}(X)$ due to the fact that every open set in X belongs to $\mathcal{E}$, since it can be represented as a countable union of finite products of open sets in X_n. □

We shall see below that the situation is not that simple for uncountable products and for countable products of more complicated spaces. Another simple, but important result concerns countable products of Radon measures.

7.6.2. Theorem. (i) *Let μ_n be a sequence of Radon probability measures on Hausdorff spaces X_n. Then their product $\mu = \bigotimes_{n=1}^{\infty} \mu_n$ uniquely extends to a Radon measure on $X = \prod_{n=1}^{\infty} X_n$.*

(ii) *Let μ_n be a sequence of tight Baire probability measures on completely regular spaces X_n. Then their product μ is a tight measure on the space $\bigotimes_{n=1}^{\infty} \mathcal{B}a(X_n)$ and uniquely extends to a Radon measure on $X = \prod_{n=1}^{\infty} X_n$.*

Proof. (i) Let $\varepsilon > 0$. The measure μ is defined on the σ-algebra $\mathcal{E} = \bigotimes_{n=1}^{\infty} \mathcal{B}(X_n)$, which contains finite products of open sets, hence contains a base of the topology in X. Every X_n contains a compact set K_n with $\mu_n(K_n) > 1 - \varepsilon 2^{-n}$. It remains to observe that $K = \prod_{n=1}^{\infty} K_n$ is compact and $\mu(K) > 1 - \varepsilon$. Thus, the measure μ is tight. In order to apply Theorem 7.3.2, we have to verify the regularity of μ on $\mathcal{E}$. According to the cited theorem, it suffices to verify the regularity of μ on the algebra $\mathcal{R}$ generated by finite products of Borel sets in the spaces X_n (we observe that $\mathcal{E}$ is the σ-algebra generated by $\mathcal{R}$). The algebra $\mathcal{R}$ consists of finite unions of finite products of sets in $\mathcal{B}(X_n)$, hence the required regularity follows by the regularity of each measure μ_n.

In case (ii) the reasoning is analogous: we take compact sets K_n such that $\mu_n(A) < \varepsilon 2^{-n}$ for every Baire set A disjoint with K_n. The set $K = \prod_{n=1}^{\infty} K_n$ is compact. If a set $A \in \bigotimes_{n=1}^{\infty} \mathcal{B}a(X_n)$ does not meet K, then $\mu(A) \leq \varepsilon$. Indeed, let ν_n be the Radon extension of μ_n to $\mathcal{B}(X_n)$. Then $\nu_n(K_n) \geq 1 - \varepsilon 2^{-n}$ since otherwise we could take a compact set $C_n \subset X_n \backslash K_n$ with $\nu_n(C_n) > \varepsilon 2^{-n}$, next find a functionally open set U_n with $C_n \subset U_n$

and $U_n \cap K_n = \varnothing$, which would give $\mu_n(U_n) = \nu_n(U_n) > \varepsilon 2^{-n}$. Hence $\mu(A) = \bigl(\bigotimes_{n=1}^{\infty} \nu_n\bigr)(A) \le 1 - \bigl(\bigotimes_{n=1}^{\infty} \nu_n\bigr)(K) \le \varepsilon$. □

For uncountable products this theorem may fail.

7.6.3. Example. Let μ_α, $\alpha \in A$, be an uncountable family of Baire probability measures on spaces X_α without compact subsets of outer measure 1. Then $\otimes_\alpha \mu_\alpha(K) = 0$ for every compact set $K \subset \prod_\alpha X_\alpha$. In particular, the measure $\otimes_\alpha \mu_\alpha$ is not tight.

Proof. By the compactness of K, there exist compact sets $K_\alpha \subset X_\alpha$ such that $K \subset \prod_\alpha K_\alpha$. Since A is uncountable, our hypothesis yields that for some $q < 1$, there is an infinite family of indices β with $\mu_\beta^*(K_\beta) \le q$. We take in this family any countable subfamily $B = \{\beta_n\}$ and obtain the set $C = \prod_{n=1}^{\infty} K_{\beta_n} \times \prod_{\alpha \notin B} K_\alpha$ of measure zero containing K. □

Obviously, it follows by the above theorem that finite products of Radon measures have Radon extensions. But when dealing with products it is often desirable to have not only the existence of a product measure, but also to be able to apply Fubini's theorem. Certainly, Fubini's theorem is applicable to all sets in the σ-algebra generated by rectangles (this has no topological specifics). However, as we have already noted, in the case of general topological spaces, there are Borel sets in the product not belonging to this σ-algebra. We shall now see that Fubini's theorem can be applied to such sets as well.

Let X_1 and X_2 be two spaces. For every set $A \subset X_1 \times X_2$, let

$$A_{x_1} = \bigl\{x_2 \in X_2\colon\ (x_1, x_2) \in A\bigr\}, \quad A_{x_2} = \bigl\{x_1 \in X_1\colon\ (x_1, x_2) \in A\bigr\}.$$

7.6.4. Lemma. *Let X_1 and X_2 be topological spaces and let ν be a τ-additive measure on X_1. Then:*

(i) *for every $B \in \mathcal{B}(X_1 \times X_2)$, the function $x_2 \mapsto \nu(B_{x_2})$ is Borel on X_2; hence for every bounded Borel function f on $X \times Y$ the function*

$$x_2 \mapsto \int_{X_1} f(x_1, x_2)\, \nu(dx_1)$$

is Borel on X_2;

(ii) *if ν is nonnegative and the set $U \subset X_1 \times X_2$ is open, then the function $x_2 \mapsto \nu(U_{x_2})$ is lower semicontinuous on X_2.*

Proof. First we verify assertion (ii). If $U = U_1 \times U_2$, then we have $\nu(U_{x_2}) = \nu(U_1) I_{U_2}$ and it remains to observe that the indicator of an open set is lower semicontinuous. Our assertion remains true for any set U that is a finite union of such products. Finally, an arbitrary open set $U \subset X_1 \times X_2$ can be represented as $U = \bigcup_\alpha U_\alpha$, where $\{U_\alpha\}$ is a net of increasing open sets that are finite unions of open rectangles. By the τ-additivity we obtain $\nu(U_{x_2}) = \sup_\alpha \nu\bigl((U_\alpha)_{x_2}\bigr)$, whence the claim follows.

It suffices to prove (i) for nonnegative measures. Denote by $\mathcal{B}'$ the class of all sets $B \in \mathcal{B}(X_1 \times X_2)$ such that the function $x_2 \mapsto \nu(B_{x_2})$ is Borel. By the above, $\mathcal{B}'$ contains the class $\mathcal{E}$ of all open sets. It is clear that any countable union of pairwise disjoint sets in $\mathcal{B}'$ belongs to $\mathcal{B}'$ as well (since the sum of the series of Borel functions is a Borel function). In addition, $B_1 \backslash B_2 \in \mathcal{B}'$ for all $B_1, B_2 \in \mathcal{B}'$ such that $B_2 \subset B_1$. According to Theorem 1.9.3, we obtain that the σ-algebra generated by $\mathcal{E}$ is contained in $\mathcal{B}'$. Therefore, the class $\mathcal{B}'$ coincides with $\mathcal{B}(X_1 \times X_2)$. $\square$

7.6.5. Theorem. *Suppose that μ_1 and μ_2 are τ-additive measures. Then the measure $\mu = \mu_1 \otimes \mu_2$ has a unique extension to a τ-additive measure μ on $\mathcal{B}(X_1 \times X_2)$ and for every $B \in \mathcal{B}(X_1 \times X_2)$ one has*

$$\mu(B) = \int_{X_2} \mu_1(B_{x_2})\, \mu_2(dx_2) = \int_{X_1} \mu_2(B_{x_1})\, \mu_1(dx_1), \tag{7.6.1}$$

where the functions $x_2 \mapsto \mu_1(B_{x_2})$ and $x_1 \mapsto \mu_2(B_{x_1})$ are Borel. If both measures μ_1 and μ_2 are Radon, then the extension by formula (7.6.1) *is Radon as well and coincides with the extension from Theorem* 7.6.2.

PROOF. According to the above lemma the integrands in (7.6.1) are Borel. Hence both integrals are well-defined and produce Borel measures on $X_1 \times X_2$. In the justification of equality (7.6.1) it is sufficient to consider nonnegative measures. Denote by $\mathcal{E}$ the class of all sets $B \in \mathcal{B}(X_1 \times X_2)$ on which these measures are equal. As in the proof of the above lemma, the class $\mathcal{E}$ is σ-additive. Hence for the proof of the equality $\mathcal{E} = \mathcal{B}(X_1 \times X_2)$ it suffices to show that every open set U belongs to $\mathcal{E}$. We represent U in the form $U = \bigcup_\alpha U_\alpha$, where $\{U_\alpha\}$ is a net of increasing open sets that are finite unions of open rectangles. Clearly, $U_\alpha \in \mathcal{E}$. The τ-additivity of μ_1, the lower semicontinuity of the functions $x_2 \mapsto \mu_1\bigl((U_\alpha)_{x_2}\bigr)$, and Lemma 7.2.6 yield

$$\begin{aligned}\int_{X_2} \mu_1(U_{x_2})\, \mu_2(dx_2) &= \int_{X_2} \lim_\alpha \mu_1\bigl((U_\alpha)_{x_2}\bigr)\, \mu_2(dx_2)\\ &= \lim_\alpha \int_{X_2} \mu_1\bigl((U_\alpha)_{x_2}\bigr)\, \mu_2(dx_2) = \lim_\alpha \mu(U_\alpha).\end{aligned}$$

The same reasoning applies to the second integral, whence we obtain $U \in \mathcal{E}$. The proof of the τ-additivity of the obtained measure μ is analogous. The uniqueness of a τ-additive extension follows from the fact that if a τ-additive measure vanishes on all open rectangles, then it vanishes on all open sets, hence on all Borel sets (this follows by Lemma 1.9.4). Finally, if the measures μ_1 and μ_2 are Radon, then so is the constructed measure μ, since it is τ-additive and obviously tight. $\square$

7.6.6. Lemma. *Let X and Y be topological spaces and let μ be a probability measure on $\mathcal{B}(X) \otimes \mathcal{B}(Y)$. Suppose that the projections of μ on X and Y are tight. Then μ is tight as well. If both projections are concentrated on countable unions of metrizable compact sets, then μ has this property as well.*

PROOF. Given $\varepsilon > 0$, we find compact sets $K \subset X$ and $S \subset Y$ such that $\mu(K \times Y) > 1 - \varepsilon/2$ and $\mu(X \times S) > 1 - \varepsilon/2$. Then $K \times S$ is compact in $X \times Y$ and $\mu(K \times S) > 1 - \varepsilon$. The last claim is obvious from the proof. $\square$

Additional information about finite and infinite products of measures is given in §7.14(i) and Exercises 7.14.100, 7.14.116, 7.14.157, and 7.14.158.

7.7. The Kolmogorov theorem

In many problems of measure theory and probability theory and their applications, one has to construct measures on products of measurable spaces that are more complicated than product measures. In this section, we prove the principal result in this direction: the Kolmogorov theorem on consistent probability distributions. The classical Kolmogorov result was concerned with measures on products of real lines, and the abstract formulation given below goes back to E. Marczewski.

Let T be a nonempty set. Suppose that we are given nonempty measurable spaces $(\Omega_t, \mathcal{B}_t)$, $t \in T$. For every nonempty set $\Lambda \subset T$, we denote by Ω_Λ the product of the spaces Ω_t, $t \in \Lambda$. The space Ω_Λ is equipped with the σ-algebra $\mathcal{B}_\Lambda$ that is the product of the σ-algebras $\mathcal{B}_t$, $t \in \Lambda$ (see §3.5 in Chapter 3).

7.7.1. Theorem. *Suppose that for every finite set $\Lambda \subset T$, we are given a probability measure μ_Λ on $(\Omega_\Lambda, \mathcal{B}_\Lambda)$ such that the following consistency condition is fulfilled: if $\Lambda_1 \subset \Lambda_2$, then the image of the measure μ_{Λ_2} under the natural projection from Ω_{Λ_2} to Ω_{Λ_1} coincides with μ_{Λ_1}. Suppose that for every $t \in T$, the measure μ_t on $\mathcal{B}_t$ possesses an approximating compact class $\mathcal{K}_t \subset \mathcal{B}_t$. Then, there exists a probability measure μ on the measurable space $\bigl(\Omega := \prod_{t\in T}\Omega_t, \mathcal{B} := \bigotimes_{t\in T}\mathcal{B}_t\bigr)$ such that the image of μ under the natural projection from Ω to Ω_Λ is μ_Λ for each finite set $\Lambda \subset T$.*

Proof. Every set $B \in \mathcal{B}_\Lambda$ can be identified with the cylindrical set $C_\Lambda = B \times \prod_{t\in T\backslash\Lambda}\Omega_t$. It is clear that the family of such sets forms an algebra $\mathcal{R}$. This algebra is generated by the semialgebra of finite products

$$\prod_{i=1}^n B_{t_i} \times \prod_{t\notin\{t_1,\ldots,t_n\}} \Omega_t.$$

On the algebra $\mathcal{R}$, we have the set function $\mu(C_\Lambda) = \mu_\Lambda(B)$. The consistency condition yields that this function is well-defined, i.e., $\mu(C_\Lambda)$ is independent of the representation of C_Λ in the above form. Indeed, if we replace B with some other set $B' \in \mathcal{B}_{\Lambda'}$, where $\Lambda \subset \Lambda'$, then B is the image of B' under projecting $\Omega_{\Lambda'}$ to Ω_Λ, hence $\mu_\Lambda(B) = \mu_{\Lambda'}(B')$.

We verify the countable additivity of the set function μ on the algebra $\mathcal{R}$. Let us recall that the class $\mathcal{K}$ of all finite unions of products of the form $\prod_{i=1}^n K_{t_i} \times \Omega_{T\backslash\{t_1,\ldots,t_n\}}$, where $K_{t_i} \in \mathcal{K}_{t_i}$, is compact (see Lemma 3.5.3). We prove that this class approximates μ. It suffices to show that for every product $B = \prod_{i=1}^n B_{t_i} \times \prod_{t\notin\{t_1,\ldots,t_n\}}\Omega_t$ and every $\varepsilon > 0$, there exists a set $K_{t_i} \in \mathcal{K}_{t_i}$ such that the set $K = \prod_{i=1}^n K_{t_i} \times \prod_{t\notin\{t_1,\ldots,t_n\}}\Omega_t$ approximates B with respect to μ up to ε. We take $K_{t_i} \in \mathcal{K}_{t_i}$ such that $\mu_{t_i}(B_{t_i} \backslash K_{t_i}) < \varepsilon n^{-1}$ and observe that one has the easily verified inclusion

$$B\backslash K \subset \bigcup_{i=1}^{n} \Big((B_{t_i}\backslash K_{t_i}) \times \prod_{t\neq t_i} \Omega_t \Big),$$

whence we obtain

$$\mu(B\backslash K) \le \sum_{i=1}^{n} \mu_{t_i}(B_{t_i}\backslash K_{t_i}) = \sum_{i=1}^{n} \mu\Big((B_{t_i}\backslash K_{t_i}) \times \prod_{t\neq t_i} \Omega_t \Big) < \varepsilon,$$

which completes the proof. □

The measure μ is called the projective limit of the measures μ_Λ.

It is clear that the Kolmogorov theorem is applicable if μ_Λ are consistent Radon measures.

7.7.2. Corollary. *Let X_t, $t \in T$, be Souslin spaces and let $\mathcal{B}_t = \mathcal{B}(X_t)$. Suppose that for every finite set $\Lambda \subset T$, we are given a probability measure μ_Λ on $(\Omega_\Lambda, \mathcal{B}_\Lambda)$ such that the consistency condition from Theorem* 7.7.1 *is fulfilled. Then, there exists a probability measure μ on the measurable space $\big(\Omega = \prod_{t\in T} \Omega_t, \mathcal{B} = \bigotimes_{t\in T} \mathcal{B}_t\big)$ such that the image of μ under the natural projection from Ω to Ω_Λ is μ_Λ for all finite sets $\Lambda \subset T$.*

Proof. It suffices to use the fact that all Borel measures on Souslin spaces are Radon. □

Certainly, the same result is true for measurable spaces that are isomorphic to Souslin spaces with the Borel σ-algebras.

We remark that a particular case of the above theorem is the existence of the product of the measures μ_t. Indeed, in this case one takes for μ_Λ, where Λ is a finite set, the finite product $\bigotimes_{t\in\Lambda} \mu_t$ on $\bigotimes_{t\in\Lambda} \mathcal{B}_t$. However, in this particular case, as we know, no approximating compact class is needed (see §3.5). Let us show that in Theorem 7.7.1 one cannot omit this condition.

7.7.3. Example. Let us take sets $X_n \subset [0,1]$ such that all X_n have outer Lebesgue measure 1, $X_{n+1} \subset X_n$ and $\bigcap_{n=1}^\infty X_n = \varnothing$ (see Exercise 1.12.58). Let $\mathcal{B}_n$ be the Borel σ-algebra of X_n and let μ_n be the trace of Lebesgue measure on $\mathcal{B}_n$ (see Chapter 1, Definition 1.12.11). For every n, let

$$\pi_n\colon\ X_n \to \prod_{i=1}^{n} X_i, \quad \pi_n(x) = (x, \ldots, x).$$

On $\bigotimes_{i=1}^n \mathcal{B}_i$ we obtain the measure $\mu_{(1,\ldots,n)} = \mu_n \circ \pi_n^{-1}$. Then the family of probability measures $\{\mu_{(1,\ldots,n)}, n \ge 1\}$ is consistent, but there is no measure on the product $(\prod_{i=1}^\infty X_i, \bigotimes_{i=1}^\infty \mathcal{B}_i)$ whose images under the projections to $\prod_{i=1}^n X_i$ coincide with the measures $\mu_{(1,\ldots,n)}$ for all n.

Proof. Since X_n are separable metric spaces, one has

$$\bigotimes_{i=1}^{n} \mathcal{B}_i = \mathcal{B}\Big(\prod_{i=1}^{n} X_i\Big),$$

in particular, the diagonal $\Delta_n := \{x = (x_1, \ldots, x_n)\colon\ x_1 = \ldots = x_n\}$ belongs to $\bigotimes_{i=1}^n \mathcal{B}_i$. It is clear from the construction that $\mu_{(1,\ldots,n)}(\Delta_n) = 1$ for all n. If we had a measure μ on X with the projections $\mu_{(1,\ldots,n)}$, then we would obtain $\mu(\Omega_n) = 1$ for all sets $\Omega_n = \Delta_n \times \prod_{k=n+1}^\infty X_k$. However, this is impossible for a countably additive measure, since $\bigcap_{n=1}^\infty \Omega_n = \varnothing$ due to the equality $\bigcap_{n=1}^\infty X_n = \varnothing$. □

In books on probability theory and random processes, the Kolmogorov theorem appears in the context of the distributions of random processes. We recall the corresponding terminology. A random process $\xi = (\xi_t)_{t\in T}$ on a nonempty set T is just a family of measurable functions ξ_t indexed by points $t \in T$ and defined on a probability space $(\Omega, \mathcal{A}, P)$. For every ordered finite collection of distinct points $t_1, \ldots, t_n \in T$, one obtains a Borel probability measure on $\mathbb{R}^n$ defined by

$$P_{t_1,\ldots,t_n}(B) := P\Big(\omega\colon\ \big(\xi_{t_1}(\omega), \ldots, \xi_{t_n}(\omega)\big) \in B\Big).$$

This measure is called a finite-dimensional distribution of the process ξ. The finite-dimensional distributions are consistent in the following sense:

(1) the image of the measure $P_{t_1,\ldots,t_n,s_1,\ldots,s_k}$ under the projection from $\mathbb{R}^{n+k}$ to $\mathbb{R}^n$ coincides with $P_{t_1,\ldots,t_n}$ for all t_i and s_j,

(2) for every permutation σ of the set $\{1, \ldots, n\}$, one has

$$P_{t_{\sigma(1)},\ldots,t_{\sigma(n)}} = P_{t_1,\ldots,t_n} \circ T^{-1},$$

where $T\colon\ \mathbb{R}^n \to \mathbb{R}^n$, $T(x_1, \ldots, x_n) = (x_{\sigma(1)}, \ldots, x_{\sigma(n)})$.

The latter property enables one to define the measures μ_Λ for subsets Λ in T consisting of all (not ordered) collections t_i, $i = 1, \ldots, n$ (note that Theorem 7.7.1 deals merely with subsets of T without any ordering or numbering, so that $\{t_1, t_2\}$ is the same subset as $\{t_2, t_1\}$). Namely, if we fix an arbitrary enumeration of the points $t_1, \ldots, t_n$, then every set $B \in \mathcal{B}(\mathbb{R}^\Lambda)$ is identified with some set $B' \in \mathcal{B}(\mathbb{R}^n)$. Hence one can set

$$P_\Lambda\big(x \in \mathbb{R}^\Lambda\colon\ x \in B\big) := P_{t_1,\ldots,t_n}(B'),$$

which gives a well-defined object due to the foregoing consistency condition. Certainly, it is possible to consider the distributions $P_{t_1,\ldots,t_n}$ with multiple points t_i, but this is not necessary for applying Theorem 7.7.1.

The Kolmogorov theorem states the converse: given a nonempty set T and a family of consistent (in the sense of conditions (1) and (2)) measures $P_{t_1,\ldots,t_n}$ on the spaces $\mathbb{R}^n$ for all distinct $t_i \in T$, there exist a probability space and a random process ξ on it whose finite-dimensional distributions are $P_{t_1,\ldots,t_n}$. For a probability space Ω one can take the space $\mathbb{R}^T$, and for P the measure μ from the Kolmogorov theorem, in which for any $\Lambda = \{t_1, \ldots, t_n\}$ we set $\mu_\Lambda := P_{t_1,\ldots,t_n}$. Any point $\omega \in \Omega$ is a function on T and we set $\xi_t(\omega) := \omega(t)$. It is clear that we obtain a random process with the required properties. The constructed measure μ on $\mathbb{R}^T$ is called the distribution of the process ξ in the path space (the space of trajectories) and is denoted by μ_ξ. The Kolmogorov theorem can be alternatively formulated as follows: an additive set function on the cylindrical algebra in $\mathbb{R}^T$ with countably additive finite-dimensional projections is itself countably additive.

In applications of Theorem 7.7.1 the following problem is typical. Usually, it is clear that the random process ξ, the distribution of which is constructed in this theorem, possesses trajectories with certain additional properties (for example, continuous), and it is desirable that the corresponding measure μ_ξ be concentrated on the set X_0 of such trajectories. However, the straightforward application of the Kolmogorov theorem does not guarantee this in most of the cases because the set X_0 turns out to be nonmeasurable with respect to μ_ξ. A trivial example: the process is identically 0 and X_0 is a point (see Example 7.3.1). The effect of this in the study of random processes is that the distribution of a process does not determine the process uniquely (in particular, does not uniquely determine the properties of its trajectories as functions of t). For example, it can occur that two processes ξ and η have equal distributions, but $\xi_t(\omega) = 0$ for all t and ω, whereas for every ω, there exists t with $\eta_t(\omega) = 1$. To this end, it suffices to take $\Omega = [0,1]$ with Lebesgue measure and set $\eta_t(t) = 1$ and $\eta_t(\omega) = 0$ if $\omega \neq t$. Several standard tricks are known to circumvent the obstacle. A natural and efficient procedure (going back to Kolmogorov) is to verify the equality $\mu_\xi^*(X_0) = 1$, which enables one to restrict μ to the set X_0 of full outer measure. Let us formulate another important theorem of Kolmogorov that gives a constructive sufficient condition of the above equality (we do not include a proof, since it is found in many textbooks, see, e.g., Wentzell [**1973**, §5.2]).

7.7.4. Theorem. *Suppose that a random process ξ on a set $T \subset \mathbb{R}^1$ satisfies the following condition*:

$$\mathbb{E}|\xi_t - \xi_s|^\alpha \le L|t-s|^{1+\beta},$$

where L, α, β are positive numbers and $\mathbb{E}$ is the expectation (*i.e., the integral*). *Then $\mu_\xi^*\bigl(C(T)\bigr) = 1$.*

By means of two Kolmogorov's theorems given above one can easily justify the existence of the Wiener measure on $C[0,1]$, i.e., a measure μ_W such that every functional $x \mapsto x(t) - x(s)$ is a Gaussian random variable with

$$\int_{C[0,1]} x(t)\,\mu_W(dx) = 0, \qquad \int_{C[0,1]} |x(t) - x(s)|^2\,\mu_W(dx) = |t-s|,$$

and, additionally, for all $t_1 < t_2 < \ldots < t_n$, the functionals $x(t_{i+1}) - x(t_i)$ are independent and $x(0) = 0$ for μ_W-a.e. x. Regarding this see Bogachev [**208**].

In the literature, one can find diverse sufficient conditions for various sets X_0 (for example, functions without discontinuities of second order); see Gikhman, Skorokhod [**685**]. We remark that certain additional problems arise in the case where for X_0 one has to take a space whose elements are equivalence classes rather than individual functions (for example, L^2). One more procedure of fighting the arising nonmeasurabilities goes back to Doob and employs his concept of a separable random process (see details in Doob [**467**] and Neveu [**1368**]).

7.8. The Daniell integral

The construction of the integral presented in this book is based on a preliminary introduction of a measure. However, it is possible to go in the opposite direction and define measures by means of integrals. The following result due to Daniell is at the basis of this approach. In the formulation we use the concept of a vector lattice of functions, i.e., a linear space of real functions on a nonempty set Ω such that $\max(f,g) \in \mathcal{F}$ for all $f,g \in \mathcal{F}$. Note that then one has $\min(f,g) = -\max(-f,-g) \in \mathcal{F}$ and $|f| \in \mathcal{F}$ for all $f \in \mathcal{F}$. Since $\max(f,g) = (|f-g| + f + g)/2$, it would be sufficient to require only that $\mathcal{F}$ be a linear space closed with respect to taking the absolute values. A vector lattice of functions is a particular case of an abstract vector lattice, i.e., a linear space with a lattice structure that is consistent with the linear structure in the sense that $\alpha x \le \beta x$ if $x \ge 0$, $\alpha, \beta \in [0,\infty)$, and $x+z \le y+z$ if $x \le y$. As an example one can take $L^p[0,1]$.

7.8.1. Theorem. *Let $\mathcal{F}$ be a vector lattice of functions on a set Ω such that $1 \in \mathcal{F}$. Let L be a linear functional on $\mathcal{F}$ with the following properties: $L(f) \ge 0$ whenever $f \ge 0$, $L(1) = 1$, and $L(f_n) \to 0$ for every sequence of functions f_n in $\mathcal{F}$ monotonically decreasing to zero. Then, there exists a unique probability measure μ on the σ-algebra $\mathcal{A} = \sigma(\mathcal{F})$ generated by $\mathcal{F}$ such that $\mathcal{F} \subset \mathcal{L}^1(\mu)$ and*

$$L(f) = \int_\Omega f\, d\mu, \quad \forall\, f \in \mathcal{F}. \tag{7.8.1}$$

PROOF. (i) Denote by $\mathcal{L}^+$ the set of all bounded functions f of the form $f(x) = \lim\limits_{n\to\infty} f_n(x)$, where $f_n \in \mathcal{F}$ are nonnegative and the sequence $\{f_n\}$ is increasing. Clearly, the sequence $\{f_n\}$ is uniformly bounded, hence the sequence $\{L(f_n)\}$ is increasing and bounded by the properties of L. Set $L(f) = \lim\limits_{n\to\infty} L(f_n)$. We show that the extended functional is well-defined, coincides on bounded nonnegative functions in $\mathcal{F}$ with the initial functional and possesses the following properties:

(1) $L(f) \le L(g)$ for all $f,g \in \mathcal{L}^+$ with $f \le g$;

(2) $L(f+g) = L(f) + L(g)$, $L(cf) = cL(f)$ for all $f,g \in \mathcal{L}^+$ and all $c \in [0,+\infty)$;

(3) $\min(f,g) \in \mathcal{L}^+$, $\max(f,g) \in \mathcal{L}^+$ for all $f,g \in \mathcal{L}^+$ and

$$L(f) + L(g) = L\bigl(\min(f,g)\bigr) + L\bigl(\max(f,g)\bigr);$$

(4) $\lim\limits_{n\to\infty} f_n \in \mathcal{L}^+$ for every uniformly bounded increasing sequence of functions $f_n \in \mathcal{L}^+$, and one has $L(\lim\limits_{n\to\infty} f_n) = \lim\limits_{n\to\infty} L(f_n)$.

We observe that if $\{f_n\}$ and $\{g_k\}$ are two increasing sequences of nonnegative functions in $\mathcal{F}$ with $\lim\limits_{n\to\infty} f_n \le \lim\limits_{k\to\infty} g_k$, then $\lim\limits_{n\to\infty} L(f_n) \le \lim\limits_{k\to\infty} L(g_k)$.

Indeed, it follows by the hypotheses of the theorem that $L(\psi_m) \to L(\psi)$ if nonnegative functions ψ_m in $\mathcal{F}$ are decreasing to $\psi \in \mathcal{F}$. The functions $\min(f_n, g_k) \in \mathcal{F}$ are increasing to f_n as $k \to \infty$, since $f_n \le \lim\limits_{k\to\infty} g_k$. Hence

$$L(f_n) = \lim_{k\to\infty} L\bigl(\min(f_n, g_k)\bigr) \le \lim_{k\to\infty} L(g_k).$$

It remains to take the limit as $n \to \infty$. This shows that L on $\mathcal{L}^+$ is well-defined, i.e., is independent of our choice of an increasing sequence convergent to an element in $\mathcal{L}^+$. In particular, we obtain that on $\mathcal{F}\cap\mathcal{L}^+$ the constructed functional coincides with the initial one. Properties (1) and (2) now follow at once from the fact that they hold for functions in $\mathcal{F}$. We have $\max(f,g) = \lim\limits_{n\to\infty} \max(f_n, g_n)$ and $\min(f,g) = \lim\limits_{n\to\infty} \min(f_n, g_n)$ if nonnegative functions $f_n, g_n \in \mathcal{F}$ are increasing to f and g, respectively. In addition, both limits are monotone. Hence Property (3) follows by definition and the obvious equality $\max(f,g) + \min(f,g) = f + g$. Let us verify (4). Suppose that nonnegative functions $f_{k,n} \in \mathcal{F}$ are increasing to $f_n \in \mathcal{L}^+$ as $k \to \infty$. Set $g_m = \max\limits_{n\le m} f_{m,n}$. Then $g_m \in \mathcal{F}$, $g_m \le g_{m+1}$ and $f_{m,n} \le g_m \le f_m$ if $n \le m$. Hence we have $L(g_m) \le L(g_{m+1})$ and $L(f_{m,n}) \le L(g_m) \le L(f_m)$ if $n \le m$. Therefore, $\lim\limits_{m\to\infty} f_m = \lim\limits_{m\to\infty} g_m \in \mathcal{L}^+$ and

$$\lim_{m\to\infty} L(f_m) = \lim_{m\to\infty} L(g_m) = L(\lim_{m\to\infty} g_m) = L(\lim_{m\to\infty} f_m).$$

(ii) Denote by $\mathcal{G}$ the class of all sets G with $I_G \in \mathcal{L}^+$. Set $\mu(G) = L(I_G)$ for all $G \in \mathcal{G}$. We observe that $I_{G_1\cap G_2} = \min(I_{G_1}, I_{G_2})$, $I_{G_1\cup G_2} = \max(I_{G_1}, I_{G_2})$. Hence by Property (3) established in (i), the class $\mathcal{G}$ is closed with respect to finite intersections and finite unions, then also with respect to countable unions by Property (4). In addition, μ is a nonnegative monotone additive function on $\mathcal{G}$, and one has

$$\mu(G_1 \cap G_2) + \mu(G_1 \cup G_2) = \mu(G_1) + \mu(G_2)$$

for all $G_1, G_2 \in \mathcal{G}$, and $\mu(G) = \lim\limits_{n\to\infty} \mu(G_n)$ if the sets $G_n \in \mathcal{G}$ are increasing to G. Note also that $\mu(\Omega) = 1$. According to Theorem 1.11.4 (applicable in view of Example 1.11.5 and the fact that $\mathcal{G}$ is closed with respect to countable unions), the function

$$\mu^*(A) = \inf\{\mu(G)\colon\ G \in \mathcal{G},\ A \subset G\}$$

is a countably additive measure on the class

$$\mathcal{B} = \bigl\{B \subset \Omega\colon\ \mu^*(B) + \mu^*(\Omega\backslash B) = 1\bigr\}.$$

We shall denote by μ the restriction of μ^* to $\mathcal{B}$.

(iii) We verify that $\mathcal{A} \subset \mathcal{B}$. If $f \in \mathcal{L}^+$, then $\{f > c\} \in \mathcal{G}$ for all c, since

$$I_{\{f>c\}} = \lim_{n\to\infty} \min\bigl(1, n\max(f - c, 0)\bigr).$$

Hence all functions in $\mathcal{L}^+$ are measurable with respect to the σ-algebra $\sigma(\mathcal{G})$. On the other hand, all such functions are measurable with respect to the σ-algebra $\mathcal{A}$ generated by the class $\mathcal{F}$. Since $\mathcal{G} \subset \sigma(\mathcal{L}^+) = \sigma(\mathcal{F})$, we obtain the equality $\mathcal{A} = \sigma(\mathcal{G})$. Thus, it suffices to show that $\mathcal{G} \subset \mathcal{B}$. Let $G \in \mathcal{G}$. We take an increasing sequence of nonnegative functions $f_n \in \mathcal{F}$ with $I_G = \lim\limits_{n\to\infty} f_n$. Then $\mu^*(G) = \mu(G) = \lim\limits_{n\to\infty} L(f_n)$. Since $\mu^*(G) + \mu^*(\Omega\backslash G) \geq 1$, in order to show the inclusion $G \in \mathcal{B}$, it suffices to prove that $\mu^*(G) + \mu^*(\Omega\backslash G) \leq 1$, which is equivalent to the inequality

$$\mu^*(\Omega\backslash G) \leq \lim_{n\to\infty} L(1 - f_n). \tag{7.8.2}$$

The functions $1 - f_n$ are decreasing to $I_{\Omega\backslash G}$. For any n and any $c \in (0, 1)$, the set $U_c = \{1 - f_n > c\}$ contains $\Omega\backslash G$ and by the above belongs to $\mathcal{G}$. Therefore, the obvious inequality $I_{U_c} \leq c^{-1}(1 - f_n)$ yields

$$\mu^*(\Omega\backslash G) \leq \mu(U_c) = L(I_{U_c}) \leq c^{-1}L(1 - f_n).$$

Letting $c \to 1$ and then $n \to \infty$, we obtain (7.8.2).

(iv) It remains to prove that $\mathcal{F} \subset \mathcal{L}^1(\mu)$ and that (7.8.1) is true. We know that all functions in $\mathcal{L}^+$ are $\mathcal{A}$-measurable. If $f = I_G$, where $G \in \mathcal{G}$, then the required equality is fulfilled by the definition of μ. Clearly, this equality remains true for any finite linear combinations of indicators of sets in $\mathcal{G}$. Let $f \in \mathcal{L}^+$ and $f \leq 1$. Then f is the limit of the increasing sequence of functions

$$f_n := \sum_{j=1}^{2^n-1} j2^{-n} I_{\{j2^{-n} < f \leq (j+1)2^{-n}\}} = 2^{-n} \sum_{j=1}^{2^n-1} I_{\{f > j2^{-n}\}}.$$

It follows that

$$L(f_n) = \int_\Omega f_n\, d\mu.$$

Property (4) established in (i) and the properties of the integral show that as $n \to \infty$, the right-hand side and left-hand side of this equality converge to $L(f)$ and

$$\int_\Omega f\, d\mu,$$

respectively. Moreover, by the same reasoning (7.8.1) extends to all nonnegative functions $f \in \mathcal{F}$, since $f = \lim\limits_{n\to\infty} \min(f, n)$ and $\min(f, n) \in \mathcal{L}^+$. Finally, for any function $f \in \mathcal{F}$, we have $f = \max(f, 0) - \max(-f, 0)$, which yields our assertion.

The uniqueness of μ satisfying (7.8.1) follows from the fact that it is uniquely determined on the class $\mathcal{G}$, which is closed with respect to finite intersections and generates $\mathcal{A}$. □

A function L with the properties listed in the above theorem is called the Daniell integral (see below the case $1 \notin \mathcal{F}$).

7.8.2. Corollary. *Suppose that in Theorem* 7.8.1 *the class $\mathcal{F}$ is closed with respect to uniform convergence. Let $\mathcal{G}_\mathcal{F}$ be the class of all sets of the form $\{f > 0\}$, $f \in \mathcal{F}$, $f \geq 0$. Then $\mathcal{G}_\mathcal{F}$ generates the σ-algebra $\mathcal{A} = \sigma(\mathcal{F})$, and one has the equalities*

$$\mu(A) = \inf\bigl\{\mu(G)\colon\ A \subset G, G \in \mathcal{G}_\mathcal{F}\bigr\}, \quad \forall\, A \in \mathcal{A}, \tag{7.8.3}$$

$$\mu(G) = \sup\bigl\{L(f)\colon\ f \in \mathcal{F}, 0 \le f \le I_G\bigr\}, \quad \forall\, G \in \mathcal{G}_{\mathcal{F}}. \tag{7.8.4}$$

PROOF. It suffices to verify that the class $\mathcal{G}_{\mathcal{F}}$ coincides with the class $\mathcal{G}$ introduced in the proof of the theorem. It has been shown in that proof that $\{f > 0\} \in \mathcal{G}$ for all nonnegative $f \in \mathcal{F}$. On the other hand, if $G \in \mathcal{G}$, then by definition there exists an increasing sequence of nonnegative functions $f_n \in \mathcal{F}$ convergent to I_G. Set $f = \sum_{n=1}^{\infty} 2^{-n} f_n$. By uniform convergence of the series we have $f \in \mathcal{F}$. It is clear that $f \ge 0$ and $G = \{f > 0\}$. □

Functionals considered in the above theorem are called positive. Thus, the expression $L \ge 0$ means that $L(f) \ge 0$ if $f \ge 0$. However, this theorem extends to not necessarily positive functionals.

7.8.3. Theorem. *Let $\mathcal{F}$ be a vector lattice of bounded functions on a set Ω such that $1 \in \mathcal{F}$. Suppose that we are given a linear functional L on $\mathcal{F}$ that is continuous with respect to the norm $\|f\| = \sup_\Omega |f(x)|$. Then L can be represented in the form $L = L^+ - L^-$, where $L^+ \ge 0$, $L^- \ge 0$, and for all nonnegative $f \in \mathcal{F}$ one has*

$$L^+(f) = \sup_{0 \le g \le f} L(g), \quad L^-(f) = -\inf_{0 \le g \le f} L(g). \tag{7.8.5}$$

In addition, letting $|L| := L^+ + L^-$, we have for all $f \ge 0$

$$|L|(f) = \sup_{0 \le |g| \le f} |L(g)|, \quad \|L\| = L^+(1) + L^-(1).$$

PROOF. Given two nonnegative functions $f, g \in \mathcal{F}$ and a function $h \in \mathcal{F}$ such that $0 \le h \le f + g$, we can write $h = h_1 + h_2$, where $h_1,\ h_2 \in \mathcal{F}$, $0 \le h_1 \le f$, $0 \le h_2 \le g$. Indeed, let $h_1 = \min(f, h)$, $h_2 = h - h_1$. Then $h_1,\ h_2 \in \mathcal{F}$, $0 \le h_1 \le f$ and $h_2 \ge 0$. Finally, $h_2 \le g$. For, if $h_1(x) = h(x)$, then $h_2(x) = 0$, and if $h_1(x) = f(x)$, then $h_2(x) = h(x) - f(x) \le g(x)$, since $h \le g + f$.

Let L^+ be defined by equality (7.8.5). Note that the quantity $L^+(f)$ is finite, since $|L(h)| \le \|L\|\,\|h\| \le \|L\|\,\|f\|$. It is clear that $L^+(tf) = tL^+(f)$ for all nonnegative numbers t and $f \ge 0$. Let $f \ge 0$ and $g \ge 0$ be in $\mathcal{F}$. Keeping the above notation we obtain

$$\begin{aligned} L^+(f+g) &= \sup\{L(h)\colon\ 0 \le h \le f + g\} \\ &= \sup\{L(h_1) + L(h_2)\colon\ 0 \le h_1 \le f,\ 0 \le h_2 \le g\} = L^+(f) + L^+(g). \end{aligned}$$

Now for all $f \in \mathcal{F}$ we set $L^+(f) = L^+(f^+) - L^+(f^-)$, where $f^+ = \max(f, 0)$, $f^- = -\min(f, 0)$. Note that if $f = f_1 - f_2$, where $f_1, f_2 \ge 0$, then

$$L^+(f) = L^+(f_1) - L^+(f_2).$$

Indeed, $f_1 + f^- = f_2 + f^+$, hence $L^+(f_1) + L^+(f^-) = L^+(f_2) + L^+(f^+)$. It is clear that $L^+(tf) = tL^+(f)$ for all $t \in \mathbb{R}^1$ and $f \in \mathcal{F}$. The additivity of the functional L^+ follows by its additivity on nonnegative functions. Indeed, given f and g, we can write $f = f^+ - f^-$, $g = g^+ - g^-$, whence we have

$f + g = (f^+ + g^+) - (f^- + g^-)$, and according to what has been said above we obtain

$$L^+(f + g) = L^+(f^+ + g^+) - L^+(f^- + g^-) = L^+(f) + L^+(g).$$

By definition, one has $L^+(f) \geq L(f)$ for nonnegative f, hence the functional $L^- := L^+ - L$ is nonnegative. It is easy to see that L^- is given by the stated formula.

Finally, $\|L\| \leq \|L^+\| + \|L^-\| = L^+(1) + L^-(1)$. On the other hand,

$$\begin{aligned} L^+(1) + L^-(1) = 2L^+(1) - L(1) &= \sup\{L(2\varphi - 1)\colon\ 0 \leq \varphi \leq 1\} \\ &\leq \sup\{L(h)\colon\ -1 \leq h \leq 1\} \leq \|L\|. \end{aligned}$$

The theorem is proven. □

7.8.4. Corollary. *Suppose that in the situation of the previous theorem the functional L has the following property: $L(f_n) \to 0$ for every sequence of functions f_n in $\mathcal{F}$ monotonically decreasing to zero. Then the functionals L^+ and L^- have this property as well. In particular, L^+ and L^- are defined by nonnegative countably additive measures on $\sigma(\mathcal{F})$ and L has representation* (7.8.1) *with some signed countably additive measure μ on $\sigma(\mathcal{F})$.*

Proof. Let $\{f_n\}$ be a sequence in $\mathcal{F}$ monotonically decreasing to zero and let $\varepsilon > 0$. By definition one can find $\varphi_n \in \mathcal{F}$ with $0 \leq \varphi_n \leq f_n$ and $L(\varphi_n) \geq L^+(f_n) - \varepsilon 2^{-n}$. Set $g_n = \min(\varphi_1, \ldots, \varphi_n)$. We verify by induction that

$$L^+(f_n) \leq L(g_n) + \varepsilon \sum_{i=1}^{n} 2^{-i}. \tag{7.8.6}$$

This is true if $n = 1$. Suppose that (7.8.6) is true for $n = 1, \ldots, m$. One has the equalities

$$g_{m+1} = \min(g_m, \varphi_{m+1}),$$
$$\max(g_m, \varphi_{m+1}) + \min(g_m, \varphi_{m+1}) = g_m + \varphi_{m+1},$$

whence

$$\begin{aligned} L\bigl(\max(g_m, \varphi_{m+1})\bigr) + L(g_{m+1}) &= L(g_m) + L(\varphi_{m+1}) \\ &\geq L(g_m) + L^+(f_{m+1}) - \varepsilon 2^{-m-1}. \end{aligned}$$

On other hand, the estimates $g_m \leq \varphi_m \leq f_m$, $\varphi_{m+1} \leq f_{m+1} \leq f_m$ and the inductive assumption yield

$$L\bigl(\max(g_m, \varphi_{m+1})\bigr) \leq L^+(f_m) \leq L(g_m) + \varepsilon \sum_{i=1}^{m} 2^{-i}.$$

Therefore,

$$L(g_m) + L^+(f_{m+1}) - \varepsilon 2^{-m-1} - L(g_{m+1}) \leq L(g_m) + \varepsilon \sum_{i=1}^{m} 2^{-i},$$

whence we obtain (7.8.6) for $n = m + 1$. Thus, (7.8.6) is established for all n. Since $g_n \leq f_n$, the sequence $\{g_n\}$ is decreasing to zero. Therefore, $L(g_n) \to 0$

and (7.8.6) yields $\limsup L^+(f_n) \le \varepsilon$. Since $\varepsilon > 0$ is arbitrary and $L^+(f_n)$ is nonnegative, we obtain that $L^+(f_n) \to 0$. The claim for L^- follows too. $\square$

7.8.5. Remark. In the above corollary, the functionals L^+ and L^- are represented by the measures μ^+ and μ^-, where μ represents L. This can be easily seen from (7.8.5) and the properties of the integral.

7.8.6. Theorem. *Let $\mathcal{F}$ be a vector lattice of functions on a set Ω such that $1 \in \mathcal{F}$. Suppose that L is a linear functional on $\mathcal{F}$ with the following properties: $L(f) \ge 0$ if $f \ge 0$, $L(1) = 1$, and $L(f_\alpha) \to 0$ for every net of functions f_α in $\mathcal{F}$ monotonically decreasing to zero. Then, there exists a unique probability measure μ on the σ-algebra $\mathcal{A} = \sigma(\mathcal{F})$ generated by $\mathcal{F}$ such that $\mathcal{F} \subset \mathcal{L}^1(\mu)$ and (7.8.1) holds. In addition, $\mu(G_\alpha) \to \mu\bigl(\bigcup_\alpha G_\alpha\bigr)$ for every increasing net of sets G_α such that $I_{G_\alpha} \in \mathcal{L}^+$, where $\mathcal{L}^+$ is the class of all bounded functions that are the limits of increasing nets of nonnegative functions in $\mathcal{F}$.*

Proof. The reasoning in the proof of Theorem 7.8.1, where we dealt with σ-additive functionals, applies with minor changes. We take for $\mathcal{L}^+$ the class of all bounded functions f representable as the limits of increasing nets of nonnegative functions f_α in $\mathcal{F}$. The extension of L to $\mathcal{L}^+$ is defined as in Theorem 7.8.1 with nets in place of sequences. All the arguments remain valid and show that the extension possesses the following property: if an increasing net of functions $f_\alpha \in \mathcal{L}^+$ converges to a function $f \in \mathcal{L}^+$, then $L(f_\alpha) \to L(f)$. As in the cited theorem, we obtain a countably additive measure on the σ-algebra $\sigma(\mathcal{L}^+)$ generated by $\mathcal{L}^+$ such that the following equalities hold:

$$\mu(G) = L(I_G),\ G \in \mathcal{G} := \{G\colon\ I_G \in \mathcal{L}^+\},\ \mu(B) = \inf\bigl\{\mu(G)\colon\ G \in \mathcal{G},\ B \subset G\bigr\},$$

$$\int_\Omega f\,d\mu = L(f) \quad \text{for all } f \in \mathcal{L}^+.$$

Moreover, $\mathcal{F} \subset \mathcal{L}^1(\mu)$ and the previous equality holds for all $f \in \mathcal{F}$. It should be noted that in this situation the σ-algebra $\mathcal{A} = \sigma(\mathcal{F})$ may be strictly smaller than $\sigma(\mathcal{L}^+)$. It is clear from the construction that if an increasing net of sets G_α gives in the union the set G, then $\mu(G_\alpha) = L(I_{G_\alpha}) \to L(I_G) = \mu(G)$. $\square$

We assumed in the above results that the lattice $\mathcal{F}$ contains 1. For this reason they are not applicable so far to constructing infinite measures. It turns out that if $1 \notin \mathcal{F}$, then the above conditions are not sufficient for the existence of a representing measure. One can construct an example of a set Ω, a vector lattice $\mathcal{F}$ of functions on Ω, and a positive τ-smooth linear functional on $\mathcal{F}$ that is not representable as the integral, see Fremlin, Talagrand [**639**], Fremlin [**635**, §439H]. We give below a similar example (borrowed from Fremlin [**619**]) with a σ-smooth functional. However, one can improve the situation by adding the *Stone condition*:

$$\min(f, 1) \in \mathcal{F} \quad \text{for all } f \in \mathcal{F}.$$

A natural example of a lattice satisfying the Stone condition and not containing 1 is the space of all continuous functions with compact support on $\mathbb{R}^n$. The proof of the following theorem is delegated to Exercise 7.14.126 (it can also be derived from the previous results).

7.8.7. Theorem. *Let $\mathcal{F}$ be a vector lattice of functions on a set Ω satisfying the Stone condition. Suppose that L is a nonnegative linear functional on $\mathcal{F}$ such that $L(f_n) \to 0$ for every sequence of functions $f_n \in \mathcal{F}$ pointwise decreasing to zero. Then, there exists a countably additive measure μ defined on $\sigma(\mathcal{F})$ and having values in $[0, +\infty]$ such that $\mathcal{F} \subset \mathcal{L}^1(\mu)$ and (7.8.1) is fulfilled.*

In place of the Stone condition one can sometimes use the following condition: there exists a sequence of nonnegative functions $\varphi_n \in \mathcal{F}$ increasing to 1 (see Hirsch, Lacombe [**834**, p. 58]). One can verify that the Stone condition is fulfilled on the space $\mathcal{L}$ of all functions f such that f^+ and f^- belong to the class $\mathcal{V}$ of all functions of the form $g = \lim\limits_{n\to\infty} g_n$, where $\{g_n\}$ is increasing, $g_n \in \mathcal{F}$ and $\sup L(g_n) < \infty$. The functional L extends to $\mathcal{V}$ by monotonicity and then to $\mathcal{L}$ by linearity. The measure μ generating L is σ-finite in this case, since $\mu(\{\varphi_n > 1/k\}) < \infty$. Hence the aforementioned condition is more restrictive than that of Stone.

As a simple corollary of the above results one obtains the existence of the Lebesgue integral on $\mathbb{R}^n$ or on a cube. To this end, we take for $\mathcal{F}$ the class of all continuous functions with bounded support (observe that every sequence of such functions pointwise decreasing to zero converges uniformly) and for L we take the Riemann integral. The same method works for constructing the Lebesgue integral on any sufficiently regular manifold (certainly, it is necessary that the Riemann integral of continuous functions be defined).

We now proceed to the aforementioned example of non-existence of representing measures.

7.8.8. Example. Let $\mathcal{F}$ be the set of all real functions f on $[0,1]$ with the following property: for some number $\alpha = \alpha(f)$, the set $\bigl\{t\colon\ f(t) \neq \alpha(1+t)\bigr\}$ is a first category set. Let $L(f) := \alpha$. Then $\mathcal{F}$ is a vector lattice of functions with the natural order on $\mathbb{R}^{[0,1]}$, L is a nonnegative linear functional on $\mathcal{F}$, and $L(f_n) \to 0$ for every sequence functions $f_n \in \mathcal{F}$ pointwise decreasing to zero, but L cannot be represented as the integral with respect to a countably additive measure.

PROOF. We observe that for each function $f \in \mathcal{F}$, there is only one number α with the indicated property, since the interval is not a first category set. Hence the function L is well-defined. Given $f \in \mathcal{F}$, we set

$$E_f := \bigl\{t\colon\ f(t) \neq \alpha(1+t)\bigr\},$$

where α is the number corresponding to f. If $f, g \in \mathcal{F}$ and $\alpha = \alpha(f)$, $\beta = \alpha(g)$ are the corresponding numbers, then $E_f \cup E_g$ is a first category set, and one has $f(t) + g(t) = (\alpha + \beta)(1 + t)$ outside it. For any real c

we have $cf(t) = c\alpha(1+t)$ outside the set E_f. Thus, $\mathcal{F}$ is a linear space. It is easily seen that $|f| \in \mathcal{F}$ if $f \in \mathcal{F}$. It is also clear that the function L is linear. If $f \geq 0$, then $L(f) \geq 0$. If functions $f_n \in \mathcal{F}$ are pointwise decreasing to zero, then the union of the sets E_{f_n} is a first category set. Hence there exists a point t such that $L(f_n) = f_n(t)/(1+t)$ simultaneously for all n, whence $\lim\limits_{n\to\infty} L(f_n) = 0$. Suppose now that there exists a measure μ on $\sigma(\mathcal{F})$ with values in $[0, +\infty]$ such that $\mathcal{F} \subset \mathcal{L}^1(\mu)$ and $L(f)$ coincides with the integral of f against the measure μ. The function $\psi\colon\ t \mapsto 1+t$ belongs to $\mathcal{F}$, whence we obtain that all open sets in $[0,1]$ belong to $\sigma(\mathcal{F})$. The estimate $\psi \geq 1$ yields that $\mu([0,1]) \leq L(\psi) = 1$. Thus, the restriction of μ to $\mathcal{B}([0,1])$ is a finite measure. Therefore, there exists a first category Borel set E such that $\mu([0,1]\backslash E) = 0$. Indeed, one can take the union of nowhere dense compact sets K_n with $\mu([0,1]\backslash K_n) < 1/n$, which can be constructed by deleting sufficiently small open intervals centered at the points of a countable dense set of μ-measure zero. Let us consider the following function f: $f(t) = 0$ if $t \in E$, $f(t) = 1+t$ if $t \notin E$. It is clear that $f \in \mathcal{F}$ and $L(f) = 1$. On the other hand, the integral of f with respect to the measure μ is zero, which is a contradiction. □

This example shows that one cannot always represent L as an integral, but a closer look at the proof of Theorem 7.8.1 reveals that even without the Stone condition one obtains the functional L with the basic properties of the integral (which explains the term “the Daniell integral”). Let $\mathcal{F}$ be some vector lattice of functions on a set Ω and let L be a nonnegative linear functional on $\mathcal{F}$ such that $L(f_n) \to 0$ for every sequence $\{f_n\} \subset \mathcal{F}$ pointwise decreasing to zero. We shall use the term an L-zero set for sets $S \subset \Omega$ with the property that for every $\varepsilon > 0$, there exists an increasing sequence of functions $f_n \geq 0$ in $\mathcal{F}$ such that $L(f_n) < \varepsilon$ and $\sup_n f_n(x) \geq 1$ on S. Let $\mathcal{L}^+$ denote the class of all functions f with values in $(-\infty, +\infty]$ for which one can find an increasing sequence $\{f_n\} \subset \mathcal{F}$ such that $f(x) = \lim\limits_{n\to\infty} f_n(x)$ outside some L-zero set and the sequence $L(f_n)$ is bounded. It is readily verified that such a function f is finite outside some L-zero set. Set $L(f) := \lim\limits_{n\to\infty} L(f_n)$. The reasoning in the proof of Theorem 7.8.1 shows that L is well-defined on $\mathcal{L}^+$. Let $\mathcal{L}$ denote the set of all functions f with $f^+, f^- \in \mathcal{L}^+$. For such functions, we set $L(f) := L(f^+) - L(f^-)$. The class $\mathcal{L}$ is equipped with the following equivalence relation: two functions are equivalent if the set on which they differ is L-zero. Then the set $\widetilde{\mathcal{L}}$ of all equivalence classes becomes a metric space with the metric $d_L(f,g) := L(|f-g|)$. In addition, $\widetilde{\mathcal{L}}$ is a linear space. It is clear by construction that $\mathcal{F}$ is everywhere dense in $\mathcal{L}$.

7.8.9. Proposition. *The functional L on $\mathcal{L}$ is linear, and the statements of the Beppo Levi, Lebesgue, and Fatou theorems are true if the integral in their formulations is replaced by L. In addition, $\widetilde{\mathcal{L}}$ is complete with respect to the metric d_L.*

PROOF. We give only a sketch of the proof; more details can be found in Shilov, Gurevich [**1699**, §2]. It is easily verified that L is linear on $\mathcal{L}$ and $L(f) \le L(g)$ if $f \le g$. We observe that the union S of L-zero sets S_n is an L-zero set. Indeed, given $\varepsilon > 0$, for every n, there is an increasing sequence of functions $f_{n,k} \ge 0$ in $\mathcal{F}$ with $L(f_{n,k}) \le \varepsilon 2^{-n}$ and $\sup_k f_{n,k}(x) \ge 1$ on S_n. Let $g_n := f_{n,1} + \cdots + f_{n,n}$. Then $\{g_n\}$ is increasing, $g_n \ge 0$, $L(g_n) \le \varepsilon$, and $\sup_n g_n(x) \ge 1$ on S. Suppose the sequence of functions $f_n \in \mathcal{L}$ is increasing outside an L-zero set and $\{L(f_n)\}$ is bounded. Passing to $f_n - f_1$ we may assume that $f_n \ge 0$. For every n, we can find $f_{n,k} \in \mathcal{F}$ increasing to f_n outside some L-zero set S_n. Let $g_n = \max_{k,m \le n} f_{m,k}$. Then $g_n \in \mathcal{F}$, $\{g_n\}$ is increasing and $\{L(g_n)\}$ is bounded. Then $f = \lim\limits_{n\to\infty} g_n \in \mathcal{L}^+$ and $L(f) = \lim\limits_{n\to\infty} L(g_n)$. Clearly, $f_n(x) \to f(x)$ outside an L-zero set and $L(f) = \lim\limits_{n\to\infty} L(f_n)$ because $L(g_n) \le L(f_n)$ and $L(f_n) = \lim\limits_{k\to\infty} L(f_{n,k})$. Fatou's theorem is deduced exactly as in the case of the Lebesgue integral.

Suppose $f_n(x) \to f(x)$ and $|f_n(x)| \le \Phi(x)$ outside an L-zero set, where $f_n, \Phi \in \mathcal{L}$. Let $\varphi_n(x) := \inf_{k\ge n} f_k(x)$, $\varphi_n(x) := \sup_{k\ge n} f_k(x)$. Then, outside an L-zero set, one has $\varphi_n \le f_n \le \psi_n$, $\varphi_n \ge -\Phi$, $\psi_n \le \Phi$, $\{\varphi_n\}$ increases to f and $\{\psi_n\}$ decreases to f. Hence $f \in \mathcal{L}$ and $L(f) = \lim\limits_{n\to\infty} L(\varphi_n) = \lim\limits_{n\to\infty} L(\psi_n)$, which gives $L(f) = \lim\limits_{n\to\infty} L(f_n)$.

Suppose $\{f_n\} \subset \mathcal{L}$ is d_L-fundamental. Passing to a subsequence we may assume that $d_L(f_n, f_{n+1}) \le 2^{-n}$. As shown above, the series of $|f_n - f_{n-1}|$, where $f_0 := 0$, converges outside some L-zero set S to an element Φ of $\mathcal{L}$. Then the sums $f_n = \sum_{k=1}^n (f_k - f_{k-1})$ converge to a finite limit f outside S. Since $|f_n| \le \Phi$, we conclude that $\{f_n\}$ converges to f in $\widetilde{\mathcal{L}}$. □

Let us now consider the class $\mathcal{R}_L$ of all sets $E \subset \Omega$ such that there exists a sequence of functions $f_n \in \mathcal{F}$ convergent to I_E outside some L-zero set. Such sets will be called measurable (although no measure is introduced). Given $E \in \mathcal{R}_L$, we set $\nu(E) := L(I_E)$ if $I_E \in \mathcal{L}$ and $\nu(E) = +\infty$ otherwise. It is readily verified that $\mathcal{R}_L$ is a σ-ring and the function ν is a countably additive measure with values in $[0, +\infty]$. One can also consider ν on the δ-ring $\mathcal{R}_L^0$ of all sets on which ν is finite. However, in the general case (without Stone's condition), the integral with respect to the measure ν does not coincide with L. Say, in Example 7.8.8, the measure ν is identically zero. Indeed, in that example the L-zero sets are precisely the first category sets, since if $\alpha(f_n) \le 1/3$, then $f_n(t) \le 2/3$ outside a first category set. The class $\mathcal{L}$ differs from $\mathcal{F}$ only in that a function may now assume the values $+\infty$ and $-\infty$ on first category sets. If functions $f_n \in \mathcal{F}$ have a finite limit outside some first category set, then this limit coincides with the function $\alpha(1+t)$ outside a first category set, hence the indicator of a set can only appear if $\alpha = 0$, i.e., only the first category sets are measurable and they are L-zero.

We note that the theorems in this section do not involve topology. The topological concepts will be employed in the next two sections.

7.9. Measures as functionals

Every Baire measure μ on a topological space X defines a continuous linear functional on the Banach space $C_b(X)$ with the norm $\|f\| = \sup_X |f(x)|$ by the formula

$$f \mapsto \int_X f(x)\,\mu(dx). \tag{7.9.1}$$

In this and the next sections, we discuss what functionals can be obtained in such a way and what can be said about the properties of measures (such as regularity) in terms of the corresponding functionals. If a net of functions $\{f_\alpha\}$ decreases pointwise to f (i.e., $f_\alpha(x) \downarrow f(x)\ \forall\, x$), we write $f_\alpha \downarrow f$.

Although we do not discuss measures other than countably additive ones, for the purposes of this section it is useful to recall certain basic concepts related to additive set functions. It should be noted that in most of the literature, additive set functions are also called measures. However, following our earlier convention, we reserve the term "measure" only for countably additive set functions. Now let X be a topological space with the algebra $\mathfrak{A}(X)$ generated by all functionally closed sets. A set function $m\colon \mathfrak{A}(X) \to \mathbb{R}$ is called an additive regular set function if it is (i) additive, (ii) uniformly bounded, and (iii) for every $A \in \mathfrak{A}(X)$ and $\varepsilon > 0$, there exists a functionally closed set F such that $F \subset A$ and $|m(B)| < \varepsilon$ for all $B \subset A\backslash F$, $B \in \mathfrak{A}(X)$. It is verified directly (Exercise 7.14.88) that such a function m can be written as the difference of two nonnegative additive regular set functions m^+ and m^-, where

$$m^+(A) = \sup\{m(B)\colon\ B \in \mathfrak{A}(X), B \subset A\},$$

$$m^-(A) = -\inf\{m(B)\colon\ B \in \mathfrak{A}(X), B \subset A\}.$$

Set $\|m\| := m^+(X) + m^-(X)$.

In analogy with the Riemann integration, one can define the integral of a bounded continuous function f on X with respect to an additive regular set function m (see §4.7(ix)). The role of additive set functions can be seen from the following fundamental result due to A.D. Alexandroff [**30**].

7.9.1. Theorem. *If m is an additive regular set function on $\mathfrak{A}(X)$, then*

$$f \mapsto \int_X f(x)\,m(dx)$$

is a bounded linear functional on $C_b(X)$ whose norm equals $\|m\|$. Conversely, for any bounded linear functional L on $C_b(X)$, there exists an additive regular set function m on $\mathfrak{A}(X)$ with $\|m\| = \|L\|$ such that

$$L(f) = \int_X f(x)\,m(dx)$$

for all $f \in C_b(X)$. In addition, m is nonnegative precisely when so is the functional L.

PROOF. The direct claim is obvious. Let us prove the converse. According to what has been said above, we can assume that L is a nonnegative functional on $C_b(X)$. Let be $\mathcal{Z}$ the class of all functionally closed sets and

$$m(Z) = \inf\bigl\{L(f)\colon\ f \in C_b(X),\ I_Z \le f \le 1\bigr\}, \quad Z \in \mathcal{Z}.$$

We show that m_* is the required set function. It is clear that $m(Z) = m_*(Z)$ for any $Z \in \mathcal{Z}$, since the class $\mathcal{Z}$ admits finite unions. Let $Z_1, Z_2 \in \mathcal{Z}$ and $Z_1 \subset Z_2$. We show that

$$m(Z_2) - m(Z_1) = m_*(Z_2 \backslash Z_1).$$

Note that $m(Z_2) - m(Z_1) \ge m_*(Z_2\backslash Z_1)$ because $Z_1 \cup Z \in \mathcal{Z}$ if $Z \in \mathcal{Z}$ and $Z \subset Z_2\backslash Z_1$. Let $\varepsilon > 0$, $f \in C_b(X)$ and $f \ge I_{Z_1}$. Let $Y = \{x\colon\ f(x) \le 1 - \varepsilon\}$. Then $Y \cap Z_1 = \varnothing$. We fix a function $g \in C_b(X)$ with $g \ge I_{Z_2 \cap Y}$. For all $x \in Z_2$ we have $f(x) + g(x) > 1 - \varepsilon$, since if $x \in Y$, then $g(x) \ge 1$, and if $x \notin Y$, then $f(x) > 1 - \varepsilon$. Since $f + g \ge 0$, we obtain $(1-\varepsilon)^{-1}(f + g) \ge I_{Z_2}$, whence $L(f) + L(g) \ge (1 - \varepsilon)m(Z_2)$. Taking the infimum in g, we obtain the inequality $L(f) + m(Z_2 \cap Y) \ge (1-\varepsilon)m(Z_2)$. By using that $Z_2 \cap Y \subset Z_2\backslash Z_1$, we arrive at the estimate $L(f) + m_*(Z_2\backslash Z_1) \ge (1 - \varepsilon)m(Z_2)$. Therefore,

$$m(Z_1) + m_*(Z_2\backslash Z_1) \ge (1 - \varepsilon)m(Z_2),$$

which yields $m(Z_1) + m_*(Z_2\backslash Z_1) \ge m(Z_2)$, since ε is arbitrary. Thus, we have $m(Z_2) - m(Z_1) = m_*(Z_2\backslash Z_1)$.

Now let $Z \in \mathcal{Z}$ and let E be an arbitrary set. Let us verify the equality $m_*(E) = m_*(E \cap Z) + m_*(E\backslash Z)$, which means the Carathéodory measurability of Z with respect to m_*. Since one always has $m_*(E) \ge m_*(E\cap Z) + m_*(E\backslash Z)$, we have to verify the reverse inequality. Let $Z_0 \subset E$, $Z_0 \in \mathcal{Z}$. By the above we have $m(Z_0) = m(Z_0 \cap Z) + m_*(Z_0\backslash(Z_0 \cap Z))$. The right-hand side does not exceed $m_*(E\cap Z) + m_*(E\backslash Z)$, which yields the required inequality. According to Theorem 1.11.4, the class $\mathfrak{M}_{m_*}$ is an algebra, contains $\mathcal{Z}$, and the function m_* is additive on $\mathfrak{M}_{m_*}$. Hence the restriction of m_* to $\mathfrak{A}(X)$ is the required function. □

It is clear that in the general case the set function m may not be countably additive. In this and the next sections we clarify what functionals correspond to countably additive, Radon, and τ-additive measures. Let us introduce the following classes of functionals.

7.9.2. Definition. *Let $L \in C_b(X)^*$.*

(i) *The functional L is called σ-smooth if for every sequence $\{f_n\} \subset C_b(X)$ with $f_n \downarrow 0$, one has $L(f_n) \to 0$.*

(ii) *The functional L is called τ-smooth if for every net $\{f_\alpha\} \subset C_b(X)$ with $f_\alpha \downarrow 0$, one has $L(f_\alpha) \to 0$.*

(iii) *The functional L is called tight if for every net $\{f_\alpha\} \subset C_b(X)$ such that $\|f_\alpha\| \le 1$ and $f_\alpha \to 0$ uniformly on compact subsets of X, one has $L(f_\alpha) \to 0$.*

Let $\mathcal{M}_\sigma(X)$, $\mathcal{M}_\tau(X)$, $\mathcal{M}_t(X)$ denote the spaces of σ-smooth, τ-smooth, and tight functionals, respectively.

7.9.3. Theorem. *The following properties are equivalent*:
(i) $L \in \mathcal{M}_\sigma(X)$; (ii) $L^+, L^- \in \mathcal{M}_\sigma(X)$; (iii) $|L| \in \mathcal{M}_\sigma(X)$.

PROOF. Clearly, (ii) yields (i) and (iii), and (iii) yields (i). We show that (i) implies (ii). Let us verify that $L^+ \in \mathcal{M}_\sigma(X)$. If this is not true, then there is a sequence of functions $f_n \in C_b(X)$ decreasing to zero such that $L^+(f_n) > c > 0$. By the definition of L^+ one can find $g_1 \in C_b(X)$ with $0 \le g_1 \le f_1$ and $L(g_1) > c/2$. We observe that the functions $\max(f_n, g_1)$ are decreasing to g_1. Hence $L(\max(f_n, g_1)) \to L(g_1)$ and there exists n_1 with $L(\max(f_{n_1}, g_1)) > c/2$. Set $h_1 := \max(f_{n_1}, g_1)$. Then $0 \le f_{n_1} \le h_1 \le f_1$ and $L(h_1) > c/2$. Repeating the same reasoning we find $n_2 \in \mathbb{N}$ and $h_2 \in C_b(X)$ with $0 \le f_{n_2} \le h_2 \le f_{n_1}$ and $L(h_2) > c/2$. By induction, we obtain indices n_k and functions $h_k \in C_b(X)$ with the following properties: $n_{k+1} > n_k$, $f_{n_{k+1}} \le h_{k+1} \le f_{n_k}$, and $L(h_k) > c/2$. Then $\{h_k\}$ is decreasing to zero, which leads to a contradiction. The case of L^- is similar. □

7.9.4. Theorem. *The following properties are equivalent*:
(i) $L \in \mathcal{M}_\tau(X)$; (ii) $L^+, L^- \in \mathcal{M}_\tau(X)$; (iii) $|L| \in \mathcal{M}_\tau(X)$.

PROOF. As in Theorem 7.9.3, the main step is a verification of the inclusion $L^+ \in \mathcal{M}_\tau(X)$ for any $L \in \mathcal{M}_\tau(X)$. Suppose that there exists a net of functions $f_\alpha \in C_b(X)$ decreasing to zero such that $L^+(f_\alpha) > c > 0$. Without loss of generality we can assume that $|f_\alpha| \le 1$. The set T of all pairs (α, β) with $\beta > \alpha$ will be equipped with the following partial order: $(\alpha_1, \beta_1) \ge (\alpha_2, \beta_2)$ if either $\alpha_1 \ge \beta_2$ or $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$. If $(\alpha_3, \beta_3) \ge (\alpha_2, \beta_2)$ and $(\alpha_2, \beta_2) \ge (\alpha_1, \beta_1)$, where the three pairs are distinct, then $\alpha_3 \ge \beta_2$, $\beta_2 > \alpha_2$ and $\alpha_2 \ge \beta_1$, hence $\alpha_3 > \beta_1$, i.e., $(\alpha_3, \beta_3) \ge (\alpha_1, \beta_1)$. As in the case of sequences, for every α we find $g_\alpha \in C_b(X)$ with $0 \le g_\alpha \le f_\alpha$ and $L(g_\alpha) > c/2$. Taking T as a new index set, we observe that the net $\varphi_{\alpha,\beta} := \max(g_\alpha, f_\beta)$, $(\alpha, \beta) \in T$, is decreasing to zero. Indeed, if $(\alpha, \beta) \ge (\alpha_1, \beta_1)$ and $\alpha \ne \alpha_1$, then $\alpha \ge \beta_1$ and $\beta > \alpha \ge \beta_1$, so $g_\alpha \le f_\alpha \le f_{\beta_1}$ and $f_\beta \le f_{\beta_1}$. Hence $L(\varphi_{\alpha,\beta}) \to 0$. Let us take an index (α_0, β_0) such that $|L(\varphi_{\alpha,\beta})| < c/2$ if $(\alpha, \beta) \ge (\alpha_0, \beta_0)$. Then for all $\beta > \beta_0$ we obtain $|L(\varphi_{\beta_0,\beta})| < c/2$. Note that the net $\varphi_{\beta_0,\beta}$ is decreasing to g_{β_0}. By hypothesis, we have $L(\varphi_{\beta_0,\beta}) \to L(g_{\beta_0}) > c/2$. Then for some $\beta > \beta_0$ we have $|L(\varphi_{\beta_0,\beta})| > c/2$, which is a contradiction. □

7.9.5. Theorem. *The following properties are equivalent*:
(i) $L \in \mathcal{M}_t(X)$; (ii) $L^+, L^- \in \mathcal{M}_t(X)$; (iii) $|L| \in \mathcal{M}_t(X)$.

PROOF. As in the two previous theorems, everything reduces to the proof of the inclusion $L^+ \in \mathcal{M}_t(X)$ for $L \in \mathcal{M}_t(X)$. Suppose we are given a net of functions $f_\alpha \in C_b(X)$ that converges to zero uniformly on compact sets and $|f_\alpha| \le 1$. It is clear from the definition of L^+ that there exists $g_\alpha \in C_b(X)$ such that $0 \le g_\alpha \le f_\alpha$ and $0 \le L^+(|f_\alpha|) \le 2L(g_\alpha)$. Then the net $\{g_\alpha\}$ also converges to zero uniformly on compact sets and $|g_\alpha| \le 1$. Hence we obtain $L(g_\alpha) \to 0$, whence the assertion follows. □

7.10. The regularity of measures in terms of functionals

Now we show that the functionals in the classes mentioned in the last three theorems correspond one-to-one to Baire, τ-additive, and Radon measures.

7.10.1. Theorem. *Let X be a topological space. The formula*

$$L(f) = \int_X f(x)\,\mu(dx) \tag{7.10.1}$$

establishes a one-to-one correspondence between Baire measures μ on X and continuous linear functionals L on $C_b(X)$ with the following property:

$$\lim_{n\to\infty} L(f_n) = 0$$

for every sequence $\{f_n\}$ pointwise decreasing to zero.

PROOF. Any measure $\mu \in \mathcal{B}a(X)$ defines a continuous linear functional on the space $C_b(X)$. The converse follows by Theorem 7.8.1 and Corollary 7.8.4. $\square$

7.10.2. Remark. It is clear that every nonnegative linear functional L on $C_b(X)$ (i.e., nonnegative on nonnegative functions) is automatically continuous, since it satisfies the estimate $|L(f)| \le L(1)\sup|f|$.

Certainly, not every continuous linear functional satisfies the condition of Theorem 7.10.1.

7.10.3. Example. Let $X = \mathbb{N}$ be equipped with the usual discrete topology. Set

$$LIM(f) = \lim_{n\to\infty} f(n)$$

on the space $C_0(\mathbb{N})$ of all functions f on $\mathbb{N}$ for which this limit exists and is finite. The functional LIM is continuous on the space $C_0(\mathbb{N})$ by the estimate $|LIM(f)| \le \sup|f|$. By the Hahn–Banach theorem LIM extends to a continuous linear functional on the space $C_b(\mathbb{N})$. It is clear that even on the subspace $C_0(\mathbb{N})$ the functional LIM cannot be represented as the integral with respect to a countably additive measure on the space $\mathbb{N}$.

Such a situation is impossible for compact spaces. The following result is called the *Riesz representation theorem.*

7.10.4. Theorem. *Let K be a compact space. Then, for every continuous linear functional L on the Banach space $C(K)$, there exists a unique Radon measure μ such that*

$$L(f) = \int_K f(x)\,\mu(dx), \quad \forall\, f \in C(K).$$

PROOF. By Dini's theorem, any sequence of continuous functions monotonically decreasing to zero on a compact set is uniformly convergent (see Engelking [**532**, 3.2.18]). Hence, in our case, every continuous linear functional satisfies the hypothesis of Theorem 7.10.1. It remains to observe that every Baire measure on a compact space extends uniquely to a Radon measure according to Theorem 7.3.2. □

The Riesz theorem yields at once a Radon extension of the product of Radon measures μ and ν on compact spaces X and Y: the integral with respect to $\mu\otimes\nu$ defines a continuous functional on $C(X\times Y)$ (we recall that for all compact spaces one has $\mathcal{B}a(X\times Y) = \mathcal{B}a(X)\otimes\mathcal{B}a(Y)$).

7.10.5. Corollary. *Let X be a compact space. Then formula* (7.10.1) *establishes a one-to-one correspondence between nonnegative linear functionals on the space $C(X)$ and nonnegative Radon measures on X.*

The following two theorems characterize functionals generated by Radon and τ-additive measures.

7.10.6. Theorem. *Let X be a completely regular space. Formula* (7.10.1) *establishes a one-to-one correspondence between Radon measures μ on X and continuous linear functionals L on $C_b(X)$ satisfying the following condition: for every $\varepsilon > 0$, there exists a compact set K_ε such that if $f \in C_b(X)$ and $f\,|_{K_\varepsilon}= 0$, then*

$$|L(f)| \le \varepsilon \sup |f|.$$

PROOF. If μ is a Radon measure, then this condition is satisfied. Let us prove the converse. Let $\{f_n\}$ be a sequence of bounded continuous functions monotonically decreasing to zero. Let us verify the hypotheses of Theorem 7.10.1. We may assume that $|f_n| \le 1$ and $\|L\| \le 1$. Let us fix $\varepsilon \in (0, 1)$ and take the corresponding compact set K_ε. By Dini's theorem, there exists a number n_0 such that $\sup_{K_\varepsilon} |f_n| < \varepsilon$ for all $n > n_0$. For every $n \ge n_0$, we find a function $g_n \in C_b(X)$ such that $g_n = f_n$ on K_ε and $|g_n| \le \varepsilon$. Then $|L(g_n)| \le \varepsilon$. By hypothesis, $|L(f_n - g_n)| \le 2\varepsilon$, since $f_n - g_n = 0$ on K_ε and $|f_n - g_n| \le 2$. Hence $|L(f_n)| \le 3\varepsilon$. Therefore, L is generated by a Baire measure μ.

Let us verify that μ is tight. We observe that it suffices to consider positive functionals L (this corresponds to nonnegative measures μ), since the functional $|L|$ generated by the measure $|\mu|$ satisfies the condition mentioned in the formulation of the theorem. Indeed, if a compact set K_ε is taken for ε and L, and a function $f \in C_b(X)$ vanishes outside K_ε, then by Theorem 7.8.3 we have $|L(f)| \le |L|(|f|) \le \varepsilon \sup |f|$, since $|f| = 0$ on K_ε. Thus, we may assume that μ is nonnegative. In order to show that μ is tight, suppose that a Baire set B does not meet K_ε. By the regularity of μ we can find a functionally closed set $Z \subset B$ such that $\mu(B\backslash Z) < \varepsilon$, and then a neighborhood U of K_ε disjoint with Z. By the complete regularity of X, there exists a continuous function $f\colon X \to [0, 1]$ such that $f = 0$ on K_ε and $f = 1$ outside U, in particular, $f = 1$ on Z. Then

$$\mu(Z) \le \int_X f\, d\mu < \varepsilon,$$

whence we obtain $\mu(B) < 2\varepsilon$. □

7.10.7. Theorem. *Let X be a completely regular space. Formula* (7.10.1) *establishes a one-to-one correspondence between τ-additive measures μ on X and continuous linear functionals L on $C_b(X)$ satisfying the following condition*: *if a net $\{f_\alpha\}$ of bounded continuous functions is decreasing to zero pointwise, then $L(f_\alpha) \to 0$.*

PROOF. According to Corollary 7.2.7, the functionals defined by τ-additive measures satisfy the above condition. In view of Theorem 7.9.4, in the proof of the converse assertion we can assume that the functional L is nonnegative. It remains to apply Theorem 7.8.6. □

Thus, the classes of functionals $\mathcal{M}_\sigma(X)$, $\mathcal{M}_\tau(X)$, and $\mathcal{M}_t(X)$ can be identified with the respective classes of measures.

If an additive set function $m \geq 0$ on $\mathcal{B}a(X)$ is such that there is no nonzero countably additive measure $m_1 \geq 0$ with $m_1 \leq m$, then m is called purely finitely additive. If m is countably additive, but there is no nonzero τ-additive $m_1 \geq 0$ with $m_1 \leq m$, then m is called purely countably additive. Finally, if m is τ-additive, but there is no nonzero tight measure $m_1 \geq 0$ with $m_1 \leq m$, then m is called purely τ-additive. Let us mention the following decomposition theorem obtained in Knowles [**1015**] (the existence of the compact regular and τ-additive components was proved by Alexandroff [**30**], who raised the question about the purely countably additive component).

7.10.8. Theorem. *Every nonnegative additive set function m on the Baire σ-algebra of a completely regular space X has a unique representation*

$$m = m_c + m_\tau + m_\sigma + m_a,$$

where $m_c \geq 0$ is a tight measure, $m_\tau \geq 0$ is a purely τ-additive measure, $m_\sigma \geq 0$ is a purely countably additive measure, and $m_a \geq 0$ is a purely finitely additive set function on $\mathcal{B}a(X)$. An analogous result is true for signed additive set functions of bounded variation on $\mathcal{B}a(X)$.

This result, excluding, possibly, the presence of the m_τ-component, holds for general Borel measures as well.

In connection with the Riesz representation theorem the following useful condition of weak compactness in the space $C(K)$ should be mentioned (see Dunford, Schwartz [**503**, IV.6.14] for a proof and related references).

7.10.9. Theorem. *Let K be a compact space and let $F \subset C(K)$. Then the following conditions are equivalent*:

(i) *the closure of F in the weak topology is compact,*

(ii) *every sequence in F has a weakly convergent subsequence,*

(iii) *F is norm bounded and is contained in a set in $C(K)$ that is compact in the topology of pointwise convergence.*

7.11. Measures on locally compact spaces

Consideration of locally compact spaces brings some specific features in the theory of integration. We recall that a Hausdorff topological space X is called locally compact if every point in X possesses an open neighborhood with compact closure. A locally compact space is completely regular (see Engelking [**532**, Theorem 3.3.1]). By Lemma 6.1.5, for any compact set K in a locally compact space X and any open set $U \supset K$, one can find a continuous function $f\colon X \to [0,1]$ such that $f|_K = 1$ and f vanishes outside some compact set contained in U. The set of all continuous functions on X with compact support is denoted by $C_0(X)$. On typical non-locally compact spaces, for example, infinite-dimensional normed spaces, the class $C_0(X)$ consists only of the zero function. In the locally compact case, this class separates points, which turns out to be of great importance in the theory of integration. Apart from compact spaces, standard locally compact spaces encountered in applications are finite-dimensional manifolds and locally compact groups. Denote by $\mathcal{K}(X)$ the class of all compact sets in X.

7.11.1. Theorem. *Suppose that X is a locally compact space and that $\tau\colon \mathcal{K}(X) \to [0,+\infty)$ is a set function such that for all $K_1, K_2 \in \mathcal{K}(X)$, one has*

$$\tau(K_1 \cup K_2) \le \tau(K_1) + \tau(K_2), \quad \tau(K_1 \cup K_2) = \tau(K_1) + \tau(K_2) \quad \text{if } K_1 \cap K_2 = \varnothing,$$

and $\tau(K_1) \le \tau(K_2)$ if $K_1 \subset K_2$. Then, there exists a unique measure μ on $\mathcal{B}(X)$ with values in $[0,+\infty]$ that is outer regular in the sense that the measure of every Borel set is the infimum of measures of the enclosing open sets, and the value on every open set U is the supremum of measures of compact subsets of U, and one has

$$\mu(U) = \sup\bigl\{\tau(K)\colon\ K \subset U,\, K \in \mathcal{K}(X)\bigr\}. \tag{7.11.1}$$

In addition,

$$\mu(K^o) \le \tau(K) \le \mu(K), \quad \forall\, K \in \mathcal{K}(X), \tag{7.11.2}$$

where K^o is the interior of K.

If $\tau(K) = \inf\bigl\{\tau(S)\colon\ S \in \mathcal{K}(X),\, K \subset S^o\bigr\}$ for all sets $K \in \mathcal{K}(X)$, then μ coincides with τ on $\mathcal{K}(X)$.

Finally, the restrictions of μ to all Borel sets of finite measure are Radon measures, and the formula

$$\mu'(B) = \sup\bigl\{\mu(K)\colon\ K \subset B,\, K \in \mathcal{K}(X)\bigr\}, \quad B \in \mathcal{B}(X), \tag{7.11.3}$$

defines the Borel measure μ' with values in $[0,+\infty]$ that coincides with μ on compact sets, in particular, every function in $C_0(X)$ has equal integrals with respect to μ and μ' (the completion of μ' is an infinite Radon measure in the sense of §7.14(xviii)*).*

Proof. For every open set U, we define $\mu(U)$ by formula (7.11.1). We obtain a monotone and additive function μ with values in $[0,+\infty]$ on the class $\mathcal{U}$ of all open sets. Indeed, if $U, V \in \mathcal{U}$ are disjoint, then for every compact set $K \subset U \cup V$, the sets $U \cap K$ and $V \cap K$ are compact. This yields $\mu(U \cup V) \le \mu(U) + \mu(V)$ by the additivity of τ. The reverse inequality is easily verified as well. Further, one has

$$\mu(U) = \sup\{\mu(V)\colon\ V \in \mathcal{U},\, \overline{V} \subset U,\, \overline{V} \in \mathcal{K}(X)\}.$$

This follows from the fact that for every compact set $K \subset U$, one can find a set $V \in \mathcal{U}$ with the compact closure $\overline{V}$ such that $K \subset V \subset \overline{V} \subset U$. Finally, the function μ is countably subadditive. Indeed, if $U = \bigcup_{i=1}^{\infty} U_i$, $U_i \in \mathcal{U}$, then, given $\varepsilon > 0$, there exists a set $V \in \mathcal{U}$ with compact closure such that $\mu(V) > \mu(U) - \varepsilon$ and $V \subset \overline{V} \subset U$. Then $\overline{V} \subset \bigcup_{i=1}^{n} U_i$ for some n, whence $\mu(V) \le \mu\Bigl(\bigcup_{i=1}^{n} U_i\Bigr)$. Hence it suffices to establish the finite subadditivity of μ on $\mathcal{U}$. Now we can consider only two sets U_1 and U_2. For every compact set $K \subset U_1 \cup U_2$, according to Exercise 7.14.71, there are continuous nonnegative functions f_1 and f_2 with the compact supports $K_1 \subset U_1$ and $K_2 \subset U_2$, respectively, such that $f_1 + f_2 = 1$ on K. The sets $Q_i = \{f_i \ge 1/2\}$ with $i = 1, 2$ are compact in U_i and $K = (K \cap Q_1) \cup (K \cap Q_2)$. Hence

$$\tau(K) \le \tau(K \cap Q_1) + \tau(K \cap Q_2) \le \mu(U_1) + \mu(U_2),$$

whence $\mu(U) \le \mu(U_1) + \mu(U_2)$. Then $\mu = \mu^*$ on $\mathcal{U}$ (Exercise 1.12.125). It is readily seen that $\mu(A) = \mu(A \cap B) + \mu^*(A \backslash B)$ if $A, B \in \mathcal{U}$, hence $\mathcal{U} \subset \mathfrak{M}_{\mu^*}$ (Exercise 1.12.126). The restriction of μ^* to $\mathfrak{M}_{\mu^*}$ will be denoted by μ as well. Thus, we obtain an outer regular measure. For every $K \in \mathcal{K}(X)$, we have $\mu(K^o) \le \tau(K)$ by (7.11.1). Hence $\mu(K^o) \le \tau(K) \le \mu(K)$. The uniqueness of μ follows by construction.

If for all $K \in \mathcal{K}(X)$ the condition $\tau(K) = \inf\{\tau(S)\colon\ S \in \mathcal{K}(X),\, K \subset S^o\}$ is fulfilled, then

$$\begin{aligned}\mu(K) &= \inf\{\mu(U)\colon U \in \mathcal{U}, K \subset U\} \le \inf\{\mu(S^o)\colon S \in \mathcal{K}(X), K \subset S^o\}\\ &\le \inf\{\tau(S)\colon S \in \mathcal{K}(X), K \subset S^o\} = \tau(K).\end{aligned}$$

Note that under the aforementioned condition we could also apply Theorem 1.12.33, which would give us the measure μ'.

If $B \in \mathcal{B}(X)$ and $\mu(B) < \infty$, then the restriction of μ to B is a Radon measure. Indeed, the outer regularity of μ yields that the restrictions of μ to compact sets are Radon. Now, given $\varepsilon > 0$, we take an open set $U \supset B$ with $\mu(U \backslash B) < \varepsilon/4$, next we find a compact set $K_1 \subset U$ with $\mu(U \backslash K_1) < \varepsilon/4$. Since μ is Radon on K_1, there exists a compact set $K_2 \subset K_1 \cap B$ with $\mu\bigl((K_1 \cap B) \backslash K_2\bigr) < \varepsilon/3$. Hence $K_2 \subset B$ and $\mu(B \backslash K_2) < \varepsilon$. Finally, for any Borel set B with compact closure, we have $\mu'(B) = \mu(B)$, since by the above this is true for all sets of finite measure. The countable additivity of μ' follows by the additivity verified as follows. If A and B are disjoint and have finite measures, then μ' coincides with μ on A, B and $A \cup B$, and if A or B has the infinite measure, then $A \cup B$ also does. □

7.11.2. Remark. (i) The measure μ constructed in the theorem may not be inner compact regular, and the measure μ' may not be outer regular, i.e., one cannot always combine both regularity properties (this happens for some Haar measures, see also Example 7.14.65 and Exercise 7.14.160). Certainly, for finite measures this problem does not arise. The property of inner compact regularity is more useful than the outer regularity, and in our discussion of Haar measures in Chapter 9 we shall employ the measure μ'.

(ii) The assertions of the theorem remain valid if $\mathcal{K}(X)$ is a certain class of compact sets in X that is closed with respect to finite unions and intersections and contains all compact G_δ-sets. This is easily seen from the proof.

7.11.3. Theorem. *Let X be a locally compact space and let L be a linear function on $C_0(X)$ such that $L(f) \geq 0$ if $f \geq 0$. Then, there exists a Borel measure μ on X with values in $[0, +\infty]$ such that*

$$L(f) = \int_X f\,d\mu, \quad \forall\, f \in C_0(X). \tag{7.11.4}$$

In addition, one can choose μ in such a way that it will be Radon on all sets of finite measure (and even inner compact regular on $\mathcal{B}(X)$, and there is only one measure with this property).

Proof. Here Theorem 7.8.7 is applicable, since if $f_n \in C_0(X)$ and $f \downarrow 0$, then convergence is uniform. This theorem gives a measure on $\sigma\bigl(C_0(X)\bigr)$ that can be extended to $\mathcal{B}(X)$ by the previous theorem and remark. Let us give an alternative justification. For every open set V with the compact closure $\overline{V}$, let $C_0(V)$ be the set of continuous functions on X with compact support in V. Since V is open, the class $C_0(V)$ can be identified with the set of all continuous functions on $\overline{V}$ with compact support in V, extended to X by zero outside the support. Thus, $C_0(V)$ can be regarded as a linear subspace in the space $C(\overline{V})$. The functional L on $C_0(V)$ satisfies the condition $L(f) \leq M \max_{\overline{V}} |f|$ with some $M \geq 0$. Indeed, let us find $\theta \in C_0(X)$ with $\theta \geq 0$ and $\theta|_{\overline{V}} = 1$. Let $M = L(\theta)$. Then $L(f) \leq L(\theta)$ if $f \in C_0(V)$ and $|f| \leq 1$. By the Hahn–Banach theorem L extends to a continuous linear functional on $C(\overline{V})$, which by the Riesz theorem gives a Radon measure ν on $\overline{V}$ such that

$$L(f) = \int_{\overline{V}} f\,d\nu, \quad \forall\, f \in C_0(V).$$

Let $\mu_V = \nu|_V$. Then

$$L(f) = \int_V f\,d\mu_V, \quad \forall\, f \in C_0(V). \tag{7.11.5}$$

It is clear that $\mu_V \geq 0$ and that if V, W are two open sets with compact closure, then $\mu_V|_{V\cap W} = \mu_W|_{V\cap W}$. This follows by (7.11.5) due to the fact that every Radon measure τ on $V \cap W$ is uniquely determined by the values on compact sets $S \subset V \cap W$ and if $\tau \geq 0$, then $\tau(S)$ is the infimum of the integrals with respect to τ of functions $f \in C_0(V \cap W)$ with $0 \leq f \leq 1$ and $f|_S = 1$. Thus, the required measure μ is constructed on the δ-ring of Borel sets whose closures are compact. Given such a set B, we find its neighborhood V with compact closure and set $\mu(B) := \mu_V(B)$. It follows by the above that $\mu(B)$ is well-defined. It remains to extend μ to all Borel sets. This can be done by formula (7.11.3). Certainly, one can also refer to the previous theorem and remark. The uniqueness assertion is clear from the proof. □

If L is a nonnegative linear functional on the space $C(X)$, then one might hope to find a Borel measure μ such that (7.11.4) is true for all $f \in C(X)$. However, this is not always possible for general locally compact spaces. If X is locally compact and σ-compact, then such a measure exists (details are found in Exercise 7.14.161).

7.12. Measures on linear spaces

In this section, some of the general results obtained above are applied to measures on linear spaces. If X is a linear space and G is some linear space of linear functions on X, then sets of the form

$$C(f_1,\ldots,f_n,B) = \bigl\{x \in X\colon\ \bigl(f_1(x),\ldots,f_n(x)\bigr) \in B\bigr\},$$

where $f_1,\ldots,f_n \in G$ and $B \in \mathcal{B}(\mathbb{R}^n)$, are called G-cylindrical. The family of all G-cylindrical sets is denoted by $Cyl(X,G)$. It is clear that the smallest σ-algebra containing $Cyl(X,G)$ is $\sigma(G)$, i.e., the σ-algebra generated by G.

Any cylindrical set has the following representation. Suppose that the functionals are f_i linearly independent. Then, one can find linearly independent vectors $e_1,\ldots,e_n$ with $f_i(e_j) = 0$ if $i \neq j$ and $f_i(e_i) = 1$. The isomorphism $(x_1,\ldots,x_n) \mapsto \sum_{i=1}^n x_i e_i$ takes the set B to a set B' in X. Then the set $C(f_1,\ldots,f_n,B)$ is the cylinder $B' + L$, where L is the intersection of the kernels of the functionals f_i, i.e., $L = \bigcap_{i=1}^n f_i^{-1}(0)$. Geometrically, one can think of $B' + L$ as a cylinder with a base B'.

The most interesting case in applications is where X is a locally convex space, X^* is the space of all continuous linear functions on X, and $G \subset X^*$ is a linear subspace. If $G = X^*$, then the sets in $Cyl(X,X^*)$ are called cylindrical. Exercise 7.14.132 proposes to verify that the class $Cyl(X,G)$ is the algebra generated by G. The base of the topology $\sigma(X,G)$ (see §4.7(ii)) consists of cylinders. Applying the general results from §7.1 to measures on $\sigma(X^*)$, where X is a locally convex space with the dual X^*, we see that every measure μ on $\sigma(X^*)$ is regular: for every $A \in \sigma(X^*)$ and $\varepsilon > 0$, there exists a closed set $F \in \sigma(X^*)$ with $F \subset A$ and $|\mu|(A\backslash F) < \varepsilon$. We recall that by Corollary 7.3.6 every tight nonnegative regular additive set function on $Cyl(X,X^*)$ has a unique extension to a nonnegative Radon measure on X. Hence every Radon measure on a locally convex space is uniquely determined by its values on $Cyl(X,X^*)$. However, we shall prove this useful fact directly in a different formulation.

7.12.1. Proposition. *Let μ be a Radon measure on a locally convex space X. Then, for every μ-measurable set A, there exists a set $B \in \sigma(X^*)$ such that $|\mu|(A \triangle B) = 0$. Moreover, if $G \subset X^*$ is an arbitrary linear subspace separating the points in X, then such a set B can be chosen in $\sigma(G)$.*

PROOF. Let us verify that for every $\varepsilon > 0$, there exists a set C in $Cyl(X,G)$ such that $|\mu|(A \triangle C) < \varepsilon$. Since μ is Radon, it suffices to do this for compact sets A. We find an open set $U \supset A$ with $|\mu|(U \backslash A) < \varepsilon/4$ and a compact set S with $|\mu|(X \backslash S) < \varepsilon/4$. Now we use that on the compact set S, the original topology of X coincides with the topology $\sigma(X,G)$ (in particular, if $G = X^*$, then with the weak topology). By the compactness of $A \cap S$ one can find finitely many open G-cylindrical sets $C_1, \ldots, C_k$ such that $A \cap S \subset (C_1 \cup \cdots \cup C_k) \cap S \subset U \cap S$. Let $C = C_1 \cup \cdots \cup C_k$. Then $C \in Cyl(X,G)$ and

$$|\mu|(A \triangle C) \leq |\mu|\bigl((A \cap S) \triangle (C \cap S)\bigr) + \varepsilon/4 \leq |\mu|\bigl((U \cap S) \backslash (A \cap S)\bigr) + \varepsilon/4 < \varepsilon,$$

as required. $\square$

Let us explain why this proposition is not identical to Corollary 7.3.6. The point is that the Lebesgue completion of $\sigma(X^*)$ may not include $\mathcal{B}(X)$. For example, we have already seen that if μ is Dirac's measure at the point 0 on the product of the continuum of real lines, then this point does not belong to $\sigma\bigl((\mathbb{R}^c)^*\bigr)_\mu$. Hence the assertion of the proposition cannot be obtained by using only the outer measure generated by the values of μ on $Cyl(X,X^*)$ or on $\sigma(X^*)$. It is important that in this proposition the measure is already defined on $\mathcal{B}(X)$.

7.12.2. Corollary. *Let μ be a Radon measure on a locally convex space X. Then the class of all bounded cylindrical functions on X is dense in $L^p(\mu)$ for any $p > 0$. In the case of complex-valued functions, the same is true for the linear space T generated by the functions $\exp(if)$, $f \in X^*$. Moreover, this assertion is true if we replace X^* with any linear subspace $G \subset X^*$ separating the points in X.*

Let μ be a set function on an algebra $Cyl(X,G)$, where X is a locally convex space and $G \subset X^*$. For every continuous linear operator $P\colon X \to \mathbb{R}^n$ of the form $Px = \bigl(f_1(x), \ldots, f_n(x)\bigr)$, where $f_i \in G$, one has the set function

$$\mu \circ P^{-1}(B) := \mu\bigl(P^{-1}(B)\bigr) = \mu\bigl(C(f_1, \ldots, f_n, B)\bigr), \quad B \in \mathcal{B}(\mathbb{R}^n),$$

called the projection of μ generated by P.

7.12.3. Definition. *An additive real function μ on $Cyl(X,G)$ such that all finite-dimensional projections $\mu \circ P^{-1}$ are bounded and countably additive is called a G-cylindrical quasi-measure. If $G = X^*$, then such a function is called a cylindrical quasi-measure. A probability quasi-measure is a nonnegative quasi-measure μ with $\mu(X) = 1$.*

It is clear that any countably additive measure on $Cyl(X,G)$ is a G-cylindrical quasi-measure, but the converse is false. Let us consider the following simple example. Let $X = l^2$, $G = X^* = l^2$, and let γ be the quasi-measure defined as follows: if $C = P^{-1}(B)$, where P is the orthogonal projection to a linear subspace $L \subset X$ of dimension n and B is a Borel set in L, then

$\gamma(C) = \gamma_n(B)$, where γ_n is the standard Gaussian measure on L (with density $(2\pi)^{-n/2}e^{-|x|^2/2}$ with respect to Lebesgue measure on L generated by the inner product in X). It is clear that every cylinder can be written in such a form. If the measure γ were countably additive on the algebra of cylinders, then it would have a unique extension to a countably additive measure on the σ-algebra generated by all cylinders (which coincides with the Borel σ-algebra of X). However, direct computations show that in this case every ball has measure zero. Indeed, if $U_{n,R}$ is the ball of radius R centered at the origin in $\mathbb{R}^n$, then $\lim\limits_{n\to\infty} \gamma_n(U_{n,R}) = 0$ for all R. This is a contradiction. Corollary 7.3.6 states that a sufficient (in the case of a complete separable metric space also necessary) condition of the countable additivity of a nonnegative cylindrical quasi-measure is its tightness. In the next section we shall give sufficient conditions in terms of characteristic functionals.

In applications, one usually deals with measures on separable Banach spaces and also on some special nonnormable spaces such as the spaces $\mathcal{S}'$ and $\mathcal{D}'$ of distributions. Measures on Fréchet spaces (i.e., complete metrizable locally convex spaces) are concentrated on separable Banach spaces. The proof of this fact employs the following construction, which is useful in diverse problems of infinite-dimensional analysis. Let X be a locally convex space and let K be a convex and symmetric compact set (the symmetry means that $-x \in K$ if $x \in K$). Denote by E_K the linear subspace in X generated by K, i.e., E_K is the union of the sets nK. It turns out that E_K can be made a Banach space if we declare K to be the unit ball. More precisely, E_K is complete with respect to the norm $p_K(x) = \inf\{\lambda > 0\colon\ x/\lambda \in K\}$, called the Minkowski functional of the set K. Moreover, in place of the compactness of K it suffices that K be a bounded convex symmetric and sequentially complete set (see Edwards [**518**, Lemma 6.5.2, p. 609].

7.12.4. Theorem. *Let μ be a Radon probability measure on a Fréchet space X. Then, there exists a linear subspace $E \subset X$ such that $\mu(E) = 1$ and E with some norm $\|\cdot\|_E$ is a separable reflexive Banach space whose closed balls are compact in X.*

Proof. The topology of X is generated by a metric ϱ. For every n, we take a compact set K_n with $\mu(X\backslash K_n) < 1/n$. Then $\mu\bigl(\bigcup_{n=1}^{\infty} K_n\bigr) = 1$. Let us pick a number $c_n > 0$ such that $c_n K_n$ belongs to the ball of radius $1/n$ centered at the origin. It is easily verified that the closure S of the set $\bigcup_{n=1}^{\infty} c_n K_n$ is compact. There is a convex symmetric compact set K_0 containing S (see Schaefer [**1661**, Corollary in p. 80, §4, Ch. II]). This set may not be what we want, since E_{K_0} may not be even separable (just look at the embedding of l^∞ to $\mathbb{R}^\infty$). But according to Edwards [**518**, Lemma 9.6.4, p. 922], one can take a larger convex symmetric compact set K_1 such that K_0 is compact as a subset of E_{K_1}. The closure E_0 of the linear span of K_0 in E_{K_1} is already a separable Banach space of full μ-measure. However, it may not be reflexive, although its closed unit ball is compact in X (since K_1 is the unit ball in E_{K_1}). The measure μ can now be restricted to E_0, since all Borel sets in E_0 are Borel in X (see Chapter 6). Repeating this procedure once again, we obtain a separable Banach space $E_2 \subset E_0$ of full μ-measure whose closed unit ball is compact in E_0. According to a well-known result in the theory of Banach spaces (see Diestel [**442**, p. 124]), there exists a reflexive Banach space E such that $E_2 \subset E \subset E_0$ and the unit ball from E is bounded in E_0. Note that E is automatically separable (Exercise 7.14.134), although one can simply deal with the closure of E_2 in E. The closed balls in E are compact in X. This follows from the fact that they are closed in X, being convex and weakly closed by their weak compactness in E (see [**1661**, Ch. IV]). □

The question arises as to which Banach spaces can be taken for E.

It is shown in Fonf, Johnson, Pisier, Preiss [**596**] that one cannot always take for E a space with a Schauder basis or with the approximation property. A Hilbert space E can be found even more rarely. Moreover, if in a Banach space X every Radon measure is concentrated on a continuously embedded Hilbert space, then X itself is linearly homeomorphic to a Hilbert space (see Mouchtari [**1337**] and Sato [**1651**]). This assertion does not extend to Fréchet spaces: for example, it is obvious that on $\mathbb{R}^\infty$ every Radon measure is concentrated on a continuously embedded Hilbert space $\bigl\{(x_n)\colon\ \sum_{n=1}^\infty c_n x_n^2 < \infty\bigr\}$, where the numbers $c_n > 0$ decrease to zero sufficiently fast. Indeed, the unit ball in the space E from the previous theorem is coordinate-wise bounded in $\mathbb{R}^\infty$ and hence is contained in some Hilbert space of the indicated type. The following interesting generalization of Theorem 7.12.4 is obtained in Matsak, Plichko [**1271**]: one can take for E a closed subspace in the l^2-sum of finite-dimensional Banach spaces. Herer [**819**] and Okazaki [**1397**] considered the so-called stochastic bases in a separable Fréchet space X with a Borel probability measure μ. A stochastic basis is a system of vectors $\varphi_n \in X$ with the following property: there exist $f_n \in X^*$ with $f_n(\varphi_k) = \delta_{nk}$ such that letting $P_n x := \sum_{i=1}^n f_i(x)\varphi_i$, one has $P_n x \to x$ μ-a.e. It is shown in [**1397**] that such a basis exists provided that all continuous seminorms are in $L^2(\mu)$, the elements of X^* have zero means, and there is a sequence $\{f_n\} \subset X^*$ whose elements are independent random variables with respect to μ such that their linear span is dense in X^* with the metric from $L^2(\mu)$. It is also shown in the same work that the existence of a stochastic basis yields a Banach space of full measure possessing a Schauder basis. Hence, by the above-mentioned result, stochastic bases do not always exist.

7.13. Characteristic functionals

This section is devoted to the conditions of countable additivity of additive set functions on certain algebras of subsets of a linear space. Our main tool is the concept of a characteristic functional introduced by A.N. Kolmogorov. However, we start our discussion with the following theorem of Bochner, giving the description of characteristic functionals of probability measures on $\mathbb{R}^n$. We already know that the characteristic functionals of probability measures are positive definite, continuous and equal to 1 at the origin. It turns out that these properties completely identify the characteristic functionals of probability measures.

7.13.1. Theorem. *A function $\varphi\colon\ \mathbb{R}^n \to \mathbb{C}$ coincides with the characteristic functional of a probability measure on $\mathbb{R}^n$ precisely when it is continuous, positive definite and $\varphi(0) = 1$. Hence the class of all characteristic functionals of nonnegative measures on $\mathbb{R}^n$ coincides with the class of all continuous positive definite functions.*

PROOF. The necessity of the indicated conditions has already been established. In the proof of sufficiency we suppose first that the function φ is integrable. It was shown in the proof of Theorem 3.10.20 that φ coincides with the characteristic functional of a probability measure possessing a density with respect to Lebesgue measure. In the general case, we consider the integrable functions $\varphi_k(x) = \varphi(x)\exp[-k^{-1}|x|^2/2]$, which are positive definite, since so are the functions $\exp[-k^{-1}|x|^2/2]$ that are the Fourier transforms of Gaussian densities. In addition, $\varphi_k(0) = 1$. Hence there exist probability measures μ_k with $\widetilde{\mu_k} = \varphi_k$. We show that for every $\delta > 0$, there exists $R > 0$ such that

$$\mu_k\bigl(x\colon\ |x| \geq R\bigr) < \delta, \quad \forall\, k \in \mathbb{N}. \tag{7.13.1}$$

Since $\varphi(x) = \lim\limits_{k\to\infty} \varphi_k(x)$, for the standard Gaussian measure γ_n on $\mathbb{R}^n$ and any $t > 0$, we have

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} [1 - \varphi_k(y/t)]\, \gamma_n(dy) = \int_{\mathbb{R}^n} [1 - \varphi(y/t)]\, \gamma_n(dy).$$

By (3.8.6) we obtain

$$\limsup_{k\to\infty} \mu_k\bigl(x\colon\ |x| \geq R\bigr) \leq 3 \int_{\mathbb{R}^n} [1 - \varphi(y/R)]\, \gamma_n(dy).$$

It remains to observe that as $R \to \infty$, the right-hand side tends to zero by the dominated convergence theorem and continuity of φ. It follows by (7.13.1) that for every bounded continuous function f on $\mathbb{R}^n$, the integrals of f against the measures μ_k converge. Indeed, such integrals have a limit for every smooth function f with bounded support, since by the Parseval equality one has

$$\int_{\mathbb{R}^n} f\, d\mu_k = \int_{\mathbb{R}^n} (2\pi)^{n/2} \widehat{f} \varphi_k\, dx,$$

where $\widehat{f} \in L^1(\mathbb{R}^n)$. This yields that such integrals converge for every continuous function f with bounded support, and then (7.13.1) implies the existence of a limit for every bounded continuous function. Moreover, (7.13.1) and Theorem 7.11.3 yield the existence of a probability measure μ the integral with respect to which of every bounded continuous function f equals the limit of the above integrals (this also follows by a general theorem on the sequential completeness in §8.7). It is easily verified that μ is the required measure. $\square$

We remark that by Theorem 3.10.20 and the Bochner theorem, every measurable positive definite function φ almost everywhere equals the characteristic functional of a nonnegative measure (however, even under the condition $\varphi(0) = 1$ it is not always true that this measure is probability, since the continuous modification of φ may not equal 1 at zero). As has already been noted, one cannot omit the measurability of φ.

We now proceed to infinite-dimensional analogs of the Bochner theorem.

7.13.2. Definition. *The characteristic functional* (*the Fourier transform*) *of a quasi-measure μ on $Cyl(X, G)$ is the function $\widetilde{\mu}\colon G \to \mathbb{C}$ defined by the equality*

$$\widetilde{\mu}(f) = \int_{\mathbb{R}^1} e^{it}\, \mu \circ f^{-1}(dt).$$

We remark that the function e^{it} is integrable with respect to the bounded measure $\mu \circ f^{-1}$ on the real line.

The most important case for applications is where X is a locally convex space and $G = X^*$ is its dual.

7.13.3. Definition. *Let G be a linear space. A function $\varphi\colon G \to \mathbb{C}$ is called positive definite if $\sum_{i,j=1}^{k} c_i \overline{c_j} \varphi(y_i - y_j) \geq 0$ for all $y_i \in G$, $c_i \in \mathbb{C}$, $i = 1, \ldots, k$, $k \in \mathbb{N}$.*

The Bochner theorem yields the following.

7.13.4. Proposition. *A function $\varphi\colon G \to \mathbb{C}$ is the characteristic functional of a probability quasi-measure precisely when it is positive definite, continuous on finite-dimensional linear subspaces in the space G and $\varphi(0) = 1$.*

We note that if a quasi-measure μ is symmetric, i.e., $\mu(A) = \mu(-A)$ for every set $A \in Cyl(X, G)$, then $\widetilde{\mu}$ is real.

7.13.5. Lemma. *If μ and ν are measures on $\sigma(X^*)$ and $\widetilde{\mu} = \widetilde{\nu}$, then one has $\mu = \nu$. The same is true for Radon measures.*

PROOF. For all functionals $f_1, \ldots, f_n \in X^*$ by Proposition 3.8.6 we have $\mu \circ (f_1, \ldots, f_n)^{-1} = \nu \circ (f_1, \ldots, f_n)^{-1}$. Hence $\mu = \nu$ on $\sigma(X^*)$, which for Radon measures yields the equality on $\mathcal{B}(X)$. $\square$

It is clear that by the dominated convergence theorem the characteristic functional of any measure on $\sigma(X^*)$ is sequentially continuous. Hence if μ is a measure on a normed space X, then the function $\widetilde{\mu}$ is continuous with respect to the norm on X^*. In the general case, the characteristic functional of a Radon measure is not continuous in the weak* topology $\sigma(X^*, X)$. For example, if X is an infinite-dimensional locally convex space, then the function $\widetilde{\mu}$ is $\sigma(X^*, X)$-continuous only in the case, where μ is concentrated on the union of a sequence of finite-dimensional subspaces (Exercise 7.14.133).

Let us give a sufficient condition of continuity of the Fourier transform of a measure. Recall that a locally convex space X is called barrelled if every closed symmetric convex set whose multiples cover X contains a neighborhood of zero. The Mackey topology $\tau(X^*, X)$ on the dual X^* to a locally convex space X is the topology of uniform convergence on convex symmetric weakly compact sets in X. Regarding X as the dual to $\bigl(X^*, \sigma(X^*, X)\bigr)$, we obtain the Mackey topology $\tau(X, X^*)$ on X. If the space X is barrelled, then its topology is exactly the Mackey topology. A locally convex space is quasi-complete if all closed bounded sets in it are complete, i.e., all fundamental nets have limits.

7.13.6. Proposition. (i) *Let μ be a Radon measure on a locally convex space X. Then the function $\widetilde{\mu}$ is uniformly continuous in the topology of uniform convergence on compact sets in X, and if X is quasi-complete, then also in the Mackey topology $\tau(X^*, X)$.*

(ii) *If a measure μ is defined on the dual X^* to a barrelled space X and is Radon in the weak* topology, then the function $\widetilde{\mu}$ is uniformly continuous on X.*

Moreover, the characteristic functionals of Radon measures in a uniformly tight bounded family are uniformly equicontinuous in both cases.

PROOF. Let $\|\mu\| \le 1$ and $\varepsilon > 0$. We can find a compact set K such that $|\mu|(X\backslash K) < \varepsilon$. Let us take in X^* the following neighborhood of zero: $U := \{y \in X^*\colon\ \sup_{x\in K} |y(x)| < \varepsilon\}$. Then, by the estimate $\bigl|\exp\bigl(iy(x)\bigr) - 1\bigr| \le |y(x)|$ we have for all $y \in U$

$$\int_X |\exp(iy) - 1|\, d|\mu| \le 2|\mu|(X\backslash K) + \int_K |\exp(iy) - 1|\, d|\mu| \le 2\varepsilon + \varepsilon.$$

It remains to use the estimate

$$\bigl|\widetilde{\mu}(y_1) - \widetilde{\mu}(y_2)\bigr| \le \int_X |\exp(iy_1) - \exp(iy_2)|\, d|\mu| \le \int_X |\exp[i(y_1 - y_2)] - 1|\, d|\mu|.$$

If X is quasi-complete, then the closed convex envelope of any compact set is compact, hence K can be made convex. In particular, this is the case if X is the dual to a barrelled space (see Schaefer [**1661**, Ch. II, Corollary in §4.3, Ch. IV, §6.1]). The last claim of the proposition is clear from our reasoning. $\square$

In general, $\widetilde{\mu}$ may not be continuous in the Mackey topology (see Kwapień, Tarieladze [**1095**]).

We note the following simple estimate useful in the study of characteristic functionals: if μ is a probability quasi-measure on $Cyl(X, G)$, then for all $l \in G$ we have

$$|\widetilde{\mu}(l) - 1| \le \int_X |l(x)|\, \mu(dx) \le \biggl(\int_X l(x)^2\, \mu(dx)\biggr)^{1/2}. \tag{7.13.2}$$

One can ask under what conditions a function $\varphi\colon X^* \to \mathbb{C}$ is the characteristic functional of a (Radon) measure on X. In the case of a nonnegative measure on $\mathbb{R}^n$, the Bochner theorem asserts that this is so if and only if φ is continuous and positive definite. This is not true in general infinite-dimensional spaces. For example, the function $e^{-(x,x)}$ on the infinite-dimensional Hilbert space $X = l^2$ is not the characteristic functional of a Borel measure because it is not sequentially continuous in the weak topology. Important infinite-dimensional generalizations of the Bochner theorem are given by the Minlos and Sazonov theorems. The Sazonov theorem [**1655**] states that a function φ on a Hilbert space X is the characteristic functional of a nonnegative Radon measure on X if and only if it is positive definite and continuous in the topology generated by all seminorms of the form $x \mapsto |Tx|$, where T is a Hilbert–Schmidt operator on X. According to the Minlos theorem [**1320**], if X is the dual to a barrelled nuclear space Y, then the same is true for the Mackey topology on X. The role of Hilbert-Schmidt operators in both theorems was clarified by Kolmogorov [**1031**].

A continuous linear operator on a Hilbert space X is called a Hilbert–Schmidt operator if for some orthonormal basis $\{e_\alpha\}$, the sum of the series $\sum_\alpha |Te_\alpha|^2$ is finite (then this sum is independent of the basis). An operator S on H is called nonnegative nuclear if S is a symmetric operator such that $(Sx,x) \geq 0$ for all x and $\sum_\alpha (Se_\alpha, e_\alpha) < \infty$ for some (and then for all) orthonormal basis $\{e_\alpha\}$. Given a locally convex space X, we denote by $\mathcal{LS}(X^*, X)$ the class of all operators $R\colon X^* \to X$ of the form $R = ASA^*$, where S is a symmetric nonnegative nuclear operator in some separable Hilbert space H and $A\colon H \to X$ is a continuous linear operator. Let $\mathcal{T}(X^*, X)$ be the locally convex topology on X^* generated by all seminorms $y \mapsto \sqrt{\langle y, Ry\rangle}$, $R \in \mathcal{LS}(X^*, X)$. This topology is called the Sazonov topology. Similarly, one defines the topology $\mathcal{T}(X, X^*)$ on X. The Sazonov topology on a Hilbert space X is generated by the seminorms $x \mapsto |Tx|$, where T is a Hilbert–Schmidt operator on X.

If X is a locally convex space, then the set $M \subset X^*$ is called $\sigma(X^*, X)$-bounded if $\sup_{l\in M} |l(x)| < \infty$ for every $x \in X$. The strong topology $\beta(X, X^*)$ on X is the topology of uniform convergence on all $\sigma(X^*, X)$-bounded sets in X^*.

7.13.7. Theorem. *Let X be a locally convex space and let φ be a positive definite function on X^* that is continuous in the topology $\mathcal{T}(X^*, X)$ with $\varphi(0) = 1$. Then φ is the characteristic functional of a probability measure on X that is Radon with respect to the strong topology $\beta(X, X^*)$.*

PROOF. By the finite-dimensional Bochner theorem, the function φ is the characteristic functional of a cylindrical quasi-measure μ. We have to verify that the measure μ is tight when X is considered with the strong topology. The main idea of the proof is to apply the following estimate. Let μ be a probability measure on $\mathbb{R}^n$, and let A and B be symmetric nonnegative operators on $\mathbb{R}^n$ such that B is invertible. Similarly to Corollary 3.8.16 one proves that if $1 - \mathrm{Re}\widetilde{\mu}(y) \leq \varepsilon$ whenever $(Ay, y) \leq 1$, then for all $C > 0$ one has

$$\mu\bigl(x\colon\ (Bx, x) \geq C\bigr) \leq \frac{\sqrt{e}}{\sqrt{e} - 1}\bigl(\varepsilon + 2C^{-1}\mathrm{trace}AB\bigr).$$

Now one can verify that for every $\varepsilon > 0$, there exists a compact ellipsoid K_ε in X such that $\mu^*(K_\varepsilon) > 1 - \varepsilon$. This ellipsoid is constructed in the following way. Given $\delta > 0$, there exists a seminorm $q_\delta \in \mathcal{T}(X^*, X)$ with the property that $1 - \mathrm{Re}\widetilde{\mu}(y) \le \delta$ whenever $q_\delta(y) < 1$. Let $S := \{y \in X^*\colon\ q_\delta(y) < C\}$ and

$$K_\varepsilon := \bigl\{x \in X\colon\ \sup_{y\in S} |y(x)| \le 1\bigr\}.$$

By using the aforementioned inequality one can choose δ and C such that the set K_ε will be as required. Since the corresponding arguments are presented in detail in Bourbaki [**242**, Ch. IX, §6], Vakhania, Tarieladze, Chobanyan [**1910**, Ch. VI, §4], Daletskii, Fomin [**394**, Ch. III, §1], and Smolyanov, Fomin [**1755**, §4], we do not reproduce them here. □

7.13.8. Corollary. *A function φ on a Hilbert space X with $\varphi(0) = 1$ is the characteristic functional of a Radon probability measure on X if and only if it is positive definite and continuous in the Sazonov topology generated by all seminorms of the form $x \mapsto |Tx|$, where T is a Hilbert–Schmidt operator on X.*

Proof. The sufficiency of continuity in the Sazonov topology is clear from the theorem, since $R = \sqrt{S}$ is a Hilbert–Schmidt operator for any nonnegative nuclear operator S on X. Now let μ be a Radon probability measure on X. It suffices to verify the continuity in the Sazonov topology in the case where μ is concentrated on the ball of radius M centered at the origin, since the measures $I_{U_n} \cdot \mu$, where U_n is the ball of radius n centered at the origin, converge in the variation norm to μ, and their characteristic functionals converge uniformly to $\widetilde{\mu}$. The nonnegative operator S defined by the equality

$$(Su, v) = \int_X (u, x)(v, x)\, \mu(dx),$$

is nuclear, since for any orthonormal basis $\{e_j\}$ one has

$$\sum_{j=1}^{\infty} (Se_j, e_j) = \int_X |x|^2\, \mu(dx) \le M^2.$$

It remains to apply (7.13.2), which yields $|\widetilde{\mu}(y) - 1| \le |\sqrt{S}y|$. □

In general Banach spaces, the condition of Theorem 7.13.7 is not necessary (see Vakhania, Tarieladze, Chobanyan [**1910**], Mushtari [**1348**]). Moreover, the Radon measures on a Banach space X with $\mathcal{T}(X^*, X)$-continuous characteristic functionals are precisely the measures concentrated on continuously embedded separable Hilbert spaces. In order to obtain the Minlos theorem, one has to consider the case where X is the dual to a nuclear space. Namely, by using Theorem 7.13.7 one proves the following.

7.13.9. Theorem. *Let E be a nuclear locally convex space.*

(i) *Let φ be a positive definite function on E with $\varphi(0) = 1$ that is continuous in the topology $\mathcal{T}(E, E^*)$. Then φ is the characteristic functional of* *a probability measure on E^* that is Radon with respect to the strong topology $\beta(E^*, E)$.*

(ii) *If E^* is metrizable or barrelled, then the characteristic functional of any probability measure on E^* that is Radon in the weak* topology $\sigma(E^*, E)$ (e.g., is Radon in the strong topology $\beta(E^*, E)$) satisfies the conditions in* (i).

It should be noted that in the above theorem, it is not enough to have only the sequential continuity of the characteristic functional. For example, for any compact symmetric nonnegative operator S on l^2 that has no finite trace, the function $\exp\bigl(-(Sx, x)\bigr)$ is the characteristic functional of a non-countably additive Gaussian cylindrical quasi-measure on l^2 and is sequentially continuous even in the weak topology (which is weaker than the Sazonov topology).

The analysis of the proof of Theorem 7.13.7 yields at once the following statement (see details in Daletskii, Fomin [**394**, Ch. III], Smolyanov, Fomin [**1755**, §4]).

7.13.10. Corollary. (i) *Let M be a family of probability measures on the σ-algebra $\sigma(X^*)$ in a locally convex space X such that their characteristic functionals are equicontinuous at the origin in the topology $\mathcal{T}(X^*, X)$. Then the family M is uniformly tight with respect to the strong topology $\beta(X, X^*)$.*

(ii) *If a locally convex space X is barrelled and nuclear, then the characteristic functionals of any uniformly tight family of Radon* (*with respect to the topology $\sigma(X^*, X)$*) *probability measures on X^* are equicontinuous at the origin in the topology of space X.*

It is important for applications that the above analogs of the Bochner theorem are valid for such spaces as $\mathbb{R}^\infty$, $\mathcal{S}(\mathbb{R}^n)$, $\mathcal{S}'(\mathbb{R}^n)$, $\mathcal{D}(\mathbb{R}^n)$, $\mathcal{D}'(\mathbb{R}^n)$.

7.14. Supplements and exercises

7.14(i). Extensions of product measure

Let X_1 and X_2 be topological spaces with σ-algebras of one of our standard classes (say, Borel or Baire). The space $X = X_1 \times X_2$ is topological as well and can be equipped with the corresponding σ-algebra. If the inclusions $\mathcal{B}(X_1)\otimes\mathcal{B}(X_2) \subset \mathcal{B}(X)$, $\mathcal{B}a(X_1)\otimes\mathcal{B}a(X_2) \subset \mathcal{B}a(X)$ are strict, then the question arises about extensions of a product measure μ to these larger σ-algebras (see §7.6). There are trivial cases, where μ is defined on $\mathcal{B}(X)$ or $\mathcal{B}a(X)$. For example, if the spaces X_i have countable bases, then $\mathcal{B}(X) = \mathcal{B}(X_1)\otimes\mathcal{B}(X_2)$, and if both X_1 and X_2 are compact, then $\mathcal{B}a(X) = \mathcal{B}a(X_1)\otimes\mathcal{B}a(X_2)$ (see Lemma 6.4.2).

According to Fremlin [**622**], $\mathcal{B}(X_1\times X_2)$ may not belong to the Lebesgue completion of $\mathcal{B}(X_1)\otimes\mathcal{B}(X_2)$ with respect to the measure $\mu_1\otimes\mu_2$ even if both measures μ_1 and μ_2 are completion regular (see Definition 7.14.17) Radon measures on compact spaces. However, as we know, the product measure admits a Radon extension. It remains an open problem whether the product of two Borel measures on topological spaces can be always extended to a Borel measure (this problem is not solved even for purely atomic measures on compact spaces). It is not known whether there exists a non-Radon Borel extension of the product of two Radon measures on compact spaces. The following result shows that the condition in Theorem 7.6.5 can be partly relaxed.

7.14.1. Theorem. *Let μ_1 and μ_2 be Borel measures on topological spaces X_1 and X_2, respectively. Then, the product measure $\mu=\mu_1\otimes\mu_2$ extends to a Borel measure on $X=X_1\times X_2$ in either of the following cases*:

(i) *at least one of the measures μ_1 and μ_2 is τ-additive* (*for example, is Radon*);

(ii) *either X_1 or X_2 is a first countable space.*

Assertion (i) is obvious from Lemma 7.6.4 (it was noted in Godfrey, Sion [**703**], Ressel [**1555**], Johnson [**911**]), and (ii) can be found in Johnson [**907**]. As observed by R.A. Johnson (see Gardner [**660**, Section 26]), in case (i) there may exist two different Borel extensions of $\mu_1\otimes\mu_2$. The proof of (i) employs the following natural construction of a product of two probability Borel measures μ and ν on topological spaces X and Y. Given a set $B\in\mathcal{B}(X\times Y)$, the sets $B_x := \{y\colon (x,y)\in B\}$ are Borel in Y. Hence the function $x\mapsto\nu(B_x)$ is well-defined. If this function is μ-measurable (as is the case if ν is τ-additive), then we shall say that the measure $\nu\mu$ is defined and set

$$\nu\mu(B) := \int_X \nu(B_x)\,\mu(dx).$$

It is clear that such a measure is a Borel extension of $\mu\otimes\nu$. However, Johnson [**908**] constructed examples where the measure $\nu\mu$ is not defined. In addition, he constructed an example where the measure $\nu\mu$ is defined whereas the measure $\mu\nu$ is not. Finally, there is an example (Exercise 7.14.111) where $X=Y$, and both measures $\nu\mu$ and $\mu\nu$ are defined, but are not equal.

We close this subsection with two interesting results on infinite products.

7.14.2. Theorem. *Let μ_n be τ-additive probability measures on topological spaces X_n, $n\in\mathbb{N}$. Then the measure $\mu=\bigotimes_{n=1}^\infty \mu_n$ on the σ-algebra $\bigotimes_{n=1}^\infty \mathcal{B}(X_n)$ in the space $X=\prod_{n=1}^\infty X_n$ is τ-additive as well and extends to a τ-additive measure on $\mathcal{B}(X)$.*

The proof is delegated to Exercise 7.14.70.

We have seen in Example 7.3.1 that the Lebesgue completion of an uncountable product of Dirac measures is not defined on all Borel sets. The following theorem shows that this effect is caused by open sets of zero measure in the factors.

7.14.3. Theorem. *Let T be a nonempty set and let X_t, $t \in T$, be separable metric (or Souslin) spaces with Radon probability measures μ_t such that for every t, the measure μ_t does not vanish on nonempty open sets. Let $X := \prod_{t\in T} X_t$ and $\mu = \bigotimes_{t\in T} \mu_t$. Then $\mathcal{B}(X)$ belongs to the Lebesgue completion of $\bigotimes_{t\in T} \mathcal{B}(X_t)$ with respect to μ, and μ is τ-additive. In particular, in the case of separable metric spaces or completely regular Souslin spaces, μ is completion regular in the sense of Definition* 7.14.17 *below.*

PROOF. (1) Let U_α, where α belongs to some index set, be nonempty open sets of the form $V_\alpha \times Y_\alpha$, where V_α is an open set in the product of finitely many spaces X_t and Y_α is the product of the remaining X_t. Let $U := \bigcup_\alpha U_\alpha$. We show that there exists a finite or countable set of indices α_n such that $\mu\bigl(U\backslash \bigcup_{n=1}^\infty U_{\alpha_n}\bigr) = 0$. By Corollary 4.7.3, there exists a countable set of indices α_n such that $\mu\bigl(U_\alpha\backslash \bigcup_{n=1}^\infty U_{\alpha_n}\bigr) = 0$ for each α. We show that this is the required set. Since each U_{α_n} depends only on finitely many coordinates, one can find a finite or countable set $S \subset T$ with the property that every U_{α_n} has the form $U_{\alpha_n} = W_n \times Y$, where W_n is an open set in $\prod_{s\in S} X_s$ and $Y := \prod_{t\in T\backslash S} X_t$. Denote by π the projection to the countable product $\prod_{s\in S} X_s$ and set $U' := \bigcup_{n=1}^\infty U_{\alpha_n}$. The set $\pi(U)$ is open in $\prod_{s\in S} X_s$. Since $U' \subset U \subset \pi^{-1}\bigl(\pi(U)\bigr)$, where the open sets U' and $\pi^{-1}\bigl(\pi(U)\bigr)$ belong to $\bigotimes_{t\in T} \mathcal{B}(X_t)$, it suffices to show that $\mu(U') = \mu\bigl(\pi^{-1}\bigl(\pi(U)\bigr)\bigr)$. Suppose that $\mu(U') < \mu\bigl(\pi^{-1}\bigl(\pi(U)\bigr)\bigr)$, i.e., $\mu\circ\pi^{-1}\bigl(\pi(U')\bigr) < \mu\circ\pi^{-1}\bigl(\pi(U)\bigr)$. In the case of separable metrizable spaces, the product $\prod_{s\in S} X_s$ is separable metrizable as well, and the set $\pi(U)$ is the union of open (in this space) sets $\pi(U_\alpha)$. Therefore, $\pi(U)$ coincides with some finite or countable union of these sets. The same is true in the case of Souslin spaces. Hence, there exists α such that

$$\mu\circ\pi^{-1}\bigl(\pi(U_\alpha)\backslash\pi(U')\bigr) > 0. \tag{7.14.1}$$

The set U_α can be written in the form $U_\alpha = W_1 \cap W_2$, where

$$W_1 = G\times \prod_{s\in S\backslash F} X_s \times \prod_{t\in T\backslash S} X_t,\ W_2 = \prod_{s\in S} X_s \times W \times \prod_{t\in T\backslash(S\cup N)} X_t,$$

$F \subset S$ and $N \subset T\backslash S$ are finite sets, G is open in $\prod_{s\in F} X_s$, W is open in $\prod_{t\in N} X_t$. It is clear that $\mu(U_\alpha\backslash U') = \mu(W_2)\mu(W_1\backslash U')$ by the definition of the product measures (in this case everything reduces to the countable product over the indices in $S\cup N$). Our hypothesis yields that $\mu(W_2) > 0$, since this number equals the measure of the nonempty open set W in the finite product of the spaces X_t, $t \in N$. By the construction of U' we have $\mu(U_\alpha\backslash U') = 0$. Hence $\mu(W_1\backslash U') = 0$. This contradicts (7.14.1), since we have $\pi(U_\alpha) = G\times\prod_{s\in S\backslash F} X_s = \pi(W_1)$ and $\mu(W_1\backslash U') = \mu\circ\pi^{-1}\bigl(\pi(W_1)\backslash\pi(U')\bigr)$.

(2) By the above, all open sets belong to the completion of $\bigotimes_{t\in T} \mathcal{B}(X_t)$, hence it contains $\mathcal{B}(X)$. In addition, we obtain the τ-additivity of μ. $\square$

7.14.4. Remark. It is clear from the proof that this theorem extends to more general spaces, for example, hereditary Lindelöf. One could also require the validity of the conclusion for all finite products of τ-additive measures μ_t that are positive on nonempty open sets.

7.14(ii). Measurability on products

When one considers functions on the product $X\times Y$ of topological spaces, the following two questions frequently arise:

(a) the measurability of the function $f(x,y)$ in the situation where the functions $x\mapsto f(x,y)$ and $y\mapsto f(x,y)$ possess certain nice properties,

(b) the measurability or continuity of the function

$$x\mapsto \int_Y f(x,y)\,\nu(dy), \tag{7.14.2}$$

where ν is a measure on Y,

(c) the measurability of the function $f\bigl(x,\varphi(x)\bigr)$ for a mapping $\varphi\colon\ X\to Y$.

In Lemma 6.4.6 and Exercise 6.10.43 we have already encountered question (a); Corollary 3.4.6 and Lemma 7.6.4 were concerned with question (b). In this subsection, some additional related facts are mentioned. Exercises 7.14.102–7.14.106 contain information on question (a). In particular, it turns out that if X and Y are equipped with Radon measures μ and ν, and all compact sets in Y are metrizable (for example, Y is a Souslin space), then the continuity of f in y and its μ-measurability in x yield the measurability with respect to $\mu\otimes\nu$. However, one cannot omit the requirement of metrizability of compact sets in Y. Under the continuum hypothesis, Fremlin [**621**] constructed a counter-example (see Exercise 7.14.106). Let us mention an interesting result from Johnson [**905**] and Moran [**1329**], extended in Fremlin [**621**] to arbitrary finite products.

7.14.5. Theorem. *Let μ and ν be Radon probability measures on X and Y and let a function $f\colon\ X\times Y\to\mathbb{R}^1$ be continuous in every argument separately. Then f is measurable with respect to the Radon measure on $X\times Y$ that is the extension of $\mu\otimes\nu$.*

The proof and a more general assertion can be found in Exercise 7.14.105. We remark that this theorem follows at once from Proposition 5.2 in Burke, Pol [**285**], according to which every separately continuous function on the product of two compact spaces is jointly Borel measurable.

Concerning question (b) we note that if a function f is bounded and continuous in every argument separately, then in the case of a metrizable space X, the function (7.14.2) is continuous on X by the dominated convergence theorem. If X is Souslin (or compact sets in X are metrizable), then such functions are μ-measurable due to the sequential continuity. In the general case, the function (7.14.2) may fail to be continuous.

7.14.6. Example. Let $X=[0,1]$ be equipped with Lebesgue measure and let Y be the space of all continuous functions from $[0,1]$ to $[0,1]$ with the topology of pointwise convergence. Set $F(x,y) = y(x)$, $x \in X$, $y \in Y$. The function F is continuous in every argument separately, but the function

$$\varphi(y) = \int_0^1 F(x,y)\,dx = \int_0^1 y(x)\,dx$$

is discontinuous on Y since for any $x_1, \ldots, x_n \in [0,1]$, there is $y \in Y$ with $y(x_i) = 0$ and $\varphi(y) > 1/2$, although φ is sequentially continuous.

Now we give a positive result from Glicksberg [**697**].

7.14.7. Theorem. *Let X be a compact space, let Y be a Hausdorff space, and let $f\colon X \times Y \to \mathbb{R}^1$ be a bounded function that is continuous in every argument separately. Then, for every Radon measure ν on Y, the function* (7.14.2) *is continuous.*

PROOF. Since ν is a limit of a sequence of Radon measures with compact support convergent in variation, it suffices to consider the case where Y is compact. For every $x \in X$, we consider the function $f_x\colon y \mapsto f(x,y)$. By the continuity of f in the second argument, we have $f_x \in C(Y)$. By the continuity of f in the first argument, the mapping $x \mapsto f_x$ from X to the space $C(Y)$ with the topology of pointwise convergence is continuous. Hence the image Φ of this mapping is compact in the pointwise topology. By the boundedness of f the set Φ is norm bounded in $C(Y)$. By Theorem 7.10.9 the topology of pointwise convergence coincides on Φ with the weak topology. Therefore, the considered mapping is continuous if we equip $C(Y)$ with the weak topology, which proves our assertion. □

Exercise 7.14.107 gives some generalization of this theorem. For jointly continuous functions the situation simplifies; the proof of the next result is left as Exercise 7.14.108.

7.14.8. Proposition. *Suppose that X and Y are Hausdorff spaces. Let μ be a τ-additive measure on Y and let $f\colon X{\times}Y \to \mathbb{R}^1$ be a bounded continuous function. Then the function* (7.14.2) *is continuous.*

Concerning question (c), see Exercise 7.14.113. The measurability of separately continuous functions is also considered in Janssen [**883**].

7.14(iii). Mařík spaces

Mařík [**1267**] obtained the following result.

7.14.9. Theorem. *If a space X is normal and countably paracompact, then every Baire measure μ on X has a regular Borel extension ν that, for every open set $U \subset X$, satisfies the condition*

$$|\nu|(U) = \sup\bigl\{|\mu|(F)\colon\ F \subset U,\, F = f^{-1}(0),\, f \in C_b(X)\bigr\}.$$

This nice result gave rise to the problem of characterization of topological spaces with the Mařík property.

7.14.10. Definition. *Let X be a completely regular space.*

(i) *The space X is called a Mařík space if every Baire measure on X extends to a regular Borel measure.*

(ii) *The space X is called a quasi-Mařík space if every Baire measure on X extends to a Borel measure* (*not necessarily regular*)*.*

(iii) *The space X is called measure-compact* (*or almost Lindelöf*) *if every Baire measure on X has a τ-additive Borel extension.*

By definition, every normal countably paracompact space is a Mařík space. It has already been noted (see Example 7.3.9 and Exercise 7.14.69) that not all completely regular spaces are Mařík. A general result, which gives a lot of examples with additional interesting properties, is proved in Ohta, Tamano [**1394**]. In particular, according to [**1394**, Example 3.5], there exists a countably paracompact space X with a Baire measure μ without Borel extensions. Under some additional set-theoretic assumptions, there exists a normal space X with a Baire measure without Borel extensions (see Fremlin [**635**, §439N]). Thus, both conditions in Mařík's theorem are essential. Trivial examples of Mařík spaces are perfectly normal spaces. Compact spaces are less trivial examples, since we know that a Baire measure on a compact space may possess Borel extensions that are not regular. It is clear by Theorem 7.3.2(ii) that any measure-compact space is Mařík. As shown in Fremlin [**623**], under Martin's axiom and the negation of the continuum hypothesis, the space $\mathbb{N}^{\omega_1}$ is measure-compact (hence Mařík), but is neither normal nor countably paracompact. As shown in Moran [**1328**] and Kemperman, Maharam [**980**], such standard spaces of measure theory as $\mathbb{R}^{\mathfrak{c}}$ and $\mathbb{N}^{\mathfrak{c}}$, where $\mathfrak{c}$ is the cardinality of the continuum, are not measure-compact. Under some additional set-theoretic axiom, Aldaz [**19**] established the existence of a normal quasi-Mařík space that is not Mařík. On the other hand, it is shown in [**19**] that a quasi-Mařík space X is Mařík if every countable open cover of X has a pointwise finite refinement. It is known that the product of any family of metric spaces is a quasi-Mařík space (Ohta, Tamano [**1394**]). It is unknown whether such a product is always Mařík (in particular, it is even unknown whether any power of $\mathbb{N}$ is a Mařík space). According to [**1394**, Example 3.16], the union of two Mařík spaces may not be a quasi-Mařík space even if one of them is a functionally open set and the other one is a functionally closed set. There exists a first countable locally compact space X possessing a Baire probability measure μ that has no Borel extensions (see Fremlin [**635**, §439L]). Aldaz [**19**] has shown that the union $X = Y \cup K$ and the product $X = Y \times K$, where Y is a Mařík space and K is compact, are Mařík spaces. Gale [**651**] proved that the union of a compact space and a measure-compact space is measure-compact. It is worth noting that every $\mathcal{F}$-analytic set (hence every Baire set) in a measure-compact space is measure-compact, see Fremlin [**635**, §436G]. Some additional information can be found in Adamski [**7**], Aldaz [**19**], Bachman, Sultan [**89**], Gale [**651**], Gardner, Gruenhage [**664**], Kirk [**1004**], Koumoullis [**1046**], Ohta, Tamano [**1394**], Wheeler [**1978**], [**1979**].

7.14(iv). Separable measures

In applications it is often desirable to deal with separable measures. By definition (see §1.12(iii)), a bounded measure μ on $(X, \mathcal{B})$ is separable if there exists an at most countable family $\mathcal{C} \subset \mathcal{B}$ such that for every $B \in \mathcal{B}$ and every $\varepsilon > 0$, one can find a set $C \in \mathcal{C}$ with $|\mu|(B \triangle C) < \varepsilon$ (in other words, the countable family $\mathcal{C}$ is dense in the measure algebra associated with $|\mu|$). It is easily verified that μ is separable if and only if all spaces $L^p(\mu)$, where $p \in (0, \infty)$, are separable (in fact, the separability of either of these spaces is enough, see Exercise 4.7.63). The connections between the separability of a measure and its topological regularity properties are not very strong. For example, the product μ of the continuum of copies of Lebesgue measure on $I = [0, 1]$ is a nonseparable Radon measure on a separable compact space $I^{\mathfrak{c}}$ (the mutual distances in $L^2(\mu)$ between the coordinate functions are equal positive numbers). On the other hand, let us consider an example of a Radon measure μ on a compact space X that vanishes on every metrizable compact set, hence on every Souslin set in X (according to Exercise 7.14.156, so does the above-mentioned product), but has separable $L^1(\mu)$.

7.14.11. Example. Let X be the space "two arrows" (see Example 6.1.20). The space X is compact, separable, perfectly normal, hereditary Lindelöf and satisfies the first axiom of countability, but every metrizable subspace in X is at most countable. In addition:

(i) the Borel σ-algebra of X is generated by a countable family and singletons, and every Borel measure on X is separable;

(ii) there exists a Radon probability measure μ on X (the natural normalized linear Lebesgue measure on X) such that its image under the natural projection coincides with Lebesgue measure on $[0, 1]$, and μ vanishes on all metrizable subspaces in X (hence on all Souslin subsets in X).

PROOF. The topological properties of X are listed in Example 6.1.20. We recall that $\mathcal{B}(X)$ is contained in the Borel σ-algebra generated by the standard topology of $\mathbb{R}^2$, since X is hereditary Lindelöf and every open set in X is an at most countable union of elements of the base. According to Exercise 6.10.36, $\mathcal{B}(X)$ consists of all sets B such that for some Borel set $E \subset [0, 1]$, the set $B \triangle \pi^{-1}(E)$ is at most countable, where $\pi \colon X \to [0, 1]$ is the natural projection. It is clear from this description that $\mathcal{B}(X)$ is generated by a countable family and singletons and that every measure on $\mathcal{B}(X)$ is separable. The measure μ is given by the formula $\mu(B) = \lambda(E)$. The Radon property of μ is obvious from the fact that the set $S := \pi\bigl(B \triangle \pi^{-1}(E)\bigr)$ is at most countable, hence for every $\varepsilon > 0$, the set $E \backslash S$ contains a compact subset K with $\lambda(K) > \lambda(E) - \varepsilon$, and the set $\pi^{-1}(K)$ is compact in X. By construction, μ vanishes on all countable sets, hence by property (i) on all metrizable subsets (which yields that it vanishes on all Souslin subset in X). Note that μ is a unique probability measure on $\mathcal{B}(X)$ with the projection λ. We observe that every measure on $\mathcal{B}(X)$ is Radon (the proof is similar). $\square$

The following result (its proof is delegated to Exercise 7.14.147) gives some sufficient conditions of separability.

7.14.12. Proposition. *Either of the following conditions is sufficient for separability of a Borel measure μ on a space X:*

(i) *the space X is hereditary Lindelöf and there exists a countable family of measurable sets approximating with respect to μ every element of some base of the topology in X;*

(ii) *for each $\varepsilon > 0$, there exists a metrizable compact set K_ε such that one has $|\mu|(X \backslash K_\varepsilon) < \varepsilon$.*

7.14.13. Example. Suppose that all compact subsets in X are metrizable. Then every Radon measure on X is separable.

We recall that a simple necessary and sufficient condition of the metrizability of a compact space K is the existence of a countable family of continuous functions separating the points in K.

7.14(v). Diffused and atomless measures

7.14.14. Definition. *A Borel measure on a Hausdorff space is called diffused or continuous if it vanishes on all singletons.*

Let us recall a concept already encountered in §1.12(iii).

7.14.15. Definition. *Let $(M, \mathcal{M}, \mu)$ be a space with a nonnegative measure. An element $A \subset M$ is called an atom of the measure μ if $\mu(A) > 0$ and every element B in $\mathcal{M}$ that is contained in A, has measure either zero or $\mu(A)$. A measure without atoms is called atomless.*

It is clear that any atomless Borel measure is diffused. The following assertion is obvious (see Exercise 7.14.148).

7.14.16. Lemma. *Every diffused τ-regular (for example, Radon) measure is atomless.*

There exist diffused Borel measures with atoms. An example is the Dieudonné measure (see Example 7.1.3), for which the whole space is an atom (since this measure assumes only two values).

It is shown in Grzegorek [**751**] that there exist two countably generated σ-algebras $\mathfrak{S}_1$ and $\mathfrak{S}_2$ such that on each of them there exist atomless probability measures, but there are no such measures on $\sigma(\mathfrak{S}_1 \cup \mathfrak{S}_2)$.

7.14(vi). Completion regular measures

7.14.17. Definition. (i) *A Baire measure is called completion regular if its Lebesgue extension contains the Borel σ-algebra. A Borel measure is called completion regular if its restriction to the Baire σ-algebra is completion regular; in other words, for every $B \in \mathcal{B}(X)$, there exist $B_1, B_2 \in \mathcal{B}a(X)$ with*

$$B_1 \subset B \subset B_2 \quad \textit{and} \quad |\mu|(B_1 \backslash B_2) = 0.$$

(ii) *A Baire measure is called monogenic if it has a unique regular Borel extension. A Borel measure is called monogenic if so is its Baire restriction.*

It is clear that any completion regular measure is monogenic, but the converse is not true (for example, for the Dieudonné measure). There exists a Radon measure on a Radon space (a space on which every Borel measure is Radon, see the next subsection) such that it is not completion regular. See references and additional results in Gardner [**660**, §21].

According to Theorem 7.14.3, the product of any family of Radon probability measures on separable metric (or Souslin) spaces is completion regular, provided these measures are positive on nonempty open sets. An important example of a completion regular measure is the Haar measure on any locally compact group (see Theorem 9.11.6).

It is unknown whether in the ZFC there exists an example of a completion regular, but not τ-additive measure on a completely regular space. Moran [**1328**] constructed an example of a Baire measure on $\mathbb{R}^{\mathfrak{c}}$ that is not τ-additive, but his measure is not completion regular. Assuming that there is a measurable cardinal, we obtain a Baire measure on a metric space that is not τ-additive (but is completion regular, of course). Let us consider a class of spaces on which any completion regular measure is τ-additive.

A space X is called *dyadic* if it is a continuous image of the space $\{0,1\}^I$ for some set I. The following spaces are dyadic: (i) compact metric spaces, (ii) finite unions and arbitrary products of dyadic spaces, (iii) functionally closed sets in dyadic spaces, (iv) compact topological groups. Fremlin and Grekas [**637**] introduced the larger class of *quasi-dyadic spaces*, i.e., continuous images of arbitrary products of separable metric spaces. According to [**637**], continuous images, arbitrary products, and countable unions of quasi-dyadic spaces are quasi-dyadic. In addition, the Baire subsets of quasi-dyadic spaces are quasi-dyadic. The following two results are obtained in [**637**].

7.14.18. Theorem. *Let X be a quasi-dyadic space with a completion regular Borel probability measure μ. Then μ is τ-additive. If, in addition, ν is a τ-additive Borel probability measure on a space Y, then every open subset in $X\times Y$ is measurable with respect to the usual product measure $\mu\otimes\nu$.*

7.14.19. Corollary. *Let X_α, $\alpha\in A$, be a family of quasi-dyadic spaces equipped with completion regular Borel probability measures μ_α. Suppose that all, with the exception of at most countably many, measures μ_α are positive on nonempty open sets. Then the measure $\otimes_\alpha\mu_\alpha$ on the space $\prod_{\alpha\in A} X_\alpha$ is defined on the Borel σ-algebra and is completion regular.*

It is worth noting in this connection that according to Gryllakis, Koumoullis [**750**], if μ_α are τ-additive Borel probability measures such that all τ-additive finite sub-products are completion regular and all measures μ_α, with the exception of at most countably many of them, are positive on nonempty open sets, then the usual product measure is defined on the Borel σ-algebra of $\prod X_\alpha$ and is τ-additive, i.e., μ is completion regular.

7.14(vii). Radon spaces

Let us consider the following classes of topological spaces.

7.14.20. Definition. (i) *A topological space X is called a Radon space if every Borel measure on X is Radon.*

(ii) *A topological space X is called Borel measure-complete if every Borel measure on X is τ-additive.*

Any Radon space is Borel measure-complete, but the converse is false (example: a nonmeasurable subset of an interval). Exercise 7.14.128 lists some properties of Radon spaces. Not all compact spaces are Radon (example: the Dieudonné measure). There exists a first countable compact space that is not Radon (see Fremlin [**635**, §439J]). The class of Radon spaces is not closed with respect to weakening the topology, taking continuous (even injective) images and, under the continuum hypothesis, the product of two compact Radon spaces may not be a Radon space (see Wage [**1955**]). It is unknown whether every continuous image of a Radon compact space in a Hausdorff space is Radon. All known examples of Radon compact spaces are sequentially compact. Some special classes of spaces (for example, Eberlein compacts or Corson compacts) are known to be Radon under additional set-theoretic axioms (see Fremlin [**635**], Gardner [**660**], Schachermayer [**1660**]). Although the definition of Radon spaces is simple and the membership in this class may be important, it appears, on the basis of the above facts, that it would be unlikely that a complete characterization of Radon spaces, were it to be found, could be of great use in applications.

7.14.21. Remark. Sometimes, considering a measure μ on a completely regular space X, it is useful to extend it to the Stone–Čech compactification βX by the formula $\mu_\beta(B) := \mu(B \cap X)$. This is possible for Borel or Baire measures, but X may be nonmeasurable with respect to the corresponding extension μ_β of the measure μ (i.e., may fail to belong to $\mathcal{B}(\beta X)_{\mu_\beta}$ or $\mathcal{B}a(\beta X)_{\mu_\beta}$). Then one of the following additional assumptions may be useful: (1) $X \in \mathcal{B}a(\beta X)$, (2) $X \in \mathcal{B}(\beta X)$, (3) X is measurable with respect to all Radon measures on βX, (4) X is measurable with respect to all Borel measures on βX.

For example, if X is locally compact, then it is open in βX, in particular, $X \in \mathcal{B}(\beta X)$.

7.14.22. Example. (see Alexandroff [**30**], Knowles [**1015**]) Let X be completely regular. Every τ-additive measure on X is Radon if and only if X is measurable with respect to every Radon measure on βX (i.e., is universally Radon measurable in βX).

Proof. If X is universally Radon measurable in βX and μ is a τ-additive measure on X, then its extension μ_β to βX is Radon, which yields that μ is Radon. In order to obtain the inverse implication, it suffices to consider the case where ν is a Radon measure on βX such that X is a set of full outer ν-measure. Then the measure μ on X defined by $\mu(B \cap X) = \nu(B)$, $B \in \mathcal{B}(\beta X)$, is τ-additive. By our hypothesis, it is Radon on X, whence the ν-measurability of X follows. $\square$

7.14(viii). Supports of measures

In connection with supports of measures, questions arise concerning:

(a) the existence of a non-trivial atomless (in the sense of §7.14(v)) Borel measure μ on a given space X,

(b) the existence of μ with the additional property $\operatorname{supp} \mu = X$,

(c) the properties of the support of a given measure (for example, the metrizability).

We recall that for Radon measures the absence of atoms is equivalent to the absence of points of positive measure, but in the general case the first property is strictly stronger. In §9.12(iii), there is a simple proof of the fact that on every nonempty compact space without isolated points, there is an atomless Radon probability measure (but its support may be smaller than the whole space). The following more general result is obtained in Knowles [**1014**].

7.14.23. Theorem. (i) *If X is Čech complete and has no isolated points, then there exists a non-trivial regular atomless Borel measure on X.*

(ii) *If every subset of X contains an isolated point and X is Borel measure-complete* (*see Definition* 7.14.20)*, then there is no non-trivial regular atomless Borel measure on X.*

Babiker [**83**] constructed an example (under the continuum hypothesis) of a completely regular space without isolated points on which there is no non-trivial atomless Borel measure. Necessary and sufficient conditions for the existence of a Radon measure μ with full support on a compact space are obtained in Hebert, Lacey [**805**]. However, such a measure may be atomic. As shown in [**805**], if X is compact and first countable and has no isolated points, then the existence of a Radon measure μ with support X implies the existence of an atomless Radon measure ν with support X. In particular, such a measure ν exists if X is a separable first countable compact space without isolated points. In such problems, various additional set-theoretic assumptions may be essential. For example, under the continuum hypothesis, Kunen [**1077**] constructed a compact, hereditary Lindelöf first countable space X that is nonseparable, but is the support of a Radon measure μ (see also Haydon [**802**]). On the other hand, under Martin's axiom and the negation of the continuum hypothesis, such a space cannot exist (see Juhász [**921**], Fremlin [**627**]). On some spaces, Radon measures are concentrated on subspaces with nice properties. For example, according to the Phillips–Grothendieck theorem, every Radon measure on a weakly compact set in a Banach space has a norm metrizable support. A more general result is given in §7.14(xvii). Let us give a simple result in this direction (see its application in Exercise 7.14.131).

7.14.24. Proposition. *Let μ be a Radon measure on a topological space X such that there exists a sequence of μ-measurable functions f_n separating the points in X. Then, for every $\varepsilon > 0$, there exists a metrizable compact set K_ε with $|\mu|(X\backslash K_\varepsilon) < \varepsilon$.*

PROOF. The hypothesis yields the existence of an injective μ-measurable function g. Since μ is Radon, for every $\varepsilon > 0$, there is a compact set K_ε such that $|\mu|(X\backslash K_\varepsilon) < \varepsilon$ and g is continuous on K_ε. By the injectivity of g the compact sets K_ε are metrizable. □

7.14(ix). Generalizations of Lusin's theorem

The classical Lusin's theorem states that a measurable function f on the space $X = [0,1]$ is *almost continuous* in the sense that given $\varepsilon > 0$, one can find a compact set K_ε such that $\lambda([0,1]\backslash K_\varepsilon) < \varepsilon$ and f is continuous on K_ε. There are a number of generalizations of this theorem: to more general spaces X or to more general spaces of values Y (or both). One can construct an example of a Borel mapping from $X = [0,1]$ to a compact space Y that is not almost continuous with respect to Lebesgue measure (Exercise 7.14.76). A standard generalization (Theorem 7.1.13) covers the case where X is a space with a Radon measure μ and Y is a separable metric space. If, in addition, X is completely regular and Y is a Fréchet space, then as in the classical Lusin theorem, given $\varepsilon > 0$, there exists a continuous mapping $f_\varepsilon\colon X \to Y$ with $|\mu|(f \neq f_\varepsilon) < \varepsilon$. Further generalizations are obtained in Fremlin [**625**] and Koumoullis, Prikry [**1049**] (the latter deals with multivalued mappings), where it is shown that for every Radon measure μ on a space X and every μ-measurable mapping f from X to a metric space Y, there exists a separable subspace Y_0 in Y such that $f(x) \in Y_0$ for μ-a.e. x. In particular, the following generalization of Lusin's theorem is obtained in [**625**]; for simplicity we formulate it for finite measures (for another proof, see Kupka, Prikry [**1081**]).

7.14.25. Theorem. *Let μ be a Radon measure on a topological space X and let Y be a metric space. A mapping $f\colon X \to Y$ is measurable with respect to μ if and only if it is almost continuous.*

In the case of Lebesgue measure the proof is simplified (Exercise 7.14.75). It is shown in Burke, Fremlin [**288**] that under certain additional set-theoretical assumptions, there exists a measurable mapping $f\colon [0,1] \to [0,\omega_1]$ that is not almost continuous, but there are some other set-theoretic assumptions making this impossible according to Fremlin [**625**]; see also Fremlin [**628**].

The next result is a generalization of a theorem obtained in Scorza Dragoni [**1686**] and Krasnosel'skiĭ [**1055**] in the case $X = Y = [a,b]$, in which it is a direct corollary of Lusin's theorem for $C[a,b]$-valued mappings. We follow Berliocchi, Lasry [**159**] (see also Castaing [**317**]). Kucia [**1069**] gives an extension to the case of f with values in a topological space Z with a countable base and to the case of multivalued mappings. The latter case under various assumptions is discussed in many papers on multivalued analysis (see, e.g., Averna [**81**]). Other important results and a survey can be found in Bouziad [**247**].

7.14.26. Theorem. *Let X and Y be two topological spaces such that Y has a countable base, let μ be a regular Borel probability measure on X and let a function $f\colon X\times Y\to\mathbb{R}^1$ be such that, for μ-a.e. $x\in X$, the function $y\mapsto f(x,y)$ is continuous, and for every $y\in Y$, the function $x\mapsto f(x,y)$ is μ-measurable. Then, for every $\varepsilon>0$, there exists a closed set $F\subset X$ such that $\mu(X\backslash F)<\varepsilon$ and $f|_{F\times Y}$ is continuous.*

PROOF. It suffices to consider functions with values in $(0,1)$. Let $\{U_n\}$ be a countable topology base in Y, let $\{y_k\}$ be a dense sequence in Y, and let $\varphi_{n,q}=qI_{U_n}$, $q\in\mathbb{Q}\cap(0,1)$. Set $E_{n,q,k}:=\{x\in X\colon\ f(x,y_k)\geq\varphi_{n,q}(y_k)\}$. Then $E_{n,q}=\bigcap_{k=1}^\infty E_{n,q,k}\in\mathcal{B}(X)_\mu$. It is readily seen that

$$E_{n,q}=\{x\in X\colon\ f(x,y)\geq\varphi_{n,q}(y)\ \forall y\in Y\}.$$

Letting $\psi_{n,q}(x,y)=I_{E_{n,q}}(x)\varphi_{n,q}(y)$, we obtain $f=\sup_{n,q}\psi_{n,q}$. Therefore, arranging the pairs (n,q) in a single sequence, we can write $f(x,y)=\sup I_{A_k}(x)g_k(y)$, where $A_k\in\mathcal{B}(X)_\mu$ and each g_k is a lower semicontinuous function. For every k, there exist a closed set F_k and an open set G_k such that $F_k\subset A_k\subset G_k$ and $\mu(G_k\backslash F_k)<\varepsilon 2^{-k-2}$. The restriction of I_{A_k} to the closed set $B_k=F_k\cup(X\backslash G_k)$ is lower semicontinuous, hence the restriction of $I_{A_k}g_k$ to $B_k\times Y$ is lower semicontinuous. The set $F'=\bigcap_{k=1}^\infty B_k$ is closed, $\mu(X\backslash F')<\varepsilon/2$, and $f|_{F'}$ is lower semicontinuous. Applying the same reasoning to $1-f$ we find a closed set F'' such that $\mu(X\backslash F'')<\varepsilon/2$ and $-f$ is lower semicontinuous on $F''\times Y$. Finally, letting $F=F'\cap F''$, we obtain a desired set. Note that if Y is a compact metric space, then the result follows immediately by Lusin's theorem applied to the following mapping: $\Phi\colon X\to C(Y)$, $\Phi(x)(y)=f(x,y)$. □

It is clear from the proof that an analogous theorem holds for lower semicontinuous functions (see also the papers cited above).

The existence of a countable base in Y is essential and cannot be replaced, for example, by the assumption that Y is a Lusin space. Indeed, let Y be $C[0,1]$ with the pointwise convergence topology (this is a Lusin space with the same Borel σ-algebra as for the standard norm on $C[0,1]$), $X=[0,1]$ with Lebesgue measure, $f(x,y)=y(x)$. Suppose we have a positive measure set F such that f is continuous on $F\times Y$. Then F contains an infinite convergent sequence $\{x_n\}$. One can find a sequence of continuous functions y_n convergent to zero pointwise with $y_n(x_n)\to\infty$, which leads to a contradiction.

Let us see how the above theorem works.

7.14.27. Example. Let μ be a Radon probability measure on a topological space X and let a function $\Phi\colon X\times\mathbb{R}^1\times\mathbb{R}^1\to\mathbb{R}^1$ be measurable in the first variable and continuous in the couple of the last variables. Suppose that a sequence of μ-measurable functions f_n converges in measure to a μ-measurable function f and a sequence of μ-measurable functions g_n is bounded in measure in the sense that $\lim\limits_{M\to\infty} \sup_n \mu\bigl(x\colon\ |g_n(x)| \ge M\bigr) = 0$ (which is fulfilled, e.g., if $\{g_n\}$ is bounded in $L^1(\mu)$). Then the sequence $\psi_n(x) := \Phi\bigl(x, f_n(x), g_n(x)\bigr) - \Phi\bigl(x, f(x), g_n(x)\bigr)$ converges to zero in measure. Observe that in general one cannot replace the functions $\Phi\bigl(x, f(x), g_n(x)\bigr)$ by $\Phi\bigl(x, f(x), g(x)\bigr)$.

PROOF. It suffices to show that any subsequence in $\{\psi_n\}$ contains a further subsequence for which the claim is true because convergence in measure is metrizable. Hence we may assume that $\{f_n\}$ converges a.e. Given $\varepsilon > 0$, we combine Theorem 7.14.26 and Egoroff's theorem to find a compact set $K \subset X$ such that $\mu(K) > 1-\varepsilon$, the restriction of Φ to $K\times\mathbb{R}^1\times\mathbb{R}^1$ is continuous, f is bounded on K, and the sequence $\{f_n\}$ converges to f uniformly on K. There is M such that $\mu\bigl(x\colon\ |g_n(x)| \ge M\bigr) \le \varepsilon$ for all n. Hence for some $N \ge M$ one has $|f_n(x)| \le N$ for all $x \in K$ and all $n \ge N$. By the compactness of $K\times[-N, N]$, there exists $\delta > 0$ such that $|\Phi(x, t, s) - \Phi(x, t', s)| \le \varepsilon$ whenever $t, t' \in [-N, N]$ and $|t - t'| \le \delta$. Hence $|\psi_n(x)| \le \varepsilon$ if $x \in K$, $|f_n(x)| \le N$, and $|g_n(x)| \le N$. Let $n \ge N$. Then

$$\mu\bigl(x\colon\ |\psi_n(x)| \ge \varepsilon\bigr) \le \mu(X\backslash K) + \mu\bigl(x\colon\ |g_n(x)| \ge N\bigr) \le 2\varepsilon,$$

which completes the proof. □

Yet another aspect of Lusin's theorem is related to the approximate continuity. Approximately continuous functions on topological spaces are considered in Sion [**1733**]. Let X be a topological space equipped with a finite nonnegative regular Borel measure μ, let $x \in X$, and let $\mathcal{N}(x)$ denote a basis of neighborhoods of x. A mapping f on X with values in a topological space Y is said to be μ-continuous at x if, for every $\varepsilon > 0$ and every neighborhood V of $f(x)$, there exists a neighborhood U_x of x such that for every W belonging to $\mathcal{N}(x)$ and contained in U_x, we have $\mu\bigl(W - f^{-1}(V)\bigr) \le \varepsilon\mu(W)$. Let us consider the following property (V) (Vitali's property): there exists $\alpha > 0$ such that, for every $A \in \mathcal{B}(X)_\mu$ and every family $\mathcal{U}$ of open sets with the property that every neighborhood W of every $x \in A$ contains some $U \in \mathcal{U} \cap \mathcal{N}(x)$, one can find a countable subfamily $\{U_n\}$ of $\mathcal{U}$ such that (1) $\mu\bigl(A - \bigcup_{n=1}^\infty U_n\bigr) = 0$, (2) for every μ-measurable set $B \subset \bigcup_{n=1}^\infty U_n$ one has $\sum_{W\in F'}(B \cap W) \le \alpha\mu(B)$. The following result is proved in [**1733**].

7.14.28. Theorem. *Let Y have a countable base and let μ have property* (V). *Then $f\colon X \to Y$ is μ-measurable if and only if f is μ-continuous at μ-almost all x. In addition, for every $A \in \mathcal{B}(X)_\mu$, one has the equality $\lim\limits_{W\in\mathcal{N}(x)} \mu(A \cap W)/\mu(W) = 0$ for μ-almost all x.*

The last assertion remains true if in place of property (V) the measure μ possesses property (V') that is defined as follows: only (1) in the definition of (V) is required for some disjoint family $\{U_n\}$.

7.14(x). Metric outer measures

We shall discuss here an application of Carathéodory's method to constructing the so-called metric outer measures on metric spaces, including certain generalizations of Hausdorff measures. A metric outer measure on a metric space (X,d) is a Carathéodory outer measure $\mathfrak{m}$ such that

$$\mathfrak{m}(A\cup B)=\mathfrak{m}(A)+\mathfrak{m}(B)\quad \text{if } \operatorname{dist}(A,B)>0, \tag{7.14.3}$$

where $\operatorname{dist}(A,B) := \inf_{a\in A,b\in B} d(a,b)$, $\operatorname{dist}(A,\varnothing) := +\infty$. We have already encountered this condition in Chapter 1, where we have in fact proved the following result (see Theorem 1.11.10).

7.14.29. Theorem. *A Carathéodory outer measure $\mathfrak{m}$ on a metric space X is a metric outer measure precisely when all Borel sets are $\mathfrak{m}$-measurable.*

We know that the Hausdorff measures H^s satisfy this condition. The measures H^s are obtained as a special case of the measure H^h generated by a set function $h\colon \mathcal{F}\to[0,+\infty]$ defined on some class $\mathcal{F}$ of subsets of X and satisfying the condition $h(\varnothing)=0$. By means of this function one defines the Carathéodory outer measures

$$H^{h,\varepsilon}(A)=\inf\Bigl\{\sum_{j=1}^{\infty}h(F_j)\colon\ F_j\in\mathcal{F}, \operatorname{diam}F_j\le\varepsilon, A\subset\bigcup_{j=1}^{\infty}F_j\Bigr\},\quad \varepsilon>0.$$

If there are no such F_j, then we set $H^{h,\varepsilon}(A)=\infty$. According to the terminology of Chapter 1, the function $H^{h,\varepsilon}$ is the Carathéodory outer measure generated by the function h with the domain consisting of all sets in the class $\mathcal{F}$ of diameter at most ε. Now let

$$H^h(A) := \lim_{\varepsilon\to 0} H^{h,\varepsilon}(A)=\sup_{\varepsilon>0}H^{h,\varepsilon}(A).$$

Letting $h(F)=\alpha(s)2^{-s}(\operatorname{diam}F)^s$, $\alpha(s)=\Gamma(1+s/2)^{-1}$, and $\mathcal{F}=2^X$, we obtain the r-dimensional Hausdorff measure H^s. One can take more general functions $h(F)=\psi(\operatorname{diam}F)$. Certainly, H^h also depends on the choice of the class $\mathcal{F}$.

The proof of the following theorem is the subject of Exercise 7.14.85.

7.14.30. Theorem. *The above-defined Carathéodory outer measure H^h is a metric outer measure.*

Howroyd [**856**] established the following important fact.

7.14.31. Theorem. *Let X be a Souslin metric space and let H^r be the r-dimensional Hausdorff measure on X. Then, for every Borel set $B\subset X$ and every $\alpha<H^r(B)$, there exists a compact set $K\subset B$ with $\alpha\le H^r(K)<\infty$.*

According to a theorem of Davies (see Davies [**410**], Rogers [**1587**]), in the case of Souslin subspaces of $\mathbb{R}^n$ the analogous assertion is true for the measure H^h with an arbitrary strictly increasing continuous function h such that $h(0)=0$. However, for general compact metric spaces, this is not true, as an example in Davies, Rogers [**417**] shows.

7.14.32. Proposition. *Let X be a separable metric space, let $\mathcal{F}$ be a family of subsets of X containing $\mathcal{B}(X)$, and let $h\colon \mathcal{F} \to [0,+\infty]$ be a monotone countably subadditive set function. Then, for every H^h-measurable set A, one has*

$$H^h(A) = \sup_{\Pi} \sum_{B\in\Pi} h(B),$$

where Π runs through the family of all partitions of X into countably many disjoint Borel sets.

In addition, $H^h(A) = \lim_{j\to\infty} \sum_{B\in\Pi_j} h(B)$ for every sequence of partitions Π_j of the set A into countably many disjoint Borel parts of diameter at most δ_j, where $\delta_j \to 0$.

PROOF. Let A_k be Borel sets of diameter at most δ covering A. Then $h(A) \le \sum_{k=1}^{\infty} h(A_k)$, whence we obtain $h(A) \le H^{h,\delta}(A)$ for all $\delta > 0$. Hence $h(A) \le H^h(A)$. For every sequence of pairwise disjoint Borel sets $E_k \subset A$, we obtain $H^h(E) \ge \sum_{k=1}^{\infty} H^h(E_k) \ge \sum_{k=1}^{\infty} h(E_k)$. Thus, $H^h(A)$ is not smaller than the indicated supremum denoted by S. On the other hand, for every $\varepsilon > 0$, there exists $\delta > 0$ such that $H^h(A) \le H^{h,\delta}(A) + \varepsilon$. It is clear from the definition of $H^{h,\delta}(A)$ that the right-hand side is estimated by $S + \varepsilon$ because we can consider partitions of A into Borel parts E_k of diameter at most δ. Therefore, $H^h(A) \le S$. The last claim is clear from the estimate $H^{h,\delta_j}(A) \le \sum_{B\in\Pi_j} h(B)$. $\square$

7.14.33. Theorem. (i) *Let X be a separable metric space, let $(Y, \mathcal{A}, \mu)$ be a measure space, and let $f\colon X \to Y$ satisfy the condition $f\big(\mathcal{B}(X)\big) \subset \mathcal{A}$. We set $\mathcal{F} := \mathcal{B}(X)$ and $h(B) := \mu\big(f(B)\big)$, $B \in \mathcal{B}(X)$. Then, for every H^h-measurable set A, one has*

$$H^h(A) = \int_Y \mathrm{Card}\,\big(A \cap f^{-1}(y)\big)\,\mu(dy).$$

(ii) *If X is a complete separable metric space, Y is a metric space, and a mapping $f\colon X \to Y$ is Lipschitzian with constant L, then for all $B \in \mathcal{B}(X)$ one has*

$$H^n\big(f(B)\big) \le \int_Y \mathrm{Card}\,\big(B \cap f^{-1}(y)\big)\,H^n(dy) \le L^n H^n(B), \quad n \in \mathbb{N}.$$

PROOF. (i) Since there exist Borel sets B_1 and B_2 with $B_1 \subset A \subset B_2$ and $H^h(B_1) = H^h(B_2)$, it suffices to prove our theorem for any Borel set A. Let us take a sequence of decreasing partitions Π_j of the set A into Borel parts $A_{j,k}$ of diameter at most 2^{-j}. Then the functions $\sum_{k=1}^{\infty} I_{f(A_{j,k})}(x)$ increase to $\mathrm{Card}\big(A \cap f^{-1}(x)\big)$ as $j \to \infty$. It remains to use the equalities

$$H^h(A) = \lim_{j\to\infty} \sum_{k=1}^{\infty} h(A_{j,k}) = \lim_{j\to\infty} \sum_{k=1}^{\infty} \int_X I_{f(A_{j,k})}(x)\,\mu(dx)$$

and the monotone convergence theorem.

(ii) For every $B \in \mathcal{B}(X)$, the set $f(B)$ is measurable with respect to H^m. In addition, $h(B) := H^n\bigl(f(B)\bigr) \le L^n H^n(B)$. Let us take a sequence of decreasing partitions Π_k of the set A into Borel parts $A_{k,j}$ of diameter at most 2^{-j}. Then

$$H^h(A) = \lim_{k\to\infty} \sum_{j=1}^{\infty} H^n\bigl(f(A_{k,j})\bigr) \le L^n \lim_{k\to\infty} \sum_{j=1}^{\infty} H^n(A_{k,j}) = L^n H^n(A).$$

Hence we obtain the inequality

$$H^n\bigl(f(A)\bigr) \le \int_Y \operatorname{Card}\bigl(A \cap f^{-1}(y)\bigr)\, H^n(dy) = H^h(A) \le L^n H^n(A)$$

as required. □

7.14(xi). Capacities

Let us make several remarks about capacities, an interesting class of set functions. A Choquet capacity is a function C defined on the family of all subsets of a topological space X and having values in $[0, +\infty]$ such that $C(A) \le C(B)$ if $A \subset B$, $\lim\limits_{n\to\infty} C(A_n) = C(A)$ if the sets A_n are increasing to A, and $\lim\limits_{n\to\infty} C(K_n) = C(K)$ if the sets K_n are compact and decrease to K. If μ is a nonnegative Borel measure on X, then μ^* is a Choquet capacity. Similarly to Theorem 1.10.5 one proves the following Choquet theorem.

7.14.34. Theorem. *Let C be a Choquet capacity on a Souslin space X such that $C(X) < \infty$. Then, for every $\varepsilon > 0$, there exists a compact set K_ε such that $C(K_\varepsilon) > C(X) - \varepsilon$.*

Unlike the case of measures, this property of capacities does not mean that there exist compact sets S_ε with $C(X\backslash S_\varepsilon) < \varepsilon$. Regarding capacities, see Bogachev [**208**], Choquet [**349**], Dellacherie [**424**], [**425**], Goldshtein, Reshetnyak [**709**], Meyer [**1311**], Sion [**1734**].

7.14(xii). Covariance operators and means of measures

Throughout this subsection X is a locally convex space and all measures under consideration are nonnegative. Let X^* denote the dual space to X (the space of all continuous linear functions on X).

7.14.35. Definition. (i) *A measure μ on $\sigma(X^*)$ is said to have a weak moment of order $r > 0$ (or to be of weak order r) if $X^* \subset L^r(\mu)$.*
(ii) *A Borel (or Baire) measure μ on X is said to be a measure with a strong moment of order $r > 0$ (or to be of strong order r) if $\psi \in L^r(\mu)$ for every continuous seminorm ψ on X.*

The atomic measure μ on l^2 with $\mu(ne_n) = n^{-2}$, where $\{e_n\}$ is the standard basis, has a weak first moment because $\sum_{n=1}^\infty n^{-1}|y_n| < \infty$ if $(y_n) \in l^2$, but has no strong first moment, since $\sum_{n=1}^\infty n^{-1} = \infty$.

7.14.36. Definition. *Let μ be a measure on X of weak order* 1. *We shall say that μ has the mean* (*or barycenter*) $m_\mu \in X$ *if for every $l \in X^*$, one has*

$$l(m_\mu) = \int_X l(x)\,\mu(dx).$$

In the general case, the existence of weak moments does not guarantee the existence of the mean. For example, let the measure μ be defined on the space c_0 by $\mu(2^n e_n) = 2^{-n}$, where e_n are the elements of the standard basis in c_0. Then μ has a weak first moment, but has no mean (otherwise all coordinates of the mean would equal 1). It is interesting to note that such an example is impossible in the spaces that do not contain c_0.

7.14.37. Proposition. *If a complete metrizable locally convex space X has no subspace that is linearly homeomorphic to c_0, then every Radon measure μ on X of weak order* 1 *has the mean m_μ.*

The proof is given in Vakhania, Tarieladze [**1909**].

For any measure μ of weak order p on a locally convex space X we obtain the operator $T_\mu\colon\ X^* \to L^p(\mu)$ of the natural embedding.

7.14.38. Lemma. *Let a measure μ on a normed space X have a weak moment of order p. Then the operator $T_\mu\colon\ X^* \to L^p(\mu)$ has a closed graph in the norm topologies and hence is continuous.*

PROOF. If $f_n, f \in X^*$ and $f_n(x) \to f(x)$ pointwise and $f_n \to g$ in $L^p(\mu)$, then the sequence $\{|f_n|^p\}$ is uniformly integrable, whence we obtain that $f_n \to f$ in $L^p(\mu)$ and $f = g$ a.e. The second claim follows by the closed graph theorem due to the completeness of X^*. □

7.14.39. Definition. *Let μ be a probability measure of weak order* 2. *Its covariance $C_\mu\colon\ X^* \times X^* \to \mathbb{R}$ is defined by the formula*

$$C_\mu(l_1, l_2) = \int_X l_1(x) l_2(x)\,\mu(dx) - \int_X l_1(x)\,\mu(dx) \int_X l_2(x)\,\mu(dx).$$

The covariance operator R_μ from X^* to the algebraic dual of X^* is defined by the equality $R_\mu\colon\ X^* \to (X^*)',\ R_\mu(f)(g) = C_\mu(f, g)$.

It is clear that every covariance operator R has the following properties: (1) linearity, (2) nonnegativity, i.e., $\langle f, R(f)\rangle \geq 0$ for all $f \in X^*$, (3) symmetry, i.e., $\langle R(f), g\rangle = \langle R(g), f\rangle$ for all $f, g \in X^*$.

Under broad assumptions, the covariance operators have values in such subspaces of the algebraic dual of X^* as X^{**} or X and are continuous in reasonable topologies. This question is thoroughly investigated in Vakhania, Tarieladze [**1909**]. We mention only few results.

7.14.40. Theorem. *Let μ be a Radon probability measure on a complete* (*or quasi-complete*) *locally convex space X and let μ have a weak second moment. Then $R_\mu(X^*) \subset X$.*

7.14.41. Theorem. *The class of covariance operators of measures of weak second order on a separable Fréchet space X coincides with the class of all symmetric nonnegative operators from X^* to X.*

Typically, the class of covariance operators of measures of strong second order is smaller.

7.14.42. Proposition. *Let H be a separable Hilbert space and let μ be a measure of weak order* 2*. Then μ has a strong second moment if and only if its covariance operator R_μ is nuclear.*

On non-Hilbert spaces, the covariance operators do not characterize the existence of strong moments.

7.14.43. Theorem. *Let X be a Banach space. The following two conditions are equivalent*: (i) *X is linearly homeomorphic to a Hilbert space*; (ii) *for every two Radon probability measures μ and ν with $R_\mu = R_\nu$, the existence of the strong second moment of μ implies the existence of the strong second moment of ν.*

There exists extensive literature on the covariance operators of Gaussian measures (see references in Bogachev [**208**], Vakhania, Tarieladze [**1909**], Vakhania, Tarieladze, Chobanyan [**1910**]). The consideration of strong moments is especially efficient for measures on Banach spaces. Given a Borel probability measure μ on a separable Banach space with a strong first moment, it is often necessary in applications to be able to approximate in the mean the identity operator by "finite-dimensional mappings", i.e., to construct mapping F_n such that

$$\int_X \|x - F_n(x)\|\, \mu(dx) \to 0, \tag{7.14.4}$$

where F_n is finite-dimensional in a reasonable sense, for example, has a finite-dimensional range or depends on finitely many linear functionals (has the form $F_n = G_n(l_1, \ldots, l_k)$, where $l_i \in X^*$ and $G_n\colon \mathbb{R}^k \to X$).

7.14.44. Proposition. *Let X be a separable Banach space, let μ be a Borel probability measure on X, and let $F\colon X \to X$ be a measurable mapping with*

$$\int_X \|F(x)\|^p\, \mu(dx) < \infty,$$

where $p \in [1, \infty)$. Then, for every $\varepsilon > 0$, there exist continuous linear functions $l_1, \ldots, l_n$ on X and a continuous mapping $\varphi\colon \mathbb{R}^n \to X$ with compact support and values in a finite-dimensional subspace such that

$$\int_X \Bigl\| F(x) - \varphi\bigl(l_1(x), \ldots, l_n(x)\bigr) \Bigr\|^p \mu(dx) < \varepsilon.$$

The proof can be found in Exercise 7.14.145.

The obtained approximation is a function of finitely many functionals and has values in a finite-dimensional subspace, but is not linear even for linear continuous F. If X has a Schauder basis $\{e_i\}$ and F is a continuous linear operator, then one can easily construct finite-dimensional linear approximations of F by setting $F_n(x) = \sum_{i=1}^n l_i\bigl(F(x)\bigr)e_i$, where l_i are the coefficients in the expansion with respect to the basis $\{e_i\}$. Corollary 7.14.46 below uses a weaker requirement on X, namely, the approximation property. This property means that for every compact set $K \subset X$ and every $\varepsilon > 0$, there exists a continuous linear operator $T\colon X \to X$ with a finite-dimensional range such that $\|x - Tx\| < \varepsilon$ for all $x \in K$. It is known that not every Banach space possesses such a property.

7.14.45. Theorem. *Let μ be a Borel probability measure on a separable Banach space X with the strong moment of some order $r > 0$. Then, there exists a linear subspace $E \subset X$ with the following properties*:

(i) *E with some norm $\|\cdot\|_E$ is a separable reflexive Banach space whose closed balls are compact in X*; (ii) *$\mu(E) = 1$ and*

$$\int_E \|z\|_E^r \,\mu(dz) < \infty.$$

If μ on X has all strong moments, then E can be chosen with such a property. Finally, these assertions are true for separable Fréchet spaces.

The proof can be found in Exercise 7.14.146 (see also Exercise 8.10.127).

7.14.46. Corollary. *Let μ be a Borel probability measure on a separable Banach space X having the strong moment of order r. Suppose that X has the approximation property. Then, for every $\varepsilon > 0$, there exists a continuous linear operator T with a finite-dimensional range such that*

$$\int_X \|x - Tx\|^r \,\mu(dx) < \varepsilon.$$

PROOF. Let E be the space from the above theorem and let K be its unit ball. We find $\varepsilon_0 > 0$ such that the integral of the function $\|z\|_E^r$ on E is less than $\varepsilon/\varepsilon_0$. Take a finite-dimensional operator T with $\sup_K \|z - Tz\| \le \varepsilon_0$. Then we have $\|z - Tz\| \le \varepsilon_0 \|z\|_E$ if $z \in E$. Thus,

$$\int_E \|z - Tz\|^r \,\mu(dz) \le \varepsilon_0 \int_E \|z\|_E^r \,\mu(dz) < \varepsilon.$$

The assertion is proven. □

This corollary does not extend to arbitrary Banach spaces (see Fonf, Johnson, Pisier, Preiss [**596**]).

7.14(xiii). The Choquet representation

Let K be a compact set in a locally convex space X. Then, for every element b in the closed convex envelope of K, there exists a Radon probability measure μ on K for which b is the barycenter, i.e.,

$$l(b) = \int_K l\,d\mu \quad \text{for all } l \in X^*.$$

See Exercise 7.14.144 for a proof. In this case μ is called a representing measure for b. For convex compact sets, it is useful to have a representing measure concentrated on the set of extreme points. The existence of such measures is established by the following Choquet–Bishop–de Leeuw theorem. Choquet proved this theorem for metrizable K. In this case the set $\mathrm{ext}K$ of extreme points of K is a G_δ-set, in particular, it belongs to $\mathcal{B}(K)$. This is not true in the general case, which leads to modifications in the formulation. See Phelps [**1448**] for a proof.

7.14.47. Theorem. *Let K be a convex compact set in a locally convex space X. Then for every $k \in K$, there exists a Radon probability measure μ on K representing k and vanishing on all Baire sets in $K\backslash\mathrm{ext}K$. If K metrizable, then $\mu(\mathrm{ext}K) = 1$.*

Let K be a convex metrizable compact set in a locally convex space X. Denote by E the set of its extreme points. Let us consider the mapping $\beta\colon \mathcal{P}_r(E) \to K$ that associates to every Radon probability measure μ on E its barycenter $\beta(\mu)$. By the Choquet theorem this mapping is surjective. It is clear that β is affine and continuous if $\mathcal{P}_r(E)$ is equipped with the weak topology, in which $\mathcal{P}_r(E)$ is a Souslin space. Hence there exists a universally measurable mapping $\psi\colon K \to \mathcal{P}_r(E)$ such that k is the barycenter of the measure $\psi(k)$ for all $k \in K$.

There is extensive literature devoted to representation theorems of the Choquet type, see, for example, Alfsen [**35**], Edwards [**518**], Meyer [**1311**], Phelps [**1448**], von Weizsäcker [**1968**], von Weizsäcker, Winkler [**1971**].

7.14(xiv). Convolution

Let us observe that if μ and ν are two measures defined on the σ-algebra $\sigma(X^*)$ in a locally convex space X, then their product $\mu\otimes\nu$ is a measure on $\sigma\bigl((X\times X)^*\bigr)$. It follows by Theorem 7.6.2 that if μ and ν are Radon (or τ-additive) measures, then their product $\mu\otimes\nu$ has a unique extension to a Radon (respectively, τ-additive) measure on $X\times X$. The same is true if X is a Hausdorff topological vector space. Under the product of Radon measures we shall always understand this extension.

7.14.48. Definition. *Let μ and ν be Radon (or τ-additive) measures on a locally convex (or Hausdorff topological vector) space X. Their convolution $\mu*\nu$ is defined as the image of the measure $\mu\otimes\nu$ (extended to a Radon measure as stated above) on the space $X\times X$ under the mapping $(x,y)\mapsto x+y$ from $X\times X$ to X.*

7.14.49. Theorem. *Let μ and ν be Radon measures on a locally convex space X. Then for every Borel set $B \subset X$, the function $x \mapsto \mu(B-x)$ is ν-measurable and one has*

$$\mu * \nu(B) = \int_X \mu(B-x)\,\nu(dx).$$

In addition, $\mu * \nu = \nu * \mu$ and $\widetilde{\mu * \nu} = \widetilde{\mu}\widetilde{\nu}$.

The proof is left as Exercise 7.14.151.

It is clear that by analogy one can define the convolution of two cylindrical quasi-measures.

7.14.50. Proposition. *Let μ and λ be Radon probability measures on a locally convex space X. Suppose that there exists a positive definite function $\varphi\colon\ X^* \to \mathbb{C}$ such that $\widetilde{\lambda} = \varphi\widetilde{\mu}$. Then, there exists a Radon probability measure ν on X with $\widetilde{\nu} = \varphi$. In addition, $\lambda = \nu * \mu$.*

PROOF. It follows by our hypothesis that the restrictions of the function φ to finite-dimensional subspaces are continuous at the origin, hence at any other point. Therefore, φ is the characteristic functional of a nonnegative quasi-measure ν on the algebra of cylindrical sets. It remains to show that the set function ν is tight because then the equality $\widetilde{\lambda} = \widetilde{\nu}\widetilde{\mu}$ will give the equality $\lambda = \nu*\mu$. Let $\varepsilon > 0$ and let S be a compact set with $\mu(X\backslash S)+\lambda(X\backslash S) < \varepsilon/2$. One can assume that $0 \in S$. The set $K := S - S$ is compact and $S \subset K$. Let be C be a cylindrical set with $C \cap K = \varnothing$. The set C has the form $C = P^{-1}(B)$, where $B \in \mathcal{B}(\mathbb{R}^n)$ and $P\colon X \to \mathbb{R}^n$ is a continuous linear mapping. We observe that $B \cap P(K) = \varnothing$. Indeed, if $x \in C$, then $x + h \in C$ for all $h \in \operatorname{Ker} P$. In particular, $B \cap P(S) = \varnothing$, whence $C \cap P^{-1}\bigl(P(S)\bigr) = \varnothing$. The set $C_0 := P^{-1}\bigl(P(S)\bigr)$ is cylindrical, and we have $S \subset C_0$ and

$$1 - \varepsilon/2 \le \lambda(S) \le \lambda(C_0) = \int_X \nu(C_0 - x)\,\mu(dx) \le \int_S \nu(C_0 - x)\,\mu(dx) + \varepsilon/2,$$

whence we obtain the existence of $x_0 \in S$ such that $\nu(C_0 - x_0) \ge 1 - \varepsilon$. In addition, $(C_0 - x_0) \cap C = \varnothing$, since $P(C_0 - x_0) \subset P(S - S)$ because $x_0 \in S$. Thus, $\nu(C) \le \varepsilon$, i.e., the quasi-measure ν is tight. □

For the proof of the following result, see Vakhania, Tarieladze, Chobanyan [**1910**, §VI.3].

7.14.51. Proposition. *Let μ_1 and μ_2 be two nonnegative cylindrical quasi-measures on the algebra of cylindrical sets in a locally convex space X such that μ_1 is symmetric, i.e., $\mu_1(A) = \mu_1(-A)$. If $\mu := \mu_1 * \mu_2$ admits a Radon extension, then both measures μ_1 and μ_2 admit Radon extensions.*

The assumption that μ_1 is symmetric cannot be omitted. Indeed, let l be a discontinuous linear functional on X^* (which exists, for example, if X is an infinite-dimensional Banach space). Then the functionals $\exp(il)$ and $\exp(-il)$ are the Fourier transforms of cylindrical quasi-measures on $Cyl(X, X^*)$ without Radon extensions, but their convolution is the Dirac measure δ. This example is typical: according to Rosiński [**1611**], if μ and ν are nonnegative cylindrical quasi-measures on $Cyl(X, X^*)$ such that $\mu * \nu$ is tight, then there exists an element l in the algebraic dual of X^* with the property that the cylindrical quasi-measures $\mu * \delta_l$ and $\nu * \delta_{-l}$ (where δ_l and δ_{-l} are cylindrical quasi-measures with the Fourier transforms $\exp(il)$ and $\exp(-il)$, respectively) are tight on X (and hence have Radon extensions). These results can be generalized to families of measures as follows (see Vakhania, Tarieladze, Chobanyan [**1910**, Proposition I.4.8]).

7.14.52. Proposition. *Let $\{\mu_\lambda\}$ and $\{\nu_\lambda\}$ be two families of τ-additive probability measures on a Hausdorff topological vector space X. Suppose that the family $\{\mu_\lambda * \nu_\lambda\}$ is uniformly tight, i.e., for every $\varepsilon > 0$, there is a compact set K_ε such that $\mu_\lambda * \nu_\lambda(X\backslash K_\varepsilon) < \varepsilon$ for all λ. Then, there exists a family $\{x_\lambda\}$ of points in X such that $\{\mu_\lambda * \delta_{x_\lambda}\}$ is a uniformly tight family. If, in addition, the measures μ_λ are symmetric, then both families $\{\mu_\lambda\}$ and $\{\nu_\lambda\}$ are uniformly tight.*

In a similar manner one defines the convolution of measures on a topological group. Namely, let $(G, \mathcal{B})$ be a measurable group (i.e., the mappings $x \mapsto -x$ and $(x, y) \mapsto x + y$ are measurable with respect to $\mathcal{B}$ and $\mathcal{B}\otimes\mathcal{B}$, respectively). Let μ and ν be two measures on $\mathcal{B}$. The image of the measure $\mu\otimes\nu$ on $G\times G$ under the mapping $\varrho : (x, y) \mapsto x + y$ is called the convolution of μ and ν and is denoted by $\mu * \nu$.

One can verify that for every $B \in \mathcal{B}$ one has

$$\mu * \nu(B) = \int_G \mu(B - x)\, \nu(dx) = \int_G \nu(-x + B)\, \mu(dx). \tag{7.14.5}$$

If G is commutative, then so is the convolution.

Let G be a topological group. Then, as we have seen above in the case of a locally convex space, G may not be a measurable group with the σ-algebra $\mathcal{B} = \mathcal{B}(G)$. However, if μ and ν are τ-additive or Radon, then $\mu\otimes\nu$ admits a τ-additive (respectively, Radon) extension to $G\times G$. Therefore, in this case the convolution can be defined as the image of this extension under the mapping ϱ, which is continuous. Then (7.14.5) remains valid for $B \in \mathcal{B}(G)$.

Equipped with the operation of convolution, the space of Radon (or τ-additive) probability measures on a topological group G becomes a topological semigroup; its neutral element is Dirac's measure at the neutral element of G.

It is shown in [**1910**, Corollary of Lemma I.4.3] that if $\{\mu_\lambda\}$ and $\{\nu_\lambda\}$ are two families of τ-additive probability measures on a topological group G such that the family $\{\mu_\lambda * \nu_\lambda\}$ is uniformly tight, then there exists a family $\{x_\lambda\}$ of elements of G such that the family $\{\mu_\lambda * \delta_{x_\lambda}\}$ is uniformly tight.

According to [**1910**, Proposition I.4.6], if μ and ν are two τ-additive probability measures on a topological group G, then the support of $\mu * \nu$ coincides with the closure of the set $S_\mu + S_\nu$. This means that the Dirac measures δ_x, $x \in G$, are the only invertible elements in the topological semigroup $\mathcal{P}_\tau(G)$.

Finally, let us make a remark about random vectors. Let X be a locally convex space and let $(\Omega, \mathcal{F}, P)$ be a probability space. A measurable mapping $\xi\colon \Omega \to \bigl(X, \sigma(X)\bigr)$ is called a *random vector* in X. The measure $P_\xi(C) = P\bigl(\xi^{-1}(C)\bigr)$ is called the *distribution* (*law*) of ξ. It is clear that every probability measure on $\sigma(X^*)$ has such a form (with the identity mapping $\xi(x) = x$). If we have a family of probability measures μ_n on X, then there exists a family of independent random vectors ξ_n on a common probability space Ω such that $P_{\xi_n} = \mu_n$ (we take $\Omega = \prod_{n=1}^{\infty} X_n$, $X_n = X$, $P = \bigotimes_{n=1}^{\infty} \mu_n$, $\xi_n(\omega) = \omega_n$); see §10.10(i) about independent random elements. In particular, two random vectors ξ and η with values in X are called independent if

$$P(\xi \in A, \eta \in B) = P(\xi \in A)P(\eta \in B), \quad \forall\, A, B \in \sigma(X^*).$$

Several interesting classes of measures on infinite-dimensional spaces are defined by means of independent random vectors or convolutions. For example, a random vector ξ with values in a locally convex space X is called (see Tortrat [**1888**]) stable of order $\alpha \in (0, 2]$ if for every n, there exists a vector $a_n \in X$ such that, given independent random vectors $\xi_1, \ldots, \xi_n$ with the same distribution μ of the vector ξ, the random vector $n^{-1/\alpha}(\xi_1 + \cdots + \xi_n) - a_n$ has the distribution μ as well. The stable of order 2 random vectors are precisely the Gaussian vectors. The distributions of stable vectors are mixtures of Gaussian measures (see Sztencel [**1820**]). One-dimensional stable distributions are studied in depth in Zolotarev [**2033**].

7.14(xv). Measurable linear functions

Let μ be a Radon probability measure on a locally convex space X with the topological dual X^*. A function $l\colon X \to \mathbb{R}^1$ is called proper linear μ-measurable if it is linear on all of X in the usual sense and is μ-measurable. The collection of all such functions is denoted by $\Lambda(\mu)$. Let $\widetilde{\Lambda}(\mu)$ denote the class of all functions having modifications in the class $\Lambda(\mu)$. However, there is another natural way of defining measurable linear functions. Namely, let $\Lambda_0(\mu)$ be the closure of X^* in $L^0(\mu)$, i.e., $l \in \Lambda_0(\mu)$ if there exists a sequence of functions $l_n \in X^*$ convergent to l in measure. Since $\{l_n\}$ contains an almost everywhere convergent subsequence, we may assume that $l_n \to l$ a.e.

7.14.53. Lemma. *One has $\Lambda_0(\mu) \subset \widetilde{\Lambda}(\mu)$.*

The proof is left as Exercise 7.14.152. There are examples where $\widetilde{\Lambda}(\mu)$ does not coincide with $\Lambda_0(\mu)$ even for symmetric measures μ, see Kanter [**949**], [**950**], Urbanik [**1902**]. One such example is the distribution of the stable of order $\alpha < 2$ random process with independent increments.

7.14(xvi). Convex measures

The convexity of a Radon probability measure μ on a locally convex space X is defined exactly as in $\mathbb{R}^n$. Namely, it is required that

$$\mu_*\bigl(\alpha A + (1-\alpha)B\bigr) \geq \mu(A)^{\alpha}\mu(B)^{1-\alpha}$$

for all nonempty Borel sets A and B and all $\alpha \in [0, 1]$. Convex measures are also called logarithmically concave.

If X is a Souslin space, then the algebraic sum of two Borel sets is Souslin, hence there is no need to consider the inner measure.

7.14.54. Lemma. *A Radon probability measure μ is convex precisely when all its finite-dimensional projections are convex.*

PROOF. If we take for A and B cylindrical sets, then we obtain the convexity of finite-dimensional projections. Conversely, suppose that all such projections are convex and let A and B be Borel sets. Since μ is Radon, it suffices to consider the case where A and B are compact. In that case, since in the weak topology μ is Radon and A and B are compact, given $\varepsilon > 0$ and $\alpha \in (0,1)$, one can find an open cylindrical set C such that

$$\alpha A + (1-\alpha)B \subset C \quad \text{and} \quad \mu(C) < \mu\bigl(\alpha A + (1-\alpha)B\bigr) + \varepsilon.$$

By using the compactness of A and B once again, we find a convex cylindrical neighborhood of the origin V such that $\alpha(A+V) + (1-\alpha)(B+V) \subset C$. As one can easily see, $A+V$ and $B+V$ are cylinders. The required estimate is true for all cylinders by the convexity of the finite-dimensional projections. Hence we obtain

$$\begin{aligned}\mu(C) &\geq \mu\bigl(\alpha(A+V) + (1-\alpha)(B+V)\bigr) \geq \mu(A+V)^{\alpha}\mu(B+V)^{1-\alpha} \\ &\geq \mu(A)^{\alpha}\mu(B)^{1-\alpha},\end{aligned}$$

which yields the required estimate because ε is arbitrary. □

7.14.55. Corollary. (i) *If μ is a convex Radon probability measure on a locally convex space X and $T\colon X \to Y$ is a continuous linear mapping to a locally convex space Y, then the measure $\mu \circ T^{-1}$ is convex.*

(ii) *If μ is a convex Radon probability measure on a locally convex space X and ν is a convex Radon measure on a locally convex space Y, then $\mu \otimes \nu$ is a convex measure on $X \times Y$. In particular, if $X = Y$, then $\mu * \nu$ is a convex measure.*

7.14.56. Theorem. (Borell [**236**]) *Let μ be a convex Radon probability measure on a locally convex space X and let p be a seminorm on X that is measurable with respect to μ. Then, there exists $c > 0$ such that* $\exp(cp)$ *is μ-integrable. In particular, $p \in \mathcal{L}^r(\mu)$ for all $r \in (0,\infty)$.*

7.14.57. Theorem. (Borell [**238**]) *Let μ be a convex Radon probability measure on a locally convex space X, $h \in X$ a nonzero vector and Y a closed hyperplane such that $X = Y \oplus \mathbb{R}^1 h$. Then, on the straight lines $y + \mathbb{R}^1 h$, $y \in Y$, there exist convex probability measures μ^y such that*

$$\mu(B) = \int_Y \mu^y(B)\,\nu(dy), \quad B \in \mathcal{B}(X),$$

where ν is the image of μ under the natural projection $X \to Y$.

Bobkov [**194**] proved that for any convex measure μ, as in the well-known Gaussian case, convergence in measure in the space of polynomials of degree at most d in continuous linear functionals is equivalent to convergence in all $L^p(\mu)$, $p \in [1,+\infty)$. On convex measures, see also Exercise 8.10.115.

7.14(xvii). Pointwise convergence

We know that pointwise convergence of a sequence of measurable functions yields convergence in measure, but this is no longer true for nets. The inverse implication also is false in the general case. Here we consider conditions under which the topology of convergence in measure on a given class of functions coincides with the topology of pointwise convergence. The main results were obtained in Ionescu Tulcea [**862**], [**863**] and reinforced in Edgar [**515**], Fremlin [**621**], Talagrand [**1831**]. A detailed presentation of these results is given in Fremlin [**635**, v. 4].

7.14.58. Proposition. *Let $(X,\mathcal{F},\mu)$ be a complete probability space and let $M\subset\mathcal{L}^\infty(\mu)$ be a set such that if two functions in M are equal a.e., then they coincide everywhere. Then the following assertions are true.*

(i) *If the set M is countably compact in the topology τ_p of pointwise convergence, then for every $x\in X$, the function $f\mapsto f(x)$ is continuous on M with the topology τ_μ of convergence in measure, i.e., the identity mapping $(M,\tau_\mu)\to(M,\tau_p)$ is continuous. In addition, M is closed in $L^0(\mu)$.*

(ii) *If the set M is sequentially compact in the topology τ_p, then the identity mapping $(M,\tau_p)\to(M,\tau_\mu)$ is continuous and is a homeomorphism, and M is a metrizable compact space in these topologies.*

(iii) *If the set M is countably compact in the topology τ_p and is convex and uniformly integrable, then the following topologies coincide on M: τ_p, τ_μ, the weak topology $\sigma(L^1,L^\infty)$, and the norm topology of L^1.*

(iv) *If the set M is compact in the topology τ_p and convex, then the topology τ_p coincides on M with the metrizable topology τ_μ.*

PROOF. (i) Let $x\in X$ and let $f_n\to f$ in measure, $f_n,f\in M$. If $f_n(x)\not\to f(x)$, then, by pointwise boundedness, which follows by countable compactness, there exists a subsequence $\{f_n'\}$ such that $f_n'(x)$ converges, but not to $f(x)$. Let us take an a.e. convergent subsequence $\{f_n''\}$ in $\{f_n'\}$. By countable compactness, $\{f_n''\}$ has a limit point $g\in M$ in the topology τ_p. Then $g(x)=f(x)$ a.e. (at all points x where $\{f_n''(x)\}$ converges), but $g(x)=\lim\limits_{n\to\infty}f_n''(x)\neq f(x)$, contrary to our hypothesis on M. It is easily verified that M is closed in $L^0(\mu)$.

(ii) The set M is compact in the metrizable topology τ_μ because every sequence $\{f_n\}$ in M contains a subsequence that is pointwise convergent to a function from M, hence in measure. Now (i) applies, since M is countably compact by sequential compactness, and the continuous images of compact sets are compact.

(iii) We observe that M is closed in $L^1(\mu)$ by virtue of (i). By convexity M is closed in the weak topology. On account of uniform integrability this yields the weak compactness of M in $L^1(\mu)$. Let us show that the mapping $\bigl(M,\sigma(L^1,L^\infty)\bigr)\to(M,\tau_p)$ is continuous, i.e., for every fixed $x\in X$, the function $f\mapsto f(x)$ is continuous on $\bigl(M,\sigma(L^1,L^\infty)\bigr)$. To this end, it suffices to verify that for every real number c, the sets $\{f\in M\colon\ f(x)\le c\}$ and

$\{f \in M\colon\ f(x) \geq c\}$ are closed in the topology $\sigma(L^1, L^\infty)$. Since these sets are closed in the topology of pointwise convergence, it follows by (i) that they are closed in the topology τ_μ, hence in the norm topology on M. Since M is closed in $L^1(\mu)$, both sets are closed subsets of $L^1(\mu)$, which by convexity yields that they are weakly closed. Thus, the mapping $\bigl(M, \sigma(L^1, L^\infty)\bigr) \to (M, \tau_p)$ is continuous, hence by the weak compactness of M it is a homeomorphism. By the uniform integrability of M the norm topology coincides on M with τ_μ. The already-established weak compactness of M by the Eberlein–Šmulian theorem gives weak sequential compactness, which means (by the equality of τ_p and $\sigma(L^1, L^\infty)$ on the set M) sequential compactness in τ_p. According to (ii) all the indicated topologies coincide on M.

(iv) According to assertion (i) the identity mapping $(M, \tau_\mu) \to (M, \tau_p)$ is continuous and the set M is closed in $L^0(\mu)$. Hence it suffices to prove the compactness of M in the metrizable topology τ_μ. Suppose we are given a sequence $\{f_n\} \subset M$. Let us show that it contains a convergent subsequence. The function $g(x) := 1 + \sup_n |f_n(x)|$ is finite and measurable, since the sequence $\{f_n(x)\}$ is bounded for every x by compactness in the topology τ_p. The measure $\nu := g^{-2} \cdot \mu$ is finite and equivalent to the measure μ. Hence it also satisfies our principal condition on M. Let M_0 be the closed convex envelope of $\{f_n\}$ in the topology τ_p. Then M_0 is a convex compact set in this topology. For every function $f \in M_0$, we have $|f(x)| \leq g(x)$, $x \in X$, since this inequality is fulfilled for all f_n and is preserved by convex combinations and the pointwise limits. Therefore, the integral of $|f|^2$ with respect to the measure ν does not exceed 1 for all $f \in M_0$. Thus, the set M_0 is uniformly integrable with respect to the measure ν. By assertion (iii) the topology τ_p coincides on M_0 with τ_ν and M_0 is compact in these topologies. Hence $\{f_n\}$ contains a subsequence $\{f_{n_k}\}$ convergent in measure μ. □

A typical example where condition (ii) is fulfilled is the case where M is a set of continuous functions on a topological space X such that M is sequentially compact in the topology of pointwise convergence and X is the support of a Radon measure μ.

7.14.59. Corollary. *Let X be a normed space and let μ be a probability measure on $\sigma(X^*)$ such that $\mu\bigl(x\colon\ l(x) = 0\bigr) < 1$ for every nonzero $l \in X^*$. Then X is separable, and on the closed unit ball of X^* the weak* topology coincides with the topology of convergence in measure μ.*

Proof. The set $M := \{f \in X^*\colon\ \|f\| \leq 1\}$ is compact in the weak* topology by the Banach–Alaoglu theorem and is convex. The metrizability of M in the weak* topology yields the separability of X. □

7.14.60. Example. (i) Let X be a normed space, let μ be a probability measure on $\sigma(X^*)$, and let V_μ be the intersection of all closed linear subspaces of outer measure 1. Suppose that $\mu^*(V_\mu) = 1$. Then μ has a τ-additive extension in the norm topology (Radon if X is Banach).

(ii) Every τ-additive in the weak topology (in particular, every Radon in the weak topology) probability measure on a Banach space has a Radon extension in the norm topology.

(iii) If X is a reflexive Banach space, then every measure on $\sigma(X^*)$ has a Radon extension with respect to the norm topology.

PROOF. (i) Let us consider the restriction of μ to V_μ (in the sense of Definition 1.12.11). Let f be a nonzero element in V_μ^*. We extend f to a functional $f_0 \in X^*$. If $\mu(f = 0) = 1$, then $\mu(f_0 = 0) = 1$. This contradicts the choice of V_μ, since $V_\mu \cap f_0^{-1}(0)$ is a proper closed subspace in V_μ. Then V_μ is separable by the above corollary, whence the claim follows. (ii) By the τ-additivity in the weak topology, the measure μ has the topological support S in the weak topology, whence $\mu(V_\mu) = 1$, since $S \subset V_\mu$. (iii) By the weak compactness of balls in reflexive spaces μ is tight in the weak topology, hence is τ-additive. □

The reader is warned that this example does not extend to locally convex spaces (Exercise 7.14.149): there exists a measure that is Radon in the weak topology, but is not tight in the original topology.

The proofs of the following interesting and deep facts can be found in Fremlin [**621**], [**635**, §463].

7.14.61. Theorem. *Let $(X, \mathcal{A}, \mu)$ be a complete probability space with a perfect measure μ and let a set $M \subset \mathcal{L}^0(\mu)$ be countably compact in the topology of pointwise convergence. Then every sequence in M has a subsequence convergent a.e. and M is compact in the topology of convergence in measure. If every two distinct* (*i.e., not identically equal*) *functions in M differ on a set of positive measure, then the topology of pointwise convergence and the topology of convergence in measure coincide on the set M, which turns out to be a metrizable compact set.*

It is unclear how essential the assumption of perfectness of the measure is. Talagrand [**1834**] showed that if the set M is compact in the topology τ_p and a.e. equal functions in M are equal pointwise, then under Martin's axiom the topologies τ_p and τ_μ coincide on M.

7.14.62. Theorem. *Let $(X, \mathcal{A}, \mu)$ be a complete probability space and let an infinite set $M \subset \mathcal{L}^0(\mu)$ be compact in the topology of pointwise convergence. Suppose that every two different functions in M differ on a set of positive measure. Then M contains a pointwise convergent subsequence.*

Let us mention the following Fremlin alternative (see Fremlin [**621**], [**635**, §463H], and also Talagrand [**1834**]).

7.14.63. Theorem. *Let $(X, \mathcal{A}, \mu)$ be a complete probability space with a perfect measure μ and let f_n, $n \in \mathbb{N}$, be μ-measurable functions. Then, either $\{f_n\}$ contains an a.e. convergent subsequence or $\{f_n\}$ contains a subsequence for which no μ-measurable function is a limit point in the topology of pointwise convergence.*

Talagrand [**1832**] obtained sufficient conditions (including the continuum hypothesis or Martin's axiom) for the closed convex envelope of a set of measurable functions in the topology of pointwise convergence to consist of measurable functions.

7.14(xviii). Infinite Radon measures

All Radon measures discussed in this book are finite by definition. However, in some applications it is useful to enlarge this concept (which has already been done in §7.11). Obvious examples are Lebesgue measure on $\mathbb{R}^n$, Hausdorff measures, and Haar measures on noncompact groups. Yet, the first of them is σ-finite and there is no need to develop a special terminology to deal with it (although the classical work of Radon was concerned with infinite, in general, measures on $\mathbb{R}^n$). But Hausdorff and Haar measures are not always σ-finite. Thus, what should one understand by a "Radon measure with values in $[0,+\infty]$"? Different definitions are possible, leading to the same object in the case of a finite measure. The following definition appears to be reasonable (see Fremlin [**619**], [**635**]).

7.14.64. Definition. *Let X be a Hausdorff space. A measure μ with values in $[0,+\infty]$ defined on a σ-algebra $\mathfrak{S}$ of subsets of X is called a Radon measure with values in $[0,+\infty]$ if μ is complete, locally determined (see Exercise 1.12.135), all open sets belong to $\mathfrak{S}$, every point has a neighborhood of finite measure, and for all $E \in \mathfrak{S}$ one has*

$$\mu(E) = \sup\{\mu(K)\colon\ K \subset E,\ K \textit{ is compact}\}.$$

In the case of a finite measure, this definition corresponds to the completion of a Radon (in our usual meaning) measure on the Borel σ-algebra. According to another definition frequently used in the literature, a Radon measure with values in $[0,+\infty]$ is defined on the Borel σ-algebra, every point has a neighborhood of finite measure, and one has the inner compact regularity condition from the above definition. Such a measure extends uniquely to a Radon measure in the sense of the above definition (see [**619**]). The product of two infinite Radon measures extends uniquely to an infinite Radon measure (see [**619**]). If X is locally compact, then every positive linear functional on $C_0(X)$ is given as the integral with respect to a Radon measure with values in $[0,+\infty]$ (Theorem 7.11.3). An infinite Radon measure may not be outer regular (i.e., may not satisfy the condition $\mu(B) = \inf \mu(U)$, where $U \supset B$ is open).

7.14.65. Example. Let us consider the metric space $X = \Omega \times \mathbb{R}^1$, where Ω is the real line with the discrete metric and $\mathbb{R}^1$ is equipped with the standard metric. Then X with the product topology is locally compact and $\mathcal{B}(X) = \mathcal{B}(\Omega)\otimes\mathcal{B}(\mathbb{R}^1)$. For every $B \in \mathcal{B}(X)$, we set

$$\mu(B) := \sum_{\omega\in\Omega} \lambda(B_\omega),$$

where $B_\omega = \{t\colon (\omega, t) \in B\}$ and λ is Lebesgue measure, i.e., μ is the product of the counting measure on Ω and Lebesgue measure. Then $\mu(\Omega\times\{0\}) = 0$, but $\mu(U) = +\infty$ for every open set $U \supset \Omega\times\{0\}$. It is readily seen that μ is inner compact regular. Indeed, given $B \in \mathcal{B}(X)$ and $c < \mu(B)$, we can find points $\omega_1, \ldots, \omega_n \in \Omega$ and compact sets $K_i \subset B_{\omega_i}$ such that the μ-measure of the compact set $\bigcup_{i=1}^n \{\omega_i\}\times K_i$ is greater than c.

A more general example: a non-σ-finite inner compact regular Haar measure (see §9.11). However, there exist σ-finite measures that are inner compact regular but not outer regular; see Exercise 7.14.160.

A system $\mathfrak{C}$ of nonempty pairwise disjoint compact sets in a space X is called a concassage for a Radon measure μ on X with values in $[0, +\infty]$ if the intersections of the sets in $\mathfrak{C}$ with open sets are either empty or have positive measures, and for every set E in the domain of definition of μ one has

$$\mu(E) = \sum_{C\in\mathfrak{C}} \mu(C \cap E).$$

Every Radon measure with values in $[0, +\infty]$ possesses a concassage (this is readily verified by Zorn's lemma, see details in Gardner, Pfeffer [**666**, Proposition 12.10]). Any saturated (see Chapter 1) Radon measure with values in $[0, +\infty]$ is decomposable, hence is Maharam (it is easy to verify that a concassage of such a measure μ gives its decomposition, see Gardner, Pfeffer [**667**]), however, neither completeness nor the property to be saturated can be omitted (see Fremlin [**620**]).

It is shown in Bauer [**132**] that in the situation of Theorem 7.8.7 in the Daniell–Stone approach, there exists a locally compact space T with a Radon measure ν with values in $[0, +\infty]$ such that Ω is embedded into T as a dense subset, the measure ν naturally extends μ, every function $f \in \mathcal{F}$ extends to a continuous function f' on T decreasing to zero at the infinity, and such extensions separate the points in T and do not vanish at any point in T, provided the latter two properties hold for $\mathcal{F}$. On infinite Radon measures, see also Fremlin [**635**], Gardner, Pfeffer [**666**], [**667**], Gruenhage, Pfeffer [**747**].

Exercises

7.14.66.° Show that every regular τ_0-additive Borel measure is τ-additive.

Hint: given a net of increasing open sets U_α whose union is U, we fix $\varepsilon > 0$, take a closed set $F \subset U$ with $|\mu|(U\backslash F) < \varepsilon$, and consider the sets $U_\alpha \cup (X\backslash F)$ that are open and increase to X.

7.14.67.° Let X be an uncountable space, let $\mathcal{A}$ be the σ-algebra in X consisting of finite and countable sets and their complements, and let the measure μ equal 0 on all countable sets and 1 on their complements. Show that μ is perfect. Deduce that any measure on $\mathcal{A}$ is perfect.

Hint: use that every $\mathcal{A}$-measurable function assumes at most countably many values; any measure on $\mathcal{A}$ has at most countably many points of positive measure.

7.14.68. (Adamski [**6**]) Construct an example of a non-regular τ-additive measure on some non-regular second countable space (in particular, assertion (ii) of Proposition 7.2.2 may be false for non-regular spaces).

HINT: let S be a subset of $[0,1]$ with $\lambda_*(S) = 0 < \lambda^*(S)$, where λ is Lebesgue measure. Let X be $[0,1]$ with the topology generated by the standard topology of $[0,1]$ together with the set S (the open sets in X have the form $[0,1]\cap\bigl(U\cup(V\cap S)\bigr)$, where U and V are open in $\mathbb{R}^1$). It is clear that the space X satisfies the second axiom of countability, but is not regular. Let μ be the image of the restriction λ_S of λ to S (see Definition 1.12.11) under the natural embedding $S \to X$ (which is continuous). The measure μ is τ-additive by the last assertion in Proposition 7.2.2. But it is not regular, since $\mu(S) > 0$, whereas $\mu(F) = 0$ for every set $F \subset S$ that is closed in X, since such a set is compact in the standard topology of $[0,1]$, hence $\lambda(F) = 0$ due to our choice of S.

7.14.69. (i) (Wheeler [**1978**], [**1979**]) There exist a completely regular space X and a Baire probability measure on X that has no countably additive extensions to the Borel σ-algebra.

(ii) (Ohta, Tamano [**1394**]) There exists a locally compact space X with the property indicated in (i). In addition, there exists a countably paracompact space with such a property.

HINT: for constructing an example in (i) it suffices to have a Baire probability measure μ on X that assumes only the values 0 and 1, has a discrete Baire set T of full measure and cardinality of the continuum $\mathfrak{c}$, but vanishes on all singletons. A Borel extension of μ would be a measure defined on all subsets in T and vanishing on all singletons (which contradicts the fact that $\mathfrak{c}$ is not two-valued measurable). Concrete examples are discussed in the cited papers. It is also possible to replace in Example 7.3.9 the set I by a set $I_0 \subset I$ of the least cardinality among all sets of outer measure 1 and equip I_0 with the restriction of Lebesgue measure and the Sorgenfrey topology.

7.14.70. Prove Theorem 7.14.2.

HINT: let $\mathcal{A}_n$ be the σ-algebra of all cylindrical sets with bases in $\mathcal{B}\bigl(\prod_{i=1}^n X_i\bigr)$. The union of all $\mathcal{A}_n$ is an algebra; μ extends to this algebra as a countably additive measure, which is verified similarly to the proof of the theorem on countable products of measures. The τ_0-additivity of μ follows from this. To this end, a given net of open cylinders is split into parts containing the cylinders with bases in $\prod_{i=1}^n X_i$. See also Ressel [**1555**], Amemiya, Okada, Okazaki [**46**].

7.14.71.° Suppose that a compact set K in a completely regular space is covered by two open sets U_1 and U_2. Show that there exist continuous nonnegative functions f_1 and f_2 with the compact supports $K_1 \subset U_1$ and $K_2 \subset U_2$, respectively, such that $f_1 + f_2 = 1$ on K.

7.14.72.° Let μ be a nonnegative Baire measure on a normal space X. Prove that for every closed set $C \subset X$ and every $\varepsilon > 0$, there exists a functionally closed set Z such that $C \subset Z$ and $\mu(Z) \le \mu^*(C) + \varepsilon$.

HINT: there exists a functionally open set U such that $C \subset U$ and $\mu(U) \le \mu^*(C)+\varepsilon$; since X is normal, there exists a functionally closed set Z with $C \subset Z \subset U$.

7.14.73.° Let f_n be measurable mappings from a space with a finite measure μ to a separable metric space (Y, ϱ_Y) convergent in measure to a measurable mapping f, i.e., for all $c > 0$ we have $\lim\limits_{n\to\infty} \mu\bigl(\varrho_Y(f_n, f) > c\bigr) = 0$. Show that there exists a subsequence $\{f_{n_i}\}$ that converges a.e.

HINT: consider the completion $\overline{Y}$ of Y and use the reasoning from the scalar case.

7.14.74. Let f_n be measurable mappings from a probability space (X, μ) to a separable metric space S convergent in measure to a mapping f. Let $\Psi\colon\ S \to M$ be a continuous mapping with values in a metric space (M, d). Show that the mappings $\Psi \circ f_n$ converge in measure to $\Psi \circ f$.

HINT: show that the integrals of $\min\bigl(1, d(\Psi\circ f_n, \Psi\circ f)\bigr)$ converge to zero; to this end, use that any subsequence in $\{f_n\}$ contains a further subsequence convergent to f almost everywhere.

7.14.75. Let Y be a metric space and let a function $f\colon\ [0, 1] \to Y$ be measurable with respect to Lebesgue measure. Prove that there exists a separable subspace $Y_0 \subset Y$ such that $f(x) \in Y_0$ for a.e. x and deduce that for every $\varepsilon > 0$, there exists a compact set K_ε of measure at least $1 - \varepsilon$ on which f is continuous.

HINT: apply Theorem 1.12.19.

7.14.76. (i) Let $X = [0, 1]$ be equipped with the standard topology and Lebesgue measure μ and let $Y = [0, 1]$ be equipped with the topology generated by all intervals $[a, b) \cap [0, 1]$, $a < b$ (i.e., the Sorgenfrey interval with the added point 1 as an open set, see Example 7.2.4). Show that the identity mapping $f\colon\ X \to Y$ is Borel, but its restriction to any uncountable set is not continuous.

(ii) Construct an example of a Borel mapping from the interval $[0, 1]$ with the standard topology and Lebesgue measure to a compact space such that the analog of Lusin's theorem fails for it.

(iii) Let μ be the measure on $(0, 1) \times \{0, 1\}$ that is the product of Lebesgue measure and the measure on $\{0, 1\}$ assigning $1/2$ to the points 0 and 1. Let X be the space "two arrows" from Example 6.1.20 equipped with its natural normalized Lebesgue measure λ. Consider the natural mapping f from $(0, 1) \times \{0, 1\}$ to X. Show that f is measurable and $\mu \circ f^{-1} = \lambda$, but there is no compact set of positive μ-measure on which f is continuous.

(iv) Let $X = [0, 1]^{\mathfrak{c}}$ be the product of the continuum of intervals and let X be equipped with the Radon measure μ that is the extension of the product of the continuum of Lebesgue measures. Let $f\colon\ X \to X$ be defined as follows: $f(x)(s) = x(s)$ if $0 < x(s) < 1$, $f(x)(s) = 1 - x(s)$ if $x(s) = 0$ or $x(s) = 1$. Show that f is measurable with respect to μ, but is not almost continuous.

HINT: (i) is verified directly; (ii) consider the compactification of Y from (i); (iii) any continuous image of a metrizable compact space is metrizable, but any metrizable set in X is at most countable; (iv) see Fremlin [**625**, example 3G].

7.14.77. Construct an example of a Borel probability measure ν on a compact space X and a Borel function $f\colon\ X \to \mathbb{R}$ such that for every continuous function $g\colon\ X \to \mathbb{R}$, one has $\nu\bigl(x\colon\ f(x) \neq g(x)\bigr) \geq 1/2$.

HINT: let μ be the Dieudonné measure from Example 7.1.3, let $\nu = (\mu + \delta_{\omega_1})/2$ and $f = I_{\{\omega_1\}}$; use Exercise 6.10.75; see also Wise, Hall [**1993**, Example 4.48].

7.14.78. (i) Show that $C_b(X)$ is dense in $L^1(\mu)$ for every Radon measure μ on a completely regular space X.

(ii) Construct an example of a Borel probability measure ν on a compact space X such that $C(X)$ is not dense in $L^1(\nu)$.

(iii) (Hart, Kunen [**789**].) There is a Radon probability measure μ on a compact space such that $L^2(\mu)$ has no orthonormal basis consisting of continuous functions.

HINT: (i) apply Lusin's theorem; (ii) use the previous exercise.

7.14.79.° Let μ and ν be two Radon measures on a topological space X and let $\mathcal{F}$ be a family of bounded continuous 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. Suppose that $1 \in \mathcal{F}$ and $\mathcal{F}$ separates the points in X. Show that $\mu = \nu$.

HINT: the mapping $T\colon x \mapsto \bigl(f(x)\bigr)_{f\in\mathcal{F}}$ from X to the compact space Y that is the product of the closed intervals $I_f = [\inf_x f(x), \sup_x f(x)]$ is continuous, the measures $\mu' := \mu \circ T^{-1}$ and $\nu' := \nu \circ T^{-1}$ on Y are Radon and assign equal integrals to any polynomial in finitely many coordinate functions on Y. By the Stone–Weierstrass theorem $\mu' = \nu'$. Since T is injective, we obtain $\mu(K) = \nu(K)$ for every compact set K in X, hence $\mu = \nu$.

7.14.80.° Let $(X, \mathcal{A}, \mu)$ be a measure space with a perfect measure μ, let $(Y, \mathcal{B})$ be a measurable space such that $\mathcal{B}$ is countably generated and countably separated, and let $f\colon X \to Y$ be a μ-measurable mapping. Prove that a real function g on Y is measurable with respect to the measure $\mu \circ f^{-1}$ (i.e., $\mathcal{B}_{\mu\circ f^{-1}}$-measurable) precisely when the function $g \circ f$ is measurable with respect to μ.

HINT: the $\mu\circ f^{-1}$-measurability of g yields the μ-measurability of $g\circ f$. In order to prove the converse, recall that $(Y, \mathcal{B})$ is isomorphic to a subset of an interval with the Borel σ-algebra. Hence $\mathcal{B}_{\mu\circ f^{-1}} = \{B \subset Y\colon f^{-1} \in \mathcal{A}_\mu\}$. Now the μ-measurability of $g \circ f$ yields the $\mu \circ f^{-1}$-measurability of the sets $\{g \le c\}$.

7.14.81.° Give an example of a probability measure μ on a σ-algebra $\mathcal{A}$ in a space X and a mapping $F\colon X \to Y$ with values in a compact space Y such that $F^{-1}(B) \in \mathcal{A}_\mu$ for all $B \in \mathcal{B}(Y)$, but there is no mapping G which μ-a.e. equals F and $G^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}(Y)$ (in other words, F is measurable with respect to completion of μ, but is not equivalent to any $\mathcal{A}$-measurable mapping).

HINT: consider $X = [0,1]^{\mathfrak{c}}$, $\mathcal{A} = \mathcal{B}a([0,1]^{\mathfrak{c}})$, take for μ on $\mathcal{A}$ the product of the continuum of Lebesgue measures, and let F be the identity mapping from X to X. The measurability of F follows by Theorem 7.14.3. If there exists a $\bigl(\mathcal{B}a(X), \mathcal{B}(X)\bigr)$-measurable modification G of the mapping F, then there is a set $A \in \mathcal{B}a(X)$ of full measure, dependent only on countably many coordinates x_{t_i}, such that $A \cap B \in \mathcal{B}a(X)$ for all $B \in \mathcal{B}(X)$. This leads to a contradiction if we take for B the set $\bigl\{x\colon x_t = 0, \forall\, t \notin \{t_i\}\bigr\}$.

7.14.82.° Give an example of a regular Borel probability measure μ on a locally compact Hausdorff space that has no support, in particular, is not τ-additive.

HINT: consider the measure constructed in Example 7.1.3 on X_0.

7.14.83.° Let $X = [0,1]$ be equipped with the following topology: all singletons in $(0,1]$ are open and all sets of the form $[0,1]\backslash\{x_1, \ldots, x_n\}$, where $x_i \in (0,1]$, are open. Verify that the generated topology is Hausdorff and is X compact in this topology. Show that X cannot be the support of a Radon measure.

7.14.84. Let X be a Hausdorff space, let $\mathcal{A}$ be an algebra of subsets in X, and let m be a nonnegative finitely additive set function on $\mathcal{A}$ such that

$$m(A) = \sup\{m(Z)\colon\ Z \subset A, Z \in \mathcal{A},\ Z \text{ is closed}\}$$

for all $A \in \mathcal{A}$ and

$$m(X) = \sup_{K \in \mathcal{K}} \bigl\{\inf\{m(E)\colon\ E \in \mathcal{A},\, K \subset E\}\bigr\},$$

where $\mathcal{K}$ is the class of all compact sets. Prove that m extends to a Radon measure on X.

HINT: see Fremlin [**635**, §416O].

7.14.85. Prove Theorem 7.14.30.

7.14.86. Prove that the algebra $\mathfrak{A}(X)$ generated by all functionally closed subsets of a topological space X consists of finite unions of the form $\bigcup_{i=1}^{n}(F_i \backslash F_i')$, where F_i, F_i' are functionally closed and $F_i' \subset F_i$. Prove the analogous assertion for the algebra generated by all closed sets.

HINT: see Exercise 1.12.51.

7.14.87. Let X be a completely regular space, let βX be its Stone–Čech compactification, and let L be a continuous linear functional on the space $C_b(X)$. Let the functional $L'(g) = L(g \circ j)$ on $C_b(\beta X)$, where $j\colon X \to \beta X$ is the canonical embedding, be represented by a Baire measure ν on βX. Prove that L is represented by some Baire measure μ on X precisely when $|\nu|^*(X) = |\nu|(\beta X)$, where the outer measure is defined by means of $\mathcal{B}a(\beta X)$; in addition, ν extends μ to βX. The analogous assertion is true for τ-additive measures if the outer measure is defined by means of $\mathcal{B}(\beta X)$.

HINT: if $|\nu|^*(X) = |\nu|(\beta X)$ in the case of the Baire σ-algebra, then ν can be restricted to X by means of the standard construction of restricting to a set of full outer measure, and the induced σ-algebra coincides with $\mathcal{B}a(X)$. The obtained measure μ on X represents the functional L, since any function $f \in C_b(X)$ extends uniquely to a function $\widehat{f} \in C_b(\beta X)$, whence one has

$$L(f) = L'(\widehat{f}) = \int_{\beta X} \widehat{f}\, d\nu = \int_X f\, d\mu.$$

7.14.88. Prove that every additive regular set function m on the algebra $\mathfrak{A}(X)$ generated by all functionally closed subsets in a topological space X (see the definition in §7.9) is the difference of two nonnegative additive regular set functions defined before Theorem 7.9.1.

7.14.89. Let X and Y be topological spaces and let μ be a Borel probability measure on Y. Prove that given a continuous mapping $f\colon X \to Y$, the equality $\kappa(A) = \mu^*\bigl(f(A)\bigr)$ defines a Choquet capacity on X.

7.14.90. (Shortt [**1703**]) We shall say that a separable metric space is universally measurable if it is measurable with respect to every Borel measure on its completion. Suppose that a set X is equipped with two metrics d_1 and d_2 with respect to which X is separable and the corresponding Borel σ-algebras coincide. (i) Prove that X is universally measurable with the metric d_1 precisely when it is universally measurable with the metric d_2. (ii) Deduce from (i) and Theorem 7.5.7 that a separable metric space X is universally measurable precisely when every Borel probability measure on X is perfect.

7.14.91. (Sazonov [**1656**]) Prove without the continuum hypothesis that on the set of all subsets in $[0, 1]$ there is no perfect probability measure vanishing on all singletons.

HINT: let μ be such a measure and let $F(x) = \mu\bigl([0,x)\bigr)$; then the measure $\nu = \mu \circ F^{-1}$ is defined on the set of all subsets of the interval and extends Lebesgue measure; in addition, ν is perfect being the image of a perfect measure; we take a set A with outer Lebesgue measure 1 and inner Lebesgue measure 0; let $\nu(A) > 0$ (otherwise $\nu([0,1]\backslash A) = 1$ and we can deal with $[0,1]\backslash A$); the restriction of ν to A is a perfect measure, which is impossible (it suffices to take $f(x) = x$). See a generalization in Pachl [**1415**].

7.14.92. (Sazonov [**1656**]) Let X be a metric space containing no system of disjoint nonempty open sets of cardinality greater than that of the continuum. Prove that a Borel measure on X is perfect precisely when it is tight. See a generalization in Pachl [**1415**].

7.14.93. (Zink [**2031**], Saks, Sierpiński [**1643**] for $Y = \mathbb{R}$) Let (X, S, μ) be a probability space and let (Y, d) be a separable metric space. Let $f\colon X \to Y$ be an arbitrary mapping. Prove that for every $\varepsilon > 0$, there exists a $\bigl(\mathcal{B}(Y), S\bigr)$-measurable mapping $g\colon X \to Y$ such that $d\bigl(f(x), g(x)\bigr) < \varepsilon$ for every x, with the exception of points of a set of inner measure zero.

7.14.94. Let X be an uncountable Souslin space. Prove that there exists a family of mutually singular atomless Radon probability measures on X having the cardinality of the continuum.

HINT: find in X a collection of cardinality of the continuum of disjoint Borel sets of cardinality of the continuum.

7.14.95. (Plebanek [**1465**]) Let $\mathcal{K}$ be some compact class of subsets of a set X such that to every $K \in \mathcal{K}$ there corresponds a number r_K. Denote by $\mathcal{A}_{\mathcal{K}}$ the algebra generated by $\mathcal{K}$. Suppose that for every finite collection $K_1, \ldots, K_n \in \mathcal{K}$, there is an additive set function $\mu_{K_1,\ldots,K_n}\colon \mathcal{A}_{\mathcal{K}} \to [0,1]$ with $\mu_{K_1,\ldots,K_n}(K_i) \geq r_{K_i}$, $i = 1, \ldots, n$. Then there exists a probability measure μ on $\sigma(\mathcal{K})$ with $\mu(K) \geq r_K$ for all $K \in \mathcal{K}$.

7.14.96.$^\circ$ Let μ be an atomless Radon measure on a metric space X. Prove that for every $\varepsilon > 0$, there exists $\delta > 0$ such that $\mu(B) < \varepsilon$ for every Borel set B of diameter less than δ.

HINT: it suffices to consider the restriction of μ to a compact set K of a sufficiently large measure; for every point $x \in K$, there exists $r(x) > 0$ such that one has $\mu\bigl(K\bigl(x, 2r(x)\bigr)\bigr) < \varepsilon/2$; hence there is a finite cover of K by the balls $K(x_i, r)$ of some radius $r > 0$ with $\mu\bigl(K(x_i, 2r)\bigr) < \varepsilon/2$; let $\delta = r/2$.

7.14.97. (Davies, Schuss [**418**]) Let μ be a Radon probability measure on a topological space X, let f be a μ-integrable function, and let J be its integral. Prove that for every $\varepsilon > 0$, every point $x \in X$ can be associated with an open set $G(x)$ containing x such that given measurable sets B_i having pairwise intersections of measure zero and covering X up to a measure zero set and any given points $x_i \in B_i$ satisfying the condition $B_i \subset G(x_i)$, one has $\Bigl|\sum_{i=1}^\infty f(x_i)\mu(B_i) - J\Bigr| < \varepsilon$.

7.14.98. Let M be a metric space, let $\mathcal{F}$ be some σ-algebra in M containing all singletons, and let μ be a probability measure on $\mathcal{F}$. Prove that the following conditions are equivalent: (1) $\mu(p) = 0$ for all $p \in M$; (2) for every $p \in M$ and $\varepsilon > 0$, there exists $r > 0$ such that if a set E in $\mathcal{F}$ is contained in the ball of radius r centered at p, then $\mu(E) < \varepsilon$.

HINT: see Hahn [**770**, p. 409] or Sierpiński [**1717**].

7.14.99. (Rao, Rao [**1536**]) Show that on the Borel σ-algebra of the space $[0,\omega_1)$, where ω_1 is the first uncountable ordinal, there exists no atomless countably additive probability measure (see generalizations in Mauldin [**1275**]).

7.14.100. (Marczewski, Ryll-Nardzewski [**1259**]) (i) Let μ be a countably additive probability measure on an algebra $\mathcal{A}$ of subsets of a space X and let ν be a countably additive probability measure on an algebra $\mathcal{B}$ of subsets of a space Y possessing a compact approximating class. Suppose that on the algebra $\mathcal{E}$ generated by the rectangles $A\times B$, where $A\in\mathcal{A}$, $B\in\mathcal{B}$, one has a nonnegative additive set function σ such that $\sigma(A\times Y)=\mu(A)$ for all $A\in\mathcal{A}$ and $\sigma(X\times B)=\nu(B)$ for all $B\in\mathcal{B}$. Prove that σ is countably additive.

(ii) Construct an example showing that assertion (i) may be false if one does not require the existence of a compact approximating class for at least one of the measures μ or ν.

7.14.101. Construct an example of a function f on $[0,1]^\infty$ that is constant in every variable if the remaining variables are fixed, but is not measurable with respect to the countable product of Lebesgue measures.

Hint: see Marczewski, Ryll-Nardzewski [**1257**].

7.14.102°. (Ursell [**1904**]) Let μ be a finite nonnegative measure on a space X. (i) Let a function $f\colon X\times[0,1]\to\mathbb{R}^1$ be such that for every fixed t, the function $x\mapsto f(x,t)$ is μ-measurable, and for μ-a.e. x, the function $t\mapsto f(x,t)$ is increasing. Prove that the function f is measurable with respect to the measure $\mu\otimes\lambda$, where λ is Lebesgue measure.

(ii) Let $E\subset X\times[0,1]$ be such that the sections $E_t:=\{x\colon (x,t)\in E\}$ are μ-measurable and $E_t\subset E_s$ if $t<s$. Prove that E is measurable with respect to $\mu\otimes\lambda$.

(iii) (S. Hartman) Let A be a non-Borel set on the line $\{(x,y)\colon x+y=0\}$ in $\mathbb{R}^2$ and let E be the union of A and the open half-plane $\{(x,y)\colon x+y>0\}$. Show that the function I_E has the following properties: it is nondecreasing and one-sided continuous in every variable separately, but is not Borel in both variables. Hence in assertion (i) one cannot assert the Borel measurability of f even if $X=[0,1]$ with the Borel σ-algebra.

Hint: (i) follows by considering the approximations

$$f_n(x,t)=f(x,j2^{-n})\quad \text{if } t\in\bigl[j2^{-n},(j+1)2^{-n}\bigr),\ j=0,\ldots,2^n-1,$$

$$g_n(x,t)=f(x,(j+1)2^{-n})\quad \text{if } t\in\bigl[j2^{-n},(j+1)2^{-n}\bigr),\ j=0,\ldots,2^n,$$

with $f_n(x,1)=g_n(x,1)=f(x,1)$. One has $f_n\le f\le g_n$. The set Ω of all points where both sequences $\{f_n(x,t)\}$ and $\{g_n(x,t)\}$ converge to a common limit $\varphi(x,t)$ is $\mu\otimes\lambda$-measurable. It follows from our hypotheses that, for μ-a.e. x, one has $\varphi(x,t)=f(x,t)$, hence the section Ω_x may differ from $[0,1]$ only in an at most countable set. By Fubini's theorem $\mu\otimes\lambda(\Omega)=1$, i.e., $\varphi(x,t)=f(x,t)$ for $\mu\otimes\lambda$-a.e. (x,t). Clearly, φ is $\mu\otimes\lambda$-measurable. Assertion (ii) follows from (i). Assertion (iii) is readily verified.

7.14.103. (Marczewski, Ryll-Nardzewski [**1257**]) Let $(X,\mathcal{F})$ be a measurable space, T a separable metric space, Y a metric space, and let a mapping $f\colon X\times T\to Y$ be such that for every fixed t, the mapping $x\mapsto f(x,t)$ is measurable with respect to $\mathcal{F}$, provided that Y is equipped with the Borel σ-algebra.

(i) Let $T = \mathbb{R}^1$. Suppose additionally that for every x, the mapping $t \mapsto f(x,t)$ is right continuous. Then f is $\mathcal{F}\otimes\mathcal{B}(T)$-measurable.

(ii) Let $Y = \mathbb{R}^1$, ν a measure on $\mathcal{F}$, λ a measure on $\mathcal{B}(T)$, $\mu = \nu\otimes\lambda$, and let Q be some countable everywhere dense set in T. Suppose additionally that for every x, the set of discontinuity points of the function $t \mapsto f(x,t)$ has λ-measure zero and $\liminf_{r\to t, r\in Q} f(r,x) \le f(x,t) \le \limsup_{r\to t, r\in Q} f(r,x)$. Prove that the function f is measurable with respect to $\bigl(\mathcal{F}\otimes\mathcal{B}(T)\bigr)_\mu$.

7.14.104.° Let μ be a finite nonnegative measure on a measurable space $(X, \mathcal{A})$. Let a function $f\colon\ X\times[0,1] \to \mathbb{R}^1$ be such that for every fixed t, the function $x \mapsto f(x,t)$ is μ-measurable, and for μ-a.e. x, the function $t \mapsto f(x,t)$ is Riemann integrable. Set

$$f_0(x,t) := \frac{d}{dt}F(x,t), \quad F(x,t) := \int_0^t f(x,s)\,ds,$$

where $f_0(x,t) = 0$ if $f(x,s)$ is not Riemann integrable or the derivative does not exist. Prove that the function f_0 is measurable with respect to the measure $\mu\otimes\lambda$, where λ is Lebesgue measure, and that for a.e. x, one has $f_0(t,x) = f(t,x)$ for a.e. t, although f may not be $\mu\otimes\lambda$-measurable.

Hint: observe that the function $F(x,t)$ is measurable in x for any fixed t and is continuous in t for a.e. x.

7.14.105. (Talagrand [**1834**, p. 140]) Let X_i, $i = 1,\ldots,n$, be compact spaces with Radon probability measures μ_i which for all $i \ge 2$ do not vanish on nonempty open sets. Suppose that the function $f\colon\ \prod_{i=1}^n X_i \to \mathbb{R}$ is continuous in every variable separately. Then, there exist metrizable compact sets K_i, continuous surjections $h_i\colon\ X_i \to K_i$, and a function $g\colon\ \prod_{i=1}^n K_i \to \mathbb{R}$, continuous in every variable separately, such that $f(x_1,\ldots,x_n) = g\bigl(h_1(x_1),\ldots,h_n(x_n)\bigr)$. In particular, f is a Baire function.

Hint: we consider only the simpler case $n = 2$ and take a mapping $h_1\colon\ x \mapsto f_x$, where $f_x(y) = f(x,y)$, from X_1 to the space $C(X_2)$ with the topology of pointwise convergence. This mapping is continuous and its range $K_1 := h_1(X_1)$ is compact. By Theorem 7.10.9, any sequence in K_1 has a pointwise convergent subsequence. Since K_1 consists of continuous functions and the support of μ_2 coincides with X_2, the set K_1 is compact in the topology τ_{μ_2} of convergence in measure μ_2. Analogous arguments show that the topology τ_{μ_2} on K_1 is stronger than the topology of pointwise convergence and hence coincides with the latter, which means the metrizability of K_1. Let $h_2\colon\ X_2 \to C(K_1)$ be given by the formula $h_2(y)(f_x) = f(x,y)$, $y \in X_2$, $f_x \in K_1$, where $C(K_1)$ is equipped with the topology of pointwise convergence. Then $K_2 := h_2(X_2)$ is compact in this topology, which implies the metrizability of K_2. Finally, let $g(u,v) = v(u)$, $u \in K_1$, $v \in K_2 \subset C(K_1)$. Then we have $f(x,y) = g\bigl(h_1(x), h_2(y)\bigr)$.

7.14.106. (i) (Fremlin [**621**, Proposition 4J]) Prove that under the continuum hypothesis there exist a compact space X with a Radon measure μ and a function f on $X\times[0,1]$ that is continuous in the first argument, is Lebesgue measurable in the second argument, but is not measurable with respect to the Radon extension of the measure $\mu\otimes\lambda$, where λ is Lebesgue measure.

(ii) Let μ be a probability measure on a space $(X,\mathcal{A})$, let ν be a Radon probability measure on a compact metric space Y, and let a function $f\colon\ X\times Y \to \mathbb{R}^1$ be continuous in the second argument and μ-measurable in the first argument. Prove that f is measurable with respect to $\mathcal{A}_\mu \otimes \mathcal{B}(Y)$, in particular, measurable with respect to the measure $\mu\otimes\nu$. Extend the latter assertion to the case, where Y is a space with metrizable compacts.

HINT: (ii) let $\varphi_x(y) := f(x,y)$. The mapping $x \mapsto \varphi_x$, $X \mapsto C(Y)$, where $C(Y)$ is equipped with its usual norm and the σ-algebra $\mathcal{B}\bigl(C(Y)\bigr)$, is $\mathcal{A}_\mu$-measurable, since by hypothesis the functions $x \mapsto \varphi_x(y)$, $y \in Y$, are $\mathcal{A}_\mu$-measurable and $\mathcal{B}\bigl(C(Y)\bigr)$ is generated by the functions $y \mapsto \varphi(y)$. Then the mapping $\Psi\colon (x,y) \mapsto (\varphi_x, y)$, $X\times Y \to C(Y)\times Y$, is $\mathcal{A}_\mu\otimes\mathcal{B}(Y)$-measurable, and f is the composition of Ψ and the continuous function $(\varphi, y) \mapsto \varphi(y)$ on $C(Y)\times Y$. An alternative proof: approximate f by simple functions as in Lemma 6.4.6.

7.14.107. (Nussbaum [**1387**]) Let X be a compact (or locally compact) space, let Y be a Hausdorff space, and let ν be a Radon measure on X. Suppose that a function $f\colon X\times Y \to \mathbb{R}^1$ is continuous in every variable separately. Suppose also there exists a ν-integrable function g such that $|f(x,y)| \le g(y)$ for all $x \in X$, $y \in Y$. Prove that the function

$$x \mapsto \int_Y f(x,y)\,\nu(dy)$$

is continuous.

7.14.108.° Prove Proposition 7.14.8.

HINT: one can assume that $\|\nu\| \le 1$ and $|f| \le 1$; given $x_0 \in X$ and $\varepsilon > 0$, for every point $y \in Y$, one can find open sets $U(y)$ and V_y such that $y \in U(y)$, $x_0 \in V_y$ and $|f(x,z) - f(x_0,y)| < \varepsilon$ for all $(x,z) \in V_y \times U(y)$. By the τ-additivity, there exists a finite collection $U(y_1),\ldots,U(y_k)$ such that $|\nu|\bigl(\bigcup_{i=1}^k U(y_i)\bigr) > 1-\varepsilon$. Letting $V := V_{y_1} \cap\cdots\cap V_{y_k}$, one obtains

$$\int_Y |f(x,y) - f(x_0,y)|\,\nu(dy) \le 2\varepsilon$$

for all $x \in V$.

7.14.109. (Babiker, Knowles [**88**]) (i) Let X be a completely regular space and let μ be a Baire probability measure on X. Suppose that for every completely regular space Y and every function $f \in C_b(X\times Y)$, the function

$$g(y) = \int_X f(x,y)\,\mu(dx)$$

is continuous. Prove that the measure μ is τ-additive.

(ii) Let X and Y be completely regular spaces such that $\mathcal{B}a(X)\otimes\mathcal{B}a(Y) = \mathcal{B}a(X\times Y)$. Prove that for every Baire measure μ on X and every function f in $C_b(X\times Y)$, the function g in (i) is continuous.

7.14.110. (Johnson [**905**]) Let X and Y be compact spaces, let μ be a Radon measure on X, and let f be a bounded function on $X\times Y$ that is separately continuous in every argument.

(i) Prove that the set of functions $f_x\colon y \mapsto f(x,y)$, $x \in \operatorname{supp}\mu$, is separable in the Banach space $C(Y)$.

(ii) Give an example showing that in (i) the set of all functions f_x, $x \in X$, may be nonseparable.

(iii) Prove that the function f is measurable with respect to every Radon measure on $X\times Y$.

(iv) Prove that if $X = \operatorname{supp}\mu$, then the function f is Borel.

7.14.111. Construct a Borel probability measure μ on a topological space X with the following property: for every set $B \in \mathcal{B}(X\times X)$, the functions

$$x_2 \mapsto \mu\bigl(\{x_1\colon (x_1,x_2)\in B\}\bigr) \quad \text{and} \quad x_1 \mapsto \mu\bigl(\{x_2\colon (x_1,x_2)\in B\}\bigr)$$

are measurable with respect to μ, but have different integrals for some B.

HINT: see Johnson, Wilczynski [**914**]; take for X the set of all infinite ordinals smaller than the first uncountable ordinal ω_1 with the Dieudonné measure μ from Example 7.1.3 (in the notation of that example, the space is X_0). Finally, take the set $B = \bigl\{(x_1,x_2)\colon x_1 \geq x_2\bigr\}$.

7.14.112. Construct an example of completely regular first countable spaces X and Y equipped with Baire probability measures μ and ν such that the Baire measures ζ and ζ' on $X\times Y$ defined by the formulas

$$\int_{X\times Y} f\,d\zeta = \int_X\int_Y f(x,y)\,\mu(dx)\,\nu(dy),\ \int_{X\times Y} f\,d\zeta' = \int_Y\int_X f(x,y)\,\nu(dy)\,\mu(dx)$$

do not coincide.

HINT: see Fremlin [**635**, §439Q].

7.14.113.° (Carathéodory [**308**]) Let μ be a finite nonnegative measure on a measurable space X.

(i) Let a function $f\colon X\times\mathbb{R}^1 \to \mathbb{R}^1$ be such that for every fixed t, the function $x\mapsto f(x,t)$ is μ-measurable, and for μ-a.e. x, the function $t\mapsto f(x,t)$ is continuous. Prove that for every μ-measurable function φ, the function $x\mapsto f\bigl(x,\varphi(x)\bigr)$ is μ-measurable.

(ii) Let $T_1,\ldots,T_k$ be separable metric spaces and let $f\colon X\times T_1\times\cdots\times T_k \to \mathbb{R}^1$ be a function that is separately continuous in every t_i for μ-a.e. x and, for all fixed $(t_1,\ldots,t_k)$, is μ-measurable in x. Prove that for every μ-measurable mapping $\varphi\colon X\to T_1\times\cdots\times T_k$, the function $f\bigl(x,\varphi(x)\bigr)$ is μ-measurable.

HINT: (i) if φ assumes finitely many values c_i on measurable sets A_i, then $f\bigl(x,\varphi(x)\bigr) = f(x,c_i)$ for all $x\in A_i$, whence the measurability follows. In the general case, there is a sequence of simple functions φ_n convergent a.e. to φ, hence $\lim\limits_{n\to\infty} f\bigl(x,\varphi_n(x)\bigr) = f\bigl(x,\varphi(x)\bigr)$ a.e. (for all x at which one has the continuity of f in t and $\varphi(x) = \lim\limits_{n\to\infty}\varphi_n(x)$). (ii) For $k=1$ the reasoning from (i) is applicable, the general case follows by induction on k, since for a.e. x the function $f\bigl(x,\varphi_1(x),t_2,\ldots,t_k\bigr)$ is separately continuous in t_i.

7.14.114. (Grande, Lipiński [**728**]) Assuming the continuum hypothesis, construct a nonmeasurable function $F\colon \mathbb{R}^2\to\mathbb{R}^1$ such that for every Lebesgue measurable function $f\colon \mathbb{R}^1\to\mathbb{R}^1$, the function $F\bigl(x,f(x)\bigr)$ is measurable (cf. Exercise 9.12.62).

7.14.115. Let $(X,\mathcal{A},\mu)$ be a complete probability space, let $L^0(\mu)$ be equipped with the metric d_0 of convergence in measure, and let $(T,\mathcal{T})$ be a measurable space. Prove that the following conditions on a mapping $F\colon T\to L^0(\mu)$ are equivalent: (i) $F(T)$ is separable and F is measurable, (ii) there exists a $\mathcal{T}\otimes\mathcal{A}$-measurable function G on $T\times X$ such that for every $t\in T$, one has $F(t)(x) = G(t,x)$ for a.e. x.

HINT: if one has (i), then there is a sequence of points $t_k\in T$ such that the set $\{F(t_k)\}$ is dense in $F(T)$. Fix some representatives f_k of the classes $F(t_k)\in L^0(\mu)$. By the measurability of F we have $T_{n,k} := \bigl\{t\in T\colon d_0\bigl(F(t),F(t_k)\bigr) < 2^{-n}\bigr\}\in\mathcal{T}$ for all $n,k\in\mathbb{N}$. The sets $D_{n,k} := T_{n,k}\backslash\bigcup_{i=1}^{k-1} T_{n,i}$ for every fixed n form a measurable partition of T. Set $G_n(t,x) = f_k(x)$ if $t \in D_{n,k}$. It is clear that we obtain $\mathcal{T}\otimes\mathcal{A}$-measurable functions. Let $G(t,x) := \lim\limits_{n\to\infty} G_n(t,x)$ at all points (t,x) where this limit exists and is finite (this set belongs to $\mathcal{T}\otimes\mathcal{A}$), and let $G(t,x) = 0$ at all other points. Now one can verify that G is the required function. If one has (ii), then it suffices to verify (i) for bounded functions G, which by using uniform approximations reduces the claim to the case of the indicator of a set E in $\mathcal{T}\otimes\mathcal{A}$. The separability of $F(T)$ and measurability of F are verified directly for sets E in the algebra generated by the products $S\times A$, $S \in \mathcal{T}$, $A \in \mathcal{A}$. Now the monotone class theorem yields the claim for all $E \in \mathcal{T}\otimes\mathcal{A}$.

7.14.116.° Let X and Y be compact spaces with Radon probability measures μ and ν and let $\mu^*(A) = \nu^*(B) = 1$. Show that $(\mu\otimes_R\nu)^*(A\times B) = 1$, where $\mu\otimes_R\nu$ is the Radon extension of $\mu\otimes\nu$ to $X\times Y$.

HINT: let K be compact in $X\times Y$; if $\mu\otimes_R\nu(K) > 0$, then there exists $x \in A$ such that $\nu(K_x) > 0$, i.e., there exists $y \in B$ such that $(x,y) \in K$. Hence $K \cap (A\times B)$ is nonempty.

7.14.117. (Talagrand [**1834**, p. 121]) Let $(X,\mathcal{A},\mu)$ be a probability space, let Y be a compact space with a Radon probability measure ν, and let a function f on $X\times Y$ be continuous in y and measurable with respect to $\mu\otimes\nu$. Show that for every $\varepsilon > 0$, there exist two sequences of sets $A_n \in \mathcal{A}$ and $B_n \in \mathcal{B}(Y)$ such that $\mu\otimes\nu\bigl(\bigcup_{n=1}^\infty A_n\times B_n\bigr) = 1$ and on every $A_n\times B_n$ the oscillation of f does not exceed ε.

7.14.118.° (Tolstoff [**1863**]) Let $(X,\mathcal{A},\mu)$ be a probability space, let (Y,d) be a complete separable metric space, and let $y_0 \in Y$ be a fixed point. Suppose that a function $f\colon X\times Y \to \mathbb{R}$ is measurable with respect to $\mathcal{A}\otimes\mathcal{B}(Y)$ and the equality $\lim_{y\to y_0} f(x,y) = f(x,y_0)$ holds for every $x \in X$. Prove that for every $\varepsilon > 0$, there exists a set $A_\varepsilon \in \mathcal{A}$ such that $\mu(A_\varepsilon) > 1-\varepsilon$ and $\lim_{y\to y_0} f(x,y) = f(x,y_0)$ uniformly in $x \in A_\varepsilon$.

HINT: in the solution to Exercise 2.12.46, use Theorem 6.10.9 in place of Proposition 1.10.8.

7.14.119.° Let K be a compact space and let μ be a Radon probability measure on K with support K.

(i) Prove that the following conditions on a bounded function f are equivalent: (a) there exists a bounded function g such that the set of all discontinuity points of g has μ-measure zero and $f(x) = g(x)$ μ-a.e., (b) there exists a set Z of measure zero such that the restriction of f to $K\backslash Z$ is continuous.

(ii) Construct an example showing that in (i) one cannot always find a continuous function g.

HINT: (i) if (b) is fulfilled, then the set $A = K\backslash Z$ is everywhere dense in K and one can define $g(x) = \limsup_{y\to x, y\in A} f(y)$ if $x \in Z$.

7.14.120. One says that a Radon measure μ has a metrizable-like support if there exists a sequence of compact sets $K_n \subset X$ such that for every open set $U \subset X$ and $\varepsilon > 0$, there exists n with $K_n \subset U$ and $|\mu|(U\backslash K_n) < \varepsilon$. Show that this property is strictly stronger than the separability of μ. Show that the existence of a metrizable-like support follows from the existence of a sequence of metrizable compact sets K_n with $|\mu|(X) = |\mu|(\bigcup_{n=1}^\infty K_n)$, but is weaker than the latter condition.

HINT: see Gardner [**660**, Section 24]; see also Example 9.5.3.

7.14.121. (Lozanovskiĭ [**1195**]) Let K_1 and K_2 be compact spaces without isolated points. Prove that there is no Radon probability measure μ on $K_1\times K_2$ such that $\mu(K)=0$ for every nowhere dense compact set K.

7.14.122. (Fremlin [**630**]) Let μ be a Radon probability measure on a topological space X, regarded on $\mathcal{B}(X)_\mu$, let μ^* and μ_* be the corresponding outer and inner measures, and let ν be the measure generated by the Carathéodory outer measure $\mathfrak{m} := (\mu^*+\mu_*)/2$ (see Exercise 1.12.143). Prove that $\mu=\nu$. Show that the same is true for every complete perfect atomless probability measure.

7.14.123. (i) (Varadarajan [**1918**]) Let μ be a τ-additive Borel measure on a paracompact space X. Prove that the topological support of μ is Lindelöf.

(ii) (Plebanek [**1469**]) Show that there exists a τ-additive Baire measure without Lindelöf subspaces of full measure.

HINT: (i) let S be the support of μ and let U_t, where $t\in T$, be an open cover of S. Since S is closed, the space S is paracompact too. Hence one can inscribe in the given cover an open cover $\mathcal{V}$ representable in the form $\mathcal{V}=\bigcup_{n=1}^\infty \mathcal{V}_n$, where every subfamily $\mathcal{V}_n$ consists of disjoint sets $V_{n,\alpha}$ (see Engelking [**532**, Theorem 5.1.12]). It is clear from the definition of S that $|\mu|(V_{n,\alpha}\cap S)>0$. Therefore, for every fixed n, one has at most countably many nonempty sets $V_{n,\alpha_k}\cap S$, which gives a countable cover of S by the sets V_{n,α_k}, consequently, a countable subcover in $\{U_t\}$.

7.14.124. (Aldaz, Render [**20**]) Let X be a $\mathcal{K}$-analytic Hausdorff space in the sense of Definition 6.10.12, let $\mathcal{F}$ be the class of all closed sets in X, and let μ be a probability measure on some σ-algebra $\mathcal{E}$ such that for every $E\in\mathcal{E}$, one has $\mu(E)=\sup\{\mu(F)\colon\ F\subset E, F\in\mathcal{F}\cap\mathcal{E}\}$. Prove that μ extends to a Radon measure on X.

7.14.125. Let X be a $\mathcal{K}$-analytic space in the sense of Definition 6.10.12. Prove that every Borel probability measure on X is tight.

HINT: let Ψ be a mapping from $\mathbb{N}^\infty$ to the set of compact subsets of X, representing X. Let $\Psi(A)=\bigcup_{a\in A}\Psi(a)$ and $C(A)=\mu^*\bigl(\Psi(A)\bigr)$, $A\subset\mathbb{N}^\infty$. It is verified directly that $\Psi(K)$ is compact for every compact K. Next we verify that C is a Choquet capacity on $\mathbb{N}^\infty$. A direct proof is found in Fremlin [**635**, §432B].

7.14.126. (i) (Kindler [**998**]) Let $\mathcal{F}$ be a vector lattice of functions on a set Ω and let L be a nonnegative linear functional on $\mathcal{F}$ such that $L(f_n)\to 0$ for each sequence $\{f_n\}\subset\mathcal{F}$ that decreases pointwise to zero. Given $f,g\in\mathcal{F}$ with $f\le g$, let

$$[f,g[:=\{(x,t)\in\Omega\times\mathbb{R}^1\colon\ f(x)\le t<g(x)\}\quad\text{and}\quad \nu([f,g[):=L(g-f).$$

Prove that the class $\mathcal{R}$ of all such sets $[f,g[$ is a semiring and the function ν is well-defined and countably additive on $\mathcal{R}$.

(ii) Apply (i) to prove Theorem 7.8.7 letting

$$\mu\bigl(\{f>1\}\bigr):=\nu\bigl(\{f>1\}\times[0,1)\bigr),\quad f\in\mathcal{F}.$$

(iii) Show that the measure μ is uniquely defined on the σ-ring generated by the sets $\{f>1\}$, where $f\in\mathcal{F}$, and give an example where μ is not unique on the σ-algebra generated by $\mathcal{F}$.

HINT: see Dudley [**495**, §4.5].

7.14.127. Show that the Sorgenfrey line is measure-compact, but its square is not.

HINT: see Fremlin [**635**, §439P].

7.14.128. Show that the class of all Radon spaces is closed under the following operations: (i) countable topological sums, (ii) countable unions of Radon subspaces, (iii) countable intersections of Radon subspaces, (iv) passage to universally Borel measurable subspaces. In addition, the countable product of Radon spaces in each of which all compact sets are metrizable, is Radon as well.

HINT: if μ is a Borel probability measure on the product of Radon spaces X_n with metrizable compacts, then its projections are Radon, which yields that μ is concentrated on a countable union of metrizable compacts (note that countable products of metrizable compacts are metrizable). Other assertions are immediate.

7.14.129.° Let X be a Radon space that is homeomorphically embedded into a topological space Y. Prove that X is measurable with respect to every Borel measure on Y.

HINT: let μ be a nonnegative Borel measure on Y and let $Y_0 \in \mathcal{B}(Y)$ be a measurable envelope of X in Y. Consider the measure $\nu(B \cap X) := \mu(B \cap Y_0)$, $B \in \mathcal{B}(Y)$, and observe that $\mathcal{B}(X) = \mathcal{B}(Y) \cap X$.

7.14.130.° Prove that $\mathbb{R}^{\mathfrak{c}}$ and $[0,1]^{\mathfrak{c}}$ are not Radon spaces.

HINT: the space from Example 7.1.3 can be embedded into $[0,1]^{\mathfrak{c}}$.

7.14.131.° Let two Radon measures μ and ν on a space X coincide on a countable algebra $\mathcal{A}$ that is contained in $\mathcal{B}(X)_{|\mu|} \cap \mathcal{B}(X)_{|\nu|}$ and separates the points in X. Prove that $\mu = \nu$.

HINT: in view of Proposition 7.14.24, one can assume that X is a countable union of metrizable compact sets. Let $\eta = |\mu| + |\nu|$. For every $A \in \mathcal{A}$ we find Borel sets A', A'' with $A' \subset A \subset A''$, $\eta(A') = \eta(A'')$. We obtain a countable collection $\mathcal{B}_0$ of Borel sets separating points. Hence the generated algebra $\mathcal{A}_0$ is countable and $\sigma(\mathcal{A}_0) = \mathcal{B}(X)$ (see Theorem 6.8.9). Finally, one has $\mu = \nu$ on $\mathcal{A}_0$. Indeed, it suffices to observe that given $B_1, \ldots, B_k$ in $\mathcal{B}_0$, we find sets $A_1, \ldots, A_k \in \mathcal{A}$ such that B_i is associated to A_i as A_i' or A_i'', whence $\eta\bigl((B_1 \cap \cdots \cap B_k) \triangle (A_1 \cap \cdots \cap A_k)\bigr) = 0$, and consequently $\mu(B_1 \cap \cdots \cap B_k) = \nu(B_1 \cap \cdots \cap B_k)$, since $A_1 \cap \cdots \cap A_k \in \mathcal{A}$. The assertion of this exercise is found in the literature in close formulations (see, e.g., Stegall [**1775**]).

7.14.132.° Prove that $Cyl(X, G)$ is the algebra of sets generated by G (see §7.12).

7.14.133. Let μ be a Radon measure on an infinite-dimensional locally convex space X. Show that its characteristic functional $\widetilde{\mu}$ is continuous in the weak* topology $\sigma(X^*, X)$ only in the case where μ is concentrated on the union of a sequence of finite-dimensional subspaces.

HINT: see Vakhania, Tarieladze, Chobanyan [**1910**, Ch. VI, §3, Theorem 3.3].

7.14.134.° Let X and Y be Banach spaces such that Y is separable and X is reflexive and let $T\colon X \to Y$ be a continuous injective linear mapping. Prove that X is separable.

HINT: embed Y injectively into l^2; in the case $Y = l^2$ verify that the range of the adjoint mapping $T^*\colon l^2 \to X^*$ is dense.

7.14.135. Construct an example of a cylindrical quasi-measure of unbounded variation on l^2 such that its characteristic functional is bounded and continuous in the Sazonov topology.

HINT: see Bogachev, Smolyanov [**225**, Remark 4.2].

7.14.136. (Kwapień [**1093**]) Let $\{\xi_n\}$ be a sequence of random variables on a probability space $(\Omega, \mathcal{F}, P)$ such that the series $\sum_{n=1}^{\infty} \lambda_n \xi_n$ converges in probability for every sequence of numbers $\lambda_n \to 0$. Prove that $\sum_{n=1}^{\infty} |\xi_n|^2 < \infty$ a.e. Deduce that the embedding $l^1 \to l^2$ is a radonifying operator, i.e., it takes every nonnegative cylindrical quasi-measure on l^1 with a continuous characteristic functional to a Radon measure on l^2.

7.14.137. (Schwartz [**1679**]) Let a linear function l on a Banach space X be measurable with respect to every Radon measure on X. Prove that l is continuous.

Hint: see Christensen [**355**] and Kats [**962**], where more general results are proven.

7.14.138. (Talagrand [**1829**]) Prove that under Martin's axiom every infinite-dimensional separable Banach space X contains a hyperplane X_0 that is not closed, but is measurable with respect to every Borel measure on X. By the previous exercise X_0 cannot be the kernel of a universally measurable linear function.

7.14.139. (Talagrand [**1834**, p. 184]) Let E be the Banach space of all bounded functions on $[0,1]$ that are nonzero on an at most countable set and let E be equipped with the norm $\sup |f(t)|$. Denote by w the weak topology of the space E. Then, there exists a probability measure on the weak Borel σ-algebra $\mathcal{B}\big((E,w)\big)$ that assumes only two values 0 and 1, but is not Radon.

7.14.140. (von Weizsäcker [**1969**]) (i) Let X be the space of all Borel measures on $[0,1]$ equipped with the weak topology (i.e., the weak* topology of $C[0,1]^*$) and let K be the convex compact set in X consisting of all probability measures. Let λ be Lebesgue measure and let μ be the image of λ under the continuous (in the indicated topology) mapping $\pi\colon\ t \mapsto \delta_t$, $[0,1] \to K$, where δ_t is Dirac's measure at the point t. Let

$$C := \bigcap_{n=1}^{\infty} \big\{ m \in K\colon\ \lambda + n^{-1}(\lambda - m) \in K \big\}.$$

Prove that C is a convex G_δ-set in K and $\mu(C) = 1$, but $\mu(S) = 0$ for every convex compact set $S \subset C$.

(ii) Let K be a convex compact set in a locally convex space X such that the linear span of K is infinite-dimensional. Prove that there exist a convex set $C \subset K$ and a Radon probability measure μ on K such that C is a G_δ-set in some metrizable convex compact set $K_0 \subset K$ and $\mu(C) = 1$, but $\mu(S) = 0$ for every convex compact set $S \subset C$.

Hint: (i) it is easily verified that C is convex and can be represented as the intersection of a sequence of open sets in K with the weak topology. In addition,

$$C = K \backslash \bigcup_{\varepsilon > 0} \big\{ m \in K\colon\ \lambda + \varepsilon(\lambda - m) \in K \big\}.$$

Let D be the set of all Dirac measures. Then D is compact in K and $\mu(D) = 1$. If $S \subset C$ is a convex compact set with $\mu(C) > 0$, then $\mu(S \cap D) > 0$. Then $A := \pi^{-1}(S \cap D)$ is compact and $\lambda(A) > 0$. Since $\delta_t \in S$ if $t \in A$, by using that S is convex and closed we obtain that every probability measure ν on A belongs to S. In particular, $\nu := \lambda(A)^{-1} \lambda|_A \in S$. The measure $\lambda + \lambda(A)(\lambda - \nu)$ is probability, hence belongs to K. According to the above equality for C we obtain that $\nu \notin C$. This contradicts the fact that $\nu \in S \subset C$. Claim (ii) is deduced from (i) by using a suitable mapping (see details in [**1969**]).

7.14.141. Let μ and ν be τ-additive measures on a locally convex space X with equal Fourier transforms. Prove that $\mu = \nu$.

HINT: let p be a continuous seminorm on X, let X_p be the normed space obtained by the factorization of X with respect to $p^{-1}(0)$, and let $\pi_p\colon X \to X_p$ be the natural projection. Sets of the form $\pi_p^{-1}(U)$, where p is a continuous seminorm and U is open in X_p, form a topology base in X. Hence it suffices to show the equality of μ and ν on such sets. The measures $\mu \circ \pi_p^{-1}$ and $\nu \circ \pi_p^{-1}$ on X_p have equal Fourier transforms and are τ-additive. Both properties are preserved for the natural extensions of the two measures to the completion of X_p. Since on a Banach space all τ-additive measures are Radon, we obtain the equality of the indicated extensions on the completion of X_p, hence the equality $\mu\circ\pi_p^{-1} = \nu\circ\pi_p^{-1}$. Therefore, $\mu\bigl(\pi_p^{-1}(U)\bigr) = \nu\bigl(\pi_p^{-1}(U)\bigr)$ for all open sets $U \subset X_p$.

7.14.142. (i) Let μ be a Radon probability measure on a convex compact set K in a locally convex space X. Show that μ has a barycenter $b \in K$.

(ii) Let X be a complete locally convex space and let μ be a τ-additive probability measure on X with bounded support. Prove that μ has a barycenter.

HINT: (i) it is not difficult to show that there is a net of probability measures μ_α with finite support in K possessing the following property:

$$\lim_\alpha \int_K f\,d\mu_\alpha = \int_K f\,d\mu \quad \text{for every } f \in X^*.$$

It is obvious that the measures μ_α have barycenters $b_\alpha \in K$ that possess an accumulation point $b \in K$, which is the barycenter of μ. (ii) See Fremlin [**635**, §461E].

7.14.143.° Let K be a convex compact set in a locally convex space X. Suppose that a sequence of Radon probability measures μ_n on K converges to μ in the weak* topology on $C(K)^*$. Prove that the barycenters of the measures μ_n converge to the barycenter of μ.

HINT: the weak topology on K coincides with the original topology.

7.14.144.° Let K be a compact set in a locally convex space X. Prove that its closed convex envelope $\widehat{K}$ coincides with the set of barycenters of all Radon probability measures on K.

HINT: if μ is a Radon probability measure on K and b is its barycenter, then for every $l \in X^*$ we have $l(b) \le \sup_{x\in K} l(x)$, whence by the Hahn–Banach theorem we obtain $b \in \widehat{K}$. The converse is verified first for finite sets. Then the convex envelope of K belongs to the set of barycenters of probability measures on K. Let $b \in \widehat{K}$. There is a net of points b_α in the convex envelope K convergent to b. Let us take a probability measure μ_α on K with the barycenter b_α. The net $\{\mu_\alpha\}$ has a limit point μ in the weak topology on $\mathcal{P}_t(K)$ (see Chapter 8). Then b is the barycenter of μ.

7.14.145.° Prove Proposition 7.14.44.

HINT: it is clear from Lemma 6.2.3 that there exists a Borel mapping G with a finite range such that the integral of the function $\|F(x)-G(x)\|^p$ is less than $\varepsilon^p 2^{-p}$. Since G takes values in some finite-dimensional subspace E, by Corollary 7.12.2 there exists a mapping F_0 of the form $F_0 = \varphi \circ P$, where $P\colon X \to \mathbb{R}^n$ is a continuous linear mapping and $\varphi\colon \mathbb{R}^n \to E$ is a continuous mapping with compact support such that the integral of $\|F_0(x) - G(x)\|^p$ is less than $\varepsilon^p 2^{-p}$.

7.14.146. Prove Theorem 7.14.45.

HINT: one can modify the proof of Theorem 7.12.4 (see Bogachev [**209**]; another proof is outlined in Ostrovskiĭ [**1406**]). Let $\varphi \geq 0$ be a decreasing function on $[0, \infty)$ with $\sum_{n=1}^{\infty} \varphi(n) < \infty$. One can find $\alpha_n \downarrow 0$ with $\alpha_n n \to \infty$ and $\sum_{n=1}^{\infty} \varphi(\alpha_n n) < \infty$. Let $\varphi(R) = \mu(x\colon \|x\| > R^{1/r})$. Then $\sum_{n=1}^{\infty} \varphi(n) < \infty$. Take α_n as above. For every n, there is a compact set K_n in the ball U_n of radius $n^{1/r}$ centered at the origin such that $\mu(\alpha_n^{1/r} K_n) \geq \mu(\alpha_n^{1/r} U_n) - 2^{-n}$. The set $K = \bigcup_{n=1}^{\infty} c_n K_n$, where $c_n := \alpha_n^{1/r} n^{-1/r}$, has compact closure. The closed convex envelope V of the set K is compact too. Let p_V be the Minkowski functional of V and E_V the associated Banach space. Since $\{p_V \leq c\} = cV$ as $c \geq 0$, the function p_V is measurable. One has $\alpha_n^{1/r} K_n \subset n^{1/r} K \subset \{p_V \leq n^{1/r}\} = n^{1/r} V$. By our choice of K_n we obtain $p_V^r \in L^1(\mu)$, since

$$\mu\bigl(x\colon\ p_V^r(x) > n\bigr) = 1 - \mu\bigl(x\colon\ p_V(x) \leq n^{1/r}\bigr) \leq 1 - \mu\bigl(\alpha_n^{1/r} K_n\bigr)$$
$$\leq 1 + 2^{-n} - \mu\bigl(\alpha_n^{1/r} U_n\bigr) = 2^{-n} + \mu\bigl(x\colon\ \|x\| > \alpha_n^{1/r} n^{1/r}\bigr) = 2^{-n} + \varphi(\alpha_n n).$$

It is clear that $\mu(E_V) = 1$ and $\mu\bigl(\alpha_n^{1/r} K_n\bigr) \to 1$, since the balls $\alpha_n^{1/r} U_n$ have radii $(\alpha_n n)^{1/r} \to \infty$ (note that $\alpha_n n \to \infty$). We argue further as in Theorem 7.12.4. The case of a Fréchet space reduces to the considered case by passing to the subspace $X_0 := \{q < \infty\}$, where $q^r := \sum_{n=1}^{\infty} c_n q_n^r$, $\{q_n\}$ is a sequence of seminorms defining the topology, and $c_n = 2^{-n}(\|q_n^r\|_{L^1(\mu)} + 1)^{-1}$.

7.14.147. Prove Proposition 7.14.12.

7.14.148. Prove Lemma 7.14.16.

7.14.149. Construct an example of a probability measure on a locally convex space (X, τ) that is defined on $\sigma(X^*)$ and is tight in the weak topology $\sigma(X, X^*)$, but is not tight in the original topology τ.

HINT: let $E = C[0,1]$, let $X := E^*$ be equipped with the topology $\sigma(E^*, E)$, and let μ be the image of Lebesgue measure on $[0,1]$ under the mapping $t \mapsto \delta_t$. Then μ is a tight Baire measure with respect to the topology $\sigma(E^*, E)$. Let us take for τ the Mackey topology $\tau(E^*, E)$. If μ were tight in this topology, then it would be tight in the topology $\sigma(E^*, E^{**})$ according to Exercise 8.10.124. Then μ would have a Radon extension in the norm topology, hence it would have a norm separable support. This leads to a contradiction since $\|\delta_t - \delta_s\| = 2$ if $t \neq s$.

7.14.150.° Let K be a compact space and let a set $X \subset K$ be measurable with respect to all Radon measures on K. Prove that $\mathcal{M}_\tau(X) = \mathcal{M}_r(X)$.

HINT: see Example 7.14.22.

7.14.151. Prove Theorem 7.14.49.

7.14.152.° Prove Lemma 7.14.53.

HINT: if $l \in \Lambda_0(\mu)$, then there is a sequence $\{l_n\} \subset X^*$ convergent to l a.e. The set X_0 of all points of convergence of $\{l_n\}$ is a Borel linear subspace in X and $\mu(X_0) = 1$. Let $f(x) = \lim\limits_{n\to\infty} l_n(x)$ if $x \in X_0$. Then f is a Borel linear function on X_0 and $f = l$ a.e. on X_0. There is a linear subspace $X_1 \subset X$ such that X is the direct algebraic sum of X_0 and X_1. Hence f can be extended to a linear function on all of X by letting $f|_{X_1} = 0$. The extension is μ-measurable since $\mu(X_0) = 1$, i.e., we obtain a version of l in the class $\Lambda(\mu)$.

7.14.153. Show that there exists a probability measure μ on some compact metric space K such that $\sum_{n=1}^{\infty}\mu(B_n) < 1/2$ for every sequence of disjoint closed balls with radii at most 1.

HINT: see Davies [**412**] or Wise, Hall [**1993**, Example 4.49].

7.14.154. Let $(X,\mathcal{A},\mu)$ be a space with a complete locally determined (see Exercise 1.12.135) measure with values in $[0,+\infty]$ and let $\mathcal{K}$ be a family of sets such that $\mu(A)=\sup\{\mu(K)\colon\ K\in\mathcal{K}\cap\mathcal{A}, K\subset A\}$ for all $A\in\mathcal{A}$. Prove that the following conditions on a set $A\subset X$ are equivalent:

(i) $A\in\mathcal{A}$,

(ii) $A\cap K\in\mathcal{A}$ for all $K\in\mathcal{K}\cap\mathcal{A}$,

(iii) $\mu^*(K\cap A)+\mu^*(K\backslash A)=\mu^*(K)$ for all $K\in\mathcal{K}$,

(iv) $\mu^*(K\cap A)=\mu_*(K\cap A)$ for all $K\in\mathcal{K}\cap\mathcal{A}$.

HINT: see, e.g., Fremlin [**635**, §413F].

7.14.155. Show that the property to be Radon or τ-additive for a Borel probability measure on a product of two compact spaces does not follow from the fact that its projections on the factors are Radon.

HINT: according to Wage [**1955**], under the continuum hypothesis there exist Radon compact spaces X and Y such that there is a non-Radon Borel probability measure on their product. Both projections of this measure are Radon.

7.14.156. (i) Let T be an uncountable set, let $X=\mathbb{R}^T$, and let μ be a separable probability measure on $\mathcal{B}a(X)$ (i.e., $L^2(\mu)$ is separable). Prove that there exist a countable set $\{t_n\}\subset T$ and a probability measure ν on $\mathbb{R}^\infty$ such that $\mu=\nu\circ\pi^{-1}$, where $\pi=(\pi_t)\colon\ \mathbb{R}^\infty\to X$, π_t are measurable functions on $\mathbb{R}^\infty$, $\pi_{t_n}(x)=x_n$, and for every $t\notin\{t_n\}$, the function x_t is a.e. the limit of a subsequence in $\{x_{t_n}\}$.

(ii) Let T be an uncountable set, let $X=\mathbb{R}^T$, and let $\mu=\bigotimes_{t\in T}\mu_t$, where all measures μ_t coincide with an atomless Borel probability measure σ on the real line. Show that the restriction of μ to every set of positive measure is not separable.

(iii) Let μ be the Radon extension of the product of an uncountable family of copies of Lebesgue measure on $[0,1]$. Prove that $\mu(S)=0$ for every Souslin set S, in particular, for every metrizable compact set S.

HINT: (i) one can deal with the space $(0,1)^T$, then the coordinate functions x_t belong to $L^2(\mu)$. By the separability of $L^2(\mu)$, there exists a countable set $\{t_n\}\subset T$ such that the sequence of functions x_{t_n} is everywhere dense in the set of all functions x_t, $t\in T$, with the metric from $L^2(\mu)$. Hence for every $t\notin\{t_n\}$, there exists a sequence of indices $s_k\in\{t_n\}$ such that $x_t=\lim\limits_{k\to\infty}x_{s_k}$ in $L^2(\mu)$. Passing to a subsequence, we can assume that this relationship is true μ-a.e. Its right-hand side we take for π_t. Let $\pi_{t_n}(x)=x_n$. Let ν be the projection of μ under the mapping $(x_t)_{t\in T}\mapsto(x_{t_n})_{n=1}^\infty$. (ii) As in (i) we can deal with $(0,1)^T$ in place of $\mathbb{R}^T$. If $\mu(E)>0$, then the measure $\nu:=\mu|_E$ is positive. If this measure is separable, then according to (i), there exist a countable set $\{t_n\}\subset T$ and an index $t\notin\{t_n\}$ such that for some subsequence $\{s_k\}\subset\{t_n\}$, we have $x_t=\lim\limits_{k\to\infty}x_{s_k}$ ν-a.e. This leads to a contradiction, since the set $\Omega:=\bigl\{(u_0,u_1,\ldots)\in(0,1)^\infty\colon\ u_0=\lim\limits_{k\to\infty}u_k\bigr\}$ has measure zero with respect to the product of countably many copies of σ. This follows by Fubini's theorem because if we fix numbers u_k with $k\geq 1$, the set $\{u\colon\ (u,u_1,u_2,\ldots)\in\Omega\}$ is either empty or consists of a single point and has measure zero with respect to σ. Assertion (iii) follows from (ii) by the separability of any Borel measure on a Souslin space.

7.14.157. (Bourbaki [**242**, Ch. V, §8.5, exercise 13]) Let T be an uncountable set and let μ_t, where $t \in T$, be a family of Radon probability measures on compact spaces X_t such that the support of μ_t coincides with X_t. Denote by μ the Radon measure obtained as the extension of the product of μ_t. Let $E = \prod_{t\in T} E_t$, where $E_t \neq X_t$ for every t.

(i) Prove that $\mu_*(E) = 0$.

(ii) Let $\mu_t(E_t) = 1$ for all t. Prove that E does not belong to the Lebesgue completion of $\mathcal{B}\bigl(\prod_{t\in T} X_t\bigr)$ with respect to μ, in particular, is not Borel. For example, in the case of uncountable T, the sets $(0,1)^T$ and $(0,1]^T$ in $[0,1]^T$ are not measurable with respect to the Radon extension of the product of T copies of Lebesgue measure on $[0,1]$.

HINT: (i) for any compact $K \subset E$, its projections K_t to X_t are compact and differ from X_t. Hence there exists $n \in \mathbb{N}$ such that the set of all t for which $\mu_t(K_t) \le 1 - n^{-1}$ is infinite. We take a countable set of such points t_j and obtain the set $\prod_{j=1}^{\infty} K_{t_j} \times \prod_{t\notin\{t_j\}} X_t$ that contains K and has μ-measure zero. (ii) Suppose that E is measurable. Then, according to (i), we have $\mu(E) = 0$, hence there is a Borel set B with $E \subset B$ and $\mu(B) = 0$. On the other hand, one can consider the product of the measures μ_t on the space E. It has a τ-additive extension μ' to $\mathcal{B}(E)$, hence a τ-additive extension μ'' to $\prod_{t\in T} X_t$, which coincides with μ by the equality on all cylinders. This leads to a contradiction, since $\mu''(B) = 1$. An alternative reasoning: we take a compact set S in the complement of E with $\mu(S) > 0$, find a set $A \supset S$ that depends on countably many indices t_j such that $\mu(A) = \mu(S)$, apply to S Fubini's theorem and use the compactness of the sections of S.

7.14.158. (Chentsov [**335**], [**337**]) Let $X = [0,1]^T$, $T = [0,1]$, and let $\Omega = [0,1]$ be equipped with Lebesgue measure λ. Set $\xi_1(t,\omega) = t - \omega + 1$ if $t \in [0,\omega)$, $\xi_1(t,\omega) = t - \omega$ if $t \in [\omega,1]$, $\xi_2(t,\omega) = t - \omega + 1$ if $t \in [0,\omega]$, $\xi_2(t,\omega) = t - \omega$ if $t \in (\omega,1]$. Let $f_1, f_2\colon \Omega \to X$, $f_1(\omega)(t) = \xi_1(t,\omega)$, $f_2(\omega)(t) = \xi_1(t,\omega)$. Finally, let us consider two probability measures $\mu_1 = \lambda\circ f_1^{-1}$ and $\mu_2 = \lambda\circ f_2^{-1}$ on the σ-algebras $\mathcal{A}_1$ and $\mathcal{A}_2$ consisting of all sets in X whose preimages with respect to f_1 and f_2, respectively, are Lebesgue measurable. Show that μ_1 and μ_2 are Radon on $\mathcal{B}(X)$ and coincide on all cylinders, whence one has their equality on $\mathcal{B}(X)$. However, $\mu_1\bigl((0,1]^T\bigr) = 0$, $\mu_2\bigl((0,1]^T\bigr) = 1$.

7.14.159. Let X be a Hausdorff topological vector space and let μ and ν be Radon probability measures with $\mu = \mu * \nu$. Show that ν is Dirac's measure at the origin.

HINT: see Vakhania, Tarieladze, Chobanyan [**1910**, Proposition I.4.7].

7.14.160. Let X be the union of all open sets G in $\beta\mathbb{N}$ with $\sum_{n\in\pi(G)} n^{-1} < \infty$, where $\pi(E) := E \cap \mathbb{N}$ for $E \subset \beta\mathbb{N}$. Show that the measure $\mu(B) := \sum_{n\in\pi(B)} n^{-1}$ on $\mathcal{B}(X)$ is σ-finite and inner compact regular, but is not outer regular.

HINT: $\mu(X\backslash\mathbb{N}) = 0$, but there is no open set $U \supset X\backslash\mathbb{N}$ with $\mu(U) < \infty$. Indeed, if $U \supset X\backslash\mathbb{N}$ is open, the set $X\backslash U \subset \mathbb{N}$ is closed and finite because otherwise it would contain an infinite sequence $S = \{s_n\}$ with $\sum_{n=1}^{\infty} s_n^{-1} < \infty$. This is impossible, since the closure G of S in $\beta\mathbb{N}$ is open, hence $G \subset X$, but this closure must contain a point from $\beta\mathbb{N}\backslash\mathbb{N}$. Thus, $\mu(U)$ is infinite.

7.14.161. (i) Let $X = \beta\mathbb{N}\backslash\{a\}$, where $a \in \beta\mathbb{N}\backslash\mathbb{N}$. Show that X is locally compact, but there is a nonnegative linear functional on $C(X)$ that is not represented by a measure.

(ii) Show that if X is any locally compact and σ-compact space, then for every nonnegative linear functional L on $C(X)$, there is a measure μ on $\mathcal{B}(X)$ with values in $[0,+\infty]$ such that $C(X)\subset\mathcal{L}^1(\mu)$ and L is represented by μ.

(iii) Construct an example of a Radon probability measure μ on a locally compact space with a noncompact support and $C(X)\subset\mathcal{L}^1(\mu)$.

HINT: (i) $C(X)=C_b(X)$ according to Engelking [**532**, Example 3.10.18]; take $L(f)=\widehat{f}(a)$, where $\widehat{f}$ is the extension of f to $C(\beta\mathbb{N})$. (ii) Show that there is a compact set $K\subset X$ such that $L(f)=0$ if $f|_K=0$. Otherwise we could construct compact sets K_n and functions $f_n\in C(X)$ such that K_n is contained in the interior of K_{n+1}, $\bigcup_{n=1}^\infty K_n=X$, $f_n\geq 0$, $f_n|_{K_n}=0$, and $L(f_n)>1$. This is impossible since the function $\sum_{n=1}^\infty f_n$ is continuous. (iii) Take in (i) the measure $\mu=\sum_{n=1}^\infty 2^{-n}\delta_n$.

7.14.162. Prove that the product of a family of perfect probability measures is perfect. See also Exercise 9.12.70.

HINT: apply Theorem 7.5.6(ii) and Corollary 3.5.4.

7.14.163. (Pachl [**1413**]) Let $(X,\mathcal{A},\mu)$ be a probability space. Prove that μ has a compact approximating class if and only if μ is weakly compact in the sense of Erohin [**537**], i.e., there exists a family $\mathcal{U}$ of subsets of X that contains X and $\varnothing$ and is closed with respect to finite intersections and countable unions (such a family is called a σ-topology) and has the property that for every $\varepsilon>0$, there is a set $K_\varepsilon\subset X$ such that $X\backslash K_\varepsilon\in\mathcal{U}$, $\mu^*(K_\varepsilon)>1-\varepsilon$ and K_ε is $\mathcal{U}$-compact, i.e., every countable family of sets from $\mathcal{U}$ covering K_ε contains a finite subfamily covering K_ε.

7.14.164. Suppose that a Borel probability measure μ on a topological space X assumes only the values 0 and 1 and is $\aleph$-compact in the sense explained at the end of §7.5. Prove that μ is Dirac's measure at some point.

HINT: let $\mathcal{K}\subset\mathcal{B}(X)$ be an $\aleph$-compact approximating class for μ. The subclass $\mathcal{K}_0$ in $\mathcal{K}$ consisting of all sets of positive measure is $\aleph$-compact. The class $\mathcal{K}_0$ is not empty. Since all sets in $\mathcal{K}_0$ have measure 1, any finite intersection of such sets is not empty, whence we obtain that the intersection of all sets in $\mathcal{K}_0$ contains at least one point x. Then $\mu(\{x\})=1$, since otherwise $\mu(\{x\})=0$, which leads to a contradiction due to the existence of a set $K\in\mathcal{K}_0$ that is contained in $X\backslash\{x\}$.

7.14.165. A probability measure μ on a σ-algebra $\mathcal{A}$ in a space X is called pure (see Rao [**1537**]) if there exists a subalgebra $\mathcal{A}_0\subset\mathcal{A}$ such that

$$\mu(A)=\inf\Bigl\{\sum_{n=1}^\infty\mu(B_n)\colon\ B_n\in\mathcal{A}_0,\ A\subset\bigcup_{n=1}^\infty B_n\Bigr\}$$

for every $A\in\mathcal{A}$, and for every decreasing sequence of sets $A_n\in\mathcal{A}_0$ whose intersection is empty, there exists a number k such that $\mu(A_k)=0$. If there exists an algebra $\mathcal{A}_0\subset\mathcal{A}$ that is a compact class and satisfies the indicated equality, then the measure μ is called purely $\aleph_0$-compact.

(i) (Frolík, Pachl [**643**]) Prove that a probability measure μ on a countably generated σ-algebra $\mathcal{A}$ is pure if and only if it is compact. In particular, every pure measure is perfect.

(ii) (Aniszczyk [**54**]) Construct a measure with values in $\{0,1\}$ that is not pure.

7.14.166. (Krupa, Zięba [**1065**]) Let $(\Omega,\mathcal{B},P)$ be a probability space, X a Polish space, and let a sequence of measurable mappings $\xi_n\colon\ \Omega\to X$ converge a.e. to a mapping ξ. Prove that for every $\varepsilon>0$, there exists a compact set $K_\varepsilon\subset X$ such that $P\bigl(\bigcap_{n=1}^\infty\xi_n^{-1}(K_\varepsilon)\bigr)>1-\varepsilon$.

HINT: we may assume that X is a complete separable metric space with a metric d. Let us take a set $\Omega_\varepsilon \in \mathcal{B}$ such that $P(\Omega_\varepsilon) > 1 - \varepsilon/2$ and the mappings ξ_n converge uniformly on Ω_ε. There is a compact set $K_0 \subset X$ such that $P\bigl(\xi^{-1}(K_0)\bigr) > 1 - \varepsilon/4$. There exist strictly increasing numbers n_k with $d\bigl(\xi_n(\omega), \xi(\omega)\bigr) < 2^{-k}$ for all $n \geq n_k$ and $\omega \in \Omega_\varepsilon$. Finally, we can find a compact set K_1 such that $P\bigl(\xi_n^{-1}(K_1)\bigr) > 1 - \varepsilon/8$ whenever $n \leq n_1$, next we find a compact set K_2 in the 1/2-neighborhood of K_0 such that $P\bigl(\xi_n^{-1}(K_2)\bigr) > 1 - \varepsilon/16$ whenever $n_1 < n \leq n_2$ and so on: a compact set K_p is chosen by induction in the 2^{1-p}-neighborhood of K_0 in such a way that $P\bigl(\xi_n^{-1}(K_2)\bigr) > 1 - \varepsilon 2^{-p}/8$ whenever $n_{p-1} < n \leq n_p$. It remains to observe that the set $S := \bigcup_{n=0}^{\infty} K_n$ is completely bounded and $P\bigl(\bigcap_{n=1}^{\infty} \xi_n^{-1}(S)\bigr) > 1 - \varepsilon$. For K_ε we take the closure of S.

7.14.167. (Iwanik [**873**]) Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be probability spaces and let the measure μ be perfect. Prove that for every continuous linear operator $T\colon\, L^1(\mu) \to L^1(\nu)$, there is a countably additive measure σ on $\mathcal{A}\otimes\mathcal{B}$ such that

$$\int_Y g(y) Tf(y)\, \nu(dy) = \int_{X\times Y} f(x) g(y)\, \sigma(dxdy)$$

for all $\mathcal{A}$-measurable $f \in \mathcal{L}^1(\mu)$ and $\mathcal{B}$-measurable $g \in \mathcal{L}^\infty(\nu)$. In addition, the projections of σ on X and Y are absolutely continuous with respect to μ and ν, in particular, σ extends to $\mathcal{A}_\mu\otimes\mathcal{B}_\nu$.

7.14.168.° Let X and Y be Souslin spaces with Borel probability measures μ and ν, where ν has no atoms, and let π_X denote the projection on X. Suppose a set $E \subset X\times Y$ is measurable with respect to $\mu\otimes\nu$. Show that there is a set $Z \subset E$ such that $\mu\otimes\nu(Z) = 0$, $\pi_X(Z) = \pi_X(E)$, and for every $x \in \pi_X(Z)$, the section $\{y\colon\, (x, y) \in Z\}$ consists of a single point.

HINT: take a Borel set $A \subset E$ with $\mu\otimes\nu(A) = \mu\otimes\nu(E)$; by Corollary 6.9.17 there is a coanalytic set $S \subset A$ which is projected one-to-one on $\pi_X(A)$. By Fubini's theorem we have $\mu\otimes\nu(S) = 0$, since ν has no points of positive measure. The set $E' = E\backslash\bigl(\pi_X(A)\times Y\bigr) \subset E\backslash A$ has $\mu\otimes\nu$-measure zero, hence it is trivial to find a required set Z' for it. Let us set $Z = S \cup Z'$.

7.14.169.° Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\{\mu^\omega\}_{\omega\in\Omega}$ and $\{\nu^\omega\}_{\omega\in\Omega}$ be two families of probability measures on a measurable space $(X, \mathcal{A})$, such that for every $A \in \mathcal{A}$, the functions $\omega \mapsto \mu^\omega(A)$ and $\omega \mapsto \nu^\omega(A)$ are P-measurable. Let

$$\mu(A) := \int_\Omega \mu^\omega(A)\, P(d\omega), \quad \nu(A) := \int_\Omega \nu^\omega(A)\, P(d\omega).$$

Show that if $\mu \perp \nu$, then $\mu^\omega \perp \nu^\omega$ for P-a.e. ω. Show that the converse is false.

HINT: take $B \in \mathcal{A}$ with $\mu(B) = \nu(X\backslash B) = 0$; then we have $\mu^\omega(B) = \nu^\omega(X\backslash B) = 0$ for P-a.e. ω. In order to see that the converse is false, write Lebesgue measure λ on $[-1/2, 1/2]$ as the integral of the Dirac measures δ_ω, $\omega \in [-1/2, 1/2]$, with respect to λ and also as the integral of the measures $\delta_{-\omega}$, $\omega \in [-1/2, 1/2]$.

7.14.170. Show that there is no injective Borel function on the space $[0, \omega_1)$ of all countable ordinals equipped with the order topology with values in $[0, 1)$. No such such functions exist on $[0, 1]^{\mathfrak{c}}$.

HINT: if such a function f exists, then the image of the Dieudonné measure must be a Borel measure ν on $[0, 1)$ with values 0 and 1, hence ν is Dirac's measure at some point x_0, which is impossible for an injective function. The second claim follows since $[0, \omega_1)$ can be embedded into $[0, 1]^{\mathfrak{c}}$.

CHAPTER 8

Weak convergence of measures

> The linkage of general ideas exposed here arose, however, not by itself, but from the investigation of weak convergence of additive set-functions.
>
> *A.D. Alexandroff.* Additive set functions in abstract spaces.

8.1. The definition of weak convergence

Let $\{\mu_\alpha\}$ be a net (for example, a countable sequence) of finite measures defined on the Baire σ-algebra $\mathcal{B}a(X)$ of a topological space X. In this section, we introduce one of the most important modes of convergence of such nets. We recall that the space of all Baire measures on X is denoted by $\mathcal{M}_\sigma(X)$. Other notation frequently used in this chapter can be found in §§6.1, 6.2, 7.1, and 7.2.

8.1.1. Definition. *A net $\{\mu_\alpha\} \subset \mathcal{M}_\sigma(X)$ is called weakly convergent to a measure $\mu \in \mathcal{M}_\sigma(X)$ if for every bounded continuous real function f on X, one has*

$$\lim_\alpha \int_X f(x)\,\mu_\alpha(dx) = \int_X f(x)\,\mu(dx). \tag{8.1.1}$$

Notation: $\mu_\alpha \Rightarrow \mu$.

We shall say that a sequence of Baire measures μ_n on a space X is weakly fundamental if, for every bounded continuous function f on X, the sequence of the integrals

$$\int_X f\,d\mu_n$$

is fundamental (hence converges).

Weak convergence of Borel measures is understood as weak convergence of their Baire restrictions. In §8.10(iv) we discuss another natural convergence of Borel measures (convergence in the A-topology), which in the general case is not equivalent to weak convergence, but is closely related to it.

Weak convergence can be defined by a topology.

8.1.2. Definition. *Let X be a topological space. The weak topology on the space $\mathcal{M}_\sigma(X)$ of Baire measures on X is the topology $\sigma\bigl(\mathcal{M}_\sigma(X), C_b(X)\bigr)$, i.e., the base of the weak topology consists of the sets*

$$U_{f_1,\ldots,f_n,\varepsilon}(\mu) = \biggl\{\nu\colon\ \biggl|\int_X f_i\,d\mu - \int_X f_i\,d\nu\biggr| < \varepsilon,\ i = 1,\ldots,n\biggr\}, \tag{8.1.2}$$

where $\mu \in \mathcal{M}_\sigma(X)$, $f_i \in C_b(X)$, $\varepsilon > 0$. A set of such a form is called a fundamental neighborhood of μ in the weak topology.

In fact, the weak topology is the weak* topology in the terminology of functional analysis (however, following the tradition, we call it "*weak topology*"). Convergence in this topology is also called w^*-*convergence* (or *narrow convergence*). Random elements are called *convergent in distribution* if their distributions converge weakly.

8.1.3. Example. If a net of measures μ_α converges in the variation norm to a measure μ, then it weakly converges to μ. More generally, if there exists α_1 such that $\sup_{\alpha\geq\alpha_1}\|\mu_\alpha\| < \infty$, and $\lim\limits_{\alpha}\mu_\alpha(B) = \mu(B)$ for every $B \in \mathcal{B}a(X)$ or at least for every B of the form $B = \{f < c\}$, where $f \in C_b(X)$ and $|\mu|(\{f = c\}) = 0$, then $\mu_\alpha \Rightarrow \mu$.

PROOF. It suffices to prove the last assertion. Let $\|\mu_\alpha\| \leq C$, $\|\mu\| \leq C$, let $f \in C_b(X)$, and let $\varepsilon > 0$. We may assume that $|f| \leq 1$. We can find $c_i \in [-1, 1]$, $i = 1, \ldots, n$, such that $0 < c_{i+1} - c_i < \varepsilon$, $c_1 = -1$, $c_n = 1$ and $|\mu|(\{f = c_i\}) = 0$. Let $g(x) = c_i$ if $c_i \leq f(x) < c_{i+1}$. Then $|f(x) - g(x)| < \varepsilon$. For all indices α larger than some α_0 we have the estimate

$$\biggl|\int_X g\,d\mu_\alpha - \int_X g\,d\mu\biggr| < \varepsilon$$

because $\lim_\alpha \mu_\alpha(\{c_i \leq f < c_{i+1}\}) = \mu(\{c_i \leq f < c_{i+1}\})$ by our hypothesis and the equality $\{c_i \leq f < c_{i+1}\} = \{f < c_{i+1}\}\backslash\{f < c_i\}$. Hence for all $\alpha \geq \alpha_0$ the absolute value of the difference between the integrals of f with respect to the measures μ and μ_α does not exceed $(2C + 1)\varepsilon$. □

However, weak convergence does not imply convergence even on open Baire sets. The following simple example is very typical.

8.1.4. Example. Let p be a probability density on the real line and let ν_n be probability measures defined by the densities $p_n(t) = np(nt)$. Then the measures ν_n converge weakly to Dirac's measure δ at zero, although there is no convergence on $\mathbb{R}\backslash\{0\}$. Indeed, if $f \in C_b(R)$, then

$$\lim_{n\to\infty}\int_{-\infty}^{+\infty} f(t)p_n(t)\,dt = \lim_{n\to\infty}\int_{-\infty}^{+\infty} f(s/n)p(s)\,ds = f(0).$$

8.1.5. Example. A net $\{x_\alpha\}$ of elements of a completely regular space X converges to an element $x \in X$ if and only if the Dirac measures δ_{x_α} converge weakly to δ_x (we recall that $\delta_x(A) = 1$ if $x \in A$, $\delta_x(A) = 0$ if $x \notin A$).

A justification of this example is Exercise 8.10.66.

8.1.6. Example. (i) The set of all measures of the form $\sum_{j=1}^n c_j\delta_{x_j}$, where $c_j \in \mathbb{R}^1$, $x_j \in X$, is everywhere dense in $\mathcal{M}_\sigma(X)$ in the weak topology.

(ii) Let μ be a Borel measure on a separable Hilbert space X and let $P_n(x) = \sum_{i=1}^n (x, e_i)e_i$, where $\{e_n\}$ is an orthonormal basis. Then the measures $\mu\circ P_n^{-1}$ converge weakly to μ.

PROOF. (i) Suppose we are given a neighborhood U of the form (8.1.2). We may assume that $\|\mu\| \le 1$. There are simple functions g_i such that $\sup_x |f_i(x) - g_i(x)| < \varepsilon/4$ for all $i = 1, \ldots, n$. We show that U contains a measure ν of the required form with $\|\nu\| \le 1$. It suffices to find a finite linear combination ν of Dirac's measures such that $\|\nu\| \le 1$ and every g_i has equal integrals with respect to μ and ν. Now, given a finite partition of X into disjoint Baire sets A_i, $i = 1, \ldots, k$, everything reduces to finding points x_i and numbers c_i such that $\nu(A_i) = \mu(A_i)$. It remains to take a point x_i in every A_i and set $c_i := \mu(A_i)$. Assertion (ii) is obvious from the dominated convergence theorem, since $f\bigl(P_n(x)\bigr) \to f(x)$ for all continuous f. □

8.1.7. Proposition. *Let $\mathcal{M} \subset \mathcal{M}_\sigma(X)$ be a family of measures such that*

$$\sup_{\mu \in \mathcal{M}} \int_X f\, d\mu < \infty \quad \textit{for all } f \in C_b(X).$$

Then $\sup_{\mu \in \mathcal{M}} \|\mu\| < \infty$. In particular, every weakly convergent sequence of Baire measures is bounded in the variation norm.

PROOF. We apply the Banach–Steinhaus theorem and the fact that the variation of a Baire measure μ equals the norm of the functional on $C_b(X)$ generated by μ (see §7.9). □

The analogous assertion is true, of course, for complex-valued measures if we consider the absolute values of integrals (in the real case this gives an equivalent condition because in place of f one can take $-f$).

8.1.8. Proposition. *A sequence of signed measures μ_n on the interval $[a, b]$ converges weakly to a measure μ precisely when $\sup_n \|\mu_n\| < \infty$ and every subsequence in the sequence of the distribution functions F_{μ_n} of the measures μ_n contains a further subsequence convergent to F_μ at all points, with the exception of points of an at most countable set. In the case of nonnegative measures, the whole sequence F_{μ_n} converges to the function F_μ at all continuity points of the latter.*

An equivalent condition: $\sup_n \|\mu_n\| < \infty$ and for every closed interval $[c, d] \subset [a, b]$ and every $\varepsilon > 0$, there exists N such that

$$\inf_{t \in [c,d]} |F_\mu(t) - F_{\mu_n}(t)| < \varepsilon \quad \textit{for all } n \ge N.$$

In the case of measures on $\mathbb{R}^1$, the conditions listed above must be complemented by the following one: for every $\varepsilon > 0$ there is a compact interval $[a, b]$ such that $|\mu_n|(\mathbb{R}^1 \backslash [a, b]) < \varepsilon$ for all n.

PROOF. Suppose that the measures μ_n are uniformly bounded and satisfy the indicated condition with subsequences, but do not converge weakly to μ. Since every continuous function f can be uniformly approximated by smooth functions, we obtain, taking into account the boundedness of $\|\mu_n\|$, that there exists a smooth function f such that the integrals of f against the measures μ_n do not converge to the integral of f against μ. Passing to a subsequence, we may assume that the difference between the indicated integrals remains greater than some $\delta > 0$. Passing to a subsequence once again we can assume that $\lim\limits_{n\to\infty} F_{\mu_n}(t) = F_\mu(t)$ everywhere, with the exception of finitely or countably many points. The functions F_μ and F_{μ_n} are constant on $(b, +\infty)$, hence $\mu([a,b]) = \lim\limits_{n\to\infty} \mu_n([a,b])$. Then the integration by parts formula (see Exercise 5.8.112) yields that the right-hand side of the equality

$$\int_a^b f(t)\,\mu_n(dt) = f(b)F_{\mu_n}(b+) - \int_a^b f'(t)F_{\mu_n}(t)\,dt \tag{8.1.3}$$

converges to

$$f(b)F_\mu(b+) - \int_a^b f'(t)F_\mu(t)\,dt = \int_a^b f(t)\,\mu(dt),$$

which leads to a contradiction. In the case of nonnegative measures, the functions F_{μ_n} are increasing. Hence by Exercise 5.8.67, every subsequence in $\{F_{\mu_n}\}$ contains a subsequence convergent to F_μ at the continuity points of F_μ, whence we obtain convergence of the whole sequence at such points.

Conversely, let measures μ_n converge weakly to μ. Then, by the above we have $\sup_n \|\mu_n\| < \infty$. This yields the uniform boundedness of variations of the functions F_{μ_n}. Every subsequence in $\{F_{\mu_n}\}$ contains a further subsequence that converges at every point. Indeed, $F_{\mu_n} = \varphi_n - \psi_n$, where the functions φ_n and ψ_n are increasing. Hence we can apply Exercise 5.8.67. Thus, we may assume that the sequence F_{μ_n} converges pointwise to some function G. Now (8.1.3) and the equality

$$\lim_{n\to\infty} F_{\mu_n}(b+) = \lim_{n\to\infty} \mu_n([a,b]) = \mu([a,b]) = F_\mu(b+)$$

yield by weak convergence and the dominated convergence theorem that

$$\int_a^b f(t)\,\mu(dt) = f(b)F_\mu(b+) - \int_a^b f'(t)G(t)\,dt.$$

Hence

$$\int_a^b f'(t)G(t)\,dt = \int_a^b f'(t)F_\mu(t)\ dt$$

for every polynomial f. Hence $G(t) = F_\mu(t)$ a.e. Therefore, the functions G and F_μ coincide at all points where both are continuous, i.e., on the complement of an at most countable set (which depends on G, in particular, on the initial subsequence).

Let us turn to the second condition. If it is not fulfilled, then either our measures are not uniformly bounded and then there is no weak convergence, or there exist an interval $[c,d]$, a number $\varepsilon > 0$, and a subsequence $\{n_k\}$ with $|F_\mu(t) - F_{\mu_{n_k}}(t)| > \varepsilon$ for all $t \in [c,d]$, which contradicts the condition with subsequences. Conversely, suppose that the second condition is fulfilled. Since $C[a,b]$ is separable, every bounded sequence in $C[a,b]^*$ contains a weakly* convergent subsequence, i.e., every bounded sequence of measures contains a weakly convergent subsequence. Thus, if there is no weak convergence of μ_n to μ, then $\{\mu_n\}$ contains a subsequence that is weakly convergent to some measure ν on $[a,b]$ distinct from μ. According to what has already been proven, this subsequence contains a further subsequence with indices $\{n_k\}$ such that the functions $F_{\mu_{n_k}}$ converge to F_ν on the complement of an at most countable set. Passing to another subsequence we may assume that the sequences $\{\mu_{n_k}^+\}$ and $\{\mu_{n_k}^-\}$ have weak limits ν_1 and ν_2 and that $n_k = k$. The set of convergence of F_{μ_k} contains a point τ at which the functions F_μ, F_ν, F_{ν_1}, and F_{ν_2} are continuous and $|F_\mu(\tau_1) - F_\nu(\tau_1)| = \varepsilon > 0$ (otherwise $\mu = \nu$). There exist $\tau_2 \in [a,b]$ and $N \in \mathbb{N}$ such that we have $|F_\nu(t) - F_\mu(t)| > \varepsilon/2$ and $|F_{\nu_i}(t) - F_{\nu_i}(s)| \le \varepsilon/16$ whenever $t, s \in I = [\tau_1, \tau_2]$, $|F_{\mu_k^+}(\tau_i) - F_{\nu_1}(\tau_i)| \le \varepsilon/16$ and $|F_{\mu_k^-}(\tau_i) - F_{\nu_2}(\tau_i)| \le \varepsilon/16$ whenever $k \ge N$. Then $\sup_{t\in I} |F_{\mu_k^+}(t) - F_{\nu_1}(t)| \le 3\varepsilon/16$, $\sup_{t\in I} |F_{\mu_k^-}(t) - F_{\nu_2}(t)| \le 3\varepsilon/16$, hence $\sup_{t\in I} |F_{\mu_k}(t) - F_\nu(t)| \le 3\varepsilon/8$, i.e., $\inf_{t\in I} |F_\mu(t) - F_{\mu_k}(t)| \ge \varepsilon/8$ if $k \ge N$. The case of the whole real line is similar. $\square$

An alternative proof along with some useful similar results can be found in Exercise 8.10.135 (see also Exercise 8.10.137).

The reader is warned that in the case of signed measures weak convergence does not imply pointwise convergence of the distribution functions on a dense set (Exercise 8.10.69).

A.D. Alexandroff [**30**, §15] gave the following criterion of weak convergence. Let $\mathcal{Z}$ be the class of all functionally closed sets and let $\mathcal{G}$ be the class of all functionally open sets in a given space.

8.1.9. Theorem. *A sequence of Baire measures μ_n is fundamental in the weak topology precisely when it is bounded in the variation norm and for every $Z \in \mathcal{Z}$ and $U \in \mathcal{G}$ with $U \supset Z$ and every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n, k > N$ one has*

$$\inf\Bigl\{|\mu_n(V) - \mu_k(V)|\colon\ V \in \mathcal{G},\, Z \subset V \subset U\Bigr\} < \varepsilon.$$

In addition, weak convergence of μ_n to μ is equivalent to the following: $\{\mu_n\}$ is bounded in variation and for every $Z \in \mathcal{Z}$ and $U \in \mathcal{G}$ with $U \supset Z$ one has

$$\lim_{n\to\infty} \inf\Bigl\{|\mu_n(V) - \mu(V)|\colon\ V \in \mathcal{G},\, Z \subset V \subset U\Bigr\} = 0.$$

Finally, in the case of weak convergence of nonnegative measures, there exists $V \in \mathcal{G}$ with $Z \subset V \subset U$ and $\lim\limits_{n\to\infty} \mu_n(V) = \mu(V)$.

PROOF. Suppose that the sequence $\{\mu_n\}$ is fundamental in the weak topology. By Lemma 6.3.2, there exists a function $f \in C_b(X)$ such that $Z = f^{-1}(0)$, $X \backslash U = f^{-1}(1)$. In order to find a set $V \in \mathcal{G}$ with $Z \subset V \subset U$ and $|\mu_n(V) - \mu_k(V)| < \varepsilon$ for all sufficiently large n and k, one can take some of the sets $\{f < t\}$ with a suitable $t \in (0,1)$. This follows by the second condition in Proposition 8.1.8 and the fact that the measures $\mu_n \circ f^{-1}$ on $[0,1]$ form a fundamental sequence and hence converge weakly (we recall that the dual to $C[0,1]$ is identified with the space of measures). If the measures μ_n are nonnegative and the distribution functions of $\mu\circ f^{-1}$ are continuous at the point t, then Proposition 8.1.8 yields that $\lim\limits_{n\to\infty}\mu_n(V)=\mu(V)$.

Conversely, suppose that the condition of the theorem is fulfilled. We may assume that $\|\mu_n\|\le 1$. We observe that the sequence $\mu_n(X)$ converges, for one can take $Z=U=X$. Let $\varphi\in C_b(X)$, $0\le\varphi<1$, and let $\varepsilon=1/p$, where $p\in\mathbb{N}$. We consider the sets $U_j=\{\varphi<\varepsilon j\}$, $Z_j=\{\varphi\le\varepsilon(j-1)\}$, $j=1,\ldots,p$. By hypothesis, there exists N such that for every $j\le p$ and every $n,k>N$, there exist functionally open sets $V_{j,n,k}$ such that $Z_j\subset V_{j,n,k}\subset U_j$ and one has $|\mu_n(V_{j,n,k})-\mu_k(V_{j,n,k})|<\varepsilon p^{-2}$. One can also assume that $|\mu_n(X)-\mu_k(X)|<\varepsilon p^{-2}$ for all $n,k>N$. For all fixed n and k, the sets

$$W_{1,n,k}:=V_{1,n,k},W_{2,n,k}:=V_{2,n,k}\backslash V_{1,n,k},\ldots,W_{p+1,n,k}:=X\backslash V_{p,n,k}$$

form a partition of X. It is easily seen that the values of the measures μ_n and μ_k on these sets differ in absolute value in at most ε/p; for example, we have $|\mu_n(V_{1,n,k})-\mu_k(V_{1,n,k})|<\varepsilon p^{-2}$,

$$|\mu_n(V_{2,n,k}\backslash V_{1,n,k})-\mu_k(V_{2,n,k}\backslash V_{1,n,k})|<2\varepsilon p^{-2},$$

and so on. It remains to observe that

$$\biggl|\int_X\varphi\,d\mu_n-\sum_{j=1}^{p+1}(j-1)p^{-1}\mu_n(W_{j,n,k})\biggr|\le\varepsilon$$

and that $\Bigl|\sum_{j=1}^{p+1}(j-1)p^{-1}\bigl(\mu_n(W_{j,n,k})-\mu_k(W_{j,n,k})\bigr)\Bigr|\le\varepsilon(p+1)/p$. The assertion about convergence to μ is proved in a similar way. □

A very important property of weak convergence is described in the following result due to A.D. Alexandroff [**30**, §18].

8.1.10. Proposition. *Suppose that a sequence of Baire measures μ_n on a topological space X converges weakly to a measure μ. Then this sequence has no "eluding load" in Alexandroff's sense, i.e., $\lim\limits_{n\to\infty}\sup_k|\mu_k|(Z_n)=0$ for every sequence of pairwise disjoint functionally closed sets Z_n with the property that the union of every subfamily in $\{Z_n\}$ is functionally closed.*

Proof. Suppose the contrary. Taking a subsequence, we may assume that $|\mu_n(Z_n)|\ge c>0$. By Exercise 6.10.79, there exist pairwise disjoint functionally open sets U_n with $Z_n\subset U_n$ and $|\mu_n|(U_n\backslash Z_n)\le c/2$. Let us show that there exist functions $f_n\in C_b(X)$ with the following properties: $0\le f_n\le 1$, $f_n=0$ outside U_n,

$$\biggl|\int_X f_n\,d\mu_n\biggr|\ge c/2, \qquad (8.1.4)$$

and for every bounded sequence $\{c_n\}$ of real numbers, the function $\sum_{n=1}^\infty c_nf_n$ is bounded and continuous. This will lead to a contradiction. Indeed, by our hypothesis the sequence of integrals of such a function with respect to the measures μ_n is convergent, i.e., the sequence

$$l_n := \Bigl\{ \int f_k \, d\mu_n \Bigr\}_{k=1}^{\infty}$$

of elements in l^1 is weakly convergent, which contradicts (8.1.4) by Corollary 4.5.8. In order to construct the required functions f_n, we take (applying Lemma 6.3.2) a continuous function f such that $0 \le f \le 1$, $f = 1$ on $\bigcup_{n=1}^{\infty} Z_n$ and $f = 0$ outside $\bigcup_{n=1}^{\infty} U_n$. Set $f_n = f$ on U_n and $f_n = 0$ outside U_n. The nonnegative function f_n is continuous, since for every $c \ge 0$, we have

$$\{f_n > c\} = \{f > c\} \cap U_n, \quad \{f_n < c\} = \{f < c\} \bigcup \Bigl(\bigcup_{k \ne n} U_k \Bigr),$$

and the sets on the right-hand side are open. For the same reason we have the continuity of any function $h = \sum_{n=1}^{\infty} c_n f_n$, $|c_n| < 1$ because we have

$$\{h > c\} = \bigcup_{n\colon\, c_n > 0} \bigl(U_n \cap \{f > c/c_n\} \bigr), \ c \ge 0,$$

$$\{h > c\} = \bigcup_{n\colon\, c_n < 0} \bigl(U_n \cap \{f < c/c_n\} \bigr) \cup \{f < |c|\} \cup \Bigl(\bigcup_{n\colon\, c_n \ge 0} U_n \Bigr), \ c < 0.$$

Similarly, one proves that the sets $\{h < c\}$ are open. $\square$

8.1.11. Remark. A.D. Alexandroff [**30**, §17] introduced the following terminology. A set M of Borel measures on a normal topological space X has an eluding load equal to the number $a \ne 0$ if M contains an infinite sequence of measures μ_n such that for some sequence of pairwise disjoint functionally closed sets Z_n with the property that the union of every subfamily in $\{Z_n\}$ is functionally closed (such sequences are called by Alexandroff *divergent*), we have $\mu_n(Z_n)/a \ge 1$. If for some $a \ne 0$ the set M has an eluding load equal to a, then we say that M has an eluding load. It is clear that the absence of eluding load is equivalent to that $\lim\limits_{n\to\infty} \sup_{\mu\in M} |\mu|(Z_n) = 0$ for every divergent sequence of functionally closed sets Z_n. Indeed, if $|\mu_n|(Z_n) \ge a > 0$, then, taking a subsequence, we may assume that $\mu_n^+(Z_n) \ge a/2$ (otherwise we have $\mu_n^-(Z_n) \le -a/2$). Then, there exist functionally closed sets $F_n \subset Z_n \cap X_n^+$, where $X = X_n^+ \cup X_n^-$ is the Hahn decomposition for μ_n, such that one has $\mu_n(F_n) \ge a/4$. It remains to observe that the sequence F_n is divergent as well. The condition on the sets Z_n used above coincides with Alexandroff's condition for normal spaces (see Exercise 6.10.79).

The next result is due to A.D. Alexandroff [**30**, §18].

8.1.12. Proposition. *A family $\mathcal{M}$ of Baire measures on a topological space X has no eluding load precisely when for every sequence of functionally closed sets Z_n with $Z_n \downarrow \varnothing$, one has*

$$\lim_{n\to\infty} \sup_{\mu\in\mathcal{M}} |\mu|(Z_n) = 0. \tag{8.1.5}$$

PROOF. Suppose that $\mathcal{M}$ has no eluding load and Z_n are decreasing functionally closed sets with empty intersection. If (8.1.5) is not fulfilled, then, taking a subsequence, we may assume that we are given measures $\mu_n \in \mathcal{M}$ with $|\mu_n|(Z_n) > c > 0$. Taking a subsequence once again, we reduce everything to the case where $\mu_n^+(Z_n) > c/2$. In view of Exercise 6.10.80, there exist decreasing functionally open sets G_n with empty intersection and $Z_n \subset G_n$. We can find n_1 with $\mu_1^+(G_{n_1}) < c/4$. Then $\mu_1^+(Z_1 \backslash G_{n_1}) > c/4$. Next we find $n_2 > n_1$ with $\mu_{n_1}^+(G_{n_2}) < c/4$, whence $\mu_{n_1}^+(Z_{n_1} \backslash G_{n_2}) > c/4$. By induction, we obtain strictly increasing numbers n_k with $\mu_{n_k}^+(Z_{n_k} \backslash G_{n_{k+1}}) > c/4$. By the definition of $\mu_{n_k}^+$, for every k, one can find a functionally closed set F_k in $Z_{n_k} \backslash G_{n_{k+1}}$ such that $\mu_{n_k}(F_k) > \mu_{n_k}^+(Z_{n_k} \backslash G_{n_{k+1}}) - c/8 > c/8$. By assertion (ii) in Exercise 6.10.80 the sets $Z_{n_k} \backslash G_{n_{k+1}}$, hence also the sets F_k, form a divergent sequence. Thus, $\mathcal{M}$ has an eluding load, which is a contradiction.

Let $\mathcal{M}$ have an eluding load. Then, there exist a divergent sequence of functionally closed sets F_n, measures $\mu_n \in \mathcal{M}$, and a number $a \neq 0$ with $\mu_n(F_n)/a \geq 1$. The sets $Z_n := \bigcup_{k=n}^{\infty} F_k$ are functionally closed and decrease to the empty set; in addition, one has $|\mu_n|(Z_n) \geq |\mu_n(F_n)| \geq |a|$. $\square$

We discuss below many other properties of weak convergence of measures, but it is worth noting already now that, excepting trivial cases, the weak topology on the space of signed measures on X is not metrizable (for example, if X is an infinite metric space, see Exercise 8.10.72). It may occur, yet, that although the weak topology on $\mathcal{M}_\sigma(X)$ is not metrizable, but there is a metric on $\mathcal{M}_\sigma(X)$ in which convergence of *sequences* is equivalent to weak convergence. For example, this is the case if $X = \mathbb{N}$ with the usual metric (Exercise 8.10.68). It will be shown later that for any separable metric space X, the weak topology is metrizable on the set $\mathcal{M}_\sigma^+(X)$ of nonnegative measures.

8.2. Weak convergence of nonnegative measures

A base of the weak topology on the set of probability measures can be defined by means of values on certain sets. Let us consider the following two classes of sets in the space $\mathcal{P}_\sigma(X)$ of Baire probability measures:

$$W_{F_1,\ldots,F_n,\varepsilon}(\mu) = \Bigl\{\nu \in \mathcal{P}_\sigma(X)\colon\ \nu(F_i) < \mu(F_i) + \varepsilon,\ i = 1,\ldots,n\Bigr\},$$
$$F_i = f_i^{-1}(0),\ f_i \in C(X),\ \varepsilon > 0, \qquad (8.2.1)$$

$$W_{G_1,\ldots,G_n,\varepsilon}(\mu) = \Bigl\{\nu \in \mathcal{P}_\sigma(X)\colon\ \nu(G_i) > \mu(G_i) - \varepsilon,\ i = 1,\ldots,n\Bigr\},$$
$$G_i = X \backslash f_i^{-1}(0),\ f_i \in C(X),\ \varepsilon > 0. \qquad (8.2.2)$$

We recall that in the case of a metrizable space, the F_i's represent arbitrary closed sets and the G_i's represent arbitrary open sets.

8.2.1. Theorem. *The above-mentioned bases generate the weak topology on the set of probability measures $\mathcal{P}_\sigma(X)$.*

PROOF. The coincidence of the bases (8.2.1) and (8.2.2) is obvious from the defining formulas. Let U be a neighborhood of the form (8.1.2). We may assume that $0 < f_i < 1$. Let us fix $k \in \mathbb{N}$ with $k^{-1} < \varepsilon/4$. For every $i = 1, \ldots, n$, there exist points $c_{i,j} \in [0,1]$ such that $0 = c_{i,0} < \cdots < c_{i,m} = 1$, $c_{i,j+1} - c_{i,j} < \varepsilon/4$ and $\mu\bigl(f_i^{-1}(c_{i,j})\bigr) = 0$. Set $A_{i,1} = \{0 \le f_i < c_{i,1}\}$, $\ldots$, $A_{i,m} = \{c_{i,m-1} \le f_i < c_{i,m}\}$. Let us show that there is a neighborhood V of the form (8.2.1) such that for all i, j and $\nu \in V$, we have the estimate $|\mu(A_{i,j}) - \nu(A_{i,j})| < \delta$, where $\delta = (4m)^{-1}\varepsilon$. Then we shall have $V \subset U$ by the inequality $\Bigl|f_i - \sum_{j=1}^m c_{i,j} I_{A_{i,j}}\Bigr| < \varepsilon/4$. Indeed,

$$\biggl|\int_X f_i\, d\mu - \int_X f_i\, d\nu\biggr| \le \sum_{j=1}^m c_{i,j}|\mu(A_{i,j}) - \nu(A_{i,j})| + \varepsilon/2 < \varepsilon.$$

The required neighborhood V can be taken as the intersection of the neighborhood V_1 of the form (8.2.1), where we take the functionally closed sets $F_{i,j} = \{c_{i,j-1} \le f_i \le c_{i,j}\}$, $i \le n$, $j \le k$, and δ in place of ε, and the analogous neighborhood V_2 of the form (8.2.2), where we take the functionally open sets $G_{i,j} = \{c_{i,j-1} < f_i < c_{i,j}\}$. It is clear that

$$\nu(A_{i,j}) \ge \nu(G_{i,j}) > \mu(G_{i,j}) - \delta = \mu(A_{i,j}) - \delta$$

for all $\nu \in V_2$. Similarly, one has $\nu(A_{i,j}) < \mu(A_{i,j}) + \delta$ for all $\nu \in V_1$.

Let us show that every neighborhood of the form (8.2.1) contains a neighborhood in the weak topology. It suffices to consider neighborhoods defined by a single closed set F_1. We can assume that $F_1 = f_1^{-1}(0)$, where $0 \le f_1 \le 1$. Let us find $c > 0$ such that $\mu\bigl(\{0 < f < c\}\bigr) < \varepsilon/2$. Let ζ be a continuous function on the real line, $\zeta(t) = 1$ if $t \le 0$, $\zeta(t) = 0$ if $t \ge c$ and $0 < \zeta(t) < 1$ if $t \in (0, c)$. Set $f = \zeta \circ f_1$. It remains to observe that $\nu(F_1) < \mu(F_1) + \varepsilon$ if

$$\int_X f\, d\nu < \int_X f\, d\mu + \varepsilon/2.$$

Indeed,

$$\nu(F_1) \le \int_X f\, d\nu,$$

since $f = 1$ on F_1. On the other hand,

$$\int_X f\, d\mu \le \mu\bigl(f^{-1}(1)\bigr) + \mu\bigl(\{0 < f < 1\}\bigr) = \mu(F_1) + \mu\bigl(\{0 < f < c\}\bigr),$$

which is less than $\mu(F_1) + \varepsilon/2$. $\square$

8.2.2. Remark. A similar reasoning shows that the neighborhoods of the form (8.2.1) or (8.2.2) together with the neighborhoods

$$\bigl\{\nu\colon\ |\mu(X) - \nu(X)| < \varepsilon\bigr\}$$

form a base of the weak topology in the space of all nonnegative Baire measures $\mathcal{M}_\sigma^+(X)$.

We observe that the closed set $\{0\}$ in Example 8.1.4 has measure zero with respect to every measure ν_n, but is a full measure set for δ, whereas the situation with the open set $\mathbb{R}\backslash\{0\}$ is the opposite. Thus, there is no convergence on sets, but for every Borel set B whose boundary does not contain zero, one has $\nu_n(B) \to \delta(B)$. We shall see below that this example is typical. Having it in mind, one can easily remember the formulation of the following classical theorem of A.D. Alexandroff on weak convergence (see [**30**]), which is an immediate corollary of Theorem 8.2.1.

Given a net of numbers $(c_\alpha)_{\alpha\in\Lambda}$, the quantity $\limsup_\alpha c_\alpha$ is defined as the supremum of numbers c such that for every $\alpha_0 \in \Lambda$, there exists $\alpha > \alpha_0$ with $c_\alpha \geq c$; $\liminf_\alpha c_\alpha := -\limsup_\alpha -c_\alpha$. We note that even for countable nets, these quantities may differ from the upper and lower limits of the set of numbers c_α because the set $\{\alpha < \alpha_0\}$ may be infinite.

8.2.3. Theorem. *Suppose we are given a topological space X, a net of Baire probability measures $\{\mu_\alpha\}$, and a Baire probability measure μ on X. Then the following conditions are equivalent*:

(i) *the net $\{\mu_\alpha\}$ converges weakly to μ*;

(ii) *for every functionally closed set F one has*

$$\limsup_\alpha \mu_\alpha(F) \leq \mu(F); \tag{8.2.3}$$

(iii) *for every functionally open set U one has*

$$\liminf_\alpha \mu_\alpha(U) \geq \mu(U). \tag{8.2.4}$$

In the case of not necessarily probability measures μ_α, $\mu \in \mathcal{M}_\sigma^+(X)$, condition (i) *is equivalent to either of conditions* (ii) *and* (iii) *complemented by the condition $\lim\limits_\alpha \mu_\alpha(X) = \mu(X)$.*

Since a Baire measure may fail to have a Borel extension (or may have several Borel extensions), the discussion of relationships (8.2.3) and (8.2.4) for arbitrary closed sets F and open sets U requires additional assumptions. Certainly, no additional conditions are needed if all closed sets are functionally closed (i.e., if X is perfectly normal).

8.2.4. Corollary. (a) *If X is metrizable* (*or at least is perfectly normal*), *then condition* (i) *is equivalent to condition* (ii) *for every closed set F and condition* (iii) *for every open set U. The same is true if X is completely regular, the measures μ_α are Borel and the measure μ is τ-additive* (*for example, is Radon*).

(b) *If in Theorem* 8.2.3 *the space X is completely regular and the limit measure μ is τ_0-additive, then condition* (i) *implies condition* (ii) *for all closed Baire sets F* (*not necessarily functionally closed*) *and condition* (iii) *for all open Baire sets U. In particular, this is true if the measure μ is tight.*

PROOF. The first claim in (a) is obvious. The second one follows by the fact that in the case of a completely regular space X, the value of a τ-additive measure μ on every open set U is the supremum of measures of functionally open sets inscribed in U. For the proof of assertion (b), it suffices to apply Theorem 7.3.2 on the existence of a τ-additive extension of the measure μ and assertion (a). $\square$

8.2.5. Corollary. *Suppose that a net of Borel probability measures μ_α on a completely regular space X converges weakly to a Borel probability measure μ that is τ-additive (for example, is Radon). If f is a bounded upper semicontinuous function, then*

$$\limsup_{\alpha} \int_X f\,d\mu_\alpha \le \int_X f\,d\mu.$$

If f is a bounded lower semicontinuous function, then

$$\liminf_{\alpha} \int_X f\,d\mu_\alpha \ge \int_X f\,d\mu.$$

PROOF. We may assume that $0 < f < 1$. For every fixed n, let us set $U_k := \{x\colon\ f(x) > k/n\}$, $k = 1, \ldots, n$. In the case of a lower semicontinuous function f the sets U_k are open. Hence, letting $f_n := n^{-1} \sum_{k=1}^n I_{U_k}$, we have

$$\liminf_{\alpha} \int_X f_n\,d\mu_\alpha \ge \int_X f_n\,d\mu.$$

It remains to observe that $|f(x) - f_n(x)| \le n^{-1}$ for all $x \in X$. Indeed, if $m/n < f(x) \le (m+1)/n$, where $m \ge 1$, then $I_{U_k}(x) = 1$ for all $k \le m$ and $I_{U_k}(x) = 0$ for all $k > m$, whence $f_n(x) = m/n$. If $m = 0$, then $f_n(x) = 0$. $\square$

In general, weak convergence of measures does not yield any reasonable convergence of their densities with respect to a common dominating measure (see, e.g., Exercise 8.10.70). Note, however, the following simple fact.

8.2.6. Example. Suppose that Baire probability measures μ_n on a topological space X converge weakly to a Baire probability measure μ. Let ν be a Baire probability on X such that the measures μ_n and μ are absolutely continuous with respect to ν, i.e., $\mu_n = \varrho_n \cdot \nu$, $\mu = \varrho \cdot \nu$. Then the functions $\varrho_n I_{\{\varrho=0\}}$ converge to zero in measure ν.

In particular, if a Baire probability measure λ on X is mutually singular with μ, then the densities of the absolutely continuous parts of μ_n with respect to λ converge to zero in measure λ.

PROOF. Given $\varepsilon > 0$, we find a functionally closed set $F \subset E := \{\varrho = 0\}$ with $\nu(E\backslash F) < \varepsilon$. Since $\mu(F) = 0$, we have $\mu_n(F) \to 0$, i.e., $\|\varrho_n I_F\|_{L^1(\nu)} \to 0$. Hence $\varrho_n I_F \to 0$ in measure ν, which proves the first claim. The second claim follows by choosing ν such that $\mu_n \ll \nu$, $\mu \ll \nu$ and $\lambda \ll \nu$. $\square$

One can see from Theorem 8.2.1 that weak convergence ensures convergence on certain "sufficiently regular" sets (see also Example 8.1.3). Let us discuss this phenomenon in greater detail.

8.2.7. Theorem. *A net $\{\mu_\alpha\}$ of Baire probability measures on a topological space X converges weakly to a Baire probability measure μ if and only if the equality*

$$\lim_\alpha \mu_\alpha(E) = \mu(E) \tag{8.2.5}$$

is fulfilled for every set $E \in \mathcal{B}a(X)$ with the following property: there exist a functionally open set W and a functionally closed set F such that $W \subset E \subset F$ and $\mu(F\backslash W) = 0$.

PROOF. In the case of weak convergence we have

$$\limsup_\alpha \mu_\alpha(E) \le \limsup_\alpha \mu_\alpha(F) \le \mu(F) = \mu(E).$$

Similarly, $\liminf_\alpha \mu_\alpha(E) \ge \mu(E)$, whence we obtain (8.2.5). Suppose now that we have (8.2.5). Let $U = \{f > 0\}$, where $f \in C(X)$, and let $\varepsilon > 0$. It is easily seen that there exist $c > 0$ such that one has $\mu(U) < \mu\bigl(\{f > c\}\bigr) + \varepsilon$ and $\mu\bigl(\{f > c\}\bigr) = \mu\bigl(\{f \ge c\}\bigr)$. Then the set $E = \{f > c\}$ satisfies (8.2.5), since one can take the sets $W = E$ and $F = \{f \ge c\}$, the first of which is functionally open and the second one is functionally closed. Thus, we have the inequality $\liminf_\alpha \mu_\alpha(U) \ge \mu(U) - \varepsilon$, which yields (8.2.4) because ε is arbitrary. □

It is clear that in the case where X is a metric space, the sets E with the foregoing property are exactly the sets with the boundaries of μ-measure zero. Let us formulate an analogous assertion in the case of Borel measures. Let μ be a nonnegative Borel measure on a topological space X. Denote by Γ_μ the class of all Borel sets $E \subset X$ with boundaries of μ-measure zero. The boundary ∂E of any set E is defined as the closure of E without the interior of E, hence is a Borel set for arbitrary E. The sets in Γ_μ are called the continuity sets of μ or μ-continuity sets.

8.2.8. Proposition. (i) *Γ_μ is a subalgebra in $\mathcal{B}(X)$.*
(ii) *If X is completely regular, then Γ_μ contains a base of the topology of X.*

PROOF. Claim (i) follows from the fact that E and $X\backslash E$ have a common boundary, and the boundary of the union of two sets is contained in the union of their boundaries. In order to prove (ii), given a bounded continuous function f on X, we set $U(f,c) = \{x\colon\ f(x) > c\}$ and observe that the set

$$M_f = \Bigl\{c \in \mathbb{R}\colon\ \mu\bigl(\partial U(f,c)\bigr) > 0\Bigr\}$$

is at most countable, since $\partial U(f,c) \subset f^{-1}(c)$ and the measure $\mu \circ f^{-1}$ has at most countably many atoms. The sets $U(f,c)$, $c \in \mathbb{R}\backslash M_f$, belong to the class Γ_μ. By the complete regularity of X these sets form a base of the topology. Indeed, for every point x and every open set U containing x, there exists a continuous function $f\colon\ X \to [0,1]$ with $f(x) = 1$ that equals 0 outside U. Thus, U contains the set $U(f,c)$ for some $c \in \mathbb{R}\backslash M_f$. □

8.2.9. Theorem. *Let $\{\mu_\alpha\}$ be a net of Borel probability measures on a topological space X and let μ be a Borel probability measure on X. Then the following assertions are true.* (i) *If we have*

$$\lim_\alpha \mu_\alpha(E) = \mu(E) \quad \textit{for all} \quad E \in \Gamma_\mu, \tag{8.2.6}$$

then the net $\{\mu_\alpha\}$ converges weakly to μ.

(ii) *Let X be completely regular. If the net $\{\mu_\alpha\}$ converges weakly to μ and μ is τ-additive, then one has* (8.2.6). *If X is metrizable* (*or at least perfectly normal*), *then the τ-additivity of μ is not required.*

Proof. In order to prove (i) it suffices to observe that any set E with the property indicated in Theorem 8.2.7 is contained in Γ_μ. Assertion (ii) follows by Corollary 8.2.4 and the arguments used in the proof of Corollary 8.2.7. It is worth noting that for weak convergence of signed measures relation (8.2.6) is sufficient, but not necessary (Example 8.1.3 and Exercise 8.10.69). $\square$

An immediate corollary of the above results is the following assertion.

8.2.10. Corollary. *Let X be metrizable* (*or at least perfectly normal*). *Then the following conditions are equivalent*:

(i) *a net $\{\mu_\alpha\}$ of Borel probability measures converges weakly to a Borel probability measure μ*;

(ii) $\limsup\limits_\alpha \mu_\alpha(F) \le \mu(F)$ *for every closed set F*;

(iii) $\liminf\limits_\alpha \mu_\alpha(U) \ge \mu(U)$ *for every open set U*;

(iv) $\lim\limits_\alpha \mu_\alpha(E) = \mu(E)$ *for every set $E \in \Gamma_\mu$.*

These conditions remain equivalent for an arbitrary completely regular space X if the measure μ is τ-additive (*for example, is Radon*).

8.2.11. Corollary. *A net $\{\mu_\alpha\}$ of probability measures on the real line converges weakly to a probability measure μ precisely when the corresponding distribution functions F_{μ_α} converge to the distribution function F_μ of the measure μ at the points of continuity of F_μ, where $F_\mu(t) = \mu\bigl((-\infty, t)\bigr)$.*

Proof. The necessity of the foregoing condition follows by assertion (iv) of the previous corollary because the boundary of $(-\infty, t)$, i.e., the point t, has μ-measure zero if the function F_μ is continuous at this point. The sufficiency is clear from representation (8.2.2) of neighborhoods of the measure μ. Indeed, given $\varepsilon > 0$ and open sets $G_1, \ldots, G_n$ on the real line, one can find open sets $G_1', \ldots, G_n'$ consisting of finite collections of intervals with the endpoints at the continuity points of F_μ such that $G_i' \subset G_i$ and $\mu(G_i') > \mu(G_i) - \varepsilon/2$, $i = 1, \ldots, n$. Then the neighborhood (8.2.2) contains the measure μ_α for all α such that $\mu_\alpha(G_i') > \mu(G_i') - \varepsilon/2$, i.e., for all α greater than some index, since the net $\{\mu_\alpha(G_i')\}$ converges to $\mu(G_i')$. $\square$

8.2.12. Example. Suppose that Borel probability measures μ_n on $\mathbb{R}^d$ converge weakly to a Borel probability measure μ that is absolutely continuous. Then $\lim\limits_{n\to\infty} \mu_n(E) = \mu(E)$ for every Jordan measurable Borel set E. In particular, if μ_n and μ are Borel probability measures on $[0,1]$ such that μ is absolutely continuous, and $\lim\limits_{n\to\infty} \mu_n([a,b]) = \mu([a,b])$ for every interval $[a,b]$, then convergence holds on every Jordan measurable Borel set.

The last assertion in the case of absolutely continuous measures μ_n was proved by Fichtenholz [**575**], who also constructed an example when there is no convergence on some Borel set E. The existence of such an example is easily derived from the basic properties of weak convergence. Namely, let $E \subset [0,1]$ be a nowhere dense compact set of positive Lebesgue measure. It is clear from the previous results that one can find probability measures ν_n on $[0,1]$ that converge weakly to Lebesgue measure λ on $[0,1]$ and are concentrated on finite sets in the complement of E. Hence one can find probability measures μ_n that converge weakly to λ and are given by smooth densities vanishing on E (it suffices to take such a measure μ_n in the ball of radius $1/n$ and center ν_n with respect to the metric determining weak convergence, which is discussed in the next section).

One more sufficient condition of weak convergence in terms of convergence on certain sets is given in the following theorem from Prohorov [**1497**].

8.2.13. Theorem. *Let $\mathcal{E}$ be a class of Baire sets in a topological space X such that $\mathcal{E}$ is closed with respect to finite intersections and every functionally open set is representable as a finite or countable union of sets from $\mathcal{E}$. Suppose that μ and μ_n, where $n \in \mathbb{N}$, are Baire probability measures on X such that $\mu_n(E) \to \mu(E)$ for all $E \in \mathcal{E}$. Then $\{\mu_n\}$ converges weakly to μ. The analogous assertion is true for Radon (or τ-additive) measures and Borel sets.*

PROOF. We observe that

$$\lim_{n\to\infty} \mu_n\Bigl(\bigcup_{j=1}^k E_j\Bigr) = \mu\Bigl(\bigcup_{j=1}^k E_j\Bigr) \quad \text{for all } E_1, \ldots, E_k \in \mathcal{E}.$$

Indeed, if $k = 2$, then by hypothesis we have convergence on E_1, E_2 and $E_1 \cap E_2$, which yields convergence on $E_1 \backslash (E_1 \cap E_2)$ and $E_2 \backslash (E_1 \cap E_2)$, hence also on the set $E_1 \cup E_2$ that equals the disjoint union of $E_1 \cap E_2$, $E_1 \backslash (E_1 \cap E_2)$, and $E_2 \backslash (E_1 \cap E_2)$. By induction on k we obtain our assertion. Indeed, if it is true for some k, then it is true for $k+1$, since the set $(E_1 \cup \ldots \cup E_k) \cap E_{k+1}$ is the union of the sets $E_i \cap E_{k+1} \in \mathcal{E}$, $i = 1, \ldots, k$, which gives convergence on this set. Suppose we are given a set $U = \{f > 0\}$, where $f \in C_b(X)$. It can be represented as an at most countable union of sets $E_j \in \mathcal{E}$. Hence one has

$$\mu(U) = \lim_{k\to\infty} \mu\Bigl(\bigcup_{j=1}^k E_j\Bigr) = \lim_{k\to\infty} \lim_{n\to\infty} \mu_n\Bigl(\bigcup_{j=1}^k E_j\Bigr) \le \liminf_{n\to\infty} \mu_n(U),$$

whence the assertion follows. □

It is easily seen from the proof that this theorem remains valid if U is representable as an at most countable union of sets from $\mathcal{E}$ up to a set of μ-measure zero (Exercise 8.10.78).

Let us consider examples of classes $\mathcal{E}$ satisfying the hypotheses of this theorem.

8.2.14. Corollary. *Let $\mathcal{E}$ be some class of Borel sets in a separable metric space X such that $\mathcal{E}$ is closed with respect to finite intersections. Suppose that for every point $x \in X$ and every neighborhood U of x, one can find a set $E_x \in \mathcal{E}$ containing some neighborhood of the point x and contained in U. Then convergence of a sequence of Borel probability measures on all sets in $\mathcal{E}$ yields its weak convergence. The same is true if X is completely regular and hereditary Lindelöf.*

PROOF. Let U be open. By hypothesis, for every point $x \in U$, there exists $E_x \in \mathcal{E}$ such that $x \in E_x \subset U$, and x has a neighborhood $V_x \subset E_x$. By the separability of X, the cover of U by the sets V_x contains an at most countable subcover $\{V_{x_n}\}$, which means that $U = \bigcup_{n=1}^{\infty} E_{x_n}$. The second claim is proven by the same reasoning. □

8.2.15. Corollary. *Let X be a separable metric space and let μ and μ_n, where $n \in \mathbb{N}$, be Borel probability measures on X such that $\mu_n(E) \to \mu(E)$ for every set E that is a continuity set for μ (i.e., $E \in \Gamma_\mu$) and is representable as a finite intersection of open balls. Then $\mu_n \Rightarrow \mu$.*

PROOF. The family of sets with the indicated properties satisfies the hypotheses of the previous corollary. Indeed, finite intersections of such sets are continuity sets as well. In addition, for every point x and every $\varepsilon > 0$, there exists $r \in (0, \varepsilon)$ such that the boundary of the ball of radius r centered at x has μ-measure zero because for different r these boundaries have empty intersections (note that the boundary of the ball of radius r is contained in the sphere of radius r with the same center). □

8.2.16. Example. A sequence $\{\mu_n\}$ of Borel probability measures on $\mathbb{R}^\infty$ converges weakly to a Borel probability measure μ if and only if the finite-dimensional projections of the measures μ_n, i.e., the images of μ_n under the projections $\pi_d\colon \mathbb{R}^\infty \to \mathbb{R}^d$, $(x_i) \mapsto (x_1, \ldots, x_d)$, converge weakly to the corresponding projections of μ for every fixed d.

PROOF. The necessity of weak convergence of projections is obvious. Its sufficiency follows by Corollary 8.2.14 applied to the class of open cylinders of the form

$$C = \{x\colon (x_1, \ldots, x_d) \in U\}, \quad \text{where } U \subset \mathbb{R}^d \text{ is open},$$

with boundaries of μ-measure zero. The equality $\lim\limits_{n\to\infty} \mu_n(C) = \mu(C)$ follows by the equality $\mu \circ \pi_d^{-1}(\partial U) = \mu(\partial C) = 0$. □

A generalization of Theorem 8.2.13 given in Exercise 8.10.78 yields the following result due to Kolmogorov and Prohorov [**1034**]. However, we shall give a simple direct proof.

8.2.17. Theorem. *Let $\{\mu_\alpha\}$ be a net of Borel probability measures on a topological space X and let μ be a τ-additive probability measure on X. Suppose that the equality $\lim\limits_\alpha \mu_\alpha(U) = \mu(U)$ is fulfilled for all elements U of some base $\mathcal{O}$ of the topology of X that is closed with respect to finite intersections. Then the net of measures μ_α converges weakly to μ.*

PROOF. Denote by $\mathcal{U}$ the family of all finite unions of sets in $\mathcal{O}$. Convergence on $\mathcal{O}$ and stability of $\mathcal{O}$ with respect to finite intersections yields that $\lim\limits_\alpha \mu_\alpha(U) = \mu(U)$ for all $U \in \mathcal{U}$. For every open set G and every set $U \in \mathcal{U}$ with $U \subset G$, we have

$$\mu(U) = \lim_\alpha \mu_\alpha(U) \le \liminf_\alpha \mu_\alpha(G),$$

whence by the τ-additivity of μ we obtain that

$$\mu(G) = \sup\bigl\{\mu(U)\colon\ U \subset G,\, U \in \mathcal{U}\bigr\} \le \liminf_\alpha \mu_\alpha(G),$$

since G is the union of the directed family of all sets $U \subset G$ from $\mathcal{U}$ (we recall that $\mathcal{O}$ is a topology base). As we know, the obtained estimate is equivalent to weak convergence of μ_α to μ. $\square$

The following theorem of R. Rao [**1544**] gives a useful effective sufficient condition of uniform convergence of integrals with respect to weakly convergent measures.

8.2.18. Theorem. *Suppose that a net $\{\mu_\alpha\}$ of Baire probability measures on a completely regular Lindelöf space X (for example, on a separable metric space) converges weakly to a Baire probability measure μ. If $\Gamma \subset C_b(X)$ is a uniformly bounded and pointwise equicontinuous family of functions (i.e., for every x and $\varepsilon > 0$, there exists a neighborhood U of the point x with $|f(x) - f(y)| < \varepsilon$ for all $y \in U$ and $f \in \Gamma$), then*

$$\lim_\alpha \sup_{f\in\Gamma} \biggl| \int_X f\,d\mu_\alpha - \int_X f\,d\mu \biggr| = 0. \qquad (8.2.7)$$

PROOF. We may assume that the measures μ_α and μ are Borel and τ-additive, since by the Lindelöf property of X they satisfy the hypothesis of Corollary 7.3.3(ii), which yields the existence and uniqueness of a τ-additive extension. One can also assume that $|f| \le 1$ for all $f \in \Gamma$. Let $\varepsilon > 0$. In view of the complete regularity of X and our hypothesis, every point x has a functionally open neighborhood U_x such that $\mu(\partial U_x) = 0$ and $|f(x)-f(y)| < \varepsilon$ for all $y \in U_x$ and $f \in \Gamma$. Since X is Lindelöf, some countable collection of sets U_{x_n} covers X. Let $V_n = U_{x_n} \backslash \bigcup_{i=1}^{n-1} V_i$, $V_1 = U_{x_1}$. It is readily verified that the pairwise disjoint sets V_n cover X and $\mu(\partial V_n) = 0$. Let $\nu = \sum_{n=1}^\infty \mu(V_n)\delta_{x_n}$, $\nu_\alpha = \sum_{n=1}^\infty \mu_\alpha(V_n)\delta_{x_n}$. We observe that

$$\lim_\alpha \sup_{f\in\Gamma} \biggl| \int_X f\,d\nu_\alpha - \int_X f\,d\nu \biggr| \le \lim_\alpha \sum_{n=1}^\infty |\mu_\alpha(V_n) - \mu(V_n)| = 0. \qquad (8.2.8)$$

The last equality in (8.2.8) follows by the equality $\lim\limits_{\alpha} \mu_\alpha(V_n) = \mu(V_n)$ for every fixed n (which holds according to Theorem 8.2.9(ii)) and the equality $\sum_{n=1}^\infty \mu_\alpha(V_n) = \sum_{n=1}^\infty \mu(V_n) = 1$. We observe that

$$\begin{aligned}
&\biggl|\int_X f\,d\mu_\alpha - \int_X f\,d\mu\biggr| \\
&\le \biggl|\int_X f\,d\mu_\alpha - \int_X f\,d\nu_\alpha\biggr| + \biggl|\int_X f\,d\nu_\alpha - \int_X f\,d\nu\biggr| + \biggl|\int_X f\,d\nu - \int_X f\,d\mu\biggr| \\
&\le \sum_{n=1}^\infty \int_{V_n} |f(x) - f(x_n)|\,(\mu_\alpha + \mu)(dx) + \biggl|\int_X f\,d\nu_\alpha - \int_X f\,d\nu\biggr| \\
&\le 2\varepsilon + \biggl|\int_X f\,d\nu_\alpha - \int_X f\,d\nu\biggr|,
\end{aligned}$$

since $|f(x)-f(x_n)| \le \varepsilon$ for all $x \in V_n$ because $V_n \subset U_{x_n}$. Now equality (8.2.7) follows by (8.2.8), since ε is arbitrary. □

Concerning signed measures, see Exercises 8.10.133 and 8.10.134.

8.3. The case of a metric space

In this section X is a metric space with a metric ϱ. Thus, the classes of Borel and Baire measures coincide and, as we have already seen, the formulations of some results are simplified. Nevertheless, there still remains some difference between the case where X is separable and the general case. We shall see below that the situation is most favorable for complete separable metric spaces.

We have already noted in §8.1 that except for the case of finite X, the weak topology on $\mathcal{M}_\sigma(X)$ is not metrizable, hence is not normable. But $\mathcal{M}_\sigma(X)$ can be equipped with a norm such that the generated topology coincides with the weak topology on the set of τ-additive nonnegative measures (hence on the set of probability measures).

Let us equip the space $\mathcal{M}_\sigma(X)$ with the following Kantorovich–Rubinshtein norm:

$$\|\mu\|_0 = \sup\biggl\{\int_X f\,d\mu\colon\ f \in \mathrm{Lip}_1(X), \sup_{x\in X} |f(x)| \le 1\biggr\},$$

where

$$\mathrm{Lip}_1(X) := \bigl\{f\colon\ X \to \mathbb{R}^1,\ |f(x) - f(y)| \le \varrho(x,y),\ \forall\, x,y \in X\bigr\}.$$

It is clear that $\|\mu\|_0 \le \|\mu\|$. The metric generated by the Kantorovich–Rubinshtein norm is called the Kantorovich–Rubinshtein metric (in § 8.10(viii) we consider a modification of this metric).

If the space X contains an infinite convergent sequence, then the norm $\|\cdot\|_0$ is strictly weaker than the total variation norm $\|\cdot\|$. Indeed, if $x_n \to x$, then the measures δ_{x_n} converge in the norm $\|\cdot\|_0$ to the measure δ_x, since $|f(x_n) - f(x)| \le \varrho(x_n, x)$ for all $f \in \mathrm{Lip}_1(X)$, but $\|\delta_x - \delta_{x_n}\| = 2$ if $x_n \ne x$.

In particular, in this case the space $\mathcal{M}_\sigma(X)$ cannot be complete with respect to the norm $\|\cdot\|_0$ because it is complete in the variation norm and then by the Banach theorem both norms would be equivalent. If $\varrho(x,y) \ge \delta > 0$ whenever $x \ne y$, then the norms $\|\cdot\|_0$ and $\|\cdot\|$ are equivalent, since in that case we have $f \in \mathrm{Lip}_1(X)$ provided that $|f(x)| \le \delta/2$. It is shown below that the topology generated by the norm $\|\cdot\|_0$ coincides with the weak topology on the set of nonnegative τ-additive measures. One frequently employs the equivalent norm

$$\|\mu\|^*_{\mathrm{BL}} := \sup\biggl\{\int_X f\,d\mu\colon\ f \in \mathrm{BL}(X), \|f\|_{\mathrm{BL}} \le 1\biggr\},$$

where $\mathrm{BL}(X)$ is the space of all bounded Lipschitzian functions on X with the norm

$$\|f\|_{\mathrm{BL}} := \sup_{x\in X} |f(x)| + \sup_{x\ne y} \frac{|f(x)-f(y)|}{\varrho(x,y)}.$$

It is readily verified that $\mathrm{BL}(X)$ with this norm is complete. It is clear that

$$\|\mu\|^*_{\mathrm{BL}} \le \|\mu\|_0 \le 2\|\mu\|^*_{\mathrm{BL}},$$

since $\|f\|_{\mathrm{BL}} \le 2$ whenever $f \in \mathrm{Lip}_1(X)$ and $\sup_X |f(x)| \le 1$.

8.3.1. Remark. It follows by Theorem 8.2.3 that weak convergence of a net $\{\mu_\alpha\}$ of *nonnegative* measures to a measure μ is equivalent to the equality

$$\lim_\alpha \int_X f(x)\,\mu_\alpha(dx) = \int_X f(x)\,\mu(dx)$$

for all bounded uniformly continuous functions f on X (this is also true for uniform spaces, hence for completely regular spaces equipped with suitable uniformities, see Topsøe [**1873**]). Indeed, given a closed set F and $\varepsilon > 0$, one can find a bounded uniformly continuous (even Lipschitzian) function f such that $0 \le f \le 1$, $f|_F = 1$ and the integral of f against μ is estimated by the number $\mu(F)+\varepsilon$. Then $\limsup_\alpha \mu_\alpha(F) \le \mu(F)+\varepsilon$, hence $\limsup_\alpha \mu_\alpha(F) \le \mu(F)$. Clearly, the same is true for Lipschitzian functions in place of uniformly continuous ones (this is also seen from Exercise 8.10.71). In particular, convergence of a net of nonnegative measures in the Kantorovich–Rubinshtein metric implies weak convergence. However, if X is not compact, then one can choose a metric on X generating the same topology such that there exist a sequence of signed measures μ_n and a measure μ such that the integrals with respect to μ_n of every bounded uniformly continuous function f converge to the integral of f against the measure μ, but the measures μ_n do not converge weakly to μ (Exercise 8.10.77). The original metric does not always have such a property (for example, take $X = \mathbb{N}$ with the usual metric), but in the case $X = \mathbb{R}^1$ the standard metric also works: it suffices to have two sequences $\{x_n\}$ and $\{y_n\}$ with $x_n \ne y_n$ which have no limit points, but the distance between x_n and y_n tends to zero.

For every $B \subset X$, we let $B^\varepsilon = \{x\colon\ \mathrm{dist}(x,B) < \varepsilon\}$.

8.3.2. Theorem. *The topology generated by the norm $\|\cdot\|_0$ coincides with the weak topology on the set $\mathcal{M}_\tau^+(X)$ of nonnegative τ-additive measures. In addition, on the set $\mathcal{P}_\tau(X)$ of probability τ-additive measures the weak topology is generated by the following Lévy–Prohorov metric*:

$$d_P(\mu,\nu) = \inf\Bigl\{\varepsilon > 0\colon\ \nu(B) \le \mu(B^\varepsilon) + \varepsilon,\ \mu(B) \le \nu(B^\varepsilon) + \varepsilon,\ \forall\, B \in \mathcal{B}(X)\Bigr\}.$$

In particular, if X is separable, then the weak topology on the set $\mathcal{M}_\sigma^+(X)$ is generated by the metric $d_0(\mu,\nu) := \|\mu-\nu\|_0$.

Finally, if $\mathcal{P}_\sigma(X) \ne \mathcal{P}_\tau(X)$, then the weak topology is not metrizable on $\mathcal{P}_\sigma(X)$.

PROOF. We verify that d_P is a metric on $\mathcal{P}_\sigma(X)$. It is clear that we have $d_P(\mu,\nu) = d_P(\nu,\mu)$. If $d_P(\mu,\nu) = 0$, then $\mu(B) = \nu(B)$ for every closed set B and hence $\mu = \nu$. Let $\nu(B) \le \mu(B^\varepsilon)+\varepsilon$, $\mu(B) \le \nu(B^\varepsilon)+\varepsilon$, $\mu(B) \le \eta(B^\delta)+\delta$, $\eta(B) \le \mu(B^\delta) + \delta$ for all $B \in \mathcal{B}(X)$. Then $\nu(B) \le \eta(B^{\varepsilon+\delta}) + \varepsilon + \delta$ and $\eta(B) \le \nu(B^{\varepsilon+\delta}) + \varepsilon + \delta$, whence $d_P(\nu,\eta) \le d_P(\nu,\mu) + d_P(\mu,\eta)$. Let us show that every neighborhood W of the form (8.2.1) contains a ball of a positive radius with respect to the Lévy–Prohorov metric. To this end, we pick $\delta \in (0,\varepsilon/2)$ such that $\mu(F_i^\delta) < \mu(F_i) + \varepsilon/2$ for all $i = 1,\ldots,n$. If $d_P(\mu,\nu) < \delta$, then $\nu(F_i) < \mu(F_i^\delta) + \delta < \mu(F_i) + \varepsilon$, i.e., ν belongs to the neighborhood W. We note that at this stage no separability of measures is used.

Now we show that every ball with respect to the Lévy–Prohorov metric with the center μ and radius ε contains a neighborhood of the form (8.2.1). We pick $\delta > 0$ such that $3\delta < \varepsilon$. Let us cover the separable support of the measure μ by countably many open balls V_n of diameter less than δ having the boundaries of μ-measure zero (by the τ-additivity the support exists and is separable). We construct pairwise disjoint sets A_n that have boundaries of μ-measure zero and cover the support of μ. To this end, let $A_n = \bigcup_{i=1}^n V_i \backslash \bigcup_{i=1}^{n-1} V_i$, $A_1 = V_1$. There is k such that

$$\mu\Bigl(\bigcup_{i=1}^k A_i\Bigr) > 1-\delta. \tag{8.3.1}$$

By Corollary 8.2.10 there exists a neighborhood W of the form (8.2.1) such that

$$|\mu(A) - \nu(A)| < \delta \tag{8.3.2}$$

for all $\nu \in W$ and every set A that is a union of some of the sets $A_1,\ldots, A_k$. We verify that $d_P(\mu,\nu) < \varepsilon$ for all $\nu \in W$. Let $B \in \mathcal{B}(X)$. Let us consider the set A that is the union of those sets $A_1,\ldots,A_k$ that do not meet B. Then $B \subset A \bigcup \bigcup_{i=k+1}^\infty A_i$ and $A \subset B^\delta$, since the diameter of every A_i is less than δ. Given $\nu \in W$, we obtain by (8.3.1) and (8.3.2) that

$$\mu(B) < \mu(A) + \delta < \nu(A) + 2\delta \le \nu(B^\delta) + 2\delta.$$

Since (8.3.1) and (8.3.2) yield that $\nu\Bigl(\bigcup_{i=1}^k A_i\Bigr) > 1-2\delta$, we obtain similarly that $\nu(B) < \mu(B^\delta) + 3\delta$. Thus, $d_P(\mu,\nu) < 3\delta < \varepsilon$.

According to the remark above, convergence in the Kantorovich–Rubinshtein metric yields weak convergence for nets in $\mathcal{M}^+_\sigma(X)$. On the other hand, if a sequence of nonnegative τ-additive measures μ_n converges weakly to a τ-additive measure μ, then there exists a separable closed subspace X_0 on which all measures μ_n and μ are concentrated. Hence by Theorem 8.2.18 we have

$$\sup_{f\in \mathrm{Lip}_1(X), |f|\le 1} \biggl| \int_X f\,d\mu - \int_X f\,d\mu_n \biggr| \to 0.$$

Thus, the families of convergent sequences in the weak topology on $\mathcal{M}^+_\tau(X)$ and in the metric d_0 coincide. It follows by the already obtained results that on $\mathcal{P}_\tau(X)$ the metrics d_P and d_0 generate one and the same topology, namely, the weak topology. So all the three topologies have the same convergent nets. Suppose now that a net $\{\mu_\alpha\} \subset \mathcal{M}^+_\tau(X)$ converges to a measure $\mu \in \mathcal{M}^+_\tau(X)$ in the weak topology. If $\mu = 0$, then $\mu_\alpha(X) \to 0$ and hence $d_0(\mu_\alpha, 0) \to 0$. If $\mu \neq 0$, then we may assume that $c_\alpha := \mu_\alpha(X) > 0$. Since $c_\alpha \to \mu(X)$, one has $\mu_\alpha/c_\alpha \to \mu/\mu(X)$ in the weak topology. By the already established assertion for probability measures, $\mu_\alpha/c_\alpha \to \mu/\mu(X)$ in the metric d_0, whence one has that $\|\mu_\alpha - \mu\|_0 \to 0$.

Finally, if the weak topology is metrizable on $\mathcal{P}_\sigma(X)$, then in view of Example 8.1.6, every measure $\mu \in \mathcal{P}_\sigma(X)$ is the limit of a sequence of measures μ_n with finite supports, hence has a separable support. $\square$

8.3.3. Example. If $\mu \in \mathcal{P}_\sigma(X)$ has no atoms and $\alpha \in (0,1)$, then there exist sets $B_n \in \mathcal{B}(X)$ with $\mu(B_n) = \alpha$ such that the measures $\mu_n := \alpha^{-1} I_{B_n}\cdot\mu$ converge to μ in the norm $\|\cdot\|_0$. Indeed, let us partition X into Borel parts $E_{n,i}$ of diameter less than $1/n$ with $\mu(E_{n,i}) > 0$. Next we find Borel sets $B_{n,i} \subset E_{n,i}$ with $\mu(B_{n,i}) = \alpha\mu(E_{n,i})$ and take $B_n := \bigcup_{i=1}^\infty B_{n,i}$. Let f belong to $\mathrm{Lip}_1(X)$. The absolute value of the integral of f against the measure $\mu_n - \mu$ does not exceed $2/n$, since taking $x_i \in B_{n,i}$ we obtain that the integral of $fI_{E_{n,i}}$ against the measure μ differs from $f(x_i)\mu(E_{n,i})$ in at most $\mu(E_{n,i})/n$ and the same is true for the measure μ_n.

In §8.9 we discuss the completeness of $\mathcal{M}^+_\tau(X)$ in the metric d_0.

8.4. Some properties of weak convergence

In this section, we discuss the behavior of weak convergence under some operations on measures: transformation of measures, restrictions to sets, multiplication by functions, and products of measures.

8.4.1. Theorem. *Suppose a net of Baire measures μ_α on a topological space X converges weakly to a measure μ. Then the following assertions are true.*

(i) *For every continuous mapping $F\colon X \to Y$ to a topological space Y, the net of measures $\mu_\alpha \circ F^{-1}$ converges weakly to the measure $\mu \circ F^{-1}$.*

(ii) *Suppose that X is a completely regular space, the measures μ_α and μ are nonnegative Borel, and the measure μ is τ-additive. Let F be a Borel mapping from X to a topological space Y such that F is continuous μ-almost everywhere. Then $\mu_\alpha \circ F^{-1} \Rightarrow \mu \circ F^{-1}$.*

(iii) *Let X be a separable metric space, let the measures μ_α be nonnegative, and let F_α be pointwise equicontinuous mappings from X to a metric space Y such that the measures $\mu \circ F_\alpha^{-1}$ converge weakly to the measure $\mu \circ F^{-1}$, where $F\colon X \to Y$ is some Borel mapping. Then the measures $\mu_\alpha \circ F_\alpha^{-1}$ also converge weakly to $\mu \circ F^{-1}$.*

Proof. Assertion (i) is obvious. Let us verify (ii). Let Z be a closed set in Y. Denote by D_F the set of discontinuity points of F. We observe that $\overline{F^{-1}(Z)} \subset F^{-1}(Z) \cup D_F$, where $\overline{A}$ is the closure of A. Then by Corollary 8.2.4 one has

$$\limsup_\alpha \mu_\alpha \circ F^{-1}(Z) \le \limsup_\alpha \mu_\alpha\Bigl(\overline{F^{-1}(Z)}\Bigr) \le \mu\Bigl(\overline{F^{-1}(Z)}\Bigr) = \mu\Bigl(F^{-1}(Z)\Bigr),$$

which yields that $\mu_\alpha \circ F^{-1} \Rightarrow \mu \circ F^{-1}$

For the proof of assertion (iii) we fix a uniformly continuous bounded function φ on Y. The functions $\varphi \circ F_\alpha$ on X are uniformly bounded and pointwise equicontinuous. By Theorem 8.2.18, for every $\varepsilon > 0$, there exists an index α_0 such that for all $\alpha \ge \alpha_0$ one has

$$\biggl|\int_X \varphi \circ F_\alpha \, d\mu_\alpha - \int_X \varphi \circ F_\alpha \, d\mu\biggr| < \varepsilon/2.$$

It follows by our hypothesis that there exists an index $\alpha_1 \ge \alpha_0$ such that for all $\alpha \ge \alpha_1$ we have

$$\biggl|\int_X \varphi \circ F_\alpha \, d\mu - \int_X \varphi \circ F \, d\mu\biggr| < \varepsilon/2.$$

These two estimates yield the claim. □

8.4.2. Corollary. *Let $\{\mu_\alpha\}$ be a net of Borel probability measures on a completely regular space X and let μ be a τ-additive probability measure. Then $\{\mu_\alpha\}$ converges weakly to μ if and only if the equality*

$$\lim_\alpha \int_X f \, d\mu_\alpha = \int_X f \, d\mu$$

is true for every bounded Borel function f that is continuous μ-almost everywhere.

Proof. The sufficiency of the above condition is obvious. Its necessity follows by assertion (ii) in the previous theorem in view of the equality

$$\int_X f \, d\mu_\alpha = \int_{\mathbb{R}^1} h \, d(\mu_\alpha \circ f^{-1}),$$

where $h \in C_b(\mathbb{R}^1)$ and $h(t) = t$ if $|t| \le \sup|f|$, and the analogous equality for μ. □

8.4.3. Lemma. *If a net $\{\mu_\alpha\}$ of Baire probability measures on a topological space X converges weakly to a Baire measure μ, then for every continuous function f on X satisfying the condition*

$$\lim_{R\to\infty} \sup_\alpha \int_{|f|\geq R} |f|\, d\mu_\alpha = 0,$$

one has

$$\lim_\alpha \int_X f\, d\mu_\alpha = \int_X f\, d\mu$$

If X is completely regular, μ_α and μ are Radon and for every $\varepsilon > 0$, there exists a compact set K_ε such that $\mu_\alpha(X\backslash K_\varepsilon) < \varepsilon$ for all α, then the continuity of f can be relaxed to the continuity on each K_ε.

PROOF. We observe that $f \in L^1(\mu)$. Indeed, let $f_n = \min(|f|, n)$. Then $f_n \leq |f|$ and hence by the hypothesis of the lemma we obtain

$$M := \sup_{n,\alpha} \int_X f_n\, d\mu_\alpha < \infty.$$

Since $f_n \in C_b(X)$, one has

$$\int f_n\, d\mu \leq M$$

for all n, whence $f \in L^1(\mu)$. Let $\varepsilon > 0$. Pick $R > 0$ such that for all α

$$\int_{|f|\geq R} |f|\, d\mu_\alpha + \int_{|f|\geq R} |f|\, d\mu < \varepsilon.$$

Let $g = \max\bigl(\min(f, R), -R\bigr)$. For all α with

$$\Bigl|\int_X g\, d\mu_\alpha - \int_X g\, d\mu\Bigr| < \varepsilon,$$

we obtain

$$\Bigl|\int_X f\, d\mu_\alpha - \int_X f\, d\mu\Bigr| \leq 3\varepsilon,$$

since $|g(x)| \leq |f(x)|$ and $g(x) = f(x)$ whenever $|f(x)| \leq R$. The second assertion is proved similarly. Letting $A = \{|f| \leq R\}$, we find a compact set $K \subset A$ on which f is continuous and $\mu_\alpha(A\backslash K) + \mu(A\backslash K) < \varepsilon R^{-1}$ for all α. Then $f|_K$ can be extended to a continuous function g on all of the space such that $|g| \leq R$ (Exercise 6.10.22). □

Let us consider the behavior of weak convergence under restricting measures to subsets. It is clear that in the general case there is no convergence of restrictions: in Example 8.1.4, the convergent measures vanish on the set $\{0\}$, but the limit Dirac measure is concentrated on that set.

The next result follows by the last assertion in Corollary 8.2.10.

8.4.4. Proposition. *Suppose that a net $\{\mu_\alpha\}$ of Borel probability measures on a completely regular space X converges weakly to a τ-additive Borel probability measure μ. Let a set $X_0 \subset X$ be equipped with the induced topology. Then the induced measures μ_α^0 on X_0 converge weakly to the measure μ^0 induced by μ in either of the following cases*:

(i) X_0 *is a set of full outer measure for all measures* μ_α *and* μ;

(ii) X_0 *is either open or closed and* $\lim\limits_\alpha \mu_\alpha(X_0) = \mu(X_0)$.

This assertion remains valid for nonnegative, not necessarily probability, measures provided that $\lim\limits_\alpha \mu_\alpha(X) = \mu(X)$.

It is easy to see that in the general case weak convergence is not preserved by the elements of the Jordan–Hahn decomposition and does not commute with taking the total variation. Let us consider some examples.

8.4.5. Example. (i) Let μ_n be measures on the interval $[0, 2\pi]$ defined as follows: $\mu_n = 0$ if n is odd and $\mu_n = \sin(nx)\,dx$ if n is even. It is readily seen that the measures μ_n converge weakly to the zero measure, but the measures $|\mu_n|$ have no weak limit.

(ii) The measures $\delta_0 - \delta_{1/n}$ on the real line converge weakly to the zero measure, but their total variations $|\delta_0 - \delta_{1/n}| = \delta_0 + \delta_{1/n}$ converge weakly to $2\delta_0$.

The next example due to Le Cam [**1138**] exhibits another interesting aspect of this phenomenon.

8.4.6. Example. Let X be a subset of $[0, 1]$ containing all numbers of the form $k2^{-n}$ with $n, k \in \mathbb{N}$ and having the inner measure zero and outer measure 1. We equip X with the induced topology and the measure μ that is the restriction of Lebesgue measure λ to X (see Definition 1.12.11). Set

$$\nu_n(k2^{-n}) = 2^{-n} \quad \text{for } k = 1, \ldots, 2^n, \quad \mu_n = \nu_{n+1} - \nu_n.$$

The sequence $\{\mu_n\}$ of Radon measures converges weakly to zero, but the sequence of measures $|\mu_n| = \nu_{n+1}$ converges weakly to the measure μ, which is τ-additive, but not Radon.

The following result from Varadarajan [**1918**, Part 2, Theorem 3] is useful for the study of weak convergence of signed measures.

8.4.7. Theorem. *Suppose that a net of Baire measures μ_α converges weakly to a Baire measure μ. Then, for every functionally open set U we have*

$$\liminf_\alpha |\mu_\alpha|(U) \geq |\mu|(U).$$

In this situation, the net of measures $|\mu_\alpha|$ converges weakly to $|\mu|$ precisely when $|\mu_\alpha|(X) \to |\mu|(X)$.

PROOF. Let $\varepsilon > 0$. By Lemma 7.1.10, one can find a function $g \in C_b(X)$ such that $0 \leq g \leq 1$, $g = 0$ on $X\backslash U$, and

$$\int_X g\, d|\mu| > |\mu|(U) - \varepsilon.$$

It is readily seen that there exists a function $h \in C_b(X)$ such that $|h| \le g$ and

$$\Bigl|\int_X h\,d\mu\Bigr| > \int_X g\,d|\mu| - \varepsilon.$$

It is clear that $|h| \le 1$ and $h = 0$ on $X\backslash U$. In addition,

$$\Bigl|\int_X h\,d\mu\Bigr| > |\mu|(U) - 2\varepsilon.$$

Since

$$\int_X h\,d\mu_\alpha \to \int_X h\,d\mu,$$

one has

$$\liminf_\alpha |\mu_\alpha|(U) \ge \lim_\alpha \Bigl|\int_X h\,d\mu_\alpha\Bigr| = \Bigl|\int_X h\,d\mu\Bigr| > |\mu|(U) - 2\varepsilon.$$

Letting $\varepsilon \to 0$, we obtain the first assertion. If $|\mu_\alpha|(X) \to |\mu|(X) > 0$, then weak convergence of $|\mu_\alpha|$ to $|\mu|$ follows by the first claim. If $|\mu|(X) = 0$, then one has convergence in the variation norm. □

8.4.8. Corollary. *Suppose that a net of Baire measures μ_α converges weakly to a Baire measure μ and that*

$$\lim_\alpha |\mu_\alpha|(X) = |\mu|(X).$$

Let $\mu_\alpha = \mu_\alpha^+ - \mu_\alpha^-$ and $\mu = \mu^+ - \mu^-$. Then, the nets $\{\mu_\alpha^+\}$ and $\{\mu_\alpha^-\}$ converge weakly to μ^+ and μ^-, respectively.

Proof. We apply the equalities $\mu_\alpha^+ = (|\mu_\alpha| + \mu_\alpha)/2$, $\mu_\alpha^- = (|\mu_\alpha| - \mu_\alpha)/2$ and the theorem proven above. □

Now we can investigate the problem of preservation of weak convergence under multiplication by a function. It follows by definition that if measures μ_α converge weakly to a measure μ, then for every bounded continuous function f, the measures $f\cdot\mu_\alpha$ converge weakly to the measure $f\cdot\mu$. However, there are less trivial results of this sort. For example, Proposition 8.4.4 and Corollary 8.4.2 yield the following assertion.

8.4.9. Proposition. *Suppose that a net of Borel probability measures μ_α on a completely regular space X converges weakly to a τ-additive Borel probability measure μ and a bounded Borel function f is continuous at μ-almost all points of a set X_0 that has full measure with respect to all measures μ_α and μ. Then, the measures $f\cdot\mu_\alpha$ converge weakly to the measure $f\cdot\mu$.*

8.4.10. Theorem. *Let $\{\mu_\alpha\}$ and $\{\nu_\alpha\}$ be two nets of τ-additive probability measures on completely regular spaces X and Y convergent weakly to τ-additive measures μ and ν, respectively. Then the τ-additive extensions of the measures $\mu_\alpha\otimes\nu_\alpha$ converge weakly to the τ-additive extension of the measure $\mu\otimes\nu$.*

PROOF. Denote by $\mathcal{U}_\mu$ and $\mathcal{U}_\nu$ the classes of open sets in X and Y with boundaries of zero measure with respect to μ and ν, correspondingly. By Proposition 8.2.8 these families form topology bases in X and Y. Hence the family $\mathcal{U} = \{U\times V\colon\ U \in \mathcal{U}_\mu, V \in \mathcal{U}_\nu\}$ is a topology base in $X\times Y$. The family $\mathcal{U}$ is closed with respect to finite intersections because, as one can easily see, $\mathcal{U}_\mu$ and $\mathcal{U}_\nu$ have such a property. For all $U \in \mathcal{U}_\mu$ and $V \in \mathcal{U}_\nu$ one has

$$\lim_\alpha \mu_\alpha\otimes\nu_\alpha(U\times V) = \lim_\alpha \mu_\alpha(U)\lim_\alpha \nu_\alpha(V) = \mu\otimes\nu(U\times V).$$

Hence Theorem 8.2.17 yields the claim. $\square$

8.5. The Skorohod representation

Suppose that P is a probability measure on some measurable space $(\Omega, \mathcal{F})$ and $\{\xi_n\}$ is a sequence of $\bigl(\mathcal{F}, \mathcal{B}a(X)\bigr)$-measurable mappings from Ω to a topological space X equipped with the Baire σ-algebra $\mathcal{B}a(X)$. Assume also that there exists a $\bigl(\mathcal{F}, \mathcal{B}a(X)\bigr)$-measurable mapping $\xi\colon\ \Omega \to X$ such that $\xi(\omega) = \lim\limits_{n\to\infty} \xi_n(\omega)$ for P-a.e. $\omega \in \Omega$. It is clear that the measures $\mu_n = P\circ\xi_n^{-1}$ converge weakly to the measure $\mu = P \circ \xi^{-1}$ because, for all $\varphi \in C_b(X)$, we have

$$\lim_{n\to\infty} \int_\Omega \varphi\bigl(\xi_n(\omega)\bigr)\, P(d\omega) = \int_\Omega \varphi\bigl(\xi(\omega)\bigr)\, P(d\omega)$$

by the dominated convergence theorem. Skorohod [**1739**], [**1740**] discovered that every weakly convergent sequence of probability measures on a complete separable metric space X admits the above representation and that one can take for P Lebesgue measure on $[0, 1]$ (for measures on $X = \mathbb{R}^d$ this was shown in Hammersley [**783**]). Blackwell and Dubins [**184**] and Fernique [**566**] established that one can simultaneously parameterize all probability measures on X by mappings from $[0, 1]$ in such a way that to weakly convergent sequences of measures there will correspond almost everywhere convergent sequences of mappings. This section contains a simple derivation of this result by means of functional-topological arguments. The following concept introduced in Bogachev, Kolesnikov [**211**] will be useful in our discussion. This concept is of independent interest.

8.5.1. Definition. *We shall say that a topological space X has the strong Skorohod property for Radon measures if to every Radon probability measure μ on X, one can associate a Borel mapping $\xi_\mu\colon\ [0, 1] \to X$ such that μ is the image of Lebesgue measure under the mapping ξ_μ and $\xi_{\mu_n}(t) \to \xi_\mu(t)$ a.e. whenever the measures μ_n converge weakly to μ.*

If such a parameterization exists for the class of all Borel probability measures on X, then the obtained property will be called the strong Skorohod property for Borel measures. By analogy one can define the strong Skorohod property for other classes of measures (for example, discrete).

8.5.2. Lemma. *Let X be a space with the strong Skorohod property for Radon measures. Then*:

(i) *every subset* Y *of* X *has this property as well;*

(ii) *if* F *is a continuous mapping from* X *to a topological space* Y *and there exists a mapping* $\Psi\colon \mathcal{P}_r(Y) \to \mathcal{P}_r(X)$ *continuous in the weak topology such that* $\Psi(\nu)\circ F^{-1} = \nu$ *for all* $\nu \in \mathcal{P}_r(Y)$*, then* Y *has the strong Skorohod property for Radon measures.*

PROOF. (i) Every Radon measure μ on Y extends uniquely to a Radon measure on X, and Y is measurable with respect to this extension, since Y contains compact sets K_n (these sets are also compact in X) whose union has full measure. Let $\xi_\mu\colon\ [0,1] \to X$ be a Borel mapping corresponding to μ in the definition of the strong Skorohod property (i.e., we fix some parameterization). As noted above, there exists a set $B \subset Y$ of full μ-measure that is σ-compact in X and Y. Let $\eta_\mu(t) = \xi_\mu(t)$ if $t \in \xi_\mu^{-1}(B)$ and $\eta_\mu(t) = z$ if $t \notin \xi_\mu^{-1}(B)$, where z is an arbitrary point in Y. Then $\lambda\bigl(\xi_\mu^{-1}(B)\bigr) = 1$ and hence $\eta_\mu(t) = \xi_\mu(t)$ for almost all t in $[0,1]$, whence $\lambda\circ\eta_\mu^{-1} = \lambda\circ\xi_\mu^{-1}$. If probability measures μ_n on Y converge weakly to the measure μ, then their extensions to X converge weakly to the extension of μ, whence one has $\lim\limits_{n\to\infty} \xi_{\mu_n}(t) = \xi_\mu(t)$ almost everywhere. Therefore, $\lim\limits_{n\to\infty} \eta_{\mu_n}(t) = \eta_\mu(t)$ almost everywhere.

(ii) Given $\nu \in \mathcal{P}_r(Y)$, let $\eta_\nu(t) = F\bigl(\xi_{\Psi(\nu)}(t)\bigr)$, where ξ is a parameterization of measures in $\mathcal{P}_r(X)$ by Borel mappings from the interval $[0,1]$ to X. Then

$$\lambda \circ \eta_\nu^{-1} = \bigl(\lambda \circ \xi_{\Psi(\nu)}^{-1}\bigr) \circ F^{-1} = \Psi(\nu) \circ F^{-1} = \nu.$$

If measures ν_n converge weakly to the measure ν on Y, then the measures $\Psi(\nu_n)$ converge weakly to the measure $\Psi(\nu)$ on X, hence $\xi_{\Psi(\nu_n)}(t) \to \xi_{\Psi(\nu)}(t)$ for almost all t in $[0,1]$, whence $\eta_{\nu_n}(t) \to \eta_\nu(t)$ for such points t due to the continuity of F. □

The mapping Ψ in assertion (ii) of this lemma is called a continuous right inverse to the induced mapping $\widehat{F}\colon\ \mathcal{P}_r(X) \to \mathcal{P}_r(Y)$, $\mu \mapsto \mu \circ F^{-1}$.

Let $F\colon\ X \to Y$ be a continuous surjection of compact spaces X and Y. A linear operator $U\colon\ C(X) \to C(Y)$ is called a regular averaging operator for F if $U\psi \geq 0$ whenever $\psi \geq 0$ and $U(\varphi \circ F) = \varphi$ for all $\varphi \in C(Y)$. Such an operator is automatically continuous and has the unit norm. It is easy to see that the operator $V = U^*\colon\ \mathcal{M}_r(Y) = C(Y)^* \to \mathcal{M}_r(X) = C(X)^*$ takes $\mathcal{P}_r(Y)$ to $\mathcal{P}_r(X)$ and that $\widehat{F} \circ V$ is the identity mapping on $\mathcal{M}_r(Y)$, i.e., V is a continuous right inverse for $\widehat{F}$. Indeed, for all $\nu \in \mathcal{M}_r(Y)$ and $\varphi \in C(Y)$, we have

$$\begin{aligned}\int_Y \varphi(y)\, \widehat{F}\bigl(V(\nu)\bigr)(dy) &= \int_X \varphi\bigl(F(x)\bigr)\, V(\nu)(dx) \\ &= \int_Y U(\varphi \circ F)(y)\, \nu(dy) = \int_Y \varphi(y)\, \nu(dy).\end{aligned}$$

A compact space S is called a *Milyutin space* if for some cardinality τ, there exists a continuous surjection $F\colon \{0,1\}^\tau \to S$, where $\{0,1\}$ is the two-point space, such that F has a regular averaging operator. According to the celebrated Milyutin lemma (see Pełcziński [**1430**, Theorem 5.6], Fedorchuk, Filippov [**561**, Ch. 8, §4]), the closed interval is a Milyutin space. In addition, it is known that the direct product of an arbitrary family of compact metric spaces is a Milyutin space. In particular, $S = [0,1]^\infty$ is a Milyutin space, and for τ one can take $\mathbb{N}$. Since the space $\{0,1\}^\infty$ is homeomorphic to the classical Cantor set $C \subset [0,1]$, consisting of all numbers in the interval $[0,1]$ whose ternary expansions do not contain 1 (see Engelking [**532**, Example 3.1.28]), we arrive at the following result.

8.5.3. Lemma. *Let S be a nonempty metrizable compact space and let C be the Cantor set. Then, there exists a continuous surjection $F\colon C \to S$ such that the mapping $\widehat{F}$ has a linear continuous right inverse.*

8.5.4. Theorem. *Let X be a universally measurable set in a complete separable metric space. Then, to every Borel probability measure μ on X, one can associate a Borel mapping $\xi_\mu\colon [0,1] \to X$ such that $\mu = \lambda \circ \xi_\mu^{-1}$, where λ is Lebesgue measure, and $\xi_{\mu_n}(t) \to \xi_\mu(t)$ for almost all $t \in [0,1]$ whenever the measures μ_n converge weakly to the measure μ. If X is an arbitrary subset of a complete separable metric space, then the analogous assertion is true for Radon probability measures.*

PROOF. Every Polish space is homeomorphic to a G_δ-set in $[0,1]^\infty$ (Theorem 6.1.12). Hence in view of Lemma 8.5.2(i) we may assume that X is contained in $[0,1]^\infty$. By part (ii) of the cited lemma and Lemma 8.5.3 it suffices to verify our claim only for subsets in $[0,1]$, which reduces everything to the case $X = [0,1]$. In the latter case, the required mapping is given by the explicit formula

$$\xi_\mu(t) = \sup\bigl\{x \in [0,1]\colon\ \mu\bigl([0,x)\bigr) \le t\bigr\}. \tag{8.5.1}$$

Indeed, it is easy to see that for every point c, one has $\lambda\circ\xi_\mu^{-1}\bigl([0,c)\bigr) = \mu\bigl([0,c)\bigr)$. Hence $\lambda \circ \xi_\mu^{-1} = \mu$. If measures μ_n converge weakly to the measure μ, then their distribution functions $F_{\mu_n}(t) = \mu_n\bigl([0,t)\bigr)$ converge to the distribution function F_μ of the measure μ at all continuity points of F_μ. Let $t \in [0,1]$ and $\varepsilon > 0$. If

$$\limsup_{n\to\infty} \xi_{\mu_n}(t) > \xi_\mu(t) + 2\varepsilon,$$

then there is a point x_0 in the interval $\bigl(\xi_\mu(t)+\varepsilon, \xi_\mu(t)+2\varepsilon\bigr)$ such that $F_\mu(x_0) = \lim\limits_{n\to\infty} F_{\mu_n}(x_0)$. For some infinite sequence of n_k we have $\xi_{\mu_{n_k}}(t) > x_0$, i.e., $F_{\mu_{n_k}}(x_0) \le t$, whence $F_\mu(x_0) \le t$. Hence $\xi_\mu(t) \ge x_0$, which is a contradiction. Similarly, one considers the case $\liminf\limits_{n\to\infty} \xi_{\mu_n}(t) \le \xi_\mu(t) - 2\varepsilon$. Therefore, we have $\lim\limits_{n\to\infty} \xi_{\mu_n}(t) = \xi_\mu(t)$. In the case of Radon measures a similar reasoning applies to arbitrary subsets of Polish spaces. □

Thus, any subspace of a Polish space possesses the strong Skorohod property for Radon measures (and the universally measurable subspaces have the strong Skorohod property for Borel measures). It is shown in Bogachev, Kolesnikov [**211**] that all complete metric spaces possess the strong Skorohod property for Radon measures.

Additional results on this property can be found in the cited work and in Banakh, Bogachev, Kolesnikov [**115**], [**114**], [**116**], where, in particular, it is shown that there are non-metrizable spaces with the strong Skorohod property, for example, the countable subspace $X = \mathbb{N} \cup \{p\}$ in the Stone–Čech compactification $\beta\mathbb{N}$ of the space of natural numbers, where $p \in \beta\mathbb{N}\backslash\mathbb{N}$.

In relation to the material of this section see also §8.10(v).

8.6. Weak compactness and the Prohorov theorem

The conditions for the weak compactness of families of measures, i.e., compactness in the weak topology $\sigma\bigl(\mathcal{M}, C_b(X)\bigr)$, are very important for the most diverse applications. The following problem is especially frequent: can one select a weakly convergent subsequence in a given sequence of measures? It turns out that for reasonable spaces the problem reduces to the study of the uniform tightness of the given family of measures. In this section, we discuss the principal results in this direction.

8.6.1. Definition. *A family $\mathcal{M}$ of Radon measures on a topological space X is called uniformly tight if for every $\varepsilon > 0$, there exists a compact set K_ε such that $|\mu|(X\backslash K_\varepsilon) < \varepsilon$ for all $\mu \in \mathcal{M}$.*

A family $\mathcal{M}$ of Baire measures on a topological space X is called uniformly tight if for every $\varepsilon > 0$, there exists a compact set K_ε such that $|\mu|_*(X\backslash K_\varepsilon) < \varepsilon$ for all $\mu \in \mathcal{M}$.

For a completely regular space, the uniform tightness of a family of Baire measures is equivalent to the existence of uniformly tight Radon extensions of these measures.

Sometimes, for brevity, uniformly tight families are called tight families.

The following fundamental theorem due to Yu.V. Prohorov [**1497**] (who considered probability measures) is the most important result for applications.

8.6.2. Theorem. *Let X be a complete separable metric space and let $\mathcal{M}$ be a family of Borel measures on X. Then the following conditions are equivalent*:

(i) *every sequence $\{\mu_n\} \subset \mathcal{M}$ contains a weakly convergent subsequence*;

(ii) *the family $\mathcal{M}$ is uniformly tight and uniformly bounded in the variation norm.*

The above conditions are equivalent for any complete metric space X if $\mathcal{M} \subset \mathcal{M}_t(X)$.

PROOF. Let (i) be fulfilled. The uniform boundedness of measures in $\mathcal{M}$ follows by the Banach–Steinhaus theorem. Suppose that $\mathcal{M}$ is not uniformly tight. We show that there exists $\varepsilon > 0$ with the following property: for every compact set $K \subset X$, one can find a measure $\mu^K \in \mathcal{M}$ such that

$$|\mu^K|(X\backslash K^\varepsilon) > \varepsilon, \tag{8.6.1}$$

where K^ε is the closed ε-neighborhood of K. Indeed, otherwise for every $\varepsilon > 0$, there exists a compact set $K(\varepsilon) \subset X$ such that

$$|\mu|\bigl(X\backslash K(\varepsilon)^\varepsilon\bigr) \le \varepsilon, \quad \forall \mu \in \mathcal{M}.$$

For any fixed number $\delta > 0$ we let $K_n = K(\delta 2^{-n})^{\delta 2^{-n}}$ and obtain the set $K = \bigcap_{n=1}^\infty K_n$, which is compact and satisfies the inequality

$$|\mu|(X\backslash K) \le \sum_{n=1}^\infty |\mu|(X\backslash K_n) \le \delta, \quad \forall \mu \in \mathcal{M},$$

which is a contradiction. Now by using (8.6.1) we find by induction pairwise disjoint compact sets K_j and measures $\mu_j \in \mathcal{M}$ with the following properties:

(1) $|\mu_j|(K_j) > \varepsilon$,

(2) $K_{j+1} \subset X\backslash \bigcup_{i=1}^j K_i^\varepsilon$.

Let $\mu_1 \in \mathcal{M}$ be an arbitrary measure with $\|\mu_1\| > \varepsilon$ (which exists due to (8.6.1)) and let K_1 be a compact set with $|\mu_1|(K_1) > \varepsilon$. By applying (8.6.1) to K_1 we find μ_2. Next we take a compact set $K_2 \subset X\backslash K_1^\varepsilon$ with $|\mu_2|(K_2) > \varepsilon$. By using $Q_2 = K_1 \cup K_2$ we find a measure μ_3 with $|\mu_3|(X\backslash Q_2^\varepsilon) > \varepsilon$ and so on. Property (2) yields that the sets $U_j := K_j^{\varepsilon/4}$ are pairwise disjoint. There exist continuous functions f_j such that $f_j = 0$ outside U_j, $|f_j| \le 1$ and

$$\int_{U_j} f_j \, d\mu_j > \varepsilon.$$

By hypothesis, the sequence $\{\mu_j\}$ contains a weakly convergent subsequence. For notational simplicity we shall assume that the whole sequence $\{\mu_j\}$ is weakly convergent. Let

$$a_n^i = \int_X f_i(x)\, \mu_n(dx).$$

Then $a_n = (a_n^1, a_n^2, \ldots) \in l^1$, since $\sum_{i=1}^\infty |f_i| \le 1$. For every $\lambda = (\lambda_i) \in l^\infty$, the function $f^\lambda = \sum_{i=1}^\infty \lambda_i f_i$ is continuous on X and $|f^\lambda| \le \sup_i |\lambda_i|$. Since the sequence of numbers

$$\langle \lambda, a_n\rangle = \int_X f^\lambda \, d\mu_n$$

converges, the sequence $\{a_n\}$ is Cauchy in the topology $\sigma(l^1, l^\infty)$. According to Corollary 4.5.8 the sequence $\{a_n\}$ converges in the norm of l^1. Hence $\lim\limits_{n\to\infty} a_n^n = 0$, which contradicts our choice of f_n. Thus, $\mathcal{M}$ is uniformly tight.

Suppose that (ii) is fulfilled, $\sup_{\mu\in\mathcal{M}} \|\mu\| = C$ and $\{\mu_n\} \subset \mathcal{M}$. We recall that every norm bounded sequence of linear functionals on a separable normed space contains a pointwise convergent subsequence. Hence every uniformly bounded sequence of measures on a metrizable compact space K contains a weakly convergent subsequence. Let us take an increasing sequence of compact sets K_j such that $|\mu_n|(X\backslash K_j) < 2^{-j}$ for all n. It is clear from what has been said above that by the diagonal process one can find a sequence of measures μ_{n_i} whose restrictions to every K_j converge weakly. Let $f \in C_b(X)$. We show that the sequence

$$\int f\, d\mu_{n_i}$$

is fundamental. Let $\varepsilon > 0$. We may assume that $|f| \le 1$. Let us pick j with $2^{-j} < \varepsilon$. Then

$$\left|\int_X f\, d\mu_{n_i} - \int_X f\, d\mu_{n_m}\right| \le \left|\int_{K_j} f\, d\mu_{n_i} - \int_{K_j} f\, d\mu_{n_m}\right| + 2\varepsilon,$$

whence our claim follows. □

The following assertion is implicitly contained in the proof of the Prohorov theorem.

8.6.3. Corollary. *Every weakly fundamental sequence of Radon measures μ_n on a complete metric space X is uniformly tight. Moreover, if the measures μ_n are nonnegative, then for their uniform tightness it is sufficient that for every bounded Lipschitzian function f the sequence of the integrals*

$$\int_X f\, d\mu_n$$

be fundamental.

Proof. The first assertion has actually been proven. We shall explain the necessary changes in our reasoning in order to cover the second assertion as well. It suffices to take functions f_j such that they are Lipschitzian with a common constant and satisfy the following conditions: $0 \le f_j \le 1$ on X, $f_j = 1$ on K_j, and $f_j = 0$ outside U_j. This is possible, since $U_j = K_j^{\varepsilon/4}$. Moreover, the functions f^λ are Lipschitzian. As μ_n and f_n are nonnegative, the integral of f_n against μ_n is at least $\mu_n(K_n) > \varepsilon$. □

For nonnegative measures, Prohorov's theorem can be proved more concisely. Moreover, as it was first observed by Le Cam (see his theorem below), in the case of nonnegative measures the completeness of X is not needed provided that the limit measure is tight as well. The nonnegativity of measures is essential: we recall that in Example 8.4.6 we constructed a sequence of signed measures μ_n on a separable metric space X (a subset of an interval) that converges weakly to zero such that the measures $|\mu_n|$ converge weakly to a measure that is not tight. It is clear that such a sequence $\{\mu_n\}$ cannot be uniformly tight.

8.6.4. Theorem. *If a sequence of nonnegative Radon measures μ_n on a metric space X converges weakly to a Radon measure μ, then this sequence is uniformly tight.*

PROOF. Let $\varepsilon > 0$. There is a compact set K such that $\mu(X\backslash K) < \varepsilon/4$. Set $G_k = \{x\colon \operatorname{dist}(x, K) < 1/k\}$. By Theorem 8.4.7, there exists an increasing sequence of indices n_k such that

$$\mu_n(X\backslash G_k) < \mu(X\backslash G_k) + \varepsilon/4 < \varepsilon/2, \quad \forall\, n \geq n_k. \tag{8.6.2}$$

For every n with $n_k \leq n \leq n_{k+1}$, we find a compact set $K_n \subset G_k$ such that

$$\mu_n(G_k\backslash K_n) < \varepsilon/4.$$

Let $Q_k = K \cup \bigl(\bigcup_{n=n_k}^{n_{k+1}} K_n\bigr)$ and $K_\varepsilon = \bigcup_{k=1}^\infty Q_k$. We observe that the sets Q_k are compact, $K \subset Q_k \subset G_k$ and $\mu_n(G_k\backslash Q_k) < \varepsilon/4$ if $n_k \leq n \leq n_{k+1}$. It follows by (8.6.2) that $\mu_n(X\backslash Q_k) < \varepsilon$ if $n_k \leq n \leq n_{k+1}$, whence we obtain $\mu_n(X\backslash K_\varepsilon) < \varepsilon$ for all n. It remains to verify that K_ε is compact. Indeed, let $\{x_j\} \subset K_\varepsilon$. If one of the sets Q_k contains an infinite part of $\{x_j\}$, then in Q_k, hence in K_ε, there is a limit point of this sequence. If there is no such Q_k, then there exist two infinite sequences of indices j_m and i_m such that $x_{j_m} \in Q_{i_m}$. Since $Q_{i_m} \subset G_{i_m}$, there exist points $z_m \in K$ such that the distance between x_{j_m} and z_m does not exceed i_m^{-1}. The sequence $\{z_m\}$ has a limit point $z \in K$, which is obviously a limit point of $\{x_{j_m}\}$. □

Prohorov's theorem gives a criterion of the weak sequential compactness of a set of measures on a complete separable metric space. It is natural to ask about weak compactness in the usual topological sense (we recall that in nonmetrizable spaces, compactness is not equivalent to sequential compactness) and about the situation in more general topological spaces. However, before going further, we consider several examples which may help to verify the uniform tightness of measures.

8.6.5. Example. (i) A family $\mathcal{M}$ of probability measures on a complete separable metric space X is uniformly tight precisely when there exists a Borel function $V\colon\ X \to [0, +\infty]$ such that the sets $\{V \leq c\}$, $c < +\infty$, are compact, $\mu(V = +\infty) = 0$ for all $\mu \in \mathcal{M}$, and

$$\sup_{\mu\in\mathcal{M}} \int_X V(x)\,\mu(dx) < \infty.$$

(ii) A family $\mathcal{M}$ of Borel probability measures on a separable reflexive Banach space X is uniformly tight on X with the weak topology precisely when there exists a function $V\colon\ X \to [0, \infty)$ continuous in the norm topology such that

$$\lim_{\|x\|\to\infty} V(x) = \infty \quad \text{and} \quad \sup_{\mu\in\mathcal{M}} \int_X V(x)\,\mu(dx) < \infty.$$

PROOF. The sufficiency of the condition in (i) follows by Chebyshev's inequality:

$$\mu(V > c) \leq c^{-1} \int_X V\,d\mu.$$

In order to see its necessity, we take an increasing sequence of compact sets K_n with $\mu(K_n) > 1 - 2^{-n}$ for all $\mu \in \mathcal{M}$ and define $V = +\infty$ on the complement to the union of K_n, $V = 1$ on K_1, $V = n$ on $K_{n+1}\backslash K_n$, $n \geq 1$. Then, for all $\mu \in \mathcal{M}$, we have

$$\int_X V\, d\mu = \mu(K_1) + \sum_{n=1}^{\infty} n\mu(K_{n+1}\backslash K_n) \leq 1 + \sum_{n=1}^{\infty} n2^{-n}.$$

Claim (ii) is proved similarly, taking into account the compactness of closed balls in any reflexive Banach space with the weak topology. In this case, the function V can be taken in the form $V(x) = f(\|x\|)$ for some increasing to infinity (even concave) positive continuous function f on $[0, +\infty)$. □

8.6.6. Example. A subset K of a metric space X has compact closure if and only if the family of measures $\{\delta_x,\, x \in K\}$ has compact closure in the weak topology.

Now we prove the following reinforced version of one implication in Prohorov's theorem.

8.6.7. Theorem. *Let $\mathcal{K} \subset \mathcal{M}_r(X)$ be a uniformly bounded in the variation norm and uniformly tight family of Radon measures on a completely regular space X. Then $\mathcal{K}$ has compact closure in the weak topology.*

If, in addition, for every $\varepsilon > 0$, there exists a metrizable compact set K_ε such that $|\mu|(X\backslash K_\varepsilon) < \varepsilon$ for all $\mu \in \mathcal{K}$ (which is the case if all compact subsets of X are metrizable), then every sequence in $\mathcal{K}$ contains a weakly convergent subsequence.

PROOF. We consider $\mathcal{K}$ as a subset of the dual space of the Banach space $C_b(X)$ equipped with the weak* topology. By the Banach–Alaoglu theorem (which is applicable by the norm boundedness of $\mathcal{K}$) any infinite set $\mathcal{K}' \subset \mathcal{K}$ has a limit point F. We have to verify that F is representable as the integral with respect to a Radon measure. It is here that we need the uniform tightness. We can assume that $\|\mu\| \leq 1$ for all $\mu \in \mathcal{K}$. Let $\varepsilon > 0$ and let K_ε be a compact set such that $|\mu|(X\backslash K_\varepsilon) < \varepsilon$ for all $\mu \in \mathcal{K}$. If $f \in C_b(X)$, $|f| \leq 1$ and $f = 0$ on K_ε, then

$$|F(f)| \leq \limsup_{\mu \in \mathcal{K}} \biggl| \int_X f\, d\mu \biggr| \leq \varepsilon.$$

By Theorem 7.10.6 the functional F is represented by some Radon measure ν, which is the required limit point of $\mathcal{K}'$ in the weak topology. The second claim has in fact been obtained in the proof of Prohorov's theorem, since we have used there only the metrizability of compact sets K_ε on which the considered sequence of measures is uniformly concentrated. □

In many spaces the uniform tightness is a necessary condition of the weak compactness of families of measures. We shall discuss such spaces in §8.10(ii). Here we establish only the following fact.

8.6.8. Theorem. *Let X be a complete metric space. Then every weakly compact subset of $\mathcal{M}_r(X)$ is uniformly tight.*

PROOF. Suppose that we have a weakly compact set M in $\mathcal{M}_r(X)$ that is not uniformly tight. Let us consider the functions f_j and measures μ_n constructed in the proof of Theorem 8.6.2 (in their construction, we only used the failure of uniform tightness, the fact that all measures in M are Radon and that X is complete). Now, however, we only have the relative weak compactness of $\{\mu_n\}$, which does not mean the existence of a convergent subsequence. Nevertheless, by the relative weak compactness of $\{\mu_n\}$ the sequence $a_n = (a_n^i) \in l^1$, where a_n^i is the integral of f_i against μ_n, is relatively weakly compact in l^1. Indeed, the mapping from $\mathcal{M}_r(X)$ to l^1 that to every measure μ associates the sequence of the integrals of f_i against μ is continuous provided that $\mathcal{M}_r(X)$ and l^1 are equipped with the weak topology. This is clear from the fact that, as observed in the proof of Theorem 8.6.2, for every element $\lambda = (\lambda_i)_{i=1}^{\infty} \in l^{\infty} = (l^1)^*$, the function $f^{\lambda} = \sum_{i=1}^{\infty} \lambda_i f_i$ is continuous and bounded. Therefore, the image of M under this mapping is weakly compact in l^1. It follows that $a_n^n \to 0$, i.e., we arrive again at a contradiction. $\square$

In the general case, unlike the case of a complete metric space, the condition in Theorem 8.6.7 is not necessary: even on a countable nonmetrizable space, a weakly convergent sequence of probability measures may not be uniformly tight.

8.6.9. Example. Let $X = \mathbb{N} \cup \{\infty\}$, where all points in $\mathbb{N}$ are open and the neighborhoods of ∞ have the form $U \cup \{\infty\}$, where U is a subset of $\mathbb{N}$ with density 1, i.e., $\lim\limits_{n\to\infty} N(U,n)/n = 1$, where $N(U,n)$ is the number of points in U not exceeding n. Then the sequence $n^{-1}\sum_{i=1}^{n} \delta_i$ of the arithmetic means of the Dirac measures at the points i converges weakly to Dirac's measure δ_{∞}, but is not uniformly tight. The proof is left as Exercise 8.10.92.

In applications, various special conditions of weak compactness are often useful. For example, for the distributions of random processes in function spaces such conditions can be expressed in terms of the covariance functions, sample moduli of continuity, etc., and for measures on linear spaces, there are efficient conditions in terms of the Fourier transform (see §8.8).

8.6.10. Example. Let $X = \bigcup_{n=1}^{\infty} X_n$ be a locally convex space that is the strict inductive limit of an increasing sequence of closed subspaces X_n, i.e., every X_n is a proper closed subspace in a locally convex space X_{n+1}, and the convex neighborhoods of the origin in X are convex sets V such that $V \cap X_n$ is a neighborhood of the origin in X_n. If a sequence $\{\mu_i\}$ of nonnegative τ-additive (for example, Radon) measures on X converges weakly to a τ-additive measure μ, then for every $\varepsilon > 0$, there exists $n \in \mathbb{N}$ such that $\mu_i(X \backslash X_n) < \varepsilon$ for all $i \in \mathbb{N}$.

Moreover, if a family $\{\mu_\alpha\}$ of nonnegative τ-additive measures on X has compact closure in the weak topology in the space $\mathcal{M}_\tau(X)$, then for every $\varepsilon > 0$, there exists $n \in \mathbb{N}$ such that $\mu_\alpha(X \backslash X_n) < \varepsilon$ for all α.

PROOF. Without loss of generality we may assume that μ_i and μ are probability measures (if $\mu_i(X) \to 0$, then the claim is trivial). If our claim is false, then for every $n \in \mathbb{N}$, there exists $i(n) \in \mathbb{N}$ with $\mu_{i(n)}(X_n) < 1 - \varepsilon$. Passing to a new sequence of measures, we may assume that $i(n) = n$. We pick $m \in \mathbb{N}$ such that $\mu(X_m) > 1 - \varepsilon/2$. Set $k_1 := m$. Next we find $k_2 > m$ with $\mu_m(X_{k_2}) > 1 - \varepsilon/2$. Then we find a convex symmetric open set U_1 in X_{k_2} such that $X_m \subset U_1$ and $\mu_m(U_1) < 1 - \varepsilon$. Such a set U_1 indeed exists. To show this, we observe that by the Hahn–Banach theorem the subspace X_m is the intersection of all closed hyperplanes containing it. By the τ-additivity of μ_m, there exists a finite collection of closed hyperplanes $L_1, \ldots, L_p$ in X_{k_2} such that $X_m \subset \bigcap_{i=1}^p L_i$ and $\mu_m\bigl(\bigcap_{i=1}^p L_i\bigr) < 1 - \varepsilon$. Then $L_i = l_i^{-1}(0)$ for some $l_i \in X_{k_2}^*$, and the set $\bigcap_{i=1}^p l_i^{-1}(-\delta, \delta)$ can be taken for U_1 provided $\delta > 0$ is sufficiently small. Next we take $k_3 \geq k_2$ with $\mu_{k_2}(X_{k_3}) > 1 - \varepsilon/2$. There exists a convex symmetric neighborhood of zero $W \subset X_{k_3}$ such that $W \cap X_{k_2} = U_1$ (see Schaefer [**1661**, II.6.4, Lemma]). As above, there exists a convex symmetric open set V in the space X_{k_3} such that $X_{k_2} \subset V$ and $\mu_{k_2}(V) < 1 - \varepsilon$. Set $U_2 := W \cap V$. Continuing the described process by induction, we obtain an increasing sequence of indices $k_n \geq n$ such that every space $X_{k_{n+1}}$ contains a convex symmetric open set U_n with the following properties: (1) $U_n \cap X_{k_n} = U_{n-1}$, (2) $\mu_{k_n}(U_n) < 1 - \varepsilon$, $\mu_{k_n}(X_{k_{n+1}}) > 1 - \varepsilon/2$. By the definition of the strict inductive limit, the set $U = \bigcup_{n=1}^\infty U_n$ is a neighborhood of zero in X. By construction, for every n one has

$$\mu_{k_n}(U) < \mu_{k_n}(U \cap X_{k_{n+1}}) + \varepsilon/2 = \mu_{k_n}(U_n) + \varepsilon/2 < 1 - \varepsilon/2,$$

which contradicts weak convergence (see Corollary 8.2.10), since we have the estimate $\mu(U) > 1 - \varepsilon/2$. In the case of a relatively weakly compact family $\{\mu_\alpha\}$ the reasoning is similar. We construct a sequence $\{\mu_{\alpha(n)}\}$ as above and denote by μ its weak limit point. The previous choice of U leads again to a contradiction with Corollary 8.2.10, since there exists a subnet $\{\mu_\beta\}$ in $\{\mu_{\alpha(n)}\}$ convergent weakly to μ. □

Now we give a simple criterion of relative weak compactness in the space of nonnegative Baire measures on an arbitrary space X. We shall say that a sequence of functionally closed sets Z_n in a topological space X is regular if $X = \bigcup_{n=1}^\infty Z_n$, $Z_n \subset Z_{n+1}$, and there exist functionally open sets U_n such that $Z_n \subset U_n \subset Z_{n+1}$.

8.6.11. Theorem. *A bounded set $M \subset \mathcal{M}_\sigma^+(X)$ has compact closure in the weak topology precisely when*

$$\lim_{n\to\infty} \sup_{\mu \in M} \int_X f_n \, d\mu = 0$$

for every sequence of functions $f_n \in C_b(X)$ pointwise decreasing to 0. An equivalent condition: for every regular sequence of functionally closed sets Z_n

$$\lim_{n\to\infty} \sup_{\mu \in M} \mu(X \backslash Z_n) = 0.$$

PROOF. Let the first condition be fulfilled. The bounded set M has the compact closure M' in the space $C_b(X)^*$. Every element $\mu \in M'$ belongs to $\mathcal{M}_\sigma^+(X)$ by Theorem 7.10.1. Conversely, if M has the compact closure M' in the weak topology, then all measures in M' are nonnegative and the functions

$$\mu \mapsto \int_X f_n \, d\mu$$

on M' decrease to 0. By Dini's theorem they converge to 0 uniformly on M'.

If $\{Z_n\}$ is a regular sequence, then there exists a sequence $\{f_n\} \subset C_b(X)$ with $f_n \downarrow 0$ such that $f_n = 1$ on $X \backslash Z_n$ (see Lemma 6.3.2). Then for all $\mu \in M$ one has

$$\mu(X \backslash Z_n) \leq \int_X f_n \, d\mu.$$

Conversely, if functions $f_n \in C_b(X)$ decrease to 0, then, given $\varepsilon > 0$, let $U_n = \{f_n < \varepsilon\}$. It is readily verified that there exists a regular sequence of functionally closed sets Z_n with $Z_n \subset U_n$; one can take sets $Z_n := \{\min(f_n, \varepsilon) \leq \varepsilon - 1/n\}$. Then

$$\int_X f_n \, d\mu \leq \varepsilon \mu(X) + \mu(X \backslash Z_n),$$

which shows the equivalence of both conditions. □

8.7. Weak sequential completeness

In this section, we show that any weakly fundamental sequence of Baire measures converges weakly to some Baire measure, i.e., the space of Baire measures is weakly sequentially complete.

8.7.1. Theorem. *Suppose that a sequence of Baire measures μ_n on a topological space X is weakly fundamental. Then $\{\mu_n\}$ converges weakly to some Baire measure on X.*

PROOF. By the Banach–Steinhaus theorem the formula

$$L(\varphi) = \lim_{n\to\infty} \int \varphi \, d\mu_n, \quad \varphi \in C_b(X),$$

defines a continuous linear functional on $C_b(X)$. According to Theorem 7.10.1, this functional is represented by a Baire measure under the following condition: $L(\varphi_j) \to 0$ for every sequence of functions $\varphi_j \in C_b(X)$ that decreases pointwise to zero. Suppose that this condition is not fulfilled, i.e., the sequence $L(\varphi_j)$ does not converge to zero. We may assume that $0 \leq \varphi_n \leq 1$ for all n. Set $I = [0,1]^\infty$ and consider the mapping $F\colon X \to I$, $F(x) = \bigl(\varphi_j(x)\bigr)_{j=1}^\infty$. We equip the space $Y = F(X)$ with the topology induced from I (since I is metrizable, then Y is metrizable as well). It is clear that F is continuous as a mapping from X to Y, hence the sequence of measures $\nu_n := \mu_n \circ F^{-1}$ on Y is weakly fundamental (for all $\psi \in C_b(Y)$ we have $\psi \circ F \in C_b(X)$). The natural extensions of the measures ν_n to I will again be denoted by ν_n. It is clear that on the compact space I the measures ν_n converge weakly to some measure ν. One has

$$\int_I x_j\,\nu(dx) = \lim_{n\to\infty}\int_I x_j\,\nu_n(dx) = \lim_{n\to\infty}\int_X \varphi_j(x)\,\mu_n(dx) = L(\varphi_j).$$

In order to obtain a contradiction with the fact that $L(\varphi_j) \not\to 0$, it suffices to establish that the measure ν is concentrated on the set

$$I_0 := \bigl\{x = (x_j) \in I\colon\ \lim_{j\to\infty} x_j = 0\bigr\}.$$

This will be done if we verify that $|\nu|(K) = 0$ for every compact set K in $I\backslash I_0$. Let $\varepsilon > 0$. The set $U = I\backslash K$ is open. Since $Y \subset I_0 \subset U$, it follows that the measures ν_n on U also form a weakly fundamental sequence. We recall that U is a Polish space (as an open subset in a Polish space). By Prohorov's theorem, the sequence $\{\nu_n\}$ is uniformly tight on U, i.e., one can find a compact set $Q \subset U$ such that $|\nu_n|(U\backslash Q) < \varepsilon$ for all n. Then $|\nu|(K) \le |\nu|(I\backslash Q) \le \liminf\limits_{n\to\infty} |\nu_n|(I\backslash Q) \le \varepsilon$ by weak convergence on I (see Theorem 8.4.7) and the equality $|\nu_n|(I\backslash Q) = |\nu_n|(U\backslash Q)$. Since ε is arbitrary, one has $|\nu|(K) = 0$, as required. □

The proof of the next assertion is left as Exercise 8.10.67.

8.7.2. Example. Let $\{x_n\}$ be a sequence in a metric space X such that the sequence of measures δ_{x_n} is weakly fundamental. Then $\{x_n\}$ converges in the space X.

It should be noted that although the weak topology on $\mathcal{P}_\tau(X)$ is generated by a metric d (for example, by the Lévy–Prohorov and Kantorovich–Rubinshtein metrics), the collections of Cauchy sequences in this topology and such a metric may be different. For example, if a separable metric space X does not admit a complete metric, then there exists a sequence of measures $\mu_n \in \mathcal{P}_\sigma(X)$ that is fundamental with respect to d, but has no limit (otherwise $\mathcal{P}_\sigma(X)$ and hence X would be Polish). This sequence $\{\mu_n\}$ is not fundamental in the weak topology, since the latter is sequentially complete.

8.8. Weak convergence and the Fourier transform

In this section, we are concerned with characterizations of weak convergence and weak compactness in terms of characteristic functionals (Fourier transforms). We begin with the following theorem due to P. Lévy.

8.8.1. Theorem. (i) *A sequence $\{\mu_j\}$ of probability measures on $\mathbb{R}^d$ converges weakly precisely when the sequence of their characteristic functionals $\widetilde{\mu}_j$ converges at every point and the function $\varphi(x) := \lim\limits_{j\to\infty}\widetilde{\mu}_j(x)$ is continuous at the origin. In that case, φ is the characteristic functional of a probability measure μ that is the limit of the measures μ_j in the weak topology.*

(ii) *A family M of probability measures on $\mathbb{R}^d$ is uniformly tight if and only if the family of functions $\widetilde{\mu}$, $\mu \in M$, is uniformly equicontinuous on $\mathbb{R}^d$ (the uniform equicontinuity at the origin is enough).*

PROOF. (i) Weak convergence of measures yields pointwise convergence of their characteristic functionals. Let us prove the converse. It is easy to observe that estimates (3.8.6) and (3.8.7), obtained in Chapter 3, along with pointwise convergence of the characteristic functionals and the dominated convergence theorem ensure the uniform tightness of the sequence $\{\mu_j\}$. This yields weak convergence of μ_j to μ. Claim (ii) is proven similarly by using the same estimates (3.8.6) and (3.8.7). □

8.8.2. Remark. In assertion (i), one cannot omit the assumption of continuity of φ. Indeed, for every n, the function $(\cos x)^{2n}$ is the characteristic functional of the $2n$-fold convolution of the probability measure ν that assigns the value $1/2$ to the points -1 and 1. These functions converge pointwise to the function φ equal to 1 at the points πk and 0 at all other points. It is clear that φ is not a characteristic functional because of its discontinuity. Let us also note that the function φ in (i) always has a continuous modification which is the characteristic functional of some nonnegative measure μ (since it is measurable and positive definite), but this measure may not be a probability measure (in the above example $\mu = 0$). Hence in place of continuity of φ one can require that φ be almost everywhere equal to the characteristic functional of some probability measure.

Now we turn to infinite-dimensional spaces. Corollary 7.13.10 yields the following assertion.

8.8.3. Theorem. *Let X be a locally convex space equipped with the strong topology $\beta(X, X^*)$. Let a family M of Radon probability measures on X be such that their characteristic functionals are equicontinuous at the point 0 in the topology $\mathcal{T}(X^*, X)$. Then M has compact closure in the weak topology.*

8.8.4. Corollary. *Let X be a reflexive nuclear space and let M be a family of Radon probability measures on X^* such that their characteristic functionals are equicontinuous at zero. Then M has compact closure in the weak topology.*

This corollary is applicable to such spaces X^* as the classical spaces of distributions $\mathcal{S}'(\mathbb{R}^d)$ and $\mathcal{D}'(\mathbb{R}^d)$ (see the definition in Exercise 6.10.27).

8.9. Spaces of measures with the weak topology

In this section, we discuss some basic topological properties of spaces of measures on a topological space X, in particular, connections between the properties of X and the corresponding properties of the spaces of measures. The most natural connections with topological concepts arise when the spaces of measures are equipped with the weak topology. In applications, the following problems related to spaces of measures are most important:

(1) completeness and sequential completeness;

(2) compactness conditions;

(3) metrizability and separability;

(4) existence of some additional properties, for example, the membership in the class of Souslin spaces.

Since we are interested in the weak topology, it is reasonable to consider completely regular spaces. For the metric case, see also §8.10(viii).

8.9.1. Remark. Suppose that a completely regular space X is homeomorphically embedded into a completely regular space Y. For every measure $\mu \in \mathcal{M}_\tau(X)$, let $\widehat{\mu}$ denote its extension to $\mathcal{B}(Y)$ defined by $\widehat{\mu}(B) := \mu(B \cap X)$, $B \in \mathcal{B}(Y)$. Then $\widehat{\mu} \in \mathcal{M}_\tau(Y)$. The mapping $\mu \mapsto \widehat{\mu}$ on $\mathcal{M}_\tau^+(X)$ is a homeomorphic embedding, which is clear from Corollary 8.2.4 and the fact that the open sets in X are precisely the intersections of X with open sets in Y. Moreover, by Theorem 8.4.7, the same is true for the space $\mathcal{M}_\tau^1(X)$ of all signed measures in $\mathcal{M}_\tau(X)$ whose total variation is 1. However, this mapping need not be a homeomorphic embedding of the whole space $\mathcal{M}_\tau(X)$. For example, if $X = (0, 1]$ and $Y = [0, 1]$ with the their standard topologies, then the sequence of measures $\delta_{1/(2n)} - \delta_{1/(2n+1)}$ weakly converges to zero on Y, but not on X because there is a bounded continuous function f on X such that $f\bigl(1/(2n)\bigr) = 1$ and $f\bigl(1/(2n + 1)\bigr) = 0$ for all n. On the space $\mathcal{P}_\sigma(X)$, the mapping $\mu \mapsto \widehat{\mu}$ need not be even injective (see Wheeler [**1979**, §14]). If X is closed and Y is normal, then the embedding of $\mathcal{M}_\tau(X)$ into $\mathcal{M}_\tau(Y)$ is homeomorphic, which is straightforward.

8.9.2. Lemma. *Let X be completely regular. Then X is homeomorphic to the set of all Dirac measures on X and this set is closed in $\mathcal{M}_\tau(X)$ and in $\mathcal{M}_t(X)$ as well as in the corresponding subspaces of nonnegative and probability measures equipped with the weak topology.*

PROOF. Let $j(x) = \delta_x$. Then the mapping $j\colon X \to \mathcal{M}_\sigma(X)$ is a topological embedding. Indeed, according to Example 8.1.5, a net $\{x_\alpha\}$ converges to x precisely when the net $\{j(x_\alpha)\}$ converges to $j(x)$.

Suppose now that a τ-additive measure μ is a limit point of the set of Dirac measures in the weak topology. Then, there exists a net $\{\delta_{x_\alpha}\}$ weakly convergent to μ, in particular, μ is a probability measure. Let us take an arbitrary point x in the topological support of μ (which exists by its τ-additivity). We show that the net $\{x_\alpha\}$ converges to x. If this is not the case, then outside some neighborhood U of the point x, there is a subnet $\{x'_\alpha\}$ of the initial net. There exists a bounded nonnegative continuous function f that equals 1 on some neighborhood V of the point x and vanishes outside U. Since $f(x'_\alpha) = 0$, one has

$$\int_X f\, d\mu = 0,$$

whence we obtain $\mu(V) = 0$ contrary to the fact that x belongs to the support. Thus, $x_\alpha \to x$, whence it follows that $\mu = \delta_x$. □

In Exercise 8.10.80 it is proposed to construct an example of a completely regular space X such that the set of all Dirac measures is not closed in the space $\mathcal{M}_\sigma^+(X)$.

It is worth recalling that if X is completely regular, then the spaces $\mathcal{M}_t(X)$ and $\mathcal{M}_r(X)$ regarded as subspaces in $C_b(X)^*$ coincide because every tight Baire measure has a unique Radon extension. Certainly, in general $\mathcal{M}_t(X)$ and $\mathcal{M}_r(X)$ may not coincide as spaces of measures: the point is that $\mathcal{M}_t(X)$ consists of Baire measures (but the distinction disappears if we consider only Baire sets).

8.9.3. Theorem. (i) *Let X be a compact space. Then the spaces $\mathcal{P}_\sigma(X) = \mathcal{P}_t(X)$ and $\mathcal{P}_\tau(X) = \mathcal{P}_r(X)$ are compact in the weak topology.*

(ii) *If X is completely regular and $\mathcal{P}_t(X)$ (or $\mathcal{P}_\tau(X)$) is compact in the weak topology, then X is compact as well.*

PROOF. The compactness of $\mathcal{P}_t(X)$ is an immediate corollary of the Banach–Alaoglu theorem on the weak* compactness of balls in the dual space and the Riesz theorem identifying the dual of $C(X)$ with $\mathcal{M}_t(X)$. The compactness of the space $\mathcal{P}_r(X)$ (which coincides with $\mathcal{P}_\tau(X)$ by the compactness of X, see Proposition 7.2.2) is clear from the above remark. The necessity of compactness of X in the second assertion follows by Lemma 8.9.2. $\square$

We observe that in (ii) one cannot replace $\mathcal{P}_t(X)$ by $\mathcal{P}_\sigma(X)$. One can verify that the space in Exercise 8.10.80 gives a counter-example.

8.9.4. Theorem. *Let X be completely regular.*

(i) *The space $\mathcal{M}_\tau^+(X)$ with the weak topology is metrizable if and only if X is metrizable. In that case, the metrizability of $\mathcal{M}_\tau^+(X)$ by a complete metric is necessary and sufficient for the metrizability of X by a complete metric. The analogous assertions are valid for $\mathcal{P}_\tau(X)$, $\mathcal{P}_t(X)$, and $\mathcal{M}_t^+(X)$ in place of $\mathcal{M}_\tau^+(X)$.*

(ii) *If X is separable, then the spaces of measures $\mathcal{M}_\sigma(X)$, $\mathcal{M}_\tau(X)$ and $\mathcal{M}_t(X)$ are separable in the weak topology as well as the corresponding subspaces of nonnegative and probability measures.*

PROOF. (i) Lemma 8.9.2 yields that the aforementioned properties of the spaces of measures imply the respective properties of X. Let us show the converse assertion. Theorem 8.3.2 gives at once the metrizability of $\mathcal{M}_\tau^+(X)$ with the weak topology. In order to verify the completeness of $\mathcal{M}_\tau^+(X)$ in the metric d_0 from Theorem 8.3.2 in the case of a complete space X, suppose that a sequence of nonnegative Radon (which in this case is equivalent to the τ-additivity) measures μ_n is fundamental in the metric d_0. Then the sequence

$$\int_X f\,d\mu_n$$

converges for every bounded Lipschitzian function f. According to Corollary 8.6.3 the measures μ_n are uniformly tight. Therefore, the measures μ_n converge weakly to some Radon measure μ. Hence $d_0(\mu_n, \mu) \to 0$. The case of the spaces $\mathcal{M}_t^+(X)$ and $\mathcal{P}_t(X)$ follows by the same reasoning (note that if X is a complete metric space, then $\mathcal{M}_\tau(X) = \mathcal{M}_t(X)$, and $\mathcal{M}_t(X) \subset \mathcal{M}_\tau(X)$ for any metric space).

(ii) If X contains an everywhere dense countable set of points x_j, then the countable set of all finite linear combinations of the measures δ_{x_j} with rational coefficients is everywhere dense in $\mathcal{M}_\sigma(X)$, and its subset corresponding to nonnegative coefficients is everywhere dense in $\mathcal{M}_\sigma^+(X)$. Linear combinations with nonnegative coefficients whose sum is 1 give a countable everywhere dense set in $\mathcal{P}_\sigma(X)$. This also shows the separability of $\mathcal{M}_\tau(X)$ and $\mathcal{M}_t(X)$ and their subspaces $\mathcal{M}_\tau^+(X)$, $\mathcal{M}_t^+(X)$, $\mathcal{P}_\tau(X)$, and $\mathcal{P}_t(X)$. □

The reader is warned that the separability of $\mathcal{P}_t(X)$ with the weak topology does not yield the separability of X, and the separability of the whole space $\mathcal{M}_t(X)$ with the weak topology does not guarantee the separability of $\mathcal{P}_t(X)$ even if X is compact (see §8.10(vi)).

8.9.5. Theorem. *If E is a Polish space, then so is the subspace $\mathcal{M}^1(E)$ in $\mathcal{M}(E) := \mathcal{M}_\sigma(E)$ consisting of all measures μ with $\|\mu\| = 1$.*

Proof. We recall that the space E is homeomorphic to a G_δ-set in the compact space $Q = [0,1]^\infty$. Hence it suffices to consider the case where E is a G_δ-set in Q. Let $\mathcal{P}(E) = \mathcal{P}_\sigma(E)$, $\mathcal{M}(Q) = \mathcal{M}_\sigma(Q)$. The unit ball

$$T = \bigl\{\mu \in \mathcal{M}(Q)\colon\ \|\mu\| \le 1\bigr\}$$

in $\mathcal{M}(Q)$ is compact and metrizable in the weak topology. Our set $\mathcal{M}^1(E)$ in the metrizable compact space T is the union of the following three sets: $\mathcal{P}(E)$, $-\mathcal{P}(E)$, and $D := \mathcal{M}^1(E)\backslash\bigl(\mathcal{P}(E) \cup \bigl(-\mathcal{P}(E)\bigr)\bigr)$. The first two sets are Polish spaces and hence are G_δ-sets (see §6.1). We verify that D is a G_δ-set as well. Then the union of three G_δ-sets will be a set of the same type in the metrizable compact space T, hence a Polish space. We recall that as noted in Remark 8.9.1, the weak topology on $\mathcal{M}^1(E)$ coincides with the induced weak topology of T.

Since the space $\mathcal{P}(E)$ is Polish, the space $Z := \mathcal{P}(E)\times\mathcal{P}(E)\times(0,1)$ is Polish as well. Let us consider the mapping $\psi\colon\ (\mu,\nu,\alpha) \mapsto \alpha\mu - (1-\alpha)\nu$ from Z to $\mathcal{M}(E)$. This mapping is continuous if the spaces of measures are equipped with the weak topology. Let $U_r = \{\mu \in \mathcal{M}(E)\colon\ \|\mu\| \le r\}$. The sets U_r are closed in the weak topology. Let

$$H := \bigl\{(\mu,\nu,\alpha) \in Z\colon\ \|\alpha\mu - (1-\alpha)\nu\| = 1\bigr\}.$$

The set H is the intersection of the sequence of open sets $\psi^{-1}\bigl(\mathcal{M}(E)\backslash U_{1-1/n}\bigr)$, i.e., is a G_δ-set, hence a Polish space. Now it is important to observe that the mapping ψ homeomorphically maps H onto the set D. Indeed, if measures $\mu,\nu \in \mathcal{P}(E)$ are such that $\|\alpha\mu - (1-\alpha)\nu\| = 1$, then it is easy to see that they are mutually singular (see Exercise 3.10.33). It is clear from this that if $\alpha\mu - (1-\alpha)\nu = \alpha'\mu' + (1-\alpha')\nu'$ has the variation 1 for some $\alpha,\alpha' \in (0,1)$ and $\mu,\mu',\nu,\nu' \in \mathcal{P}(E)$, then $\alpha = \alpha'$, $\mu = \mu'$ and $\nu = \nu'$. Thus, ψ maps H one-to-one onto D (that $\psi(H) = D$ is obvious from the decomposition $\mu = \mu^+ - \mu^-$, where $\mu^+(E) + \mu^-(E) = 1$ and $\mu^+(E) > 0$, $\mu^-(E) > 0$). Finally, the mapping $\psi^{-1}\colon\ D \to H$ is continuous. Indeed, let a net of measures μ_τ from D converge weakly to a measure μ in D. By Theorem 8.4.7 we obtain

$\mu_\tau^+ \to \mu^+$ and $\mu_\tau^- \to \mu^-$ in the weak topology. This yields weak convergence of the measures $\psi^{-1}(\mu_\tau)$ to the measure $\psi^{-1}(\mu)$ since $\mu_\tau^+/\mu_\tau^+(X) \to \mu^+/\mu^+(X)$ and $\mu_\tau^-/\mu_\tau^-(X) \to \mu^-/\mu^-(X)$ due to $\mu^+(X) > 0$ and $\mu^-(X) > 0$. Thus, D is homeomorphic to the G_δ-set H in a Polish space, which completes the proof. □

8.9.6. Theorem. *Let X be completely regular. If X is a Souslin (or Lusin) space, then so are the spaces $\mathcal{M}_\sigma(X)$, $\mathcal{M}_\sigma^+(X)$ and $\mathcal{P}_\sigma(X)$ with the weak topology (note that these spaces consist of Radon measures). Conversely, if one of the spaces $\mathcal{M}_t(X)$, $\mathcal{M}_t^+(X)$ or $\mathcal{P}_t(X)$ is Souslin (or Lusin), then so is the space X.*

PROOF. By assumption, we have a Polish space E and a continuous surjection $\varphi\colon E \to X$. The induced mapping $\widehat{\varphi}\colon \mathcal{M}_\sigma(E) \to \mathcal{M}_\sigma(X)$ is continuous. It will be shown in Chapter 9 (see Theorem 9.1.5) that the mapping $\widehat{\varphi}$ is surjective. If φ is injective, then $\widehat{\varphi}$ is injective as well. Hence it remains to prove that the space $\mathcal{M}_\sigma(E)$ is Lusin. This follows by the previous theorem, since $\mathcal{M}_\sigma(E) = 0 \cup \bigl(\mathcal{M}^1(E)\times(0,\infty)\bigr)$. □

If X is not completely regular, then an analogous theorem is valid for the A-topology considered in §8.10(iv).

8.9.7. Proposition. *Let X be completely regular. The space X^∞ is homeomorphic to a closed subset in $\mathcal{M}_\tau^+(X)$ and to a subset in $\mathcal{M}_t^+(X)$.*

The proof is delegated to Exercise 8.10.96.

Thus, every topological property that is inherited by closed sets but is not preserved by countable products does not extend from X to the spaces $\mathcal{M}_\tau^+(X)$, and $\mathcal{M}_t^+(X)$. The normality and the Lindelöf property deliver such examples. For the same reason the spaces $\mathcal{M}_\tau^+(X)$ and $\mathcal{M}_t^+(X)$ may not be Radon spaces for a Radon space X (even for compact X).

Now we prove a useful result on measurability in spaces of measures established in Hoffmann-Jørgensen [**844**].

8.9.8. Proposition. *Suppose that f is a bounded Baire function on a topological space X. Then the following functions on the space $\mathcal{M}_\sigma(X)$ with the weak topology are Borel measurable:*

$$F_1(\mu) = \int_X f\,d\mu, \quad F_2(\mu) = \int_X f\,d\mu^+,$$

$$F_3(\mu) = \int_X f\,d\mu^-, \quad F_4(\mu) = \int_X f\,d|\mu|.$$

If X is completely regular, then these functions are Borel on $\mathcal{M}_\tau(X)$ and $\mathcal{M}_t(X)$ with the weak topology for every bounded Borel function f. Finally, if, in addition, f is nonnegative and lower semicontinuous, then the functions F_2, F_3, and F_4 are lower semicontinuous on $\mathcal{M}_\tau(X)$ and $\mathcal{M}_t(X)$.

PROOF. It is readily seen that it suffices to verify our claim for F_2. Clearly, it reduces to the case of a simple function and then to the case of an indicator function. Let $f = I_U$, where the set U is functionally open. Then

$$\mu^+(U) = \sup\Bigl\{\int_X \varphi\, d\mu\colon\ \varphi \in C_b(X),\, 0 \le \varphi \le I_U\Bigr\}.$$

Indeed, given $\varepsilon > 0$, one can find a functionally open set $W \subset U$ such that for the set X^+ from the Hahn decomposition we obtain $U \cap X^+ \subset W$ and $|\mu|\bigl(W\backslash(U \cap X^+)\bigr) < \varepsilon$. Next we find in $U \cap X^+$ a functionally closed set Z for which $|\mu|\bigl((U\cap X^+)\backslash Z\bigr) < \varepsilon$. There exists a function $\varphi \in C_b(X)$ such that

$$0 \le \varphi \le 1, \quad \varphi|_Z = 1, \quad \varphi|_{X\backslash W} = 0.$$

Then

$$\Bigl|\int \varphi\, d\mu - \mu^+(U)\Bigr| \le 3\varepsilon.$$

The functions

$$\mu \mapsto \int \varphi\, d\mu, \quad \text{where } \varphi \in C_b(X),$$

are continuous on $\mathcal{M}_\sigma(X)$. Hence the function F_2 is lower semicontinuous. The class $\mathcal{E}$ of all sets $E \in \mathcal{B}a(X)$ for which the function F_2 generated by $f = I_E$ is Borel is σ-additive. By Theorem 1.9.3 we obtain $\mathcal{E} = \mathcal{B}a(X)$, since the class of all functionally open sets admits finite intersections and the σ-algebra generated by it is $\mathcal{B}a(X)$.

Let us consider the space $\mathcal{M}_\tau(X)$ in the case of a completely regular space X. The preceding reasoning remains valid if we take arbitrary open sets U. The indicated equality for $\mu^+(U)$ remains true by the τ-additivity of μ, since $\mu^+(U)$ equals $\sup\{\mu^+(V)\}$, where sup is taken over all functionally open sets $V \subset U$. Finally, the assertion about the lower semicontinuity is clear from the proof, since any lower semicontinuous nonnegative function f can be uniformly approximated by finite linear combinations of the indicators of open sets with nonnegative coefficients (see the proof of Lemma 7.2.6). □

8.9.9. Corollary. *Let X be a completely regular space. Then for every τ-additive measure Ψ on $\mathcal{M}_\tau(X)$ with respect to which the function $q \mapsto \|q\|$ is integrable, the measures*

$$\sigma(B) := \int_{\mathcal{M}_\tau(X)} q(B)\, \Psi(dq), \quad \eta(B) := \int_{\mathcal{M}_\tau(X)} |q|(B)\, |\Psi|(dq)$$

on $\mathcal{B}(X)$ are defined and τ-additive. Hence, for any $B \in \mathcal{B}(X)$ and $\varepsilon > 0$, there is an open set $U \supset B$ such that $|\Psi|\bigl(q\colon\ |q|(U\backslash B\bigr) > \varepsilon) < \varepsilon$.

PROOF. According to Proposition 8.9.8, for every $B \in \mathcal{B}(X)$, the functions $q \mapsto q(B)$ and $q \mapsto |q|(B)$ are Borel measurable on $\mathcal{M}_\tau(X)$. By the integrability of $q \mapsto \|q\|$ the measures σ and η are defined. Let us show that $\eta \in \mathcal{M}_\tau(X)$. Suppose a net of open sets $U_\lambda \subset X$ increases to an open set U. Then the net of functions $q \mapsto |q|(U_\lambda)$ increases to the function

$q \mapsto |q|(U)$ by the τ-additivity of $|q|$, and these functions are lower semicontinuous on $\mathcal{M}_\tau(X)$. Now we can use Lemma 7.2.6. The same reasoning applies to q^+ and q^- in place of $|q|$, which yields the τ-additivity of σ. $\square$

8.10. Supplements and exercises

(i) Weak compactness (217). (ii) Prohorov spaces (219). (iii) Weak sequential completeness of spaces of measures (226). (iv) The A-topology (226). (v) Continuous mappings of spaces of measures (227). (vi) The separability of spaces of measures (230). (vii) Young measures (231). (viii) Metrics on spaces of measures (232). (ix) Uniformly distributed sequences (237). (x) Setwise convergence of measures (241). (xi) Stable convergence and ws-topology (246). Exercises (249).

8.10(i). Weak compactness

A useful technical result characterizing weak compactness for nonnegative measures was obtained in Topsøe [**1874**].

8.10.1. Theorem. *Let X be a completely regular space. Then a set $M \subset \mathcal{M}_t^+(X)$ has compact closure in the weak topology if and only if:*

(i) *M is uniformly bounded,*

(ii) *for every $\varepsilon > 0$ and every collection $\mathcal{U}$ of open sets with the property that every compact set is contained in a set from $\mathcal{U}$, there exist $U_1, \ldots, U_n \in \mathcal{U}$ such that $\inf\bigl\{\mu(X\backslash U_i)\colon\ 1 \le i \le n\bigr\} < \varepsilon$ for all $\mu \in M$.*

8.10.2. Corollary. *Let $Y \subset X$ be closed and let a set $\mathcal{M} \subset \mathcal{M}_t^+(X)$ have compact closure in the weak topology in $\mathcal{M}_t^+(X)$. Then the family of restrictions of the measures from $\mathcal{M}$ to Y has compact closure in the weak topology in $\mathcal{M}_t^+(Y)$.*

This corollary is rather unexpected (although for Polish spaces it is obvious from Prohorov's criterion and for normal spaces it follows from Theorem 8.6.11), since weak convergence does not imply convergence on closed sets. In particular, the limit of restrictions of measures from a weakly convergent sequence to a closed set may not coincide with the restriction of the limit of that sequence (as in Example 8.1.4). In the case of a complete metric space, the previous corollary holds for signed measures as well due to Theorem 8.6.8, but it fails for signed measures on general spaces.

8.10.3. Example. Let $X = \bigl([0,\omega_1]\times[0,\omega_0]\bigr)\backslash(\omega_1,\omega_0)$, where ω_0 is the ordinal corresponding to $\mathbb{N}$, ω_1 is the first uncountable ordinal, and both intervals of ordinals are equipped with the natural order topology. Let

$$Y = \{(\omega_1, 2n)\}_{n=1}^{\infty}, \quad M = \{\delta(\omega_1, 2n) - \delta(\omega_1, 2n+1)\}_{n=1}^{\infty} \cup \{0\}.$$

The set M is weakly compact in $\mathcal{M}_t(X)$, but the restrictions of measures from M to Y form a discrete set in $\mathcal{M}_t(Y)$ without accumulation points.

The next three theorems are proved in Hoffmann-Jørgensen [**844**].

8.10.4. Theorem. *Let X be a completely regular space that admits a continuous injective mapping to a metric space. Then, for every set M in the space $\mathcal{M}_t^+(X)$ with the weak topology, the following conditions are equivalent*: (i) *every infinite sequence in M has a limit point in $\mathcal{M}_t^+(X)$*; (ii) *every infinite sequence in M has a convergent subsequence in $\mathcal{M}_t^+(X)$*; (iii) *the closure of M is compact*; (iv) *the closure of M is compact and metrizable.*

PROOF. It suffices to show that (i) implies (iv). Let $h\colon X \to Y$ be a continuous injective mapping to a metric space Y. Then the mapping $\widehat{h}\colon \mathcal{M}_t^+(X) \to \mathcal{M}_t^+(Y)$ is continuous in the weak topology and injective (because any measure in $\mathcal{M}_t(X)$ has a unique Radon extension, and h is a homeomorphism on any compact set). Since $\mathcal{M}_t^+(Y)$ with the weak topology is metrizable, the claim follows by Exercise 6.10.82. $\square$

Every Souslin completely regular space satisfies the above hypothesis on X. On the other hand, under this hypothesis, all compact sets in X are metrizable.

8.10.5. Theorem. *Let X be a completely regular space. Then, every weakly compact set M in $\mathcal{M}_\tau(X)$ is contained in a centrally symmetric convex weakly compact set. In particular, the closed convex envelope of M is weakly compact.*

PROOF. According to Corollary 8.9.9, for every Radon measure Ψ on the compact set M, the measure

$$T(\Psi)(B) := \int_M \mu(B)\,\Psi(d\mu)$$

is τ-additive on X. For every function $f \in C_b(X)$, one has

$$\int_X f(x)\,T(\Psi)(dx) = \int_M \int_X f(x)\,\mu(dx)\,\Psi(d\mu).$$

Hence the mapping $T\colon \mathcal{M}_r(M) \to \mathcal{M}_\tau(X)$ is continuous in the weak topology. The closed unit ball K in $\mathcal{M}_r(M)$ is compact in the weak topology. Hence $T(K)$ is a centrally symmetric convex compact set. It remains to observe that $M \subset T(K)$, since one has $\mu = T(\delta_\mu)$ for all $\mu \in M$. $\square$

8.10.6. Theorem. *Let X be a completely regular space such that one has $\mathcal{M}_\tau(X) = \mathcal{M}_r(X)$, and let Ψ be a Radon measure on the space $\mathcal{M}_\tau(X)$ with the weak topology. Then, for every Borel set M in $\mathcal{M}_r(X)$ and every $\varepsilon > 0$, there exists a compact uniformly tight set $M_\varepsilon \subset M$ such that $|\Psi|(M\backslash M_\varepsilon) \le \varepsilon$.*

PROOF. By hypothesis there exists a compact set $K \subset M$ such that $|\Psi|(M\backslash K) < \varepsilon/2$. By Corollary 8.9.9 the measure

$$\eta(B) = \int_K |\mu|(B)\,|\Psi|(d\mu), \quad B \in \mathcal{B}(X),$$

is defined and τ-additive. By hypothesis this measure is Radon. Hence there exist compact sets $C_n \subset X$ such that $\eta(X\backslash C_n) \le \varepsilon 8^{-n}$. Let

$$K_n := \{\mu \in K\colon\ |\mu|(X\backslash C_n) \le 2^{-n}\}.$$

The sets K_n are closed in the weak topology according to Proposition 8.9.8. Then the set $M_\varepsilon := \bigcap_{n=1}^\infty K_n$ is compact in the weak topology and uniformly tight. By the Chebyshev inequality we have

$$|\Psi|(K\backslash K_n) \le 2^n \int_K |\mu|(X\backslash C_n)\,|\Psi|(d\mu) = 2^n \eta(X\backslash C_n) \le \varepsilon 4^{-n}.$$

Hence $|\Psi|(K\backslash M_\varepsilon) \le \sum_{n=1}^\infty |\Psi|(K\backslash K_n) \le \varepsilon/2$, so $|\Psi|(M\backslash K) < \varepsilon$. □

This theorem is valid, for example, for completely regular Souslin spaces. We observe that in this case not every weakly compact set in $\mathcal{M}_t(X)$ is uniformly tight.

8.10.7. Remark. Pachl [**1416**] studied the duality between the space $\mathcal{M}_t(X)$ and the space $C_{bu}(X)$ of bounded uniformly continuous real functions on X in the case where X is a complete metric space. He proved that $\bigl(\mathcal{M}_t, \sigma(\mathcal{M}_t, C_{bu})\bigr)$ is sequentially complete and that a norm bounded subset $\mathcal{M}_t$ is relatively $\sigma(\mathcal{M}_t, C_{bu})$-compact (or countably compact) if and only if its restriction to the class $\mathrm{Lip}_1(X)$ of all functions on X with Lipschitz constant 1, where $\mathrm{Lip}_1(X)$ is equipped with the topology of pointwise convergence, is pointwise equicontinuous. As a corollary one obtains generalizations to uniform measures on uniform spaces.

8.10(ii). Prohorov spaces

8.10.8. Definition. (i) *A completely regular topological space X is called a Prohorov space if every set in the space of measures $\mathcal{M}_t^+(X)$ that is compact in the weak topology is uniformly tight.*

(ii) *A completely regular topological space X is called sequentially Prohorov if every sequence of nonnegative tight Baire measures weakly convergent to a tight measure is uniformly tight.*

We could speak of Radon measures in this definition because every tight Baire measure on X admits a unique Radon extension.

The Prohorov and Le Cam theorems proved above can be reformulated as follows.

8.10.9. Theorem. *Every complete separable metric space is a Prohorov space. An arbitrary metric space is sequentially Prohorov.*

It is clear that a Prohorov space is sequentially Prohorov. We shall see below that the space $\mathbb{Q}$ of rational numbers is sequentially Prohorov, but not Prohorov. We observe that the sequential Prohorov property is weaker than the requirement that weakly convergent sequences of tight Baire measures be uniformly tight (because their limits may not be tight).

If in the definition of a Prohorov space one allows signed measures, then we shall say that X is *strongly Prohorov* (respectively, *strongly sequentially Prohorov*). Theorem 8.6.8 says that all complete metric spaces are strongly Prohorov. Some remarks on various related options are made in the bibliographic comments.

8.10.10. Theorem. *The class of Prohorov spaces is stable under formation of countable products and countable intersections, and passing to closed subspaces and open subspaces, hence to G_δ-subsets.*

In addition, a space is Prohorov provided that every point has a neighborhood that is a Prohorov space (for example, if the space admits a locally finite cover by closed Prohorov subspaces).

The proof can be found in Hoffmann-Jørgensen [**843**] (see also Exercise 8.10.91).

We recall that a space X is called hemicompact if it has a fundamental sequence of compact sets K_n (i.e., every compact set in X is contained in one of the sets K_n). If the continuity of a function on X is ensured by its continuity on all compact sets, then X is called a k_R-space. The latter property is fulfilled for every k-space, i.e., a space in which the closed sets are exactly the sets having closed intersections with all compact sets.

8.10.11. Corollary. *Every Čech complete space X is Prohorov. Hence all locally compact spaces and all hemicompact k_R-spaces are Prohorov.*

We have seen in Example 8.6.9 that the union of two Prohorov subspaces, one of which is a point, may not be Prohorov. The same example shows that a countable union of closed Prohorov subspaces is not always Prohorov.

Let us give several results and examples that enable one to construct broader classes of Prohorov and sequentially Prohorov spaces by means of the operations mentioned in Theorem 8.10.10.

8.10.12. Proposition. *Let X be a completely regular space possessing a countable collection of closed subspaces X_n with the following property: a function on X is continuous if and only if its restriction to every X_n is continuous.*

(i) *Suppose that every X_n is Prohorov. Then so is X.*

(ii) *Suppose that all the spaces X_n are either complete metrizable or compact. Then every weakly fundamental sequence in $\mathcal{M}_r(X)$ is uniformly tight. In particular, X is a strongly sequentially Prohorov space.*

Proof. We may assume that $X_n \subset X_{n+1}$, considering a new system $X_n' = \bigcup_{i=1}^n X_i$. Let $Y = \bigcup_{n=1}^\infty X_n$. It follows from our hypothesis that an arbitrary extension of a continuous function on Y to all of X is continuous on X. Hence $X\backslash Y$ is a functionally closed discrete subspace and its compact subsets are finite. Moreover, every subset of $X\backslash Y$ is Baire in X. Hence, for every weakly compact set M in $\mathcal{M}_r(X)$, the restrictions of measures from M to Y and $X\backslash Y$ form weakly compact families in $\mathcal{M}_r(Y)$ and $\mathcal{M}_r(X\backslash Y)$, respectively. This reduces everything to the case where $X = Y$, which we further assume.

(i) Let $M \subset \mathcal{M}_r^+(X)$ be weakly compact. Let us show that for every $\varepsilon > 0$, there exists a number $n = n(\varepsilon)$ such that $\mu(X\backslash X_n) \le \varepsilon$ for all $\mu \in M$. Indeed, otherwise for every n, there exists a measure $\mu_{i_n} \in M$ such that $\mu_{i_n}(X\backslash X_n) > \varepsilon$. Passing to subsequences, we may assume that there are two increasing sequences of indices i_n and j_n with

$$\mu_{i_n}(X_{j_{n+1}}\backslash X_{j_n}) > \varepsilon, \quad \mu_{i_n}(X\backslash X_{j_{n+1}}) < \varepsilon/2.$$

The sequence $\{\mu_{i_n}\}$ has a limit point $\mu \in M$. Let us pick a number m with $\mu(X\backslash X_m) < \varepsilon/2$. For every n, there exists a compact set $K_n \subset X_{j_{n+1}}\backslash X_{j_n}$ with $\mu_{i_n}(K_n) \ge \varepsilon$. We may assume that $j_1 > m$. There is a continuous function $f_n\colon X \to [0,1]$ such that $f_n|_{K_n} = 1$ and $f_n = 0$ on X_{j_n}. Let us set $f(x) = \sup_n f_n(x)$. Then $0 \le f \le 1$ and f is continuous because the restriction of f to every X_k coincides with the maximum of finitely many functions f_n, hence is continuous. Then

$$\int_X f\, d\mu < \varepsilon/2,$$

whereas

$$\int_X f\, d\mu_{i_n} \ge \varepsilon.$$

This contradiction shows that there exists $n = n(\varepsilon)$ with $\mu(X\backslash X_n) < \varepsilon$ for all $\mu \in M$. According to Corollary 8.10.2, the family M restricted to X_n is relatively weakly compact. Hence by the Prohorov property for X_n it is uniformly tight.

(ii) Let $\{\mu_n\} \subset \mathcal{M}_r(X)$ be a weakly fundamental sequence. Then it converges weakly to a Baire measure μ. Every measure μ_n is purely atomic on $X\backslash Y$. Let $A = \{a_n\}$ be the set of all their atoms in $X\backslash Y$. We observe that $|\mu|\big(X\backslash(Y \cup A)\big) = 0$. Indeed, otherwise there is a set $B \subset X\backslash(Y \cup A)$ on which μ is either strictly positive or strictly negative. The function I_B is continuous on X, its integrals against all the measures μ_n vanish, but the integral against μ is not zero, which leads to a contradiction. The same reasoning shows that the measures μ_n converge to μ on every set in A. Thus, we may assume that $X = Y$. A reasoning similar to the one employed in the proof of Theorem 8.6.2 shows that, for every $\varepsilon > 0$, there is a number $n = n(\varepsilon)$ such that $|\mu_i|(X\backslash X_n) \le \varepsilon$ for all i. Indeed, otherwise one can find increasing sequences of indices i_n and j_n such that

$$|\mu_{i_n}|(X_{j_{n+1}}\backslash X_{j_n}) > \varepsilon.$$

For every n, there is a compact set $K_n \subset X_{j_{n+1}}\backslash X_{j_n}$ with $|\mu_{i_n}|(K_n) > \varepsilon$. There exists a continuous function ξ on X_{j_2} with values in $[1, 1/2]$ that equals 1 on K_1. This function can be extended to a continuous function on X_{j_3} that takes values in $[1, 1/2]$ and equals $1/2$ on K_2. Consequently extending ξ from X_{j_n} to $X_{j_{n+1}}$ in such a way that the extension is continuous, takes values in $[1, 1/n]$ and equals $1/n$ on K_n, we obtain a function on all of X with values in $[0,1]$. By hypothesis, this function is continuous. It is clear that the sets $U_n = \{1/n - \delta_n < \xi < 1/n + \delta_n\}$, where $\delta_n = (2n+1)^{-2}$, are open and disjoint. In addition, every point $x \in X$ possesses a neighborhood that meets at most finitely many sets U_n. Hence for any choice of continuous functions φ_n with support in U_n the series $\sum_{n=1}^{\infty} \varphi_n$ converges and defines a continuous function. For every n, we take a continuous function f_n with values in $[-1,1]$ and support in U_n such that the integral of f_n against the measure μ_n be greater than ε. Let us denote the integral of f_i against the measure μ_n by a_n^i. Then $a_n = (a_n^1, a_n^2, \ldots) \in l^1$, since $\sum_{i=1}^{\infty} |f_i| \leq 1$. For every $\lambda = (\lambda_i) \in l^\infty$ the function $f^\lambda = \sum_{i=1}^{\infty} \lambda_i f_i$ is bounded and continuous. By hypothesis, the sequence of integrals of f_λ with respect to the measures μ_n converges. This means that the sequence $\{a_n\}$ is fundamental in the weak topology of l^1. By Corollary 4.5.8 the sequence $\{a_n\}$ converges in the norm of l^1, whence we obtain $\lim\limits_{n\to\infty} a_n^n = 0$, a contradiction. In the case where all the spaces X_n are compact, the proof is complete. In the case where every X_n is a Polish space, it suffices to verify the uniform tightness of the restrictions of the measures μ_n to every space X_k. Suppose these restrictions are not uniformly tight. The reasoning from the proof of Prohorov's theorem shows that for some $\varepsilon > 0$ there exist a subsequence of measures μ_{i_n} and a sequence of pairwise disjoint compact sets $K_n \subset X_k$ with the following properties: $|\mu_{i_n}|(K_n) > \varepsilon$ and the ε-neighborhoods of K_n (with respect to a complete metric defining the topology of X_k) are disjoint. Let us take a continuous function ξ on X_k with values in $[0,1]$ that equals $1/n$ on K_n for every n. Now the same reasoning as above leads to a contradiction. □

8.10.13. Example. In either of the following cases every weakly fundamental sequence of tight measures on X is uniformly tight:

(i) X is a hemicompact k_R-space.

(ii) X is a locally convex space that is the inductive limit of an increasing sequence of separable Banach spaces E_n such that the embedding of every E_n into E_{n+1} is a compact operator.

PROOF. Claim (i) follows from Proposition 8.10.12. (ii) We observe that X is a k-space possessing a fundamental sequence of compact sets. To this end, one can take an increasing sequence of closed balls U_n in the spaces X_n with $\bigcup_{n=1}^{\infty} U_n = X$ and denote by K_n the compact closure of U_n in X_{n+1}. Suppose a set $A \subset X$ has closed intersections with all K_n. It is readily seen that the sets $A \cap E_n$ are closed in E_n. Suppose A has a limit point $a \notin A$. By induction we construct an increasing sequence of convex sets $V_n \subset E_n$ that are open in E_n such that $a \in V_n$ and $\overline{V_n} \cap A = \varnothing$. To this end, we observe that if a convex compact set K in a Banach space does not meet a closed set M, then K has a convex neighborhood whose closure does not meet M. By definition the set $V = \bigcup_{n=1}^{\infty} V_n$ is open in X. As $a \in V$ and $A \cap V = \varnothing$, we arrive at a contradiction. □

8.10.14. Example. Let X be a locally convex space that is the strict inductive limit of an increasing sequence of its closed subspaces X_n. Then X is a Prohorov space if all spaces X_n are Prohorov. In particular, if each X_n is a separable Fréchet space, then every weakly fundamental sequence of nonnegative Baire measures on X is uniformly tight.

PROOF. According to Example 8.6.10, given a weakly compact family M of nonnegative Radon measures on X, for every $\varepsilon > 0$, all measures in M are concentrated up to ε on some subspace X_n. By Corollary 8.10.2, the restrictions of measures from M to X_n form a relatively weakly compact family. In order to prove the last assertion, it suffices to recall that the union of a sequence of separable Fréchet spaces is Souslin, hence every Baire measure on such a space is Radon. □

Obviously, one can multiply the number of such examples by taking countable products and passing to closed subsets. We observe that many classical spaces of functional analysis such as $\mathcal{D}(\mathbb{R}^d)$, $\mathcal{D}'(\mathbb{R}^d)$, $\mathcal{S}(\mathbb{R}^d)$, and $\mathcal{S}'(\mathbb{R}^d)$ are Prohorov spaces, since they can be obtained by means of the indicated operations.

8.10.15. Remark. The space $\mathcal{D}(\mathbb{R}^1)$ is a Prohorov space, but is neither a k_R-space (Exercise 6.10.27) nor a semicompact space (in addition, it is not σ-compact). The absence of a countable family of compact sets that would be either fundamental or exhausting follows by Baire's theorem applied to the subspaces $\mathcal{D}_n(\mathbb{R}^1)$ and the fact that every compact set in $\mathcal{D}(\mathbb{R}^1)$ is contained in one of the subspaces $\mathcal{D}_n(\mathbb{R}^1)$.

The following result is proved in Wójcicka [**1995**].

8.10.16. Theorem. *Let X be a Prohorov space. Then $\mathcal{P}_r(X)$ with the weak topology is Prohorov.*

PROOF. Let $S = \beta X$ be the Stone–Čech compactification of X. Then $\mathcal{P}_r(S)$ is a compact set in the weak topology, hence the mapping

$$T\colon \mathcal{P}_r\bigl(\mathcal{P}_r(S)\bigr) \to \mathcal{P}_r(S), \quad T(\Psi)(B) = \int_{\mathcal{P}_r(S)} q(B)\,\Psi(dq),$$

considered in Theorem 8.10.5 is well-defined (it is clear that $T(\Psi)$ is the barycenter of Ψ). The spaces $\mathcal{P}_r(X)$ and $\mathcal{P}_r\bigl(\mathcal{P}_r(X)\bigr)$ are naturally embedded into the spaces $\mathcal{P}_r(S)$ and $\mathcal{P}_r\bigl(\mathcal{P}_r(S)\bigr)$, respectively. If $\Psi \in \mathcal{P}_r\bigl(\mathcal{P}_r(X)\bigr)$, then $T(\Psi) \in \mathcal{P}_r(X)$. Indeed, for every $\varepsilon > 0$, there is a compact set $Q \subset \mathcal{P}_r(X)$ with $\Psi(Q) > 1 - \varepsilon$. By hypothesis, there is a compact set $K \subset X$ such that $q(K) > 1-\varepsilon$ for all $q \in Q$. This yields $T(\Psi)(K) \geq (1-\varepsilon)^2$, i.e., $T(\Psi) \in \mathcal{P}_r(X)$. Suppose M is compact in $\mathcal{P}_r\bigl(\mathcal{P}_r(X)\bigr)$ and $\varepsilon > 0$. Then the compact set $T(M)$ is contained in $\mathcal{P}_r(X)$ as shown above, which by hypothesis gives compact sets $K_n \subset X$ with $K_n \subset K_{n+1}$ and $T(\Psi)(K_n) \geq 1 - \varepsilon^2 4^{-n}$ for all $\Psi \in M$. It is readily seen that the sets $Q_n := \{q \in \mathcal{P}_r(S)\colon\ q(K_n) \geq 1 - \varepsilon 2^{-n}\}$ are compact and $Q := \bigcap_{n=1}^{\infty} Q_n \subset \mathcal{P}_r(X)$. For every $\Psi \in M$ one has $\Psi(Q_n) \geq 1 - \varepsilon 2^{-n}$, as $T(\Psi)(X\backslash K_n) \le \varepsilon^2 4^{-n}$, whence $\Psi\bigl(q\colon\ q(X\backslash K_n) \ge \varepsilon 2^{-n}\bigr) \le \varepsilon 2^{-n}$. Finally, we obtain $\Psi(Q) \ge 1-\varepsilon$, and Q is compact in $\mathcal{P}_r(X)$. □

No topological description of Prohorov spaces is known. The following two examples show that the class of Prohorov spaces is not closed with respect to formation of countable unions. The first of them has already been described in Example 8.6.9. The countable space X constructed there is hemicompact, is Baire and is an F_σ-set in the Prohorov space βX, but is not Prohorov itself.

The second example is due to Preiss [**1486**]. This deep and difficult theorem is a fundamental achievement of the topological measure theory.

8.10.17. Theorem. *The space of rational numbers $\mathbb{Q}$ with its standard topology is not a Prohorov space.*

We recall that by Theorem 8.10.9, $\mathbb{Q}$ is a sequentially Prohorov space.

The first examples of separable metric spaces that are not Prohorov spaces were constructed in Choquet [**353**] and Davies [**413**]. A simplified (but still highly non-trivial) proof of Theorem 8.10.17 is given in Topsøe [**1874**].

Fernique [**563**] observed that the space l^2 with its weak topology is not Prohorov.

8.10.18. Example. A sequence of measures $\mu_n = n^{-3}\sum_{i=1}^{n^3} \delta_{ne_i}$, where $\{e_i\}$ is the standard orthonormal basis in l^2, converges weakly to Dirac's measure at zero if l^2 is equipped with the weak topology, but obviously is not tight. For the verification of weak convergence, it suffices to observe that for every set of the form $S = \{x\colon\ |(x,v)| < 1\}$, $v \in l^2$, one has $\mu_n(S) \to 1$, which is obvious from the estimates

$$\mu_n(l^2\backslash S) \le \int_{l^2} |(x,v)|^2\, \mu_n(dx) \le n^{-3}\sum_{i=1}^{n^3} n^2 v_i^2 \le n^{-1}(v,v).$$

The space l^2 with the weak topology provides an example of a hemicompact σ-compact space that is not Prohorov. In Fremlin, Garling, Haydon [**636**], this example was generalized as follows.

8.10.19. Proposition. *Let X be an infinite-dimensional Banach space. Then, the spaces $\bigl(X, \sigma(X,X^*)\bigr)$ and $\bigl(X^*, \sigma(X^*,X)\bigr)$ are not Prohorov spaces.*

According to Fernique [**563**], the strong dual to a locally convex Fréchet–Montel space X is Prohorov. In particular, the dual to $X = \mathbb{R}^\infty$ is $\mathbb{R}^\infty_0$, which is a countable union of finite-dimensional subspaces (here $\mathbb{R}^\infty_0$ is equipped with the topology of inductive limit). Thus, a nonmetrizable Prohorov space may not be a Baire space.

Another result from Fremlin, Garling, Haydon [**636**] improves on assertion (i) in Example 8.10.13 with a close proof.

8.10.20. Theorem. *Let X be a hemicompact k_R-space. Then every weakly compact subset of $\mathcal{M}_t(X)$ is uniformly tight, i.e., X is strongly Prohorov.*

Proof. There are compact sets $X_n \subset X_{n+1}$ such that every compact set in X is contained in one of the sets X_n, and the continuity of a function on every X_n yields its continuity on all of X. Suppose a weakly compact set $M \subset \mathcal{M}_t(X)$ is not uniformly tight. As in the proof of assertion (ii) of Proposition 8.10.12, one can find measures $\mu_n \in M$ and increasing numbers j_n with $|\mu_n|(X_{j_{n+1}} \backslash X_{j_n}) > \varepsilon$. Let us take the functions f_i constructed in that proof such that $\sum_{i=1}^{\infty} |\lambda_i||f_i| \leq \|\lambda\|$ if $\lambda = (\lambda_i) \in l^\infty$, and the integral of f_n against μ_n is greater than ε. The mapping from M to l^1 that takes the measure μ to the sequence of the integrals of f_i against μ is continuous with respect to the weak topologies on M and l^1. The image of M under this mapping is weakly compact in l^1, which yields its norm compactness. This contradicts the fact that the integral of f_n against μ_n is greater than ε. $\square$

Note that any σ-compact locally compact space is a hemicompact k_R-space. Let us mention the following important result due to Preiss [**1486**].

8.10.21. Theorem. (i) *A first category metric space cannot be Prohorov (unlike the above-mentioned space* $\mathbb{R}_0^\infty$ *with the topology of inductive limit).*

(ii) *Let X be a separable coanalytic metric space. Then X is a Prohorov space if and only if X is metrizable by a complete metric. An equivalent condition: the space X contains no countable G_δ-set dense in itself.*

(iii) *Under the continuum hypothesis, there exists a separable metric Prohorov space that does not admit a complete metric.*

Since every countable space dense in itself is homeomorphic to $\mathbb{Q}$, assertion (ii) explains the role of $\mathbb{Q}$ in Theorem 8.10.17.

Under some additional set-theoretic assumptions, there exists a Souslin Prohorov subset of $[0,1]$ that is not Polish (see Cox [**379**], Gardner [**660**]). It is an open question whether it is consistent with ZFC that every universally measurable Prohorov space $X \subset [0,1]$ is topologically complete (i.e., is Polish).

Bouziad [**246**] and Choban [**342**] constructed examples showing that the image Y of a Prohorov space X under a continuous open mapping may not be Prohorov (such a space X may be even countable and a mapping may be compact). This answers a question raised in Topsøe [**1874**], where the following result was proved (see [**1874**, Corollary 6.2].

8.10.22. Proposition. *Let $\pi\colon X \to Y$ be a perfect surjection. Then X is a Prohorov space if and only if so is Y.*

It is interesting to compare the Prohorov and Skorohod properties (defined in §8.5). It was shown in Bogachev, Kolesnikov [**211**] that the space $\mathbb{R}_0^\infty$ of all finite sequences (with its natural topology of the inductive limit of an increasing sequence of finite-dimensional spaces) does not have the Skorohod property, although is Souslin and Prohorov. On the other hand, in Banakh, Bogachev, Kolesnikov [**114**], the class of almost metrizable spaces was considered (i.e., spaces X for which there exists a bijective continuous proper mapping from a metric space onto X) and it was shown that an almost metrizable space is sequentially Prohorov precisely when it has the strong Skorohod property for Radon measures.

8.10(iii). The weak sequential completeness of spaces of measures

Several remarks on the weak sequential completeness of the space $\mathcal{M}_t(X)$ are in order. First of all, two obvious observations.

8.10.23. Example. Let X be a completely regular space. The space of measures $\mathcal{M}_t(X)$ is weakly sequentially complete provided that either $\mathcal{M}_\sigma(X) = \mathcal{M}_t(X)$ or every weakly fundamental sequence in $\mathcal{M}_t(X)$ is uniformly tight.

PROOF. It suffices to use the weak sequential completeness of $\mathcal{M}_\sigma(X)$ and Theorem 8.6.7. □

8.10.24. Example. For every σ-compact completely regular space X, the space $\mathcal{M}_t(X)$ is weakly sequentially complete.

PROOF. The claim follows by the weak sequential completeness of the space $\mathcal{M}_\sigma(X)$, since every Baire measure on X is tight. □

Proposition 8.10.12 (ii) and Example 8.10.23 give one more example.

8.10.25. Example. Let X be a completely regular space possessing a sequence of compact subspaces K_n such that any function on X continuous on every K_n is continuous on all of X. Then the space $\mathcal{M}_t(X)$ is weakly sequentially complete.

The following result is obtained in Moran [**1331**].

8.10.26. Theorem. *Let X be a normal and metacompact space (i.e., in every open cover of X one can inscribe a pointly finite open cover). Then the space $\mathcal{M}_\tau(X)$ is weakly sequentially complete. The same is true for $\mathcal{M}_t(X)$ if, additionally, X is Čech complete.*

8.10(iv). The A-topology

There is another natural way to topologize the space of probability measures inspired by the Alexandroff theorem, which is used if X is not completely regular or if the class of Borel measures does not coincide with the class of Baire measures. Let $\mathcal{G}$ be the class of all open sets in X.

The A-topology on the space $\mathcal{P}(X)$ of all Borel probability measures (or its subspaces $\mathcal{P}_r(X)$ and $\mathcal{P}_\tau(X)$) is defined by means of neighborhoods of the form

$$U(\mu, G, \varepsilon) = \bigl\{\nu\colon\ \mu(G) < \nu(G) + \varepsilon\bigr\},$$

where $\mu \in \mathcal{P}(X)$, $G \in \mathcal{G}$, $\varepsilon > 0$. A net $\{\mu_\alpha\}$ converges in this topology to μ if and only if $\liminf_\alpha \mu_\alpha(G) \geq \mu(G)$ for every $G \in \mathcal{G}$. By Lemma 7.1.2 the

A-topology is Hausdorff. It follows from §8.2 that in the case of a completely regular space the A-topology coincides with the weak topology on $\mathcal{P}_\tau(X)$. Certainly, in the general case the A-topology is stronger than the weak topology (which may be trivial if there are no non-trivial continuous functions on X). Another possible advantage of the A-topology is that it is applicable to Borel measures, whereas the weak topology is naturally connected with Baire measures (it may not be Hausdorff on Borel measures). In order to define the A-topology on the space $\mathcal{M}^+(X)$ of all nonnegative Borel measures, in addition to the indicated neighborhoods one adds the neighborhoods $U'(\mu,\varepsilon) = \bigl\{\nu\colon\ |\mu(X)-\nu(X)|<\varepsilon\bigr\}$. Many results proved above for the weak topology have natural analogs for the A-topology (see, for example, Topsøe [**1873**] and Exercise 8.10.123). In particular, X is homeomorphic to the set of all Dirac measures with the A-topology, which is closed in the spaces $\mathcal{P}_r(X)$ and $\mathcal{P}_\tau(X)$ with the A-topology. In addition, the following holds.

8.10.27. Theorem. *The space $\mathcal{M}^+_\tau(X)$ with the A-topology is regular, completely regular or second countable if and only if X has the corresponding property.*

Note the following result from Holický, Kalenda [**851**].

8.10.28. Theorem. (i) *Let Y be a Hausdorff space and let $X\subset Y$. Suppose that X is a set of one of the following types: G_δ, Borel, $\mathcal{F}$-Souslin, $\mathcal{B}$-Souslin (i.e., is obtained from Borel sets by the A-operation). Then $\mathcal{M}^+(X)$ and $\mathcal{M}^+_r(X)$ are sets of the corresponding type in $\mathcal{M}^+(Y)$ and $\mathcal{M}^+_r(Y)$ with the A-topology.*

(ii) *If X is Čech complete, then so is $\mathcal{M}^+_r(X)$ with the A-topology.*

Certainly, for completely regular spaces the assertions for $\mathcal{M}^+_r(X)$ hold for the weak topology.

8.10(v). Continuous mappings of spaces of measures

A continuous mapping $f\colon X\to Y$ generates the mapping

$$\widehat{f}\colon \mathcal{M}_r(X)\to\mathcal{M}_r(Y), \quad \mu\mapsto\mu\circ f^{-1},$$

which is continuous in the weak topology. One also obtains the mappings

$$\widehat{f}\colon \mathcal{M}_t(X)\to\mathcal{M}_t(Y), \quad \widehat{f}\colon \mathcal{M}_\tau(X)\to\mathcal{M}_\tau(Y), \quad \widehat{f}\colon \mathcal{M}_\sigma(X)\to\mathcal{M}_\sigma(Y),$$

and the mappings between the corresponding spaces of nonnegative or probability measures. It is readily verified that if f is injective, then so is the mapping $\widehat{f}\colon \mathcal{M}_r(X)\to\mathcal{M}_r(Y)$ (see a more general assertion in Exercise 9.12.39). Certainly, this is also true for the classes $\mathcal{M}_t$, but not always for $\mathcal{M}_\tau$.

8.10.29. Example. Let $S\subset[0,1]$ be a set with $\lambda^*(S)=1$, $\lambda_*(S)=0$, where λ is Lebesgue measure (see Example 1.12.13). Let us consider the natural projection $f\colon \bigl(S\times\{0\}\bigr)\cup\bigl(([0,1]\backslash S)\times\{1\}\bigr)\to[0,1]$. Then f is continuous and injective, but Lebesgue measure on $[0,1]$ is the image of two different

τ-additive probability measures μ_1 and μ_2 that are induced by Lebesgue measure on $S\times\{0\}$ and $([0,1]\backslash S)\times\{1\}$, respectively.

See also Remark 8.9.1 made above. Perfect mappings between spaces induce perfect mappings between spaces of measures (Koumoullis [**1044**]):

8.10.30. Theorem. *Let $f\colon X\to Y$ be a continuous surjection of completely regular spaces. Then, for $s=t$ and $s=\tau$, the induced mapping $\widehat{f}\colon \mathcal{M}_s^+(X)\to\mathcal{M}_s^+(Y)$, $\mu\mapsto\mu\circ f^{-1}$, is perfect if and only if f is perfect.*

As observed in [**1044**], this result may fail for $s=\sigma$ and for spaces of signed measures.

This theorem and Frolík's result that the space X is Lindelöf and Čech complete precisely when it admits a perfect surjection onto a complete separable metric space, was employed in [**1044**] to obtain the following result.

8.10.31. Corollary. *Let X be completely regular. The space $\mathcal{M}_s^+(X)$, where $s=t$ or $s=\tau$, is Lindelöf and Čech complete if and only if so is X. In addition, $\mathcal{M}_s^+(X)$ is paracompact and Čech complete precisely when so is X.*

In Ditor, Eifler [**457**], Eifler [**524**], Schief [**1669**], [**1670**], Banakh [**113**], Banakh, Radul [**120**], and Bogachev, Kolesnikov [**211**], open mappings between spaces of measures are studied. Let us mention a result from [**1670**]. Set $\mathcal{M}^+(X):=\mathcal{M}_{\mathcal{B}}^+(X)$, $\mathcal{P}(X):=\mathcal{P}_{\mathcal{B}}(X)$.

8.10.32. Theorem. *Let X and Y be Hausdorff spaces and let $f\colon X\to Y$ be a Borel surjection that is open, i.e., takes open sets to open sets. Suppose that for every open set $G\subset X$ we have $\widehat{f}\bigl(\mathcal{M}^+(G)\bigr)=\mathcal{M}^+\bigl(f(G)\bigr)$. Then the mapping $\widehat{f}\colon\mathcal{M}^+(X)\to\mathcal{M}^+(Y)$ is open in the A-topology.*

8.10.33. Corollary. *Let X and Y be Souslin spaces and let $f\colon X\to Y$ be a Borel surjection. If f is an open mapping, then the induced mappings $\widehat{f}\colon\mathcal{M}^+(X)\to\mathcal{M}^+(Y)$ and $\widehat{f}\colon\mathcal{P}(X)\to\mathcal{P}(Y)$ are open in the A-topology.*

A close result is obtained in Bogachev, Kolesnikov [**211**] for the spaces of Radon probability measures: the mapping $\widehat{f}\colon\mathcal{P}_r(X)\to\mathcal{P}_r(Y)$ is a continuous open surjection if f is a continuous open surjection of completely regular spaces X and Y such that $\widehat{f}\bigl(\mathcal{P}_r(G)\bigr)=\mathcal{P}_r\bigl(f(G)\bigr)$ for every open set $G\subset X$. In particular, the next result is proved in [**211**].

8.10.34. Proposition. *Let $f\colon X\to Y$ be an open continuous surjective mapping between complete metric spaces. Then, the mapping $\mathcal{P}_r(X)\to\mathcal{P}_r(Y)$ is an open surjection.*

Interesting connections between the Skorohod representation, open mappings, and selection theorems are discussed in Bogachev, Kolesnikov [**211**]. We formulate some results of this work. We recall the following classical result, called Michael's selection theorem (see Michael [**1314**] or Repovš, Semenov [**1552**, p. 190]). Let M be a metrizable space, let P be a complete metrizable closed subset of locally convex space E, and let $\Phi\colon M \to 2^P$ be a lower semicontinuous mapping with values in the set of nonempty convex closed subsets of P. Then, there exists a continuous mapping $f\colon M \to P$ such that $f(x) \in \Phi(x)$ for all x. For our purposes, it will be enough to deal with the case where E is a normed space; a short proof for this case can be found in Repovš, Semenov [**1552**, A§1] (note that Filippov [**585**] constructed an example showing that one cannot omit the requirement that P is closed even if P is a G_δ-set in a Hilbert space). Namely, we shall deal with the situation where P and M are the sets of all Radon probability measures on Polish spaces X and Y; the weak topology on these sets is generated by the Kantorovich–Rubinshtein norm on $\mathcal{M}_r(X)$ and $\mathcal{M}_r(Y)$. A typical application of this theorem is this: let $T\colon P \to M$ be a continuous open affine mapping of a complete metrizable convex closed set P in a locally convex space to a metrizable set M in a locally convex space. Then $\Phi(x) = T^{-1}(x)$ satisfies the hypotheses of Michael's theorem. Hence T has a continuous right inverse, and Theorem 8.10.32 yields the following assertion.

8.10.35. Theorem. *Let $f\colon X \to Y$ be a continuous open surjection of Polish spaces. Then the induced mapping $\widehat{f}\colon \mathcal{P}(X) \to \mathcal{P}(Y)$ has a right inverse continuous in the weak topology. In the case of arbitrary complete metric spaces, the same is true for $\mathcal{P}_r(X)$ and $\mathcal{P}_r(Y)$.*

8.10.36. Corollary. *For every universally measurable set Y in a Polish space Z, there exist a universally measurable subset X of the space $\mathcal{R}$ of irrational numbers in $[0,1]$ and a continuous surjection $f\colon X \to Y$ such that the mapping $\widehat{f}\colon \mathcal{P}(X) \to \mathcal{P}(Y)$ has a right inverse $g\colon \mathcal{P}(Y) \to \mathcal{P}(X)$ continuous in the weak topology. For an arbitrary set $Y \subset Z$, the analogous assertion, but without universal measurability of X, is true for the spaces $\mathcal{P}_r(X)$ and $\mathcal{P}_r(Y)$.*

In the general case, it may occur in the situation of the preceding theorem that there is no linear continuous right inverse operator. However, as shown in Michael [**1313**], if X and Y are metrizable compact spaces, then every continuous open surjection $f\colon X \to Y$ has a regular averaging operator, hence the mapping $\widehat{f}\colon \mathcal{M}_r(X) \to \mathcal{M}_r(Y)$ has a linear continuous in the weak topology right inverse. We remark that according to §8.5, the assumption that f is open is not necessary for the existence of a regular averaging operator and a linear continuous right inverse of $\widehat{f}$. For example, in Lemma 8.5.3, the Cantor set cannot be mapped onto $[0,1]$ by an open mapping. The proof of the next result is given in Bogachev, Kolesnikov [**211**].

8.10.37. Proposition. *For any universally measurable set Y in a Polish space, there exist a universally measurable subset X of the Cantor set C and a continuous surjective mapping $f\colon X \to Y$ such that the associated mapping $\widehat{f}\colon \mathcal{P}(X) \to \mathcal{P}(Y)$ has a linear right inverse $g\colon \mathcal{P}(Y) \to \mathcal{P}(X)$ continuous in the weak topology. In the case of compact Y, the set X can be chosen compact. For an arbitrary set Y, the analogous assertion, but without universal measurability of X, is true for the spaces $\mathcal{P}_r(X)$ and $\mathcal{P}_r(Y)$.*

8.10(vi). Separability of spaces of measures

The separability properties of spaces of measures with the weak topology are investigated in Koumoullis, Sapounakis [**1051**], Pol [**1473**], Talagrand [**1830**]. We recall that if X is separable, then the space $\mathcal{M}_t(X)$ with the weak topology is separable as well (Theorem 8.9.4). The converse is false even for compact spaces (see [**1830**]). As shown in [**1830**] under the continuum hypothesis, there exists a compact space K such that the space $\mathcal{M}_t(K)$ is separable in the weak topology, but the unit ball of $\mathcal{M}_t(K)$ is not. In addition, the separability of the unit ball in $\mathcal{M}_t(K)$ in the weak topology does not imply the metrizability of K: according to [**1830**] (also under CH), it may even occur that there is no separable measure with support K.

A set of measures $M \subset \mathcal{M}_\sigma(X)$ is called *countably separated* if there exists a sequence $\{f_n\} \subset C_b(X)$ such that, whenever μ and ν are in M, the equality

$$\int_X f_n(x)\,\mu(dx) = \int_X f_n(x)\,\nu(dx), \quad \forall n \in \mathbb{N},$$

implies that $\mu = \nu$.

A subset $M \subset \mathcal{M}_\sigma(X)$ is called *countably determined* in $\mathcal{M}_\sigma(X)$ if there exists a sequence $\{f_n\} \subset C_b(X)$ such that, whenever $\mu \in M$ and $\nu \in \mathcal{M}_\sigma(X)$, the equality

$$\int_X f_n(x)\,\mu(dx) = \int_X f_n(x)\,\nu(dx), \quad \forall n \in \mathbb{N},$$

implies the inclusion $\nu \in M$. By analogy one defines the property to be countably determined in $\mathcal{M}_\sigma^+(X)$.

It is easy to see that for a compact space X, the set $\mathcal{M}_\sigma^+(X)$ is countably separated if and only if X is metrizable (see Exercise 8.10.81). The following simple lemma from Koumoullis [**1044**] is useful in such considerations.

8.10.38. Lemma. *Let H be a countable family of bounded Baire functions on a topological space X. Then, there exists a countable set $K \subset C_b(X)$ with the following property: if for a pair of Baire measures μ and ν on X one has the equality*

$$\int_X \varphi(x)\,\mu(dx) = \int_X \varphi(x)\,\nu(dx), \quad \forall\,\varphi \in K,$$

then this equality is fulfilled for all $h \in H$ in place of φ.

PROOF. It suffices to consider the case where H consists of a single function h. The class $\mathcal{H}$ of all bounded Baire functions h for which our claim is true contains $C_b(X)$, is a linear space and is closed with respect to the pointwise limits of uniformly bounded sequences. By Theorem 2.12.9 the class $\mathcal{H}$ coincides with the class of all bounded Baire functions. $\square$

It is clear from the lemma that in the definitions of countably separated and countably determined sets one can consider bounded Baire functions (or even sequences of Baire sets).

Since a compact space K is metrizable precisely when there is a countable family of continuous functions separating the points in K, it is clear that a compact (in the weak topology) set $M \subset \mathcal{M}_\sigma(X)$ is countably separated if and only if it is metrizable. According to [**1051**, Proposition 2.3], a compact set $M \subset \mathcal{M}_\sigma(X)$ is countably determined if and only if it is a G_δ-set in $\mathcal{M}_\sigma(X)$ (and similarly for sets in $\mathcal{M}_\sigma^+(X)$). It is clear that these assertions may fail for noncompact sets (for example, typically $\mathcal{M}_\sigma(X)$ is not metrizable in the weak topology). The following result (see Koumoullis, Sapounakis [**1051**, Theorem 4.1]) describes the situation for the whole space of measures. We recall that a space Y is called *countably submetrizable* if there exists a sequence of continuous functions separating the points in Y (in other words, a continuous injection $Y \to \mathbb{R}^\infty$).

8.10.39. Theorem. *Let X be a Hausdorff space and let s be one of the symbols σ, τ or t. The following assertions are equivalent*:

(i) $\mathcal{M}_s(X)$ *is countably separated*;

(ii) $\mathcal{M}_s^+(X)$ *is countably separated*;

(iii) $C_b(X)$ *is separable in the topology* $\sigma\bigl(C_b(X), \mathcal{M}_s(X)\bigr)$;

(iv) $\mathcal{M}_s(X)$ *is countably submetrizable*;

(v) *every point in* $\mathcal{M}_s(X)$ *is a* G_δ*-set.*

In addition, for $s = t$ conditions (i)–(v) *are equivalent to the submetrizability of the space X.*

8.10(vii). Young measures

Let $(\Omega, \mathcal{B})$ and $(S, \mathcal{A})$ be two measurable spaces and let μ be a bounded positive measure on $\mathcal{B}$. Denote by $\mathcal{Y}(\Omega, \mu, S)$ the set of all positive measures ν on $\mathcal{B}\otimes\mathcal{A}$ such that the image of ν under the natural projection $\Omega\times S \to \Omega$ coincides with μ. Measures in $\mathcal{Y}(\Omega, \mu, S)$ are called Young measures. A typical example of a Young measure: the measure $\nu := \mu\circ F^{-1}$, where $F\colon\ \Omega \to \Omega\times S$, $F(x) = \bigl(x, u(x)\bigr)$ and $u\colon\ \Omega \to S$ is a measurable mapping. Such a measure ν is called the Young measure generated by the mapping u. Young measures are useful in variational calculus; there exist some connections between convergence of mappings and convergence of the associated Young measures. A simple example of this connection is given in Exercise 8.10.86; additional information can be found in Castaing, Raynaud de Fitte, Valadier [**318**], Giaquinta, Modica, Souček [**683**], Valadier [**1912**], [**1913**]. To Young measures are partially related the next section and §9.12(vii). The proof of the following proposition is given in Valadier [**1912**, Theorem 17].

8.10.40. Proposition. *Let μ be a Radon probability measure on a Hausdorff space Ω, let u_n be measurable mappings from Ω to a separable metric space S, and let ν_n be the corresponding Young measures. Let $\Psi\colon\ \Omega\times S \to \mathbb{R}$ be a $\mathcal{B}\otimes\mathcal{B}(S)$-measurable function such that for every fixed x, the function $y \mapsto \Psi(x, y)$ is continuous, and the sequence of functions $x \mapsto \Psi\bigl(x, u_n(x)\bigr)$ is uniformly μ-integrable. Suppose that the measures ν_n converge weakly to a* *measure ν. Then the function Ψ is ν-integrable and*

$$\int_{\Omega\times S}\Psi\,d\nu=\lim_{n\to\infty}\int_{\Omega}\Psi\bigl(x,u_n(x)\bigr)\,\mu(dx).$$

8.10.41. Lemma. *Let $\mu\ge 0$ be a Radon measure on a topological space Ω and let $U\subset L^1(\mu)$ be a norm bounded set. Then the corresponding set of Young measures ν_u, $u\in U$, is uniformly tight on $\Omega\times\mathbb{R}$. If μ is concentrated on a countable union of metrizable compact sets, then for every $\varepsilon>0$, there is a metrizable compact $K_\varepsilon\subset\Omega\times\mathbb{R}$ such that $\sup_{u\in U}\nu_u\bigl((\Omega\times\mathbb{R})\backslash K_\varepsilon\bigr)\le\varepsilon$.*

PROOF. Let π_u the projection of ν_u to $\mathbb{R}$. We observe that

$$\sup_{u\in U}\int_{\mathbb{R}}|t|\,\pi_u(dt)=\sup_{u\in U}\int_{\Omega\times\mathbb{R}}|t|\,\nu_u(d\omega dt)=\sup_{u\in U}\int_{\Omega}|u|\,d\mu\le\sup_{u\in U}\|u\|_{L^1(\mu)}.$$

Hence the projections of the measures ν_u on $\mathbb{R}$ form a tight family. The projection to Ω is the tight measure μ. Now we can apply Lemma 7.6.6. □

The next result shows that Young measures that are not generated by mappings arise naturally as the limits of sequences of Young measures generated by mappings.

8.10.42. Proposition. *Suppose that a Radon probability measure μ on a completely regular space Ω is concentrated on a countable union of metrizable compact sets. Suppose that a sequence $\{u_n\}$ converges weakly in $L^1(\mu)$ to a function u, but does not converge in the norm. Then the sequence of the associated Young measures ν_n on $\Omega\times\mathbb{R}^1$ has a subsequence that converges weakly to some Young measure ν that cannot be generated by a function.*

PROOF. The sequence $\{u_n\}$ is uniformly integrable. According to our hypothesis, there exist $c>0$ and a subsequence $\{u_{n_k}\}$ with $\|u-u_{n_k}\|_{L^1(\mu)}\ge c$ for all k. By using Lemma 8.10.41 and Theorem 8.6.7, one can find a further subsequence (again denoted by u_{n_k}) such that the corresponding Young measures converge weakly to some measure ν. It is clear that ν is a Young measure. Suppose that ν is generated by some measurable function v. According to Exercise 8.10.86, the sequence $\{u_{n_k}\}$ converges in measure, hence by the Lebesgue–Vitali theorem it converges in the norm. Then its limit in $L^1(\mu)$ must coincide with u, which is a contradiction. □

8.10(viii). Metrics on spaces of measures

In §8.3 we have already discussed the Lévy–Prohorov and Kantorovich–Rubinshtein metrics on the space of probability measures on a given metric space (X,d). Here some additional results on these and related metrics are presented. The definitions of d_P, $\|\cdot\|_0$ and $\|\cdot\|^*_{\mathrm{BL}}$ are given in §8.3.

8.10.43. Theorem. *For every two Borel probability measures μ and ν on a metric space X, the following relationship between the Lévy–Prohorov* *and Kantorovich–Rubinshtein metrics holds*:

$$\frac{2d_P(\mu,\nu)^2}{2+d_P(\mu,\nu)} \le \|\mu-\nu\|^*_{\mathrm{BL}} \le \|\mu-\nu\|_0 \le 3d_P(\mu,\nu). \qquad (8.10.1)$$

In addition, $\|\mu-\nu\|^*_{\mathrm{BL}} \le 2d_P(\mu,\nu)$. *If* X *is complete, then the space* $\mathcal{P}_r(X)$ *with any of the above-mentioned metrics is complete as well.*

PROOF. Let $d_P(\mu,\nu) > r > 0$. It is clear from the definition of d_P that we may assume that there exists a closed set B with $\mu(B) > \nu(B^r)+r$. There exists a Lipschitzian function f with $|f| \le 1$ and $|f(x)-f(y)| \le 2d(x,y)/r$ such that f equals 1 on B and -1 outside B^r, e.g., $f(x) = \theta\bigl(\operatorname{dist}(x,B)\bigr)-1$, $\theta(t) = 2(1-t/r)I_{[0,r]}(t)$. Then

$$\begin{aligned}(1+2/r)\|\mu-\nu\|^*_{\mathrm{BL}} &\ge \int_X f\,d(\mu-\nu)\\ &= \int_X (f+1)\,d(\mu-\nu) \ge 2\mu(B)-2\nu(B^r) \ge 2r.\end{aligned}$$

Therefore, $\|\mu-\nu\|^*_{\mathrm{BL}} \ge 2r^2/(2+r)$. Since $r < d_P(\mu,\nu)$ is arbitrary, we obtain the first inequality in (8.10.1). Now suppose that $f \in \mathrm{Lip}_1(X)$, $|f| \le 1$ and

$$\int_X f\,d(\mu-\nu) > 3r > 0.$$

Set $\Phi_\mu(t) = \mu(f<t)$, $\Phi_\nu(t) = \nu(f<t)$. Then, integrating by parts and taking into account the equalities $\Phi_\mu(1+) = \Phi_\nu(1+) = 1$ (see Exercise 5.8.112) and applying the change of variable formula (3.6.3) we find

$$\int_{-1}^1 \bigl[\Phi_\nu(t)-\Phi_\mu(t)\bigr]\,dt = \int_{-1}^1 t\,d(\Phi_\mu-\Phi_\nu)(t) = \int_X f\,d(\mu-\nu) > 3r. \qquad (8.10.2)$$

Let us show that there exists $\tau \in [-1,1]$ such that

$$\Phi_\nu(\tau) > \Phi_\mu(\tau+r)+r. \qquad (8.10.3)$$

Indeed, otherwise $\Phi_\nu(t) \le \Phi_\mu(t+r)+r$ for all t. The integration yields

$$\int_{-1}^1 \Phi_\nu(t)\,dt \le \int_{-1+r}^{1+r} \Phi_\mu(t)\,dt + 2r.$$

Since $\Phi_\mu(t) = 1$ for all $t > 1$, we obtain by the previous inequality

$$\int_{-1}^1 \Phi_\nu(t)\,dt \le \int_{-1}^1 \Phi_\mu(t)\,dt + 3r$$

contrary to (8.10.2). Set $B := f^{-1}\bigl([-1,\tau)\bigr)$. Then $B^r \subset f^{-1}\bigl([-1,\tau+r)\bigr)$ since $f \in \mathrm{Lip}_1(X)$. Hence (8.10.3) yields $\nu(B) > \mu(B^r)+r$, which gives the estimate $d_P(\mu,\nu) \ge r$. Now the last estimate in (8.10.1) follows by choosing $3r$ sufficiently close to $\|\mu-\nu\|_0$. The estimate $\|\mu-\nu\|^*_{\mathrm{BL}} \le 2d_P(\mu,\nu)$ is proved similarly, taking into account that the equality $\|f\|_{\mathrm{BL}} = 1$ yields that the function f is Lipschitzian with constant $1-\sup_x |f(x)|$. The last assertion of the theorem has been verified in the proof of Theorem 8.9.4 for $d_0(\mu,\nu) = \|\mu-\nu\|_0$, hence it holds for the other metrics mentioned in the formulation. $\square$

Denote by $M_1(X)$ the set of all Borel probability measures on X such that the functions $x \mapsto d(x, x_0)$ are integrable for all $x_0 \in X$ (by the triangle inequality, it suffices to have the integrability for some x_0). On the set $M_1(X)$ we define the following modified Kantorovich–Rubinshtein metric:

$$\|\mu - \nu\|_0^* := \sup\biggl\{\int_X f\,d(\mu-\nu)\colon\ f \in \mathrm{Lip}_1(X)\biggr\}.$$

It is clear that $\|\mu - \nu\|_0 \le \|\mu - \nu\|_0^* \le \max(\operatorname{diam} X, 1)\|\mu - \nu\|_0$ and

$$\|\mu - \nu\|_0^* \le \int d(x,a)\,(\mu+\nu)(dx)$$

for every $a \in X$, since $f(x)$ can be replaced by $f(x) - f(a)$ due to the equality $\mu(X) = \nu(X)$ and the estimate $|f(x) - f(a)| \le d(x, a)$. If $\operatorname{diam} X \le 1$, then $\|\mu - \nu\|_0 = \|\mu - \nu\|_0^*$. Note that $\|\delta_x - \delta_y\|_0^* = d(x, y)$ and $\|\delta_x - \delta_y\|_0 \le 2$. The mapping $x \mapsto \delta_x$ is an isometry from X to the space $M_1(X) \cap \mathcal{P}_\tau(X)$ with metric $\|\cdot\|_0^*$ and its image is closed. The space $\bigl(M_1(X) \cap \mathcal{P}_\tau(X), \|\cdot\|_0^*\bigr)$ is complete precisely when so is (X, d) (the proof is similar to the case of d_0).

The quantity $\|\mu - \nu\|_0^*$ is indeed a norm of the measure $\mu - \nu$ if we consider the linear space $M_0(X)$ of all signed Borel measures σ on X such that $\sigma(X) = 0$ and the function $x \mapsto d(x, x_0)$ is integrable with respect to $|\sigma|$ (equivalently, $\mathrm{Lip}_1(X) \in L^1(|\sigma|)$). The above formula defines the norm $\|\sigma\|_0^*$ on $M_0(X)$. We observe that σ can be written as $\|\sigma^+\|\mu - \|\sigma^-\|\nu$, where $\mu, \nu \in M_1(X)$, $\mu = \sigma^+/\|\sigma^+\|$, $\nu = \sigma^-/\|\sigma^-\|$. The Kantorovich–Rubinshtein norm $\|\cdot\|_0^*$ can be extended to the linear space of all bounded Borel measures on X that integrate all Lipschitzian functions. To this end, we set

$$\|\sigma\|_0^* = |\sigma(X)| + \sup\biggl\{\int_X f\,d\sigma\colon\ f \in \mathrm{Lip}_1(X), f(x_0) = 0\biggr\}.$$

In nontrivial cases $\bigl(M_0(X), \|\cdot\|_0^*\bigr)$ is not complete (see p. 192).

8.10.44. Lemma. *For all $\mu, \nu \in M_1(X)$, one has*

$$\|\mu - \nu\|_0^* = \widehat{W}(\mu,\nu) := \sup\biggl\{\int_X f\,d\mu + \int g\,d\nu\colon \quad f, g \in C(X), f(x) + g(y) \le d(x,y)\biggr\}. \tag{8.10.4}$$

Proof. We have $\|\mu - \nu\|_0^* \le \widehat{W}(\mu, \nu)$, since $f(x) - f(y) \le d(x, y)$ for all $f \in \mathrm{Lip}_1(X)$ and one can take $g(y) = -f(y)$. On the other hand, if f and g are such that $f(x) + g(y) \le d(x, y)$, then, letting $h(x) = \inf_y[d(x, y) - g(y)]$, we obtain $f \le h \le -g$ and $h(x) - h(x') \le \sup_y[d(x, y) - d(x', y)] \le d(x, x')$ for all x, x', whence we have $h \in \mathrm{Lip}_1(X)$. In addition,

$$\int f\,d\mu + \int g\,d\nu \le \int h\,d(\mu-\nu).$$

Thus, equality (8.10.4) is proven. □

The next result gives another expression for $\|\mu - \nu\|_0^*$.

8.10.45. Theorem. *The Kantorovich–Rubinshtein distance $\|\mu-\nu\|_0^*$ between Radon probability measures μ and ν in the class $M_1(X)$ can be represented in the form*

$$\|\mu-\nu\|_0^* = W(\mu,\nu) := \inf_{\lambda\in M(\mu,\nu)} \int_{X\times X} d(x,y)\,\lambda(dx,dy), \tag{8.10.5}$$

where $M(\mu,\nu)$ is the set of all Radon probability measures λ on $X\times X$ such that the projections of λ to the first and second factors are μ and ν. In addition, there exists a measure $\lambda_0\in M(\mu,\nu)$ at which the value $W(\mu,\nu)$ is attained.

PROOF. We observe that $\|\mu-\nu\|_0^*\le W(\mu,\nu)$, since for all $\lambda\in M(\mu,\nu)$ and every function $f\in \mathrm{Lip}_1(X)$ we have

$$\int_X f\,d(\mu-\nu) = \int_{X\times X} [f(x)-f(y)]\,\lambda(dx,dy) \le \int_{X\times X} d(x,y)\,\lambda(dx,dy).$$

The case of a finite space X is left as Exercise 8.10.111. In the general case, we find two sequences of probability measures μ_n and ν_n that have finite supports X_n and converge weakly to μ and ν, respectively, such that both sequences are uniformly tight. We may assume that all sets X_n contain some point a. Let $\lambda_n\in M(\mu_n,\nu_n)$ be a probability measure on $X_n\times X_n$ with

$$\int_{X\times X} d(x,y)\,\lambda_n(dx,dy) = \|\mu_n-\nu_n\|_0^*.$$

The sequence of measures λ_n with uniformly tight projections is uniformly tight on X. Passing to a subsequence, we may assume that the measures λ_n converge weakly to a measure λ on $X\times X$. It is clear that $\lambda\in M(\mu,\nu)$. Since the measures μ and ν are Radon, we can assume that the space X is separable. For every n, there is a function f_n on X_n that is Lipschitzian with constant 1, $f_n(a)=0$ and

$$\begin{aligned}\int_X f_n\,d(\mu_n-\nu_n) &= \|\mu_n-\nu_n\|_0^* = W(\mu_n,\nu_n)\\ &= \int_{X\times X} d(x,y)\,\lambda_n(dx,dy).\end{aligned} \tag{8.10.6}$$

The functions f_n can be extended to the whole space X with the same Lipschitzian constant (see Exercise 8.10.71). We denote the extension again by f_n and find in $\{f_n\}$ a subsequence convergent on a countable everywhere dense set. By the uniform Lipschitzness this subsequence, denoted again by $\{f_n\}$, converges at every point. It is clear that the limit f of this subsequence is Lipschitzian with constant 1 and $f(a)=0$. By Theorem 8.2.18 we obtain

$$\lim_{n\to\infty}\int_X f_n\,d(\mu_n-\nu_n) = \int_X f\,d(\mu-\nu). \tag{8.10.7}$$

In addition, one has

$$\int_{X\times X} d(x,y)\,\lambda(dx,dy) \le \liminf_{n\to\infty}\int_{X\times X} d(x,y)\,\lambda_n(dx,dy)$$

by the continuity of the function d (see Exercise 8.10.73). Therefore, taking into account (8.10.6) and (8.10.7) we obtain

$$W(\mu,\nu) \le \int_{X\times X} d(x,y)\,\lambda(dx,dy) \le \int_X f\,d(\mu-\nu) \le \|\mu-\nu\|_0^*.$$

Since $\|\mu-\nu\|_0^* \le W(\mu,\nu)$, one has equalities in this chain of inequalities. $\square$

Note that if (X,d_0) is any metric space, then the metric $d = d_0/(1+d_0)$ (or the metric $d=\min(1,d_0)$) generates the same topology and is bounded, so $M_1\bigl((X,d)\bigr) = \mathcal{P}_\mathcal{B}(X)$. Hence the function W defined in (8.10.5) is a metric on $\mathcal{P}_r(X)$ that generates the weak topology. But this is not true for the original metric d_0 if $\operatorname{diam}(X,d_0)=\infty$: taking $c_n := d_0(x_n,x_0)\to\infty$, we obtain $\mu_n := c_n(c_n+1)^{-1}(\delta_{x_0} + c_n^{-1}\delta_{x_n}) \Rightarrow \delta_{x_0}$ and $\|\mu_n-\delta_{x_0}\|_0^* = c_n(c_n+1)^{-1}\to 1$.

In many interesting cases, the measure λ_0, at which the extremum is attained, can be obtained in the form $\mu\circ\Psi^{-1}$ with some measurable mapping $\Psi\colon X\to X\times X$. Moreover, under certain assumptions (but not always, of course), the mapping Ψ can be even obtained in the form $\Psi(x) = \bigl(x,F(x)\bigr)$ with some mapping $F\colon X\to X$. This is one of the links between this problem and the study of transformations of measures, since $\nu=\mu\circ F^{-1}$. It should be also added that under broad assumptions, F turns out to be sufficiently regular (for example, it is the differential or subdifferential of a convex function). Unfortunately, it is not possible to provide here more details on this interesting direction at the intersection of measure theory, variational calculus, and the theory of nonlinear differential equations. The interested reader can consult Ambrosio [**42**], Bogachev, Kolesnikov [**214**], Brenier [**252**], Caffarelli [**300**], Feyel, Üstünel [**571**], Kolesnikov [**1019**], Lipchius [**1175**], McCann [**1285**], Rachev, Rüschendorf [**1508**], Sudakov [**1803**], Villani [**1928**].

We shall mention an interesting theorem due to Strassen [**1791**] (its proof can also be found in the book Dudley [**495**, §11.6]).

8.10.46. Theorem. *Let μ and ν be two Radon probability measures on a metric space (X,d). Then, there exist a probability space $(\Omega,\mathcal{F},P)$ and two measurable mappings ξ and η from Ω to X such that $\mu=P\circ\xi^{-1}$, $\nu=P\circ\eta^{-1}$ and $d_P(\mu,\nu)=K(\mu,\nu)$, where $K(\mu,\nu)$ is the Ky Fan metric defined by the formula*

$$K(\xi,\eta) := \inf\bigl\{\varepsilon>0\colon\ P\bigl(d(\xi,\eta)>\varepsilon\bigr)\le\varepsilon\bigr\}.$$

Regarding measures with given projections to the factors, see also the results in §9.12(vii).

Now we briefly discuss a concept of merging of measures. Let us say that two sequences of Baire measures $\{\mu_n\}$ and $\{\nu_n\}$ on a topological space X are weakly merging if the sequence of measures $\mu_n-\nu_n$ converges weakly to zero. If $\{\mu_n\}$ and $\{\nu_n\}$ are weakly merging sequences of Borel probability measures on a separable metric space (X,d), then according to Exercise 8.10.134 we have $\|\mu_n-\nu_n\|_0\to 0$ and $d_P(\mu_n,\nu_n)\to 0$. However, the fact that $d_P(\mu_n,\nu_n)\to 0$ (or, equivalently, $\|\mu_n-\nu_n\|_0\to 0$) does not imply that $\{\mu_n\}$ and $\{\nu_n\}$ are weakly merging. For example, let μ_n be Dirac's measure at the point n on the real line and let ν_n be Dirac's measure at the point $n+1/n$. It is clear that $\|\mu_n - \nu_n\|_0 \to 0$, but the measures $\delta_n - \delta_{n+1/n}$ do not converge weakly. See Dudley [**495**, § 11.7] for the proof of the following result.

8.10.47. Proposition. *Suppose that (X,d) is a separable metric space and that $\{\mu_n\}, \{\nu_n\} \subset \mathcal{P}_\sigma(X)$. Then the following conditions are equivalent*:

(a) $d_P(\mu_n, \nu_n) \to 0$, (b) $\|\mu_n - \nu_n\|_0 \to 0$,

(c) *there exist a probability space (Ω, P) and measurable mappings ξ_n, η_n from Ω to X such that $P\circ\xi_n^{-1} = \mu_n$, $P\circ\eta_n^{-1} = \nu_n$ and $d(\xi_n, \eta_n) \to 0$ a.e.*

In the case of a separable metric space X a stronger concept of merging of measures, called F-merging, is considered in D'Aristotile, Diaconis, Freedman [**404**], where it is required that $\varrho(\mu_n, \nu_n) \to 0$ for every metric ϱ on $\mathcal{P}_\sigma(X)$ that metrizes the weak topology. In order to see that this is indeed a stronger condition, consider the following example. Let a measure μ_n on the real line assign the value $1/n$ to the points $1, \ldots, n$ and let $\nu_n = \mu_{n+1}$. Then the measures $\mu_n - \nu_n$ converge to zero even in the variation norm, but are not F-merging. Indeed, the sets $\{\mu_{2n}\}$ and $\{\nu_{2n}\}$ are closed in $\mathcal{P}(\mathbb{R}^1)$ and disjoint, which yields a function $\Phi \in C_b\bigl(\mathcal{P}(\mathbb{R}^1)\bigr)$ such that the numbers $\Phi(\mu_n) - \Phi(\nu_n)$ do not approach zero (then one can take the metric $d_P(\mu,\nu) + |\Phi(\mu) - \Phi(\nu)|$ on $\mathcal{P}_\sigma(X)$). In [**404**] among other things the following result is established.

8.10.48. Theorem. *Let X be a separable metric space and let $\{\mu_n\}$ and $\{\nu_n\}$ be two sequences in $\mathcal{P}_\sigma(X)$. The following conditions are equivalent*:

(a) *the sequences $\{\mu_n\}$ and $\{\nu_n\}$ are F-merging,*

(b) *for every function $\Phi \in C_b\bigl(\mathcal{P}_\sigma(X)\bigr)$, one has $\Phi(\mu_n) - \Phi(\nu_n) \to 0$,*

(c) *for every function $\Psi \in C_b\bigl(\mathcal{P}_\sigma(X)\times\mathcal{P}_\sigma(X)\bigr)$ vanishing on the diagonal, one has $\Psi(\mu_n, \nu_n) \to 0$.*

If $\mu_n \neq \nu_n$ for all n, then yet another equivalent condition is:

(d) *every subsequence in $\{\mu_n\}$ contains a further subsequence that converges weakly, and the corresponding subsequence in $\{\nu_n\}$ converges weakly to the same limit.*

Finally, if X is complete, then the latter condition is equivalent to that both sequences are uniformly tight and are weakly merging.

Weak merging is equivalent to F-merging precisely when X is compact. Analogous problems are studied in [**404**] for nets.

8.10(ix). Uniformly distributed sequences

An interesting concept related to weak convergence of measures is that of a uniformly distributed sequence. We shall give several basic facts related to this concept and refer the reader to detailed accounts in the books Hlawka [**836**] and Kuipers, Niederreiter [**1074**], which contain extensive bibliographies. Note only that as early as at the beginning of the 20th century, P. Bol, W. Sierpiński, and H. Weyl (see [**1976**]) studied uniformly distributed sequences of numbers, and at the beginning of the 1950s the study of their analogs in topological spaces began (see Hlawka [**835**]).

8.10.49. Definition. *A sequence of points x_n in a topological space X is called uniformly distributed with respect to a Borel (or Baire) probability measure μ on X if the measures $(\delta_{x_1} + \cdots + \delta_{x_n})/n$ converge weakly to μ.*

Thus, it is required that for all $f \in C_b(X)$

$$\lim_{n\to\infty} \frac{f(x_1) + \cdots + f(x_n)}{n} = \int_X f(x)\,\mu(dx).$$

An important example of a uniformly distributed sequence was indicated independently by P. Bol, W. Sierpiński, and H. Weyl (its justification is left as Exercise 8.10.103). Let $[x]$ denote the integer part of a real number x.

8.10.50. Example. (i) For every irrational number $\theta \in (0,1)$, the sequence $x_n := n\theta - [n\theta]$ is uniformly distributed with respect to Lebesgue measure on $[0,1]$.

It is clear from the properties of weak convergence that for every uniformly distributed sequence $\{x_n\}$ in $[0,1]$ with Lebesgue measure, the quantities $n^{-1}\sum_{i=1}^n f(x_i)$ converge to the integral of f for every Riemann integrable function f.

We observe that if $\{x_n\}$ is a uniformly distributed sequence for a Radon measure μ on a completely regular space X and $T\colon\, X \to Y$ is a Borel mapping to a space Y such that the set of discontinuity points of T has μ-measure zero, then $\mu\circ T^{-1}$ is a Radon measure and the sequence $\{T(x_n)\}$ is uniformly distributed with respect to $\mu\circ T^{-1}$ (see Theorem 8.4.1). This simple fact along with Theorem 9.12.29 enables one to construct uniformly distributed sequences in many spaces. The existence of such sequences can be deduced from a general theorem due to Niederreiter [**1370**], proven below. The proof is based on the following combinatorial lemma.

8.10.51. Lemma. *Let X be a nonempty set. For every probability measure ν with a finite support $\{z_1,\ldots,z_k\} \subset X$, there exists a sequence $\{y_n\}$ with $y_n \in \{z_1,\ldots,z_k\}$ such that, for every set $M \subset X$ and every $N \in \mathbb{N}$, one has*

$$\Bigl|\frac{S_N(M,\{y_n\})}{N} - \nu(M)\Bigr| \le \frac{C(\nu)}{N}, \tag{8.10.8}$$

where $S_N(M,\{y_n\}) := \sum_{n=1}^N I_M(y_n)$ *and* $C(\nu) = (k-1)k$.

PROOF. Suppose that we have found a sequence $\{y_n\}$ such that

$$\Bigl|\frac{S_N(z_i,\{y_n\})}{N} - \nu(z_i)\Bigr| \le \frac{k-1}{N}, \quad \forall\, i \le k, \forall\, N \ge 1. \tag{8.10.9}$$

Then one can take $C(\nu) = (k-1)k$. Indeed, since $y_n \in \{z_1,\ldots,z_k\}$ and μ is concentrated at $\{z_1,\ldots,z_k\}$, it suffices to verify (8.10.8) for sets M in $\{z_1,\ldots,z_k\}$. Then the left-hand side of (8.10.8) is estimated by $k(k-1)N^{-1}$ in view of (8.10.9). Now we show by induction on k that one can obtain (8.10.9). If $k=1$, then we take the sequence $y_n \equiv z_1$. Suppose that our claim is true for $k-1$. Let $\nu(z_i) = \lambda_i > 0$, $i = 1,\ldots,k$. Let us consider a probability measure ν' with support at the points $z_1,\ldots,z_{k-1}$ and

$\nu'(z_i) = \lambda_i(1-\lambda_k)^{-1}$. By the inductive assumption there exists a sequence $\{y'_n\}$ such that $y'_n \in \{z_1, \ldots, z_{k-1}\}$ and

$$\Bigl|\frac{S_N(z_i, \{y'_n\})}{N} - \nu'(z_i)\Bigr| \le \frac{k-2}{N}, \quad \forall\, i \le k-1, \forall\, N \ge 1.$$

Now we define the sequence $\{y_n\}$ as follows: if $n = \bigl[m(1-\lambda_k)^{-1}\bigr]$ for some $m \in \mathbb{N}$, where $[p]$ is the integer part of p, then we set $y_n := y'_m$, otherwise we set $y_n := z_k$. Note that such a number m is unique. We verify (8.10.9). Let us consider the case $i \le k-1$. Then $S_N(z_i, \{y_n\})$ equals the cardinality of the set of natural numbers m such that $\bigl[m(1-\lambda_k)^{-1}\bigr] \le N$ and $y'_m = z_i$. Hence $S_N(z_i, \{y_n\}) = S_L(z_i, \{y'_n\})$, where $L = [(N+1)(1-\lambda_k)] - \varepsilon$ and $\varepsilon = 1$ or 0 depending on whether the number $(N+1)(1-\lambda_k)$ is an integer or not. Thus,

$$\begin{aligned}
&\Bigl|\frac{S_N(z_i, \{y_n\})}{N} - \nu(z_i)\Bigr| = \Bigl|\frac{L}{N}\frac{S_L(z_i, \{y'_n\})}{L} - (1-\lambda_k)\nu'(z_i)\Bigr| \\
&\le \frac{L}{N}\Bigl|\frac{S_L(z_i, \{y'_n\})}{L} - \nu'(z_i)\Bigr| + \nu'(z_i)\Bigl|\frac{L}{N} - (1-\lambda_k)\Bigr| \\
&\le \frac{k-2}{N} + \frac{\nu'(z_i)}{N}\bigl|N(1-\lambda_k) - [(N+1)(1-\lambda_k)] + \varepsilon\bigr|.
\end{aligned}$$

It remains to observe that the second summand on the right-hand side is estimated by N^{-1}, since the number $\bigl|N(1-\lambda_k) - [(N+1)(1-\lambda_k)] + \varepsilon\bigr|$ equals λ_k if $(N+1)(1-\lambda_k)$ is an integer and this number does not exceed 1 otherwise. Finally, let us consider the point z_k. It is readily seen that we have the equality $S_N(z_k, \{y_n\}) = N - L$, where L is defined above. Hence one has

$$\Bigl|\frac{S_N(z_k, \{y_n\})}{N} - \nu(z_k)\Bigr| = \Bigl|\lambda_1 + \cdots + \lambda_{k-1} - \frac{L}{N}\Bigr| \le \frac{1}{N},$$

which completes the proof. □

Now we prove a criterion of the existence of uniformly distributed sequences.

8.10.52. Theorem. *Let μ be a Radon (or τ-additive) probability measure on a completely regular space X. The existence of a sequence uniformly distributed with respect to μ is equivalent to the existence of a sequence of probability measures with finite supports weakly convergent to μ.*

PROOF. If $\{x_n\}$ is a uniformly distributed sequence in the space X, then the measures $n^{-1}(\delta_{x_1} + \cdots + \delta_{x_n})$ have finite supports and converge weakly to μ. The converse is not that simple. Suppose that probability measures μ_j with finite supports converge weakly to μ. By the above lemma, for every j, there exist a number $C_j := C(\mu_j)$ and a sequence $\{y^j_n\}$ such that for all $M \subset X$ and $N \in \mathbb{N}$ one has the inequality

$$\Bigl|\frac{S_N(M, \{y^j_n\})}{N} - \mu_j(M)\Bigr| \le \frac{C_j}{N}.$$

For every j we take a natural number $r_j \geq j(C_1 + \cdots + C_{j+1})$. Now we construct the required sequence $\{x_n\}$ as follows. Every natural number n is uniquely written in the form $n = r_1 + \cdots + r_{j-1} + s$, where $j \in \mathbb{N}$, $0 < s \leq r_j$, and $r_0 := 0$. Let $x_n := y_s^j$. The obtained sequence is as required. Indeed, let a set M have the boundary of μ-measure zero. Every natural number $N > r_1$ is written in the form $N = r_1 + \cdots + r_k + r$, $0 < r \leq r_{k+1}$. Then, as one can easily verify, we have

$$S_N(M, \{x_n\}) = \sum_{j=1}^{k} S_{r_j}(M, \{y_n^j\}) + S_r(M, \{y_n^{k+1}\}).$$

Therefore,

$$\frac{S_N(M, \{x_n\})}{N} - \mu(M) = \sum_{j=1}^{k} \frac{r_j}{N}\Bigl(\frac{S_{r_j}(M, \{y_n^j\})}{r_j} - \mu_j(M)\Bigr)$$

$$+ \frac{r}{N}\Bigl(\frac{S_r(M, \{y_n^{k+1}\})}{r} - \mu_{k+1}(M)\Bigr) + \sum_{j=1}^{k} \frac{r_j}{N}\mu_j(M) + \frac{r}{N}\mu_{k+1}(M) - \mu(M),$$

which is bounded in the absolute value by

$$\sum_{j=1}^{k} \frac{r_j}{N}\frac{C_j}{r_j} + \frac{r}{N}\frac{C_{k+1}}{r} + \Bigl|\frac{1}{N}\Bigl(\sum_{j=1}^{k} r_j\mu_j(M) + r\mu_{k+1}(M)\Bigr) - \mu(M)\Bigr|$$

$$\leq \frac{1}{r_k}\sum_{j=1}^{k+1} C_j + \Bigl|\frac{1}{N}\Bigl(\sum_{j=1}^{k} r_j\mu_j(M) + r\mu_{k+1}(M)\Bigr) - \mu(M)\Bigr|$$

$$\leq \frac{1}{k} + \Bigl|\frac{1}{N}\Bigl(\sum_{j=1}^{k} r_j\mu_j(M) + r\mu_{k+1}(M)\Bigr) - \mu(M)\Bigr|.$$

Letting $N \to \infty$ we have $k \to \infty$. Hence the first term on the right-hand side of the obtained estimate tends to zero. The second term tends to zero as well, since we have $\mu_j(M) \to \mu(M)$ by weak convergence and the equality $N = \sum_{j=1}^{k} r_j + r$. □

8.10.53. Corollary. *Let X be a completely regular space. The following conditions are equivalent*:

(i) *for every Radon probability measure on X, there exists a uniformly distributed sequence,*

(ii) *the sequential closure of the set of probability measures with finite support coincides with $\mathcal{M}_r(X)$.*

In particular, for every Borel probability measure on a completely regular Souslin space, there exists a uniformly distributed sequence.

We emphasize that it is important in this corollary to deal with the sequential closure (the set of the limits of all convergent sequences), but not with the larger closure in the usual topological sense, which, as we know, always coincides with $\mathcal{M}_r(X)$. Not every Radon measure on an arbitrary compact space has a uniformly distributed sequence. Let us consider an example constructed by Losert [**1187**].

8.10.54. Example. Let $X = \beta\mathbb{N}$ be the Stone–Čech compactification of $\mathbb{N}$. Then, there exists a Radon probability measure on X that has no uniformly distributed sequences.

PROOF. We show that any atomless Radon probability measure μ on $\beta\mathbb{N}$ has no uniformly distributed sequences. The existence of atomless measures on $\beta\mathbb{N}$ follows by Theorem 9.1.9, since $\beta\mathbb{N}$ can be mapped continuously onto $[0,1]$. To this end, we set $f(n) = r_n$, where $\{r_n\}$ is the set of all rational numbers in $[0,1]$. Next we extend f to a continuous function on $\beta\mathbb{N}$ with values in $[0,1]$. Suppose that there is a sequence of discrete measures weakly convergent to μ. Then, by Proposition 8.10.59 below, the measure μ is concentrated on a countable set, which contradicts the fact that it has no atoms. □

The following result is due to Losert [**1188**] too.

8.10.55. Proposition. *Let X be a compact space such that there exists a continuous mapping from the space $\{0,1\}^{\aleph_1}$ onto X, where $\aleph_1$ is the least uncountable cardinal. Then every Radon probability measure on X has a uniformly distributed sequence. In particular, this is true for $[0,1]^{\mathfrak{c}}$ under the continuum hypothesis.*

Additional information on uniformly distributed sequences can be found in the above cited works and in Losert [**1189**], Mercourakis [**1303**], Plebanek [**1471**], and Sun [**1806**], [**1807**], as well as in Exercises 8.10.104–8.10.109.

8.10(x). Setwise convergence of measures

As early as in 1916, G.M. Fichtenholz (see [**576**, §30], [**578**] and the comments in V. 1 related to §§4.5–4.6) discovered a remarkable fact: if the integrals of functions f_n over every open set in the interval $[0,1]$ converge to zero, then the integrals over every Borel set converge to zero as well. Thirty-five years later Dieudonné [**448**] proved that if a sequence of measures on a compact metric space converges on every open set, then it converges on every Borel set. Grothendieck [**744**] extended the Dieudonné theorem to locally compact spaces. The method used by Fichtenholz can be modified for Radon measures; moreover, in view of Theorem 9.6.3, his result yields easily the Dieudonné result. So the assertion that a sequence of Radon measures convergent on open sets converges on all Borel sets can naturally be called the Fichtenholz–Dieudonné–Grothendieck theorem. Later several authors extended the result to more general cases. We shall give a proof of a useful generalization obtained in Pfanzagl [**1442**], and then mention a number of other results.

8.10.56. Theorem. *Let a topology base $\mathcal{U}_0$ in a Hausdorff space X be closed with respect to countable unions and let a sequence of Radon measures μ_n converge on every set in $\mathcal{U}_0$. Then it converges on every Borel set.*

PROOF. The assertion reduces to the case where the measures μ_n converge to zero on every set in $\mathcal{U}_0$. Indeed, if the assertion is false, then there exist $B \in \mathcal{B}(X)$ and $\varepsilon > 0$ such that for every n there exists $k(n) > n$ with $|\mu_n(B) - \mu_{k(n)}(B)| > \varepsilon$. Then the sequence of measures $\mu_n - \mu_{k(n)}$ converges to zero on all sets in $\mathcal{U}_0$, but not on B.

We assume further that $\mu_n(U) \to 0$ for all $U \in \mathcal{U}_0$. Let C be compact. We show that for every $\varepsilon > 0$, there exists $U \in \mathcal{U}_0$ such that

$$C \subset U, \quad |\mu_n|(U\backslash C) \le \varepsilon \quad \text{for all } n. \tag{8.10.10}$$

Otherwise for some $\varepsilon > 0$ and all $U \in \mathcal{U}_0$ with $C \subset U$ we have $|\mu_n|(U\backslash C) > \varepsilon$ for infinitely many n. Indeed, if the set of such numbers n were finite and consisted of the elements $n_1, \ldots, n_k$, then due to the assumption that $\mathcal{U}_0$ is a topology base closed with respect to finite unions, one could find a set $V \in \mathcal{U}_0$ such that $C \subset V \subset U$ and $\sum_{i=1}^k |\mu_{n_i}|(V\backslash C) < \varepsilon$. Let us verify that there exist a decreasing sequence of sets $U_i \in \mathcal{U}_0$ with $C \subset U_i$, sets $V_i \in \mathcal{U}_0$ with $V_i \subset U_{i-1}\backslash U_i$, and an increasing sequence of numbers n_i such that $|\mu_{n_i}(V_i)| > \varepsilon/4$ for all i. We argue by induction. Let U_0 be any set in $\mathcal{U}_0$ containing C. Suppose that U_i, V_i, and n_i are constructed for $i = 1, \ldots, j-1$. As noted above, there exists $n_j > n_{j-1}$ with $|\mu_{n_j}|(U_{j-1}\backslash C) > \varepsilon$. Let us take a compact set $C_j \subset U_{j-1}\backslash C$ with $|\mu_{n_j}(C_j)| > \varepsilon/2$. The compact sets C and C_j do not meet and hence possess disjoint neighborhoods. Hence one can find sets $U_j, V_j \in \mathcal{U}_0$ such that $U_j \cap V_j = \varnothing$, $C \subset U_j \subset U_{j-1}$, $C_j \subset V_j \subset U_{j-1}$, and $|\mu_{n_j}|(V_j\backslash C_j) < \varepsilon/4$. It is then clear that $V_j \subset U_{j-1}\backslash U_j$ and

$$|\mu_{n_j}(V_j)| \ge |\mu_{n_j}(C_j)| - |\mu_{n_j}|(V_j\backslash C_j) > \varepsilon/4.$$

The constructed sets V_i are disjoint, since $V_i \subset U_{i-1}\backslash U_i$. According to Exercise 8.10.112, there is an infinite set $S \subset \mathbb{N}$ with $\inf_{i\in S}\bigl|\mu_{n_i}\bigl(\bigcup_{j\in S} V_j\bigr)\bigr| > 0$. Since $\bigcup_{j\in S} V_j \in \mathcal{U}_0$, we arrive at a contradiction, which proves (8.10.10).

Now we show that for every $B \in \mathcal{B}(X)$ and every $\varepsilon > 0$, there exists a compact set $C \subset B$ such that $|\mu_n|(B\backslash C) \le \varepsilon$ for all $n \in \mathbb{N}$. Together with (8.10.10) this will yield that $\lim\limits_{n\to\infty} \mu_n(B) = 0$. Suppose that for some $B \in \mathcal{B}(X)$ and $\varepsilon > 0$, there is no such compact set. It is then clear that for every compact set $C \subset B$, we obtain $|\mu_n|(B\backslash C) > \varepsilon$ for infinitely many numbers n. We show that this gives a sequence of disjoint compact sets $C_i \subset B$ and a sequence of numbers n_i with $|\mu_{n_i}(C_i)| > \varepsilon/2$ for all i. These sequences are constructed inductively, by setting $C_0 = \varnothing$. If C_i and n_i are already found for all $i \le j-1$, then by the compactness of $K_j := C_1 \cup \cdots \cup C_{j-1}$ and the above observation, there exists $n_j > n_{j-1}$ with $|\mu_{n_j}|(B\backslash K_j) > \varepsilon$. Next we find a compact set $C_j \subset B\backslash K_j$ with $|\mu_{n_j}(C_j)| > \varepsilon/2$. Relationship (8.10.10) implies that the values of the measures μ_n on every compact set tend to zero. Applying Exercise 8.10.112 once again, we obtain an infinite set

$D \subset \mathbb{N}$ such that

$$\delta := \inf_{i \in D} \Big| \mu_{n_i}\Big(\bigcup_{j \in D} C_j\Big)\Big| > 0.$$

By (8.10.10) for every j, there exists $V_j \in \mathcal{U}_0$ with $C_j \subset V_j$ and

$$|\mu_n|(V_j \backslash C_j) < \delta 2^{-j-1}$$

for all $n \in \mathbb{N}$. Then $|\mu_n|\big(\bigcup_{j\in D} V_j \backslash \bigcup_{j \in D} C_j\big) \le \delta/2$ for all $n \in \mathbb{N}$. Hence $\big|\mu_{n_i}\big(\bigcup_{j \in D} V_j\big)\big| \ge \delta/2$ for all $i \in D$, which is a contradiction. □

The measures μ_n have densities f_n with respect to some bounded Radon measure ν (for example, of the form $\sum_{n=1}^{\infty} c_n|\mu_n|$), and it follows by Theorem 4.5.6 that the functions f_n are uniformly integrable and converge to some function $f \in L^1(\nu)$ in the weak topology of $L^1(\nu)$. In particular, convergence of μ_n takes place on even a larger class than $\mathcal{B}(X)$. The limit of $\{\mu_n\}$ is a Radon measure. Finally, the above theorem yields the fact (which is not obvious) that the measures μ_n are uniformly bounded. However, this fact can be obtained under a weaker hypothesis.

8.10.57. Corollary. *Suppose that a topology base $\mathcal{U}_0$ in a Hausdorff space X is closed with respect to countable unions. Let a family M of Radon measures on X be such that $\sup\big\{|\mu(U)|\colon\ \mu \in M\big\} < \infty$ for all $U \in \mathcal{U}_0$. Then the family M is bounded in the variation norm.*

Proof. It suffices to deal with a sequence of measures μ_n bounded on every $U \in \mathcal{U}_0$. If it is not bounded in the variation norm, then we may assume that $\|\mu_n\| \ge n$. Then the sequence $n^{-1/2}\mu_n$ converges to zero on $\mathcal{U}_0$. By the above theorem it converges to zero on every Borel set, which by Corollary 4.6.4 yields the boundedness in the variation norm contrary to the estimate $\|n^{-1/2}\mu_n\| \ge n^{1/2}$. □

8.10.58. Theorem. *Let M be a bounded set of Radon measures on a Hausdorff space X. Then M has compact closure in the topology of convergence on Borel sets precisely when $\lim\limits_{n\to\infty} \sup_{\mu \in M} |\mu(K_n)| = 0$ for every sequence of pairwise disjoint compact sets K_n. If X is regular, then this is equivalent to the condition that for every sequence of pairwise disjoint open sets U_n one has $\lim\limits_{n\to\infty} \sup_{\mu\in M} |\mu(U_n)| = 0$.*

Proof. The first claim follows by Lemma 4.6.5 and the Radon property of our measures. The necessity of the second condition is also clear from that lemma. For the proof of sufficiency we observe that for every compact set K and every $\varepsilon > 0$, there exists an open set $U \supset K$ such that $|\mu|(U\backslash K) \le \varepsilon$ for all $\mu \in M$. Otherwise we let $V_1 = X$ and take a measure $\mu_1 \in M$ with $|\mu_1|(V_1\backslash K) > \varepsilon$. The set $V_1\backslash K$ contains a compact set S with $|\mu_1|(S) > \varepsilon$. The compact sets S and K have disjoint neighborhoods U_1 and V_2. Then we repeat the construction for V_2 and continue it inductively, which gives a sequence of pairwise disjoint open sets U_n and measures μ_n with $|\mu_n|(U_n) > \varepsilon$ contrary to the hypothesis. It remains to verify that $\lim\limits_{n\to\infty} \sup_{\mu\in M} |\mu|(A_n) = 0$

for every disjoint sequence of compact sets A_n. If this is not the case, then for some $\varepsilon > 0$, there exist measures $\mu_k \in M$ and indices n_k with $|\mu_k|(A_{n_k}) > \varepsilon$. As we have shown, there exists a neighborhood W_1 of the compact set A_{n_1} such that $|\mu|(W_1 \backslash A_{n_1}) < \varepsilon/4$ for all $\mu \in M$. By the regularity of X, there exists a neighborhood V_1 of the compact set A_{n_1} such that $\overline{V_1} \subset W_1$. The sets $A_{n_k} \backslash W_1$ are compact and disjoint and $|\mu_k|(A_{n_k} \backslash W_1) > 3\varepsilon/4$ for all $k \geq 2$. By induction we construct pairwise disjoint open sets V_k with $|\mu_k|(V_k) > \varepsilon/2$. The obtained contradiction completes the proof. □

It should be noted that Theorem 8.10.56 fails for arbitrary Borel measures (Exercise 8.10.113). However, the Radon property of measures can be somewhat weakened at the expense of certain restrictions on the space. For example, if X is regular and the measures μ_n are τ-additive, then, as shown in Adamski, Gänssler, Kaiser [**11**], convergence on every open set implies convergence on every Borel set (moreover, it suffices to have convergence on every regular open set, i.e., a set that is the interior of its closure). In this case, one says that the class of open sets is a convergence class. If X is completely regular and the measures μ_n are τ-additive, then the class of functionally open sets is a convergence class, see [**11**]. If we deal with Baire measures μ_n, then, according to Landers, Rogge [**1103**], convergence on functionally open sets implies convergence on all Baire sets for every topological space. More special results in this direction and additional references can be found in Adamski, Gänssler, Kaiser [**11**], Gänssler [**652**], [**653**], Landers, Rogge [**1103**], Rogge [**1591**], Sazhenkov [**1654**], Stein [**1780**], Topsøe [**1872**], Wells [**1972**].

We know that setwise convergence implies weak convergence of measures, but the converse is false in general. However, there is a class of spaces for which the converse is true as well. We recall that a compact space X is called extremally disconnected if the closure of every open set is open (see Engelking [**532**, §6.2]). This is equivalent to saying that the closures of disjoint open sets in X do not meet. Note that X has a topology base consisting of sets that are simultaneously open and closed (such sets are called clopen). The following result is due to Grothendieck [**744**].

8.10.59. Proposition. *Let X be an extremally disconnected compact space. Then every weakly convergent sequence of Radon measures converges on every Borel set. In particular, this is true if $X = \beta\mathbb{N}$ is the Stone–Čech compactification of $\mathbb{N}$.*

Proof. We may assume that our sequence of measures μ_n converges weakly to zero. Suppose that we are given a sequence of pairwise disjoint sets V_k that are open and closed. By weak convergence of our measures to zero and continuity of I_{V_k} for every k, we have $\lim\limits_{n\to\infty} \mu_n(V_k) = 0$. In addition, for every subset $S \subset \mathbb{N}$, the closure $Z(S)$ of the open set $\bigcup_{k\in S} V_k$ is open (by the definition of extremal disconnectedness) and hence $\lim\limits_{n\to\infty} \mu_n\bigl(Z(S)\bigr) = 0$. If the sets S_1 and S_2 are disjoint, then $Z(S_1)$ and $Z(S_2)$ are disjoint as well and $Z(S_1 \cup S_2) = Z(S_1) \cup Z(S_2)$. Thus, on the set of all subsets of $\mathbb{N}$ we obtain the additive functions $\nu_k(S) := \mu_k\bigl(Z(S)\bigr)$ such that $\lim\limits_{n\to\infty} \nu_n(S) = 0$ for all $S \subset \mathbb{N}$. By Lemma 4.7.41 one has $\lim\limits_{n\to\infty} \sum_{k=1}^\infty |\nu_n(k)| = 0$. So $\lim\limits_{n\to\infty} \mu_n\bigl(\bigcup_{k=1}^\infty V_k\bigr) = 0$. It remains to observe that for every open set U in X, one can find a sequence of disjoint clopen sets $V_k \subset U$ with $|\mu_n|\bigl(U\backslash\bigcup_{k=1}^\infty V_k\bigr) = 0$ for all n. To this end, we take the measure $\nu := \sum_{n=1}^\infty 2^{-n}(\|\mu_n\|+1)^{-1}|\mu_n|$, find a clopen set $V_1 \subset U$ with $\nu(U\backslash V_1) < 1/2$ (which is possible because ν is Radon and there is a base of topology consisting of clopen sets), then we find a clopen set V_2 in the open set $U_1 := U\backslash V_1$ with $\nu(U_1\backslash V_1) < 1/4$ and so on. It follows that $\lim\limits_{n\to\infty} \mu_n(U) = 0$. □

The measurability of mappings of the form $\mu \mapsto \mu(A)$ was investigated in Ressel [**1555**]. Here are two results from his work. Let X be a Hausdorff space and let $\mathcal{K}(X)$ be the set of all its compact subsets. The space $\mathcal{K}(X)$ can be equipped with a natural topology (the Vietoris topology, see Fedorchuk, Filippov [**561**, Ch. 4]) that is generated by all sets of the form $\{K \in \mathcal{K}(X)\colon\ K \subset U\}$ and $\{K \in \mathcal{K}(X)\colon\ K \cap U \neq \varnothing\}$, where $U \subset X$ is open. If X is a Polish space, then so is $\mathcal{K}(X)$ with the Vietoris topology.

8.10.60. Theorem. *Let X be a Souslin space and let the space $\mathcal{M}^+(X)$ of nonnegative Radon measures be equipped with the weak topology* (*or the A-topology if X is not completely regular*).

(i) *If Y is a Polish space and $f\colon Y \to X$ is a continuous mapping, then the function $(\mu, K) \to \mu\bigl(f(K)\bigr)$ on $\mathcal{M}^+(X)\times\mathcal{K}(Y)$ is upper semicontinuous.*

(ii) *If $A \subset X$ is a Souslin set, then the function $\varphi_A\colon\ \mu \mapsto \mu(A)$ on $\mathcal{M}^+(X)$ is an S-function, i.e., the sets $\{\varphi_A > t\}$ are Souslin for all $t \in \mathbb{R}^1$. If A is a set in the σ-algebra generated by Souslin sets, then the function φ_A is measurable with respect to the σ-algebra generated by Souslin sets.*

8.10.61. Theorem. (i) *Let X, Y, and Z be Souslin spaces and let a mapping $f\colon\ X\times Y \to Z$ be universally measurable* (*i.e., $f^{-1}(B)$ is measurable with respect to all Borel measures on $X\times Y$ for all $B \in \mathcal{B}(Z)$*). *Let us set $f_y(x) := f(x,y)$. We equip the space of measures with the A-topology* (*the weak topology in the case of completely regular spaces*). *Then the mapping*

$$F\colon\ \mathcal{M}^+(X)\times Y \to \mathcal{M}^+(Z),\ (\mu, y) \mapsto \mu\circ f_y^{-1},$$

is universally measurable. In addition, if f is continuous or Borel, then so is F. Finally, if f is measurable with respect to the σ-algebra generated by Souslin sets, then F has the same property.

(ii) *If, additionally, $Z = \mathbb{R}^1$ and the function f is bounded, then the following function is universally measurable*:

$$\Psi\colon\ \mathcal{M}^+(X)\times Y \to \mathbb{R}^1,\ (\mu, y) \mapsto \int_X f(x,y)\,\mu(dx).$$

If f is $\mathcal{A}$-measurable (*or is, respectively, an S-function, Borel measurable, upper semicontinuous, continuous*), *then Ψ has the respective property.*

8.10(xi). Stable convergence and ws-topology

Here we discuss one more mode of convergence of measures, which is useful in applications and combines weak convergence and setwise convergence. Suppose we are given a measurable space $(\Omega, \mathcal{A})$ and a topological space T. Let us consider the space $\mathcal{M}(\Omega\times T)$ of all bounded measures on the product $\Omega\times T$ equipped with one of the σ-algebras $\mathcal{A}\otimes\mathcal{B}(T)$ or $\mathcal{A}\otimes\mathcal{B}a(T)$. The set of all nonnegative measures in $\mathcal{M}(\Omega\times T)$ is denoted by $\mathcal{M}^+(\Omega\times T)$. We say that a net of measures $\mu_\alpha \in \mathcal{M}(\Omega\times T)$ converges to a measure μ in the ws-topology if, for every bounded $\mathcal{A}$-measurable function ψ and every function $\varphi \in C_b(T)$, one has

$$\lim_\alpha \int_{\Omega\times T} \psi(\omega)\varphi(t)\,\mu_\alpha(d\omega dt) = \int_{\Omega\times T} \psi(\omega)\varphi(t)\,\mu(d\omega dt). \tag{8.10.11}$$

It is clear that this convergence is indeed generated by a topology: we equip the space $\mathcal{M}(\Omega\times T)$ with the seminorms

$$\left|\int_{\Omega\times T} \psi(\omega)\varphi(t)\,\mu(d\omega dt)\right|.$$

Fundamental neighborhoods of the element $\mu \in \mathcal{M}(\Omega\times T)$ have the form

$$U_{\psi_1,\ldots,\psi_n;\varphi_1,\ldots,\varphi_n;\varepsilon}(\mu) := \left\{\nu\colon\ \left|\int \psi_j\varphi_j\, d(\nu-\mu)\right| < \varepsilon, j=1,\ldots,n\right\}, \tag{8.10.12}$$

where $\varepsilon > 0$, $\varphi_j \in C_b(T)$, and ψ_j is a bounded $\mathcal{A}$-measurable function. Convergence of a uniformly bounded net (e.g., consisting of probability measures) in the ws-topology is equivalent to equality (8.10.11) with ψ of the form $\psi = I_A$, $A \in \mathcal{A}$. The same is true for nets of nonnegative measures on $\Omega\times T$. If $\mathcal{A} = \{\Omega, \varnothing\}$, then the ws-topology reduces to the weak topology $\mathcal{M}(T)$ and if T is a singleton, then we obtain the topology of convergence on bounded $\mathcal{A}$-measurable functions.

8.10.62. Theorem. *Let T be a completely regular space in which all compact subsets are metrizable and let a net of measures $\mu_\alpha \in \mathcal{M}(\Omega\times T)$ converge to a measure $\mu \in \mathcal{M}(\Omega\times T)$ in the ws-topology and be uniformly bounded in the variation norm. If the projections of the measures $|\mu_\alpha|$ and $|\mu|$ on T are uniformly tight and the projections of the measures $|\mu_\alpha|$ on Ω are uniformly countably additive, then*

$$\lim_\alpha \int f\,d\mu_\alpha = \int f\,d\mu$$

for every bounded $\mathcal{A}\otimes\mathcal{B}(T)$-measurable function f with the property that for every $\omega \in \Omega$, the function $t \mapsto f(\omega, t)$ is continuous.

PROOF. Without loss of generality we may assume that $|f| \le 1$ and $\|\mu_\alpha\| \le 1$, $\|\mu\| \le 1$. Let us fix $\varepsilon > 0$. Let π_T and π_Ω denote the projection mappings on T and Ω, respectively. By hypothesis, there exists a compact set $K \subset T$ such that

$$|\mu_\alpha| \circ \pi_T^{-1}(T\backslash K) + |\mu| \circ \pi_T^{-1}(T\backslash K) \le \varepsilon \quad \text{for all } \alpha.$$

The space $C(K)$ is separable because K is metrizable. For every $\omega \in \Omega$, we denote by g_ω the continuous function $t \mapsto f(\omega, t)$ on K. It is clear that the mapping $g\colon \Omega \to C(K)$, $\omega \mapsto g_\omega$, is Borel. Since the projections of our measures on Ω are uniformly countably additive, there is a probability measure ν on $\mathcal{A}$ with respect to which they have uniformly integrable densities. By using the separability of the Banach space $C(K)$ and applying Lusin's theorem to the mapping g and the measure ν, we can find a finite partition of Ω into sets $A_1, \ldots, A_p, A_{p+1} \in \mathcal{A}$ and functions $f_1, \ldots, f_p \in C(K)$ such that $\|f_i\|_{C(K)} \le 1$, $\|g_\omega - f_i\|_{C(K)} \le \varepsilon$ whenever $\omega \in A_i$, $i \le p$, and

$$|\mu_\alpha| \circ \pi_\Omega^{-1}(A_{p+1}) + |\mu| \circ \pi_\Omega^{-1}(A_{p+1}) \le \varepsilon \quad \text{for all } \alpha.$$

Since T is completely regular, every function f_i extends to T with the preservation of the maximum of the absolute value. The extension is denoted again by f_i. By hypothesis, there exists an index α_0 such that the absolute value of the difference between the integrals of $h(\omega, t) := \sum_{i=1}^p f_i(t) I_{A_i}(\omega)$ against the measures μ_α and μ does not exceed ε for all $\alpha \ge \alpha_0$. We observe that $\sup_x |f(x) - h(x)| \le 2$, $|f(x) - h(x)| \le \varepsilon$ on $\bigcup_{i=1}^p A_i \times K$ and

$$|\mu_\alpha|\bigl(\Omega \times (T\backslash K)\bigr) + |\mu_\alpha|(A_{p+1} \times T) \le 2\varepsilon.$$

It remains to use the estimate

$$\int_{\Omega\times T} |f - h|\, d|\mu_\alpha| \le \int_{\bigcup_{i=1}^p A_i \times K} |f - h|\, d|\mu_\alpha| + 4\varepsilon \le 5\varepsilon,$$

and a similar estimate for μ. □

8.10.63. Corollary. *Suppose that a sequence of nonnegative measures μ_n on $\Omega\times T$ converges to a measure μ in the ws-topology and that T is a Polish space. Then the conclusion of Theorem* 8.10.62 *is valid. More generally, the same is true if T is a Prohorov space in which all compact sets are metrizable, and the projections of the measures μ_n and μ on T are Radon.*

PROOF. We have $\mu_\alpha = |\mu_\alpha|$. The projections of the measures μ_α on Ω are uniformly countably additive, which follows by setwise convergence on $\mathcal{A}$. The projections of the measures μ_α on T converge weakly, hence are uniformly tight (in the case where the space is Prohorov and the projections are Radon, this follows by the hypotheses). □

Under broad assumptions, compact sets in the ws-topology are metrizable, although on the whole space this topology is not metrizable in non-trivial cases.

8.10.64. Proposition. *Let T be a Polish space and let $\mathcal{A}$ be a countably generated σ-algebra. Then any set $M \subset \mathcal{M}^+(\Omega\times T)$ that is compact in the ws-topology is metrizable.*

PROOF. There exists a countable algebra $\mathcal{A}_0 = \{A_n\}$ generating $\mathcal{A}$. In addition, there exists a countable collection of functions $\mathcal{F} = \{f_j\} \subset C_b(T)$ such that the weak topology on $\mathcal{P}(T)$ is generated by the metric

$$d(\mu,\nu) := \sum_{j=1}^{\infty} 2^{-j}\psi_j(\mu-\nu)\bigl(1+\psi_j(\mu-\nu)\bigr)^{-1}, \quad \psi_j(\mu-\nu) = \Bigl|\int f_j\, d(\mu-\nu)\Bigr|.$$

We may assume that $f_1 = 1$. Let us equip $\mathcal{M}^+(\Omega\times T)$ with the metric

$$\varrho(\mu,\nu) := \sum_{n=1}^{\infty} 2^{-n} d\bigl((I_{A_n}\cdot\mu)\circ\pi_T^{-1}, (I_{A_n}\cdot\nu)\circ\pi_T^{-1}\bigr).$$

We observe that the sets of measures obtained from M by projecting on Ω and T are compact in the topology of setwise convergence and in the weak topology respectively. It is readily seen from this that every neighborhood U of $\mu \in M$ in the ws-topology that has the form (8.10.12) with functions $\varphi_j \in C_b(T)$ and $\psi_j = I_{B_j}$, where $B_j \in \mathcal{A}$, contains some ball with respect to the metric ϱ. To this end, we first inscribe in U a neighborhood U' of the form $U_{\psi_1,\ldots,\psi_n;h_1,\ldots,h_n;\varepsilon'}(\mu)$ with $\varepsilon' < \varepsilon$ and $h_j \in \mathcal{F}$. Next we find in U' a neighborhood $U_{g_1,\ldots,g_n;h_1,\ldots,h_n;\varepsilon''}(\mu)$ with $\varepsilon'' < \varepsilon'$ and $g_j = I_{A_{n_j}}$. Note that, without explicit construction of a metric, we could use just as well the fact that the compact set M_Ω is metrizable in the setwise convergence topology (Exercise 4.7.148 in Ch. 4), the compact set M_T is metrizable in the weak topology, and the compact set M is homeomorphic to its image under the natural mapping into the metrizable compact set $M_\Omega\times M_T$. $\square$

The following result is obtained in Raynaud de Fitte [**1546**].

8.10.65. Theorem. *Let T be a metrizable Souslin space with a metric d. Any of the following conditions is equivalent to convergence of a net of measures $\mu_\alpha \in \mathcal{M}^+(\Omega\times T)$ to a measure $\mu \in \mathcal{M}^+(\Omega\times T)$ in the ws-topology:*

(i) *for every bounded $\mathcal{A}\otimes\mathcal{B}(T)$-measurable function f such that the function $t \mapsto f(\omega,t)$ is lower semicontinuous for every $\omega \in \Omega$, one has*

$$\liminf_\alpha \int f\, d\mu_\alpha \ge \int f\, d\mu;$$

(ii) *for every bounded $\mathcal{A}\otimes\mathcal{B}(T)$-measurable function f with the property that the function $t \mapsto f(\omega,t)$ is continuous for every $\omega \in \Omega$, one has*

$$\lim_\alpha \int f\, d\mu_\alpha = \int f\, d\mu;$$

(iii) *the equality in* (ii) *holds for every function f of the form $f(\omega,t) = I_A(\omega)\varphi(t)$, where $A \in \mathcal{A}$ and φ is a bounded Lipschitzian function on T.*

It is not clear whether convergence in the ws-topology implies property (ii) in the case of an arbitrary completely regular space. The ws-topology is also called the stable topology, and the corresponding convergence is called stable convergence (see Rényi [**1551**]). However, in many works this terminology is attached to property (ii), which is equivalent to ws-convergence in the case of a Polish space T.

According to Castaing, Raynaud de Fitte, Valadier [**318**, Theorem 2.2.3], if T is a completely regular space in which all compact sets are metrizable and a measure $\mu \in \mathcal{P}(\Omega\times T)$ is such that its projection μ_T to T is Radon and its projection μ_Ω to Ω has no atoms, then μ is the limit in the ws-topology of a net of measures of the form $\mu_\Omega \circ F_\alpha^{-1}$, $F_\alpha(x) = \bigl(x, \varphi_\alpha(x)\bigr)$, for some measurable mappings $\varphi_\alpha\colon \Omega \to T$. It would be interesting to know whether a convergent sequence in place of a net can be found.

Additional information on the ws-topology can be found in Balder [**96**], Castaing, Raynaud de Fitte, Valadier [**318**], Jacod, Mémin [**877**], Lebedev [**1117**], Letta [**1159**], Raynaud de Fitte [**1546**], Schäl [**1663**].

Exercises

8.10.66°. Prove that a net $\{x_\alpha\}$ of elements of a completely regular space X converges to an element $x \in X$ if and only if the measures δ_{x_α} converge weakly to δ_x.

Hint: observe that if a net $\{x_\alpha\}$ does not converge to x, then there exists its subnet $\{x'_\alpha\}$ such that $f(x'_\alpha) = 0$ and $f(x) = 1$ for some function $f \in C_b(X)$.

8.10.67°. Let X be a completely regular space and let $\{x_n\}$ be a sequence in X such that the sequence of measures δ_{x_n} is weakly fundamental. Show that the sequence $\{x_n\}$ converges in X.

Hint: first observe that $\{x_n\}$ has limit points. Otherwise one can find pairwise disjoint neighborhoods U_n of the points x_n such that U_n contains the closure of some smaller neighborhood W_n of the point x_n. For every n, there is a continuous function f_n with $0 \le f_n \le 1$, $f(x_{2n+1}) = 1$ and $f = 0$ outside W_{2n+1}. The function f that equals f_n on W_{2n+1} and 0 outside the union of the sets W_{2n+1}, is bounded and continuous, but its integrals with respect to $\delta_{x_{2n+1}}$ equal 1, whereas the integrals with respect to $\delta_{x_{2n}}$ equal 0, which contradicts the weak fundamentality. It is readily verified that there is only one limit point. Finally, the same applies to any subsequence in $\{x_n\}$.

8.10.68°. Show that a sequence of measures μ_n on the space $\mathbb{N}$ converges weakly to a measure μ precisely when $\|\mu - \mu_n\| \to 0$.

Hint: see Corollary 4.5.8.

8.10.69°. Give an example of a weakly convergent sequence of signed measures μ_n on $[0,1]$ for which the distribution functions converge at no point of $(0,1)$.

Hint: consider the measures $\mu_n := \delta_{x_n} - \delta_{y_n}$, where the sequence of intervals $[x_n, y_n]$ is obtained in the following way: for every $m \in \mathbb{N}$, we take the consecutive intervals of length 2^{-m} with the endpoints of the form $k2^{-m}$ and arrange all such intervals in a single sequence such that the intervals obtained for $m+1$ are preceded by those obtained for m. The measures μ_n converge weakly to the zero measure, but the functions F_{μ_n} converge at no point of $(0,1)$.

8.10.70°. Give an example of a sequence of probability measures μ_n on the interval $[0,1]$ that are defined by smooth uniformly bounded densities ϱ_n with respect to Lebesgue measure and converge weakly to a measure μ with a smooth density ϱ, but the functions ϱ_n do not converge in measure.

HINT: consider $\varrho_n(x) = 1 + \sin(2\pi nx)$ and $\varrho(x) = 1$.

8.10.71°. Let (X, d) be a metric space. (i) Let f be a bounded function on a set $A \subset X$ with $|f(x) - f(y)| \le d(x, y)$ for all $x, y \in A$. Let

$$g(x) := \max\bigl\{\sup_{y\in A}\bigl(f(y) - d(x, y)\bigr), \inf_A f\bigr\}.$$

Verify that $g(x) = f(x)$ if $x \in A$, $\sup_{y\in X} |g(y)| = \sup_{x\in A} |f(x)|$, and $|g(x) - g(y)| \le d(x, y)$ for all $x, y \in X$. (ii) Prove that every bounded uniformly continuous function on X is uniformly approximated by bounded Lipschitzian functions.

8.10.72°. Let X be an infinite metric space. Show that the weak topology on the space $\mathcal{M}_\sigma(X)$ of signed measures is not metrizable.

HINT: consider the case of a countable space that is either discrete (i.e., the distances between distinct points are separated from zero) or is a Cauchy sequence. The first case reduces to the weak topology of $X = l^1$. In the second case, if a Cauchy sequence $\{x_n\}$ has no limit, then it is homeomorphic to $\mathbb{N}$, hence the first case applies; if $\{x_n\}$ converges to x, then $K = \{x_n\} \cup \{x\}$ is compact, hence $C(K)^*$ is not metrizable in the $*$-weak topology.

8.10.73°. Let Baire probability measures μ_n on a topological space X converge weakly to a measure μ and let $f \ge 0$ be a continuous function. Show that

$$\int_X f\,d\mu \le \liminf_{n\to\infty} \int_X f\,d\mu_n.$$

HINT: let $f_k = \min(f, k)$. Then $f_k \in C_b(X)$ and for all $k \in \mathbb{N}$ one has

$$\int_X f_k\,d\mu = \lim_{n\to\infty} \int_X f_k\,d\mu_n \le \liminf_{n\to\infty} \int_X f\,d\mu_n.$$

8.10.74°. (A.D. Alexandroff [**30**, §17]) Suppose that a sequence of Baire measures $\mu_n \ge 0$ converges weakly to a measure μ and that Z and Z_n, $n \in \mathbb{N}$, are functionally closed sets such that $\mu(Z) = \lim\limits_{n\to\infty} \mu(Z_n)$ and for every n, there exists m with $Z_{n+k} \subset Z_n$ for all $k \ge m$. Prove that $\limsup_{n\to\infty} \mu_n(Z_n) \le \mu(Z)$.

8.10.75. (Varadarajan [**1918**]) Let X be a paracompact space and let τ-additive measures μ_n, $n \in \mathbb{N}$, converge weakly to a Baire measure μ. Prove that μ has a unique τ-additive Borel extension.

HINT: according to Exercise 7.14.123 the topological supports S_n of the measures μ_n are Lindelöf. Let Z be the closure of $\bigcup_{n=1}^\infty S_n$. Then Z is Lindelöf. Indeed, let $\{U_t\}$ be an open cover of Z. As in Exercise 7.14.123, there is a finer open cover $\mathcal{V}$ consisting of a sequence of families $\mathcal{V}_k = \{V_{k,\alpha}\}$, where for each fixed k the sets $V_{k,\alpha}$ are open and disjoint. For every k, there is an at most countable set of indices α_j with $Z \cap V_{k,\alpha_j} \neq \varnothing$, since this is true for every S_n in place of Z, and the union of all S_n is everywhere dense in Z. We obtain a countable cover of Z by the sets V_{k,α_j}, which implies the existence of a countable subcover in $\{U_t\}$. By Exercise 7.14.72 we have $|\mu|_*(X\backslash Z) = 0$. Hence the measure μ is τ_0-additive. Indeed, if X is the union of an increasing net of functionally open sets G_α, we can find a countable sequence $\{G_{\alpha_n}\}$ covering Z, which by the above gives $|\mu|\bigl(X\backslash\bigcup_{n=1}^\infty G_{\alpha_n}\bigr) = 0$. Hence μ has a unique τ-additive Borel extension by Corollary 7.3.3.

8.10.76. (A.D. Alexandroff [**30**], Varadarajan [**1918**]) Let X be a paracompact space and let τ-additive measures μ_n, $n \in \mathbb{N}$, converge weakly to a Baire measure μ. Prove that for every net of open sets U_α increasing to X, one has $\lim\limits_{\alpha} |\mu_n|(X\backslash U_\alpha) = 0$ uniformly in n.

Hint: by the previous exercise the measure μ is τ-additive and there exists a Lindelöf closed subspace $Z \subset X$ with $|\mu|(X\backslash Z) = |\mu_n|(X\backslash Z) = 0$ for all n. The restrictions of the measures μ_n to Z converge weakly to the restriction of the measure μ (every continuous function on Z extends to a continuous function on X because X is normal, and our measures are concentrated on Z). Hence everything reduces to a Lindelöf space, which by the complete regularity of X reduces the claim to the case of a countable increasing sequence of functionally open sets U_k, when Proposition 8.1.12 is applicable.

8.10.77. Let (X, d) be a noncompact metric space. Show that one can find a new metric $\widetilde{d}$ on X defining the same topology and possessing the following property: there exist a sequence of signed Radon measures μ_n and a Radon measure μ such that the integrals of every bounded function f, uniformly continuous in the metric $\widetilde{d}$, with respect to the measures μ_n converge to the integral of f with respect to the measure μ, but the measures μ_n do not converge weakly to μ. The original metric has such a property provided that there are two sequences $\{x_n\}$ and $\{y_n\}$ with $x_n \neq y_n$ which have no limit points and the distance between x_n and y_n tends to zero.

Hint: if the latter condition is fulfilled, then take the measures $\mu_n = \delta_{x_n} - \delta_{y_n}$ and observe that the integrals of any uniformly continuous function against these measures tend to zero. Every point x_n has a neighborhood V_n which contains no point from both sequences distinct from x_n. There is a bounded continuous function f such that $f(x_n) = 1$ for all n and $f = 0$ outside $\bigcup_{n=1}^\infty V_n$. Hence there is no weak convergence of $\{\mu_n\}$ to zero. In the general case we can find a metric d_0 which generates the original topology and $d_0(x, y) \le 1$ for all x, y. Either X contains a sequence $\{x_n\}$ that is Cauchy but not convergent, i.e., the aforementioned condition is fulfilled, or there is a countable set of points x_n whose mutual distances are separated from zero. It suffices to consider the case where for some $r > 0$, there are no points x with $d_0(x, x_n) \le r$ (otherwise we are in the already-considered situation). Now we define a new metric on X as follows: $\widetilde{d}(x, y) = d_0(x, y)$ if $x, y \notin \{x_n\}$, $\widetilde{d}(x, x_n) = r + 1$ if $x \notin \{x_n\}$, $\widetilde{d}(x_n, x_k) := r|1/n - 1/k|$. See also Varadarajan [**1918**, Part 2, Theorem 4].

8.10.78.° Let μ be a Radon probability measure on a completely regular space X and let $\mathcal{E}$ be some class of Borel sets that is closed with respect to finite intersections. Suppose that for every open set U and every $\varepsilon > 0$, one can find sets $E_1, \ldots, E_k \in \mathcal{E}$ such that $\bigcup_{i=1}^k E_i \subset U$ and $\mu\bigl(U \backslash \bigcup_{i=1}^k E_i\bigr) < \varepsilon$. Prove that if a sequence of Radon probability measures μ_n is such that $\lim\limits_{n\to\infty} \mu_n(E) = \mu(E)$ for all $E \in \mathcal{E}$, then the measures μ_n converge weakly to μ. Prove the analogous assertion for Baire measures and Baire sets.

Hint: observe that in the proof of Theorem 8.2.13 it suffices to represent U as the union of a sequence of sets in $\mathcal{E}$ up to a set of μ-measure zero.

8.10.79. (Wichura [**1982**]) Let (X, d) be a metric space, (Ω, P) a probability space, $\xi_n, \xi\colon \Omega \to X$ measurable mappings, and let $T_n\colon X \to X$ be Borel mappings such that for every n, the measures $P \circ (T_n \circ \xi_k)^{-1}$ converge weakly to the measure

$P\circ(T_n\circ\xi)^{-1}$ as $k\to\infty$. Suppose that the sequence $d(\xi, T_n\circ\xi)$ converges to 0 in probability and that for every $\varepsilon>0$ one has

$$\lim_{n\to\infty}\limsup_{k\to\infty} P\bigl(d(\xi_k, T_n\circ\xi_k)\ge\varepsilon\bigr)=0.$$

Prove that the measures $P\circ\xi_k^{-1}$ converge weakly to the measure $P\circ\xi^{-1}$.

8.10.80. Construct an example of a completely regular space X such that the set of all Dirac measures is not closed in $\mathcal{M}_\sigma^+(X)$.

HINT: let ω_1 be the least uncountable ordinal and let $X=[0,\omega_1)$ be equipped with the order topology. For every continuous function f on X, there exists $\tau<\omega_1$ such that f is constant on $[\tau,\omega_1)$ (Exercise 6.10.75). Let the measure μ equal 0 on all countable sets and 1 on their complements. Then μ is defined on all Baire sets. The net of Dirac measures δ_α, $\alpha<\omega_1$, converges weakly to μ. Indeed, if a continuous function f equals 1 on $[\tau,\omega_1)$, then it has the integral 1 with respect to the measure μ (because the set $[0,\tau)$ is countable) and the measures δ_α, $\alpha\ge\tau$.

8.10.81.° Let X be a compact space. Prove that the set $\mathcal{M}_\sigma^+(X)$ is countably separated if and only if $C_b(X)$ is norm separable, which, in turn, is equivalent to the metrizability of X.

HINT: use Exercise 6.10.24.

8.10.82.° Let Baire probability measures μ_n on a topological space X be given by densities ϱ_n with respect to a fixed Baire probability measure ν. Suppose that for some $p\in[1,+\infty)$ the sequence $\{\varrho_n\}$ is bounded in $L^p(\nu)$.

(i) Show that in the case $1<p<\infty$, the functions ϱ_n converge in the weak topology of the space $L^p(\nu)$ to a function $\varrho\in L^p(\nu)$ precisely when the measures μ_n converge weakly to the measure $\varrho\cdot\nu$.

(ii) Show that weak convergence of the functions ϱ_n to the function ϱ in $L^1(\nu)$ implies weak convergence of the measures μ_n to the measure $\varrho\cdot\nu$, but the converse is false.

(iii) Give an example showing that in the case where ν is Lebesgue measure on the whole line and ϱ is a probability density, weak convergence of ϱ_n to ϱ in $L^p(\nu)$, $p>1$, is not sufficient for weak convergence of the measures μ_n to $\varrho\cdot\nu$, i.e., the assumption that ν is bounded is essential in (i).

8.10.83. Suppose that bounded (possibly signed) measures μ_n on the real line are given by densities ϱ_n and converge weakly to a measure μ with a density ϱ such that one has

$$\int_{-\infty}^{+\infty}\sqrt{1+\varrho_n^2}\,dx\to\int_{-\infty}^{+\infty}\sqrt{1+\varrho^2}\,dx.$$

Prove that $\|\mu-\mu_n\|\to 0$.

HINT: see Reshetnyak [**1553**], Giaquinta, Modica, Souček [**683**, v. 2, §3.4, Proposition 1].

8.10.84.° Let X be a locally compact space and let $\{\mu_n\}$ be a sequence of Radon measures of bounded variation on X such that there exists a bounded Radon measure μ satisfying the equality

$$\lim_{n\to\infty}\int_X\varphi\,d\mu_n=\int_X\varphi\,d\mu$$

for every continuous function φ with compact support. Suppose that $\|\mu_n\|\to\|\mu\|$. Prove that the sequence $\{\mu_n\}$ is uniformly tight and converges weakly to μ.

HINT: given $\varepsilon > 0$, find a compact set K and a number n_ε such that one has $|\mu|(K) > \|\mu\| - \varepsilon$ and $\|\mu_n\| < \|\mu\| + \varepsilon$ for all $n \geq n_\varepsilon$; let f be a continuous function with compact support S containing K such that $|f| \leq 1$ and the integral of f over X is greater than $|\mu|(K) - \varepsilon$. There exists $N \geq n_\varepsilon$ such that

$$\int_X f\,d\mu_n \geq \int_X f\,d\mu - \varepsilon$$

for all $n \geq N$; then $|\mu_n|(S) \geq \|\mu_n\| - 4\varepsilon$ whenever $n \geq N$; now it is easy to verify weak convergence to μ.

8.10.85.° Let μ_n be Borel measures on $\mathbb{R}^d$ with $\sup_n \|\mu_n\| < \infty$. Assume that there exists a bounded Borel measure μ such that the characteristic functionals $\widetilde{\mu}_n$ of the measures μ_n converge pointwise to the characteristic functional $\widetilde{\mu}$ of the measure μ. Prove that

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \varphi\,d\mu_n = \int_{\mathbb{R}^d} \varphi\,d\mu$$

for every continuous function φ with compact support.

HINT: it suffices to prove the indicated equality for every function $\varphi \in C_0^\infty(\mathbb{R}^d)$; in that case it remains to observe that the Fourier transform $\widehat{\varphi}$ of the function φ is integrable and

$$\int_{\mathbb{R}^d} \varphi\,d\mu_n = (2\pi)^{d/2} \int_{\mathbb{R}^d} \widetilde{\mu_n}(y)\widehat{\varphi}(y)\,dy \to (2\pi)^{d/2} \int_{\mathbb{R}^d} \widetilde{\mu}(y)\widehat{\varphi}(y)\,dy = \int_{\mathbb{R}^d} \varphi\,d\mu$$

by the dominated convergence theorem.

8.10.86. Let (S,d) be a separable metric space and let Ω be a Hausdorff space with a Radon probability measure μ. Let $\{u_n\}$ be a sequence of measurable mappings from Ω to S and let u_∞ be a measurable mapping from Ω to S. Prove that the mappings u_n converge to u_∞ in measure if and only if the associated Young measures ν_n converge weakly to the Young measure ν_∞ generated by u_∞.

HINT: convergence in measure implies convergence of integrals for every bounded continuous function ψ on $\Omega \times S$, since the functions $\psi\bigl(x, u_n(x)\bigr)$ converge in measure to $\psi\bigl(x, u(x)\bigr)$ according to Exercise 7.14.74. Conversely, suppose we have weak convergence of the measures ν_n. Let $\psi(x,y) = \min\bigl(1, d\bigl(u_\infty(x), y\bigr)\bigr)$. Then the integral of ψ with respect to ν_∞ vanishes and the integral with respect to ν_n equals

$$\int \min\bigl(1, d(u_n, u_\infty)\bigr)\,d\mu.$$

Therefore, in order to show that $u_n \to u_\infty$ in measure, it suffices to prove that the integrals of ψ against ν_n converge to the integral of ψ against ν_∞. This convergence holds if we replace u_∞ by a continuous mapping v. In the general case, we may assume that $S = \mathbb{R}^\infty$ because S is homeomorphic to a set in $\mathbb{R}^\infty$. It remains to apply Lusin's theorem, which for every $\varepsilon > 0$ gives a set $E \subset \Omega$ with $\mu(\Omega \backslash E) < \varepsilon$ and a continuous mapping $v\colon \Omega \to S$ such that $v = u_\infty$ on E. The difference between

$$\int_{\Omega\times S} \min\Bigl(1, d\bigl(v(x), y\bigr)\Bigr)\,\nu_n(dx) = \int_\Omega \min\bigl(1, d(v, u_n)\bigr)\,d\mu$$

and

$$\int_\Omega \min\bigl(1, d(u_n, u_\infty)\bigr)\,d\mu$$

is at most 2ε and the same is true for u_∞ in place of u_n.

8.10.87. (Hartman, Marczewski [**790**]) Let $(X,\mathcal{A},\mu)$ be a probability space, let (Y,d) be a separable metric space, and let $f, f_n\colon\ X\to Y$ be μ-measurable mappings. Prove that $f_n\to f$ in measure, i.e., $\lim\limits_{n\to\infty}\mu\Bigl(d\bigl(f_n(x),f(x)\bigr)>\varepsilon\Bigr)=0$ for all $\varepsilon>0$, precisely when $\lim\limits_{n\to\infty}\mu\bigl(f_n^{-1}(E)\bigtriangleup f^{-1}(E)\bigr)=0$ for every set $E\in\mathcal{B}(Y)$ such that the boundary of E has measure zero with respect to $\mu\circ f^{-1}$.

8.10.88.° (G. Pólya) Let μ be a probability measure and let f and f_n, where $n\in\mathbb{N}$, be measurable functions such that the measures $\mu\circ f_n^{-1}$ converge weakly to the measure $\mu\circ f^{-1}$, which has no atoms. Prove that the corresponding distribution functions converge uniformly.

8.10.89.° Let μ be a probability measure and let f_n, f be μ-integrable functions such that the measures $\mu\circ f_n^{-1}$ converge weakly to the measure $\mu\circ f^{-1}$. Show that if the sequence $\{f_n\}$ is uniformly integrable, then

$$\int f_n\,d\mu\to\int f\,d\mu.$$

HINT: given $\varepsilon>0$, find $C>0$ with

$$\int_{\{|f_n|\geq C\}}|f_n|\,d\mu<\varepsilon/3,\ \int_{\{|f|\geq C\}}|f|\,d\mu<\varepsilon/3,$$

set $\varphi(t):=\operatorname{sign}(t)\min(|t|,C)$, take N such that whenever $n\geq N$, the integrals of $\varphi\circ f_n$ and $\varphi\circ f$ differ at most in $\varepsilon/3$, and then observe that

$$\int f_n\,d\mu-\int f\,d\mu=\int t\,\mu\circ f_n^{-1}(dt)-\int t\,\mu\circ f^{-1}(dt),$$

$$\int_{|t|\geq C}|t|\,\mu\circ f_n^{-1}(dt)<\varepsilon/3,\quad \int_{|t|\geq C}|t|\,\mu\circ f^{-1}(dt)<\varepsilon/3.$$

8.10.90. (Borel [**233**], Gâteaux [**672**]) Let μ_n, $n\in\mathbb{N}$, be a probability measure on $\mathbb{R}^n$ obtained by normalizing the surface measure on the sphere of radius $\sqrt{n}$ centered at the origin. Prove that the sequence of measures μ_n regarded as measures on $\mathbb{R}^\infty$ (by means of the natural embedding of $\mathbb{R}^n$ into $\mathbb{R}^\infty$) converges weakly to the countable product of the standard Gaussian measures on $\mathbb{R}^1$.

HINT: it suffices to verify weak convergence of the projections on each $\mathbb{R}^d$ with fixed d. The coordinate functions x_n on $(\mathbb{R}^\infty,\gamma)$ form a sequence of independent standard Gaussian random variables. Let $\zeta_n:=\sqrt{x_1^2+\cdots+x_n^2}$. One can verify that the image of the measure γ under the mapping $\sqrt{n}(x_1,\ldots,x_n)/\zeta_n$ is the normalized surface measure on the sphere of radius $\sqrt{n}$ in $\mathbb{R}^n$. Hence the projection of this surface measure to $\mathbb{R}^d$ coincides with the image of the measure γ under the mapping $f_n=\sqrt{n}(x_1,\ldots,x_d)/\zeta_n$ from $\mathbb{R}^\infty$ to $\mathbb{R}^d$. Letting $n\to\infty$, by the law of large numbers we have $\zeta_n/\sqrt{n}\to 1$ a.e. (see Chapter 10). Hence the measures $\gamma\circ f_n^{-1}$ on $\mathbb{R}^d$ converge weakly to the projection of the measure γ on $\mathbb{R}^d$.

8.10.91. (Hoffmann-Jørgensen [**844**]) Suppose we are given Prohorov spaces X_n and continuous mappings f_n from a completely regular space X to X_n such that if sets K_n are compact in X_n, then $\bigcap_{n=1}^\infty f_n^{-1}(K_n)$ is compact in X. Prove that X is a Prohorov space and derive from this assertions (i)–(ii) of Theorem 8.10.10.

8.10.92. Justify Example 8.6.9.

HINT: $n^{-1}\sum_{i=1}^n\delta_i(U)\to 1$ for every open neighborhood U of the point ∞, one has $\delta_\infty(U)=0$ for other open sets. Any compact set in the indicated topology is finite. Indeed, any infinite sequence $\{n_k\}$ contains an infinite subsequence $\{n_{k_i}\}$ such that the complement U of $\{n_{k_i}\}$ is open in the regarded topology. Then U and the points n_{k_i} form an open cover of $\{n_k\} \cup \{\infty\}$ that has no finite subcovers.

8.10.93. (Choquet [**353**], Fremlin, Garling, Haydon [**636**]) Let X be a metric space. Prove that every countable set in $\mathcal{M}_t^+(X)$ that is compact in the weak topology is uniformly tight.

8.10.94. Let X be a completely regular space possessing a sequence of closed subspaces X_n such that $\mathcal{M}_\sigma(X_n) = \mathcal{M}_t(X_n)$ and every function on X that is continuous on each X_n, is continuous on all of X. Suppose that all Baire subsets of X_n are Baire in X. Prove that the space $\mathcal{M}_t(X)$ is weakly sequentially complete.

HINT: as in the proof of Proposition 8.10.12, the complement of the set $Y = \bigcup_{n=1}^\infty X_n$ is discrete and all its subsets are Baire in X. One can replace the measures μ_n by their (unique) Radon extensions. All measures μ_n are purely atomic on $X\backslash Y$, and the collection of their atoms in $X\backslash Y$ is an at most countable discrete subset A of X. As in the proof of the cited proposition, $|\mu|\bigl(X\backslash(Y \cup A)\bigr) = 0$. In particular, the limit Baire measure μ is tight on $X\backslash Y$. It follows by our hypotheses that the restriction of μ is tight on every X_m (it is well-defined due to our hypothesis). Therefore, the measure μ is tight on Y, hence on X.

8.10.95.° (i) (Dembski [**428**]) Let X be a separable metric space. A class $\mathcal{D}$ of Borel sets is called determining weak convergence if, for any Borel probability measures μ_n and μ on X, the relation $\lim\limits_{n\to\infty} \mu_n(D) = \mu(D)$ for all $D \in \mathcal{D}$ with $\mu(\partial D) = 0$ yields weak convergence of μ_n to μ. Show that if $\mathcal{D}$ is a class determining weak convergence, then, for every Borel probability measure ν, the class $\mathcal{D}^\nu$ consisting of all Borel sets in $\mathcal{D}$ that have boundaries of ν-measure zero is a class determining weak convergence.

(ii) Given a Borel probability measure ν, let $\mathcal{D}_\nu$ be the class of all Borel sets that have boundaries of ν-measure zero. Show that convergence of a sequence of Borel probability measures μ_n to a Borel probability measure μ on every set in $\mathcal{D}_\nu$ yields weak convergence.

(iii) Let $\mathcal{D}$ be the class of all compact sets in $[0,1]$ with boundaries of positive Lebesgue measure. Show that convergence of a sequence of Borel probability measures μ_n on $[0,1]$ to a Borel probability measure μ on every set in $\mathcal{D}$ implies weak convergence, although $\mathcal{D}$ is not a class determining weak convergence.

HINT: (i) if $\lim\limits_{n\to\infty} \mu_n(D) = \mu(D)$ for all $D \in \mathcal{D}^\nu$ with $\mu(\partial D) = 0$, then we have $\lim\limits_{n\to\infty} (\mu_n + \nu)(D) = (\mu + \nu)(D)$ for all $D \in \mathcal{D}^\nu$ with $(\mu + \nu)(\partial D) = 0$, hence the measures $(\mu_n + \nu)/2$ converge weakly to $(\mu + \nu)/2$, which yields weak convergence of $\{\mu_n\}$ to μ. Clearly, (ii) follows from (i). (iii) Let μ_n and μ be Borel probability measures on $[0,1]$ such that $\lim\limits_{n\to\infty} \mu_n(D) = \mu(D)$ for all $D \in \mathcal{D}$. We have to show that $\lim\limits_{n\to\infty} F_{\mu_n}(t) = F_\mu(t)$ for every continuity point t of the distribution function F_μ of the measure μ. If there is $\varepsilon > 0$ such that $F_{\mu_n}(t) > F_\mu(t) + \varepsilon$ for infinitely many n, then we can find $s > t$ such that $F_\mu(s) < F_\mu(t) + \varepsilon/2$. Clearly, (t, s) contains a compact set K with boundary of positive Lebesgue measure. Then $D = [0, t] \cup K \in \mathcal{D}$ and we obtain a contradiction because $\mu_n(D) \geq F_{\mu_n}(t)$. If $F_{\mu_n}(t) < F_\mu(t) - \varepsilon$ for infinitely many n, then there is $s < t$ such that $F_\mu(s) > F_\mu(t) - \varepsilon/2$. Again we find a set $D \in \mathcal{D}$ of the form $D = [0, s] \cup K$, $K \subset (s, t)$, which gives $\mu_n(D) \leq F_{\mu_n}(t) \leq \mu(D) - \varepsilon/2$, since $\mu(D) \geq F_\mu(s)$.

8.10.96. Prove Proposition 8.9.7.

HINT: consider the map $(x_n) \mapsto \sum_{n=1}^{\infty} 2^{-n}\delta_{x_n}$; see Grömig [**740**], Koumoullis [**1044**].

8.10.97.° Let X be a completely regular space. Prove that the set $M \subset \mathcal{P}_\tau(X)$ has compact closure in the weak topology precisely when for every net of open sets U_α increasing to X, one has

$$\sup_\alpha \inf_{\mu\in M} \mu(U_\alpha) = 1.$$

In addition, this is equivalent to the following property: for every net of bounded continuous functions f_α on X pointwise decreasing to zero, one has

$$\inf_\alpha \sup_{\mu\in M} \int_X f_\alpha\, d\mu = 0.$$

HINT: the necessity is easily verified. The sufficiency follows from the compactness of balls in $C_b(X)^*$ in the weak* topology and Theorem 7.10.7.

8.10.98. (Pachl [**1416**]) Let X be a complete metric space and let $U_b(X)$ be the set of all bounded uniformly continuous functions on X. (i) Prove that the space $\mathcal{M}_r(X)$ of all Radon measures on X is sequentially complete in the topology $\sigma\bigl(\mathcal{M}_r(X), U_b(X)\bigr)$.

(ii) Prove that for every bounded set $M \subset \mathcal{M}_r(X)$, the following conditions are equivalent: (a) M has compact closure in the Kantorovich–Rubinshtein norm $\|\cdot\|_0$; (b) the closure of M in the topology $\sigma\bigl(\mathcal{M}_r(X), U_b(X)\bigr)$ is countably compact.

8.10.99. (Haydon [**801**]) Show that the Stone–Čech compactification of $\mathbb{N}$ contains a set Z such that $\mathcal{P}_t(Z) \neq \mathcal{P}_\tau(Z)$, but every weakly compact set of measures in $\mathcal{P}_t(Z)$ is uniformly tight, i.e., Z is Prohorov.

8.10.100. (Lange [**1107**]) Let X be a Polish space and $\mu \in \mathcal{P}_r(X)$.

(i) Prove that the sets $\{\nu \in \mathcal{P}_r(X)\colon\ \nu \ll \mu\}$ and $\{\nu \in \mathcal{P}_r(X)\colon\ \nu \sim \mu\}$ are Borel in $\mathcal{P}_r(X)$ with the weak topology.

(ii) If X is locally compact, then the following subsets of $\mathcal{P}_r(X)$ are Borel as well: (a) measures with compact supports, (b) measures with compact connected supports, (c) measures with a given closed support, (d) measures with supports contained in a given closed set, (e) measures with supports containing a given closed set, (f) measures with supports without inner points, (g) measures with supports without isolated points, (h) measures with supports consisting of at most k points. However, this may be false for a non-locally compact space.

(iii) If $X = \mathbb{R}^n$, then the set of all probability measures with convex supports and the set of all probability measures having the finite moment of a fixed order p are Borel.

8.10.101. Suppose we are given a sequence of measurable spaces $(X_n, \mathcal{A}_n)$ and for every n, there are two probability measures P_n and Q_n on $\mathcal{A}_n$. The sequences $\{Q_n\}$ and $\{P_n\}$ are called contigual (or mutually contigual) if for all $A_n \in \mathcal{A}_n$, the condition $P_n(A_n) \to 0$ is equivalent to $Q_n(A_n) \to 0$. Let

$$\lambda_n = (P_n + Q_n)/2,\ f_n = dP_n/d\lambda_n,\ g_n = dQ_n/d\lambda_n,$$

and $\Lambda_n = \log(g_n/f_n)$ if $f_n g_n > 0$ and $\Lambda_n = 0$ otherwise. Prove that the following conditions are equivalent: (i) $\{Q_n\}$ and $\{P_n\}$ are contigual, (ii) $\{P_n \circ \Lambda_n^{-1}\}$ is uniformly tight, (iii) $\{Q_n \circ \Lambda_n^{-1}\}$ is uniformly tight.

Hint: see Roussas [**1616**, Ch. 1].

8.10.102. Let X be a separable Banach space and let μ be a Borel probability measure on X. For every compact set $K \subset X$, we define the concentration function $C_\mu(K)$ by the formula $C_\mu(K) = \sup_{x\in X} \mu(K + x)$. Prove that the following conditions are equivalent for every sequence of Borel probability measures μ_n on X:

(i) $\sup\limits_{K\in\mathcal{K}} \inf\limits_n C_{\mu_n}(K) = 1$, where $\mathcal{K}$ is the family of all compact sets in X,

(ii) every subsequence in $\{\mu_n\}$ contains a further subsequence $\{\nu_n\}$ such that for some vectors $x_n \in X$ the sequence of measures $\nu_n(\,\cdot\, + x_n)$ is uniformly tight.

Hint: see Hengartner, Theodorescu [**810**, Ch. 5], where one can find additional information about concentration functions.

8.10.103. Justify Example 8.10.50.

8.10.104.$^\circ$ (Weyl [**1976**]) Let $\{x_n\} \subset [0, 1)$. Prove that the following conditions are equivalent: (i) the sequence $\{x_n\}$ is uniformly distributed with respect to Lebesgue measure on $[0, 1)$,

(ii) for all $[\alpha, \beta] \subset [0, 1]$, one has $\lim\limits_{N\to\infty} N^{-1}F(N, \alpha, \beta) = \beta-\alpha$, where $F(N, \alpha, \beta)$ is the number of all $n \le N$ such that $\alpha \le x_n < \beta$,

(iii) $\lim\limits_{N\to\infty} \sup_{\alpha,\beta} |N^{-1}F(N, \alpha, \beta) - (\beta - \alpha)| = 0$,

(iv) for every integer $m \ne 0$, one has $\lim\limits_{N\to\infty} N^{-1} \sum_{n=1}^N \exp(2\pi i m x_n) = 0$.

Hint: the equivalence of (i)–(iii) is easily seen from the general properties of weak convergence; (iv) follows from (i); Finally, (iv) yields (i), since every measure that is a limit point of the sequence of measures $N^{-1} \sum_{n=1}^N \delta_{x_n}$ in the weak topology assigns the same integral to any finite linear combination of the functions $\exp(i2\pi mx)$ as Lebesgue measure does, hence equals Lebesgue measure.

8.10.105. (de Bruijn, Post [**267**]) Let f be a function on $[0, 1]$ such that for every uniformly distributed sequence $\{x_n\} \subset [0, 1]$, the limit $\lim\limits_{N\to\infty} N^{-1} \sum_{n=1}^N f(x_n)$ exists and is finite. Prove that the function f is Riemann integrable in the proper sense.

8.10.106. (Losert [**1188**]) Let X and Y be compact metric spaces.

(i) Let μ be a Radon probability measure on $X\times Y$ and let $\pi_X\colon X\times Y \to X$ be the natural projection. Show that if a sequence $\{x_n\} \subset X$ is uniformly distributed with respect to $\mu \circ \pi_X^{-1}$, then Y contains a sequence $\{y_n\}$ such that the sequence (x_n, y_n) is uniformly distributed with respect to μ.

(ii) Construct an example showing that (i) may fail for non-metrizable compact spaces even if μ is the product of Radon measures on X and Y.

(iii) Let $\pi\colon X \to Y$ be a continuous surjection, let d be the metric of Y, and let μ be a Radon probability measure on X. Set $\nu = \mu \circ \pi^{-1}$. Show that for every sequence $\{y_n\}$ that is uniformly distributed with respect to ν, there exists a sequence $\{x_n\}$ uniformly distributed with respect to μ such that $\lim\limits_{n\to\infty} d\bigl(\pi(x_n), y_n\bigr) = 0$.

(iv) Let $\pi\colon X \to Y$ be a continuous surjection, let μ be a Radon probability measure on X, and let $\nu = \mu\circ\pi^{-1}$. Show that the following conditions are equivalent: (a) for every sequence $\{y_n\}$ that is uniformly distributed with respect to ν, there exists a sequence $\{x_n\}$ uniformly distributed with respect to μ such that $y_n = \pi(x_n)$, (b) the set of all points x possessing neighborhoods whose images under π are not open, has μ-measure zero.

8.10.107. (Losert [**1188**]) Assuming the continuum hypothesis, show that there is a Radon probability measure μ on $[0,1]^{\mathfrak{c}}$ such that there exist sequences that are uniformly distributed with respect to μ, but such a sequence cannot be chosen in the topological support of μ.

HINT: $[0,1]^{\mathfrak{c}}$ contains a compact set homeomorphic to $\beta\mathbb{N}$; there is a Radon measure μ on $\beta\mathbb{N}$ without uniformly distributed sequences (Example 8.10.54), but this measure has uniformly distributed sequences in X by Proposition 8.10.55.

8.10.108. (Losert [**1188**]) Show that $\{0,1\}^{\mathfrak{c}}$ contains an everywhere dense set M such that M contains no uniformly distributed sequence with respect to the measure μ that is the power of the measure equal $1/2$ at the points 0 and 1.

8.10.109. (Hlawka [**835**]) Let X be a completely regular space such that there exists a countable family of functions $f_j \in C_b(X)$ with the property that if for a sequence of Radon probability measures μ_n and a Radon probability measure μ, one has

$$\lim_{n\to\infty} \int_X f_j\, d\mu_n = \int_X f_j\, d\mu$$

for all j, then the sequence $\{\mu_n\}$ converges weakly to μ. Let μ^∞ be the countable power of μ. Prove that μ^∞-almost every sequence in X^∞ is uniformly distributed with respect to μ.

HINT: by the law of large numbers (see Chapter 10), for every j, the set of sequences (x_n) such that the arithmetic means $N^{-1}\sum_{n=1}^N f_j(x_n)$ converge to the integral of f_j with respect to the measure μ has full μ^∞-measure.

8.10.110. (Kawabe [**966**]) Let X be a Hausdorff space, let Y be a completely regular space, and let the space $\mathcal{P}_\tau(Y)$ be equipped with the weak topology.

(i) Prove that a mapping $\lambda\colon X \to \mathcal{P}_\tau(Y)$, $x \mapsto \lambda(x,\,\cdot\,)$, is continuous if and only if for every open set $U \subset X\times Y$, the function $x \mapsto \lambda(x, U_x)$ is upper semicontinuous on X, where, as usual, $U_x = \{y \in Y\colon\ (x,y) \in U\}$.

(ii) Show that if the mapping λ in (i) is continuous, then for every $B \in \mathcal{B}(X\times Y)$, the function $x \mapsto \lambda(x, B_x)$ is Borel on X. Hence for every Borel measure μ on X we obtain a Borel measure

$$\mu\circ\lambda(B) := \int_X \lambda(x, B_x)\,\mu(dx), \quad B \in \mathcal{B}(X\times Y).$$

(iii) Show that if the measure μ in (ii) is τ-additive, then so is $\mu\circ\lambda$.

(iv) Let X be a k-space (e.g., a locally compact or metrizable space), let Y be a compact space, $f \in C_b(X\times Y)$, $\lambda \in C\bigl(X, \mathcal{P}_\tau(Y)\bigr)$. Prove that the function

$$x \mapsto \int_Y f(x,y)\,\lambda(x,dy)$$

is continuous on X.

(v) Let X be a completely regular k-space (for example, locally compact or metrizable). Suppose we are given a net of mappings $\lambda_\alpha\colon X \to \mathcal{P}_\tau(Y)$ that are pointwise equicontinuous on every compact set in X, and for every $x \in X$, the net of measures $\lambda_\alpha(x,\,\cdot\,)$ is uniformly tight and converges weakly to $\lambda(x,\,\cdot\,)$ for some continuous mapping $\lambda\colon X \to \mathcal{P}_\tau(Y)$. Prove that if a net of measures $\mu_\alpha \in \mathcal{P}_\tau(X)$ is uniformly tight and converges weakly to a measure $\mu \in \mathcal{P}_\tau(X)$, then the net of measures $\mu_\alpha\circ\lambda_\alpha$ converges weakly to the measure $\mu\circ\lambda$.

(vi) Let X and Y be the same as in (iv), let $P \subset \mathcal{P}_\tau(X)$ be a uniformly tight family, and let a family of mappings $Q \subset C\bigl(X, \mathcal{P}_\tau(Y)\bigr)$ be pointwise equicontinuous on every compact set in X. Assume that for every $x \in X$, the family of measures $\lambda(x,\,\cdot\,) := Q(x)$ on Y is uniformly tight. Prove that for every net of measures $\mu_\alpha \circ \lambda_\alpha$, where $\mu_\alpha \in P$, $\lambda_\alpha \in Q$, there exist a measure $\mu \in \mathcal{P}_\tau(X)$, a mapping $\lambda \in C\bigl(X, \mathcal{P}_\tau(Y)\bigr)$, and a subnet $\{\mu_{\alpha'} \circ \lambda_{\alpha'}\}$ in $\{\mu_\alpha \circ \lambda_\alpha\}$ such that one has weak convergence $\mu_{\alpha'} \Rightarrow \mu$, $\lambda_{\alpha'}(x,\,\cdot\,) \Rightarrow \lambda(x,\,\cdot\,)$ for every $x \in X$ and $\mu_{\alpha'} \circ \lambda_{\alpha'} \Rightarrow \mu \circ \lambda$. In particular, the set $P \circ Q := \{\nu \circ \zeta\colon\ \nu \in P, \zeta \in Q\}$ is relatively weakly compact in $\mathcal{P}_\tau(X \times Y)$. Show also that if, in addition, Y is Prohorov, then the family of measures $P \circ Q$ is uniformly tight.

8.10.111. Prove Theorem 8.10.45 for measures on a finite set X.

HINT: we show that $\widehat{W}(\mu,\nu) = W(\mu,\nu)$. One has $\widehat{W}(\mu,\nu) \le W(\mu,\nu)$. Let L be the linear space of all functions of the form $\varphi(x,y) = f(x) + g(y)$ on $X \times X$. We consider the functional

$$l(\varphi) = \int_X f\,d\mu + \int_X g\,d\nu$$

on L. It is easy to see that l is well-defined. The set

$$U = \bigl\{\varphi \in C(X \times X)\colon\ \varphi(x,y) < d(x,y)\bigr\}$$

is convex and open in $C(X \times X)$ and l is bounded on $U \cap L$. By the Hahn–Banach theorem l extends to a linear functional l_0 on $C(X \times X)$ with $\sup_U l_0 = \sup_{U \cap L} l$. In addition, one has $l_0(u) \ge 0$ whenever $u \ge 0$, since $d - 1 - cu \in U$ for all $c > 0$ and $\sup_{c>0} l(d - 1 - cu) < \infty$. Hence there exists a nonnegative measure λ on $X \times X$ representing l_0. Since $l_0 = l$ on L, one has

$$\int f(x)\,\lambda(dx,dy) = l(f) = \int f(x)\,\mu(dx),$$

$$\int g(y)\,\lambda(dx,dy) = l(g) = \int g(y)\,\nu(dy),$$

i.e., $\lambda \in M(\mu,\nu)$. It is easy to see that

$$\widehat{W}(\mu,\nu) = \int d(x,y)\,\lambda(dx,dy).$$

8.10.112.° Let μ_n be bounded measures on a σ-algebra $\mathcal{A}$ and let $E_k \in \mathcal{A}$ be disjoint sets such that $\lim\limits_{n\to\infty} \mu_n(E_k) = 0$ for every k and $\inf_n |\mu_n(E_n)| > 0$. Prove that there exists a sequence $\{n_j\}$ with

$$\inf_{n \in \{n_j\}} \Bigl|\mu_n\Bigl(\bigcup_{j=1}^{\infty} E_{n_j}\Bigr)\Bigr| > 0.$$

HINT: let $\inf_n |\mu_n(E_n)| = \delta$. It suffices to find a sequence $\{n_j\}$ with

$$\sum_{i=1}^{j-1} |\mu_{n_j}(E_{n_i})| < \delta/3, \qquad \sum_{n=n_j}^{\infty} |\mu_{n_{j-1}}(E_n)| < \delta/3,$$

which will give $\bigl|\mu_{n_j}\bigl(\bigcup_{i=1}^{\infty} E_{n_i}\bigr)\bigr| > \delta/3$. Letting $n_0 = 1$, we construct n_j inductively. If $n_1, \dots, n_j$ are already found, we find $n'_{j+1} \ge n_j + 1$ with $\sum_{i=1}^{j} |\mu_n(E_{n_i})| < \delta/3$ for all $n \ge n'_{j+1}$. Next we find $n_{j+1} > n'_{j+1}$ with $\sum_{n=n_{j+1}}^{\infty} |\mu_{n_j}(E_n)| < \delta/3$.

8.10.113. Construct a sequence of Borel probability measures on a Hausdorff space that converges on every open set, but does not converge on some Borel set.

HINT: see Pfanzagl [**1442**, Example 2].

8.10.114.° (i) Prove that if a set of Radon measures on a Hausdorff space is compact in the topology of convergence on Borel sets, then it is uniformly tight.

(ii) Prove that a set M of Radon measures on a Hausdorff space is relatively compact in the space of all Radon measures on X with the topology of convergence on Borel sets precisely when M is bounded and uniformly tight and for every compact set K and every $\varepsilon > 0$ there exists an open set $U \supset K$ such that $|\mu|(U\backslash K) < \varepsilon$ for all $\mu \in M$.

(iii) Prove that a sequence of Radon measures μ_n on a Hausdorff space X converges to a Radon measure μ on every Borel set precisely when it is uniformly tight and $\lim\limits_{n\to\infty} \mu_n(K) = \mu(K)$ for every compact set K.

HINT: (i) if a set M is compact in the indicated topology, then according to §4.7(v), there exists a Radon probability measure μ_0 such that all measures in M are uniformly absolutely continuous with respect to μ_0. The necessity of the conditions mentioned in (ii) follows from (i) and the proof of Theorem 8.10.58. The sufficiency reduces to the case of a compact space due to the uniform tightness, and also follows in that case from the proof of the cited theorem. (iii) The necessity of the indicated conditions is clear. The sufficiency follows by Theorem 8.10.56 applied to the restrictions of the considered measures to compact sets K_j chosen such that $|\mu|(X\backslash K_j) < 2^{-j}$ for all $\mu \in M$. For every compact set $K \subset K_j$, one has convergence on the set $K_j\backslash K$, and every set $U \subset K_j$ that is open in the induced topology has such a form.

8.10.115.° Let μ_n be convex Radon probability measures on a locally convex space X (see §7.14(xvi)) convergent weakly to a Radon measure μ. Prove that μ is convex as well.

HINT: apply Lemma 7.14.54, reduce the assertion to the case of $\mathbb{R}^n$, consider open sets A and B with boundaries of μ-measure zero.

8.10.116. (Y. Peres) Let the spaces $\mathcal{P}([0,1])$ and $\mathcal{P}([0,1]^2)$ of all Borel probability measures on $[0,1]$ and $[0,1]^2$ be equipped with the topology τ_s of convergence on all Borel sets. Show that the mapping $\mu \mapsto \mu\otimes\mu$ is sequentially continuous, but is not continuous at the point λ, where λ is Lebesgue measure (a question about this was raised by F. Götze).

HINT: the sequential continuity is obvious from Fubini's theorem and the dominated convergence theorem. In order to show the discontinuity at the point λ, we take the set $A = \bigl\{(x,y) \in [0,1]^2\colon\ x - y \in \mathbb{Q}\bigr\}$. This set is Borel and $\lambda\otimes\lambda(A) = 0$. We observe that every neighborhood of the point λ in the topology τ_s contains a measure $\nu \in \mathcal{P}([0,1])$ such that $\nu\otimes\nu(A) = 1$. To this end, it suffices to show that for every finite partition of $[0,1]$ into Borel parts B_i, there exist points $x_i \in B_i$ such that $x_i - x_j \in \mathbb{Q}$ if $i \neq j$. Then we take $\nu := \sum_{i=1}^n \lambda(B_i)\delta_{x_i}$. The required points indeed exist, since the set $B := \prod_{i=1}^n B_i$ in $\mathbb{R}^n$ has positive measure, hence $B - B$ contains a neighborhood, in particular, $B - B$ contains a point with rational coordinates.

8.10.117. (Schief [**1670**]) (i) Construct an example of locally compact spaces X and Y and a continuous open surjection $f\colon X \to Y$ such that the mapping $\widehat{f}\colon \mathcal{P}(X) \to \mathcal{P}(Y)$ is open, but not surjective.

(ii) Assuming the continuum hypothesis, construct a Hausdorff space X and a continuous open surjection $f\colon X \to \mathbb{R}^1$ such that the mapping $\widehat{f}\colon \mathcal{P}(X) \to \mathcal{P}(\mathbb{R}^1)$ is not surjective. Show also that $\widehat{f}$ may be surjective but not open.

8.10.118. (Schief [**1667**], [**1668**]) Let X be a Hausdorff space. Show that the mapping $(\mu,\nu)\mapsto\mu-\nu$ is continuous in the A-topology on the set of all pairs of nonnegative Borel measures (μ,ν) on X with $\mu-\nu\geq 0$. Prove that the mapping $(\mu,\nu)\mapsto\mu+\nu$ on the set of all nonnegative Borel measures is open in the A-topology.

8.10.119. (Ressel [**1556**]) Let X and Y be Hausdorff spaces and let $\{\mu_t\}_{t\in T}$ be a net of Radon probability measures on $X\times Y$ such that their projections on X converge weakly to a Radon measure ν, and their projections on Y converge weakly to Dirac's measure δ_a at some point $a\in Y$. Show that the net $\{\mu_t\}$ converges weakly to the Radon extension of the measure $\nu\otimes\delta_a$ to $\mathcal{B}(X\times Y)$. Prove the analogous assertion for τ-additive measures.

HINT: let $U\subset X\times Y$ be an open set whose projection on Y contains a. Given $\varepsilon>0$, one can find a compact set $K\subset X$ such that $K\times a\subset U$ and the estimate $\nu\otimes\delta_a(K\times a)>\nu\otimes\delta_a(U)-\varepsilon$ holds. There exist sets V and W that are open, respectively, in X and Y with $K\times a\subset V\times W\subset U$. In view of weak convergence of projections, there exists t_1 such that $\mu_t(X\times W)>1-\varepsilon$ whenever $t>t_1$, hence

$$\mu_t(V\times W)\geq\mu_t(V\times Y)-\mu_t\bigl(X\times(Y\backslash W)\bigr)>\mu_t(V\times Y)-\varepsilon.$$

There is $t_2>t_1$ such that $\mu_t(V\times Y)>\nu(V)-\varepsilon$ for all $t>t_2$. Then we obtain $\mu_t(U)\geq\mu_t(V\times W)>\nu(V)-2\varepsilon>\nu\otimes\delta_a(U)-3\varepsilon$.

8.10.120.° (Slutsky [**1743**]) Let $(\Omega,\mathcal{A},P)$ be a probability space, let E be a separable Banach space, and let ξ_n, ξ, and η_n be $\bigl(\mathcal{A},\mathcal{B}(E)\bigr)$-measurable mappings. Suppose that the measures $P\circ\xi_n^{-1}$ converge weakly to $P\circ\xi^{-1}$ and $\eta_n\to 0$ a.e. Show that the measures $P\circ(\xi_n+\eta_n)^{-1}$ converge weakly to $P\circ\xi^{-1}$ as well.

HINT: apply Egoroff's theorem to $\{\eta_n\}$.

8.10.121. (Dellacherie [**425**, Ch. 4, Theorem 31]) Let X be a Polish space, let M be a Souslin subset of the space of Borel probability measures $\mathcal{P}(X)$ with the weak topology, and let $A\subset X$ be a Souslin set such that $\mu(A)=0$ for all $\mu\in M$. Prove that there exists a Borel set $B\subset X$ such that $A\subset B$ and $\mu(B)=0$ for all measures $\mu\in M$.

8.10.122. Let X be a Polish space and let M be a compact subset of the space of Borel probability measures $\mathcal{P}(X)$ with the weak topology.

(i) (Dellacherie [**425**]) Prove that the function $I(E):=\sup_{\mu\in M}\mu^*(E)$ is a Choquet capacity and derive from this that for every Souslin set A and every $\varepsilon>0$, there exists a compact set $K_\varepsilon\subset A$ with $I(K_\varepsilon)>I(A)-\varepsilon$.

(ii) (Choquet [**353**]) Let S be a compact or σ-compact set in X such that $\mu(S)=0$ for all $\mu\in M$. Prove that for every $\varepsilon>0$, there exists an open set $U\supset S$ such that $\mu(U)<\varepsilon$ for all $\mu\in M$.

(iii) (Choquet [**353**]) Show that under the continuum hypothesis there exists a function $f\colon[0,1]\to[0,1]$ such that its graph S is measurable with respect to every Borel measure on $[0,1]^2$ and every atomless measure vanishes on S. Prove that the set M of all Borel probability measures on $[0,1]^2$ having Lebesgue measure as the projection to the first factor is compact, but for M and S assertion (ii) fails.

(iv) (Choquet [**353**]) Show that on an uncountable power of $[0,1]$, there exist a sequence of Radon probability measures μ_n weakly convergent to Dirac's measure $\mu_0=\delta_0$ and a G_δ-set S such that assertion (ii) fails for $M=\{\mu_n\}_{n\geq 0}$.

HINT: (i) for every compact set K, the function $\mu\mapsto\mu(K)$ is upper semicontinuous. This gives $I(K)=\lim\limits_{n\to\infty}I(K_n)$ for any sequence of compact sets K_n decreasing to K. If sets E_n are increasing to E, then the equality $I(E) = \lim\limits_{n\to\infty} I(E_n)$ is easily verified by using Proposition 1.5.12. (ii) In the case of compact S the assertion is easily deduced from (i) (or is proved directly by a similar reasoning); if $S = \bigcup_{n=1}^{\infty} S_n$, where S_n are compact sets, then one can take sets U_n corresponding to S_n and $\varepsilon 2^{-n}$, and let $U = \bigcup_{n=1}^{\infty} U_n$.

8.10.123.° Show that any uniformly tight set of Radon probability measures on a Hausdorff space X has compact closure in the A-topology.

HINT: suppose we have a uniformly tight sequence of Radon measures μ_j on X. We may assume that $X = \bigcup_{n=1}^{\infty} K_n$, where the sets K_n are compact, $K_n \subset K_{n+1}$ and $\mu_j(K_n) \geq 1-1/n$ for all n, j. Passing to a subsequence, we may assume that for every n the numbers $\mu_j(K_n)$ converge. Therefore, if $\{\mu_j\}$ has a subnet of measures whose restrictions to some K_n converge weakly, then we have weak convergence of their restrictions to K_{n-1}. Hence there exist Radon measures ν_n on K_n such that $\nu_n|_{K_{n-1}} = \nu_{n-1}$ and ν_n is a limit point of the sequence of measures $\mu_j|_{K_n}$ on K_n. One has $\nu_n(K_n) \geq 1 - 1/n$ and the measures ν_n converge in the variation norm to a Radon probability measure ν that is a limit point for $\{\mu_j\}$ in the A-topology.

8.10.124. (Grothendieck [**745**, p. 229]) Let K be a compact space and let $\mathcal{M} := \mathcal{M}_r(K) = C(K)^*$. Suppose a set $M \subset \mathcal{M}$ has compact closure in the Mackey topology $\tau\bigl(\mathcal{M}, C(K)\bigr)$. Show that M has compact closure in the topology $\sigma(\mathcal{M}, \mathcal{M}^*)$ as well.

HINT: by the Eberlein–Šmulian theorem and Theorems 8.10.58 and 4.7.25, it suffices to show that $\lim\limits_{n\to\infty} \mu_n(U_n) = 0$ for every sequence of measures $\mu_n \in M$ and every sequence of disjoint open sets $U_n \subset K$. If this is not true, then there exist functions $f_n \in C(K)$ such that $|f_n| \leq 1$, $f_n = 0$ outside U_n and the integral of f_n against μ_n is greater than some $\varepsilon > 0$. The sequence $\{f_n\}$ converges to zero pointwise, hence in the weak topology of $C(K)$. It is readily verified that its closed convex envelope is weakly compact. This contradicts the compactness of the closure of M in the topology of uniform convergence on convex weakly compact sets.

8.10.125. (Kallenberg [**939**]) Let $(X, \mathcal{A})$ be a measurable space and let $\mathcal{P}(\mathcal{A})$ be the set of all probability measures on $\mathcal{A}$ equipped with the σ-algebra $\mathcal{F}$ generated by the functions $\mu \mapsto \mu(A)$, $A \in \mathcal{A}$. Given a sequence of $\mathcal{A}\otimes\mathcal{F}$-measurable functions f_n on $X\times\mathcal{P}(\mathcal{A})$, denote by Λ the set of all measures $\mu \in \mathcal{P}(\mathcal{A})$ such that the sequence of functions $x \mapsto f_n(x, \mu)$ converges in measure μ. Prove that $\Lambda \in \mathcal{F}$.

HINT: suppose first that $|f_n| \leq 1$; observe that for any fixed n, k, m, the set of all μ with $\|f_n(\,\cdot\,, \mu) - f_k(\,\cdot\,, \mu)\|_{L^1(\mu)} < m^{-1}$ belongs to $\mathcal{F}$. The general case reduces easily to the considered one.

8.10.126. Let X and Y be Polish spaces, let $A \subset X\times Y$ be a Souslin set, and let $A_x := \{y \in Y\colon\ (x, y) \in A\}$. Prove that $\{(\mu, x, \alpha) \in \mathcal{P}_r(Y)\times X\times[0, 1]\colon\ \mu(A_x) > \alpha\}$ is a Souslin set, provided that $\mathcal{P}_r(Y)$ is equipped with the weak topology.

HINT: see Kechris [**968**, Theorem 29.26]

8.10.127.° Let M be a uniformly tight family of Radon measures on a Fréchet space X. Show that there exists a reflexive separable Banach space E continuously embedded into X such that all measures from M are concentrated on E and form there a uniformly tight family.

HINT: in the proof of Theorem 7.12.4 choose K_n common for all measures in M.

8.10.128. (Dall'Aglio [**396**, p. 42], Vallander [**1914**]) Show that the Kantorovich–Rubinshtein distance between two probability measures μ and ν on the real line with the distribution functions Φ_μ and Φ_ν equals $\|\Phi_\mu - \Phi_\nu\|_{L^1(\mathbb{R}^1)}$.

8.10.129. (Hoffmann-Jørgensen [**844**]) Let X be a completely regular space such that $\mathcal{M}_\sigma(X) = \mathcal{M}_t(X)$. Then $\mathcal{M}_t(X)$ with the Mackey topology is complete.

HINT: in place of Lemma 1 in [**844**] use Theorem 7.10.1.

8.10.130. Let X be a noncompact complete metric space. Show that the weak topology on the ball $U_1 := \{\mu \in \mathcal{M}_r(X)\colon\ \|\mu\| \le 1\}$ is not metrizable.

HINT: there exists a sequence of points $x_n \in X$ whose mutual distances are separated from zero, hence it suffices to consider the case $X = \mathbb{N}$. Then we have $\mathcal{M}_r(X) = l^1$. The unit ball is not metrizable in the weak topology because otherwise the weak topology on it would coincide with the norm topology due to the fact that every weakly convergent sequence in l^1 is norm convergent.

8.10.131. Suppose a sequence of Baire probability measures μ_n on a topological space X converges weakly to a Baire probability measure μ and $\mu_n = f_n \cdot \nu$, where ν is some Baire probability measure. Let

$$\sup_n \int \Psi \circ f_n \, d\nu \le C < \infty,$$

where Ψ is a convex function on $[0, +\infty)$ with $\lim\limits_{t\to+\infty} \Psi(t)/t = +\infty$. Show that $\mu \ll \nu$ and

$$\int \Psi \circ f \, d\nu \le C, \quad \text{where } f = d\mu/d\nu.$$

HINT: by the Komlós theorem one can find a subsequence $\{f_{n_k}\}$ such that the functions $g_k := (f_{n_1} + \cdots + f_{n_k})/k$ converge a.e. to some function f. Then

$$\sup_k \int \Psi \circ g_k \, d\nu \le C.$$

Hence $g_k \to f$ in $L^1(\nu)$ and by Fatou's theorem

$$\int \Psi \circ f \, d\nu \le C.$$

The measures $g_k \cdot \nu$ converge to $f \cdot \nu$ in variation, hence weakly. Since they converge weakly to μ, one has $\mu = f \cdot \nu$.

8.10.132. Let E be a G_δ-set in a topological space X. Show that $\mathcal{P}_r(E)$ is a G_δ-set in $\mathcal{P}_r(X)$ with the weak topology.

HINT: we can identify $\mathcal{P}_r(E)$ with the set P_E in $\mathcal{P}_r(X)$ consisting of the measures vanishing on $X\backslash E$ because the natural mapping of $\mathcal{P}_r(E)$ onto P_E is a homeomorphism. If E is open, then $\mathcal{P}_r(X)\backslash\mathcal{P}_r(E) = \bigcup_{k=1}^\infty M_k$, where M_k is defined by $M_k := \{\mu \in \mathcal{P}_r(X)\colon\ \mu(X\backslash E) \ge 1/k\}$. The set $X\backslash E$ is closed, hence M_k is closed as well. Therefore, $\mathcal{P}_r(E)$ is a G_δ set. If $E = \bigcap_{k=1}^\infty E_k$, where the sets E_k are open, then $\mathcal{P}_r(E) = \bigcap_{k=1}^\infty \mathcal{P}_r(E_k)$.

8.10.133. Show that Rao's theorem 8.2.18 does not extend to uniformly bounded nets of signed measures even if Γ is a uniformly Lipschitzian and uniformly bounded family.

HINT: since the weak topology on the unit ball U in l^1 is weaker than the norm topology, there exists a net $\{\mu_\alpha\} \subset U$ that weakly converges to zero, but is not norm convergent. Let us regard μ_α as measures on $\mathbb{N}$. The set Γ of all functions f on $\mathbb{N}$ with $\sup|f| \le 1$ is uniformly Lipschitzian with constant 2 and

$$\|\mu_\alpha\| = \sup\Bigl\{\int f\,d\mu_\alpha\colon\ f\in\Gamma\Bigr\}.$$

8.10.134. Suppose a sequence of Baire measures μ_n on a completely regular space X converges weakly to a tight Baire measure μ and, in addition, is uniformly tight. Let a family $\Gamma\subset C_b(X)$ be uniformly bounded and pointwise equicontinuous. Show that

$$\lim_{n\to\infty}\sup_{f\in\Gamma}\Bigl|\int f\,d(\mu_n-\mu)\Bigr| = 0.$$

HINT: we may assume that $|f|\le 1$ for all $f\in\Gamma$ and $\|\mu_n\|\le 1$. Suppose that for some $\varepsilon>0$ and some sequence $\{f_n\}\subset\Gamma$ we have

$$\Bigl|\int f_n\,d(\mu_n-\mu)\Bigr| > \varepsilon.$$

Let us find a compact set K such that $|\mu|(X\backslash K)+|\mu_n|(X\backslash K)<\varepsilon/4$ for all n. By the Ascoli–Arzela theorem (see Dunford, Schwartz [**503**, Theorem IV.6.7]) the sequence $\{f_n\}$ contains a subsequence that converges uniformly on K to some function f. We may assume that the whole sequence $\{f_n\}$ has this property. There is a function $g\in C_b(X)$ with $g|_K = f|_K$ and $|g|\le 1$. For all sufficiently large n we obtain

$$\sup_{x\in K}|g(x)-f_n(x)|\le\varepsilon/4 \quad\text{and}\quad \Bigl|\int g\,d(\mu_n-\mu)\Bigr|\le\varepsilon/4,$$

which leads to a contradiction.

8.10.135. (i) (A.N. Kolmogorov, see Glivenko [**699**, p. 157]) Prove that a sequence of Borel measures μ_n on a closed interval $[a,b]$ converges weakly to a Borel measure μ if and only if:

(1) the variations of the measures μ_n are uniformly bounded,

(2) $\mu([a,b]) = \lim\limits_{n\to\infty}\mu_n([a,b])$,

(3) for the corresponding distribution functions one has

$$\lim_{n\to\infty}\int_a^b |F_{\mu_n}(t)-F_\mu(t)|\,dt = 0.$$

(ii) Observe that (1) and (3) yield $\|F_{\mu_n}-F_\mu\|_{L^p[a,b]}\to 0$ for any $p\in[1,+\infty)$.

(iii) Prove analogous assertions for the cube $[a,b]^d$ in $\mathbb{R}^d$, where the distribution functions are defined by $F_{\mu_n}(t_1,\ldots,t_d) := \mu_n\bigl([a,t_1)\times\cdots\times[a,t_d)\bigr)$ for all $(t_1,\ldots,t_d)\in[a,b]^d$ and similarly for μ.

HINT: weak convergence yields conditions (1) and (2); condition (3) follows by the uniform boundedness of F_{μ_n} and Proposition 8.1.8. Let condition (1) be fulfilled. Weak convergence will follow from convergence of the integrals of each smooth function f against μ_n to the integral of f against μ. Due to the integration by parts formula and the equality $F_\mu(b+) = \lim\limits_{n\to\infty}F_{\mu_n}(b+)$ it remains to observe that the integral of $f'(F_\mu - F_{\mu_n})$ over $[a,b]$ tends to zero. Claim (ii) is trivial.

Let us give an alternative reasoning, which can be easily extended to the multidimensional case. Let $[a,b]=[0,2\pi]$, $\varphi_k(t)=\exp(ikt)$, $k\in\mathbb{Z}$. Set $f_{n,k} := (\varphi_k, F_{\mu_n})_{L^2[0,2\pi]}$, $f_k := (\varphi_k, F_\mu)_{L^2[0,2\pi]}$. If $\sup_n\|\mu_n\|\le C<\infty$, then

$$\Bigl|\int_0^{2\pi}\varphi_k\,d\mu_n\Bigr|\le C.$$

By the integration by parts formula

$$ik\int_0^{2\pi}\varphi_k(t)F_{\mu_n}(t)\,dt = F_{\mu_n}(2\pi+) - \int_0^{2\pi}\varphi_k(t)\,dF_{\mu_n}(t)\,dt.$$

Hence $|kf_{n,k}|\le 2C$. Thus, the sequence $\{F_{\mu_n}\}$ is completely bounded in $L^2[0,2\pi]$. If the measures μ_n converge weakly to μ, we have $f_{n,k}\to f_k$ for every k, which is clear from the above-mentioned integration by parts formula (if $k=0$, then we use the equality $t'=\varphi_0(t)$). This gives convergence of F_{μ_n} to F_μ in $L^2[0,2\pi]$. By the uniform boundedness of $\{F_{\mu_n}\}$, convergence in $L^2[0,2\pi]$ is equivalent to convergence in every $L^p[0,2\pi]$, $p<\infty$, and is equivalent to convergence in measure. In the case of a cube we take the basis $\varphi_{k_1,\ldots,k_d}(t_1,\ldots,t_d) := \varphi_{k_1}(t_1)\cdots\varphi_{k_d}(t_d)$ and estimate $(F_{\mu_n},\varphi_{k_1,\ldots,k_d})_{L^2([0,1]^d)}$ by $const\cdot k_1^{-1}\cdots k_d^{-1}$.

8.10.136. Suppose a sequence of signed Borel measures μ_n on a closed interval $[a,b]$ is bounded in the variation norm. Prove that a sufficient (but not necessary) condition of weak convergence of μ_n to a measure μ is convergence of $F_{\mu_n}(t)$ to $F_\mu(t)$ at the points of an everywhere dense set on the real line.

HINT: let f be a continuous function on $[a,b]$ and let $\varepsilon>0$. Let us consider the functions $f_m(t)=\sum_{k=1}^m f(a_{k,m})I_{[a_{k,m},a_{k+1,m})}(t)$, where the points $a_{k,m}$ belong to the set of convergence of F_{μ_n} to F_μ, $a_{1,m}=a$, $a_{k,m}<a_{k+1,m}$, $a_{m,m}=b+m^{-1}$ and $\sup_k|a_{k,m}-a_{k+1,m}|\to 0$ as $m\to\infty$, where we set $f(t):=f(b)$ if $t>b$. Then

$$\int f_m\,d\mu_n \to \int f_m\,d\mu$$

as $n\to\infty$ and $f_m\to f$ uniformly on $[a,b]$.

8.10.137. (cf. Fichtenholz's theorem in Glivenko [**699**, p. 154]) Prove that a sequence of bounded Borel measures μ_n on $\mathbb{R}^d$ converges weakly to a bounded Borel measure μ if and only if (1) the sequence $\{\mu_n\}$ is uniformly bounded in variation and is uniformly tight, (2) the sequence $\{F_{\mu_n}\}$ converges to F_μ in measure with respect to Lebesgue measure on every cube.

HINT: reduce the assertion to the case of measures on a cube by using the mapping $T\colon (x_1,\ldots,x_d)\mapsto(\operatorname{arctg}x_1,\ldots,\operatorname{arctg}x_d)$.

8.10.138. Construct a sequence of measures μ_n on the real line such that $\mu_n=\varrho_n\,dx$, where ϱ_n is a bounded function with support in $[n,n+1]$, $\|\mu_n\|=1$, and for the Kantorovich norm one has $\|\mu_n\|_0\le 2^{-n}$. Thus, the sequence $\{\mu_n\}$ converges to zero in the Kantorovich norm, but is not uniformly tight, in particular, does not converge weakly.

HINT: take the partition of $[n,n+1]$ into 2^n equal intervals I_k of length 2^{-n} and let $\varrho_n:=(-1)^k$ on I_k, $\varrho_n=0$ outside $[n,n+1]$. Let

$$F_n(x):=\int_0^x\varrho_n(t)\,dt.$$

Then $\displaystyle\int_n^{n+1}|\varrho_n(t)|\,dt=1$ and $\displaystyle\int_{-\infty}^{+\infty}|F_n(x)|\,dx\le 2^{-n}$. If f is Lipschitzian with constant 1, the equality $F_n(n)=F_n(n+1)=0$ yields

$$\int_{-\infty}^{+\infty}f(t)\varrho_n(t)\,dt = -\int_n^{n+1}f'(t)F_n(t)\,dt,$$

which is bounded in the absolute value by 2^{-n}.

8.10.139. Construct a net of continuous functions f_α on $[0,1]$ such that one has $0 \le f_\alpha \le 1$, $\lim\limits_\alpha f_\alpha(x) = 1$ for all x, but

$$\lim_\alpha \int_0^1 f_\alpha(x)\,dx = 0.$$

HINT: let the index set Λ consist of all finite subsets α of the interval $[0,1]$ and be partially ordered by inclusion. For every set $\alpha \in \Lambda$ consisting of n points, find $f_\alpha \in C[0,1]$ with $0 \le f_\alpha \le 1$ which equals 1 on α and has the integral less than $1/n$.

8.10.140. (Padmanabhan [**1417**]) Let $(\Omega, \mathcal{B}, P)$ be a probability space and let X be a Polish space. Prove that a sequence of measurable mappings $\xi_n\colon \Omega \to X$ converges in probability to a mapping ξ if and only if for every measure Q that is equivalent to P, the measures $Q \circ \xi_n^{-1}$ converge weakly to the measure $Q \circ \xi^{-1}$.

HINT: it is easy to reduce the assertion to the case $X = [0,1]$. Then, if $P \circ \xi_n^{-1} \Rightarrow P \circ \xi^{-1}$, we have $\|\xi_n\|_2 \to \|\xi\|_2$. Given $A \in \mathcal{B}$ with $P(A) > 0$, we have $(\xi_n, I_A)_2 \to (\xi, I_A)_2$. Indeed, otherwise we may assume that $|(\xi_n, I_A)_2 - (\xi, I_A)_2| \ge c > 0$. Let $Q(B) = (1-\varepsilon)P(B \cap A) + \varepsilon P(B \cap (\Omega \backslash A))$, $\varepsilon = c/4$. Then the integrals of ξ_n with respect to the measure Q do not converge to the integral of ξ with respect to Q, a contradiction. By Corollary 4.7.16 one has $\|\xi_n - \xi\|_2 \to 0$.

8.10.141. Let X be a Souslin space, let Y be a Polish space, and let random elements ξ, ξ_n, $n \in \mathbb{N}$, on a probability space $(\Omega, \mathcal{A}, P)$ with values in X and Borel mappings $f, f_n\colon X \to Y$ be such that the distributions of the elements $f_n \circ \xi_n$ converge weakly to the distribution of $f \circ \xi$. Show that there exist random elements $\widetilde{\xi}, \widetilde{\xi}_n$ in X such that $P \circ \xi^{-1} = P \circ \widetilde{\xi}^{-1}$, $P \circ \xi_n^{-1} = P \circ \widetilde{\xi}_n^{-1}$, and $f_n \circ \widetilde{\xi}_n \to f \circ \xi$ a.e.

HINT: there exist random elements η and η_n in Y such that $\eta_n \to \eta$ a.e. and $P \circ \eta^{-1} = P \circ (f \circ \xi)^{-1}$, $P \circ \eta_n^{-1} = P \circ (f \circ \xi_n)^{-1}$. By using the measurable choice theorem one can find Borel mappings $g, g_n\colon Y \to X$ such that $f\bigl(g(y)\bigr) = y$ for $P \circ \eta^{-1}$-a.e. y, $f_n\bigl(g_n(y)\bigr) = y$ for $P \circ \eta_n^{-1}$-a.e. y. Let $\widetilde{\xi} = g \circ \eta$, $\widetilde{\xi}_n = g_n \circ \eta_n$. Then $f \circ \widetilde{\xi} = \eta$ and $f_n \circ \widetilde{\xi}_n = \eta_n$ a.e.

8.10.142. (Bergin [**154**]) Let X and Y be separable metric spaces, $\mu \in \mathcal{P}_\sigma(X)$, $\nu \in \mathcal{P}_\sigma(Y)$, and let $\eta \in \mathcal{P}_\sigma(X \times Y)$ be such that its projections on X and Y are μ and ν. Suppose we are given two sequences $\{\mu_n\} \subset \mathcal{P}_\sigma(X)$ and $\{\nu_n\} \subset \mathcal{P}_\sigma(Y)$ weakly convergent to μ and ν, respectively. Prove that there are measures η_n in $\mathcal{P}_\sigma(X \times Y)$ weakly convergent to η such that, for each n, the projections of η_n on X and Y are μ_n and ν_n.

8.10.143. Let (X,d) be a bounded separable metric space. Prove that any continuous linear functional L on the normed space $\bigl(M_0(X), \|\cdot\|_0^*\bigr)$ of signed Borel measures σ on X with $\sigma(X) = 0$, where $\|\cdot\|_0^*$ is defined in §8.10(viii), is represented in the integral form by means of a Lipschitzian function F. Prove the same for unbounded X and the space of measures that integrate all Lipschitzian functions.

HINT: let $F(x) := L(\delta_x - \delta_a)$, where $a \in X$ is fixed. Then $|F(x) - F(y)| \le \|L\| \|\delta_x - \delta_y\|_0^* \le \|L\| d(x,y)$ and $L(\delta_x - \delta_y)$ equals the integral of F against $\delta_x - \delta_y$. This yields the same for all measures in $M_0(X)$. Indeed, given two probability measures μ and ν on X, we find finite sums $\mu_n = \sum_{i=1}^n c_{i,n} \delta_{x_{i,n}}$ and $\nu_n = \sum_{i=1}^n c_{i,n} \delta_{y_{i,n}}$ such that $\|\mu_n - \mu\|_0^* \to 0$ and $\|\nu_n - \nu\|_0^* \to 0$. The general case is similar.

CHAPTER 9

Transformations of measures and isomorphisms

> Now what is science?... It is before all a classification, a manner of bringing together facts which appearances separate, though they were bound together by some natural and hidden kinship. Science, in other words, is a system of relations.
>
> *H. Poincaré.* The value of science.

9.1. Images and preimages of measures

Let μ be a Borel measure on a topological space X and let f be a μ-measurable mapping from X to a topological space Y. Then on Y we obtain the Borel measure $\nu = \mu\circ f^{-1}\colon\, B \mapsto \mu\bigl(f^{-1}(B)\bigr)$. The measure ν is called the *image* of μ under the mapping f, and μ is called a *preimage* of μ. The same terms are used in the case of general measurable mappings of measurable spaces. The questions naturally arise about the regularity properties of the measure ν and the properties of the induced mapping $\mu \mapsto \mu\circ f^{-1}$. Such questions are important for measure theory as well as for its applications; these questions have already been touched upon in Chapter 8, see §8.5 and §8.10(v). In particular, it is interesting to know when for a given measure ν on Y, there exists a measure μ with $\nu = \mu\circ f^{-1}$, and when one of the two given measures can be transformed into the other by a transformation with certain additional properties (for example, of continuity). These questions are related to the classification problems for measures. Another important problem concerns invariant measures of a measurable transformation f on a measurable space $(X, \mathcal{A})$, i.e., measures μ on $(X, \mathcal{A})$ such that $\mu = \mu\circ f^{-1}$. In this case, one says that f preserves the measure μ. There is an inverse problem of characterization of transformations preserving a given measure μ. In the subsequent sections all these questions are discussed in detail.

9.1.1. Theorem. *Let X and Y be two Hausdorff spaces.*

(i) *Let $f\colon\, X\to Y$ be a continuous mapping. If a measure μ on X is Radon* (*or is tight or τ-additive*), *then so is $\mu\circ f^{-1}$.*

(ii) *Let Y be a Souslin space* (*for example, a complete separable metric space*) *and let $f\colon\, X\to Y$ be a Borel mapping. Then, the image of every Borel measure μ on X is a Radon measure on Y.*

PROOF. Claim (i) follows directly from the definitions. Claim (ii) follows from the fact that every Borel measure on Y is Radon. $\square$

Our next example shows that assertion (ii) may fail if Y is not Souslin even if X is a Souslin space.

9.1.2. Example. There exists a one-to-one Borel mapping from the interval $[0,1]$ with the standard topology and Lebesgue measure onto a hereditary Lindelöf topological space Y such that the image of Lebesgue measure is not a Radon measure.

PROOF. We have already encountered an example of this sort: take for Y the Sorgenfrey interval $[0,1)$ (see Examples 6.1.19 and 7.2.4) with the added isolated point 2. The Borel σ-algebra of the space Y coincides with the usual Borel σ-algebra of this set on the real line, but the image of Lebesgue measure on $[0,1]$ under the mapping $f(t)=t$, $t<1$, $f(1)=2$, is not a Radon measure on Y since any compact subset in Y is at most countable. □

In the investigation of transformations of measures it is important to be able to find one-sided inverse mappings to not necessarily injective mappings. The next theorem, which is an immediate corollary of Theorem 6.9.1, plays the main role in this circle of problems.

9.1.3. Theorem. *Let X and Y be Souslin spaces and let $f\colon X\to Y$ be a Borel mapping such that $f(X)=Y$. Then, there exists a mapping $g\colon Y\to X$ such that $f\bigl(g(y)\bigr)=y$ for all $y\in Y$ and g is measurable with respect to every Borel measure on Y.*

9.1.4. Corollary. *Suppose that in the situation of the foregoing theorem Y is equipped with a Borel measure ν. Then, there exists a Borel set $Y_0\subset Y$ such that $|\nu|(Y\backslash Y_0)=0$ and $g|_{Y_0}$ is a Borel mapping.*

PROOF. Follows by Corollary 6.7.6. □

The next important result also follows from the previous theorem.

9.1.5. Theorem. *Let X and Y be Souslin spaces and let $f\colon X\to Y$ be a Borel mapping such that $f(X)=Y$. Then, for every Borel measure ν on Y, there exists a Borel measure μ on X such that $\nu=\mu\circ f^{-1}$ and $\|\mu\|=\|\nu\|$.*

If f is a one-to-one mapping, then μ is unique.

PROOF. By the measurable selection theorem, there exists a mapping $g\colon Y\to X$, measurable with respect to the σ-algebra generated by Souslin sets in Y, such that $f\bigl(g(y)\bigr)=y$ for all $y\in Y$. Then the measure $\mu=\nu\circ g^{-1}$ is as required. Indeed, by construction we have $\mu\circ f^{-1}=\nu$. It remains to observe that $\|\nu\|=\|\mu\circ f^{-1}\|\le\|\mu\|$ and $\|\mu\|=\|\nu\circ g^{-1}\|\le\|\nu\|$. If f is one-to-one, then $\mu=\nu\circ g^{-1}$ because $g\bigl(f(x)\bigr)=x$ for all $x\in X$. □

9.1.6. Corollary. *Suppose that in Theorem 9.1.5 the following condition is fulfilled: $|\nu|\bigl(f(W)\bigr)>0$ for every nonempty open set $W\subset X$. Then the measure μ can be chosen in such a way that its support will be the whole space X.*

Proof. Suppose first that ν is a probability measure. We observe that there is a countable collection of Souslin sets $W_i \subset X$ such that $\nu\bigl(f(W_i)\bigr) > 0$ and every nonempty open set $U \subset X$ contains at least one of the sets W_i. Indeed, X is the image of a complete separable metric space E under a continuous mapping ψ. Let us take a countable base $\mathcal{U}$ in E. Set $W_i = \psi(U_i)$, $U_i \in \mathcal{U}$, where we take into account only those U_i for which $\nu\bigl(f(W_i)\bigr) > 0$. If $U \subset X$ is open and nonempty, then $\psi^{-1}(U)$ is a countable union of elements $V_j \in \mathcal{U}$, where the sets $f\bigl(\psi(V_j)\bigr)$ cannot simultaneously have ν-measure zero (otherwise $f(U)$ would have measure zero). Therefore, $\psi^{-1}(U)$ contains some set U_i from the above-chosen collection, hence $W_i \subset U$. By the foregoing theorem, there exists a nonnegative measure μ_i^1 on W_i such that $\mu_i^1 \circ f^{-1} = \nu|_{f(W_i)}$. Next we find a nonnegative measure μ_i^2 on $X\backslash W_i$ that is a preimage of the measure $\nu|_{Y\backslash f(W_i)}$. Let $\mu_i = \mu_i^1 + \mu_i^2$. Then μ_i is a probability measure, $\mu_i \circ f^{-1} = \nu$, and $\mu_i(W_i) > 0$. Let $\mu = \sum_{i=1}^\infty 2^{-i}\mu_i$. It is clear that μ is a probability measure. The support of μ coincides with X, since $\mu(W_i) > 0$ for all i, which due to our choice of W_i yields the positivity of μ on all nonempty open sets. In addition,

$$\mu \circ f^{-1} = \sum_{i=1}^\infty 2^{-i}\mu_i \circ f^{-1} = \sum_{i=1}^\infty 2^{-i}\nu = \nu.$$

If ν is a signed measure, then, as we have established, there exists a nonnegative Borel measure μ_0 with support X such that $\mu_0 \circ f^{-1} = |\nu|$. Let $\nu = \nu^+ - \nu^-$ be the Jordan–Hahn decomposition and let Borel sets Y_1 and Y_2 be such that $Y_1 \cap Y_2 = \varnothing$, $Y_1 \cup Y_2 = Y$ and $\nu^+(Y_2) = \nu^-(Y_1) = 0$. The measure $|\nu|$ can be written as $|\nu| = \zeta \cdot \nu$, where ζ is the Borel function that equals 1 on Y_1 and -1 on Y_2. Set $\mu = (\zeta \circ f) \cdot \mu_0$. Then $|\mu| = \mu_0$, hence the support of μ is X. In addition, $\|\mu\| = \|\nu\|$. Finally, for every bounded Borel function ψ on Y we have

$$\begin{aligned}\int_X \psi\bigl(f(x)\bigr)\,\mu(dx) &= \int_X \psi\bigl(f(x)\bigr)\zeta\bigl(f(x)\bigr)\,\mu_0(dx)\\ &= \int_Y \psi(y)\zeta(y)\,|\nu|(dy) = \int_Y \psi(y)\,\nu(dy),\end{aligned}$$

which gives the equality $\mu \circ f^{-1} = \nu$. □

Let us prove another useful result close to measurable selection theorems.

9.1.7. Proposition. *Let μ be a Radon probability measure on a metric (or Souslin) space X and let f be a μ-measurable function. Then, there exists a μ-measurable set $E \subset X$ such that $f(E) = f(X)$ and the function f is injective on E. The same is true for μ-measurable mappings with values in a metric space Y.*

Proof. By induction one can find compact sets K_n with $K_n \subset K_{n+1}$ whose union has full measure and the restriction of f to every K_n is continuous. By Theorem 6.9.7, every K_n contains a Borel part B_n on which f is injective and $f(B_n) = f(K_n)$. Let $K = \bigcup_{n=1}^{\infty} K_n$ and

$$B = \bigcup_{n=1}^{\infty} \Bigl(B_n \backslash f^{-1}\bigl(f(K_{n-1})\bigr)\Bigr), \quad K_0 = \varnothing.$$

The sets $K_n \cap f^{-1}\bigl(f(K_{n-1})\bigr)$ are compact by the continuity of f on K_n. Hence B is Borel. It is clear that $X\backslash K$ has μ-measure zero. We show that $f\colon B \to f(K)$ is one-to-one. Let $y \in f(K)$ and let n be the smallest number with $y \in f(K_n)$. Then

$$y \in f(K_n)\backslash f(K_{n-1}) \subset f\Bigl(B_n \backslash f^{-1}\bigl(f(K_{n-1})\bigr)\Bigr).$$

Hence there exists $x \in B_n\backslash f^{-1}\bigl(f(K_{n-1})\bigr) \subset B$ with $f(x) = y$, i.e., $y \in f(B)$. If we had another element $x_0 \in B$ with $f(x) = f(x_0)$, then for some $l > n$, we would obtain $x_0 \in B_l\backslash f^{-1}\bigl(f(K_{l-1})\bigr)$. But $f(x_0) = y \in f(K_n) \subset f(K_{l-1})$, i.e., one has $x_0 \in f^{-1}\bigl(f(K_{l-1})\bigr)$, which is a contradiction. Thus, f maps B one-to-one onto $f(K)$. In the set $X\backslash K$ of measure zero, we can choose an arbitrary subset B_0 that is mapped one-to-one onto the set $f(X\backslash K)\backslash f(B)$ if the latter is nonempty. It suffices to pick exactly one element in every set $f^{-1}(y)$, $y \in f(X\backslash K)\backslash f(B)$. The set $E = B \cup B_0$ is as required. The case where f takes values in a separable metric space follows from the considered case, but can also be proved directly by the same reasoning. In the case of a nonseparable Y we apply Theorem 7.14.25 and find a set X_0 of full measure that is mapped to a separable part of Y, find in X_0 a measurable subset mapped injectively onto $f(X_0)$, and then in $X\backslash X_0$ we choose a subset mapped injectively onto $f(X)\backslash f(X_0)$. One can give another proof by employing the measurable choice theorem. □

Clearly, this theorem admits extensions to formally more general settings. For example, it is obvious that the existence of a Souslin subspace of full measure is enough.

Now we consider more general spaces X and Y and the mapping between the spaces of measures generated by a mapping $f\colon X \to Y$. Even if f is continuous and one-to-one, the corresponding mapping from $\mathcal{M}_{\mathcal{B}}(X)$ to $\mathcal{M}_{\mathcal{B}}(Y)$ may be neither injective (as in Example 8.10.29) nor surjective. Let us consider an example of this sort assuming the continuum hypothesis.

9.1.8. Example. Under the continuum hypothesis, there exists a one-to-one continuous mapping f from some complete metric space M onto the interval $[0, 1]$ with its usual metric such that no Borel measure on M is mapped to Lebesgue measure.

PROOF. We equip $[0, 1]$ with the discrete metric. Then all subsets of this space M are closed and the natural mapping of M to $[0, 1]$ with the standard metric is continuous. Suppose there exists a measure μ on $\mathcal{B}(M)$ such that its image is Lebesgue measure. This yields a possibility to extend Lebesgue measure to a measure on the σ-algebra of all subsets of the interval vanishing on all points, which contradicts the continuum hypothesis (see

Corollary 1.12.41). In fact, we have used only that the cardinality of the continuum is not measurable. □

It is clear from this example that Radon and Baire measures may not have preimages under continuous mappings. In addition, it may occur that a Radon measure has a Borel preimage under a continuous mapping, but has no Radon preimages. To see this, it suffices to interchange the spaces in Example 9.1.2, i.e., take for X the Sorgenfrey interval with its natural Lebesgue measure λ, and take for Y the interval $[0,1)$ with the standard topology and Lebesgue measure λ_1, which is the image of λ under the continuous natural projection $X \to Y$, but has no Radon preimages, since all Radon measures on X are purely atomic.

An obvious necessary condition of the existence of a Radon preimage of a Borel measure ν is the existence for every $\varepsilon > 0$ a compact set K_ε in X such that $|\nu|^*\bigl(f(K_\varepsilon)\bigr) > \|\nu\| - \varepsilon$. It turns out that for continuous f this condition is sufficient.

9.1.9. Theorem. *Let f be a mapping from a topological space X to a topological space Y with a Radon measure ν. Suppose that there exists an increasing sequence of compact sets $K_n \subset X$ such that f is continuous on every K_n and*

$$\lim_{n\to\infty} |\nu|\bigl(f(K_n)\bigr) = \|\nu\|.$$

Then, there exists a Radon measure μ on X with $\mu \circ f^{-1} = \nu$. In addition, this measure can be chosen with the property $\|\nu\| = \|\mu\|$. In particular, this is true if X and Y are compact and f is a continuous surjection.

Proof. Suppose first that ν is a nonnegative measure on Y such that one has $\nu(Y\backslash Q) = 0$, where $Q = f(K)$, $K \subset X$ is compact and $f|_K$ is continuous. On the subspace of the space $C(K)$ consisting of all functions of the form $\varphi \circ f$, where $\varphi \in C_b(Y)$, we define a linear functional L by the formula

$$L(\varphi \circ f) = \int_Q \varphi(y)\,\nu(dy).$$

Since

$$\biggl|\int_Q \varphi(y)\,\nu(dy)\biggr| \le \nu(Y)\sup_Q |\varphi| = \nu(Y)\sup_K |\varphi \circ f|,$$

this functional is continuous and by the Hahn–Banach theorem can be extended (with the same norm) to all of $C(K)$. By the Riesz theorem, there exists a Radon measure μ on K with

$$L(\psi) = \int_K \psi\,d\mu, \quad \forall\,\psi \in C(K).$$

Therefore,

$$\int_K \varphi\bigl(f(x)\bigr)\,\mu(dx) = L(\varphi \circ f) = \int_Q \varphi(y)\,\nu(dy), \quad \forall\,\varphi \in C_b(Y).$$

It is clear that $\mu \circ f^{-1} = \nu$ because any continuous function φ has equal integrals with respect to the Radon measures $\mu \circ f^{-1}$ and ν. In addition, one has $\|\mu\| = \|\nu\|$.

Let us extend our assertion to signed measures on Q. Let $\nu = \nu^+ - \nu^-$ be the Jordan–Hahn decomposition, in which the measures ν^+ and ν^- are concentrated on disjoint Borel sets Y^+ and Y^- with $Y^+ \cup Y^- = Y$. We take the nonnegative Radon measures μ_1 and μ_2 constructed above on K such that $\nu^+ = \mu_1 \circ f^{-1}$, $\nu^- = \mu_2 \circ f^{-1}$. One has $\mu_1\bigl(f^{-1}(Y^-)\bigr) = \nu^+(Y^-) = 0$ and similarly, $\mu_2\bigl(f^{-1}(Y^+)\bigr) = 0$. Thus, the measures μ_1 and μ_2 are mutually singular. Hence, letting $\mu = \mu_1 - \mu_2$, we have the equality $\|\mu\| = \|\mu_1\| + \|\mu_2\| = \|\nu^+\| + \|\nu^-\| = \|\nu\|$. It is clear that $\nu = \mu \circ f^{-1}$. It is obvious from our construction for nonnegative measures that the obtained measure μ has the following property: if $|\nu|(C) = 0$ for some Borel set C, then $|\mu|\bigl(f^{-1}(C)\bigr) = 0$ (certainly, the measure ν may have preimages without such a property, for example, the zero measure may have a nonzero signed preimage).

Let us consider the general case. The sets $Q_n = f(K_n)$ are compact. Let $S_n = Q_n \backslash Q_{n-1}$, $Q_0 = \varnothing$. Applying the considered case to the restriction ν_n of the measure ν to the set S_n, considered in the compact space Q_n, we obtain a Radon measure μ_n on K_n such that $\nu_n = \mu_n \circ f^{-1}$. In addition, according to the above construction, the measures μ_n are concentrated on the disjoint sets $f^{-1}(S_n) \cap K_n$ and $\|\mu_n\| = \|\nu_n\|$. Therefore, the series $\sum_{n=1}^{\infty} \mu_n$ converges and defines the measure μ with the required properties. We note that the measure μ is concentrated on the union of the sets K_n, hence the behavior of f outside this union does not affect the measurability of f and the image of μ. $\square$

It is clear that the measure μ constructed above may be non-unique. However, it is unique if f is injective (Exercise 9.12.39).

Let us establish a result on the existence of a preimage of a measure on the preimage of the σ-algebra.

9.1.10. Theorem. *Let F be a mapping from a set X to a measure space $(Y, \mathcal{B}, \nu)$ with a finite measure ν such that $F(X) \in \mathcal{B}$. Let us consider the σ-algebra $\mathcal{A} := F^{-1}(\mathcal{B}) = \{F^{-1}(B), B \in \mathcal{B}\}$. Then $F(A) \in \mathcal{B}$ for all $A \in \mathcal{A}$, and the set function $\mu(A) := \nu\bigl(F(A)\bigr)$, $A \in \mathcal{A}$, is countably additive on $\mathcal{A}$, and if $Y \backslash F(X)$ has $|\nu|$-measure zero, then $\mu \circ F^{-1} = \nu$.*

Proof. If $A \in \mathcal{A}$, then by definition $A = F^{-1}(B)$, where $B \in \mathcal{B}$. Hence $F(A) = B \cap F(X) \in \mathcal{B}$. If sets $A_j \in \mathcal{A}$ are disjoint, then $F(A_j)$ are disjoint as well. Indeed, $A_j = F^{-1}(B_j)$, hence the sets $F(A_j) = B_j \cap F(X)$ do not meet. Therefore, μ is a measure on $\mathcal{A}$. If $F(X)$ has full ν-measure, then we may assume that $F(X) = Y$. Then it is clear that $\mu\bigl(F^{-1}(B)\bigr) = \nu(B)$, $B \in \mathcal{B}$. $\square$

In addition to the inclusion $F(X) \in \mathcal{B}$ required in the above theorem, its essential difference as compared to our previous results is that the measure μ is defined on a rather narrow σ-algebra. For example, if F is the projection from the plane to the real line, then $\mathcal{A}$ contains only the cylinders $B \times \mathbb{R}^1$.

Even in the case where μ extends to a larger σ-algebra, the extension may not be defined by the indicated formula because that formula on a larger σ-algebra may give a non-additive set function.

Now we discuss the question when a given probability measure can be transformed into Lebesgue measure.

9.1.11. Proposition. *Let μ be an atomless probability measure on a measurable space $(X, \mathcal{A})$. Then, there exists an $\mathcal{A}$-measurable function $f\colon X \to [0,1]$ such that $\mu\circ f^{-1}$ is Lebesgue measure.*

PROOF. We give two different proofs employing typical arguments based on two different ideas. It suffices to show that there exists an $\mathcal{A}$-measurable function $f\colon X \to [0,1]$ such that the Borel measure $\mu\circ f^{-1}$ on $[0,1]$ has no atoms because such a measure can be transformed into Lebesgue measure (see Example 3.6.2). Suppose that this is not true. The space F of all $\mathcal{A}$-measurable functions $f\colon X \to [0,1]$ is a closed subset in the Banach space of all bounded functions on X with the norm $\sup_x |f(x)|$. For every n, we consider the set F_n consisting of all $f \in F$ for which the measure $\mu\circ f^{-1}$ has an atom of measure at least n^{-1}. We observe that the sets F_n are closed, since if functions $f_j \in F_n$ converge uniformly to a function f, then the measures $\mu\circ f_j^{-1}$ converge weakly to $\mu\circ f^{-1}$. The atoms of these measures on the interval are points of positive measures. If $\mu\circ f_j^{-1}(c_j) \geq n^{-1}$ and c is a limit point of $\{c_j\}$, then $\mu\circ f^{-1}(c) \geq n^{-1}$, since otherwise one could find an interval $I = [c-\delta, c+\delta]$ with $\mu\circ f^{-1}(I) < n^{-1}$, and then $\mu\circ f_j^{-1}(I) < n^{-1}$ for all sufficiently large j, which leads to a contradiction. By the Baire theorem, some F_n contains a ball U of positive radius r in the space F. Let h be the center of this ball. We shall arrive at a contradiction if we show that U contains a function $g \in F$ such that the measure $\mu\circ g^{-1}$ does not have atoms of measure greater than or equal to $(2n)^{-1}$. The measure $\mu\circ h^{-1}$ has only finitely many different atoms $c_1,\ldots,c_k$ of measure at least $(2n)^{-1}$. Let us take $\delta < r/4$ such that the intervals $[c_i-\delta, c_i+\delta]$ are pairwise disjoint. Since the measure μ has no atoms, by Corollary 1.12.10 every set $E_i := h^{-1}(c_i)$ can be partitioned into finitely many measurable disjoint subsets $E_{i,j}$ with $\mu(E_{i,j}) < (4n)^{-1}$. Since the total number of atoms of the measure $\mu\circ h^{-1}$ is finite or countable, one can find distinct numbers $a_{i,j} \in [c_i-\delta, c_i+\delta]\cap[0,1]$ that are not atoms of this measure. Now let $g(x) = h(x)$ if $x \notin \bigcup_{i=1}^k E_i$, $g(x) = a_{i,j}$ if $x \in E_{i,j}$. For every $c \in [0,1]$, we have $\mu\circ g^{-1}(c) < (2n)^{-1}$. Indeed, if c differs from all $a_{i,j}$, then the set $g^{-1}(c) = h^{-1}(c)$ does not meet $\bigcup_{i=1}^k E_i$ and hence has μ-measure at most $(2n)^{-1}$. If $c = a_{i,j}$, then $g^{-1}(c)$ differs from $E_{i,j}$ in a set of μ-measure zero and also has μ-measure at most $(2n)^{-1}$. It is clear that $g \in U$. This reasoning is frequently used in other situations (see the following proposition).

There is a shorter reasoning based on the fact that every set of positive measure α (for an atomless measure) contains a subset of measure $\alpha/2$. By using this fact and induction, for every rational number r of the form $k2^{-n}$

with $n, k \in \mathbb{N}$, we construct a set X_r with $\mu(X_r) = r$ such that $X_r \subset X_s$ if $r < s$. Namely, we deal first with $r = 1/2$, next with $r = 1/4$ and $r = 3/4$, and so on. Then one can set $f(x) = \inf\{r\colon x \in X_r\}$. Taking into account that $X_r \subset \{f \le r\}$, it is readily verified that $\mu(\{f \le r\}) = r$ for all r of the above form, which proves the claim. □

If we are given a measure on a topological space, then it is natural to investigate the problem of transforming it into Lebesgue measure by means of a continuous mapping.

9.1.12. Proposition. *Let μ be an atomless Radon probability measure on a completely regular space X. Then, there exists a continuous function $f\colon X \to [0,1]$ such that $\mu \circ f^{-1}$ is Lebesgue measure. The same is true in the case of a Baire measure on an arbitrary space.*

PROOF. The reasoning in the first proof of the previous proposition remains valid if we take for F the set of all continuous functions and verify that a function g in U can also be chosen continuous (certainly, the function h is continuous as well). To this end, we consider the same $E_{i,j}$ and $a_{i,j}$ as above, but now $c_1, \ldots, c_k$ are all atoms of $\mu \circ h^{-1}$ of measure at most $(4n)^{-1}$, and we pick the points $a_{i,j}$ in $(c_i, c_i + \delta)$ (if $i = k$, then in $(c_k - \delta, c_k)$). Every set $E_{i,j}$ contains a compact set $K_{i,j}$ with $\mu(E_{i,j} \backslash K_{i,j}) < (8nM)^{-1}$, where M is the total number of sets $E_{i,j}$. There are pairwise disjoint neighborhoods $U_{i,j}$ of the compact sets $K_{i,j}$ such that $\mu(U_{i,j} \backslash K_{i,j}) < (4nM)^{-1}$ and $|h(x) - a_{i,j}| \le \delta$ if $x \in U_{i,j}$. Let $D := X \backslash \bigcup_{i,j} U_{i,j}$. By the complete regularity of X there exists a continuous function g on X such that $g = h$ on D, $g|_{K_{i,j}} = a_{i,j}$, and $|g(x) - a_{i,j}| \le 2\delta$ if $x \in U_{i,j}$. To this end, it suffices to take continuous functions $\zeta_{i,j}\colon X \to [0, a_{i,j} - c_i]$ such that $\zeta_{i,j} = a_{i,j} - c_i$ on $K_{i,j}$ and $\zeta_{i,j} = 0$ outside $U_{i,j}$. Now let

$$g(x) = h(x) + \sum_{i,j} \zeta_{i,j}(x).$$

It is clear that $\sup_x |g(x) - h(x)| \le \delta$. For every $c \in [0,1]$, the set $g^{-1}(c)$ is the union of the sets $g^{-1}(c) \cap K_{i,j}$, $g^{-1}(c) \cap D$, and $g^{-1}(c) \cap (U_{i,j} \backslash K_{i,j})$. If c is not equal to any $a_{i,j}$ and c_i, then

$$\mu\bigl(g^{-1}(c)\bigr) \le \mu\bigl(h^{-1}(c)\bigr) + (4n)^{-1} < (2n)^{-1}$$

since $g = h$ on D. The estimate $\mu(E_i \cap D) \le \mu(E_i \backslash \bigcup_j K_{i,j}) < (8n)^{-1}$ yields that $\mu\bigl(g^{-1}(c_i)\bigr) < (8n)^{-1} + (4n)^{-1} < (2n)^{-1}$. If $c = a_{i,j}$, then $\mu\bigl(g^{-1}(c)\bigr) < (4n)^{-1} + M(4nM)^{-1} = (2n)^{-1}$ since $\mu(K_{i,j}) \le \mu(E_{i,j}) < (4n)^{-1}$ and $\mu(h^{-1}\bigl(a_{i,j})\bigr) = 0$.

In the case of a Baire measure, in place of compact sets $K_{i,j}$ in the previous reasoning we take functionally closed sets and choose functionally open sets $U_{i,j}$ (then there exist the corresponding functions $\zeta_{i,j}$). Certainly, the claim for Radon measures can be easily derived from the claim for Baire measures, but one should remember that the absence of atoms of a Baire measure is not reduced to vanishing on singletons. □

9.2. Isomorphisms of measure spaces

9.2.1. Definition. *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be two measurable spaces with nonnegative measures.*

(i) *A point isomorphism T of these spaces is a one-to-one mapping of X onto Y such that $T(\mathcal{A}) = \mathcal{B}$ and $\mu \circ T^{-1} = \nu$.*

(ii) *The spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are called isomorphic* mod0 *if there exist sets $N \in \mathcal{A}_\mu$, $N' \in \mathcal{B}_\nu$ with $\mu(N) = \nu(N') = 0$ and a point isomorphism T of the spaces $X \backslash N$ and $Y \backslash N'$ that are equipped with the restrictions of the measures μ and ν and the σ-algebras $\mathcal{A}_\mu$ and $\mathcal{B}_\nu$.*

Usually, for brevity, isomorphic mod0 measure spaces are called isomorphic, and when one is concerned with point isomorphisms (or isomorphisms with other additional properties), this is appropriately specified.

In the case when $(X, \mathcal{A}, \mu) = (Y, \mathcal{B}, \nu)$, the isomorphisms of the above types are called *automorphisms*.

We observe that it follows by the definition of a point isomorphism that $\mu(A) = \nu\bigl(T(A)\bigr)$ for all $A \in \mathcal{A}$, since by condition we have $T(A) \in \mathcal{B}$ and $A = T^{-1}\bigl(T(A)\bigr)$. But it is important to remember that a mapping T may not be a point isomorphism even if it is one-to-one, measurable and $\mu \circ T^{-1} = \nu$. The point is that the images of sets in $\mathcal{A}$ may not be in $\mathcal{B}$. For example, this is the case if Lebesgue measure on $[0, 1]$ is considered on the σ-algebra $\mathcal{A}$ of Lebesgue measurable sets and one takes for T the identity mapping to $[0, 1]$ with the Borel σ-algebra $\mathcal{B}$. Certainly, in this example, passing to the completed σ-algebras we change the situation, but more complicated situations are possible.

9.2.2. Theorem. *Let (X, μ) be a Souslin (for example, complete separable metric) space with a Borel probability measure μ. Then (X, μ) is isomorphic* mod0 *to the space $([0, 1], \nu)$, where ν is some Borel probability measure. If μ is an atomless measure, then one can take for ν Lebesgue measure. Both assertions remain valid for Radon measures concentrated on Souslin subsets.*

PROOF. By Theorem 6.7.4, it suffices to consider the case where X is a Souslin subset of $[0, 1]$. Thus, the first claim is already contained in the cited theorem. We only need to show the existence of an isomorphism with Lebesgue measure when the measure μ on $[0, 1]$ has no atoms and is a probability. In that case, its distribution function $F(t) = \mu\bigl([0, t)\bigr) = \mu\bigl([0, t]\bigr)$ is continuous and increasing, $F(0) = 0$ and $F(1) = 1$. It has been verified in Example 3.6.2 that F takes the measure μ to Lebesgue measure λ. If this function were strictly increasing, then it would be a homeomorphism of the interval. However, it is easily seen that F is strictly increasing on the topological support S of μ and $F(S) = [0, 1]$. Sometimes it is more convenient to use the inverse function to F that takes λ to μ. Let

$$G(x) = \inf\bigl\{t \in [0, 1]\colon\ F(t) = x\bigr\}, \quad x \in [0, 1].$$

The function G is strictly increasing on $(0,1)$, since F is increasing and has no jumps. Hence G is a Borel function that maps the interval $(0,1)$ one-to-one to the Borel set $Y := G\bigl((0,1)\bigr)$. We verify that G transforms Lebesgue measure on $(0,1)$ to the measure μ. In order to prove the equality $\mu = \lambda\circ G^{-1}$, it suffices to show that $\mu\bigl((0,c]\bigr) = \lambda\circ G^{-1}\bigl((0,c]\bigr)$ for all $c\in(0,1)$. This is equivalent to the equality $F(c) = \lambda\bigl(G^{-1}(0,c]\bigr)$. Let $c_0 = G\bigl(F(c)\bigr)$. Then one has $c_0\le c$ and $F(c_0) = F(c)$. It remains to observe that we have the equality $G^{-1}(0,c] = \bigl(0, G^{-1}(c_0)\bigr] = \bigl(0, F(c)\bigr]$. □

9.2.3. Corollary. *Let μ be a nonnegative Radon measure on a space X. The following assertions are equivalent*:

(i) *there exists a nonnegative Radon measure ν on a compact metric space Y such that the spaces (X,μ) and (Y,ν) are isomorphic* mod0;

(ii) *one has $\mu(B) = \sup\bigl\{\mu(K)\colon\ K\subset B$ is a metrizable compact set$\bigr\}$ for all sets $B\in\mathcal{B}(X)$.*

Proof. If we have (i), then we may assume that $Y=[a,b]$. We observe that if a function $f\colon X\to[a,b]$ is injective and continuous on a compact set $K\subset X$, then K is metrizable. Indeed, in that case f maps K one-to-one and continuously on the compact set $f(K)\subset[a,b]$. Then it is well known that f is a homeomorphism. By Lusin's theorem on the almost continuity of measurable functions (Theorem 7.1.13) we obtain (ii). If (ii) is fulfilled, then (i) follows by Theorem 9.2.2. □

9.2.4. Lemma. *Let μ be a nonnegative Borel measure on a Souslin space X and let $F\colon X\to X$ be a Borel mapping such that*

$$\mu(B) = \mu\bigl(F(B)\bigr) = \mu\bigl(F^{-1}(B)\bigr),\quad \forall\, B\in\mathcal{B}(X). \tag{9.2.1}$$

Then, there exists a Souslin set $X_0\subset X$ of full μ-measure that is mapped by F one-to-one onto itself.

Proof. By Corollary 9.1.4, there exist a Borel set Y of full μ-measure and a Borel mapping $\Phi\colon Y\to X$ such that $F\bigl(\Phi(y)\bigr) = y$ for all $y\in Y$. It is clear that Φ maps Y one-to-one onto $\Phi(Y)$. In addition, $Z := \Phi(Y)$ is a Souslin set of full measure, since F maps it onto Y. We observe that F is injective on Z. We set $Z_0 = Z\cap F^{-1}(Z)$ and for every integer k we define inductively the sets Z_k by $Z_{k+1} = Z_0\cap F(Z_k)$, $k\ge 0$, $Z_{k-1} = Z_0\cap F^{-1}(Z_k)$, $k\le 0$. All these sets have full measure and are Souslin. Then the set $X_0 = \bigcap_{k\in\mathbb{Z}} Z_k$ is a Souslin set of full measure, F is injective on X_0 and $F(X_0) = X_0$. Indeed, let $x\in Z_k$ for all $k\in\mathbb{Z}$. Then $F(x)\in F(Z_{k-1})\subset Z_k$ if $k\le 0$ and $F(x)\in Z_0\cap F(Z_k) = Z_{k+1}$ if $k\ge 0$. Further, $x = F(z)$, where $z\in Z_0$. By the inclusion $F(z)\in Z_{k+1}$ we obtain $z\in Z_k$ if $k\ge 0$. Next we obtain $z\in Z_k = Z_0\cap F^{-1}(Z_{k+1})$ if $k<0$, since $F(z) = x\in Z_{k+1}$. Thus, $z\in X_0$. □

9.2.5. Corollary. *The statement of Lemma* 9.2.4 *remains valid for any μ-measurable mapping F satisfying condition* (9.2.1) *provided that $F(B)$ is μ-measurable for every $B\in\mathcal{B}(X)$.*

PROOF. By Corollary 6.7.6, there exist a Borel set N_0 of μ-measure zero and a Borel mapping F_0 equal to F outside N_0. We redefine F_0 on N_0 by setting $F_0|_{N_0} = a$, where a is an arbitrary point in N_0. Since by hypothesis $\mu\bigl(F(N_0)\bigr) = 0$, the mapping F_0 satisfies the hypothesis of Lemma 9.2.4, hence there exists a Souslin set $X_0 \subset X\backslash N_0$ of full measure that is mapped by F_0 (hence by F) one-to-one onto itself. □

Apart from a rather rough classification of measures by means of general measurable mappings, in many problems it is important to employ finer classifications, for example, by means of continuous or smooth mappings (or mappings with other additional special properties). Brief comments on this are given in §9.6 and §9.12(vi) (see also §5.8(x)).

9.3. Isomorphisms of measure algebras

Let $(X, \mathcal{A}, \mu)$ be a measure space with a finite nonnegative measure μ and let the σ-algebra $\mathcal{A}$ be complete with respect to μ. In this case we shall call the metric Boolean algebra $\mathcal{A}/\mu$ considered in Chapter 1 *a measure algebra* and denote it by E_μ. The elements of this algebra are equivalence classes of μ-measurable sets with the metric $\varrho(A, B) = \mu(A \triangle B)$. We recall that E_μ is a complete metric space (note that by our definition, $\mathcal{A}/\mu$ is complete even if $\mathcal{A}$ is not; completeness of $\mathcal{A}$ is assumed for convenience). One defines the operations of union, intersection and complementation for all elements of E_μ as the respective operations on representatives of the equivalence classes.

9.3.1. Definition. *Two measure algebras E_{μ_1} and E_{μ_2} generated by measure spaces $(X_1, \mathcal{A}_1, \mu_1)$ and $(X_2, \mathcal{A}_2, \mu_2)$ are called isomorphic if there exists a one-to-one mapping J from E_{μ_1} onto E_{μ_2} (called a metric Boolean isomorphism) such that J preserves the measure, i.e., $\mu_2\bigl(J(A)\bigr) = \mu_1(A)$ for all $A \in E_{\mu_1}$, and, in addition,*

$$J(A\backslash B) = J(A)\backslash J(B) \quad \textit{and} \quad J(A \cup B) = J(A) \cup J(B)$$

(*then also* $J(A \cap B) = J(A) \cap J(B)$).

It is clear from the definition that the equivalence class of X_1 corresponds to the equivalence class of X_2. We may assume that the isomorphism J maps $\mathcal{A}_1$ to $\mathcal{A}_2$ such that the correspondence of unions, intersections, and complements holds up to sets of measure zero.

In the investigation of measure algebras an important role is played by countable measurable partitions, i.e., partitions of a measure space $(X, \mathcal{A}, \mu)$ into pairwise disjoint measurable sets X_n. The diameter of the partition $\mathcal{X} = \{X_n\}$ is the number

$$\delta(\mathcal{X}) = \sup_n \mu(X_n).$$

A partition $\mathcal{X}$ called is a refinement of a partition $\mathcal{Y}$ if every element of $\mathcal{X}$ is contained in an element of $\mathcal{Y}$.

9.3.2. Lemma. *Let $\mathcal{X}_n$ be a sequence of partitions of $[0,1]$ into finite collections of intervals (open, closed or half-closed) such that $\lim\limits_{n\to\infty}\delta(\mathcal{X}_n)=0$. Then, the set of all finite unions of elements of the partitions $\mathcal{X}_n$ is everywhere dense in the measure algebra E_λ, where λ is Lebesgue measure on $[0,1]$.*

PROOF. It suffices to show that for every interval $I=[a,b]\subset[0,1]$ and every $\varepsilon>0$, one can find a finite collection $I_1,\ldots,I_k$ of elements of the partitions $\mathcal{X}_n$ with $\lambda\bigl(I\bigtriangleup\bigcup_{i=1}^k I_i\bigr)<\varepsilon$. Let us pick n such that $\delta(\mathcal{X}_n)<\varepsilon/2$. Let I_1 be the uniquely defined interval in $\mathcal{X}_n$ containing a. If $b\in I_1$, then I_1 gives the required approximation. Otherwise we take the consecutive intervals $I_1,\ldots,I_k$ in the partition $\mathcal{X}_n$ such that $b\in I_k$. It is clear that the union of I_j approximates I up to ε with respect to Lebesgue measure. □

9.3.3. Lemma. *Let μ be an atomless probability measure on a space $(X,\mathcal{A},\mu)$ and let $\{\mathcal{X}_n\}$ be a sequence of countable measurable partitions such that $\mathcal{X}_{n+1}$ is a refinement of $\mathcal{X}_n$ for all n and the set of all finite unions of elements of these partitions is everywhere dense in the measure algebra E_μ. Then $\lim\limits_{n\to\infty}\delta(\mathcal{X}_n)=0$.*

PROOF. Suppose this is not true. Let $\mathcal{X}_n$ consist of sets $A_{n,j}$. The diameters $\delta(\mathcal{X}_n)$ of our decreasing partitions are decreasing to some $\delta>0$. There exists an index k_1 such that for all n one has

$$\sup_j \mu(A_{1,k_1}\cap A_{n,j})\geq\delta-\delta/4.$$

Indeed, there are only finitely many sets $A_{1,j_1},\ldots,A_{1,j_m}$ in $\{A_{1,j}\}$ with measure not less than $\delta-\delta/4$. Among these sets, at least one, which will be denoted by A_{1,k_1}, contains sets with measure at least $\delta-\delta/4$ from infinitely many $\mathcal{X}_n$ because any set $A_{n,j}$ with $\mu(A_{n,j})\geq\delta-\delta/4$ must be entirely contained in one of $A_{1,j_1},\ldots,A_{1,j_m}$. But then A_{1,k_1} contains such sets from every $\mathcal{X}_n$, since any $A_{n,j}$ is contained in some $A_{n-1,i}$. Next, by the same reasoning, we can find $A_{2,k_2}\subset A_{1,k_1}$ such that

$$\sup_j \mu(A_{2,k_2}\cap A_{n,j})\geq\delta-\delta/4-\delta/8$$

for all n. By induction, for each $m\in\mathbb{N}$, we find $A_{m,k_m}\subset A_{m-1,k_{m-1}}$ with

$$\sup_j \mu(A_{m,k_m}\cap A_{n,j})\geq\delta-\delta\sum_{i=1}^m 2^{-1-i}\quad\text{for all } n.$$

Let $A=\bigcap_{m=1}^\infty A_{m,k_m}$. It is clear that $\mu(A)\geq\delta/2>0$. Since E_μ has no atoms, there exists a measurable set $B\subset A$ with $0<\mu(B)<\mu(A)$. Let us fix a positive number $\varepsilon<\min\bigl(\mu(B),\mu(A\backslash B)\bigr)$. By hypothesis, B is approximated in E_μ up to ε by the union of some sets $B_1,\ldots,B_k$ from the partitions $\mathcal{X}_n$. We observe that for any element C in any partition $\mathcal{X}_n$, the set B is either contained in C or does not meet C. Indeed, if B is not contained in C, then the set A_{n,k_n} is not contained in C. All elements of the partition $\mathcal{X}_n$ are disjoint, hence $C\cap A_{n,k_n}=\varnothing$, whence it follows that $C\cap B=\varnothing$. Since

$\mu(B) > \varepsilon$, some of the sets B_i contain B. We may assume that $B \subset B_1$. Then $\mu(B_1 \backslash B) < \varepsilon$, and in order to obtain a contradiction, it remains to observe that at the same time we have $\mu(B_1 \backslash B) > \varepsilon$. This follows by the inclusion $A \backslash B \subset B_1 \backslash B$, implied by the inclusion $A \subset B_1$, which is verified as follows. We have $B_1 = A_{n,j}$ for some n and j. Then $B_1 = A_{n,k_n}$. Indeed, otherwise $A_{n,j} \cap A_{n,k_n} = \varnothing$, hence $B_1 \cap A = \varnothing$, which is impossible, since we have $B \subset B_1$ and $B \subset A$. □

9.3.4. Theorem. *Every separable atomless measure algebra is isomorphic to the measure algebra of some interval with Lebesgue measure.*

Proof. Let E_μ be the separable atomless measure algebra generated by a probability measure μ on a space X and let $\{E_n\}$ be a countable everywhere dense family in E_μ. We show that there exists an isomorphism $J\colon E_\mu \to E_\lambda$, where λ is Lebesgue measure on $[0,1]$. For every fixed n, we consider the partition of X into measurable pairwise disjoint sets of the form $\bigcap_{i=1}^n A_i$, where for every $i = 1, \ldots, n$, the set A_i is either E_i or $X \backslash E_i$. The sets obtained in this way are denoted by $A_{n,j}$, $j \le 2^n$. The required isomorphism J is first defined inductively on the sets $A_{n,j}$, which will be sent to some intervals (closed or semiclosed). Let $J(A_{1,1}) = [0,a]$, $J(A_{1,2}) = (a,1]$, where $a = \mu(A_{1,1})$. If intervals $J(A_{n,j})$ are already found for some $n \ge 1$, then the choice of $J(A_{n+1,j})$ is made in the following way. Every element $A_{n,j}$ consists of two elements $A_{n+1,j'}$ and $A_{n+1,j''}$ and is already mapped to the interval $J(A_{n,j})$ of length $\mu(A_{n,j})$. We partition this interval into two subintervals of length $\mu(A_{n+1,j'})$ and $\mu(A_{n+1,j''})$, then associate the first of them to the element $A_{n+1,j'}$, and the second one to the element $A_{n+1,j''}$. Next we proceed by induction. By construction, for every fixed n, the intervals $J(A_{n,j})$ are pairwise disjoint, have lengths $\mu(A_{n,j})$, and form a partition of $[0,1]$. Now J extends to all finite unions of disjoint sets $A_{n_1,k_1}, \ldots, A_{n_m,k_m}$: such a union is mapped to the union of the corresponding intervals $J(A_{n_i,k_i})$. The constructed mapping is an isometry on the union of $A_{n,j}$, $n \in \mathbb{N}$, $j \le 2^n$. If we show that the domain of definition is everywhere dense in E_μ and the set of values is everywhere dense in E_λ, then we can extend J by continuity to E_μ. Since J on the already-existing domain of definition satisfies the conditions $J(X \backslash A) = [0,1] \backslash J(A)$ and $J(A \cap B) = J(A) \cap J(B)$, these conditions hold on all of E_μ in the result of extension by continuity (we recall that the elements of E_μ are equivalence classes, not individual sets). Finite unions of disjoint sets $A_{n,j}$ give all sets E_n, hence the initial domain of definition of J is everywhere dense in E_μ. That the range is dense follows by Lemma 9.3.2, since $\lim\limits_{n\to\infty} \max_{j\le 2^n} \mu(A_{n,j}) = 0$ by Lemma 9.3.3. □

For every set A of positive measure, the restriction μ_A of the measure μ to A defines another measure algebra $E_{\mu,A}$. The measure algebra E_μ is called *homogeneous* if all metric spaces $E_{\mu,A}$ (where $\mu(A) > 0$) have equal weights (the weight of a metric space is the least cardinality of its topology bases).

The following fundamental result on the structure of measure algebras is due to D. Maharam [**1228**]. It is valid even in the more general framework of Boolean algebras (see Vladimirov [**1947**]).

9.3.5. Theorem. (i) *Every atomless measure algebra is the direct sum of at most countably many homogeneous measure algebras.*

(ii) *Every atomless homogeneous measure algebra corresponding to a probability measure is isomorphic to the measure algebra generated by certain power of unit intervals with Lebesgue measure.*

It is important to note that an isomorphism of measure algebras does not always yield an isomorphism of the underlying measure spaces (see Example 9.5.3 below). In §9.5, we discuss certain important cases when such an implication is true. A discussion of measure algebras and further references can be found in Fremlin [**629**], [**635**].

9.4. Lebesgue–Rohlin spaces

A class of measure spaces important for applications was introduced and studied by V.A. Rohlin, who called them Lebesgue spaces. In this section, we consider only finite nonnegative measures.

We shall say that a measure space $(M, \mathcal{M}, \mu)$ has a countable basis $\{B_n\}$ if the sets $B_n \in \mathcal{M}$ separate the points in M (i.e., for every two distinct points x and y, there exists B_n such that either $x \in B_n$, $y \notin B_n$ or $x \notin B_n$, $y \in B_n$) and the Lebesgue completion of $\sigma(\{B_n\})$ coincides with the completion of $\mathcal{M}$ (i.e., $\sigma(\{B_n\})_\mu = \mathcal{M}_\mu$). In other words, every μ-measurable set is contained between two sets from $\sigma(\{B_n\})$ of equal measure.

A space with such a property will be called *separable in the sense of Rohlin*. In our earlier-introduced terminology, a measure space $(M, \mathcal{M}, \mu)$ is separable in the sense of Rohlin precisely when one can find a countably generated and countably separated σ-algebra $\mathcal{A} \subset \mathcal{M}$ with $\mathcal{A}_\mu = \mathcal{M}_\mu$.

Let $\Omega = \{0, 1\}^\infty$ be the space of all sequences $\omega = (\omega_i)$, where ω_i is 1 or 0. For every $\omega \in \Omega$, let

$$E_\omega = \bigcap_{n=1}^{\infty} B_n(\omega_n),$$

where $B_n(\omega_n) = B_n$ if $\omega_n = 1$ and $B_n(\omega_n) = M \backslash B_n$ if $\omega_n = 0$.

If the sets B_n separate the points in M, then each set E_ω contains at most one point.

The space M is called complete with respect to its basis $\{B_n\}$ if every E_ω is nonempty.

Thus, for a complete space, the set E_ω is some point $x_\omega \in M$, and every point $x \in M$ coincides with some E_ω: for $\omega = \omega(x)$ we take the sequence such that $\omega_n = 1$ if $x \in B_n$, $\omega_n = 0$ if $x \notin B_n$. The formula $\psi\colon\ x \mapsto \omega(x)$ defines a one-to-one mapping of M onto Ω. In particular, M has cardinality of the continuum.

9.4.1. Example. Let us equip the space Ω with its natural σ-algebra $\mathcal{B}$ generated by the cylinders $C_n = \{\omega \in \Omega\colon\ \omega_n = 1\}$ (i.e., $\mathcal{B} = \mathcal{B}(\Omega)$ if Ω is regarded as a topological product, in which case it becomes a compact metric space). Then, for every Borel measure ν on Ω, the space $(\Omega, \mathcal{B}, \nu)$ is complete with respect to the basis $\{C_n\}$.

PROOF. We only have to verify that the sets E_ω are nonempty. Since the complement to C_n consists of all sequences with the zero nth component, the point x_ω is found explicitly: its nth component is ω_n. $\square$

The completeness with respect to a basis means, in particular, that all B_n have a nonempty intersection (and if we replace some of them by their complements, then such sets will have a common point, too). Hence the natural basis of Lebesgue measure on $[0,1]$, consisting of all intervals with the rational endpoints, does not satisfy this condition. Generally speaking, it is not a very trivial task to construct a basis with respect to which a given space is complete, as we shall now see from an example of Lebesgue measure. For this reason, a considerably broader concept of completeness mod0 is discussed below.

9.4.2. Example. Let M be an uncountable Borel set in a complete separable metric space and let μ be a Borel measure on M. Then the space $\bigl(M, \mathcal{B}(M), \mu\bigr)$ has a countable basis with respect to which it is complete.

PROOF. By Corollary 6.8.8 the space M is Borel isomorphic to $\{0,1\}^\infty$. A basis in M with the required properties can be constructed as follows: we consider the basis in $\{0,1\}^\infty$ described above and take its image under the Borel isomorphism $J\colon\ \{0,1\}^\infty \to M$. $\square$

In some cases, one can find a basis with the completeness property in a more constructive way.

9.4.3. Example. Let M be the set of all points in $[0,1]$ whose ternary expansions do not contain 2 and let B_n be the set of all points in M that have 1 at the nth position in the ternary expansion. Then M with an arbitrary Borel measure is complete with respect to the basis $\{B_n\}$. In addition, the mapping $\omega \mapsto \sum_{n=1}^\infty \omega_n 3^{-n}$ defines an isomorphism between $\{0,1\}^\infty$ and M.

9.4.4. Example. The space $\bigl([0,1], \mathcal{B}([0,1]), \lambda\bigr)$, where λ is Lebesgue measure, has a countable basis with respect to which it is complete (this follows by Example 9.4.2, but there is no explicit construction there).

PROOF. The points in $[0,1]$ have the binary expansions $x = \sum_{n=1}^\infty \omega_n 2^{-n}$, where ω_n equals 1 or 0. With the exception of points of some countable set $S \subset [0,1]$, the indicated expansion is unique. Thus, $X = [0,1]\backslash S$ is in a one-to-one correspondence with the complement in $\Omega = \{0,1\}^\infty$ of the countable set S' consisting of all sequences whose components are constant from a certain position. The sets S and S' can also be put into a one-to-one correspondence. Let B_n be the set in $[0,1]$ corresponding to the set C_n in Example 9.4.1 under the above-described Borel isomorphism between $[0,1]$ and Ω. Thus, if we neglect the countable set S, then B_n is a finite collection of binary rational intervals in X containing all numbers with 1 at the nth place in the binary expansion. According to Example 9.4.1, the basis $\{B_n\}$ has the completeness property. □

9.4.5. Definition. *Let $(M,\mathcal{M},\mu)$ be a measure space with a countable basis $\{B_n\}$. We shall say that this space is complete* mod0 *with respect to the basis $\{B_n\}$ if there exist a measurable space $(\widetilde{M},\widetilde{\mathcal{M}},\widetilde{\mu})$, complete with respect to some basis $\{\widetilde{B_n}\}$, a set $M_0 \in \widetilde{\mathcal{M}}_{\widetilde{\mu}}$ of full $\widetilde{\mu}$-measure, and a one-to-one measurable mapping $\pi\colon\ M \to M_0$ such that*

$$\pi(B_n) = \widetilde{B_n} \cap M_0 \quad \textit{and} \quad \mu\circ\pi^{-1} = \widetilde{\mu}.$$

In fact, the property of completeness mod0 is a possibility to realize the given space M as a subset of full measure in some space $\widetilde{M}$ with a basis with respect to which $\widetilde{M}$ is complete such that the intersections of elements of this basis with M form the given basis of M. We observe that the condition $\pi(B_n) = \widetilde{B_n} \cap M_0$ yields the $\bigl(\sigma(\{B_n\}), \sigma(\{\widetilde{B_n}\})\bigr)$-measurability of π.

The following important definition uses the concept of isomorphism mod0 from Definition 9.2.1.

9.4.6. Definition. *A measure space $(M,\mathcal{M},\mu)$ is called a Lebesgue–Rohlin space if it is isomorphic* mod0 *to some measure space $(M',\mathcal{M}',\mu')$ with a countable basis with respect to which M' is complete.*

It is clear that if a space $(M,\mathcal{M},\mu)$ is complete mod0 with respect to some basis, then it is a Lebesgue–Rohlin space. Unlike the property of completeness, the property of completeness mod0 is independent of the choice of a basis.

9.4.7. Theorem. *Let $(M,\mathcal{M},\mu)$ be a Lebesgue–Rohlin space with a probability measure μ. Then it is isomorphic* mod0 *to the interval $[0,1]$ with the measure $\nu = c\lambda + \sum_{n=1}^{\infty} \alpha_n \delta_{1/n}$, where $c = 1 - \sum_{n=1}^{\infty}\alpha_n$, $\alpha_n = \mu(a_n)$ and $\{a_n\}$ is the family of all atoms of μ.*

PROOF. Suppose that M has a basis $\{B_n\}$ with respect to which it is complete. Let us consider the above-constructed one-to-one mapping $\pi\colon\ x \mapsto \omega(x)$ from M onto $\Omega = \{0,1\}^{\infty}$. It is readily verified that $\pi(B_n) = C_n$, where $\{C_n\}$ is the basis in Ω indicated in Example 9.4.1. Therefore, π is an isomorphism between $\bigl(M,\sigma(\{B_n\})\bigr)$ and $\bigl(\Omega,\mathcal{B}(\Omega)\bigr)$. Let $\nu = \mu\circ\pi^{-1}$. Then π is an isomorphism between $(M,\mathcal{M}_\mu,\mu)$ and $(\Omega,\mathcal{B}_\nu,\nu)$, since $\sigma(\{B_n\})_\mu = \mathcal{M}_\mu$. According to Theorem 9.2.2, there exists an isomorphism mod0 between the space $(\Omega,\mathcal{B}_\nu,\nu)$ and the measurable space generated on $[0,1]$ by the probability measure $c\lambda + \sum_{n=1}^{\infty} c_n\delta_{1/n}$, where $c_n = \nu(x_n)$ and $\{x_n\}$ is the family of all atoms of ν. The general case by definition reduces to the considered one. □

9.4.8. Theorem. *If a measure space $(M,\mathcal{M},\mu)$ is separable in the sense of Rohlin and complete* mod0 *with respect to some basis, then it is complete* mod0 *with respect to every basis.*

PROOF. By Theorem 9.4.7, it suffices to prove our claim for the interval $[0,1]$ with a Borel measure μ. Let $\{B_n\}$ be a basis consisting of Borel sets. To every point $x \in [0,1]$, we associate the point $\omega = \pi(x) \in \Omega = \{0,1\}^\infty$ such that $\omega_n = 1$ if $x \in B_n$ and $\omega_n = 0$ if $x \notin B_n$. Since $\{B_n\}$ separates the points, we obtain an injective mapping to Ω. We observe that $\pi(B_n) = C_n \cap \pi([0,1])$, where $\{C_n\}$ is the basis in Ω from Example 9.4.1. Hence for completing the proof it remains to verify that π is a Borel mapping because in that case $\pi([0,1])$ is a Borel set and $\nu = \mu \circ \pi^{-1}$ is a Borel measure concentrated on this set. Since $\{C_n\}$ is a basis in Ω, the inclusion $\pi^{-1}\bigl(\mathcal{B}(\Omega)\bigr) \subset \mathcal{B}([0,1])$ follows by the easily verified equality $\pi^{-1}(C_n) = B_n$. □

It is useful to introduce also the concept of a basis mod0. We shall say that a sequence of sets B_n in a measure space $(M, \mathcal{M}, \mu)$ is a basis mod0 if $B_n \in \mathcal{M}$ and there exists a set $Z \in \mathcal{M}_\mu$ of μ-measure zero such that the sets $B_n' = B_n \cap (M\backslash Z)$ form a basis in the space $M\backslash Z$ equipped with the induced σ-algebra and the restriction of the measure μ. If the latter space is complete mod0 with respect to $\{B_n'\}$, then we shall say that $(M, \mathcal{M}, \mu)$ is complete mod0 with respect to its basis mod0. It is clear that the existence of a basis mod0 with respect to which the space is complete mod0 is equivalent to saying that the given space is a Lebesgue–Rohlin space.

Let us explain why one should use the concept mod0 in dealing with bases as well as with completeness. Let us take for M a set of cardinality greater than that of the continuum with the σ-algebra of all subsets. Let μ be Dirac's measure at the point m. Here it is necessary to delete a set of measure zero in order that the remaining set could be embedded into an interval. Now suppose that only the point m is left: we may assume that we have the point 0 in $[0,1]$. The singleton (as well as any at most countable set) has no basis with the property of completeness, since the complement of the only nonempty set is empty. Hence one has to enlarge the space, embedding it, for example, in an interval. Then the basis of the singleton 0 consisting of the single set 0 can be obtained as the intersection of a basis in the interval with the point 0. The concept of a basis mod0 turns out to be much more flexible, so that many natural systems (such as the rational intervals) become such bases.

9.4.9. Lemma. *Suppose that a measure space $(M, \mathcal{M}, \mu)$ is separable in the sense of Rohlin and $\{B_n\} \subset \mathcal{M}_\mu$ is a sequence of sets such that every set in $\mathcal{M}_\mu$ coincides up to a set of measure zero with some set in $\sigma(\{B_n\})$. Then $\{B_n\}$ is a basis* mod0.

PROOF. Let $\{A_n\}$ be some basis in M. For every n, there exists a set $E_n \in \sigma(\{B_n\})$ with $\mu(A_n \triangle E_n) = 0$. We verify that in the new space $M_0 = M\backslash\bigcup_{n=1}^\infty (A_n \triangle E_n)$ the sets $B_n' = B_n \cap M_0$ form a basis. Let x and y be two distinct points in M_0. We find a set A_n separating them. We may assume that $x \in A_n$, $y \notin A_n$. Then $x \in E_n$, since $A_n\backslash E_n$ does not meet M_0. Similarly, one verifies that $y \notin E_n$. Thus, the sets $\{E_n\}$ separate the points in M_0. By Lemma 6.5.3, the sets $\{B_n'\}$ have the same property. Let $A \subset M_0$ and $A \in \mathcal{M}_\mu$. Let us show that there exist sets E, E' in $\sigma(\{B'_n\})$ with equal measures and $E \subset A \subset E'$. To this end, we observe that $A_n \cap M_0 = E_n \cap M_0$ and hence the σ-algebra generated by the sets $A_n \cap M_0$ is contained in the σ-algebra generated by the sets B'_n. By our hypothesis, there exist sets $D, D' \in \sigma(\{A_n\})$ such that $D \subset A \subset D'$ and $\mu(D) = \mu(D')$. Therefore, the sets $E := D \cap M_0$ and $E' := D' \cap M_0$ belong to $\sigma(\{A_n \cap M_0\}) \subset \sigma(\{B'_n\})$, $E \subset A \subset E'$ and $\mu(E) = \mu(E')$. □

9.4.10. Proposition. (i) *Every measurable subset in a Lebesgue–Rohlin space with the induced measurable structure is a Lebesgue–Rohlin space.*

(ii) *Let a measure space $(M, \mathcal{M}, \mu)$ be separable in the sense of Rohlin and let $A \subset M$. Suppose that the space $(A, \mathcal{M}_A, \mu_A)$, where $\mathcal{M}_A = \mathcal{M} \cap A$ and μ_A is the restriction of the outer measure to $\mathcal{M}_A$ (see* §1.12(iv)) *is a Lebesgue–Rohlin space. Then $A \in \mathcal{M}_\mu$.*

Proof. Assertion (i) is obvious from Theorem 9.4.7. Let us prove assertion (ii). According to Theorem 6.5.7, we may assume that M is contained in $[0,1]$, $\mathcal{B}(M) \subset \mathcal{M}$ and $\mathcal{M}_\mu = \mathcal{B}(M)_\mu$. In addition, we may assume that $\mu(M) = 1$ and $\mu^*(A) = 1$. By Theorem 9.4.7, there exist a set $A_0 \subset A$ with $\mu_A(A_0) = 1$, a Borel set $B_0 \subset [0,1]$, and a one-to-one mapping $f\colon B_0 \to A_0$ such that $f^{-1}(B) \in \mathcal{B}([0,1])$ for all B in $\mathcal{B}(A_0)$. This means that f is a Borel function. Hence $A_0 = f(B_0)$ is a Borel set. It is then clear that $\mu(A_0) = 1$ and hence the set A is μ-measurable. □

The discussion of Lebesgue–Rohlin spaces will be continued in §10.8, where we consider measurable partitions. Here we only note that the presented proofs of the main results on the structure of Lebesgue–Rohlin spaces are shorter than the original ones (mostly due to the use of some earlier-obtained results). In spite of this, the reader is strongly encouraged to get acquainted with the classical work of Rohlin [**1595**], where the techniques of proof correspond perfectly to the general idea and spirit of the work: to distinguish in intrinsic terms of a measurable structure the properties enjoyed by a broad class of spaces that are most diverse from the topological point of view.

9.5. Induced point isomorphisms

In this section, we consider only finite nonnegative measures.

It is clear that every isomorphism mod0 induces a metric Boolean isomorphism. As Example 9.5.3 shows, the converse is false. However, a classical result due to von Neumann [**1361**] states that any metric Boolean automorphism of the measure algebra corresponding to an interval with a Borel measure is induced by an automorphism mod0. Here is an abstract version of this important result.

9.5.1. Theorem. *Let $(M_1, \mathcal{M}_1, \mu_1)$ and $(M_2, \mathcal{M}_2, \mu_2)$ be Lebesgue–Rohlin spaces with probability measures. If the corresponding measure algebras E_{μ_1} and E_{μ_2} are isomorphic in the sense of Definition* 9.3.1*, then there exists* *an isomorphism* mod0 *between these measure spaces. In particular, this is the case if both measures are atomless.*

PROOF. Suppose first that both measures have no atoms. By the isomorphism theorem it suffices to consider the case where M_1 and M_2 coincide with $M := \{0,1\}^\infty$, $\mathcal{M}_1 = \mathcal{M}_2 = \mathcal{B}(M)$ and $\mu_1 = \mu_2 = \mu$. Let J be an automorphism of the measure algebra E_μ. Let us take the standard basis $C_k = \{(\omega_i)\colon \omega_k = 1\}$ of the space $\{0,1\}^\infty$. Let B_k be an arbitrary Borel representative of the class $J(C_k)$. By hypothesis, J preserves (up to sets of measure zero) finite unions, finite intersections, and complements and preserves the measure. Hence all μ-measurable sets are approximated mod0 by sets from the σ-algebra generated by $\{B_n\}$. By Lemma 9.4.9, $\{B_n\}$ is a basis mod0. According to Theorem 9.4.8, the space M is complete mod0 with respect to $\{B_n\}$. This means that M contains a Borel set M_0 of full μ-measure that can be embedded into some measurable space $(\widetilde{M}, \widetilde{\mathcal{M}}, \widetilde{\mu})$ with a basis $\widetilde{B_n}$ with respect to which $\widetilde{M}$ is complete, such that the sets $B_n \cap M_0$ are mapped to the sets $\widetilde{B_n}$, M_0 is mapped to a measurable set of full $\widetilde{\mu}$-measure, and the measure μ is transformed to $\widetilde{\mu}$. We may assume that $\widetilde{M}$ is obtained by adding to M_0 some Borel set $Z \subset M$ with $\mu(Z) = 0$ and that $\widetilde{\mu} = \mu$ so that the embedding is the identity mapping. Let us consider the mapping $f\colon M \to \widetilde{M}$, $\omega \mapsto \bigcap_{k=1}^\infty \widetilde{B_k}(\omega_k)$. This mapping is a Borel isomorphism between M and $\widetilde{M}$ and takes the standard basis $\{C_n\}$ of the space M to the basis $\{\widetilde{B_n}\}$. We observe that by construction one has

$$\mu(C_n) = \mu(B_n) = \mu(\widetilde{B_n}) = \mu\bigl(f(C_n)\bigr).$$

By using that f is one-to-one, that $J(C_n) = f(C_n)$ up to a set of measure zero and that J is an isometry, we obtain that for every set C in the algebra generated by $\{C_n\}$, one has the equality $\mu(C) = \mu\bigl(f(C)\bigr)$. Then this equality remains true for all sets $C \in \mathcal{M} = \sigma(\{C_n\})$. Therefore, f preserves μ and the induced mapping on E_μ coincides with J. In the general case, the measures μ_1 and μ_2 have atoms, but it is easy to see that the atoms a_n^1 of the measure μ_1 are taken by the mapping J to the atoms a_n^2 of the measure μ_2. We may assume again that both measures are realized on $\{0,1\}^\infty$. Then the atoms are points of positive measure. It is clear that J is a metric Boolean isomorphism of the measure algebras E_{ν_1} and E_{ν_2}, where ν_i is the restriction of μ_i to $N_i = M\backslash\{a_n^i\}$. It is easily seen that the measures ν_i have no atoms. As already shown, there exists an isomorphism mod0 of the spaces N_1 and N_2 generating the above-mentioned metric isomorphism. It remains to extend this isomorphism to $\{a_n^1\}$, by associating to every atom a_n^1 the atom a_n^2. □

9.5.2. Corollary. *Let (X,μ) and (Y,ν) be Souslin spaces with probability Borel measures. If the corresponding measure algebras E_μ and E_ν are isomorphic in the sense of Definition* 9.3.1, *then there exists an isomorphism* mod0 *between these measure spaces. In particular, this is the case if both measures are atomless.*

The above theorem does not extend to arbitrary topological spaces even with Radon measures.

9.5.3. Example. Let X be the space "two arrows of P.S. Alexandroff" defined in Example 7.14.11 (this space is compact) with its natural normalized Lebesgue measure μ, described in that example. Then the corresponding measure algebra is atomless and separable, therefore, is metric Boolean isomorphic to the measure algebra of the unit interval. However, there exists no isomorphism mod0 between the two spaces.

PROOF. This follows by Corollary 9.2.3, taking into account that metrizable subsets of X are at most countable and the measure λ vanishes on them. $\square$

For additional results, see §9.11(iv).

9.6. Topologically equivalent measures

9.6.1. Definition. *Let X and Y be two topological spaces, let μ be a Borel measure on X, and let ν be a Borel measure on Y.*

(i) *The measure spaces $\bigl(X, \mathcal{B}(X), \mu\bigr)$ and $\bigl(Y, \mathcal{B}(Y), \nu\bigr)$ are called homeomorphic if there exists a homeomorphism $h\colon X \to Y$ with $\mu \circ h^{-1} = \nu$.*

(ii) *The measure spaces $\bigl(X, \mathcal{B}(X), \mu\bigr)$ and $\bigl(Y, \mathcal{B}(Y), \nu\bigr)$ are called almost homeomorphic (or topologically equivalent) if there exist sets $N \subset X$, $N' \subset Y$ with*

$$|\mu|(N) = |\nu|(N') = 0$$

and a homeomorphism $h\colon X\backslash N \to Y\backslash N'$ such that $\mu \circ h^{-1} = \nu$.

Measures that are almost homeomorphic to Lebesgue measure are called topologically Lebesgue.

The next two important results on homeomorphisms of measure spaces are due to Oxtoby [**1408**].

9.6.2. Theorem. *Let X be a topological space equipped with a Borel probability measure μ that has no atoms and is positive on nonempty open sets. In order that the space (X, μ) be homeomorphic to $(\mathcal{R}, \lambda)$, where $\mathcal{R}$ is the space of all irrational numbers in the interval $(0, 1)$ and λ is Lebesgue measure, it is necessary and sufficient that X be homeomorphic to $\mathcal{R}$.*

PROOF. We have to show that if X and $\mathcal{R}$ are homeomorphic, then there exists a homeomorphism transforming μ into λ. Hence we may assume that $X = \mathcal{R}$. One can introduce a metric d on $\mathcal{R}$ defining the usual topology, but making $\mathcal{R}$ a complete space (see §6.1).

Let us prove the following auxiliary assertion: if U and V are nonempty open sets in $\mathcal{R}$ such that $\mu(U) = \lambda(V)$, then for every $\varepsilon > 0$, there exists a partition $\{U_n\}$ of the set U and a partition $\{V_n\}$ of the set V into nonempty open sets of diameter less than ε in the metric d such that $\mu(U_n) = \lambda(V_n)$ for all n. To this end, we take a partition of V into nonempty open sets W_i of diameter less than ε in the metric d. We observe that V is homeomorphic to $\mathcal{R}$ (this is clear from the fact that V is the intersection of $\mathcal{R}$ with a finite or countable union of disjoint intervals). Then by Exercise 9.12.41, there exists a partition of U into open sets G_i with $\lambda(W_i) = \mu(G_i)$ for all $i \in \mathbb{N}$. Every G_i can be partitioned into nonempty pairwise disjoint open sets G_{ij} of diameter less than ε in the metric d. As above, every W_i can be partitioned into open sets W_{ij} with $\lambda(W_{ij}) = \mu(G_{ij})$ for all $j \in \mathbb{N}$. The families $\{G_{ij}\}$ and $\{W_{ij}\}$ are desired partitions.

By this auxiliary assertion, we obtain partitions $\mathcal{U}_n = \{U(i_1, \ldots, i_n)\}$ and $\mathcal{V}_n = \{V(i_1, \ldots, i_n)\}$ consisting of nonempty open sets of diameter less than $1/n$ in the metric d, with the following properties:

$$U(i_1, \ldots, i_{n+1}) \subset U(i_1, \ldots, i_n), \ V(i_1, \ldots, i_{n+1}) \subset V(i_1, \ldots, i_n),$$

$\mu\bigl(U(i_1, \ldots, i_n)\bigr) = \lambda\bigl(V(i_1, \ldots, i_n)\bigr)$ for all $n \in \mathbb{N}$ and all indices i_j. For every $x \in \mathcal{R}$, there exists exactly one sequence $\{i_n\}$ with $x \in \bigcap_{n=1}^{\infty} U(i_1, \ldots, i_n)$. The same is true for the family of sets $V(i_1, \ldots, i_n)$. Let

$$f(x) = \bigcap_{n=1}^{\infty} V(i_1, \ldots, i_n).$$

Then f is a one-to-one mapping of $\mathcal{R}$ onto itself. One has the equality $f\bigl(U(i_1, \ldots, i_n)\bigr) = V(i_1, \ldots, i_n)$, which is readily verified by using the above-stated property of both families of sets. We observe that any nonempty open set in $\mathcal{R}$ can be represented as a finite or countable union of pairwise disjoint sets in the partitions $\mathcal{V}_n$. The same is true for the partitions $\mathcal{U}_n$. This yields that f is a homeomorphism and $\mu\bigl(f^{-1}(W)\bigr) = \lambda(W)$ for every open set $W \subset \mathcal{R}$. Hence $\mu \circ f^{-1} = \lambda$. □

9.6.3. Theorem. *Let μ be a Borel probability measure on a Polish space X without points of positive measure. Then, there is a G_δ-set $Y \subset X$ such that $\mu(X\backslash Y) = 0$ and the space (Y, μ_Y) is homeomorphic to the space $\mathcal{R}$ of irrational numbers of the interval $(0,1)$ with Lebesgue measure λ. In particular, (X, μ) and $([0,1], \lambda)$ are almost homeomorphic.*

Proof. Let d be a complete metric on X. Let us take a countable everywhere dense set $\{x_i\} \subset X$ and a sequence of numbers $r_j > 0$ such that $\lim\limits_{j\to\infty} r_j = 0$ and $\mu\bigl(\{x\colon\ d(x, x_i) = r_j\}\bigr) = 0$. This is possible, since for every i, the set of numbers r such that $\mu\bigl(\{x\colon\ d(x, x_i) = r\}\bigr) > 0$ is at most countable. Let

$$S_{ij} = \{x\colon\ d(x, x_i) = r_j\} \quad \text{and} \quad U_{ij} = \{x\colon\ d(x, x_i) < r_j\}.$$

It is clear that the collection $\{U_{ij}\}$ forms a topology base. We denote by S the union of all S_{ij}, and by G the union of all U_{ij} with $\mu(U_{ij}) = 0$. Then $\mu(S \cup G) = 0$. Let us consider the set $Z = X\backslash(S \cup G)$. It is clear that $\mu(Z) = 1$ and that Z can be represented as a countable intersection of open sets, i.e., is a G_δ-set (we recall that any closed set in a metric space is G_δ). We take in Z a countable everywhere dense set D. Let us show that the set $Y = Z\backslash D$ is as required. Indeed, $\mu(Y) = 1$. If U is an open set meeting Y, then $\mu(U) > 0$, since otherwise one could find a set U_{ij} of zero measure meeting Y. Thus, the measure μ_Y on the space Y is positive on nonempty open subsets of Y. In addition, μ_Y has no points of positive measure. In order to apply Theorem 9.6.2, it remains to verify that Y is homeomorphic to $\mathcal{R}$. By construction, Y is an everywhere dense G_δ-set in the space Z. By the Mazurkiewicz theorem (see Kuratowski [**1082**, §36, subsection II, Theorem 3]), it suffices to verify that $Z\backslash Y$ is everywhere dense in Z and that Z is a Polish space of zero dimension, i.e., every point has an arbitrarily small clopen neighborhood. Since Z is a G_δ-set in a Polish space, it is Polish as well. The set $D = Z\backslash Y$ is dense in Z by construction. Finally, the fact that Z has dimension zero follows by the property that the sets $U_{ij} \cap Z$ are closed in Z, since all sets S_{ij} are deleted from Z. □

Additional results on almost homeomorphisms are given in §9.12(vi).

We remark that there exists a Radon probability measure μ on a compact space X such that the space (X, μ) is isomorphic mod0 to the interval $[0, 1]$ with Lebesgue measure, but is not almost homeomorphic to $[0, 1]$ (hence to no compact metric space); see Exercise 9.12.60.

The following criterion of the existence of almost homeomorphisms is proved in Babiker [**85**].

9.6.4. Theorem. *Let μ be a Radon probability measure on a compact space X such that (X, μ) is isomorphic* mod0 *to the interval* $[0, 1]$ *with Lebesgue measure. Then, the measure μ is topologically Lebesgue if and only if it is completion regular on its topological support S_μ, i.e., $\mathcal{B}(S_\mu) \subset \mathcal{B}a(S_\mu)_\mu$.*

Finally, we mention two results on usual homeomorphisms of topological spaces with measures.

9.6.5. Theorem. *A Borel probability measure μ on the cube $[0, 1]^n$ is homeomorphic to Lebesgue measure λ on $[0, 1]^n$ if and only if it satisfies the following conditions*: (a) *μ is atomless*; (b) *μ is positive on all nonempty open sets in $[0, 1]^n$*; (c) *μ vanishes on the boundary of $[0, 1]^n$.*

9.6.6. Theorem. (i) *A Borel probability measure μ on $[0, 1]^\infty$ is homeomorphic to the measure λ^∞ on $[0, 1]^\infty$ that is the countable product of Lebesgue measures if and only if it is atomless and positive on all nonempty open sets in $[0, 1]^\infty$.*

(ii) *Every two atomless Borel probability measures on l^2, positive on all nonempty open sets, are homeomorphic.*

Further information, including references and proofs, can be found in Alpern, Prasad [**38**] and Akin [**17**].

9.7. Continuous images of Lebesgue measure

In this section, we discuss the following question: when can a measure μ on a topological space X be represented as the image of Lebesgue measure on the interval $[0,1]$ under a continuous mapping from $[0,1]$ to X? A simple answer to this question in terms of the topological support of the measure has been given by Kolesnikov [**1018**]. In order to formulate the principal result, we need the notions of connectedness and local connectedness.

We recall that a nonempty open set in a topological space is called connected if it cannot be represented as the union of two disjoint nonempty open sets. A topological space is called locally connected at a point x if every open neighborhood of the point x contains its connected neighborhood. A topological space is called locally connected if it is locally connected at every point.

It is known that a metrizable compact space is a continuous image of the interval precisely when it is connected and locally connected (see Engelking [**532**, 6.3.14]).

If a measure μ on a space X is the image of Lebesgue measure under a continuous mapping $f\colon [0,1] \to X$, then the topological support of μ (i.e., the smallest closed set of full measure) is the compact set $K = f([0,1])$, hence is a connected and locally connected metrizable compact space. It turns out that the converse is true as well. It should be observed, however, that if the support of a measure μ is the image of the interval $[0,1]$ under some continuous mapping φ, then this does not mean that μ is the image of Lebesgue measure under φ. For example, let $\varphi(t) = 0$ if $t \le 1/2$ and $\varphi(t) = 2(t-1/2)$ if $t \ge 1/2$. Then the image of Lebesgue measure with respect to φ does not coincide with Lebesgue measure, although $\varphi([0,1]) = [0,1]$.

9.7.1. Theorem. *Let K be a compact metric space that is the image of $[0,1]$ under a continuous mapping f and let μ be a Borel probability measure on K such that K is its support. Then, there exists a continuous mapping $g\colon [0,1] \to K$ such that $\mu = \lambda \circ g^{-1}$, where λ is Lebesgue measure on $[0,1]$.*

PROOF. (1) First we show that every point $y \in K$ has an arbitrarily small closed neighborhood that is a continuous image of $[0,1]$. Let $[0,1] = A_1 \cup A_2 \cup \ldots \cup A_n$, where $A_k = [(k-1)/n, k/n]$. Let $U = \bigcup_{k\colon\, y\in f(A_k)} f(A_k)$. It is clear that U is a continuous image of $[0,1]$. This set is a closed neighborhood, since, along with y, it contains the open set $K \setminus \bigcup_{k\colon\, y\notin f(A_k)} f(A_k)$, which follows by the equality $K = f([0,1])$. By the uniform continuity of f and the triangle inequality, the neighborhood U can be made as small as we wish.

(2) The main step of the proof is the verification of the existence of a continuous mapping φ from $[0,1]$ onto K such that $\mu\bigl(\varphi(V)\bigr) \neq 0$ for every nonempty open set $V \subset [0,1]$. Let U_0 be the union of all open sets that are taken by f to measure zero sets. Then $U_0 = \bigcup_{i=1}^{\infty} J_i$, where $J_i = (a_i, b_i)$, or $J_i = [0, b_i)$, or $J_i = (a_i, 1]$, and $J_i \cap J_j = \varnothing$ if $i \neq j$. We assume further that the length of J_i does not increase as i is increasing. Let m_1 be the smallest natural number for which there exist intervals J_1, $J_2, \ldots,$ J_{k_1} of length not less than $1/2^{m_1}$ (i.e., at least one such interval). For the middle point c_1 of the interval J_1 we find N such that the neighborhood V_1 of the point $f(c_1)$ constructed for the case $n = N$ as in the previous step is of diameter less than $1/2$ (in the metric ϱ_K of K). In addition, we may assume that $(c_1 - 1/N, c_1 + 1/N)$

belongs to J_1. We can construct a continuous surjective mapping f_1 of the interval $[c_1 - 1/N, c_1 + 1/N]$ onto V_1 such that

$$f_1(c_1 - 1/N) = f(c_1 - 1/N), \quad f_1(c_1 + 1/N) = f(c_1 + 1/N).$$

Similarly, for all $1 \le i \le k_1$, we can construct mappings f_i in neighborhoods of the points c_i, where c_i is the middle point of J_i. Let the mapping φ_1 coincide with f outside these neighborhoods and coincide with f_i on the corresponding neighborhood. Then φ_1 is continuous and $\varrho(f, \varphi_1) \le 1/2$, where $\varrho(\varphi, \psi) = \sup_{t \in [0,1]} \varrho_K\bigl(\varphi(t), \psi(t)\bigr)$. The largest open set taken by φ_1 to a measure zero set does not contain intervals of length greater than $1/2^{m_1}$, since $\mu(V_i) \ne 0$, where $1 \le i \le k_1$. Let us pick intervals J_{k_1+1}, $J_{k_1+2}, \ldots$, J_{k_2} of length greater than $1/2^{m_2}$, where $m_2 > m_1$ is the smallest natural number in $(m_1, +\infty)$ for which this is possible. As above, we construct a continuous mapping φ_2 such that $\varrho(\varphi_1, \varphi_2) \le 1/4$. Repeating the described construction countably or finitely many times, we obtain a sequence of continuous mappings φ_n such that $\varrho(\varphi_n, \varphi_{n+1}) \le 1/2^{n+1}$. In the limit we obtain a continuous mapping φ. Let $X_0 = [0,1] \setminus U_0$, $X_i = [0,1] \setminus U_i$, where U_i is the largest open set taken by φ_i to a measure zero set. By construction, every mapping φ_j coincides on X_i with φ_i whenever $j > i$, and the set $\bigcup_{i=1}^{\infty} X_i$ is everywhere dense in $[0,1]$. Hence the mapping φ takes nonempty open sets to sets of positive measure. Finally, $\varphi([0,1]) = K$, since already $f(X_0) = K$ (otherwise one would obtain a nonempty open set in K of zero μ-measure).

(3) For completing the proof it remains to apply Corollary 9.1.6 and the following simple fact: any Borel probability measure ν on the interval $[0,1]$ with support $[0,1]$ is the image of Lebesgue measure on $[0,1]$ under some continuous surjective mapping $\zeta\colon [0,1] \to [0,1]$. One can take

$$\zeta(t) = \sup_{x \in [0,1]} \bigl\{x\colon F(x) \le t\bigr\}$$

for such a mapping, where $F(t) = \nu\bigl([0,t)\bigr)$, $F(0) = 0$. It is clear that ζ is increasing, $\zeta(0) = 0$ (since $F(t) > 0$ if $t > 0$) and $\zeta(1) = 1$. It follows by the strict increasing of F that the function ζ has no jumps, hence is continuous. The fact that the image of Lebesgue measure with respect to ζ is the measure μ is verified in the same manner as in the proof of Theorem 9.2.2. □

9.7.2. Corollary. *The continuous images of Lebesgue measure on* $[0,1]$ *are precisely the Radon probability measures whose topological supports are connected and locally connected metrizable compact sets.*

9.7.3. Corollary. *Let μ be a Radon probability measure whose topological support is a connected and locally connected metrizable compact space and let ν be an atomless Radon probability measure on a compact space. Then μ is a continuous image of ν.*

Proof. We apply Proposition 9.1.12 and the above theorem. □

9.7.4. Remark. The requirement of continuity of mappings in the above considerations is, of course, an essential restriction. If we admit Borel mappings, then every Borel probability measure μ on a Souslin space X can be obtained as the image of Lebesgue measure on $[0,1]$ under a Borel mapping. This is obvious from the isomorphism theorem for atomless measures and the fact that any measure concentrated at a point is obtained by means of a constant mapping. If X is realized as a subset in $[0,1]$, then a required mapping can be defined by the following explicit formula:

$$f_\mu(x) := \inf\bigl\{t\colon\ F_\mu(t) \geq x\bigr\},$$

where F_μ is the distribution function of μ. If ν is an atomless Borel probability measure on $[0,1]$, then there is a natural monotone function φ such that $\mu = \nu \circ \varphi^{-1}$, namely, $\varphi := f_\mu \circ F_\nu$.

9.8. Connections with extensions of measures

We have already discussed the problem of extending measures. In particular, it has been shown that one can always extend a measure to the σ-algebra obtained by adding a single set or even a family of disjoint sets. In this section, we show that a measure on a countably generated sub-σ-algebra of the Borel σ-algebra of a Souslin space can be extended to the whole Borel σ-algebra. This problem is connected with finding preimages of measures.

Let $(X,\mathcal{A})$ and $(Y,\mathcal{B})$ be measurable spaces and let $f\colon X \to Y$ be an $(\mathcal{A},\mathcal{B})$-measurable mapping. Suppose we are given a probability measure ν on $\mathcal{B}$ such that $\nu^*\bigl(f(X)\bigr) = 1$. Then we obtain a probability measure on the σ-algebra $f^{-1}(\mathcal{B}) := \{f^{-1}(B)\colon\ B \in \mathcal{B}\}$ defined by the formula

$$\nu_0\bigl(f^{-1}(B)\bigr) := \nu(B).$$

Note that ν_0 is well-defined: if $B_1, B_2 \in \mathcal{B}$ and $f^{-1}(B_1) = f^{-1}(B_2)$, then $\nu(B_1) = \nu(B_2)$. Let a probability measure μ on $\mathcal{A}$ be a preimage of ν, i.e., $\nu = \mu \circ f^{-1}$. Clearly, $\nu^*\bigl(f(X)\bigr) = 1$. It follows that μ is an extension of ν_0 to the whole σ-algebra $\mathcal{A}$. Conversely, any extension of ν_0 to $\mathcal{A}$ is a preimage of ν; the uniqueness of extension corresponds to the uniqueness of a preimage.

Now we give an example where a separable measure on a sub-σ-algebra in the Borel σ-algebra of an interval has no Borel extensions.

9.8.1. Example. Let $\mathcal{A}$ be the class of all first category Borel sets (i.e., countable unions of nowhere dense sets) in the interval $[0,1]$ and their complements. Let $\mu(A) = 0$ if A is a first category Borel set and $\mu(A) = 1$ if A is the complement of such a set. Then $\mathcal{A}$ is a σ-algebra and μ is a countably additive measure (since in any collection of disjoint sets in $\mathcal{A}$, at most one can have a nonzero measure). The measure μ is separable on $\mathcal{A}$ (every set of positive μ-measure has μ-measure 1 and hence up to a measure zero set coincides with $[0,1]$). However, there exists no countably additive extension of μ to the Borel σ-algebra of the interval. Indeed, according to Exercise 1.12.50, every Borel measure on an interval is concentrated on a first category set. Another close example is described in Exercise 9.12.49.

Now we show that slightly strengthening our requirements on the σ-algebra we obtain a positive result.

9.8.2. Theorem. *Let X be a Souslin space and let $\mathcal{A}$ be a countably generated sub-σ-algebra in $\mathcal{B}(X)$. Then, every measure μ on $\mathcal{A}$ can be extended to a measure on $\mathcal{B}(X)$.*

Proof. We know that the countably generated σ-algebra $\mathcal{A}$ has the form $f^{-1}\bigl(\mathcal{B}([0,1])\bigr)$, where $f\colon X \to [0,1]$ is some function. Since $\mathcal{A} \subset \mathcal{B}(X)$, the function f is Borel measurable. Therefore, $f(X)$ is a Souslin set. By Theorem 9.1.5 there exists a Borel measure μ_0 on X with $\mu_0 \circ f^{-1} = \mu \circ f^{-1}$. Let us verify that μ_0 is an extension of μ. Indeed, if $A = f^{-1}(B)$, where $B \in \mathcal{B}([0,1])$, then

$$\mu_0(A) = \mu_0\bigl(f^{-1}(B)\bigr) = \mu_0 \circ f^{-1}(B) = \mu \circ f^{-1}(B) = \mu(A),$$

as required. □

9.8.3. Corollary. *Let X be a Souslin space and let a measure μ be defined on some σ-algebra $\mathcal{A} \subset \mathcal{B}(X)$. Suppose that there exists a countable collection of sets $A_n \in \mathcal{A}$ with $\mathcal{A} \subset \sigma(\{A_n\})_\mu$. Then, the measure μ can be extended to a measure on $\mathcal{B}(X)$.*

Proof. By the above theorem μ extends from $\sigma(\{A_n\})$ to $\mathcal{B}(X)$. This extension $\widetilde{\mu}$ coincides on $\mathcal{A}$ with the initial measure, since for every $A \in \mathcal{A}$ by our hypothesis there exist two sets $B_1, B_2 \in \sigma(\{A_n\})$ with $B_1 \subset A \subset B_2$ and $|\mu|(B_2 \backslash B_1) = 0$. □

Let us now turn to the problem of uniqueness of extensions.

9.8.4. Proposition. *In the situation of Theorem* 9.8.2*, the measure μ uniquely extends to $\mathcal{B}(X)$ precisely when $\mathcal{B}(X) \subset \mathcal{A}_\mu$.*

Proof. The only thing that is not obvious is that there exist at least two different extensions in the case where $\mathcal{B}(X)$ is not covered by the Lebesgue completion of μ. In this case, there exists a set $B \in \mathcal{B}(X)$ that does not belong to $\mathcal{A}_\mu$. Therefore, the set B has distinct inner and outer measures corresponding to μ on $\mathcal{A}$. By Theorem 1.12.14, there exist two different extensions of μ to the σ-algebra generated by $\mathcal{A}$ and the set B. As shown above, both extensions can be further extended to Borel measures. □

Additional remarks on uniqueness of extension are made in §9.12(ii).

9.9. Absolute continuity of the images of measures

In this section, we consider only bounded measures. We note that although every Borel mapping between Souslin spaces takes every Borel set to a Souslin (hence universally measurable) set, it may occur even for continuous functions on the real line that the image of a Lebesgue measurable set is not Lebesgue measurable. For example, if C_0 is the Cantor function, then the function $h(x) = \frac{1}{2}\bigl(x + C(x_0)\bigr)$ is a homeomorphism of $[0,1]$ that takes certain sets of Lebesgue measure zero to nonmeasurable sets. In order to characterize mappings taking all measurable sets to measurable ones, we consider Lusin's property (N) already encountered in §3.6 in Chapter 3 and studied in several exercises in Chapter 5, where, in particular, it is shown that absolutely continuous functions have property (N). We recall the definition and a result from Chapter 3.

9.9.1. Definition. *Let μ be a finite measure on a measurable space $(M,\mathcal{M})$. A mapping $F\colon M_0\subset M\to M$ is said to satisfy Lusin's condition* (N) *on M_0 (or to have Lusin's property* (N)) *if for every set $Z\subset M_0$ such that $Z\in\mathcal{M}$ and $|\mu|(Z)=0$, one has $F(Z)\in\mathcal{M}_\mu$ and $|\mu|\bigl(F(Z)\bigr)=0$.*

It is clear that if F satisfies Lusin's condition (N) on M, then for every $|\mu|$-zero set Z in $\mathcal{M}_\mu$ (not necessarily in $\mathcal{M}$) we have $|\mu|\bigl(F(Z)\bigr)=0$, since Z is contained in some $|\mu|$-zero set from $\mathcal{M}$.

Note that when we say that F has property (N) on M, the mapping F is supposed to be defined everywhere. Unlike many other properties of measurable mappings, property (N) may not be preserved when changing a function on a set of measure zero. For example, the identically zero function on $[0,1]$ can be redefined on a set C of measure zero and cardinality of the continuum so that it will map C onto $[0,1]$.

9.9.2. Remark. By analogy one defines Lusin's (N)-property for mappings $F\colon(M_1,\mathcal{M}_1,\mu_1)\to(M_2,\mathcal{M}_2,\mu_2)$ between two measure spaces: it is required that the equality $|\mu_2|\bigl(F(Z)\bigr)=0$ be true if $|\mu_1|(Z)=0$.

The next result has already been proved in Theorem 3.6.9 in Chapter 3 for mappings on $\mathbb{R}^n$. Clearly, the same is true for mappings on measurable sets.

9.9.3. Theorem. *Let $S\subset\mathbb{R}^1$ be a measurable set equipped with Lebesgue measure μ and let F be a measurable function on S. Then F satisfies Lusin's condition* (N) *if and only if F takes every Lebesgue measurable subset of S to a measurable set.*

9.9.4. Corollary. *Let $(M,\mathcal{M},\mu)$ be a measure space that is isomorphic* mod0 *to a measurable set $S\subset\mathbb{R}^1$ with Lebesgue measure. The following conditions are equivalent for any $(\mathcal{M}_\mu,\mathcal{M})$-measurable mapping $F\colon M\to M$:*

(i) *F satisfies Lusin's condition* (N);

(ii) *F takes every μ-measurable subset of M to a μ-measurable set* (*in other words, $F(\mathcal{M}_\mu)\subset\mathcal{M}_\mu$*).

In particular, this equivalence holds if M is a Souslin space with an atomless Borel measure μ and $\mathcal{M}=\mathcal{B}(M)$.

Proof. Let $h\colon(M,\mathcal{M},\mu)\to(S,\mathcal{L},\lambda)$ be an isomorphism mod0, where λ is Lebesgue measure. We may assume that the function h is defined on a set M_0 of full μ-measure and maps it one-to-one onto the set S with the preservation of measure. We set $g(s)=h^{-1}(s)$ and define h on $M\backslash M_0$ by zero. Let $G(s)=h\bigl(F\bigl(g(s)\bigr)\bigr)$. If F has property (N), then G also does, since g takes sets of λ-measure zero to sets of μ-measure zero and h takes sets of μ-measure zero to sets of λ-measure zero. Therefore, G takes Lebesgue measurable sets to measurable ones. For every $A \in \mathcal{M}_\mu$ we have $F(A) = F(Z) \cup F(M_0 \cap A)$, $Z = (M \backslash M_0) \cap A$. By hypothesis, $F(Z)$ has μ-measure zero. In addition,

$$F(M_0 \cap A) \cap M_0 = g\bigl(G\bigl(h(M_0 \cap A)\bigr)\bigr) \cap M_0.$$

This equality yields the μ-measurability of $F(M_0 \cap A)$, since M_0 has full μ-measure and the mappings h, G, and g take measurable sets to measurable ones (with respect to the corresponding measures). Thus, (i) implies (ii). The converse is proved as in the theorem, since we only need that every set of positive μ-measure have a nonmeasurable subset, which follows by the existence of an isomorphism mod0 with a Lebesgue measurable set in $\mathbb{R}^1$. $\square$

Regarding property (N), see also Exercise 9.12.44, 9.12.46.

The above equivalence may fail in the case where there are atoms: it suffices to take the measure μ on the set consisting of two points 0 and 1 such that $\mu(\{0\}) = 0$, $\mu(\{1\}) = 1$ and the function $F \equiv 1$. This function takes all sets to measurable ones, but the point of zero measure is taken to the point of positive measure.

9.9.5. Proposition. *Let a mapping* $F\colon (M, \mathcal{M}, \mu) \to (M, \mathcal{M}, \mu)$ *have Lusin's property* (N) *and be* $(\mathcal{M}_\mu, \mathcal{M})$*-measurable. Then, for every* μ*-measurable subset* $A \subset F(M)$*, the measure* $I_A \cdot \mu$ *is absolutely continuous with respect to the measure* $|\mu| \circ F^{-1}$*. If* $F(M) \in \mathcal{M}_{|\mu| \circ F^{-1}}$*, then* $F(M) \in \mathcal{M}_\mu$ *and* $\mu|_{F(M)} \ll \bigl(|\mu| \circ F^{-1}\bigr)|_{F(M)}$*.*

Proof. Let $B \subset F(M)$, $B \in \mathcal{M}$, $|\mu|\bigl(F^{-1}(B)\bigr) = 0$. Then $|\mu|(B) = 0$, since $B = F\bigl(F^{-1}(B)\bigr)$. Therefore, given a set $A \in \mathcal{M}_\mu$ that is contained in $F(M)$ and a set $B \in \mathcal{M}$ with $|\mu| \circ F^{-1}(B) = 0$, we find a set $E \subset A$ with $E \in \mathcal{M}$ and $|\mu|(A \backslash E) = 0$. Hence $|\mu|(A \cap B) = 0$, since $|\mu|\bigl(F^{-1}(B \cap E)\bigr) = 0$. Thus, $\mu \ll |\mu| \circ F^{-1}$ on μ-measurable sets in $F(M)$. If $F(M) \in \mathcal{M}_{|\mu| \circ F^{-1}}$, then we can find sets $E_1, E_2 \in \mathcal{M}$ such that $E_1 \subset F(M)$, $F(M) \subset E_1 \cup E_2$ and $|\mu| \circ F^{-1}(E_2) = 0$. By property (N) we have $|\mu|\bigl(F(F^{-1}(E_2))\bigr) = 0$, hence $|\mu|\bigl(E_2 \cap F(M)\bigr) = 0$, which means that $F(M) \in \mathcal{M}_\mu$. $\square$

Certainly, one does not always have $\mu \ll \mu \circ F^{-1}$ on the whole space. For example, one can take $F \equiv 0$ on $[0, 1]$ with Lebesgue measure.

9.9.6. Corollary. *Let* μ *be a finite nonnegative measure on* $(M, \mathcal{M})$ *and let* $F\colon M \to M$ *be a one-to-one* $(\mathcal{M}_\mu, \mathcal{M})$*-measurable mapping such that* $F(\mathcal{M}) \subset \mathcal{M}_{\mu \circ F^{-1}}$ *(or, more generally,* F *has a modification* $\widetilde{F}$ *such that* $\widetilde{F}(\mathcal{M}) \subset \mathcal{M}_{\mu \circ F^{-1}}$*). Then, the condition* $\mu \ll \mu \circ F^{-1}$ *is equivalent to Lusin's condition* (N)*.*

In particular, such an equivalence holds for one-to-one Borel mappings between Souslin spaces.

PROOF. Suppose that $\mu \ll \mu\circ F^{-1}$. We show that F has property (N). Observe that for any set $E\in\mathcal{M}_{\mu\circ F^{-1}}$ one has $\mu\circ F^{-1}(E)=\mu\bigl(F^{-1}(E)\bigr)$, since there exist sets $E_1,E_2\in\mathcal{M}$ such that $E_1\subset E\subset E_2$ and the equality $\mu\circ F^{-1}(E_1)=\mu\circ F^{-1}(E)=\mu\circ F^{-1}(E_2)$ holds. Let $B\in\mathcal{M}$ be such that $\mu(B)=0$. Since F is bijective, we have $B=F^{-1}\bigl(F(B)\bigr)$. Suppose that $F(B)\in\mathcal{M}_{\mu\circ F^{-1}}$. Then $\mu\circ F^{-1}\bigl(F(B)\bigr)=\mu(B)=0$. Hence $\mu\bigl(F(B)\bigr)=0$. Suppose now that the $\mu\circ F^{-1}$-measurability of $\widetilde{F}(B)$ is given just for some modification $\widetilde{F}$ with $\widetilde{F}(\mathcal{M})\subset\mathcal{M}_{\mu\circ F^{-1}}$. Take a set $M_0\in\mathcal{M}$ of full μ-measure on which $F=\widetilde{F}$. Let $M_1=M\backslash M_0$. Then $F(B\cap M_0)=\widetilde{F}(B\cap M_0)$ belongs to $\mathcal{M}_{\mu\circ F^{-1}}$. By the previous step we have $\mu\bigl(F(B\cap M_0)\bigr)=0$. It remains to show that $\mu\bigl(F(M_1)\bigr)=0$. This follows by the equality $\mu\circ F^{-1}\bigl(F(M_1)\bigr)=0$, which is clear from the fact that $F(M_0)=\widetilde{F}(M_0)\in\mathcal{M}_{\mu\circ F^{-1}}$ is a full measure set for $\mu\circ F^{-1}$ and $F(M_1)=X\backslash F(M_0)$, since F is one-to-one.

The converse has already been proven. The last claim is clear from the fact that the images of Borel sets in Souslin spaces under Borel mappings are measurable with respect to all Borel measures. $\square$

9.9.7. Lemma. *Let $(M,\mathcal{M},\mu)$ be a measure space and let $T\colon M\to M$ be a $(\mathcal{M}_\mu,\mathcal{M})$-measurable mapping such that the sets $T(N)$ and $T^{-1}(N)$ have measure zero for every set N of measure zero. Suppose that there exists a μ-measurable mapping S such that $T\bigl(S(x)\bigr)=S\bigl(T(x)\bigr)=x$ for μ-a.e. x. Then, there exists a set M_0 of full μ-measure such that T maps M_0 one-to-one onto itself (and S is its inverse) and $T(M\backslash M_0)\subset M\backslash M_0$.*

PROOF. The hypotheses yield that $\mu\circ T^{-1}\sim\mu$. Denote by Ω_0 the set of all points x such that $T\bigl(S(x)\bigr)=S\bigl(T(x)\bigr)=x$. The mappings T and S are obviously injective on Ω_0. Let $\Delta=M\backslash\Omega_0$. By hypothesis, $T(\Delta)$ has measure zero. Since $\mu\sim\mu\circ T^{-1}$, the set $T^{-1}\bigl(T(\Delta)\bigr)$ has measure zero as well. Hence T is a one-to-one mapping of the full measure set $\Omega_1=\Omega_0\backslash T^{-1}\bigl(T(\Delta)\bigr)$ and $T(\Omega_1)$. In addition, the complement of Ω_1 is taken to the complement of the set $T(\Omega_1)$. Since $\mu\sim\mu\circ T^{-1}$ and $\mu\bigl(T^{-1}(\Omega_1)\bigr)=\mu\circ T^{-1}(\Omega_1)$, the set $T^{-1}(\Omega_1)$ has full measure. Let $Z_0=\Omega_1\cap T^{-1}(\Omega_1)$. On the set Z_0 of full measure T is injective, $S\bigl(T(Z_0)\bigr)\subset\Omega_1$ and $S\bigl(T(x)\bigr)=x$. Hence for every $B\subset Z_0$ with $B\in\mathcal{M}$ we have $T(B)=S^{-1}(B)\in\mathcal{M}_\mu$. Since T takes sets of measure zero to sets of measure zero, this yields that T takes μ-measurable sets to μ-measurable sets. By the equivalence of the measures μ and $\mu\circ T^{-1}$, one can conclude that sets of full measure are taken to sets of full measure. For all integer k we define inductively sets Z_k by the equalities $Z_{k+1}=Z_0\cap T(Z_k)$ if $k\ge 0$, $Z_{k-1}=Z_0\cap T^{-1}(Z_k)$ if $k\le 0$. It follows from the above that the sets Z_k have full measure. Now let $\Omega=\bigcap_k Z_k$. This set has full measure. It is verified directly that T maps it one-to-one onto itself, whereas $T(M\backslash\Omega)\subset M\backslash\Omega$. $\square$

The following assertion has been obtained in the course of the proof.

9.9.8. Corollary. *The mapping T in Lemma 9.9.7 takes all μ-measurable sets to μ-measurable sets.*

9.9.9. Lemma. *Let T be a one-to-one mapping of a measure space $(M, \mathcal{M}, \mu)$ such that the mappings T and $S = T^{-1}$ are $(\mathcal{M}_\mu, \mathcal{M})$-measurable and $\mu \circ T^{-1} \sim \mu$. Then*

$$d(\mu \circ S^{-1})/d\mu = \frac{1}{\varrho \circ T}, \quad \text{where } \varrho = d(\mu \circ T^{-1})/d\mu.$$

PROOF. Since $T = S^{-1}$, one has $T(B) = S^{-1}(B) \in \mathcal{M}_\mu$ provided that $B \in \mathcal{M}$ and $\mu \circ S^{-1}(B) = \mu\bigl(T(B)\bigr)$. We observe that

$$\mu\bigl(T(B)\bigr) = \int_{T(B)} \frac{1}{\varrho}\, d(\mu \circ T^{-1}) = \int_M I_{T(B)} \circ T \,\frac{1}{\varrho \circ T}\, d\mu.$$

Since $I_{T(B)} \circ T = I_B$, the claim follows. □

9.9.10. Proposition. *Let μ be a measure on a measurable space $(X, \mathcal{A})$, let ν be a Radon probability measure on a completely regular space Y, and let $T, T_n\colon X \to Y$ be $\bigl(\mathcal{A}_{|\mu|}, \mathcal{B}(Y)\bigr)$-measurable mappings such that the sequence $\{T_n(x)\}$ converges $|\mu|$-a.e. to $T(x)$. Let us assume that $\mu \circ T^{-1}$ is a Radon measure, the measures $\mu \circ T_n^{-1}$ are absolutely continuous with respect to ν and that their Radon–Nikodym densities ϱ_n form a uniformly integrable sequence. Then, the measure $\mu \circ T^{-1}$ is absolutely continuous with respect to ν and its Radon–Nikodym density ϱ is the limit of the sequence $\{\varrho_n\}$ in the weak topology of the space $L^1(\nu)$.*

PROOF. Let K be a compact set of ν-measure zero and let $\varepsilon > 0$. Suppose that $|\mu \circ T^{-1}(K)| > \varepsilon$. We may assume that $\mu \circ T^{-1}(K) > \varepsilon$. The uniform integrability ensures the existence of $\delta > 0$ such that

$$\int_A |\varrho_n(y)|\, \nu(dy) \le \frac{\varepsilon}{2}, \quad \forall\, n \in \mathbb{N},$$

for every measurable set A with $\nu(A) \le \delta$. Let us find an open set $U \supset K$ with $\nu(U) < \delta$ and $|\mu \circ T^{-1}|(U \backslash K) < \varepsilon/2$. By Lemma 6.1.5, there exists a continuous function $f\colon Y \to [0,1]$ that equals 1 on K and 0 outside U. Then we have

$$\begin{aligned}\int_Y f(y)\, \mu \circ T^{-1}(dy) &= \int_X f\bigl(T(x)\bigr)\, \mu(dx) = \lim_{n\to\infty} \int_X f\bigl(T_n(x)\bigr)\, \mu(dx) \\ &= \lim_{n\to\infty} \int_Y f(y)\, \varrho_n(y)\, \nu(dy) \le \sup_n \int_U |\varrho_n(y)|\, \nu(dy) \le \frac{\varepsilon}{2},\end{aligned}$$

whence we obtain $\mu \circ T^{-1}(K) \le \varepsilon$, which is a contradiction. Therefore, $\mu \circ T^{-1}(K) = 0$, which by the Radon property of our measures yields the relation $\mu \circ T^{-1} \ll \nu$. Letting $\varrho := d(\mu \circ T^{-1})/d\nu$, we obtain

$$\int_Y f\, \varrho\, d\nu = \int_X f \circ T\, d\mu = \lim_{n\to\infty} \int_Y f\, \varrho_n\, d\nu \tag{9.9.1}$$

for every bounded continuous function f. According to Corollary 4.7.19, every subsequence of the sequence $\{\varrho_n\}$ contains a weakly convergent subsequence in $L^1(\nu)$. However, (9.9.1) shows that all such weakly convergent sequences may have only one limit ϱ, whence we obtain convergence of $\{\varrho_n\}$ to ϱ in the weak topology. □

The condition that ν and $\mu\circ T^{-1}$ are Radon can be replaced by the one that both measures are Baire provided that the mappings T are T_n measurable with respect to the pair $\bigl(\mathcal{A}_{|\mu|}, \mathcal{B}a(Y)\bigr)$ (then no complete regularity of Y is needed). The only change in the proof is that in place of a compact set K we take a functionally closed set and U must be functionally open.

9.9.11. Corollary. *Suppose that in the situation of the above proposition the measure μ is nonnegative and there is a sequence of ν-measurable functions f_n convergent in measure ν to a function f. Then, the functions $f_n\circ T_n$ converge in measure μ to $f\circ T$.*

Proof. We may assume that μ is a probability measure. In addition, we may assume that the functions f_n converge to f almost everywhere with respect to ν because it suffices to verify that every subsequence in $\{f_n\}$ contains a further subsequence for which the conclusion is true. Let $\varepsilon>0$. By using the uniform integrability of the densities ϱ_n and Lusin's and Egoroff's theorems, we find a compact set $K\subset Y$ and a number N_1 such that f is continuous on K, $\mu\bigl(T^{-1}(K)\bigr)>1-\varepsilon$, $\mu\bigl(T_n^{-1}(K)\bigr)>1-\varepsilon$ for all n, and $\sup_{y\in K}|f_n(y)-f(y)|<\varepsilon$ for all $n\ge N_1$. There is a continuous function g on Y such that $g|_K=f|_K$. By using the continuity of g and almost everywhere convergence of T_n to T, we find $N_2\ge N_1$ such that for all $n\ge N_2$ one has the estimate $\mu\Bigl(x\colon\ \bigl|g\bigl(T_n(x)\bigr)-f\bigl(T(x)\bigr)\bigr|>\varepsilon\Bigr)\le\varepsilon$. Then

$$\mu\Bigl(x\colon\ \bigl|f\bigl(T_n(x)\bigr)-f\bigl(T(x)\bigr)\bigr|>\varepsilon\Bigr)\le\mu\Bigl(x\colon\ \bigl|g\bigl(T_n(x)\bigr)-g\bigl(T(x)\bigr)\bigr|>\varepsilon\Bigr)$$
$$+\mu\circ T^{-1}(Y\backslash K)+\mu\circ T_n^{-1}(Y\backslash K)\le 3\varepsilon$$

for all $n\ge N_2$. It remains to observe that

$$\mu\Bigl(x\colon\ \bigl|f_n\bigl(T_n(x)\bigr)-f\bigl(T_n(x)\bigr)\bigr|>\varepsilon\Bigr)\le\mu\bigl(x\colon\ T_n(x)\notin K\bigr)<\varepsilon$$

whenever $n\ge N_2$. Hence $\mu\Bigl(x\colon\ \bigl|f_n\bigl(T_n(x)\bigr)-f\bigl(T(x)\bigr)\bigr|>2\varepsilon\Bigr)\le 4\varepsilon$. □

The established proposition and corollary are often applied in the situation where $X=Y$ and $\mu=\nu$, so one deals with transformations of a single space. In this case, one has to verify that the transformed measures have uniformly integrable densities with respect to the initial measure.

9.10. Shifts of measures along integral curves

Let $F\colon \mathbb{R}^n\to\mathbb{R}^n$ be a vector field for which the ordinary differential equation

$$x'(t)=F\bigl(x(t)\bigr),\quad x(0)=x,$$

for every initial condition x has a solution that we denote by $U_t(x)$, assuming that it exists on the whole real line. Thus, for every fixed t, we obtain a mapping $x \mapsto U_t(x)$. The action of this mapping consists in shifting along the integral curves of the given equation. The family of mappings U_t is called the flow generated by the vector field F because under broad assumptions, the family $\{U_t\}$ has the semigroup property: $U_tU_s = U_{t+s}$. In the theory of dynamical systems, it is often useful to know how a given measure is transformed by the flow $\{U_t\}$. An answer to this question enables one, in particular, to find measures that are invariant with respect to transformations U_t. In certain problems, one is interested in measures μ that may not be invariant with respect to U_t, but are transformed into equivalent measures. In this section, we solve the above-mentioned problems. Since complete proofs are technically involved, we consider in detail only the simplest partial case.

In this section, Lebesgue measure of a set D is denoted by $|D|$. The norm of a vector v in $\mathbb{R}^n$ is denoted by $|v|$. We recall that the divergence of a vector field $F = (F^1, \ldots, F^n)$ on $\mathbb{R}^n$, where $F^j \in W^{1,1}_{loc}(\mathbb{R}^n)$ (for example, $F^j \in C^1(\mathbb{R}^n)$), is defined by the equality

$$\mathrm{div} F = \sum_{j=1}^{n} \partial_{x_j} F^j.$$

According to the integration by parts formula, for every smooth function φ with compact support, one has the equality

$$\int_{\mathbb{R}^n} (\nabla\varphi, F)\, dx = -\int_{\mathbb{R}^n} \varphi \mathrm{div} F\, dx.$$

The divergence of a vector field determines how Lebesgue measure is transformed by the corresponding flow.

9.10.1. Theorem. *Let $\Psi\colon \mathbb{R}^n \to \mathbb{R}^n$ be a smooth vector field with compact support and let $\{U_t\}$ be the corresponding flow. Then, every mapping U_t is a diffeomorphism transforming Lebesgue measure into the measure with density*

$$\varrho_t(x) = \exp\biggl\{-\int_0^t \mathrm{div}\Psi\bigl(U_{-s}(x)\bigr)\, ds\biggr\}. \tag{9.10.1}$$

PROOF. It is known from the theory of ordinary differential equations that the corresponding global flow $\{U_t\}$ exists and that the mapping U_t is a diffeomorphism of $\mathbb{R}^n$. It is clear that $U_t(x) = x$ for all t and all $x \notin D$, where D is a ball containing the support of Ψ. The image of Lebesgue measure with respect to U_t has a density ϱ_t that is continuous in both arguments, since $U_t(x)$ is continuously differentiable in both arguments. For every $\varphi \in C_0^\infty(\mathbb{R}^n)$, we have

$$\frac{\partial}{\partial t}\varphi \circ U_t = \frac{\partial}{\partial \tau}\varphi \circ U_t \circ U_\tau|_{\tau=0} = \bigl(\nabla(\varphi \circ U_t), \Psi\bigr).$$

Therefore,

$$\begin{aligned}\int \varphi(x)\varrho_t(x)\,dx &= \int \varphi(x)\,dx + \int \int\limits_0^t \bigl(\nabla(\varphi\circ U_s)(x), \Psi(x)\bigr)\,ds\,dx \\ &= \int \varphi(x)\,dx + \int_0^t \int \bigl(\nabla(\varphi\circ U_s)(x), \Psi(x)\bigr)\,dx\,ds \\ &= \int \varphi(x)\,dx - \int_0^t \int \mathrm{div}\Psi(x)\varphi\bigl(U_s(x)\bigr)\,dx\,ds \\ &= \int \varphi(x)\,dx - \int_0^t \int \mathrm{div}\Psi\bigl(U_{-s}(y)\bigr)\varphi(y)\varrho_s(y)\,dy\,ds.\end{aligned}$$

Since φ is arbitrary, we obtain that for all t and x, one has

$$\varrho_t(x) = 1 - \int_0^t \mathrm{div}\Psi\bigl(U_{-s}(x)\bigr)\varrho_s(x)\,ds,$$

which yields the required relationship. □

It is clear that this theorem is valid in the case of Riemannian manifolds. Formula (9.10.1) yields the following assertion (the Liouville theorem).

9.10.2. Corollary. *In the situation of the above theorem, Lebesgue measure is invariant with respect to the transformations U_t clarify when the equality* $\mathrm{div}\Psi = 0$ *holds.*

In addition, one can derive from expression (9.10.1) a number of useful estimates. To this end, we need two lemmas.

9.10.3. Lemma. *Let f be an integrable function on the interval $[0,t]$, where $t \ge 0$. Then, letting $t\vee 1 := \max(t,1)$, we have*

$$\exp\biggl(\int_0^t f(s)\,ds\biggr) \le 1 + \int_0^t e^{(t\vee 1)f(s)}\,ds. \tag{9.10.2}$$

Proof. By Jensen's inequality one has

$$\exp\biggl(t^{-1}\int_0^t tf(s)\,ds\biggr) \le t^{-1}\int_0^t e^{tf(s)}\,ds.$$

If $t \ge 1$, then this immediately yields (9.10.2). If $t < 1$, then we apply the obtained estimate on the interval $[0,1]$ to the function g that equals f on $[0,t]$ and 0 on $(t,1]$. Since $e^{g(s)} = 1$ on $(t,1]$, we arrive again at (9.10.2). □

9.10.4. Lemma. *Let ν be a finite nonnegative measure on a space Ω and let $\{U_t\}_{|t|\le T}$ be a family of measurable transformations of Ω such that $\nu\circ U_t^{-1} = r_t\cdot\nu$, where*

$$r_t(x) = \exp\biggl\{\int_0^t f\bigl(U_{-s}(x)\bigr)\,ds\biggr\},$$

the function $f\bigl(U_{-s}(x)\bigr)$ *is measurable in* (s,x), *and* $\exp(|f|)\in L^p(\nu)$ *for all* $p\in(0,\infty)$. *Suppose that the estimate*

$$\int_{-T}^{T}\|r_t\|_{1+\varepsilon}\,dt<\infty \tag{9.10.3}$$

where $\|\cdot\|_\alpha:=\|\cdot\|_{L^\alpha(\nu)}$, *is true for some* $\varepsilon>0$. *Then it is true for every* $\varepsilon>0$. *In addition, for every* $p>1$ *and* $t\in[-T,T]$, *one has*

$$\|r_t\|_p\le\bigl(2+2\nu(\Omega)\bigr)e^{C(p,T)|t|}, \tag{9.10.4}$$

where

$$C(p,T)=\biggl(\int_\Omega e^{qp(T\vee 1)|f(x)|}\,\nu(dx)\biggr)^{1/q}\quad \textit{and}\quad \frac{1}{p}+\frac{1}{q}=1.$$

Proof. By Lemma 9.10.3, for every $t\in[0,T]$ we obtain

$$r_t(x)^p\le 1+\int_0^t\exp\bigl\{p(t\vee 1)f\bigl(U_{-s}(x)\bigr)\bigr\}\,ds.$$

According to Hölder's inequality with k defined by $\frac{1}{k}+\frac{1}{1+\varepsilon}=1$, we obtain

$$\begin{aligned}\int_\Omega r_t(x)^p\,\nu(dx)&\le\nu(\Omega)+\int_\Omega\int_0^t\exp\bigl\{p(t\vee 1)f\bigl(U_{-s}(x)\bigr)\bigr\}\,ds\,\nu(dx)\\&=\nu(\Omega)+\int_0^t\int_\Omega\exp\bigl\{p(t\vee 1)f(x)\bigr\}\,r_{-s}(x)\,\nu(dx)\,ds\\&\le\nu(\Omega)+\bigl\|\exp\{p(t\vee 1)f\}\bigr\|_k\int_0^t\|r_{-s}\|_{1+\varepsilon}\,ds.\end{aligned}$$

Similarly, for negative t we have

$$\int_\Omega r_t(x)^p\,\nu(dx)\le\nu(\Omega)+\bigl\|\exp\{p(|t|\vee 1)f\}\bigr\|_k\int_0^{|t|}\|r_s\|_{1+\varepsilon}\,ds.$$

Thus, the function $t\mapsto\|r_t\|_p$ is essentially bounded on $[-T,T]$, and (9.10.3) holds for every $\varepsilon>0$. Since the function $t\mapsto r_t(x)$ is continuous for ν-a.e. x, we obtain by the Lebesgue–Vitali theorem that the function $t\mapsto\|r_t\|_p$ is continuous, hence the above estimate holds for every t.

Let I_t denote the interval $[0,|t|]$. Then the above estimate is true with $1+\varepsilon=p$ (and $k=q$), so that for all $|t|\le T$ we have

$$\int_\Omega r_t(x)^p\,\nu(dx)\le\nu(\Omega)+\bigl\|\exp\{p(T\vee 1)f\}\bigr\|_q\int_{I_t}\|r_{-s}\|_p\,ds.$$

Since $a\le 1+a^p$ for $a\ge 0$, we obtain

$$\|r_t\|_p\le 1+\nu(\Omega)+C(p,T)\int_{I_t}\|r_{-s}\|_p\,ds.$$

Letting $\psi(t)=\|r_t\|_p+\|r_{-t}\|_p$, we arrive at the estimate

$$\psi(t)\le 2+2\nu(\Omega)+C(p,T)\int_0^{|t|}\psi(s)\,ds.$$

We recall that by Gronwall's inequality one has a.e.

$$u(t) \le C \exp \int_0^t v(s)\,ds$$

for any nonnegative integrable functions u and v satisfying a.e. the inequality

$$u(t) \le C + \int_0^t v(s)u(s)\,ds.$$

This yields the desired estimate. □

9.10.5. Corollary. *In the situation of Theorem* 9.10.1*, for every ball D that contains the support of Ψ and every $p > 1$, one has*

$$\|\varrho_t\|_p := \|\varrho_t\|_{L^p(D)} \le M_D e^{C(p,t)|t|},$$

where $M_D := 2(1 + |D|)$, $C(p,t) := \bigl\|\exp\{p(|t| \vee 1)|\mathrm{div}\Psi|\}\bigr\|_{L^{p/(p-1)}(D)}$. In addition,

$$\bigl\||\nabla U_t|\bigr\|_{L^p(D)} \le 2M_D \bigl\|\exp\{(|t| \vee 1)|\nabla\Psi|\}\bigr\|_{L^{2p}(D)} e^{(C(2,t)+1)|t|/p},$$

$$\Bigl\||\partial U_t/\partial t|\Bigr\|_{L^p(D)} \le M_D \bigl\||\Psi|\bigr\|^p_{L^{2p}(D)} e^{C(2,t)|t|}.$$

The proof is given in Bogachev, Mayer-Wolf [**220**].

Let us now see how more general measures are transformed. As in the case of Lebesgue measure, the answer will be expressed in terms of the divergence of the vector field with respect to the given measure. Suppose that μ is a measure on $\mathbb{R}^n$ with a positive density ϱ such that ϱ is continuously differentiable (or, more generally, on every ball is separated from zero and belongs to the Sobolev class $W^{1,1}_{loc}$). Let F be a vector field on $\mathbb{R}^n$ belonging to the Sobolev class $W^{1,1}_{loc}$ such that the function $|\nabla F|$ is locally μ-integrable. The divergence of F with respect to μ is the function denoted by the symbol $\delta_\mu F$ and defined by the formula

$$\delta_\mu F(x) := \mathrm{div}\, F(x) + \Bigl(F(x), \frac{\nabla\varrho(x)}{\varrho(x)}\Bigr).$$

By the integration by parts formula it is readily verified that the function $\delta_\mu F$ is characterized by the identity

$$\int_{\mathbb{R}^n} (\nabla\varphi, F)\,d\mu = -\int_{\mathbb{R}^n} \varphi\delta_\mu F\,d\mu, \quad \varphi \in C_0^\infty(\mathbb{R}^n).$$

9.10.6. Theorem. *Let $\Psi\colon \mathbb{R}^n \to \mathbb{R}^n$ be a smooth vector field with compact support and let μ be a probability measure on $\mathbb{R}^n$ with a positive continuously differentiable density ϱ. Then, for every $t \in \mathbb{R}^1$, the measure $\mu \circ U_t^{-1}$ is absolutely continuous with respect to μ and its Radon–Nikodym density is given by the equality*

$$\frac{d\bigl(\mu \circ U_t^{-1}\bigr)}{d\mu} = r_t(x) = \exp\Bigl\{-\int_0^t \delta_\mu\Psi\bigl(U_{-s}(x)\bigr)\,ds\Bigr\}.$$

In addition, if $\Lambda(p,t) = \big\|\exp\{(|t|\vee 1)p|\delta_\mu\Psi|\}\big\|_{L^{p/(p-1)}(\mu)}$, *then one has the following estimates*:

$$\|r_t\|_{L^p(\mu)} \le 4\exp\Big[\Lambda(p,t)|t|\Big],$$

$$\big\||\nabla U_t|\big\|_{L^p(\mu)} \le 4\Big\|\exp\big[(|t|\vee 1)|\nabla\Psi|\big]\Big\|_{L^{2p}(\mu)} \exp\big[\big(\Lambda(2,t)+1\big)|t|/p\big],$$

$$\Big\||\partial U_t/\partial t|\Big\|_{L^p(\mu)} \le 4\big\||\Psi|\big\|^p_{L^{2p}(\mu)} \exp\big[\Lambda(2,t)|t|\big].$$

PROOF. Since $\Psi = 0$ outside some ball, one has $\delta_\mu\Psi\big(U_{-s}(x)\big) = 0$ for all $s\in[0,t]$ and all x with a sufficiently large norm. The expression for $r_t(x)$ is obtained in the same manner as the formula for $\varrho_t(x)$ in Theorem 9.10.1. Then the same reasoning based on Lemma 9.10.3, Lemma 9.10.4, and Gronwall's inequality yields the stated estimates. □

We note that the hypotheses on the density of the measure μ can be weakened: it suffices that $\varrho\in W^{1,1}_{loc}(\mathbb{R}^n)$, the function ϱ be locally uniformly separated from zero and that for every $c\in\mathbb{R}^1$ the function $\exp\big(c(\Psi,\nabla\varrho/\varrho)\big)$ be locally μ-integrable (see Bogachev, Mayer-Wolf [**220**]).

Now we extend the above results to more general vector fields, in particular, not necessarily smooth and not necessarily with compact support. Let us precise what we mean by a flow generated by a more general vector field. Let μ be a measure on $\mathbb{R}^n$ and let F be a μ-measurable vector field. A mapping $(t,x)\in\mathbb{R}^1\times\mathbb{R}^n\mapsto U^F_t(x)\in\mathbb{R}^n$ is called a solution of the equation

$$U_t(x) = x + \int_0^t F\big(U_s(x)\big)\,ds, \tag{9.10.5}$$

if

(a) for μ-almost every x equality (9.10.5) is fulfilled with $U=U^F$ for all $t\in\mathbb{R}^1$ (in particular, the right-hand side must be meaningful),

(b) for every $t\in\mathbb{R}^1$, the measure $\mu\circ U_t^{-1}$ is absolutely continuous with respect to μ.

The Radon–Nikodym derivative $d(\mu\circ U_t^{-1})/d\mu$ will be denoted by r_t.

The family $(U_t)_{t\in\mathbb{R}^1} = (U^F_t)_{t\in\mathbb{R}^1}$ is called a flow if, in addition, we have for μ-a.e. x

$$U_{t+s}(x) = U_t\big(U_s(x)\big), \quad \forall s,\, t\in\mathbb{R}^1. \tag{9.10.6}$$

The quasi-invariance (condition (b) above) is essential when dealing with equivalence classes of vector fields if we want to have solutions independent of concrete representatives in the equivalence classes: according to Exercise 9.12.59, if $F(x)=G(x)$ μ-a.e. and $(U^F_t)_{t\in\mathbb{R}^1}$ is a solution for the field F, then it is a solution for the field G.

Simple examples such as the field $F(x)=x^2$ on the real line show that the smoothness of the field is not sufficient for the existence of a global solution.

The following result is obtained in Bogachev, Mayer-Wolf [**220**].

9.10.7. Theorem. *Let μ be a measure on $\mathbb{R}^n$ having a locally uniformly positive density $\varrho \in W^{1,1}_{loc}(\mathbb{R}^n)$ and let $F \in W^{1,1}_{loc}(\mathbb{R}^n, \mathbb{R}^n)$ be a vector field such that*

$$e^{|\delta_\mu F|} \in \bigcap_{p>1} L^p(\mu)$$

and either $e^{|\nabla F|} \in \bigcap_{p>1} L^p(\mu)$ or $|F| \in L^{1+\varepsilon}(\mathbb{R}^n)$ and $e^{|\nabla F|} \in \bigcap_{p>1} L^p(K, \mu)$ for every ball K. Then equation (9.10.5) *has a flow.*

This theorem ensures the existence of global solutions to ordinary differential equations for many rapidly increasing vector fields. To this end, one has to find a measure μ such that the functions indicated in the formulation are integrable. Let us consider the special case where μ is the standard Gaussian measure on $\mathbb{R}^n$, i.e.,

$$\varrho(x) = (2\pi)^{-n/2} \exp\bigl(-|x|^2/2\bigr).$$

In this case $\nabla\varrho(x)/\varrho(x) = -x$. Suppose that the field F is locally Lipschitzian. Then for the existence of a flow generated by this field we need that the functions

$$|F(x)|^{1+\varepsilon} \quad \text{and} \quad \exp\Bigl[c\bigl|\mathrm{div}F(x) - \bigl(x, F(x)\bigr)\bigr|\Bigr]$$

be μ-integrable for all c and some $\varepsilon > 0$. Effectively verified sufficient conditions are the estimates

$$|F(x)| \le C_1 e^{C_2|x|}, \quad \bigl|\mathrm{div}F(x) - \bigl(x, F(x)\bigr)\bigr| \le C_2|x|$$

with some constants C_1 and C_2. We remark that even for smooth fields F satisfying the condition $\delta_\mu F = 0$, one cannot omit the μ-integrability of F (see [**220**]). Yet, the main restriction is the exponential integrability of $\delta_\mu F$. Given a smooth (or locally Lipschitzian) field F, it is not difficult to find a measure μ with a rapidly decreasing density such that the function $|F|^2$ is μ-integrable. However, one cannot always achieve the exponential integrability of $\delta_\mu F$. Constructing a measure μ with the required properties is analogous to constructing Lyapunov functions used in the theory of differential equations.

The problems considered in this section are being intensively investigated for infinite-dimensional spaces; see [**220**].

9.11. Invariant measures and Haar measures

Let X be a locally compact topological space and let G be a locally compact topological group (as usual, we consider Hausdorff spaces). Suppose that we are given an action of the group G on X, i.e., a mapping $A\colon\ G{\times}X \to X$ such that $A(e, \cdot\,)$ is the identity mapping on X (e is the unity element of G) and one has the equality

$$A(g_1 g_2, x) = A\bigl(g_1, A(g_2, x)\bigr), \quad \forall\, g_1, g_2 \in G,\ \forall\, x \in X.$$

In particular, $A(g^{-1}, \cdot\,)$ is the inverse mapping to $A(g, \cdot\,)$. In other words, we are given a homomorphism of G to the group of transformations of X. For notational simplicity the transformation $A(g, x)$ is usually denoted by gx.

If we take G for X, then the usual left multiplication by $g \in G$ provides an important example of an action. Another example: the natural action of the group of isometries of a metric space X.

In applications, one usually deals with actions that are measurable or even continuous. In this section, we consider only *continuous actions*.

9.11.1. Definition. (i) *Let μ be a Borel measure on X with values in $[0, +\infty]$ (or a measure on the σ-ring generated by compact sets) that is finite on compact sets and inner compact regular, i.e., for every $B \in \mathcal{B}(X)$, one has $\mu(B) = \sup\bigl\{\mu(K)\colon\ K \subset B \text{ is compact}\bigr\}$. Let χ be a function on G. The measure μ is called χ-covariant if the image of μ under the mapping $x \mapsto A(g, x)$ is $\chi(g) \cdot \mu$ for all $g \in G$. In the case $\chi = 1$, the measure μ is called G-invariant.*

(ii) *If G acts on itself by the left multiplication, then nonzero G-invariant measures are called left (or left invariant) Haar measures and if G acts on itself by means of the formula $(g, x) \mapsto xg^{-1}$, then nonzero G-invariant measures are called right (or right invariant) Haar measures.*

It is clear that if $\chi = 1$, then the χ-covariance means just the invariance with respect to the action of G, i.e., for any left invariant Haar measure μ_L on a group G one has

$$\int_G f(gx)\,\mu_L(dx) = \int_G f(x)\,\mu_L(dx)$$

for all $f \in C_0(G)$. If the measure μ is not zero, then χ is a character of G, i.e., a homomorphism to the multiplicative group $\mathbb{R}\backslash\{0\}$. Note also that the mapping $j\colon\ x \mapsto x^{-1}$ on the group G takes left Haar measures to right Haar measures and vice versa. Indeed, if a measure μ_L is left invariant, then for any $g \in G$ and $f \in C_0(G)$ we have

$$\int_G f(xg^{-1})\,\mu_L \circ j^{-1}(dx) = \int_G f(x^{-1}g^{-1})\,\mu_L(dx)$$
$$= \int_G f\bigl((gx)^{-1}\bigr)\,\mu_L(dx) = \int_G f(x^{-1})\,\mu_L(dx) = \int_G f\,d(\mu_L \circ j^{-1}).$$

Usually χ-covariant measures are called quasi-invariant. Analogous notions make sense and are very interesting in the case of groups that are not locally compact (such as the group of diffeomorphisms of a manifold). However, the corresponding theory is much more involved and we do not discuss it here. The principal reason for its higher level of complexity is the absence of measures on such groups that are invariant or quasi-invariant with respect to all shifts. For this reason, one has to consider measures that are quasi-invariant with respect to the action of certain subgroups. For example, there is no Borel probability measure on the infinite-dimensional separable Hilbert space that is quasi-invariant with respect to all translations, but there are measures quasi-invariant with respect to translations from everywhere dense subspaces, and on the group of C^1-diffeomorphisms of the circle there is no

Borel probability measure quasi-invariant with respect to all shifts, but there are measures quasi-invariant with respect to subgroups of diffeomorphisms of higher smoothness. In recent decades the investigation of such measures on infinite-dimensional linear spaces and groups has been intensively developing; see references in Bogachev [**206**], Malliavin [**1243**].

Set $A(g, W) = \{A(g, y), y \in W\}$. We call the action A of G equicontinuous if, for every $x \in X$ and every neighborhood V of the point x, one can find a neighborhood W of this point such that if $A(g, W) \cap W \neq \varnothing$, then $A(g, W) \subset V$.

9.11.2. Theorem. *Suppose that G acts equicontinuously on X and that for every x the mapping $g \mapsto A(g, x)$ is surjective and open. Then, there is a nonzero G-invariant measure μ on X.*

PROOF. Let $H \subset X$ be a compact set with nonempty interior. (1) We fix a point p in the interior of H and denote by $\mathcal{U}_p$ the set of all open neighborhoods of p with compact closure. For every set E with compact closure and every set F with nonempty interior, one can cover E by finitely many translations of F. The smallest possible number of such translations is denoted by $[E : F]$. Let $[\varnothing, F] = 0$. It is clear that $[gE : F] = [E : gF] = [E : F]$ for all $g \in G$ and $[E : F] \leq [D : F]$ if $E \subset D$. In addition, $[E : F] \leq [E : A] \cdot [A : F]$ for every set A with compact closure and nonempty interior. Given $U \in \mathcal{U}_p$, we let $\xi_U(E) = [E : U]/[H : U]$. It is clear that $\xi_U(E) \leq [E : H]$, $\xi_U(gE) = \xi_U(E)$ if $g \in G$ and $\xi_U(H) = 1$. In addition, $\xi_U(E \cup F) \leq \xi_U(E) + \xi_U(F)$ and $\xi_U(E_1) \leq \xi_U(E_2)$ if $E_1 \subset E_2$. Finally, for any disjoint compact sets K_1 and K_2, one can find a neighborhood $U \in \mathcal{U}_p$ such that for all $V \in \mathcal{U}_p$ with $V \subset U$ one has

$$\xi_V(K_1 \cup K_2) = \xi_V(K_1) + \xi_V(K_2). \tag{9.11.1}$$

Indeed, by the equicontinuity of the action of G, there exists a neighborhood $U \in \mathcal{U}_p$ such that for every $g \in G$ either $gU \cap K_1 = \varnothing$ or $gU \cap K_2 = \varnothing$. Hence every cover of $K_1 \cup K_2$ by translations of U is a disjoint union of covers of K_1 and K_2, which yields $\xi_U(K_1 \cup K_2) \geq \xi_U(K_1) + \xi_U(K_2)$. The same is true for every smaller neighborhood. Since the reverse inequality is true as well, we arrive at (9.11.1).

(2) Our next step is to define $\lambda(K)$ as the limit of $\xi_U(K)$ as U is shrinking. The precise definition is this. Let Θ be the linear space of all bounded functions on the set $\mathcal{U}_p$. For every $\xi \in \Theta$ we let

$$p(\xi) = \inf_{U \in \mathcal{U}_p} \sup_{V \subset U, V \in \mathcal{U}_p} \xi(V), \quad q(\xi) = \sup_{U \in \mathcal{U}_p} \inf_{V \subset U, V \in \mathcal{U}_p} \xi(V).$$

We observe that $q(\xi) \leq p(\xi)$. Indeed, if $U_1, U_2 \in \mathcal{U}_p$, then $U = U_1 \cap U_2 \in \mathcal{U}_p$ and

$$\inf_{V \subset U_1} \xi(V) \leq \inf_{V \subset U} \xi(V) \leq \sup_{V \subset U_2} \xi(V).$$

It is easy to see that $p(0) = 0$, $p(\xi + \eta) \leq p(\xi) + p(\eta)$, and $p(\alpha\xi) = \alpha p(\xi)$ for all $\alpha \geq 0$ and $\xi, \eta \in \Theta$. In addition, $q(\xi) = -p(-\xi)$. By the Hahn–Banach theorem, the zero functional on the zero subspace of Θ extends to a linear function Λ on Θ such that $\Lambda(\xi) \le p(\xi)$. One has

$$-\Lambda(\xi) = \Lambda(-\xi) \le p(-\xi) = -q(\xi),$$

whence $q(\xi) \le \Lambda(\xi) \le p(\xi)$. Hence $\Lambda(1) = 1$. If $\xi \ge 0$, then $q(\xi) \ge 0$ and hence $\Lambda(\xi) \ge 0$. Thus, whenever $\xi \ge \eta$, we have $\Lambda(\xi) \ge \Lambda(\eta)$.

(3) We note one more property of Λ: if functions $\xi, \eta \in \Theta$ are such that for some $U \in \mathcal{U}_p$ we have $\xi(V) = \eta(V)$ for all $V \in \mathcal{U}_p$ with $V \subset U$, then $\Lambda(\xi) = \Lambda(\eta)$. To this end, we set $\zeta = \xi - \eta$ and observe that $p(\zeta) = 0$, whence $\Lambda(\zeta) \le 0$. Replacing ζ by $-\zeta$, we obtain $\Lambda(\zeta) \ge 0$, hence $\Lambda(\zeta) = 0$.

(4) For every compact set K we let

$$\lambda(K) = \Lambda\bigl(\xi_\bullet(K)\bigr),$$

where $\xi_\bullet(K)\colon U \mapsto \xi_U(K)$ is the element of Θ generated by K. It is clear that $\lambda(gK) = \lambda(K)$, since $\xi_U(gK) = \xi_U(K)$. In addition, $\lambda(H) = 1$, since $\xi_U(H) = 1$ if $U \in \mathcal{U}_p$. Finally, for any disjoint compact sets K_1 and K_2 we obtain

$$\lambda(K_1 \cup K_2) = \lambda(K_1) + \lambda(K_2).$$

This follows by (9.11.1) and the property established in (3).

(5) According to Theorem 7.11.1, there exists a countably additive measure μ on $\mathcal{B}(X)$ (the measure μ' from the cited theorem) with values in $[0, +\infty]$ and finite on all compact sets such that for every Borel set $B \subset X$ one has $\mu(B) = \sup \lambda(K)$, where sup is taken over all compact sets $K \subset B$. Then we obtain $\mu(H) = \lambda(H) = 1$. Finally, the equality $\lambda(gK) = \lambda(K)$ for all compact sets K and all $g \in G$ yields the equality $\mu(gB) = \mu(B)$ for all Borel sets B. $\square$

In some books (see Hewitt, Ross [**825**]), one constructs an outer regular Haar measure (see Remark 7.11.2), which coincides with μ on compact sets, but may differ from μ on some Borel sets if μ is not σ-finite. If μ has no atoms, then by the inner compact regularity $\mu(D) = 0$ for every discrete set D, although $\mu(U) = +\infty$ for any uncountable union U of disjoint open sets, in particular, for every neighborhood U of D. In particular, let $G = \mathbb{R} \times \mathbb{R}^1$, where $\mathbb{R}$ is the additive group of all real numbers with the discrete topology and $\mathbb{R}^1$ is the same additive group with the usual topology (as in Example 7.14.65). Then G is a locally compact commutative group and its Haar measure μ is the product of the counting measure on $\mathbb{R}$ and Lebesgue measure on $\mathbb{R}^1$. Here $\mu(\mathbb{R} \times \{0\}) = 0$, but $\mu(U) = +\infty$ for every open set $U \supset \mathbb{R} \times \{0\}$. A similar example exists in every locally compact group whose Haar measure has no atoms (i.e., the group is not discrete) and is not σ-finite: it suffices to take an uncountable set of points with pairwise disjoint neighborhoods.

9.11.3. Example. The hypotheses of Theorem 9.11.2 are fulfilled in the following cases: (i) $(g, h) \mapsto gx$ is the action of G on itself by the left multiplication; (ii) $(g, h) \mapsto xg^{-1}$ is the action of G on itself by the right multiplication; (iii) $(g, x) = g(x)$ is the natural action of the group of invertible matrices GL_n on $\mathbb{R}^n \backslash \{0\}$.

9.11.4. Corollary. *On every locally compact group, there is a unique, up to a constant factor, left invariant Haar measure. The same is true for right invariant measures.*

PROOF. Let ν be a right invariant Haar measure on G, μ a left invariant Haar measure on G, $\psi \in C_0(G)$, $\psi \geq 0$, and let ψ not vanish identically. We observe that μ and ν are positive on nonempty open sets. Let

$$\Delta(x) := \int_G \psi(y^{-1}x)\, \nu(dy). \tag{9.11.2}$$

It is easy to see that the function Δ is continuous and strictly positive. Multiplying μ and ν by constants, we may assume that

$$\int_G \psi(x)\, \mu(dx) = \int_G \psi(y^{-1})\, \nu(dy) = 1.$$

Let $\Gamma(x) = \Delta(x)^{-1}$. For any $\varphi \in C_0(G)$, by Fubini's theorem and the respective invariance of the two measures we have

$$\int_G \varphi(x)\, \mu(dx) = \int_G \varphi(x)\Gamma(x)\Delta(x)\, \mu(dx)$$
$$= \int_G \int_G \varphi(x)\Gamma(x)\psi(y^{-1}x)\, \nu(dy)\, \mu(dx)$$
$$= \int_G \int_G \varphi(x)\Gamma(x)\psi(y^{-1}x)\, \mu(dx)\, \nu(dy) = \int_G \int_G \varphi(yx)\Gamma(yx)\psi(x)\, \mu(dx)\, \nu(dy)$$
$$= \int_G \int_G \varphi(yx)\Gamma(yx)\psi(x)\, \nu(dy)\, \mu(dx) = \int_G \varphi(y)\Gamma(y)\, \nu(dy) \int_G \psi(x)\, \mu(dx).$$

Thus, any function in $C_0(G)$ has equal integrals against the measures μ and $\Gamma \cdot \nu$, which yields the coincidence of these measures on all compact sets, hence on $\mathcal{B}(G)$ by the inner compact regularity. Moreover, Γ is independent of μ, which shows the uniqueness of ν with the above-chosen normalization of the integral of $\psi(y^{-1})$. The assertion for μ is analogous. □

The function Δ defined by formula (9.11.2) with $\Delta(e) = 1$ is called the modular function of the group G. It does not depend on ψ. Indeed, if we take another function ψ' with $\Delta'(e) = 1$, then for the corresponding function Γ' we obtain $\Gamma' = c\Gamma$ with some constant, since $\Gamma = d\mu/d\nu$. In addition, $\Gamma(e) = \Gamma'(e) = 1$.

9.11.5. Corollary. *If μ is a left invariant Haar measure and ν is a right invariant Haar measure on G, then $\nu = c\Delta \cdot \mu$, where c is a constant. In addition, $\Delta(xy) = \Delta(x)\Delta(y)$.*

If $\Delta = 1$, then the group G is called unimodular. This is equivalent to the existence of two-sided invariant Haar measures on G. For example, all commutative and all compact groups are unimodular. The group of invertible matrices $n\times n$ is unimodular as well, but the group of all upper triangle 2×2 matrices with the numbers $u>0$ and 1 at the diagonal is not.

It is easy to verify that if a Haar measure is finite, then the group is compact (see Hewitt, Ross [**825**, §15]).

Note the following important fact discovered in Kakutani, Kodaira [**937**] (its proof can be read in Halmos [**779**, §64], Hewitt, Ross [**825**, §19], Fremlin [**635**, §463]).

9.11.6. Theorem. *Let G be a locally compact group and let λ be a Haar measure on G (left or right invariant). Then λ is completion regular in the following sense: for every Borel set $B\subset G$, we have $\mu(B)=\sup\mu(Z)$, where* sup *is taken over all functionally closed sets $Z\subset B$. In particular, if μ is σ-finite, then $\mathcal{B}(G)$ belongs to the Lebesgue completion of $\mathcal{B}a(G)$.*

9.12. Supplements and exercises

(i) Projective systems of measures (308). (ii) Extremal preimages of measures and uniqueness (310). (iii) Existence of atomless measures (317). (iv) Invariant and quasi-invariant measures of transformations (318). (v) Point and Boolean isomorphisms (320). (vi) Almost homeomorphisms (323). (vii) Measures with given marginal projections (324). (viii) The Stone representation (325). (ix) The Lyapunov theorem (326). Exercises (329)

9.12(i). Projective systems of measures

We have discussed above images and preimages of a measure in the situation where there is a single transformation. Now we intend to consider analogous questions for families of transformations. An especially important case is connected with the so-called projective systems of measures.

Let T be a directed set and let $\{X_\alpha\}_{\alpha\in T}$ be a projective system of spaces with mappings $\pi_{\alpha\beta}\colon X_\beta\to X_\alpha$, $\alpha\le\beta$, i.e., $\pi_{\alpha\alpha}=\mathrm{Id}$ and $\pi_{\alpha\beta}\circ\pi_{\beta\gamma}=\pi_{\alpha\gamma}$ if $\alpha\le\beta\le\gamma$. Suppose also we are given a space X with a system of mappings $\pi_\alpha\colon X\to X_\alpha$ that are consistent with the mappings $\pi_{\beta\alpha}$ in the following way: $\pi_\alpha=\pi_{\alpha\beta}\circ\pi_\beta$ if $\alpha\le\beta$. Such a space X is called the inverse limit of spaces X_α. A simple example: a decreasing countable sequence of spaces $X_n\supset X_{n+1}$ with the natural embeddings $\pi_{nk}\colon X_k\to X_n$, $X=\bigcap_{n=1}^\infty X_n$, and the natural embeddings $\pi_n\colon X\to X_n$. Another example: $X=\mathbb{R}^\infty$, $X_n=\mathbb{R}^n$ is identified with the subspace in $\mathbb{R}^\infty$ that consists of all sequences of the form $(x_1,\ldots,x_n,0,0,\ldots)$, and π_{nk} and π_n are the natural projections.

Suppose that the spaces X_α are equipped with σ-algebras $\mathcal{B}_\alpha$ and measures μ_α on $\mathcal{B}_\alpha$ such that the mappings $\pi_{\alpha\beta}$ are measurable. In typical cases (but not always) X_α is a topological space with its Borel σ-algebra and $\pi_{\alpha\beta}$ is continuous (hence Borel measurable). In the described setting, the problem arises whether there exists a measure μ on X, called a projective limit of the measures μ_α, such that

$$\mu\circ\pi_\alpha^{-1}=\mu_\alpha\quad\text{for all }\alpha. \tag{9.12.1}$$

Clearly, a necessary condition is this:

$$\pi_{\alpha\beta}(\mu_\beta) := \mu_\beta \circ \pi_{\alpha\beta}^{-1} = \mu_\alpha \quad \text{if } \alpha \le \beta. \tag{9.12.2}$$

For this reason, we shall discuss problem (9.12.1) under condition (9.12.2) (and assuming that X is nonempty).

An important example of such a situation (and the starting point of the related research) is the case where X is the space of mappings $x\colon [0,1] \to E$, where E is a topological space, A is the collection of all finite subsets of $[0,1]$ with their natural partial ordering by inclusion, $X_\alpha = \bigl\{x\colon \{t_1,\ldots,t_n\} \to E\bigr\}$, where $\alpha = \{t_1,\ldots,t_n\}$, and $\pi_{\alpha\beta}$ is the natural projection if $\alpha \subset \beta$. Thus, we are in the situation discussed in §7.7 in connection with the distributions of random processes. As has been noted there, one cannot always find a measure satisfying (9.12.1). We shall give some sufficient conditions for the existence of a solution, covering many cases important in applications. It should be noted that the idea of consideration of projective systems goes back to A.N. Kolmogorov, S. Bochner, and Yu.V. Prohorov. The main work in this direction was done in order to obtain suitable generalizations of Kolmogorov's theorem given in §7.7. The following result goes back to Prohorov [**1497**]. Now let X and X_α be topological spaces.

9.12.1. Theorem. *Let X be completely regular, let the mappings π_α and $\pi_{\alpha\beta}$ be continuous, and let* (9.12.2) *be fulfilled. Suppose that every μ_α is a Radon probability measure. A Radon probability measure μ on X satisfying* (9.12.1) *exists if and only if for every $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset X$ with $\mu_\alpha\bigl(\pi_\alpha(K_\varepsilon)\bigr) \ge 1-\varepsilon$ for all α.*

This result was extended to signed measures in Fremlin, Garling, Haydon [**636**]. We include the proof (borrowed from the cited work) for this generalization because the case of probability measures is not much simpler.

9.12.2. Theorem. *Let X be completely regular, let μ_α be Radon measures on X_α, and let the mappings π_α and $\pi_{\alpha\beta}$ be continuous and satisfy condition* (9.12.2). *A Radon measure μ on X satisfying* (9.12.1) *exists if and only if $\sup_\alpha \|\mu_\alpha\| < \infty$ and for every $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset X$ with $|\mu_\alpha|\bigl(X_\alpha\backslash\pi_\alpha(K_\varepsilon)\bigr) < \varepsilon$ for all α. If the mappings π_α separate the points in X, then such a measure μ is unique.*

PROOF. The necessity of this condition is obvious. Suppose it is fulfilled. We may assume that $\sup_\alpha \|\mu_\alpha\| \le 1$. For every $n \in \mathbb{N}$, there exists a compact set $K_n \subset X$ such that $|\mu_\alpha|\bigl(X_\alpha\backslash\pi_\alpha(K_n)\bigr) \le 1/n$ for all $\alpha \in T$. Let

$$M := \bigl\{\mu \in \mathcal{M}_r(X)\colon\ \|\mu\| \le 1,\ |\mu|(X\backslash K_n) \le 1/n,\ \forall\, n \in \mathbb{N}\bigr\}.$$

It is clear that M is a nonempty uniformly tight set in $\mathcal{M}_r(X)$ (we can assume that $K_n \subset K_{n+1}$; then any Dirac measure on K_1 is in M). Hence its closure $\overline{M}$ is compact in the weak topology. Let $M_\alpha := \bigl\{\mu \in \overline{M}\colon\ \mu\circ\pi_\alpha^{-1} = \mu_\alpha\bigr\}$, $\alpha \in T$. Every set M_α is closed in $\overline{M}$ in the weak topology and hence is compact. By Theorem 9.1.9 these sets are nonempty (since there is a Radon measure μ

with $\|\mu\| = \|\mu_\alpha\|$, $|\mu|\big(X\backslash\bigcup_n K_n\big) = 0$ and $\mu\circ\pi_\alpha^{-1} = \mu_\alpha$). Whenever $\alpha \le \beta$ we have $M_\beta \subset M_\alpha$. Indeed, let $\mu \in M_\beta$. Then

$$\mu\circ\pi_\alpha^{-1} = \big(\mu\circ\pi_\beta^{-1}\big)\circ\pi_{\alpha\beta}^{-1} = \mu_\beta\circ\pi_{\alpha\beta}^{-1} = \mu_\alpha$$

according to (9.12.2). The directed system of compact sets M_α has a nonempty intersection. Any element in this intersection is a required measure. The uniqueness assertion is delegated to Exercise 9.12.76. □

Theorem 9.12.2 has versions for measures with compact approximating classes and for perfect measures (Exercise 9.12.70).

9.12(ii). Extremal preimages of measures and uniqueness

Let $(Y,\mathcal{B},\nu)$ be a probability space, let $(X,\mathcal{A})$ be a measurable space, and let $f\colon X\to Y$ be an $(\mathcal{A},\mathcal{B})$-measurable mapping. Denote by M^ν the set of all probability measures μ on $(X,\mathcal{A})$ with $\nu = \mu\circ f^{-1}$. This set is convex, so the question arises about its extreme points (the set of extreme points is an important characteristic of a convex set). It turns out that under broad assumptions the extreme points of M^ν are precisely the images of the measure ν under measurable sections of the mapping f. This description is of interest also from another point of view: we recall that for a surjective mapping f between Souslin spaces, a preimage of the measure ν has been constructed in Theorem 9.1.5 as the image of ν with respect to a measurable section of f, which is not unique.

We shall say that a mapping $\pi\colon Y\to X$ is a $(\mathcal{B}_\nu,\mathcal{A})$-measurable weak section of f if π is measurable in the indicated sense and for every $B\in\mathcal{B}$, the set $\pi^{-1}\big(f^{-1}(B)\big)$ coincides with B up to a set of ν-measure zero. A short proof of the next assertion is given in Graf [**719**].

9.12.3. Theorem. *For every measure $\mu\in M^\nu$, the following conditions are equivalent*:

(i) *μ is an extreme point of M^ν*;

(ii) *there exists a σ-homomorphism $\Phi\colon \mathcal{A}\to\mathcal{B}/\nu$ (see* §9.12(v) *below) such that $\mu(A) = \nu\big(\Phi(A)\big)$ for all $A\in\mathcal{A}$ and $B\in\Phi\big(f^{-1}(B)\big)$ for all $B\in\mathcal{B}$*;

(iii) *the mapping $\varphi\mapsto\varphi\circ f$ from $L^1(\nu)$ to $L^1(\mu)$ is surjective*;

(iv) *for every $A\in\mathcal{A}$, there exists $B\in\mathcal{B}$ such that $\mu\big(A\,\triangle\, f^{-1}(B)\big) = 0$.*

This theorem and Theorem 9.12.23 yield easily the following fact (see Graf [**719**]).

9.12.4. Corollary. *Let X be a Hausdorff space with a Radon probability measure μ, let $\mathcal{A} = \mathcal{B}(X)$, and let $f\colon X\to Y$ be an $(\mathcal{A},\mathcal{B})$-measurable mapping. The following conditions are equivalent*:

(i) *μ is an extreme point of M^ν*;

(ii) *there exists a $(\mathcal{B}_\nu,\mathcal{A})$-measurable weak section $\pi\colon Y\to X$ of the mapping f such that $\mu = \nu\circ\pi^{-1}$.*

If f is surjective and $\mathcal{B}$ is countably separated, then conditions (i) *and* (ii) *are also equivalent to the following condition*:

(iii) *there exists a* $(\mathcal{B}_\nu, \mathcal{A})$*-measurable section* $\pi\colon Y \to X$ *of the mapping* f *with* $\mu = \nu \circ \pi^{-1}$.

Finally, if, in addition, $\mathcal{A}$ *is countably generated and for some* σ*-algebra* $\mathfrak{S}$ *with* $\mathcal{B} \subset \mathfrak{S} \subset \mathcal{B}_\nu$ *there exists an* $(\mathfrak{S}, \mathcal{A})$*-measurable section of the mapping* f*, then the indicated conditions are equivalent to the following condition:*

(iv) *there exists an* $(\mathfrak{S}, \mathcal{A})$*-measurable section* π *of the mapping* f *such that* $\mu = \nu \circ \pi^{-1}$.

9.12.5. Example. The most interesting for applications is the case where X and Y are Souslin spaces with their Borel σ-algebras and $f\colon X \to Y$ is a surjective Borel mapping. Then the conditions formulated before assertion (iv) are fulfilled if we take for $\mathfrak{S}$ the σ-algebra generated by all Souslin sets. Thus, in this situation, the extreme points of the set M^ν are exactly the measures of the form $\nu \circ \pi^{-1}$, where $\pi\colon Y \to X$ is measurable with respect to $(\mathfrak{S}, \mathcal{A})$ and $f\bigl(\pi(y)\bigr) = y$ for all $y \in Y$.

It was shown in Graf [**719**] that a parameterization of measurable sections of the mapping π by preimages of the measure μ can be made measurable in a certain natural sense. About representation of preimages in the form of images with respect to measurable sections, see also Hackenbroch [**761**]. The following generalization of Corollary 9.12.4 was obtained in Rinkewitz [**1580**].

9.12.6. Theorem. *Let* μ *be an* $\aleph$*-compact probability measure on* $\mathcal{A}$ *such that* $\mu \circ f^{-1} = \nu$*. Then the following conditions are equivalent:*

(i) μ *is an extreme point in* M^ν;

(ii) *there exists a measurable weak section* π *of* f *such that* $\nu \circ \pi^{-1} = \mu$.

Moreover, the measure ν *is* $\aleph$*-compact as well.*

The condition of $\aleph$-compactness in this theorem cannot be weakened to the compactness in our sense. For example, one can take for $(X, \mathcal{A}, \mu)$ the interval $[0,1]$ with the σ-algebra of all at most countable sets and their complements and equip it with the measure that equals 1 on the complements of countable sets. Let $Y = [0,1]$, $\mathcal{B} = \{\varnothing, [0,1]\}$, $\nu(X) = 1$, and let f be the identity mapping. Any $\mathcal{B}_\nu$-measurable function is constant, hence it transforms ν into Dirac's measure, and μ cannot be the image of ν. Here one has $\mu \in M^\nu$. Indeed, if $\mu = (\mu_1 + \mu_2)/2$, where μ_1 and μ_2 are probability measures on $\mathcal{A}$, then $\mu_1(C) = \mu_2(C) = 0$ for every countable set C, which yields $\mu_1 = \mu_2 = \mu$. It is shown in Rinkewitz [**1580**] that if a measure ν on $\mathcal{B}$ is $\aleph$-compact, then the set of all extreme points of the collection of all $\aleph$-compact probability measures μ on $\mathcal{A}$ such that $\mu \circ f^{-1} = \nu$ coincides with the set of images of ν under measurable weak sections of f.

Now we continue a discussion of the uniqueness problem for preimages of measures started in §9.8 and consider three different characterizations of uniqueness given by Ershov [**539**], Eisele [**525**], and Lehn, Mägerl [**1146**] for Souslin spaces and by Graf [**720**] in a more general situation. Our presentation follows Bogachev, Sadovnichiĭ, Fedorchuk [**224**]. Let ν be a probability measure on $(Y, \mathcal{B})$, let $f\colon (X, \mathcal{A}) \to (Y, \mathcal{B})$ be measurable, and let $\nu^*\bigl(f(X)\bigr) = 1$.

The measure ν_0 on $f^{-1}(\mathcal{B})$ is defined in §9.8 by $\nu_0(B) := \nu\bigl(f^{-1}(B)\bigr)$. Ershov [**539**] introduced the following condition of uniqueness:

$$\text{the completion of } f^{-1}(\mathcal{B}) \text{ with respect to } \nu_0 \text{ contains } \mathcal{A}. \qquad (\mathrm{E}_1)$$

In other words, for every $A \in \mathcal{A}$ there exist sets B_1 and B_2 in $\mathcal{B}$ such that $\nu(B_2) = 0$ and $A \triangle f^{-1}(B_1) \subset f^{-1}(B_2)$. It is obvious that this condition ensures the existence and uniqueness of a preimage. Indeed, let μ be the restriction of the completion of ν_0 to $\mathcal{A}$. Then μ is a preimage, and every preimage coincides on $\mathcal{A}$ with the completion of ν_0 due to (E_1).

Another condition was studied in Eisele [**525**] and Lehn, Mägerl [**1146**]:

$$\text{for } \nu\text{-almost every } y, \text{ the set } f^{-1}(y) \text{ is a singleton.} \qquad (\mathrm{E}_2)$$

Given a class of sets $\mathcal{K} \subset \mathcal{A}$, let us introduce the following condition:

$$\nu\bigl(f(K_1) \cap f(K_2)\bigr) = 0 \quad \text{if } K_1, K_2 \in \mathcal{K} \text{ and } K_1 \cap K_2 = \varnothing. \qquad (\mathrm{U}_{\mathcal{K}})$$

This condition is very close to the following condition from Graf [**720**]: for any disjoint compacts K_1 and K_2 in a topological space X, one has the equality $\nu^*\bigl(f(K_1) \cup f(K_2)\bigr) = \nu^*\bigl(f(K_1)\bigr) + \nu^*\bigl(f(K_2)\bigr)$, where the measure ν and the mapping f are subject to certain technical restrictions.

It is clear that condition $(\mathrm{U}_{\mathcal{K}})$ is weakened when we make the class $\mathcal{K}$ smaller. It becomes the most restrictive when we set $\mathcal{K} = \mathcal{A}$.

9.12.7. Theorem. (i) *Condition* (E_1) *implies condition* $(\mathrm{U}_{\mathcal{A}})$*, hence also condition* $(\mathrm{U}_{\mathcal{K}})$ *for every class* $\mathcal{K} \subset \mathcal{A}$*.*

(ii) *Condition* (E_2) *implies condition* $(\mathrm{U}_{\mathcal{A}})$*, hence also condition* $(\mathrm{U}_{\mathcal{K}})$ *for every class* $\mathcal{K} \subset \mathcal{A}$*.*

(iii) *Condition* (E_2) *implies condition* (E_1) *if the images of sets in* $\mathcal{A}$ *are* ν*-measurable. More generally, condition* (E_2) *implies condition* (E_1) *if the* σ*-algebra* $\mathcal{A}$ *is generated by a class of sets whose images are* ν*-measurable.*

(iv) *Let* $\mathcal{A} = \sigma(\mathcal{K})$ *and let the images of sets in* $\mathcal{K}$ *be* ν*-measurable. Then condition* (E_1) *is equivalent to condition* $(\mathrm{U}_{\mathcal{K}\cup\mathcal{K}^c})$*, where* $\mathcal{K}^c$ *is the class of complements of sets in* $\mathcal{K}$*. Condition* $(\mathrm{U}_{\mathcal{K}\cup\mathcal{K}^c})$ *can be written in the form*

$$\nu\bigl(f(K) \cap f(X\backslash K)\bigr) = 0 \quad \forall\, K \in \mathcal{K}.$$

In particular, conditions $(\mathrm{U}_{\mathcal{K}})$ *and* (E_1) *are equivalent if* $\mathcal{A} = \sigma(\mathcal{K})$*, the images of sets in* $\mathcal{K}$ *are* ν*-measurable, and for every* $K \in \mathcal{K}$ *there exist sets* $K_n \in \mathcal{K}$ *with* $X\backslash K = \bigcup_{n=1}^{\infty} K_n$ *(the latter is fulfilled if the class* $\mathcal{K}$ *is closed with respect to complementation).*

(v) *Let* X *and* Y *be Souslin spaces equipped with their Borel* σ*-algebras, let* $f\colon X \to Y$ *be a Borel surjection, and let* ν *be a Borel probability measure on* Y*. Then either of conditions* (E_1)*,* (E_2) *and* $(\mathrm{U}_{\mathcal{K}})$ *with the class of all compact sets is equivalent to the uniqueness of a preimage of* ν *in the class of Borel probability measures. In particular, all the three conditions are equivalent.*

Proof. (i) Let (E_1) be fulfilled and let $A_1, A_2 \in \mathcal{A}$ be disjoint. By assumption there exist sets $B_1, B_2, C_1, C_2 \in \mathcal{B}$ such that one has the equality

$\nu(C_1) = \nu(C_2) = 0$ and the inclusions $f^{-1}(B_1) \subset A_1$, $A_1 \backslash f^{-1}(B_1) \subset f^{-1}(C_1)$, $f^{-1}(B_2) \subset A_2$, and $A_2 \backslash f^{-1}(B_2) \subset f^{-1}(C_2)$. Then $\nu(B_1 \cap B_2) = 0$. Since $f(A_1) \cap f(A_2) \subset C_1 \cup C_2 \cup (B_1 \cap B_2)$, we obtain $\nu\bigl(f(A_1) \cap f(A_2)\bigr) = 0$.

(ii) If $K_1, K_2 \subset X$ and $K_1 \cap K_2 = \varnothing$, then $f(K_1) \cap f(K_2)$ is a subset of the set of points with a non-unique preimage.

(iii) Let condition (E_2) be fulfilled and let $A \in \mathcal{A}$. By assumption, in the ν-measurable set $f(A)$, the subset M of points with a non-unique preimage has ν-measure zero. Let us take a set $B \in \mathcal{B}$ with $B \subset f(A) \backslash M$ and $\nu(B) = \nu\bigl(f(A)\bigr)$. Then there exists a set $N \in \mathcal{B}$ of ν-measure zero that contains $f(A) \backslash B$. Let $E = f^{-1}(B)$. We obtain that $E \triangle A \subset f^{-1}(N)$. Thus, the set A belongs to the completion of $f^{-1}(\mathcal{B})$ with respect to the measure ν_0. The same reasoning proves a more general assertion, where one requires the existence of a class $\mathcal{K}$ of sets generating the σ-algebra $\mathcal{A}$ and having ν-measurable images. Indeed, in this case we obtain the ν_0-measurability of the sets in $\mathcal{K}$, which gives the ν_0-measurability of the sets in $\mathcal{A}$. Below we give an example showing that the ν-measurability of the images of sets from $\mathcal{A}$ (or, at least, from a class generating $\mathcal{A}$) is essential for the validity of the established implication.

(iv) Let condition ($\mathrm{U}_{\mathcal{K} \cup \mathcal{K}^c}$) be fulfilled and let $K \in \mathcal{K}$. We take sets $B, C_1, C_2 \in \mathcal{B}$ such that the relations $B = f(K) \cup C_1$, $f(K) \cap f(X \backslash K) \subset C_2$, and $\nu(C_1) = \nu(C_2) = 0$ hold. The set K differs from the set $f^{-1}(B)$ in a subset of the set $f^{-1}(C_1 \cup C_2)$ that has ν_0-measure zero. Hence K is measurable with respect to ν_0. The second claim in (iv) follows from the first one, since $f(X \backslash K) = \bigcup_{n=1}^{\infty} f(K_n)$. Note that the first claim in (iv) does not assume the ν-measurability of the images of the complements of sets in $\mathcal{K}$.

(v) We know from Proposition 9.8.4 that (E_1) is equivalent to the uniqueness of a preimage. In addition, (E_2) implies (E_1). Suppose that (E_1) is not fulfilled. The set M of all points in Y with more than one preimage is Souslin along with the set $S := f^{-1}(M)$. Since $\nu(M) > 0$, by the measurable selection theorem we can find a Borel set $B \subset M$ with $\nu(B) = \nu(M)$ and a Borel set $A \subset S$ such that f maps A onto B and is one-to-one. Clearly, $f(S \backslash A) = B$. There exist nonnegative measures σ_1 and σ_2 on A and $S \backslash A$, respectively, such that their images under f coincide with $\nu|_B$. Since f is sujective, there is some preimage μ of ν. By using the measures σ_1 and σ_2 one can redefine μ on A in two different ways and obtain two different preimages of ν. Hence (E_1) and (E_2) are equivalent. We know that either of them implies ($\mathrm{U}_{\mathcal{K}}$). Now let ($\mathrm{U}_{\mathcal{K}}$) be fulfilled. We observe that (E_2) is fulfilled as well. Indeed, otherwise it is easily seen from the above reasoning that one can find compact sets $K_1 \subset A$ and $K_2 \subset S \backslash A$ such that $\nu\bigl(f(K_1) \cap f(K_2)\bigr) > 0$, which contradicts ($\mathrm{U}_{\mathcal{K}}$). $\square$

The restrictions on f and $\mathcal{K}$ indicated in the second part of (iv) are fulfilled for continuous mappings of compact spaces and Radon measures if we take for $\mathcal{K}$ the class of functionally closed sets. Indeed, the complement of a functionally closed set is a countable union of functionally closed sets. In particular, if X is perfectly normal, then the whole class of compact sets can be taken.

In Example 9.12.14 given below, conditions (E_2) and $(\mathrm{U}_{\mathcal{A}})$ are fulfilled, but condition (E_1) is not. For the projection f of the space "two arrows" to $[0,1]$ with Lebesgue measure, condition (E_1) and condition $(\mathrm{U}_{\mathcal{A}})$ are fulfilled for the class $\mathcal{A}$ of all Borel sets of this space. Indeed, every such Borel set B differs in an at most countable set from a set of the form $f^{-1}(B_0)$, where $B_0 \in \mathcal{B}([0,1])$ (see Exercise 6.10.36). Hence the projections of two disjoint Borel sets have an at most countable intersection. In addition, the set $B \bigtriangleup f^{-1}(B_0)$ is at most countable and has measure zero with respect to ν_0. The projections of all sets from $\mathcal{A}$ are Borel in the interval. However, condition (E_2) is not fulfilled: every point in $(0,1)$ has two preimages. Thus, in assertion (iii), conditions (E_1) and (E_2) are not equivalent even when the images of all sets in $\mathcal{A}$ are measurable with respect to ν. In Example 9.12.12, a continuous surjection of a compact space onto $[0,1]$ satisfies condition $(\mathrm{U}_{\mathcal{K}})$ with the class of all compact sets, but condition (E_1) is not fulfilled. Hence in assertion (iv) one cannot omit additional assumptions that ensure the equivalence of (E_1) and $(\mathrm{U}_{\mathcal{K}})$.

A simple proof of the following result is given in Bogachev, Sadovnichiĭ, Fedorchuk [**224**].

9.12.8. Theorem. *Let a class $\mathcal{K} \subset \mathcal{A}$ be such that the images of sets in $\mathcal{K}$ are ν-measurable. Suppose that condition $(\mathrm{U}_{\mathcal{K}})$ is fulfilled. If ν has a preimage in the set of those probability measures on $\mathcal{A}$ for which $\mathcal{K}$ is an approximating class, then there are no other preimages in this set.*

9.12.9. Example. Suppose X is a topological space, $\mathcal{A} = \mathcal{B}(X)$, $\mathcal{F}$ is the class of all closed sets. Let f satisfy condition $(\mathrm{U}_{\mathcal{F}})$ with respect to ν and let the images of all closed sets be ν-measurable. If two regular Borel probability measures μ_1 and μ_2 on X are preimages of ν, then $\mu_1 = \mu_2$. If the measures μ_1 and μ_2 are Radon and the images of all compact sets are ν-measurable, then $\mu_1 = \mu_2$ provided that condition $(\mathrm{U}_{\mathcal{K}})$ is fulfilled with the class $\mathcal{K}$ of all compact sets.

9.12.10. Proposition. *Suppose that X and Y are topological spaces and that $f\colon X \to Y$ is a continuous mapping. Let μ be a Radon probability measure on X and let $\nu = \mu\circ f^{-1}$. Then condition $(\mathrm{U}_{\mathcal{K}})$ with the class $\mathcal{K}$ of all compact sets is necessary and sufficient for the uniqueness of a preimage of ν in the class of Radon probability measures.*

PROOF. This fact follows from Graf [**720**, Theorem 5.5], but can be verified directly. Indeed, suppose we are given disjoint compact sets K_1 and K_2 with $\nu\bigl(f(K_1)\cap f(K_2)\bigr) > 0$. Then the compact sets

$$S_1 := K_1 \cap f^{-1}\bigl(f(K_1)\cap f(K_2)\bigr) \quad \text{and} \quad S_2 := K_2 \cap f^{-1}\bigl(f(K_1)\cap f(K_2)\bigr)$$

do not meet and $f(S_1) = f(S_2) = f(K_1)\cap f(K_2)$. There exist nonnegative Radon measures σ_1 and σ_2 on S_1 and S_2 that are transformed by f to the restriction of the measure ν on $f(S_1) = f(S_2) = f(K_1)\cap f(K_2)$. By using σ_1 and σ_2 we can redefine μ on $S_1\cup S_2$ and obtain two distinct preimages of ν. The converse follows by Example 9.12.9. $\square$

The following simple example shows that in the general case condition (E_1) is not necessary for the uniqueness of a Radon preimage of a measure.

9.12.11. Example. Let $X = Y$ be the product of the continuum of copies of $[0,1]$, let $\mathcal{A} = \mathcal{B}(X)$, and let $\mathcal{B}$ be the Baire σ-algebra of the space X. Let us take for f the identity mapping and for ν Dirac's measure δ at zero. Then the unique Radon preimage of ν is the same Dirac measure, condition (E_2) is fulfilled, but condition (E_1) is broken because the set consisting of the single point zero does not belong to the completion of the Baire σ-algebra with respect to the measure δ (its outer measure equals one and its inner measure equals zero, since it is not a Baire set). However, in this example, there are non-regular Borel probability preimages of ν.

A situation is also possible when for a one-to-one mapping (E_1) is not fulfilled (conditions (E_2) and $(\mathrm{U}_{\mathcal{A}})$ are fulfilled, of course), and there are no Radon preimages, but a Borel preimage is not unique: see Example 8.10.29.

Now under the assumption that the cardinality of the continuum is not measurable, which means the absence of nonzero measures without points of positive measure on the class of all subsets of an interval (for which it suffices to accept the continuum hypothesis or Martin's axiom), we give an example of a continuous surjection of a compact space onto $[0,1]$ such that Lebesgue measure has a unique preimage in the whole class of Borel probability measures and condition $(\mathrm{U}_{\mathcal{K}})$ is fulfilled, but (E_1) and (E_2) are not fulfilled.

9.12.12. Example. Let the set $X = [0,1]^2$ be equipped with the order topology with respect to the lexicographic ordering as in Exercise 6.10.87 and let $\mathcal{A} = \mathcal{B}(X)$. Then X is compact and the natural projection $f\colon X \to [0,1]$ is continuous. The space "two arrows", denoted by X_0, is closed in X. Let $\nu = \lambda$ be Lebesgue measure on $[0,1]$. Condition (E_1) is broken, since the interior U of the square with the usual topology is open in the order topology, but does not belong to the completion of $f^{-1}\big(\mathcal{B}([0,1])\big)$ with respect to ν_0. Indeed, the only set from $f^{-1}\big(\mathcal{B}([0,1])\big)$ containing U is the whole space X, but their difference has full outer measure with respect to ν_0. Let us show that ν has a unique preimage in the class of Borel probability measures if there are no nonzero measures without points of positive measure on the class of all sets in $[0,1]$. This unique preimage is the Radon probability measure μ concentrated on the subspace X_0 and having the projection λ. Let μ_1 be another probability Borel preimage. The sets $\{x\}\times(0,1)$ are open in X and have zero μ_1-measure because their projections are points, which have zero Lebesgue measure. It follows from our assumption that $\mu_1(X\backslash X_0) = 0$, since otherwise on the class of all subsets of the interval we obtain a nonzero measure $\sigma(E) := \mu_1\big(E\times(0,1)\big)$ without points of positive measure. Now we have to verify that there is only one Borel probability measure on X_0 whose projection is ν. This is seen from the fact (see Exercise 6.10.36) that every Borel set B in X_0 differs in an at most countable set from a set of the form $f^{-1}(B_0)$, where $B_0 \in \mathcal{B}([0,1])$. Hence $\mu_1(B) = \nu(B_0)$, since μ_1 has no points of positive measure. It is easily seen that condition $(\mathrm{U}_\mathcal{K})$ is fulfilled with the class $\mathcal{K}$ of all compacts, but (E_2) and $(\mathrm{U}_\mathcal{A})$ are not fulfilled.

Without additional set-theoretic assumptions one can find a continuous surjection of a compact X onto a metrizable compact Y with a probability measure ν such that ν has only one Radon probability preimage, but there are non-regular Borel probability preimages.

9.12.13. Example. In Section 439J of volume 4 of the book Fremlin [**635**] the set $X = [0,1]^\infty \times \{0,1\}$ is given some topology τ with the following properties:

(1) (X,τ) is a compact space with the first countability axiom, and the natural projection $\pi\colon (x,y)\mapsto x$ of the space X onto $[0,1]^\infty$ with the usual topology τ_0 of a countable product of closed intervals (in which it is a metrizable compact) is continuous, (2) the set $[0,1]^\infty \times \{0\}$ is compact in X, and the topology τ on this set coincides with τ_0, (3) subsets of $[0,1]^\infty \times \{1\}$ that are compact in the topology τ are finite or countable, (4) there is a Borel probability measure μ on X that is not Radon, but is mapped by π to the measure ν equal the countable power of Lebesgue measure on $[0,1]$.

It is clear that besides μ, the same measure $\mu_0 = \nu$ transported to the subspace $[0,1]^\infty \times \{0\}$ is mapped to the measure ν. Thus, there are distinct preimages in the class of all probability Borel measures on X. However, the only preimage in the subclass of Radon measures is μ_0. Indeed, let μ' be another Radon preimage. Then μ' cannot have points of positive measure, which by property (3) yields the equality $\mu'([0,1]^\infty{\times}\{1\}) = 0$, i.e., the measure μ' is concentrated on $[0,1]^\infty \times \{0\}$. Therefore, $\mu' = \mu_0$ because the mapping π is a homeomorphism between $[0,1]^\infty \times \{0\}$ and $[0,1]^\infty$.

In this example, too, condition $(\mathrm{U}_\mathcal{K})$ is fulfilled with the class $\mathcal{K}$ of all compacts and conditions (E_1) and (E_2) are not fulfilled.

Thus, in the case of continuous surjections of compacts, conditions (E_1) and (E_2) are not necessary for the uniqueness of a preimage in the class of Borel probability measures. Here $(\mathrm{E}_2) \Rightarrow (\mathrm{E}_1) \Rightarrow (\mathrm{U}_\mathcal{K})$, where the implications are not invertible, and condition $(\mathrm{U}_\mathcal{K})$ is necessary and sufficient for the uniqueness of a preimage in the class of Radon probability measures.

If one does not confine oneself to continuous surjections of compacts, then one can give an example where condition $(\mathrm{U}_\mathcal{A})$ is fulfilled and ν has exactly one probability preimage, but (E_1) is not fulfilled. According to assertion (iv) of Theorem 9.12.7, this would be impossible under the additional assumption of the ν-measurability of the images of the sets in $\mathcal{A}$ (observe that in this case Theorem 9.12.8 ensures the uniqueness of a probability preimage provided that such a preimage exists).

9.12.14. Example. (i) Let $X = Y = [0,1]$, $\mathcal{B} = \mathcal{B}([0,1])$, and let $\nu = \lambda$ be Lebesgue measure. Let us take a Lebesgue nonmeasurable set $E \subset [0,1]$ of cardinality of the continuum with $\lambda_*(E) = 0$. Let $\mathcal{A}$ be the σ-algebra generated by all Borel sets in $[0,1]$ and all subsets of E. Note that the measure

λ extends to a measure μ on $\mathcal{A}$ that satisfies the condition $\mu(E) = 0$. Indeed, by the equality $\lambda_*(E) = 0$ the measure λ has an extension λ' to the σ-algebra $\mathcal{E}$ generated by $\mathcal{B}([0,1])$ and E such that $\lambda'(E) = 0$ (see Theorem 1.12.14). Then all subsets of E are measurable with respect to the completion of λ', i.e., one can take for μ the restriction of the completion of λ' to $\mathcal{A}$. Under the assumption that the cardinality of the continuum is not measurable there are no other extensions to the class of all subsets of E. Let f be the identity mapping $([0,1], \mathcal{A}) \to ([0,1], \mathcal{B})$. Then condition $(\mathrm{U}_{\mathcal{A}})$ is obviously fulfilled, but (E_1) is not.

(ii) Under the continuum hypothesis, it is easy to modify the example in (i) in such a way that X becomes a separable metric space whose identity embedding into the interval is continuous. Indeed, according to Corollary 3.10.3, under the continuum hypothesis E contains a countable collection of sets E_n such that the generated σ-algebra $\sigma(\{E_n\})$ contains $\mathcal{B}(E)$, but carries no nonzero measure vanishing on all one point sets. Let $\mathcal{A}$ be the σ-algebra generated by all Borel sets in $[0,1]$ and all E_n. By the same reasoning as above, Lebesgue measure has a unique extension to $\mathcal{A}$. Now we equip X with a countable topology base that consists of the rational intervals intersected with X and the sets E_n.

(iii) A close example is possible without additional set-theoretic assumptions. Take for $\mathcal{B}$ the σ-algebra consisting of the first category sets in $[0,1]$ and their complements. Let $\nu(B) = 0$ for all first category sets B and $\nu(B) = 1$ in the opposite case. Note that $\nu_*([0,1/2]) = 0$. Take for $\mathcal{A}$ the σ-algebra generated by $\mathcal{B}$ and all Borel subsets of $[0,1/2]$. As above, ν has an extension μ to $\mathcal{A}$ with $\mu([0,1/2]) = 0$. There are no other extensions, since every Borel measure on $[0,1/2]$ is concentrated on a first category set.

9.12(iii). Existence of atomless measures

Here we give two results on existence of atomless measures.

9.12.15. Proposition. *Let K be a nonempty compact space without isolated points. Then, there exists an atomless Radon probability measure on K.*

PROOF. We give two different proofs. The first one is based on the fact that there exists a continuous surjective mapping f from K onto $[0,1]$ (see Exercise 6.10.26). For the required measure one can take any Radon probability measure whose image is Lebesgue measure (such a measure exists according to Theorem 9.1.9).

Another reasoning, used in Knowles [**1014**], is based on the fact that the space $\mathcal{P}_r(K)$ of all Radon probability measures on K is compact in the weak topology. Hence it cannot be represented as the union of a sequence of nowhere dense closed sets. Let us consider the sets M_n consisting of all measures $\mu \in \mathcal{P}_r(K)$ that have atoms of measure at least $1/n$. The sets M_n are closed in $\mathcal{P}_r(K)$ with the weak topology. Indeed, let ν be a limit point of M_n.

There is a net of measures $\mu_\alpha \in M_n$ convergent to ν. Every measure μ_α has a point x_α of measure at least $1/n$. The points $\{x_\alpha\}$ have a limit point x, hence we may assume that the net $\{x_\alpha\}$ converges to x. If $\nu(\{x\}) < 1/n$, then there exists a closed set Z whose interior contains x and $\nu(Z) < 1/n$. There is α_0 such that $x_\alpha \in Z$ if $\alpha \geq \alpha_0$. Hence $\mu_\alpha(Z) \geq 1/n$, which by weak convergence yields $\nu(Z) \geq 1/n$, a contradiction. In addition, the sets M_n are nowhere dense. Indeed, let $\nu \in \mathcal{P}_r(K)$. Every neighborhood of ν in the weak topology contains a finite linear combination of Dirac measures. Since K has no isolated points, such a combination can be found in the form $\nu_0 = \sum_{j=1}^k c_j \delta_{a_j}$ where $c_j < (2n)^{-1}$ and the points a_j are distinct. As shown above, ν_0 has a neighborhood that does not meet M_n. □

9.12.16. Proposition. *Let K be a compact space. One can find an atomless Radon probability measure on K precisely when there exists a continuous function $f\colon K \to [0,1]$ with $f(K) = [0,1]$.*

Proof. If such a function exists, then Theorem 9.1.9 applies. If there is an atomless Radon probability measure μ on K, then the topological support of μ is a compact set K_0 without isolated points. According to Exercise 6.10.26, there exists a continuous function f on K_0 with $f(K_0) = [0,1]$. It remains to extend f to a continuous function from K to $[0,1]$. □

9.12(iv). Invariant and quasi-invariant measures of transformations

Let f be a Borel mapping from a topological space X into itself. We recall that a Borel measure μ on X is called *an invariant measure* of the transformation f if one has $\mu \circ f^{-1} = \mu$. The problem of existence of invariant measures of transformations arises in probability theory, ergodic theory, nonlinear analysis, the theory of representations of groups, statistical physics, and many other branches of mathematics and physics. The following fundamental result goes back to N.N. Bogolubov and N.M. Krylov [**227**].

9.12.17. Theorem. *Let $\{T_\alpha\}$ be a family of commuting continuous mappings of a compact space X into itself. Then, there exists a Radon probability measure λ on X that is invariant with respect to all T_α.*

Proof. According to the Riesz theorem, the space $C(X)^*$ can be identified with the space of all Radon measures on X. Any continuous mapping $T\colon X \to X$ induces a linear mapping $\widehat{T}\colon C(X)^* \to C(X)^*$, $\lambda \mapsto \lambda \circ T^{-1}$, which is continuous if $C(X)^*$ is equipped with the weak* topology. Indeed, let $U := \bigl\{\lambda\colon\ -\varepsilon < \lambda(f_i) < \varepsilon,\ i = 1,\ldots,n\bigr\}$, where $f_i \in C(X)$ and $\lambda(f)$ denotes the integral of f against the measure λ. Then $\widehat{T}^{-1}(U)$ contains the neighborhood of zero $\bigl\{m\colon\ -\varepsilon < m(f_i \circ T) < \varepsilon,\ i = 1,\ldots,n\bigr\}$ because $m(f\circ T) = \widehat{T}(m)(f)$. By the Banach–Alaoglu theorem, the closed unit ball in $C(X)^*$ is compact in the weak* topology. Its subset P consisting of functionals L such that $L(1) = 1$ and $L(f) \geq 0$ whenever $f \geq 0$ (i.e., corresponding to probability measures) is closed and convex. Therefore, it is a convex compact set. The continuous linear mappings $\widehat{T}_\alpha$ take P to P and commute. According to the well-known Markov–Kakutani theorem (see Edwards [**518**, Theorem 3.2.1]), there exists a point $\lambda \in P$ such that $\widehat{T}_\alpha(\lambda) = \lambda$ for all α. Thus, λ is a common invariant measure of all T_α. □

9.12.18. Corollary. *Every continuous mapping of a compact space into itself has an invariant Radon probability measure.*

An immediate corollary of Theorem 9.12.17 is the existence of a Haar measure on every commutative compact topological group, i.e., a Radon probability measure invariant with respect to translations.

9.12.19. Example. Let a bounded set K in a Hilbert space be closed in the weak topology. Then any continuous in the weak topology mapping $F\colon K \to K$ has an invariant probability measure.

In this example, it is important that the set is closed in the weak topology as well as that the mapping is continuous in this topology. Let us consider the following example from Bogachev, Prostov [**221**] (an analogous, but not polynomial, mapping was used by Kakutani in his example of a homeomorphism of the ball without fixed points).

9.12.20. Example. There exists a mapping f of the closed unit ball U in l^2 into itself such that f is a diffeomorphism (i.e., a diffeomorphism of some neighborhoods of U and a homeomorphism of U) and, in addition, a second-order polynomial, i.e., $f(x) = B(x,x) + A(x) + c$, where B is bilinear, A is linear, $c \in U$, but has no invariant measures.

Proof. Let us represent l^2 as the space of two-sided sequences $x = (x_n)$, $n \in \mathbb{Z}$, take its natural basis $\{e_n\}$, denote by T the isometry defined by $Te_n = e_{n-1}$ and let

$$f(x) = T\bigl(x + \varepsilon\bigl(1 - (x,x)\bigr)e_0\bigr),$$

where $\varepsilon \in (0, 1/2)$. All our claims are verified directly (see [**221**]), in particular, the absence of invariant measures follows by the fact that, as one can verify, for every x, the sequence $f^n(x)$ converges weakly to 0, but Dirac's measure at 0 is not invariant. If we consider T on the unit sphere, then we obtain a mapping that is weakly continuous, but has no invariant measures. Certainly, the reason is that the sphere is not weakly closed. □

It would be interesting to find conditions on a smooth mapping (different from its compactness) that ensure the existence of invariant measures. In some applications, the weaker property of quasi-invariance is more useful. For example, there exist no finite invariant Haar measures on noncompact topological groups. We shall say that μ is a *quasi-invariant measure* of a family of transformations $\{T_\alpha\}$ if $\mu \circ T_\alpha^{-1} \ll \mu$ for all α. It is clear that for a single transformation T, one can always find a quasi-invariant probability measure:

let $\mu = \sum_{n=1}^{\infty} 2^{-n}\mu \circ (T^n)^{-1}$, where μ is any probability measure. However, in the general case this is often a difficult problem. Certainly, there exist families that have no quasi-invariant measures at all. A non-trivial example is the additive group of an infinite-dimensional Banach space: it does not admit nonzero quasi-invariant finite Borel measures. The concepts of invariance and quasi-invariance are meaningful for transformations of spaces of measures on X that are not necessarily generated by transformations of the space X itself. For example, invariant measures of a stochastic process in a topological space X with the transition semigroup $\{T_t\}$ on the space of bounded Borel functions are defined as invariant measures of the associated operators T_t^* on $\mathcal{M}(X)$. Regarding extensions of Haar measures, see Hewitt, Ross [**825**, §16]. In the consideration of infinite Haar measures it is sometimes more convenient to deal with invariant integrals, rather than with measures. This is one of the situations where one can exploit advantages of the Daniell–Stone approach.

9.12(v). Point and Boolean isomorphisms

Many papers are devoted to generalizations of a result due to von Neumann (see Theorem 9.5.1), according to which any automorphism of a measure algebra is generated by a mapping of the measure space under some restrictions on a measure or a space; see Choksi [**345**], [**346**], Choksi, Fremlin [**347**], Maharam [**1232**]. In Maharam [**1233**], for any Radon probability measure μ, one constructs an isomorphism of the measures $\mu\otimes\lambda^\tau$ and λ^τ, where λ^τ is some power of Lebesgue measure on $[0,1]$. We mention a result from Choksi, Fremlin [**347**].

9.12.21. Theorem. *Suppose that X_α, $\alpha\in A$, are compact metric spaces. Let $X=\prod_{\alpha\in A} X_\alpha$ and let μ and ν be Radon probability measures on X. If the measure algebras E_μ and E_ν are isomorphic in the sense of Definition* 9.3.1*, then there exists an isomorphism* mod0 *of the measure spaces $\bigl(X, \mathcal{B}a(X)_\mu, \mu\bigr)$ and $\bigl(X, \mathcal{B}a(X)_\nu, \nu\bigr)$.*

In particular, if A is at most countable, then there exists an isomorphism mod0 *of the spaces $\bigl(X, \mathcal{B}(X)_\mu, \mu\bigr)$ and $\bigl(X, \mathcal{B}(X)_\nu, \nu\bigr)$.*

For uncountable products of unit intervals, the last assertion is false, as shown in Panzone, Segovia [**1421**]. According to Vinokurov [**1929**], two infinite products (of the same cardinality) of atomless Lebesgue spaces are isomorphic mod0 provided that they have equal metric structures. In addition, every power E^τ of an atomless Lebesgue space that generates a homogeneous metric measure algebra of the weight τ is point isomorphic mod0 to the compact space $[0,1]^\tau$.

Note the following result (see Fremlin [**635**, §344I]).

9.12.22. Theorem. *Let $(X,\mathcal{A},\mu)$ and $(Y,\mathcal{B},\nu)$ be atomless perfect probability measures on countably separated σ-algebras. Then the measure spaces $(X,\mathcal{A}_\mu,\mu)$ and $(Y,\mathcal{B}_\nu,\nu)$ are isomorphic.*

Let $(X, \mathcal{A}, \mu)$ be a complete probability space, Y a Hausdorff space, $\mathcal{B} = \mathcal{B}(Y)$. The next interesting result is obtained in Graf [**719**] by using important ideas from Edgar [**512**]. If $\mathfrak{A}_1$ and $\mathfrak{A}_2$ are Boolean algebras, then a mapping $\Phi\colon \mathfrak{A}_1 \to \mathfrak{A}_2$ is called a Boolean σ-homomorphism if Φ preserves the operations of intersection, complementation, and countable union.

9.12.23. Theorem. *Let $\Phi\colon \mathcal{B} \to \mathcal{A}/\mu$ be a Boolean σ-homomorphism such that $\mu \circ \Phi$ is a Radon measure on Y. Then, there exists an $(\mathcal{A}, \mathcal{B})$-measurable mapping $f\colon X \to Y$ such that $\Phi(B)$ is the equivalence class of the set $f^{-1}(B)$ for every $B \in \mathcal{B}$, i.e., the mapping f induces Φ.*

PROOF. Denote by $\mathcal{K}$ the class of all compact sets in Y. As will be shown in §10.5, there exists a lifting $L\colon \mathcal{A}/\mu \to \mathcal{A}$, i.e., a mapping L that associates to every class of μ-equivalent sets (we recall that $\mathcal{A} = \mathcal{A}_\mu$) a representative of this class in such a way that

$$L(X) = X,\ L(\varnothing) = \varnothing,\ L(A \cap B) = L(A) \cap L(B),\ L(A \cup B) = L(A) \cup L(B).$$

Then $\Psi = L \circ \Phi$ is a homomorphism of the Boolean algebras $\mathcal{B}$ and $\mathcal{A}$. Since the measure $\mu \circ \Phi$ is Radon, one has

$$\mu(X) = \sup\bigl\{\mu\bigl(\Phi(K)\bigr)\colon\ K \in \mathcal{K}\bigr\}.$$

The family of sets $\Psi(K)$ is an increasing (by inclusion) net, hence, according to Lemma 10.5.5, we obtain that the set $X_0 := \bigcup_{K \in \mathcal{K}} \Psi(K)$ is measurable and $\mu(X \backslash X_0) = 0$. For every $x \in X$, let $\mathcal{K}_x := \{K \in \mathcal{K}\colon\ x \in \Psi(K)\}$. We observe that for any $x \in X_0$ the class $\mathcal{K}_x$ is nonempty. We show that $\Pi_x := \bigcap_{K \in \mathcal{K}_x} K$ consists of exactly one point that we denote by $f(x)$. Indeed, the class $\mathcal{K}_x$ consists of nonempty compact sets every finite intersection of which is nonempty, since their images under the homomorphism Ψ contain x. Hence the intersection of all these compact sets is nonempty as well. Suppose that Π_x contains two distinct elements y_1 and y_2. Let us take an arbitrary compact set $K \in \mathcal{K}_x$. Then $y_1, y_2 \in K$. These two points possess disjoint neighborhoods U_1 and U_2. The sets $K_1 = K \backslash U_1$ and $K_2 = K \backslash U_2$ are compact and $K = K_1 \cup K_2$. Then x belongs either to $\Psi(K_1)$ or to $\Psi(K_2)$. We may assume that $x \in \Psi(K_1)$ and then $K_1 \in \mathcal{K}_x$. This shows that U_1 does not meet Π_x, since U_1 does not meet $K_1 \supset \Pi_x$, i.e., $y_1 \notin \Pi_x$, a contradiction. Now we extend f outside X_0 by any constant value $y_0 \in Y$. We obtain a required mapping. Indeed, for every open set $U \subset Y$, the set $f^{-1}(U)$ either coincides with $E := X_0 \cap f^{-1}(U)$ or differs from E in $X \backslash X_0$. Hence it suffices to show that $E \in \mathcal{A}$. It is easy to see that the inclusion $\Pi_x \subset U$ is equivalent to that $K \subset U$ for some $K \in \mathcal{K}_x$. In addition, for every compact set $K \subset U$, we have $X_0 \cap \Psi(K) \subset E$ because if $x \in X_0 \cap \Psi(K)$, then $K \in \mathcal{K}_x$ and $\Pi_x \subset K \subset U$, i.e., $x \in E$. Thus,

$$E = X_0 \cap \bigcup \bigl\{\Psi(K)\colon\ K \subset U,\ K \text{ is compact}\bigr\}.$$

As above, we obtain that $E \in \mathcal{A}$ and $\mu\bigl(\Psi(U) \backslash E\bigr) = 0$. By the equality $\mu\bigl(f^{-1}(U) \bigtriangleup E\bigr) = 0$, we conclude that $\Phi(U)$ is the equivalence class of the set

$f^{-1}(U)$. Taking into account that Φ is a σ-homomorphism, this remains true for all Borel sets in Y. □

We observe that if Y is a Souslin space, then the measure $\mu\circ\Phi$ is automatically Radon. Moreover, in this case it is not necessary to assume the completeness of the measure μ, since one can apply the theorem to $\mathcal{A}_\mu$ and then take an $(\mathcal{A},\mathcal{B})$-measurable version of the obtained mapping.

It is clear that if a measurable mapping $T\colon (X,\mathcal{A},\mu)\to(Y,\mathcal{B},\nu)$ of probability spaces has the property that $\nu=\mu\circ T^{-1}$, then T generates a measure-preserving embedding $\Psi\colon E_\nu\to E_\mu$ such that the equivalence class of the set $B\in\mathcal{B}$ is taken to the equivalence class of the set $\Psi(B)=T^{-1}(B)$. The fact that Ψ is well-defined and $\mu\bigl(\Psi(B)\bigr)=\nu(B)$ is clear from the equality $\nu=\mu\circ T^{-1}$. In this situation, T and Ψ may not be isomorphisms.

The next result was obtained in Edgar [**512**]; Fremlin [**625**] pointed out its simple derivation from Theorem 9.12.23.

9.12.24. Theorem. *Let $(X,\mathcal{A},\mu)$ be a probability space with a complete measure μ and let $\bigl(Y,\mathcal{B}(Y)_\nu,\nu\bigr)$ be a topological space with a Radon probability measure ν. Suppose that there exists a measure-preserving mapping Ψ of the measure algebra E_ν to E_μ. Then Ψ is induced by some measurable mapping $T\colon X\to Y$.*

9.12.25. Corollary. *Let $\bigl(Y,\mathcal{B}(Y)_\nu,\nu\bigr)$ be a topological space with a Radon probability measure ν. Then, there exist a cardinal κ and a measurable mapping $T\colon \{0,1\}^\kappa\to Y$, where $\{0,1\}^\kappa$ is equipped with the measure μ that is the product of the standard Bernoulli probability measures, such that the equality $\nu=\mu\circ T^{-1}$ holds.*

Proof. We apply Theorem 9.3.5 and Theorem 9.12.24. □

Every probability measure μ can be decomposed into the sum of a purely atomic measure ν and a measure μ_0 without atoms. Then $L^p(\mu)$ is the direct sum of $L^p(\nu)$ and $L^p(\mu_0)$, and $L^p(\nu)$ can be identified with $L^p(\nu_0)$ for some measure ν_0 on $\mathbb{N}$. The structure of the second component is described by the following theorem, which is a corollary of Theorem 9.3.5.

9.12.26. Theorem. *Suppose that μ is an atomless probability measure and let $1\le p<\infty$. Then, there exists a countable family of infinite cardinal numbers β_n such that $L^p(\mu)$ is linearly isometric and isomorphic in the sense of its natural order to the space $\bigl[\oplus_n L^p([0,1]^{\beta_n},\lambda^{\beta_n})\bigr]_p$ defined as the space of all sequences (f_n) with $f_n\in L^p([0,1]^{\beta_n},\lambda^{\beta_n})$ which have finite norm*

$$(f_n)_p := \Bigl(\sum_n \|f_n\|_p^p\Bigr)^{1/p}.$$

9.12.27. Corollary. *Let μ be a separable atomless probability measure and let $1\le p<\infty$. Then $L^p(\mu)$ is linearly isometric to $L^p[0,1]$. If μ has atoms, but is not purely atomic, then $L^p(\mu)$ is linearly isometric to the direct sum of $L^p[0,a]$ and $L^p(\nu)$ for some $a<1$ and some finite measure ν on $\mathbb{N}$.*

9.12(vi). Almost homeomorphisms

Almost homeomorphisms of measure spaces considered in §9.6 may be very discontinuous when extended to the whole space. The question arises about the existence of almost homeomorphisms with better properties. Two such properties are described in the following definition.

9.12.28. Definition. *Let (X,μ) and (Y,ν) be topological spaces with Borel measures μ and ν.* (i) *We shall say that these spaces are K-isomorphic if there exist mappings $S\colon X\to Y$ and $S'\colon Y\to X$ such that S is continuous μ-a.e., S' is continuous ν-a.e., $S'\bigl(S(x)\bigr)=x$ for μ-a.e. x, $S\bigl(S'(y)\bigr)=y$ for ν-a.e. y, and $\nu=\mu\circ S^{-1}$, where μ is extended to $\mathcal{B}(X)_\mu$.*

(ii) *We shall say that these spaces are S-isomorphic if there exists a one-to-one Borel mapping T from X onto Y such that $\nu=\mu\circ T^{-1}$, T is continuous μ-a.e., and T^{-1} is continuous ν-a.e.*

The names for the above types of isomorphisms are explained by the fact that they were investigated in Krickeberg [**1059**], [**1060**], Böge, Krickeberg, Papangelou [**226**] and Sun [**1806**], [**1807**], respectively. The following theorem is established in Sun [**1806**].

9.12.29. Theorem. *Let μ be a Borel probability measure on a Polish space X. Then the following assertions are true.*

(i) *There exist a Borel set $Y\subset[0,1]$ and a Borel probability measure ν on Y such that (X,μ) and (Y,ν) are S-isomorphic.*

(ii) *One can take $[0,1]$ for Y precisely when every atom of the measure μ is an accumulation point in X.*

(iii) *If μ has no atoms, then (X,μ) and $([0,1],\lambda)$, where λ is Lebesgue measure, are S-isomorphic, and given a countable set D in the topological support of μ, an isomorphism T can be chosen in such a way that D belongs to the set of the continuity points of T and $T(D)$ belongs to the set of the continuity points of T^{-1}.*

This theorem does not extend to arbitrary Borel sets in Polish spaces. As shown in Sun [**1807**], the situation is this.

9.12.30. Theorem. *Let X be a Borel set in a Polish space and let μ be a Borel probability measure on X. Then:*

(i) *the existence of a Borel probability measure ν on $[0,1]$ such that (X,μ) and $([0,1],\nu)$ are S-isomorphic is equivalent to the existence of a set $Y\subset X$ of measure 1 that is a Polish space such that all atoms of μ are accumulation points of X;*

(ii) *if μ has no atoms, then the existence of an S-isomorphism between (X,μ) and $([0,1],\lambda)$ is equivalent to the existence of a set $Y\subset X$ of measure 1 that is a Polish space. In this case, given a countable set D in the intersection of the support of μ with Y, an isomorphism T can be chosen in such a way that D belongs to the set of the continuity points of T and $T(D)$ belongs to the set of the continuity points of T^{-1}.*

Certainly, one cannot always find a Polish subspace of full μ-measure. For example, if $X = \mathbb{Q} = \{r_n\}$ is the set of all rational numbers and μ equals $\sum_{n=1}^{\infty} 2^{-n}\delta_{r_n}$, then obviously such subspaces do not exist, since $\mathbb{Q}$ is not a Polish space. This example can be easily modified in order to obtain an atomless measure (for example, take the measure $\mu\otimes\lambda$ on $\mathbb{Q}\times[0,1]$).

It is clear that every S-isomorphism is a K-isomorphism. The converse is not true at least for the reason that a K-isomorphism may be neither one-to-one nor Borel. We remark that even if a K-isomorphism S is one-to-one, one cannot always take for S' the mapping S^{-1} (see Exercise 3.10.74). Sun [**1806**] constructs simple examples where K-isomorphic spaces (X,μ) and (Y,ν) are not S-isomorphic. In such examples, it can even occur that there is a one-to-one measure-preserving Borel mapping between X and Y. Thus, different isomorphisms may possess some of the properties required in the definition of S-isomorphisms, but they cannot be obtained simultaneously for a single mapping. It may also occur that there exists a K-isomorphism, but there is no measure-preserving one-to-one Borel mapping. Finally, the existence of a Borel isomorphism between X and Y transforming μ into ν does not yield that (X,μ) and (Y,ν) are K-isomorphic.

9.12(vii). Measures with given marginal projections

Given two probability measures μ and ν on spaces X and Y, there exist measures on $X\times Y$ whose projections to the factors are μ and ν (for example, the measure $\mu\otimes\nu$). In many applications, it is important to have such a measure with certain additional properties (say, concentrated on a given set). For example, on the square $[0,1]^2$, apart from the two-dimensional Lebesgue measure, there is a measure concentrated on the diagonal $x=y$ such that its projections to the sides are Lebesgue measures: the normalized linear measure on the diagonal. However, on the set $\{(x,y)\colon\ x<y\}$, there is no Borel measure whose projections are Lebesgue measures on the sides (Exercise 9.12.79). Let us mention several typical results in this direction. The next theorem on measures with given projections to the factors (called the *marginal projections*) was found by Strassen [**1791**] in the case of Polish spaces, and then generalized by several authors (see Skala [**1738**], whence the presented formulation is borrowed).

9.12.31. Theorem. *Let X and Y be completely regular spaces and let M be a convex set in $\mathcal{P}_r(X\times Y)$, closed in the weak topology (or let X and Y be general Hausdorff spaces and let M be closed in the A-topology). The existence of a measure $\lambda\in M$ with given projections $\mu\in\mathcal{P}_r(X)$ and $\nu\in\mathcal{P}_r(Y)$ on X and Y is equivalent to the following condition: for all bounded Borel functions f on X and g on Y one has*

$$\int_X f\,d\mu+\int_Y g\,d\nu\le\sup\biggl\{\int_{X\times Y}\bigl(f(x)+g(y)\bigr)\,\sigma(dx,dy)\colon\ \sigma\in M\biggr\}.$$

In particular, if Z is a closed set in $X\times Y$, then the existence of a measure λ in $\mathcal{P}_r(X\times Y)$ with the marginals $\mu\in\mathcal{P}_r(X)$ and $\nu\in\mathcal{P}_r(Y)$ and $\lambda(Z)=1$ is equivalent to the inequality

$$\int_X f\,d\mu + \int_Y g\,d\nu \le \sup\bigl\{f(x)+g(y)\colon\ (x,y)\in Z\bigr\}$$

for all bounded Borel functions f on X and g on Y.

Let $(X_1,\mathcal{A}_1,P_1)$ and $(X_2,\mathcal{A}_2,P_2)$ be probability spaces and let $\mathcal{P}(P_1,P_2)$ be the set of all probability measures on $(X_1\times X_2,\mathcal{A}_1\otimes\mathcal{A}_2)$ whose projections on X_1 and X_2 equal P_1 and P_2, respectively. Let h be a bounded measurable function on $(X_1\times X_2,\mathcal{A}_1\otimes\mathcal{A}_2)$ and let

$$S(h)=\sup_{\mu\in\mathcal{P}(P_1,P_2)}\int_{X_1\times X_2} h\,d\mu,$$

$$I(h)=\inf\biggl\{\int_{X_1} h_1\,dP_1+\int_{X_2} h_2\,dP_2\biggr\},$$

where inf is taken over all $h_i\in\mathcal{L}^1(P_i)$ with $h(x_1,x_2)\le h_1(x_1)+h_2(x_2)$. Integrating the latter inequality with respect to $P_1\otimes P_2$ we get $I(h)\ge S(h)$. The next general result is proved in Ramachandran, Rüschendorf [**1524**].

9.12.32. Theorem. *If at least one of the measures P_1 and P_2 is perfect, then $S(h)=I(h)$.*

It is shown in Ramachandran, Rüschendorf [**1525**] that the assumption of perfectness of one of the measures cannot be omitted. The results in this subsection are strongly related to those in §8.10(viii), where we dealt with the case $h(x,y)=-d(x,y)$ for a metric d.

9.12(viii). The Stone representation

A Boolean algebra is a nonempty set X with two binary operations $(A,B)\mapsto A\cap B$ and $(A,B)\mapsto A\cup B$ and an operation $A\mapsto -A$ that are related by the same identities as the usual set-theoretic operations of intersection, union, and complement (see Sikorski [**1725**, Ch. 1]). In this case, the elements $A\cap(-A)$ and $A\cup(-A)$ are independent of A and are called, respectively, the zero and unit of the algebra. A Boolean homomorphism of Boolean algebras is a mapping h with the properties

$$h(A\cup B)=h(A)\cup h(B),\quad h(A\cap B)=h(A)\cap h(B),\quad h(-A)=-h(A).$$

A one-to-one Boolean homomorphism is called a Boolean isomorphism. Earlier we encountered special cases of these concepts when dealing with the metric Boolean algebra of a measure space (in this case, the Boolean operations on equivalence classes of sets are the usual set-theoretic operations of intersection, union, and complement on representatives of those classes). One can define a Boolean algebra in terms of partially ordered sets (see Vladimirov [**1947**]), and also in algebraic terms as an associative ring with a unit such that all elements satisfy the condition $a\cdot a=a$ (to this end, the operation of addition of sets is defined as the symmetric difference, which corresponds to the addition mod 2 of indicator functions). The next important result due to Stone identifies abstract Boolean algebras with algebras of clopen sets. A proof of the Stone theorem can be found in Dunford, Schwartz [**503**], Lacey [**1098**], Sikorski [**1725**], Vladimirov [**1947**].

9.12.33. Theorem. *Every Boolean algebra $\mathcal{A}$ is isomorphic to the Boolean algebra of all simultaneously open and closed sets in some totally disconnected compact space S (i.e., a compact space that has a base consisting of clopen sets).*

Suppose that on an algebra $\mathcal{A}$ of subsets of a space X we have a nonnegative additive set function m with $m(X) = 1$. By the Stone theorem we represent $\mathcal{A}$ as the algebra $\mathcal{A}_0$ of all clopen subsets of a compact space S. The function m corresponds to a nonnegative additive set function m_0 on $\mathcal{A}_0$ with $m_0(S) = 1$. Since $\mathcal{A}_0$ consists of compact subsets of S, the measure m_0 is countably additive and hence admits a countably additive extension to $\sigma(\mathcal{A}_0)$. Moreover, by Theorem 7.3.11, there exists a Radon probability measure μ on S that extends m_0. We emphasize that the initial measure m need not be countably additive (this seeming contradiction is explained by the fact that the above-mentioned isomorphism may not preserve countable unions).

Loomis [**1182**] and Sikorski [**1724**] obtained a sharpening of the Stone theorem for Boolean σ-algebras $\mathcal{A}$ (Boolean algebras with countable unions): they proved that there exist a σ-algebra $\mathcal{A}_0$ and its σ-ideal Δ such that the algebra $\mathcal{A}$ is isomorphic to the factor-algebra $\mathcal{A}_0/\Delta$. Moreover, one can take for $\mathcal{A}_0$ the σ-algebra generated by all clopen sets in the Stone space S of the algebra $\mathcal{A}$ and for Δ the σ-ideal of all first category sets in $\mathcal{A}_0$.

9.12(ix). The Lyapunov theorem

Here we consider a nice application of measurable transformations to vector measures given by A.A. Lyapunov. We shall see that under broad assumptions, several measures can be transformed into a given one by a common transformation. Lyapunov [**1216**] (see also Lyapunov [**1217**, p. 234]) proved the following interesting result. Let ψ be an absolutely continuous function on $[0,1]$ with $\psi(0) = 0$. Then there exists a Borel function $f\colon [0,1] \to [0,1]$ such that for all $t \in [0,1]$ one has $\lambda\bigl(s\colon f(s) \le t\bigr) = t$, where λ is Lebesgue measure, and
$$\int_{\{f\le t\}} \psi'(s)\,ds = t\psi(1).$$
We prove this result in an equivalent formulation.

9.12.34. Theorem. *Given an absolutely continuous measure ν on $[0,1]$ (possibly signed), there exists a Borel transformation f of $[0,1]$ that preserves Lebesgue measure λ and takes the measure ν to $\nu([0,1])\lambda$.*

PROOF. We define a function f by means of the sets
$$E_{nk} = \{x\colon\ k2^{-n} < f(x) < (k+1)2^{-n}\},\quad k = 0,1,\ldots,2^n-1,\ n = 0,1,\ldots,$$

which will be constructed by induction. Let $\nu([0,1]) = \alpha$ and let $E_{00} = [0,1]$. Suppose that for some $n \geq 1$ sets E_{nk} are constructed in such a way that $\lambda(E_{nk}) = 2^{-n}$, $\nu(E_{nk}) = \alpha 2^{-n}$. Let us show how to construct sets $E_{n+1,2k}$ and $E_{n+1,2k+1}$ for $k = 0,\ldots,2^n-1$. We find $x_1 \in [0,1]$ with

$$\lambda\bigl(E_{nk} \cap [0,x_1]\bigr) = 2^{-n-1}.$$

For every $x \in [0,x_1]$, there is the smallest number $\xi(x) > x$ with

$$\lambda\bigl(E_{nk} \cap [x,\xi(x)]\bigr) = 2^{-n-1}.$$

It is clear that the function ξ on $[0,x_1]$ is continuous. Hence the function $\eta\colon\, x \mapsto \nu\bigl(E_{nk} \cap [x,\xi(x)]\bigr)$ on $[0,x_1]$ is continuous as well. We observe that there is $z \in [0,x_1]$ with $\eta(z) = \alpha 2^{-n-1}$. Indeed, $\xi(0) \leq x_1$, hence the set

$$D := \bigl([0,\xi(0)] \cup [x_1,\xi(x_1)]\bigr) \cap E_{nk}$$

has Lebesgue measure 2^{-n}, i.e., coincides with $E_{n,k}$ up to a set of measure zero. Then $\eta(0) + \eta(x_1) = \nu(D) = \alpha 2^{-n}$. Therefore, the numbers $\eta(0)$ and $\eta(x_1)$ cannot be simultaneously greater than $\alpha 2^{-n-1}$ or smaller than $\alpha 2^{-n-1}$, which by the continuity of η yields the required number z. Now let

$$E_{n+1,2k} := E_{n,k} \cap [z,\xi(z)], \quad E_{n+1,2k+1} := E_{nk} \backslash E_{n+1,2k}.$$

It is clear from our inductive construction that every E_{nk} is the union of finitely many intervals (closed, open or semi-open). The function f is defined as follows: given $x \in [0,1]$, for every n there is a unique number k_n such that $x \in E_{nk_n}$; then we set $f(x) := \bigcap_{n=1}^{\infty} [k_n 2^{-n}, (k_n+1)2^{-n}]$. The set $\{k2^{-n} < f < (k+1)2^{-n}\}$ coincides with E_{nk} up to finitely many endpoints of the intervals constituting E_{nk}. Hence the function f is Borel and one has $\lambda \circ f^{-1}(E_{nk}) = \lambda(E_{nk})$ and $\nu \circ f^{-1}(E_{nk}) = \alpha\lambda(E_{nk})$, which gives the required equalities on all Borel sets. □

We apply this theorem to simultaneous transformations of measures.

9.12.35. Corollary. *Let $(X,\mathcal{A},\mu)$ be a probability space and let μ be atomless. Suppose we are given finitely many measures $\nu_1,\ldots,\nu_k$ on $\mathcal{A}$ that are absolutely continuous with respect to μ. Then there exists an $\mathcal{A}$-measurable function $f\colon\, X \to [0,1]$ such that one has $\mu \circ f^{-1} = \lambda$ and $\nu_i \circ f^{-1} = \nu_i(X)\lambda$ for all $i = 1,\ldots,k$, where λ is Lebesgue measure on $[0,1]$.*

Proof. There exists an $\mathcal{A}$-measurable function $f_1\colon\, X \to [0,1]$ such that one has $\mu \circ f_1^{-1} = \lambda$. Then $\nu_i \circ f_1^{-1} \ll \lambda$, which reduces our assertion to the case $X = [0,1]$ and $\mu = \lambda$. We prove it by induction on k. For $k = 1$ the assertion is already proven. Suppose that it is true for some $k \geq 1$ and that we are given measures $\nu_i \ll \lambda$, $i \leq k+1$. There exists a Borel function f_k such that $\lambda \circ f_k^{-1} = \lambda$ and $\nu_i \circ f_k^{-1} = \nu_i([0,1])\lambda$ for all $i \leq k$. Then $\nu_{k+1} \circ f_k^{-1} \ll \lambda$. Let us take a Borel function g such that $\lambda \circ g^{-1} = \lambda$ and $\nu_{k+1} \circ f_k^{-1} \circ g^{-1} = \nu_{k+1}([0,1])\lambda$. The function $g \circ f_k$ has the required property. □

9.12.36. Corollary. *Suppose that μ is an atomless Borel probability measure on a Souslin space X and let $\nu_1, \ldots, \nu_k$ be Borel measures absolutely continuous with respect to μ. Then there exists a Borel mapping $T\colon X \to X$ such that $\mu \circ T^{-1} = \mu$ and $\nu_i \circ T^{-1} = \nu_i(X)\mu$ for all $i = 1, \ldots, k$.*

9.12.37. Corollary. *Let $\mu_1, \ldots, \mu_n$ be atomless Borel probability measures on a Souslin space X. Then, for every Borel probability measure ν on X, there exists a Borel transformation $T\colon X \to X$ such that $\mu_i \circ T^{-1} = \nu$ for all $i \le n$.*

PROOF. By Corollary 9.12.36 we find a mapping that transforms the measures μ_i into the measure $\mu = (\mu_1 + \cdots + \mu_n)/n$; then we transform μ into ν by Theorem 9.2.2. $\square$

By using these results one can easily prove the following remarkable theorem due to A.A. Lyapunov [**1216**].

9.12.38. Theorem. *Let ν be a countably additive vector measure with values in $\mathbb{R}^n$ defined on a measurable space $(X, \mathcal{A})$, i.e., $\nu = (\nu_1, \ldots, \nu_n)$, where each ν_i is a real measure on $\mathcal{A}$. Suppose that the measures ν_i have no atoms. Then the set of values of ν is convex and compact.*

PROOF. Let $v_1 = \nu(A_1)$, $v_2 = \nu(A_2)$, where $A_i \in \mathcal{A}$, and let $t \in (0,1)$. We consider the sets

$$A = A_1 \cap A_2,\ X_1 = A_1 \backslash A_2,\ X_2 = A_2 \backslash A_1.$$

Let us show that $tv_1 + (1-t)v_2 = \nu(B)$ for some $B \in \mathcal{A}$. It is clear that

$$tv_1 + (1-t)v_2 = tu_1 + (1-t)u_2 + w,$$

where $u_1 = \nu(X_1)$, $u_2 = \nu(X_2)$, $w = \nu(A)$. The set B will be found in the form $B_1 \cup A \cup B_2$, where $B_i \subset X_i$. Let us consider the measure $\mu = |\nu_1| + \cdots + |\nu_n|$. If $\mu(X_1) = 0$, then $v_1 = 0$ and we set $B_1 = \varnothing$. Suppose $\mu(X_1) > 0$. Applying the above corollary to the measure μ on X_1, we obtain a function $f_1\colon X_1 \to [0,1]$ such that

$$\mu|_{X_1} \circ f_1^{-1} = \mu(X_1)\lambda \quad \text{and} \quad \nu_i|_{X_1} \circ f_1^{-1} = \nu_i(X_1)\lambda, \quad i = 1, \ldots, n,$$

where λ is Lebesgue measure. Letting $B_1 := f_1^{-1}([0,t])$ we have the equality $\nu_i(B_1) = t\nu_i(X_1)$, $i = 1, \ldots, n$, whence we obtain $\nu(B_1) = tu_1$. Similarly, there exists a set $B_2 \in \mathcal{A}$ with $B_2 \subset X_2$ and $\nu(B_2) = (1-t)u_2$. Now let us set $B := B_1 \cup A \cup B_2$. Since the sets B_1, A, B_2 are disjoint, one has $\nu(B) = tv_1 + (1-t)v_2$, i.e., the set K of all values of ν is convex.

Let us show that K is closed by induction on n. For $n = 1$ this is true by Corollary 1.12.10. Suppose our claim is true for $n-1$. Let v be a limit point of K. Suppose that v is not an inner point of K. Then there is an $(n-1)$-dimensional hyperplane L passing through v such that K belongs to one of the two closed half-spaces with the boundary L. Without loss of generality we may assume that $L = \{x_1 = 1\}$. For every $i = 2, \ldots, n$, there is a set $E_i \in \mathcal{A}$ such that $\nu_i|_{E_i} \ll |\nu_1|_{E_i}$ and $|\nu_1|(X \backslash E_i) = 0$. Let $X_1 := \bigcap_{i=2}^n E_i$ and $X_2 = X \backslash X_1$. Then one has $|\nu_1|(X_2) = 0$ and $\nu_i|_{X_1} \ll |\nu_1|_{X_1}$ for all $i \le n$. The restriction of the measure ν to X_2 takes values in the hyperplane $L_0 = \{x_1 = 0\}$, hence by the inductive assumption the set of values of ν on X_2 is a convex compact set K_2. Let $K_1 := \{\nu(A)\colon A \in \mathcal{A}, A \subset X_1\}$. Let us consider the Hahn decomposition $X_1 = Y^+ \cup Y^-$ for the measure ν_1. Since the set of values of ν_1 is closed, one has $\nu_1(Y^+) = 1$. We observe that if $A_j \in \mathcal{A}$ are such that $A_j \subset X_1$ and $\nu_1(A_j) \to 1$, then $\nu(A_j) \to \nu(Y^+)$. Indeed, the value 1 is maximal for ν_1, whence we obtain that $|\nu_1|(A_j \cap Y^-) \to 0$ and $\nu_1(A_j \cap Y^+) \to 1$, i.e., one has $|\nu_1|(A_j \triangle Y^+) \to 0$. By the absolute continuity of $\nu_i|_{X_1}$ with respect to $|\nu_1|$ we obtain $|\nu_i|(A_j \triangle Y^+) \to 0$ for every $i \leq n$, which gives $\nu(A_j) \to \nu(Y^+)$. By the definition of v, there exist sets $B_j \in \mathcal{A}$ with $\nu(B_j) \to v$. Then $\nu_1(B_j \cap X_1) \to 1$, whence we have $\nu(B_j \cap X_1) \to \nu(Y^+)$ as shown above. On the other hand, since K_2 is closed, there is a set $B \in \mathcal{A}$ such that $B \subset X_2$ and $\nu(B) = \lim\limits_{j\to\infty} \nu(A_j \cap X_2)$. Then $v = \nu(Y^+ \cup B)$. □

A completely different proof of Lyapunov's theorem can be found in Diestel, Uhl [**444**, Ch. IX]. However, that proof does not give the other results in this section.

Exercises

9.12.39.° Let K_n, where $n \in \mathbb{N}$, be increasing compact sets in a Hausdorff space X and let $f\colon X \to Y$ be an injective mapping to a Hausdorff space Y such that f is continuous on every K_n. Prove that if a Radon measure ν is concentrated on the union of the compact sets $f(K_n)$, then it has a unique Radon preimage with respect to f.

Hint: observe that if μ_1 and μ_2 are Radon preimages of ν, then they are concentrated on $\bigcup_{n=1}^\infty K_n$, and their restrictions to every compact set K_n coincide; in order to verify the latter, use that if two Radon measures μ_1 and μ_2 on $\bigcup_{n=1}^\infty K_n$ are not equal, then $\mu_1(S) \neq \mu_2(S)$ for some compact set S in one of the sets K_n, hence the compact set $f(S)$ has different measures with respect to their images.

9.12.40.° Let $(X, \mathcal{A}, \mu)$ be a probability space, let $(Y, \mathcal{E})$ be a measurable space, and let $\pi\colon X \to Y$ be an $(\mathcal{A}_\mu, \mathcal{E})$-measurable mapping. Suppose that the measure $\nu = \mu \circ \pi^{-1}$ on $\mathcal{E}$ (or on $\mathcal{E}_\nu$) has a compact approximating class and $\pi(X) \in \mathcal{E}_\nu$. Show that the measure μ on $\mathcal{B} = \pi^{-1}(\mathcal{E})$ also has a compact approximating class.

Hint: suppose first that $\pi(X) = Y$; let $\mathcal{K}$ be a compact approximating class for ν on $\mathcal{E}$ and let $\mathcal{K}_0 = \pi^{-1}(\mathcal{K})$. Then $\mathcal{K}_0$ is a compact class. Indeed, if $C_n = \pi^{-1}(K_n)$, $K_n \in \mathcal{K}$ and $\bigcap_{i=1}^n C_i \neq \varnothing$ for all n, then the sets $\bigcap_{i=1}^n K_i$ are nonempty. There exists $y \in \bigcap_{i=1}^\infty K_i$. There is x with $\pi(x) = y$. Then $x \in \bigcap_{i=1}^\infty C_i$. Clearly, $\mathcal{K}_0$ is an approximating class for μ on $\mathcal{B}$. In the general case, let $\mathcal{K}_1 = \{K \in \mathcal{K}\colon K \subset \pi(X)\}$. It is clear that $\mathcal{K}_1$ is a compact class of subsets of $\pi(X)$. In order to reduce our assertion to the case $\pi(X) = Y$, it suffices to verify that the class $\mathcal{K}_1$ approximates the measure ν on $\pi(X)$. Let $E \in \mathcal{E}$ and $\varepsilon > 0$. By hypothesis, there exists a set $Y_0 \subset \pi(X)$ such that $Y_0 \in \mathcal{E}$ and $\nu\bigl(\pi(X)\backslash Y_0\bigr) = 0$. In addition, there exist sets $E_0 \in \mathcal{E}$ and $K \in \mathcal{K}$ such that $E_0 \subset K \subset E \cap Y_0$ and $\nu\bigl((Y_0 \cap E)\backslash E_0\bigr) < \varepsilon$. It is clear that $K \subset \pi(X)$, i.e., $K \in \mathcal{K}_1$. Finally, $K \subset E$ and $\nu(E \backslash E_0) < \varepsilon$.

9.12.41.° Let μ be a Borel measure on the space $\mathcal{R}$ of irrational numbers in $(0, 1)$, positive on nonempty open sets and having no points of positive measure. Prove that for every sequence of numbers $\alpha_n > 0$ with $\sum_{n=1}^\infty \alpha_n = \mu(\mathcal{R})$, there exist disjoint open sets U_n such that $\mathcal{R} = \bigcup_{n=1}^\infty U_n$ and $\mu(U_n) = \alpha_n$ for all n.

HINT: let $a(i,j) = \alpha_i j/(j+1)$, $i,j \in \mathbb{N}$. Observe that one can find a sequence of rational numbers r_k, $r_0 = 0$, increasing to 1 and having the following property: if $\mathbb{N}\times\mathbb{N}$ is ordered according to the rule $(i,j) < (i',j')$ whenever $i+j < i'+j'$ or $i+j = i'+j'$ and $j < j'$, and if $(r_{k-1}, r_k)\cap\mathcal{R}$ is denoted by $I(i,j)$, where (i,j) is the element with the number k in the indicated ordering, then we have the estimates $a(i,j) < \sum_{n=1}^j \mu\bigl(I(i,n)\bigr) < a(i,j+1)$. The numbers r_k are constructed inductively by using that the function $\mu([0,x]\cap\mathcal{R})$ is strictly increasing and continuous on $[0,1]$. The sets $U_i = \bigcup_{j=1}^\infty I(i,j)$ give the required partition. An alternative reasoning is this. We find rational numbers $r_{1,n}$, $r_{1,1} = 0$, increasing to a rational number r and having the property that $2\alpha_n/3 \le \mu\bigl((r_{1,n}, r_{1,n+1})\bigr) \le \alpha_n$. Let us repeat this procedure for the interval $(r,1)$ and numbers $\alpha_n - \mu\bigl((r_{1,n}, r_{1,n+1})\bigr)$. We proceed inductively and take the sets $U_n = \bigcup_{k=1}^\infty (r_{k,n}, r_{k,n+1})$.

9.12.42.$^\circ$ Let X be a complete separable metric space and let $\mu \ge 0$ be a finite Borel measure on X without points of positive measure. Show that for every $A \in \mathcal{B}(X)$ with $\mu(A) > 0$ and every $\varepsilon > 0$, there exists a set $K \subset A$ homeomorphic to the Cantor set such that $\mu(A\backslash K) < \varepsilon$.

HINT: we may assume that μ is a probability measure; there exists a Borel set $B \subset X$ such that $\mu(X\backslash B) = 0$ and (B, μ_B) is homeomorphic to $(\mathcal{R}, \lambda)$, where $\mathcal{R}$ is the space of irrational numbers in $(0,1)$ with Lebesgue measure λ. Let h be the corresponding homeomorphism. The set $h(A\cap B)$ contains a perfect compact set C with $\lambda(C) > \lambda\bigl(h(A\cap B)\bigr) - \varepsilon$. Then $h^{-1}(C)$ is a required set (see also Gelbaum [**674**], Oxtoby [**1408**]). We could also use a Borel isomorphism and Lusin's theorem.

9.12.43.$^\circ$ Let $U \subset \mathbb{R}^n$ be an open set, Y a Souslin space, μ a Borel probability measure on Y, and let $f\colon U \to Y$ be a Borel mapping. Prove that there exists a sequence of pairwise disjoint open cubes $K_j \subset U$ with edges parallel to the coordinate axes such that $\mu\bigl(f(U)\bigr) = \mu\bigl(f\bigl(\bigcup_{j=1}^\infty K_j\bigr)\bigr)$.

HINT: take a Borel measure ν on U such that $\mu|_{f(U)} = \nu\circ f^{-1}$ and apply Exercise 1.12.72.

9.12.44. Let X and Y be metric or Souslin spaces with nonnegative Radon measures μ and ν and let $f\colon X \to Y$ be a $\bigl(\mathcal{B}(X)_\mu, \mathcal{B}(Y)\bigr)$-measurable mapping having property (N) with respect to the pair (μ,ν). Prove that for ν-a.e. $y \in Y$, the set $f^{-1}(y)$ is at most countable.

HINT: let $\mathcal{Y}$ denote the class of all Borel sets $Y' \subset Y$ such that $f^{-1}(y)$ is at most countable for every $y \in Y'$. Let α be the supremum of the ν-measures of sets in $\mathcal{Y}$. There are sets $Y_n \in \mathcal{Y}$ with $\nu(Y_n) \to \alpha$. Let $Y_0 = \bigcup_{n=1}^\infty Y_n$. Then $Y_0 \in \mathcal{Y}$ and $\nu(Y_0) = \alpha$. Suppose $\nu\bigl(f(X)\backslash Y_0\bigr) > 0$. Let $X_0 := f^{-1}(Y\backslash Y_0)$. Then $\mu(X_0) > 0$ because otherwise $\nu\bigl(f(X_0)\bigr) = 0$ by property (N). According to Proposition 9.1.7 there is a μ-measurable set $A_1 \subset f^{-1}(Y\backslash Y_0)$ such that $f(A_1) = f(X_0)$ and f is injective on A_1. Then $\mu(A_1) > 0$ by property (N). We may take A_1 in such a way that its measure is greater than one half of the supremum of μ-measures of sets with such a property. Repeating this reasoning, we obtain a finite or countable collection of disjoint μ-measurable sets A_n on each of which f is injective and the equality $\mu\bigl(X_0\backslash\bigcup_{n=1}^\infty A_n\bigr) = 0$ holds. This leads to a contradiction, since $f\bigl(X_0\backslash\bigcup_{n=1}^\infty A_n\bigr)$ has ν-measure zero, and every point in $f(X_0)\backslash f\bigl(X_0\backslash\bigcup_{n=1}^\infty A_n\bigr)$ has at most countably many preimages.

9.12.45. (Federer, Morse [**556**]) Let μ be a Radon probability measure on a metric (or Souslin) space X and let f be a μ-measurable function. Let $Y(\aleph_0)$ denote the set of all points y having infinite preimages and let $Y(\aleph_1)$ denote the set of all points y with uncountable preimages.

(i) Prove that there exists a μ-measurable set $C \subset X$ such that $f(C) = Y(\aleph_0)$ and the set $f^{-1}(y)\backslash C$ is finite for each $y \in f(X)$.

(ii) Prove that for every $\varepsilon > 0$, there exists a μ-measurable set $L \subset X$ with $\mu(L) < \varepsilon$ such that $f(L) = Y(\aleph_0)$ and f is injective on L.

(iii) Prove that there exists a set $Z \subset X$ with $\mu(Z) = 0$ such that $f(Z) = Y(\aleph_1)$ and f is injective on Z.

9.12.46.° Let X, Y be Souslin spaces with Borel probability measures μ and ν, respectively, and let $f\colon X \to Y$ be a Borel mapping. Show that f has property (N) with respect to (μ, ν) precisely when $\nu\bigl(f(K)\bigr) = 0$ for every compact set K with $\mu(K) = 0$.

HINT: let $B \in \mathcal{B}(X)$, $\mu(B) = 0$, but $\nu\bigl(f(B)\bigr) > 0$; by Theorem 7.14.34, there exists a compact set $K \subset B$ with $\nu\bigl(f(K)\bigr) > 0$, which is a contradiction.

9.12.47. (i) It is known that the constructability axiom in set theory yields the existence of a coanalytic set $X \subset [0,1]$ and a continuous function $\varphi\colon X \to [0,1]$ such that the set $\varphi(X)$ has inner measure zero and positive outer measure (see Novikov [**1384**] and Jech [**891**]). Let $\Omega = X \cup [2,3]$ be equipped with the usual topology and consider on Ω the measure μ that vanishes on X and coincides with Lebesgue measure on $[2,3]$. Let $f(x) = \varphi(x)$ if $x \in X$ and $f(x) = x$ if $x \in [2,3]$. Show that $f(K)$ has Lebesgue measure zero for every compact set $K \subset \Omega$ with $\mu(K) = 0$. In addition, $f(X)$ is nonmeasurable, although X is a closed subset of Ω. In particular, f has no property (N).

(ii) Assuming the constructability axiom prove that there exists a coanalytic set X in $[0,1]$ such that on some countably generated σ-algebra $S \subset \mathcal{B}(X)$, there is a probability measure having no countably additive extensions on $\mathcal{B}(X)$.

HINT: (i) the sets $K \cap X$ and $K \cap [2,3]$ are compact in Ω and one has

$$\lambda\bigl(f(K \cap [2,3])\bigr) = \lambda\bigl(K \cap [2,3]\bigr) = \mu\bigl(K \cap [2,3]\bigr) = 0$$

and $\lambda\bigl(f(K \cap X)\bigr) = \lambda\bigl(\varphi(K \cap X)\bigr) = 0$ by the compactness of $\varphi(K) \cap X$ and the equality $\lambda_*\bigl(\varphi(X)\bigr) = 0$.

(ii) Take a coanalytic set $X \subset [0,1]$ and a continuous function $f\colon X \to [0,1]$ such that $f(X)$ has inner measure zero and positive outer measure. Let us consider the class $S = \bigl\{f^{-1}(B), B \in \mathcal{B}\bigl(f(X)\bigr)\bigr\}$. Then S is a countably generated σ-algebra in $\mathcal{B}(X)$. The measure μ on S defined by the formula $\mu\bigl(f^{-1}(B)\bigr) = \lambda^*(B)$ is countably additive, but has no countably additive extensions to $\mathcal{B}(X)$. Indeed, we have $\mu(X) = \lambda^*\bigl(f(X)\bigr) > 0$ and at the same time $\mu(K) = 0$ for every compact set K in X because $f(K)$ is compact in $f(X)$ and hence $\lambda\bigl(f(K)\bigr) = 0$.

9.12.48.° Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be probability spaces and $f \in \mathcal{L}^1(\mu\otimes\nu)$. Show that the image of the measure $f \cdot (\mu\otimes\nu)$ under the natural projection $X\times Y$ to X is given by the density

$$\varrho(x) = \int_Y f(x,y)\,\nu(dy)$$

with respect to the measure μ.

HINT: express the integral of $I_{A\times Y}$ against the measure $f \cdot (\mu\otimes\nu)$ via ϱ.

9.12.49.° Let E be a Souslin subset of $[0,1]$ that is not Borel. Denote by $\mathcal{S}$ the class of all sets of the form $S = B \cup C$, where B and C are Borel sets with $B \subset E$ and $C \subset [0,1]\backslash E$. Let $\mathcal{E}$ be the class of all sets of the form $S_1 \cup ([0,1]\backslash S_2)$, where $S_1, S_2 \in \mathcal{S}$. Show that $\mathcal{E}$ is a σ-algebra and that the formula $\mu(S) = 0$, $\mu([0,1]\backslash S) = 1$ if $S \in \mathcal{S}$ defines a probability measure on $\mathcal{E}$ that has no countably additive extensions to $\mathcal{B}([0,1])$.

HINT: if μ' is a Borel extension of μ, then E is measurable with respect to the Lebesgue completion of μ' and $\mu'(E) = 0$, since for every compact set $K \subset E$ we have $\mu(K) = 0$ because $K \in \mathcal{S}$. Similarly, $\mu'([0,1]\backslash E) = 0$, whence $\mu([0,1]) = 0$, a contradiction.

9.12.50. (Steinhaus [**1784**]) For every point $\xi \in (0,1)$, let us consider its binary expansion $0, \xi_1, \xi_2, \ldots$. Let the mapping $\theta\colon (0,1) \to (0,1)^\infty$ be defined by the formula $\theta\colon \xi \mapsto (\theta_n)$,

$$\theta_1 = 0, \xi_1, \xi_3, \xi_6, \xi_{10}, \ldots, \quad \theta_2 = 0, \xi_2, \xi_5, \xi_9, \xi_{14}, \ldots, \quad \theta_3 = 0, \xi_4, \xi_8, \xi_{13}, \xi_{19}$$

and so on. In other words, to the point θ with $\theta_n = 0, \theta_{n1}, \theta_{n2}, \ldots$ we map the point $\xi = 0, \theta_{11}, \theta_{21}, \theta_{12}, \theta_{31}, \ldots$ Show that the image of Lebesgue measure λ is λ^∞.

9.12.51. Let μ be the measure on $X = \{0,1\}^\infty$ that is the countable power of the measure on $\{0,1\}$ assigning $1/2$ to $\{0\}$ and $\{1\}$. Prove that every measurable set of positive μ-measure contains a pair of points that differ only in one coordinate.

HINT: use an isomorphism with Lebesgue measure defined by the mapping $(x_n) \mapsto \sum_{n=1}^\infty x_n 2^{-n}$ and the fact that every set of positive measure in $[0,1]$ contains points with difference as small as we like.

9.12.52.° Let μ be an atomless perfect probability measure on a measurable space $(X, \mathcal{A})$. Prove that X contains a measure zero set of cardinality of the continuum.

HINT: there is a measurable function $f\colon X \to [0,1]$ such that $\mu\circ f^{-1}$ is Lebesgue measure. The set $f(X)$ contains a Borel set of measure 1 and this set contains a Borel set E of measure zero and cardinality of the continuum. Then the cardinality of $f^{-1}(E)$ is not less than that of the continuum.

9.12.53.° Let μ be an atomless Radon probability measure on a compact space K. Prove that there exists a set $E \subset K$ that does not belong to the Lebesgue completion of $\mathcal{B}(K)$ with respect to μ.

HINT: take a continuous function $f\colon K \to [0,1]$ transforming μ into Lebesgue measure and a set $A \subset [0,1]$ with $\lambda_*(A) = \lambda_*([0,1]\backslash A) = 0$. Then at least one of the sets $B = f^{-1}(A)$ and $C = f^{-1}([0,1]\backslash A)$ is not measurable with respect to μ, since both have zero inner measure: for example, if $S \subset B$ is compact and $\mu(S) > 0$, then $f(S)$ is a compact set in A and $\lambda\bigl(f(S)\bigr) > 0$.

9.12.54. (i) (Herz [**821**]) Let X and Y be locally compact spaces and let $f\colon X \to Y$ be a continuous mapping. Prove that for every Radon measure ν on Y, one can find Radon measures μ and ν' on X and Y, respectively, such that

$$\nu = \mu \circ f^{-1} + \nu', \quad \|\nu\| = \|\mu\| + \|\nu'\|,$$

and $\nu'\bigl(f(K)\bigr) = 0$ for every compact set $K \subset X$.

(ii) Let X and Y be Souslin spaces and let $f\colon X \to Y$ be a Borel mapping. Show that for every Borel measure ν on Y, one can find a Borel measure μ on X and a Borel measure ν' on Y such that $\nu = \mu \circ f^{-1} + \nu'$ and $|\nu'|\bigl(Y\backslash f(X)\bigr) = 0$.

HINT: (ii) take $\nu' = \nu|_{Y\setminus f(X)}$ and find μ such that $\mu\circ f^{-1} = \nu|_{f(X)}$.

9.12.55. Let $\mathcal{G}$ be a compact group with the Haar probability measure λ and let (X,μ) be a measure space such that $\mathcal{G}$ acts on X, i.e., we are given a $\lambda\otimes\mu$-measurable mapping $(G,x)\mapsto G(x)$ determining a homomorphism of $\mathcal{G}$ to the group of transformations of X. Let f be a μ-integrable function on X such that for every $G\in\mathcal{G}$, the functions f and $f\circ G$ are equal almost everywhere. Prove that there exists a function f_0 that is equal to f almost everywhere and is invariant with respect to all transformations $G\in\mathcal{G}$.

HINT: consider the function

$$f_0(x) = \int_{\mathcal{G}} f\bigl(G(x)\bigr)\,\lambda(dG).$$

9.12.56.° Prove that the standard surface measure on the sphere in $\mathbb{R}^n$ is a unique, up to a constant factor, spherically invariant finite measure on the sphere.

HINT: the unitary group acts transitively on the sphere.

9.12.57. (Beck, Corson, Simon [**140**]) Let G be a locally compact group with a Haar measure λ, $A,B\subset G$, $\lambda^*(A)>0$, $\lambda^*(B)>0$, where A is measurable. Prove that $A-B := \{ab^{-1}\colon\ a\in A, b\in B\}$ contains a neighborhood of the neutral element.

9.12.58. (Reiter [**1547**]) A locally compact group G is called amenable if on the space B of all bounded Borel functions on G, there exists a linear functional Λ (called an invariant mean) satisfying the conditions $\Lambda(1)=1$, $\Lambda(f)\ge 0$ if $f\ge 0$ and $\Lambda\bigl(f(g*\cdot)\bigr)=\Lambda(f)$ for all $g\in G$ and $f\in B$, where $f(g*\cdot)$ denotes the function $f(gx)$. In the case of a compact group, the integral with respect to the probability Haar measure can be taken for Λ. The noncompact group $\mathbb{R}^1$ is amenable. Prove that a locally compact group G is amenable precisely when for every function $f\in\mathcal{L}^1(\lambda)$, where λ is a left invariant Haar measure, one has

$$\biggl|\int_G f(x)\,\lambda(dx)\biggr| = \inf\int_G\Bigl|\sum_{i=1}^n \alpha_i f(x_i * x)\Bigr|\,\lambda(dx),$$

where inf is taken over all $n\in\mathbb{N}$, $x_i\in G$ and $\alpha_i\ge 0$ with $\alpha_1+\cdots+\alpha_n=1$.

HINT: see Greenleaf [**733**, §3.7], Reiter [**1547**].

9.12.59. Suppose that the mappings U_t^F satisfy (9.10.5) and that $F(x)=G(x)$ μ-a.e. Show that $(U_t^F)_{t\in\mathbb{R}^1}$ satisfies equation (9.10.5) with G in place of F.

9.12.60. Construct a Radon probability measure μ on a compact space X such that the space (X,μ) is isomorphic mod0 to the interval $[0,1]$ with Lebesgue measure λ, but is not almost homeomorphic to $([0,1],\lambda)$.

HINT: take a nonmetrizable countable subspace $S=\{s_n\}$ in some compact space K with the property that $\{s_n\}$ contains no sequences convergent in K. For example, let $K=\beta\mathbb{N}$ (the Stone–Čech compactification of $\mathbb{N}$), $S=\mathbb{N}\cup\{n_0\}$, where n_0 is a point in $\beta\mathbb{N}\backslash\mathbb{N}$. Let $X=K\times[0,1]$, $\nu=\sum_{n=1}^\infty 2^{-n}\delta_{s_n}$ and $\mu=\nu\otimes\lambda$. Then one can construct a Borel isomorphism between $S\times[0,1]$ and $[0,1]$ transforming μ into λ. However, there is no almost homeomorphism between (X,μ) and $([0,1],\lambda)$. Indeed, if we had homeomorphic sets $A\subset X$ and $B\subset[0,1]$ with unit measures, then for every $n\ge 0$, we could find a set $E_n\subset[0,1]$ of Lebesgue measure 1 with $(s_n,x)\in A$ for all $x\in E_n$. Let us fix a point $x_0\in E_0$. Then one can choose points $x_n\in E_n$ such that $x_n\to x_0$ as $n\to\infty$. The set $M=\{(s_n,x_n)\}$ is metrizable. One can verify that M is homeomorphic to S, which leads to a contradiction.

9.12.61. (Babiker, Knowles [**87**]) Construct an atomless Radon probability measure μ on a compact space X, a continuous mapping $\varphi\colon X \to [0,1]$ and an open set $G \subset X$ with the following properties: (i) $\mu \circ \varphi^{-1}$ is Lebesgue measure λ on $[0,1]$, (ii) $\varphi(G)$ is not Lebesgue measurable, (iii) the measure algebras generated by μ and λ are isomorphic, but the measures μ and λ are not almost homeomorphic, (iv) there exists a λ-measurable mapping $\psi\colon [0,1] \to X$ with $\varphi\bigl(\psi(t)\bigr) = t$ for all $t \in [0,1]$, but there is no almost continuous (in the sense of Lusin) mapping ψ with such a property. To this end, use Example 9.12.12.

9.12.62. Let X be a complete separable metric space and let $f\colon X \to \mathbb{R}^1$. Prove that the following conditions are equivalent:

(i) for every continuous mapping $g\colon \mathbb{R}^1 \to X$, the composition $f \circ g\colon \mathbb{R}^1 \to \mathbb{R}^1$ is Lebesgue measurable,

(ii) the function f is measurable with respect to every Borel measure on X.

Hint: let (i) be fulfilled and let μ be a Borel measure on X; it suffices to consider the case where μ is a probability atomless measure. By Theorem 9.6.3, there exist Borel sets $Y \subset X$ and $B \subset [0,1]$ with $\mu(Y) = 1$, $\lambda(B) = 1$, where λ is Lebesgue measure, and a homeomorphism $h\colon B \to Y$ with $\mu = \lambda \circ h^{-1}$. Given $\varepsilon > 0$, one can find a compact set $K \subset B$ with $\lambda(K) > 1 - \varepsilon$. Let us extend $h|_K$ to a continuous mapping $g\colon [0,1] \to X$ and choose in K a compact set Q with $\lambda(Q) > 1 - \varepsilon$ on which the function $\psi = f \circ g$ is continuous. We observe that f is continuous on the compact set $g(Q)$, since $g(Q) = h(Q) \subset Y$ and $f(x) = \psi\bigl(h^{-1}(x)\bigr)$ for all $x \in Y$. In addition, $\mu\bigl(h(Q)\bigr) = \lambda(Q) > 1 - \varepsilon$. If (ii) is fulfilled and the mapping $g\colon [0,1] \to X$ is continuous, then f is measurable with respect to the measure $\mu = \lambda \circ g^{-1}$, which yields the Lebesgue measurability of $f \circ g$.

9.12.63. Let $X = [0,1]^\infty$. Given $n \in \mathbb{N}$ and $t \in [0,1]$, we denote by $X_{n,t}$ the closed set in $[0,1]^\infty$ consisting of all points whose nth coordinate equals t. Then $X_t := \bigcup_{n=1}^\infty X_{n,t}$ is a Borel set for every t. This set generates the finite σ-algebra $\mathcal{X}_t = \{\varnothing, X, X_t, X\backslash X_t\}$. Let us define the probability measure μ_t on $\mathcal{X}_t$ by the equalities $\mu_t(X_t) = 1$, $\mu_t(X\backslash X_t) = 0$. Denote by $\mathcal{X}$ the σ-algebra generated by all $\mathcal{X}_t$, where $t \in [0,1]$.

(i) Show that there exists a unique measure μ on $\mathcal{X}$ that coincides with μ_t on $\mathcal{X}_t$ for each t and assumes only two values 0 and 1, hence is separable (E_μ contains only two classes, corresponding to $\varnothing$ and X).

(ii) Verify that μ has no countably additive extensions to $\mathcal{B}(X)$.

Hint: (i) it suffices to show that there is a countably additive measure μ on the algebra $\mathcal{A}_0$ generated by all $\mathcal{X}_t$ with $\mu(X_t) = 1$. This follows by Theorem 10.10.4, but a straightforward verification is possible. Any set in $\mathcal{A}_0$ has the form $A = \bigcup_{i=1}^n A_i$, where every A_i is the intersection $\bigcap_{j=1}^m Y_j$ with Y_j being one of the sets X_t or $X\backslash X_t$. Let $\mu(A) = 1$ if at least for one of A_i among Y_j there are no complements of the sets X_t, otherwise let $\mu(A) = 0$. Let us verify the countable additivity of μ. If $\{t_i\}$ is a finite or countable set of distinct numbers and sets $Y_{t_i} \in \mathcal{X}_{t_i}$ are nonempty, then $\bigcap_{i=1}^\infty Y_{t_i}$ is nonempty as well. Hence if a set $B \in \mathcal{A}_0$ is a finite or countable union of disjoint sets $B_j \in \mathcal{A}_0$, then at most one of them has a nonzero measure. (ii) If such an extension $\widetilde{\mu}$ exists, then $\widetilde{\mu}(X_t) = 1$ for all $t \in [0,1]$. For every t, there exist numbers $n(t)$ such that $\widetilde{\mu}(X_{n(t),t}) = \mu(X_{n(t),t}) > 0$. The cardinality arguments show that there exist n_0 and an uncountable set $T \subset [0,1]$ such that $n(t) = n_0$ for all $t \in T$. This gives a contradiction, since $X_{n_0,t} \cap X_{n_0,s} = \varnothing$ if $t \neq s$.

9.12.64. (Marczewski [**1252**]) Let μ be a Borel probability measure on a metric space (X, d) and let f_t, $t \geq 0$, be a family of one-to-one measurable transformations such that $\mu\bigl(f_t(E)\bigr) = \mu(E)$ for all measurable sets E and all t. Suppose that f_0 is the identity transformation and for every $\varepsilon > 0$, there is $\delta > 0$ such that $d\bigl(f_t(x), x\bigr) < \varepsilon$ whenever $t < \delta$ and $x \in X$. Prove that for every measurable set E, there exists $\tau > 0$ such that $E \cap f_t(E)$ is nonempty for all $t \leq \tau$.

9.12.65. (Holický, Ponomarev, Zajíček, Zelený [**852**]) Let $n \in \mathbb{N}$ and let X be a metrizable Souslin space with an atomless Radon probability measure μ. Prove that there exists a compact set $K \subset X$ with $\mu(K) > 0$ that is homeomorphic to the Cantor set and that can be mapped onto $[0,1]^n$ by means of a continuous mapping ψ with the following property: $\lambda_n(A) = 0$ precisely when $\mu\bigl(\psi^{-1}(A)\bigr) = 0$, where λ_n is Lebesgue measure.

9.12.66. Prove that there is no nonzero countably additive σ-finite measure on $\mathcal{B}(\mathbb{R}^\infty)$ that is invariant with respect to all translations.

Hint: any σ-finite Borel measure on a separable Fréchet space is concentrated on a proper subspace.

9.12.67. (Baker [**95**]) Prove that on $\mathcal{B}(\mathbb{R}^\infty)$, there exists a countably additive measure λ_∞ with values in $[0, +\infty]$ that is invariant with respect to translations and $\lambda_\infty\Bigl(\prod_{i=1}^\infty (a_i, b_i)\Bigr) = \prod_{i=1}^\infty |b_i - a_i|$ for all intervals (a_i, b_i) with the convergent product of lengths (the measure λ_∞ cannot be σ-finite).

9.12.68. (Kwapień [**1094**]) Let f be a bounded Lebesgue measurable function on $[0,1]$ with the zero integral over $[0,1]$. Prove that there exist a one-to-one transformation $T\colon\ [0,1] \to [0,1]$ preserving Lebesgue measure and a bounded measurable function g on $[0,1]$ with $f = g \circ T - g$ a.e.

9.12.69. (Anosov [**55**]) (i) Let T be a measure-preserving mapping on a probability space $(X, \mathcal{A}, \mu)$ and let $f \in \mathcal{L}^1(\mu)$. Suppose that there exists a measurable function g such that $g\bigl(T(x)\bigr) - g(x) = f(x)$ a.e. Prove that the integral of f vanishes.

(ii) Prove that for every irrational number α, there exist a continuous function f and a nonnegative measurable function g on the real line that have a period 1 and satisfy the equality $g(x + \alpha) - g(x) = f(x)$ a.e., but g is not integrable over $[0,1]$.

(iii) Prove that there exists an irrational number α such that in (ii) one can take for f an analytic function.

9.12.70. (Ryll-Nardzewski [**1631**], Marczewski [**1254**]) Suppose $(X_i, \mathcal{S}_i, \mu_i)$, $i \in I$, is an arbitrary family of measurable spaces with perfect probability measures. Let $X = \prod_i X_i$, let $\pi_i\colon\ X \to X_i$ be the natural projections, and let $\mathcal{A}$ be the algebra generated by all sets $\pi_i^{-1}(A_i)$, $A_i \in \mathcal{S}_i$. Suppose that ν is a finitely additive nonnegative set function on $\mathcal{A}$ such that its image under the projection π_i coincides with μ_i for all $i \in I$. Prove that ν is countably additive and its countably additive extension to $\mathcal{S} = \bigotimes_i \mathcal{S}_i$ is a perfect measure. In particular, every product of perfect probability measures is perfect. Prove an analogous assertion for compact measures.

9.12.71. (Plebanek [**1466**]) (i) Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be probability spaces such that at least one of them is perfect. For any $E \subset X \times Y$ let

$$\eta(E) := \sup\bigl\{\mu(A) + \nu(B)\colon\ A \in \mathcal{A}, B \in \mathcal{B}, (A \times B) \cap E = \varnothing\bigr\}.$$

Let $D \in \mathcal{A} \otimes \mathcal{B}$ and $d \geq 0$. Prove that the following conditions are equivalent:

(a) for every $\varepsilon > 0$, there exists a probability measure φ on $\mathcal{A}\otimes\mathcal{B}$ such that its projections on X and Y are, respectively, μ and ν, and $\varphi(D) \geq 1 - \varepsilon - d$;

(b) for every $\varepsilon > 0$, there exists a set L in the minimal lattice of sets that is closed with respect to countable intersections and contains $\mathcal{A}\times\mathcal{B}$, such that $L \subset D$ and $\eta(L) \leq 1 + \varepsilon + d$.

(ii) Let X and Y be Hausdorff spaces, let μ be a Radon probability measure on X, and let ν be a Borel probability measure on Y. Suppose that $D \subset X\times Y$ is a closed set and $d \geq 0$. Prove that the existence of a Borel probability measure φ on $X\times Y$ with the projections μ and ν and $\varphi(D) \geq 1-d$ is equivalent to the inequality $\mu(A) + \nu(B) \leq 1 + d$ for every Borel rectangle $A\times B \subset (X\times Y)\backslash D$.

9.12.72. Let $(X, \mathcal{A}, \mu)$ be a probability space.

(i) Let $f_n\colon X \to [a, b]$ be μ-measurable functions, $n \in \mathbb{N}$. Prove that there exists a strictly increasing sequence of μ-measurable functions $n_k\colon X \to \mathbb{N}$ such that $\lim\limits_{k\to\infty} f_{n_k(x)}(x) = \limsup\limits_{n\to\infty} f_n(x)$ for every $x \in X$.

(ii) Let K be a compact metric space and let $f_n\colon X \to K$ be μ-measurable mappings, $n \in \mathbb{N}$. Prove that there exists a strictly increasing sequence of μ-measurable functions $n_k\colon X \to \mathbb{N}$ such that, for every $x \in X$, the sequence $f_{n_k(x)}(x)$ converges in K.

HINT: (i) the function $\varphi(x) = \limsup\limits_{n\to\infty} f_n(x)$ is measurable, hence the inductively defined functions $n_k(x) = \min\bigl\{n > n_{k-1}(x)\colon f_n(x) \geq \varphi(x) - k^{-1}\bigr\}$ are measurable. Indeed, the set $\{x\colon n_k(x) = j\}$ consists of the points x such that $j > n_{k-1}(x)$, $f_j(x) \geq \varphi(x) - k^{-1}$ and either $j - 1 \leq n_{k-1}(x)$ or $j - 1 > n_{k-1}(x)$ and $f_{j-1}(x) < \varphi(x) - k^{-1}$. (ii) There exist a compact set $S \subset [0, 1]$ and a continuous mapping ψ from S onto K. According to Theorem 6.9.7, there exists a Borel set $B \subset S$ that ψ maps injectively onto K. Let $g\colon K \to B$ be the inverse mapping to $\psi|_B$. Since g is Borel, one can apply (i) to the functions $g\circ f_n$ and use that if a sequence $g\circ f_{n_k(x)}(x)$ is fundamental in B, then the sequence $f_{n_k(x)}(x)$ converges in K.

9.12.73. Let T be a Borel automorphism of a complete separable metric space E and let $C \subset E$ be a nonempty compact set.

(i) (Oxtoby, Ulam [**1411**]) Let $\limsup\limits_{n\to\infty} n^{-1}\sum_{k=1}^n I_C(T^k p) > 0$ for some $p \in C$. Prove that there exists a Borel probability measure μ on E with $\mu(C) > 0$ that is invariant with respect to T.

(ii) (Oxtoby, Ulam [**1410**]) Prove that there is a point $p \in C$ such that there exists $\lim\limits_{n\to\infty} n^{-1}\sum_{k=1}^n I_C(T^k p)$.

HINT: (i) see [**1411**]; (ii) if for some $p \in C$ condition (i) is fulfilled, then the claim follows by the ergodic theorem (see Chapter 10) applied to the measure μ and the function I_C; otherwise, for every point $p \in C$, the above limit equals zero, so that again the claim is true.

9.12.74. (Adamski [**9**]) Suppose we are given a Hausdorff space X and a continuous mapping $T\colon X \to X$. Prove that the following conditions are equivalent:

(i) there exists a Radon probability measure μ invariant with respect to T,

(ii) there exists a Radon probability measure ν such that for every open set $U \subset X$, the images of ν with respect to the functions $n^{-1}\sum_{i=0}^{n-1} I_U \circ T^i$ converge weakly,

(iii) there exist a compact set $K \subset X$ and a point $x_0 \in X$ such that the following inequality holds: $\limsup\limits_{n\to\infty} n^{-1}\sum_{i=0}^{n-1} I_K \circ T^i(x_0) > 0$.

9.12.75. (Fremlin, Garling, Haydon [**636**]) Let X and Y be topological spaces such that continuous functions separate their points. Let $f\colon X \to Y$ be a continuous mapping. Set $\widehat{f}\colon \mathcal{M}_t(X) \to \mathcal{M}_t(Y)$, $\mu \mapsto \mu \circ f^{-1}$.

(i) Suppose that for every uniformly tight set $M \subset \widehat{f}\bigl(\mathcal{M}_t(X)\bigr)$, there exists a uniformly tight set $M' \subset \mathcal{M}_t(X)$ such that $\widehat{f}(M') = M$. Show that for every compact set $K \subset f(X)$, there exists a compact set $K' \subset X$ such that $f(K') = K$.

(ii) Construct an example where the assertion inverse to (i) is false.

HINT: (i) let $D = \{\delta_y, y \in K\}$. Then $D \subset \widehat{f}\bigl(\mathcal{M}_t(X)\bigr)$ and D is uniformly tight. Let $C \subset \mathcal{M}_t(X)$ be uniformly tight and $\widehat{f}(C) = D$. Let us take a compact set $K_0 \subset X$ such that $|\mu|(X\backslash K_0) \le 1/2$ for all $\mu \in C$. Let $K' := K_0 \cap f^{-1}(K)$. If $y \in K$, then there is $\mu \in C$ with $\widehat{f}(\mu) = \delta_y$. Then $\mu\bigl(f^{-1}(y)\bigr) = 1$. Hence $f^{-1}(y)$ is not contained in $X\backslash K_0$, i.e., there exists $x \in K_0$ with $f(x) = y$. Thus, $K \subset f(K')$, whence $f(K') = K$. (ii) Let $X = \mathbb{N}$, $Y = \beta\,\mathbb{N}$, $f(n) = n$.

9.12.76. Prove the uniqueness assertion in Theorem 9.12.2.

HINT: see Fremlin, Garling, Haydon [**636**, Theorem 12].

9.12.77. (i) Let $(X, \mathcal{M}, \mu)$ be a Lebesgue–Rohlin space with a probability measure μ and let f be a finite measurable function. Prove that there exists a measurable mapping $h\colon M \to M$ that is one-to-one on a set of full measure such that the function $f \circ h$ is integrable.

(ii) Suppose that in (i) the measurable space is the unit cube with Lebesgue measure. Show that for h one can take some homeomorphism.

HINT: (i) the probability measure $\nu := c(1+|f|)^{-1}\cdot\mu$, where c is a normalization constant, is equivalent to the measure μ and $f \in L^1(\nu)$. There exists an isomorphism h of the spaces $(X, \mathcal{M}, \mu)$ and $(X, \mathcal{M}, \nu)$. It remains to observe that the integral of $|f| \circ h$ with respect to the measure μ equals the integral of $|f|$ with respect to the measure $\mu \circ h^{-1} = \nu$. (ii) The existence of a homeomorphism h in the case of the cube with Lebesgue measure follows by Theorem 9.6.5.

9.12.78. Let μ be a Haar measure on a locally compact group G. Show that $L^2(\mu)$ has an orthonormal basis consisting of continuous functions.

HINT: see Fremlin [**635**, §444X(n)].

9.12.79. Show that on the set $\{(x, y)\colon\ x < y\}$ in the square $[0, 1]^2$, there is no Borel measure whose projections to the sides are Lebesgue measures.

HINT: for any $\alpha \in (0, 1)$, the triangle $y < \alpha$, $x < y$ must have measure α with respect to such a measure, and the triangle $y > \alpha$, $\alpha < x < y$ must have measure $1 - \alpha$, which for the rectangle $x < \alpha$, $y > \alpha$ leaves only measure zero.

9.12.80. Let E be a nowhere dense Souslin set in a closed cube K in $\mathbb{R}^n$. Prove that there exists a homeomorphism $h\colon K \to K$ such that $h(E)$ has measure zero. In particular, any nowhere dense compact set is homeomorphic to a compact set of Lebesgue measure zero.

HINT: take the measure $\mu\colon B \mapsto \lambda(B\backslash E)/\lambda(K\backslash E)$ on K and apply Theorem 9.6.5.

9.12.81. (A.V. Korolev) Let Λ_k denote the set of the images of Lebesgue measure under k times continuously differentiable mappings from $[0, 1]$ to $[0, 1]$, $k \in \mathbb{N} \cup \{\infty\}$. Show that all the classes Λ_k are distinct.

HINT: show that for every measure $\mu \in \Lambda_k$, every interval contains a subinterval on which μ has a $k - 1$ times continuously differentiable density.

9.12.82. (Ochakovskaya [**1389**]) Show that there is a one-to-one transformation $\Phi \in C^\infty(\mathbb{R}^n, \mathbb{R}^n)$ with a positive Jacobian such that, for every ball $B_1(y)$ of unit radius one has $\lambda_n\bigl(\Phi(B_1(y))\bigr) = \lambda_n\bigl(B_1(y)\bigr)$, but Φ does not preserve Lebesgue measure.

9.12.83. (Burkholder [**289**]) Let μ be an atomless probability measure and let f be a μ-measurable function. Show that there is a μ-measurable function g with values in $[0,1]$ such that the measure $\mu\circ(f+g)^{-1}$ has no atoms.

HINT: the measure $\mu\circ f^{-1}$ has at most countably many atoms d_n; for every n, there is a μ-measurable function g_n on $E_n := f^{-1}(d_n)$ that transforms the measure $\mu(E_n)^{-1}\mu|_{E_n}$ into Lebesgue measure on $[0,1]$. Let $g(x) = g_n(x)$ if $x \in E_n$ and $g(x) = 0$ if $x \notin \bigcup_{n=1}^\infty E_n$. It is readily seen that $\mu\bigl((f+g)^{-1}(c)\bigr) = 0$ for all $c \in \mathbb{R}^1$.

9.12.84. (Blackwell [**179**]) Prove the following extension of Lyapunov's theorem: if $\mu_1, \ldots, \mu_n$ are atomless measures on a measurable space $(X, \mathcal{A})$, $E \subset \mathbb{R}^n$, then the set of all vectors of the form $(v_1, \ldots, v_n)$, where

$$v_i = \int_X a_i\, d\mu, \quad a_i \in L^1(\mu_i),\ (a_1, \ldots, a_n)\colon\ X \to E,$$

is convex. Lyapunov's theorem corresponds to the set E consisting of the two points $(0, \ldots, 0)$ and $(1, \ldots, 1)$.

9.12.85. Let μ be an atomless Borel probability measure on a separable metric space X. Show that there exists a sequence of sets $X_n \subset X$ such that $X_{n+1} \subset X_n$, $\mu^*(X_n) = 1$, $\bigcap_{n=1}^\infty X_n = \varnothing$.

HINT: by means of an isomorphism reduce the assertion to the case of Lebesgue measure restricted to a subset of $(0,1)$ and use the method from Exercise 1.12.58.

9.12.86. (i) Let μ be a probability measure on a σ-algebra $\mathcal{A}$ and let $\mathcal{E}$ be a family of sets from $\mathcal{A}$ with the following property: for every set $A \in \mathcal{A}$ with $\mu(A) > 0$, there is a set $E \in \mathcal{E}$ with $E \subset A$ and $\mu(E) > 0$. Show that for every set $A \in \mathcal{A}$ with $\mu(A) > 0$, there is an at most countable family of pairwise disjoint sets $E_n \in \mathcal{E}$ with $E_n \subset A$ and $\mu\bigl(\bigcup_{n=1}^\infty E_n\bigr) = \mu(A)$.

(ii) Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be probability spaces. Suppose a mapping $f\colon X \to Y$ has the following property: for every set $B \in \mathcal{B}$ with $\nu(B) > 0$, there is a set $E \in \mathcal{B}$ such that $E \subset B$, $\nu(E) > 0$, $f^{-1}(E) \in \mathcal{A}_\mu$ and $\mu\bigl(f^{-1}(E)\bigr) \ge \nu(E)$. Prove that f is $(\mathcal{A}_\mu, \mathcal{B})$-measurable and $\mu\circ f^{-1} = \nu$.

HINT: (i) take a maximal (in the sense of inclusion) family of sets from $\mathcal{E}$ that have positive measures and are contained in A. (ii) Consider the class of sets

$$\mathcal{E} = \bigl\{E \in \mathcal{B}\colon\ f^{-1}(E) \in \mathcal{A}_\mu, \mu\bigl(f^{-1}(E)\bigr) \ge \nu(E)\bigr\}.$$

By (i), for every $B \in \mathcal{B}$, there is $E \in \mathcal{E}$ with $E \subset B$ and $\nu(E) = \nu(B)$. Observe that $\nu(E) = \mu\bigl(f^{-1}(E)\bigr)$ if $E \in \mathcal{E}$. Indeed, one can find $D \in \mathcal{E}$ with $D \subset Y\backslash E$ and $\nu(D) = 1 - \nu(E)$, whence it follows that

$$\nu(E) + \nu(D) \le \mu\bigl(f^{-1}(E)\bigr) + \mu\bigl(f^{-1}(D)\bigr) \le 1 = \nu(E) + \nu(D),$$

which is only possible if $\nu(E) = \mu\bigl(f^{-1}(E)\bigr)$. It follows that $\mathcal{E}$ is closed under complementation. Hence there is $E' \in \mathcal{E}$ with $Y\backslash B \subset E'$ and $\nu(E') = 1 - \nu(B)$, which yields $B \in \mathcal{E}$.

CHAPTER 10

Conditional measures and conditional expectations

> Then look round and see that none of the uninitiated are listening. They are the ones who think nothing else exists except what they can grasp firmly in their hands, and do not allow actions, processes, or any thing that is not visible to have any share in being.
>
> *Plato.* Theaetetus.

10.1. Conditional expectations

Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and let $\mathcal{B}$ be a sub-σ-algebra in $\mathcal{A}$.

10.1.1. Definition. *Let $f \in L^1(\mu)$. A conditional expectation of f with respect to the σ-algebra $\mathcal{B}$ and the measure μ is a $\mathcal{B}$-measurable μ-integrable function $\mathbb{E}_\mu^{\mathcal{B}} f$ such that*

$$\int_\Omega gf\,d\mu = \int_\Omega g\mathbb{E}_\mu^{\mathcal{B}} f\,d\mu \tag{10.1.1}$$

for every bounded $\mathcal{B}$-measurable function g.

A conditional expectation of an individual integrable function f is defined as the conditional expectation of the corresponding class in $L^1(\mu)$.

We note that if a $\mathcal{B}$-measurable function ψ equals $\mathbb{E}_\mu^{\mathcal{B}} f$ a.e., then it is a conditional expectation of f, too; however, among functions equivalent to $\mathbb{E}_\mu^{\mathcal{B}} f$, there are functions that are not $\mathcal{B}$-measurable. This requires certain additional precautions in the usual identifications of individual functions and their equivalence classes. Clearly, if we define the conditional expectation as an equivalence class of $\mathcal{B}$-measurable functions, then it is unique. In what follows, we shall not always distinguish individual functions serving as a conditional expectation from their equivalence class of $\mathcal{B}$-measurable functions.

The defining equality (10.1.1) is equivalent to the following relationship obtained by the substitution $g = I_B$:

$$\int_B f\,d\mu = \int_B \mathbb{E}_\mu^{\mathcal{B}} f\,d\mu, \quad \forall\, B \in \mathcal{B}. \tag{10.1.2}$$

The equivalence of the two relationships follows from the fact that every bounded $\mathcal{B}$-measurable function is the uniform limit of simple $\mathcal{B}$-measurable functions. Clearly, one has

$$\int_\Omega f\,d\mu = \int_\Omega \mathbb{E}_\mu^{\mathcal{B}} f\,d\mu.$$

If $\mathcal{B} = \{\varnothing, \Omega\}$, then $\mathbb{E}^{\mathcal{B}} f$ coincides with the integral of f over Ω. If μ is a probability measure, the integral of f over the whole space Ω is denoted sometimes by $\mathbb{E}f$ and is called the expectation of f. This tradition from probability theory explains the above notation and terminology.

In the case where only one measure μ is given, for simplification of notation and terminology, in place of $\mathbb{E}_\mu^{\mathcal{B}}$ one uses the symbol $\mathbb{E}^{\mathcal{B}}$ and in the corresponding term the indication of the measure is omitted: $\mathbb{E}^{\mathcal{B}} f$ is called the conditional expectation of f with respect to $\mathcal{B}$. In the probabilistic literature one frequently uses the notation $\mathbb{E}(f|\mathcal{B})$. If ξ is an integrable function on a probability space and $\mathcal{B}$ is generated by a measurable function (or mapping) η, then one uses the notation $\mathbb{E}(\xi|\eta)$, i.e., $\mathbb{E}(\xi|\eta) = \mathbb{E}^{\sigma(\eta)}\xi$.

10.1.2. Example. Let μ be a probability measure and let Ω be partitioned into finitely or countably many pairwise disjoint measurable sets B_i with $\mu(B_i) > 0$. Denote by $\mathcal{B}$ the σ-algebra generated by the sets B_i. Then one has

$$\mathbb{E}^{\mathcal{B}} f(\omega) = \sum_{i=1}^{\infty} \int_{B_i} f\,d\mu\, \frac{I_{B_i}(\omega)}{\mu(B_i)}.$$

PROOF. It is clear that the above series defines an integrable $\mathcal{B}$-measurable function. It is easily seen that the $\mathcal{B}$-measurable functions are exactly the functions that are constant on the sets B_i. Hence it suffices to verify that both sides of the equality to be proven have equal integrals after multiplication by I_{B_i}. The integral of $I_{B_i}\mathbb{E}^{\mathcal{B}} f$ by definition equals the integral of f over the set B_i, which obviously coincides with the integral of the right-hand side multiplied by I_{B_i}, since $B_i \cap B_j = \varnothing$ if $j \neq i$. $\square$

10.1.3. Example. Let μ_n be Borel probability measures on the real line, let $\Omega = \mathbb{R}^\infty$, and let $\mu = \bigotimes_{n=1}^\infty \mu_n$. Let $\mathcal{B}_n$ be the σ-algebra generated by the first n coordinate functions. Then

$$\mathbb{E}^{\mathcal{B}_n} f(x_1, \ldots, x_n) = \int f(x_1, \ldots, x_n, x_{n+1}, \ldots) \bigotimes_{k=n+1}^{\infty} \mu_k \,\bigl(d(x_{n+1}, x_{n+2}, \ldots)\bigr),$$

where the integration is taken over the product of real lines corresponding to the variables x_k with $k \geq n+1$.

PROOF. Suppose that g is a bounded Borel function of $x_1, \ldots, x_n$. By Fubini's theorem the integral of the right-hand side of the equality to be proven multiplied by g equals the integral of fg. $\square$

10.1.4. Example. Let us consider Lebesgue measure λ on $[0,1)$ and let $T_k(x) = (x + 2^{-k})\mathrm{mod}(1)$, $k \in \mathbb{N}$, $x \in [0,1)$. Let

$$\mathcal{B}_k := \bigl\{B \in \mathcal{B}\bigl([0,1)\bigr)\colon\ T_k(B) = B\bigr\}.$$

Then a Borel function f is measurable with respect to $\mathcal{B}_k$ if and only if the equality $f = f \circ T_k$ holds. In addition, one has

$$\mathbb{E}^{\mathcal{B}_k} f = 2^{-k} \sum_{j=0}^{2^k-1} f \circ T_k^j, \quad \forall\, f \in L^1[0,1). \tag{10.1.3}$$

PROOF. The first claim is true because it is true for indicators of sets. Denote by g the function on the right-hand side of (10.1.3). It is clear that $g \circ T_k = g$ and hence g is measurable with respect to $\mathcal{B}_k$. Since

$$\int_0^1 \psi\, dx = \int_0^1 \psi \circ T_k\, dx$$

for all $\psi \in L^1[0,1)$, given $B \in \mathcal{B}_k$, in view of the equality $I_B \circ T_k = I_B$ we have

$$\begin{aligned}\int_B g(x)\, dx &= 2^{-k} \sum_{j=0}^{2^k-1} \int_0^1 I_B(x) f\bigl(T_k^j(x)\bigr)\, dx \\ &= 2^{-k} \sum_{j=0}^{2^k-1} \int_0^1 I_B\bigl(T_k^j(x)\bigr) f\bigl(T_k^j(x)\bigr)\, dx = \int_0^1 I_B(x) f(x)\, dx,\end{aligned}$$

which proves the second assertion. □

The existence of conditional expectation and its basic properties are established in the next theorem.

10.1.5. Theorem. *Suppose that μ is a probability measure. To every function $f \in \mathcal{L}^1(\mu)$, one can associate a $\mathcal{B}$-measurable function $\mathbb{E}^{\mathcal{B}} f$ such that*

(1) $\mathbb{E}^{\mathcal{B}}$ *is a conditional expectation of f with respect to $\mathcal{B}$*;

(2) $\mathbb{E}^{\mathcal{B}} f = f$ *μ-a.e. for every $\mathcal{B}$-measurable μ-integrable function f*;

(3) $\mathbb{E}^{\mathcal{B}} f \geq 0$ *μ-a.e. if $f \geq 0$ μ-a.e.*;

(4) *if a sequence of μ-integrable functions f_n converges monotonically decreasing or increasing to a μ-integrable function f, then $\mathbb{E}^{\mathcal{B}} f_n \to \mathbb{E}^{\mathcal{B}} f$ μ-a.e.*;

(5) *For every $p \in [1, +\infty]$, the mapping $\mathbb{E}^{\mathcal{B}}$ defines a continuous linear operator with norm 1 on the space $L^p(\mu)$. In addition, $\mathbb{E}^{\mathcal{B}}$ is the orthogonal projection of $L^2(\mu)$ to the closed linear subspace generated by $\mathcal{B}$-measurable functions.*

PROOF. It is clear that the restriction of the measure $f \cdot \mu$ to $\mathcal{B}$ is a measure absolutely continuous with respect to the restriction of μ to $\mathcal{B}$. By the Radon–Nikodym theorem, there exists a $\mathcal{B}$-measurable μ-integrable function $\mathbb{E}^{\mathcal{B}} f$ such that one has (10.1.2). We show that this function possesses the required properties. It is clear that $\mathbb{E}^{\mathcal{B}} f$ depends only on the equivalence class of f. The mapping $\mathbb{E}^{\mathcal{B}}$ defines a linear operator with values in $L^1(\mu)$, i.e., one has $\mathbb{E}^{\mathcal{B}}(f+g) = \mathbb{E}^{\mathcal{B}} f + \mathbb{E}^{\mathcal{B}} g$ and $\mathbb{E}^{\mathcal{B}}(cf) = c\mathbb{E}^{\mathcal{B}} f$ μ-a.e. for all $f, g \in \mathcal{L}^1(\mu)$ and $c \in \mathbb{R}^1$. This follows by the fact that the Radon–Nikodym density is defined uniquely up to equivalence. Substituting in (10.1.1) the function $g = \operatorname{sign} \mathbb{E}^{\mathcal{B}} f$ (this function is $\mathcal{B}$-measurable), we conclude that the norm of the operator $\mathbb{E}^{\mathcal{B}}$ on $L^1(\mu)$ does not exceed 1. In fact, it equals 1, since $\mathbb{E}^{\mathcal{B}} 1 = 1$. Properties (2) and (3) are obvious. If a sequence of functions $f_n \in \mathcal{L}^1(\mu)$ is increasing to a function $f \in \mathcal{L}^1(\mu)$, then by Property (3) the sequence of functions $g_n = \mathbb{E}^{\mathcal{B}} f_n$ is a.e. increasing. The function $g = \lim\limits_{n\to\infty} g_n$ is $\mathcal{B}$-measurable and μ-integrable by Fatou's theorem. It is clear by (10.1.2) and the dominated convergence theorem that g can be taken for $\mathbb{E}^{\mathcal{B}} f$. Property (5) follows by the established properties, but it can be verified directly. To this end, it suffices to observe that $\|\mathbb{E}^{\mathcal{B}} f\|_{L^p(\mu)}$ coincides with the supremum of the quantities

$$\int_\Omega \psi \mathbb{E}^{\mathcal{B}} f \, d\mu = \int_\Omega \psi f \, d\mu$$

over all $\mathcal{B}$-measurable functions ψ such that $\|\psi\|_{L^q(\mu)} = 1$ and $q^{-1} + p^{-1} = 1$. It remains to apply Hölder's inequality. Finally, if $p = 2$, then $f - \mathbb{E}^{\mathcal{B}} f \perp h$ for every $\mathcal{B}$-measurable function $h \in \mathcal{L}^2(\mu)$. If the function h is bounded, this follows by (10.1.1), and in the general case this is obtained in the limit. $\square$

The established theorem remains valid for σ-finite measures: see Exercise 3.10.31(ii) in Chapter 3. In the case $p < \infty$ it extends to arbitrary infinite measures, since every function $f \in L^p(\mu)$ is concentrated on a set with a σ-finite measure. Finally, in the obvious way $\mathbb{E}^{\mathcal{B}}$ extends to complex-valued functions.

10.1.6. Corollary. *Suppose that μ is a nonnegative measure with values in $[0, +\infty]$. Then, for every $p \in [1, +\infty)$, there exists a bounded operator $\mathbb{E}^{\mathcal{B}} \colon L^p(\mu) \to L^p(\mu)$ that possesses properties* (1)–(4) *on $\mathcal{L}^1(\mu) \cap \mathcal{L}^p(\mu)$.*

It can be observed from the proof that the constructed mapping $\mathbb{E}^{\mathcal{B}}$ may not be pointwise linear, i.e., *it is not claimed* that

$$\mathbb{E}^{\mathcal{B}}(f+g)(\omega) = \mathbb{E}^{\mathcal{B}}(f)(\omega) + \mathbb{E}^{\mathcal{B}}(g)(\omega) \quad \text{for all } f, g \in \mathcal{L}^1(\mu) \text{ and } \omega \in \Omega.$$

The problem of the existence of versions with pointwise preservation of linear relationships is discussed in §10.4, where we study the so-called regular conditional measures, by means of which one can effectively define conditional expectations.

Let us establish some other useful properties of conditional expectations. For simplification of formulations we shall extend the conditional expectation to those non-integrable functions f for which one of the functions f^+ or f^- is integrable. In this case we let $\mathbb{E}^{\mathcal{B}}_\mu f = \mathbb{E}^{\mathcal{B}}_\mu f^+ - \mathbb{E}^{\mathcal{B}}_\mu f^-$, where for any nonnegative measurable function $\varphi \colon \Omega \to [0, +\infty]$, the conditional expectation $\mathbb{E}^{\mathcal{B}}_\mu \varphi$ is defined as $\mathbb{E}^{\mathcal{B}}_\mu \varphi := \lim\limits_{n\to\infty} \mathbb{E}^{\mathcal{B}}_\mu \min(\varphi, n)$. One can also use Exercise 3.10.31 in Chapter 3, but it should be noted that even for a finite function φ, the restriction of the σ-finite measure $\nu := \varphi I_{\{\varphi<\infty\}} \cdot \mu$ to $\mathcal{B}$ need not be σ-finite (see Exercise 1.12.80 in Chapter 1).

10.1.7. Proposition. *Suppose that $\mathcal{B}$ is a sub-σ-algebra in $\mathcal{A}$ and that $f_n \in \mathcal{L}^1(\mu)$, $n \in \mathbb{N}$. Then:*

(i) *if $f_n \to f$ and $|f_n| \le F$ μ-a.e., where $F \in \mathcal{L}^1(\mu)$, then*

$$\mathbb{E}_\mu^{\mathcal{B}} f = \lim_{n\to\infty} \mathbb{E}_\mu^{\mathcal{B}} f_n \quad \mu\text{-a.e.};$$

(ii) *if $f_n \le F$ μ-a.e., where $F \in \mathcal{L}^1(\mu)$, then*

$$\limsup_n \mathbb{E}_\mu^{\mathcal{B}} f_n \le \mathbb{E}_\mu^{\mathcal{B}} \limsup_n f_n \quad \mu\text{-a.e.};$$

(iii) *if $f_n \ge G$ μ-a.e., where $G \in \mathcal{L}^1(\mu)$, then*

$$\mathbb{E}_\mu^{\mathcal{B}} \liminf_n f_n \le \liminf_n \mathbb{E}_\mu^{\mathcal{B}} f_n \quad \mu\text{-a.e.}$$

PROOF. (i) Set $h_n = \sup_{k\ge n} |f_k - f|$. Then the sequence h_n decreases a.e. to zero. The sequence $\mathbb{E}^{\mathcal{B}} h_n$ decreases a.e., hence $h := \lim\limits_{n\to\infty} \mathbb{E}^{\mathcal{B}} h_n$ exists a.e. By the already-known properties of the conditional expectation one has

$$|\mathbb{E}^{\mathcal{B}} f_n - \mathbb{E}^{\mathcal{B}} f| = |\mathbb{E}^{\mathcal{B}}(f_n - f)| \le \mathbb{E}^{\mathcal{B}}|f_n - f| \le \mathbb{E}^{\mathcal{B}} h_n \quad \text{a.e.}$$

Hence it suffices to verify that $h = 0$ a.e. It remains to observe that the integral of the nonnegative function h equals zero, since for all n we have

$$\int_\Omega h\, d\mu \le \int_\Omega \mathbb{E}^{\mathcal{B}} h_n\, d\mu = \int_\Omega h_n\, d\mu,$$

and the right-hand side of this relationship tends to zero by the dominated convergence theorem, which is applicable by the estimate $0 \le h_n \le 2F$.

(ii) Set $f = \lim\limits_{n\to\infty} \sup_{k\ge n} f_k$. The functions $\sup_{k\ge n} f_k$ are decreasing to f and majorized by F. If f is integrable, then by (i) we have a.e.

$$\mathbb{E}^{\mathcal{B}} f = \lim_{n\to\infty} \mathbb{E}^{\mathcal{B}} \sup_{k\ge n} f_k = \limsup_{n\to\infty} \mathbb{E}^{\mathcal{B}} \sup_{k\ge n} f_k \ge \limsup_{n\to\infty} \mathbb{E}^{\mathcal{B}} f_n.$$

In the general case, we have to justify the first equality in the above relationship, i.e., to show that if integrable functions g_n are decreasing to a function g and $g_n \le F$, then $\mathbb{E}^{\mathcal{B}} g_n \to \mathbb{E}^{\mathcal{B}} g$ a.e. Note that $\mathbb{E}^{\mathcal{B}} g_{n+1} \le \mathbb{E}^{\mathcal{B}} g_n$ a.e. Hence $\zeta := \lim\limits_{n\to\infty} \mathbb{E}^{\mathcal{B}} g_n$ exists a.e. Since $\zeta \le \mathbb{E}^{\mathcal{B}} F$ and $\mathbb{E}^{\mathcal{B}} g \le \mathbb{E}^{\mathcal{B}} g_n \le \mathbb{E}^{\mathcal{B}} F$ a.e., for every $B \in \mathcal{B}$ such that $\zeta I_B \in L^1(\mu)$, we have

$$\int_B \zeta\, d\mu = \lim_{n\to\infty} \int_B g_n\, d\mu = \int_B g\, d\mu.$$

It is easy to see that this equality remains valid in the case where the integral of ζI_B equals $-\infty$. Therefore, $\zeta = \mathbb{E}^{\mathcal{B}} g$ a.e. Finally, (iii) follows by (ii). $\square$

Note the following simple property of the conditional expectation: if $\mathcal{B}_1$ is a sub-σ-algebra in $\mathcal{B}$, then

$$\mathbb{E}^{\mathcal{B}_1} \mathbb{E}^{\mathcal{B}} f = \mathbb{E}^{\mathcal{B}_1} f = \mathbb{E}^{\mathcal{B}} \mathbb{E}^{\mathcal{B}_1} f. \tag{10.1.4}$$

Indeed, for every bounded $\mathcal{B}_1$-measurable function g we have

$$\int_\Omega g \mathbb{E}^{\mathcal{B}_1} \mathbb{E}^{\mathcal{B}} f\, d\mu = \int_\Omega g \mathbb{E}^{\mathcal{B}} f\, d\mu = \int_\Omega g f\, d\mu,$$

since g is $\mathcal{B}$-measurable. The second equality in (10.1.4) follows by Property (2) in Theorem 10.1.5.

10.1.8. Proposition. *Let $(X, \mathcal{A}, \mu)$ be a probability space, let $\mathcal{B} \subset \mathcal{A}$ be a sub-σ-algebra, and let a function f be measurable with respect to $\mathcal{B}$. Suppose that $g \in \mathcal{L}^1(\mu)$ and $fg \in \mathcal{L}^1(\mu)$. Then we have $\mathbb{E}^{\mathcal{B}}(fg) = f\mathbb{E}^{\mathcal{B}}g$ a.e.*

PROOF. If f is bounded, then this equality is obvious from the definition. In the general case, we consider the functions $fI_{\{|f|\leq n\}}$ convergent a.e. to f and majorized by $|f|$, and apply assertion (i) of the previous proposition. □

Let us extend Jensen's inequality to the conditional expectation.

10.1.9. Proposition. *Let μ be a probability measure, let f be a μ-integrable function, and let V be a convex function defined on an interval (a, b) (possibly unbounded) such that f takes values in (a, b) and the function $V \circ f$ is μ-integrable. Then $V(\mathbb{E}^{\mathcal{B}}f) \leq \mathbb{E}^{\mathcal{B}}(V \circ f)$ μ-a.e.*

PROOF. Suppose first that $f = \sum_{i=1}^n c_i I_{A_i}$, where $\sum_{i=1}^n I_{A_i} = 1$. Then we have $\sum_{i=1}^n \mathbb{E}^{\mathcal{B}} I_{A_i} = 1$ a.e. Therefore,

$$V\Bigl(\sum_{i=1}^n c_i \mathbb{E}^{\mathcal{B}} I_{A_i}\Bigr) \leq \sum_{i=1}^n V(c_i)\mathbb{E}^{\mathcal{B}} I_{A_i} \quad \text{a.e.}$$

The left-hand side of this inequality coincides with $V(\mathbb{E}^{\mathcal{B}}f)$ and the right-hand side equals $\mathbb{E}^{\mathcal{B}}(V \circ f)$. Let us consider the general case. If f is bounded and takes values in an interval $[c, d] \subset (a, b)$, then f is uniformly approximated by simple functions with values in $[c, d]$, which yields the required inequality in view of the above-considered case and the continuity of V on (a, b). If f is unbounded, then we set $f_n = f$ if $|f| \leq n$, $f_n = n$ if $f \geq n$, and $f_n = -n$ if $f \leq -n$. We may assume that (a, b) contains the origin. Then the functions $V \circ f_n$ are defined. For f_n the claim is already proven and one has $\mathbb{E}^{\mathcal{B}} f_n \to \mathbb{E}^{\mathcal{B}} f$ almost everywhere, whence we obtain $V(\mathbb{E}^{\mathcal{B}} f_n) \to V(\mathbb{E}^{\mathcal{B}} f)$ almost everywhere. Finally, $\mathbb{E}^{\mathcal{B}}(V \circ f_n) \to \mathbb{E}^{\mathcal{B}}(V \circ f)$ almost everywhere, since $V(f_n) \to V(f)$ almost everywhere and in $L^1(\mu)$. The latter follows by the fact that the functions $V(f_n)$ have an integrable majorant $|V(f)| + |V(f_1)| + C$, where $C = 0$ if $\inf_s V(s) = -\infty$ and $C = |\inf_s V(s)|$ otherwise. □

If a function $f \in \mathcal{L}^1(\mu)$ is fixed, then varying sub-σ-algebras of the main σ-algebra $\mathcal{A}$ we obtain a uniformly integrable family of functions.

10.1.10. Example. Let $(X, \mathcal{A}, \mu)$ be a probability space and let $\mathcal{A}_\alpha \subset \mathcal{A}$ be some family of sub-σ-algebras in $\mathcal{A}$, where $\alpha \in \Lambda$. Then $\{\mathbb{E}^{\mathcal{A}_\alpha} f\}_{\alpha \in \Lambda}$ is a uniformly integrable family.

PROOF. Since $\mathbb{E}^{\mathcal{A}_\alpha} f \leq \mathbb{E}^{\mathcal{A}_\alpha}|f|$, one can assume that $f \geq 0$. We apply the criterion of de la Vallée Poussin (Theorem 4.5.9). Let us take a nonnegative increasing convex function G on $[0, +\infty)$ with $\lim\limits_{t\to\infty} G(t)/t = +\infty$ and

$G \circ |f| \in L^1(\mu)$. By Jensen's inequality for conditional expectation we have

$$\int_X G \circ \mathbb{E}^{\mathcal{A}_\alpha} f \, d\mu \le \int_X G \circ f \, d\mu,$$

which by the cited theorem yields the uniform integrability of our family of functions. $\square$

In the case $f = I_A$, where $A \in \mathcal{A}$, the conditional expectation $\mathbb{E}^{\mathcal{B}} f$ is denoted by $\mu^{\mathcal{B}}(A)$ or $\mu(A|\mathcal{B})$ and is called the conditional measure (conditional probability in the case of probability measures) of A with respect to $\mathcal{B}$. In the case when $\mathcal{B}$ is the σ-algebra generated by a measurable function ξ, the notation $\mu(A|\xi)$ is also used. If ξ assumes only finitely or countably many values x_j on sets of positive measure, then one can express $\mu(A|\xi)$ by means of the numbers

$$\mu(A|\xi = x_i) = \frac{\mu\bigl(A \cap \{\xi = x_i\}\bigr)}{\mu\bigl(\{\xi = x_i\}\bigr)}$$

according to Example 10.1.2. In general, one can only say that for every $A \in \mathcal{A}$, there exists a Borel function ζ_A such that $\mu(A|\xi)(x) = \zeta_A\bigl(\xi(x)\bigr)$. Then one can set $\mu(A|\xi = x) := \zeta_A(x)$. The latter expression is referred to as "the measure of A under conditioning $\xi = x$". But it is not even claimed that for a fixed point x the conditional measure is indeed a measure in A (this may be false). Below we return to the question of when such a property can be achieved. In addition, we shall clarify a simple geometric meaning of conditional measures and conditional expectations.

In the case where $\mathcal{B}$ is generated by a mapping to $\mathbb{R}^n$, the conditional expectation can be evaluated by using the results on differentiation of measures obtained in §5.8(iii). Suppose that on a probability space $(\Omega, \mathcal{A}, P)$ we are given an integrable random variable ξ and a random vector η with values in $\mathbb{R}^n$. Let $B(x,r)$ denote the open ball with the center x and radius r. Then

$$\mathbb{E}(\xi|\eta) := \mathbb{E}^{\sigma(\eta)}\xi = f(\eta),$$

where the function f on $\mathbb{R}^n$ is defined by the formula

$$f(x) = \lim_{r \to 0} \frac{1}{P\bigl(\eta \in B(x,r)\bigr)} \int_{\{\eta \in B(x,r)\}} \xi \, dP$$

if this limit exists and $f(x) = 0$ if there is no finite limit. In particular, for every $B \in \mathcal{B}(\mathbb{R}^1)$ we have

$$P(\xi \in B|\eta) = f(\eta), \quad f(x) = \lim_{r \to 0} \frac{P\bigl(\xi \in B, \eta \in B(x,r)\bigr)}{P\bigl(\eta \in B(x,r)\bigr)}.$$

Indeed, let μ be the image of the measure P under the mapping η and let ν be the image of the measure $\xi \cdot P$. Since $\nu \ll \mu$, the Radon–Nikodym density $d\nu/d\mu$ equals

$$f(y) = \lim_{r \to 0} \frac{\mu\bigl(B(y,r)\bigr)}{\nu\bigl(B(y,r)\bigr)}$$

for μ-a.e. y (see Theorem 5.8.8), which by the change of variable formula coincides with the above expression. For every function of the form $\psi(\eta)$, where ψ is a bounded Borel function on $\mathbb{R}^n$, we have

$$\mathbb{E}[\xi\psi(\eta)] = \int_{\mathbb{R}^n} \psi\, d\nu = \int_{\mathbb{R}^n} \psi f\, d\mu = \mathbb{E}[\psi(\eta)f(\eta)],$$

which proves our claim.

10.2. Convergence of conditional expectations

The following theorem on convergence of conditional expectations with respect to an increasing family of σ-algebras is very important for applications. Most often one encounters increasing countable sequences of σ-algebras, but sometimes one has to deal with nets, so we prove this theorem in greater generality.

10.2.1. Theorem. *Let $(X, \mathcal{A}, \mu)$ be a probability space. Suppose we are given an increasing net of sub-σ-algebras $\mathcal{B}_\alpha \subset \mathcal{A}$. Denote by $\mathcal{B}_\infty$ the σ-algebra generated by all $\mathcal{B}_\alpha$. Then, for every $p \in [1, +\infty)$ and every $f \in L^p(\mu)$, the net $\mathbb{E}^{\mathcal{B}_\alpha} f$ converges in $L^p(\mu)$ to $\mathbb{E}^{\mathcal{B}_\infty} f$.*

PROOF. We may assume that $\mathcal{A} = \mathcal{B}_\infty$, since by the inclusion $\mathcal{B}_\alpha \subset \mathcal{B}_\infty$ one has

$$\mathbb{E}^{\mathcal{B}_\alpha} f = \mathbb{E}^{\mathcal{B}_\alpha}\mathbb{E}^{\mathcal{B}_\infty} f.$$

Let $f = I_B$, $B \in \mathcal{B}_\infty$. Given $\varepsilon > 0$, there exists a set C with $\mu(B \triangle C) < \varepsilon$ belonging to one of the σ-algebras $\mathcal{B}_\alpha$ (since such a set exists in the algebra generated by all $\mathcal{B}_\alpha$, and every set in this algebra is contained in one of the σ-algebras $\mathcal{B}_\alpha$ due to the fact that they form a directed family). Let $g = I_C$. Then $\mathbb{E}^{\mathcal{B}_\alpha} g = g$ for all α greater than some α_0 such that $C \in \mathcal{B}_{\alpha_0}$. Therefore,

$$f - \mathbb{E}^{\mathcal{B}_\alpha} f = f - g + \mathbb{E}^{\mathcal{B}_\alpha}(g - f).$$

The estimate

$$\|\mathbb{E}^{\mathcal{B}_\alpha} f - \mathbb{E}^{\mathcal{B}_\alpha} g\|_{L^p(\mu)} \le \|f - g\|_{L^p(\mu)} \le \varepsilon^{1/p}$$

shows that our claim is true for indicators. Therefore, it is true for all simple functions. Since simple functions are dense in $L^p(\mu)$, the general case follows by the fact that the operator $\mathbb{E}^{\mathcal{B}_\alpha}$ on $L^p(\mu)$ has the unit norm. $\square$

In the case of a countable sequence of σ-algebras, in addition to convergence in the mean one has almost everywhere convergence. The proof of this important fact is less elementary and is based on the following Doob inequality, which has a considerable independent interest.

10.2.2. Proposition. *Let $(X, \mathcal{A}, \mu)$ be a probability space and let $\{\mathcal{B}_n\}$ be an increasing sequence of sub-σ-algebras in $\mathcal{A}$. Then, for all f in $L^1(\mu)$ and $c > 0$, one has*

$$\mu\Bigl(x\colon\ \sup_i |\mathbb{E}^{\mathcal{B}_i} f(x)| > c\Bigr) \le \frac{1}{c}\int_X |f|\, d\mu. \qquad (10.2.1)$$

PROOF. It suffices to establish (10.2.1) for nonnegative f. Let $f_i = \mathbb{E}^{\mathcal{B}_i} f$, $E = \{x\colon\ \sup_i f_i(x) > c\}$ and

$$E_j = \{x\colon\ f_1(x) \le c,\ f_2(x) \le c, \ldots, f_{j-1}(x) \le c,\ f_j(x) > c\},$$

where $j \in \mathbb{N}$. It is clear that the sets E_j are measurable and disjoint and that their union is E. In addition, $E_j \in \mathcal{B}_j$, since $\mathcal{B}_1 \subset \mathcal{B}_2 \subset \cdots \subset \mathcal{B}_j$ and function f_i is $\mathcal{B}_i$-measurable. Therefore,

$$\int_X f\,d\mu \ge \int_E f\,d\mu = \sum_{j=1}^{\infty} \int_{E_j} f\,d\mu = \sum_{j=1}^{\infty} \int_{E_j} f_j\,d\mu \ge c \sum_{j=1}^{\infty} \mu(E_j) = c\mu(E),$$

as required. $\square$

10.2.3. Theorem. *Let $(X, \mathcal{A}, \mu)$ be a probability space, let $\{\mathcal{B}_n\}$ be an increasing sequence of sub-σ-algebras in $\mathcal{A}$, and let $f \in \mathcal{L}^1(\mu)$. Denote by $\mathcal{B}_\infty$ the σ-algebra generated by all $\mathcal{B}_n$. Then*

$$\mathbb{E}^{\mathcal{B}_\infty} f(x) = \lim_{n\to\infty} \mathbb{E}^{\mathcal{B}_n} f(x) \quad \textit{for } \mu\textit{-a.e. } x.$$

PROOF. Since $\mathbb{E}^{\mathcal{B}_n} f = \mathbb{E}^{\mathcal{B}_n} \mathbb{E}^{\mathcal{B}_\infty} f$ by the inclusion $\mathcal{B}_n \subset \mathcal{B}_\infty$, we may assume that $\mathcal{A} = \mathcal{B}_\infty$ and prove that $\mathbb{E}^{\mathcal{B}_n} f \to f$ a.e. Set

$$\psi(x) := \limsup_{n\to\infty} |\mathbb{E}^{\mathcal{B}_n} f(x) - f(x)|.$$

We show that $\psi(x) = 0$ a.e. Let $\varepsilon > 0$. For every sufficiently large n, there exists a $\mathcal{B}_n$-measurable integrable function g such that $\|f - g\|_{L^1(\mu)} < \varepsilon^2$. Then on account of the equality $\mathbb{E}^{\mathcal{B}_m} g = g$ for all $m \ge n$, we obtain

$$\begin{aligned}\psi(x) &\le \limsup_{n\to\infty} |\mathbb{E}^{\mathcal{B}_n}(f-g)(x)| + \limsup_{n\to\infty} |\mathbb{E}^{\mathcal{B}_n} g(x) - g(x)| + |f(x) - g(x)| \\ &= \limsup_{n\to\infty} |\mathbb{E}^{\mathcal{B}_n}(f-g)(x)| + |f(x) - g(x)|. \end{aligned} \tag{10.2.2}$$

By Doob's inequality we have

$$\begin{aligned}\mu\Bigl(x\colon\ \limsup_{n\to\infty} |\mathbb{E}^{\mathcal{B}_n}(f-g)(x)| > \varepsilon\Bigr) &\le \mu\Bigl(x\colon\ \sup_n |\mathbb{E}^{\mathcal{B}_n}(f-g)(x)| > \varepsilon\Bigr) \\ &\le \frac{1}{\varepsilon}\|f-g\|_{L^1(\mu)} < \varepsilon.\end{aligned}$$

Finally, according to Chebyshev's inequality

$$\mu\bigl(x\colon\ |f(x) - g(x)| > \varepsilon\bigr) \le \frac{1}{\varepsilon}\|f-g\|_{L^1(\mu)} < \varepsilon.$$

Thus, (10.2.2) yields

$$\begin{aligned}\mu\bigl(x\colon\ \psi(x) > 2\varepsilon\bigr) \le \mu\Bigl(x\colon\ &\limsup_{n\to\infty} |\mathbb{E}^{\mathcal{B}_n}(f-g)(x)| > \varepsilon\Bigr) \\ &+ \mu\bigl(x\colon\ |f(x) - g(x)| > \varepsilon\bigr) \le 2\varepsilon,\end{aligned}$$

whence we conclude that $\psi = 0$ a.e., since ε is arbitrary. $\square$

10.2.4. Corollary. *Let $(X_n, \mathcal{A}_n, \mu_n)$, $n \in \mathbb{N}$, be probability spaces and let $(X, \mathcal{A}, \mu)$ be their product. For every function $f \in \mathcal{L}^1(\mu)$ and every $n \in \mathbb{N}$, let the function f_n be defined as follows: $f_n(x_1, \ldots, x_n)$ equals the integral*

$$\int_{\prod_{k=n+1}^{\infty} X_k} f(x_1, \ldots, x_n, x_{n+1}, \ldots) \bigotimes_{k=n+1}^{\infty} \mu_k \bigl(d(x_{n+1}, x_{n+2}, \ldots)\bigr).$$

Then, the functions f_n, regarded as functions on X, converge to f a.e. and in $L^1(\mu)$.

In the next section we discuss the results of this section in a broader context of the theory of martingales.

10.3. Martingales

The theory of martingales is one of many intersection points of measure theory and probability theory. We present here a number of basic results of the theory of martingales, but our illustrating examples are typical in the first place for measure theory: in the books on probability theory the same results are presented in their more natural environment of random walks, betting systems, and options. Following the tradition, we denote by $\mathbb{E}$ the expectation (integral) on a probability space.

10.3.1. Definition. *Let $(\Omega, \mathcal{F}, P)$ be a probability space. A sequence of functions $\xi_n \in \mathcal{L}^1(P)$, where $n = 0, 1, \ldots$, is called a martingale with respect to the sequence of σ-algebras $\mathcal{F}_n$ with $\mathcal{F}_n \subset \mathcal{F}_{n+1} \subset \mathcal{F}$ if the function ξ_n is measurable with respect to $\mathcal{F}_n$ and $\mathbb{E}^{\mathcal{F}_n}\xi_{n+1} = \xi_n$ a.e. for all $n \geq 0$.*

More generally, if T is a directed set and $\{\mathcal{F}_t\}$, where $t \in T$, is a family of σ-algebras in $\mathcal{F}$ with $\mathcal{F}_s \subset \mathcal{F}_t$ whenever $s < t$, then a family of functions $\xi_t \in \mathcal{L}^1(P)$ is called a martingale with respect to $\{\mathcal{F}_t\}$ if for every s the function ξ_s is measurable with respect to $\mathcal{F}_s$ and for every pair $t \geq s$ one has $\mathbb{E}^{\mathcal{F}_s}\xi_t = \xi_s$ a.e. (where a measure zero set may depend on t, s).

If a function $\xi_s \in \mathcal{L}^1(P)$ is measurable with respect to $\mathcal{F}_s$ and for every pair $t \geq s$ one has $\mathbb{E}^{\mathcal{F}_s}\xi_t \geq \xi_s$ a.e., then $\{\xi_t\}$ is called a submartingale with respect to $\{\mathcal{F}_t\}$, and if $\mathbb{E}^{\mathcal{F}_s}\xi_t \leq \xi_s$ a.e., then $\{\xi_t\}$ is called a supermartingale with respect to $\{\mathcal{F}_t\}$.

If $T = \{0, -1, -2, \ldots\}$, $\mathcal{F}_n \subset \mathcal{F}_{n+1}$, ξ_n is $\mathcal{F}_n$-measurable, and we have $\mathbb{E}^{\mathcal{F}_n}\xi_{n+1} = \xi_n$ a.e., then the sequence $\{\xi_n\}$ is called a reversed (or backward) martingale with respect to $\{\mathcal{F}_n\}$.

We draw the reader's attention to the fact that replacing a function ξ_n by an equivalent one may destroy the $\mathcal{F}_n$-measurability, so it has to be postulated separately.

A simple, but very important example of a martingale is the family $\mathbb{E}^{\mathcal{F}_n}\xi$, where $\xi \in L^1(P)$ and $\{\mathcal{F}_n\}$ is an increasing sequence of sub-σ-algebras in $\mathcal{F}$. This follows by the properties of conditional expectation.

In this section, we prove the basic theorems on convergence of martingales. These theorems are important in measure theory. The proofs employ an interesting estimate of the number of upcrossings of a fixed level. We start with this estimate (it is instructive to compare it with an analogous result in the section on ergodic theory). Let $\{\xi_n\}$, $n = 0, 1 \ldots$, be a submartingale with respect to $\{\mathcal{F}_n\}$ and $a < b$. Set $N_0 = -1$ and for $k \geq 1$ we set

$$N_{2k-1} = \inf\{m > N_{2k-2}\colon\ \xi_m \leq a\}, N_{2k} = \inf\{m > N_{2k-1}\colon\ \xi_m \geq b\},$$

where $N_i = +\infty$ if the corresponding set is empty. Note that between the moments N_{2k-1} and N_{2k} the sequence ξ_m is crossing $[a, b]$ upwards. Let $U_n = \sup\{k\colon\ N_{2k} \leq n\} I_{\{N_2 \leq n\}}$. We recall that $f^+ := \max(f, 0)$.

10.3.2. Lemma. *For every submartingale $\{\xi_n\}$, where $n = 0, 1, \ldots$, one has $(b-a)\mathbb{E}U_n \leq \mathbb{E}(\xi_n - a)^+ - \mathbb{E}(\xi_0 - a)^+$.*

Proof. Let $\eta_n = a + (\xi_n - a)^+$. According to Exercise 10.10.58, $\{\eta_n\}$ is a submartingale. Set $h_m = 1$ if $N_{2k-1} < m \leq N_{2k}$ for some $k \in \mathbb{N}$ and $h_m = 0$ otherwise. Then the function h_m is measurable with respect to $\mathcal{F}_{m-1}$. For any two sequences of functions $g = \{g_n\}$ and $\zeta = \{\zeta_n\}$, we set

$$[g, \zeta]_n := \sum_{m=1}^{n} g_m(\zeta_m - \zeta_{m-1}).$$

It is readily verified that $(b-a)U_n \leq [h, \eta]_n$. Let $g_m = 1 - h_m$. By Exercise 10.10.59 the sequence $[g, \eta]_n$ is a submartingale, whence one has $\mathbb{E}[g, \eta]_n \geq \mathbb{E}[g, \eta]_0 = 0$ and $(b-a)\mathbb{E}U_n \leq \mathbb{E}[h, \eta]_n \leq \mathbb{E}(\eta_n - \eta_0)$, as required. □

10.3.3. Theorem. *Let $\{\xi_n\}$, $n = 0, 1, \ldots$, be a submartingale. Suppose that $\sup_n \mathbb{E}(\xi_n^+) < \infty$. Then $\xi(\omega) = \lim\limits_{n\to\infty} \xi_n(\omega)$ exists a.e. and $\mathbb{E}|\xi| < \infty$.*

Proof. By Lemma 10.3.2, we obtain $\mathbb{E}U_n \leq (b-a)^{-1}\bigl(|a| + \mathbb{E}\xi_n^+\bigr)$ for any fixed a and b. Hence $\mathbb{E}\sup_n U_n < \infty$, which yields

$$P\bigl(\omega\colon\ \liminf_{n\to\infty} \xi_n(\omega) < a < b < \limsup_{n\to\infty} \xi_n(\omega)\bigr) = 0,$$

since otherwise on a set of positive measure we would have infinitely many upcrossings of $[a, b]$. The established fact is true for all rational a and b. Hence we obtain the existence of a limit $\xi = \lim\limits_{n\to\infty} \xi_n$ a.e. By Fatou's theorem, $\xi < +\infty$ a.e. and ξ^+ is integrable because $\sup_n \mathbb{E}\xi_n^+ < \infty$. On the other hand, $\mathbb{E}\min(\xi_n, 0) = \mathbb{E}\xi_n - \mathbb{E}\xi_n^+ \geq \mathbb{E}\xi_0 - \mathbb{E}\xi_n^+$, since $\{\xi_n\}$ is a submartingale, whence by Fatou's theorem we obtain the integrability of ξ^-. □

10.3.4. Corollary. *Let functions $\xi_n \geq 0$, where $n = 0, 1, \ldots$, form a supermartingale. Then a.e. there exists a finite limit $\xi(\omega) = \lim\limits_{n\to\infty} \xi_n(\omega)$ and one has $\mathbb{E}\xi \leq \mathbb{E}\xi_0$.*

Proof. The functions $\eta_n = -\xi_n$ form a submartingale and $\eta_n^+ = 0$. Hence our claim follows by the above theorem. □

The hypotheses of Theorem 10.3.3 do not guarantee convergence in L^1 (see Example 10.3.8 below).

The next example is a good illustration of the use of this theorem in measure theory.

10.3.5. Example. Let μ and ν be probability measures on a measurable space $(X, \mathcal{F})$, where $\mathcal{F}$ is generated by a sequence of sub-σ-algebras $\mathcal{F}_n$ with $\mathcal{F}_n \subset \mathcal{F}_{n+1}$. Denote by μ_n and ν_n the restrictions of μ and ν to $\mathcal{F}_n$ and assume that $\nu_n \ll \mu_n$ for all n. Let $\varrho_n = d\nu_n/d\mu_n$ and $\varrho = \limsup_{n\to\infty} \varrho_n$. Then $\{\varrho_n\}$ is a martingale on the probability space $(X, \mathcal{F}, \mu)$ and one has

$$\nu(B) = \int_B \varrho\, d\mu + \nu\bigl(B \cap \{\varrho = \infty\}\bigr), \quad \forall\, B \in \mathcal{F}. \tag{10.3.1}$$

Proof. Let us consider the probability measure $\gamma := (\mu + \nu)/2$ and denote by γ_n the restriction of γ to $\mathcal{F}_n$. It is clear that $\mu \ll \gamma$ and $\nu \ll \gamma$ and that the Radon–Nikodym densities $\varrho_n^\mu := d\mu_n/d\gamma_n$ and $\varrho_n^\nu := d\nu_n/d\gamma_n$ are majorized by 2. We observe that $\{\varrho_n^\mu\}$ and $\{\varrho_n^\nu\}$ are martingales with respect to the sequence $\{\mathcal{F}_n\}$ on the probability space $(X, \mathcal{F}, \gamma)$, since for all $A \in \mathcal{F}_n \subset \mathcal{F}_{n+1}$ we have

$$\int_A \varrho_{n+1}^\mu\, d\gamma = \mu_{n+1}(A) = \mu(A) = \mu_n(A) = \int_A \varrho_n^\mu\, d\gamma.$$

Certainly, the same calculation with ϱ_n and μ in place of ϱ_{n+1}^μ and γ shows that $\{\varrho_n\}$ is a martingale with respect to the measures μ. The verification for $\{\varrho_n^\nu\}$ is similar. Therefore, by the uniform boundedness, the following limits exist γ-a.e. and in $L^1(\gamma)$:

$$\varrho^\mu := \lim_{n\to\infty} \varrho_n^\mu, \quad \varrho^\nu := \lim_{n\to\infty} \varrho_n^\nu.$$

The functions ϱ^μ and ϱ^ν are the Radon–Nikodym densities of the measures μ and ν with respect to γ. Indeed, it is clear by the above relationships and convergence of ϱ_n^μ to ϱ^μ in $L^1(\gamma)$ that for every $A \in \mathcal{F}_n$ the integral of ϱ^μ with respect to the measure γ equals $\mu(A)$. Since the union of $\mathcal{F}_n$ is an algebra generating $\mathcal{F}$, we obtain the above claim.

We observe that γ-a.e. one has $\varrho_n = \varrho_n^\nu/\varrho_n^\mu$, where we set $\varrho_n^\nu(x)/\varrho_n^\mu(x) = 0$ if $\varrho_n^\mu(x) = 0$. This is clear from the equality $\varrho_n \varrho_n^\mu \cdot \gamma_n = \varrho_n^\nu \cdot \gamma_n$. Thus, for γ-a.e. x, there exists a limit $\varrho(x) := \lim_{n\to\infty} \varrho_n(x)$, possibly infinite. In fact, the set $Y := \{x\colon\ \varrho(x) = \infty\}$ has μ-measure zero. This follows by Corollary 10.3.4, since the same computation as above shows that the sequence $\{\varrho_n\}$ is a nonnegative martingale with respect to $\{\mathcal{F}_n\}$ on the probability space $(X, \mathcal{F}, \mu)$ and hence μ-a.e. has a finite limit. Let us show that the restriction of ν to $X\backslash Y$ is absolutely continuous with respect to μ. Let us set $S_N := \bigl\{x\colon\ \sup_n \varrho_n(x) \le N\bigr\}$. It suffices to verify that $\nu|_{S_N} \ll \mu$ for every fixed number $N \in \mathbb{N}$. Let $B \in \mathcal{F}$, $B \subset S_N$, and $\mu(B) = 0$. For fixed $\varepsilon > 0$, we find $B_m \in \mathcal{F}_m$ with $\gamma(B \triangle B_m) < \varepsilon/N$. Since

$$\nu(B_m) = \int_{B_m} \varrho_m \, d\mu \le N\mu(B_m) < 2\varepsilon,$$

one has $\nu(B) < 4\varepsilon$, which yields $\nu(B) = 0$. Thus, $\nu|_{X\setminus Y} \ll \mu$.

By the dominated convergence theorem $\nu = f \cdot \mu + \nu_0$, where $f \in L^1(\mu)$ and the measure ν_0 is mutually singular with μ. It is clear that $\nu_0 = \nu|_Y$. It remains to show that $f(x) = \varrho(x)$ for μ-a.e. x. Let $B \in \mathcal{F}$ and $B \subset S_N$. Then for every $x \in B$, the sequence $\varrho_n(x)$ is majorized by N and converges γ-a.e. to $\varrho(x)$, which on account of convergence of ϱ_n^μ to ϱ^μ in $L^1(\gamma)$ yields the following chain of equalities:

$$\begin{aligned}\int_B f \, d\mu = \nu(B) &= \lim_{n\to\infty} \int_B \varrho_n^\nu \, d\gamma \\ &= \lim_{n\to\infty} \int_B \frac{\varrho_n^\nu}{\varrho_n^\mu} \varrho_n^\mu \, d\gamma = \lim_{n\to\infty} \int_B \varrho_n \varrho_n^\mu \, d\gamma = \int_B \varrho\varrho^\mu \, d\gamma = \int_B \varrho \, d\mu.\end{aligned}$$

It follows that $f|_{S_N} = \varrho|_{S_N}$ μ-a.e., hence $f = \varrho$ μ-a.e. □

As an application we prove the following alternative of Kakutani [**935**].

10.3.6. Theorem. *Suppose that for every n we are given a measurable space $(X_n, \mathcal{B}_n)$ with two probability measures μ_n and ν_n such that $\nu_n \ll \mu_n$ and ϱ_n is the Radon–Nikodym density of ν_n with respect to μ_n. Set $\mu = \bigotimes_{n=1}^\infty \mu_n$, $\nu = \bigotimes_{n=1}^\infty \nu_n$. Then either $\nu \ll \mu$ or $\nu \perp \mu$, and the latter is equivalent to the equality*

$$\prod_{n=1}^\infty \int_{X_n} \sqrt{\varrho_n} \, d\mu_n := \lim_{N\to\infty} \prod_{n=1}^N \int_{X_n} \sqrt{\varrho_n} \, d\mu_n = 0.$$

PROOF. We observe that

$$\int_{X_n} \sqrt{\varrho_n} \, d\mu_n \le 1$$

by the Cauchy–Bunyakowsky inequality. Hence the corresponding infinite product either diverges to zero or converges to a number in $(0,1]$. For every n, we consider the σ-algebra $\mathcal{F}_n$ consisting of all sets of the form $B = B_n \times \prod_{i=n+1}^\infty X_i$, $B_n \in \bigotimes_{i=1}^n \mathcal{B}_i$. The functions $\xi_n(\omega) = \prod_{i=1}^n \varrho_i(\omega_i)$ on the probability space

$$(X, \mathcal{B}, \mu) := \Bigl(\prod_{i=1}^\infty X_i, \bigotimes_{i=1}^\infty \mathcal{B}_i, \mu\Bigr)$$

form a martingale with respect to the σ-algebras $\mathcal{F}_n$. Indeed, for every set $B \in \mathcal{F}_n$ of the indicated form we have

$$\int_B \xi_{n+1} \, d\mu = \Bigl(\bigotimes_{i=1}^{n+1} \nu_i\Bigr)(B_n \times X_{n+1}) = \Bigl(\bigotimes_{i=1}^{n} \nu_i\Bigr)(B_n) = \int_B \xi_n \, d\mu.$$

According to Corollary 10.3.4, there exists a μ-integrable limit $\xi = \lim_{n\to\infty} \xi_n$. In Example 10.3.5, we justified equality (10.3.1), which in the present case is applied to ξ in place of ϱ. If our product diverges to zero, then

$$\int_X \sqrt{\xi_n}\, d\mu \to 0,$$

whence by Fatou's theorem we obtain the equality

$$\int_X \sqrt{\xi}\, d\mu = 0,$$

i.e., $\xi = 0$ μ-a.e. and $\nu \perp \mu$. On the other hand, on account of the Cauchy–Bunyakowsky inequality, the estimate $|\sqrt{\xi_{n+k}} + \sqrt{\xi_n}|^2 \le 2\xi_{n+k} + 2\xi_n$, and the equalities

$$\sqrt{\xi_{n+k}\xi_n} = \varrho_1 \cdots \varrho_n \sqrt{\varrho_{n+1} \cdots \varrho_{n+k}} \quad \text{and} \quad \int_X \xi_m\, d\mu = 1,$$

we obtain

$$\begin{aligned}
&\int_X |\xi_{n+k} - \xi_n|\, d\mu \\
&\le \biggl(\int_X |\sqrt{\xi_{n+k}} - \sqrt{\xi_n}|^2\, d\mu\biggr)^{1/2} \biggl(\int_X |\sqrt{\xi_{n+k}} + \sqrt{\xi_n}|^2\, d\mu\biggr)^{1/2} \\
&\le \biggl(4\int_X |\sqrt{\xi_{n+k}} - \sqrt{\xi_n}|^2\, d\mu\biggr)^{1/2} = \biggl(8 - 8\prod_{i=n+1}^{n+k} \int_{X_i} \sqrt{\varrho_i}\, d\mu_i\biggr)^{1/2},
\end{aligned}$$

which in the case of convergence of the product to a positive number shows that $\{\xi_n\}$ in $L^1(\mu)$ is fundamental. Then for any fixed m and every $B \in \mathcal{F}_m$, we have

$$\nu(B) = \int_B \xi_m\, d\mu = \lim_{n\to\infty} \int_B \xi_n\, d\mu = \int_B \xi\, d\mu.$$

Therefore, $\nu \ll \mu$ and ξ is the Radon–Nikodym density of ν with respect to μ, which completes the proof. $\square$

10.3.7. Remark. Given a martingale $\{\xi_n\}$ with respect to increasing σ-algebras $\mathcal{F}_n$, the formula $\nu(A) = \mathbb{E}(\xi_n I_A)$, $A \in \mathcal{A}_n$, defines an additive set function on the algebra $\mathcal{R} := \bigcup_{n=1}^\infty \mathcal{A}_n$. Indeed, if also $A \in \mathcal{A}_k$ with some $k > n$, then $\xi_k I_A$ and $\xi_n I_A$ have equal expectations. Hence ν is well-defined. Clearly, ν is additive. However, it may fail to be countably additive as the following example shows. Note that the restriction of ν to $\mathcal{A}_n$ is a measure absolutely continuous with respect to the restriction of P. Conversely, for any additive function ν on $\mathcal{R}$ with the latter property, the Radon–Nikodym densities $\xi_n := d\nu|_{\mathcal{F}_n}/dP|_{\mathcal{F}_n}$ form a martingale.

10.3.8. Example. Let $\Omega = \mathbb{N}$ be equipped with the σ-algebra $\mathcal{F}$ of all subsets of Ω. By letting $P(\{n\}) = 2^{-n}$ for every $n \in \mathbb{N}$, we define a probability measure on Ω. Denote by $\mathcal{F}_n$ the finite sub-σ-algebra in $\mathcal{F}$ generated by the points $1, \ldots, n$ and the set $M_n := \{n+1, n+2, \ldots\}$. Finally, let us set

$\xi_n := P(M_n)^{-1} I_{M_n}$. Each ξ_n is a probability density with respect to P. The integral of ξ_k over a set $A \in \mathcal{A}_n$ with $n < k$ coincides with the integral of ξ_n over A. Indeed, both integrals vanish if $A \subset \{1, \ldots, n\}$ and equal 1 if $A = M_n$. Hence $\{\xi_n\}$ is a martingale with respect to $\{\mathcal{F}_n\}$. However, the additive function ν defined in the remark above is not countably additive, since $\nu(\{n\}) = 0$ for every n and $\nu(\mathbb{N}) = 1$. Note also that $\lim\limits_{n\to\infty} \xi_n(\omega) = 0$ pointwise, in particular, there is no convergence in $L^1(P)$.

Let us proceed to convergence of martingales in L^p. Let $(\Omega, \mathcal{F}, P)$ be a probability space equipped with a sequence of increasing σ-algebras $\mathcal{F}_n \subset \mathcal{F}$, $n = 0, 1, \ldots$. An $\mathcal{F}$-measurable function τ with values in the set of nonnegative integer numbers is called a stopping time if $\{\tau = n\} \in \mathcal{F}_n$ for all $n \in \mathbb{N} \cup \{0\}$.

10.3.9. Proposition. *Let $\{\xi_n\}$, where $n = 0, 1, \ldots$, be a submartingale and let τ be a stopping time such that $\tau \le k$ a.e. Then $\mathbb{E}\xi_0 \le \mathbb{E}\xi_\tau \le \mathbb{E}\xi_k$.*

PROOF. By Exercise 10.10.61 the sequence $\xi_{\min(\tau,n)}$ is a submartingale, whence $\mathbb{E}\xi_0 = \mathbb{E}\xi_{\min(\tau,0)} \le \mathbb{E}\xi_{\min(\tau,k)} = \mathbb{E}\xi_\tau$. For every $m \in \{0, 1, \ldots, k\}$ one has

$$\mathbb{E}(\xi_k I_{\{\tau=m\}}) = \mathbb{E}(\mathbb{E}^{\mathcal{F}_m} \xi_k I_{\{\tau=m\}}) \ge \mathbb{E}(\xi_m I_{\{\tau=m\}}) = \mathbb{E}(\xi_\tau I_{\{\tau=m\}}),$$

which yields the claim by summing in m. □

An immediate corollary of this result is the following inequality of Doob, the derivation of which from the proposition is left as Exercise 10.10.64.

10.3.10. Corollary. *Let $\{\xi_n\}$, where $n = 0, 1, \ldots$, be a submartingale and let*

$$X_n := \max_{0\le k\le n} \xi_k^+.$$

Then, for every $r > 0$, we have

$$rP\bigl(\{X_n \ge r\}\bigr) \le \int_{\{X_n\ge r\}} \xi_n^+ \, dP \le \mathbb{E}\xi_n^+.$$

10.3.11. Corollary. *Under the hypotheses of the previous corollary, for every $p > 1$ with $\xi_n^+ \in L^p(P)$, we have*

$$\mathbb{E}X_n^p \le \bigl(p/(p-1)\bigr)^p \mathbb{E}(\xi_n^+)^p.$$

If $\{\xi_n\}$ is a martingale and $\xi_n \in L^p(P)$, then

$$\mathbb{E}\bigl[\max_{0\le k\le n} |\xi_k|\bigr]^p \le \bigl(p/(p-1)\bigr)^p \mathbb{E}|\xi_n|^p.$$

PROOF. The second claim follows from the first one by passing to $|\xi_n|$. By Doob's inequality we obtain

$$\begin{aligned}\mathbb{E}X_n^p &= p\int_0^\infty r^{p-1} P\bigl(X_n \ge r\bigr)\, dr \le p\int_0^\infty r^{p-2} \int_{\{X_n\ge r\}} \xi_n^+ \, dP\, dr \\ &= p\int_\Omega \xi_n^+ \int_0^{X_n} r^{p-2}\, dr\, dP = \frac{p}{p-1}\int_\Omega \xi_n^+ X_n^{p-1}\, dP.\end{aligned}$$

Set $q = p/(p-1)$. By Hölder's inequality the right-hand side of the above inequality is estimated by $q\bigl(\mathbb{E}(\xi_n^+)^p\bigr)^{1/p}\bigl(\mathbb{E}X_n^p\bigr)^{1/q}$. This yields our claim. □

The boundedness of $\{\xi_n\}$ in $L^1(P)$ does not imply the boundedness of $\{X_n\}$ in $L^1(P)$: see Example 10.3.8 (but also see Exercise 10.10.65).

10.3.12. Theorem. *Let $\{\xi_n\}$, where $n = 0, 1, \ldots$, be a martingale such that $\sup_n \mathbb{E}|\xi_n|^p < \infty$, where $1 < p < \infty$. Then $\{\xi_n\}$ converges a.e. and in the space $L^p(P)$.*

PROOF. Almost everywhere convergence to some limit $\xi \in L^p(P)$ is clear from Theorem 10.3.3. It follows by Corollary 10.3.11 that $\sup_n |\xi_n| \in L^p(P)$. Since we have $|\xi_k - \xi|^p \le 2^p \sup_n |\xi_n|^p$, it remains to apply the dominated convergence theorem. □

Example 10.3.8 shows that the statement on convergence in L^p may be false for $p = 1$. In the case $p = 1$ the situation is this.

10.3.13. Theorem. *Let $\{\xi_n\}$ be a submartingale with respect to a sequence of σ-algebras $\mathcal{F}_n$, $n = 0, 1, \ldots$. The following conditions are equivalent:*
(i) *the sequence $\{\xi_n\}$ is uniformly integrable;*
(ii) *the sequence $\{\xi_n\}$ converges in $L^1(P)$;*
(iii) *the sequence $\{\xi_n\}$ converges a.e. and in $L^1(P)$.*
If $\{\xi_n\}$ is a martingale, then (i)–(iii) *are equivalent to the existence of a function $\xi \in \mathcal{L}^1(P)$ with $\xi_n = \mathbb{E}^{\mathcal{F}_n}\xi$ for all n. Then $\{\xi_n\}$ is called a closable martingale.*

PROOF. The uniform integrability implies boundedness in $L^1(P)$, which by Theorem 10.3.3 yields the existence of a limit $\xi = \lim\limits_{n\to\infty} \xi_n$ a.e. Then we also obtain convergence in $L^1(P)$. This shows that (i) yields (ii) and (ii) yields (iii). It is clear that (iii) implies (i). It remains to show that in the case of a martingale we have $\xi_n = \mathbb{E}^{\mathcal{F}_n}\xi$ (the fact that such a sequence is a martingale has already been noted and the uniform integrability follows by Example 10.1.10). If $B \in \mathcal{F}_n$, then we have $\mathbb{E}(\xi I_B) = \lim\limits_{k\to\infty} \mathbb{E}(\xi_k I_B)$. However, for all $k \ge n$ we have $\mathbb{E}(\xi_k I_B) = \mathbb{E}(\xi_n I_B)$, whence the desired equality follows. □

It follows from this theorem that the measure ν associated with the martingale $\{\xi_n\}$ in Remark 10.3.7 is countably additive and absolutely continuous with respect to P if and only if $\{\xi_n\}$ is closable. However, it may happen that ν is countably additive, but not absolutely continuous with respect to P (it suffices to take mutually singular measures μ and ν in Example 10.3.5). A necessary and sufficient condition for the countable additivity of ν is given in Exercise 10.10.62.

Let us derive Theorem 10.2.3 on convergence of conditional expectations from the martingale convergence theorem.

10.3.14. Example. Let $(X, \mathcal{F}, \mu)$ be a space with a finite nonnegative measure, let $\mathcal{F}_n$ be an increasing sequence of sub-σ-algebras in $\mathcal{F}$, and let $\mathcal{F}_\infty$ be the σ-algebra generated by $\{\mathcal{F}_n\}$. Then, for every function $f \in L^1(\mu)$, we have $\mathbb{E}^{\mathcal{F}_n} f \to \mathbb{E}^{\mathcal{F}_\infty} f$ a.e. and in $L^1(\mu)$. If $\varphi_n \to \varphi$ in $L^1(\mu)$, then $\mathbb{E}^{\mathcal{F}_n}\varphi_n \to \mathbb{E}^{\mathcal{F}_\infty}\varphi$ in $L^1(\mu)$.

PROOF. We may assume that $\mathcal{F}_\infty = \mathcal{F}$. Then we have $\mathbb{E}^{\mathcal{F}_\infty} f = f$. The sequence $f_n := \mathbb{E}^{\mathcal{F}_n} f$ is a uniformly integrable martingale. Hence it converges a.e. and in $L^1(\mu)$ to some function g. We show that $f = g$ a.e. It suffices to show that f and g have equal integrals over every set $B \in \mathcal{F}_n$. The integral of fI_B equals the integral of $f_m I_B$ for all $m \geq n$, which coincides with the integral of gI_B. The last claim is obvious from the fact that for all n we have $\|\mathbb{E}^{\mathcal{F}_n}\psi\|_{L^1(\mu)} \leq \|\psi\|_{L^1(\mu)}$. □

10.3.15. Example. If $A \in \mathcal{F}_\infty$, then $\mathbb{E}^{\mathcal{F}_n} I_A \to I_A$ a.e.

Finally, let us consider reversed martingales.

10.3.16. Theorem. *Let $\{\xi_n\}$ be a reversed martingale with respect to $\{\mathcal{F}_n\}$, $n = 0, -1, \ldots$. Then $\xi_{-\infty} := \lim\limits_{n\to-\infty} \xi_n$ exists a.e. and in $L^1(P)$. In addition, one has $\xi_{-\infty} = \mathbb{E}^{\mathcal{F}_{-\infty}}\xi_0$, where $\mathcal{F}_{-\infty} = \bigcap_{n\leq 0} \mathcal{F}_n$.*

PROOF. As in the case of a direct martingale, for every fixed a and b, we denote by U_n the number of upcrossings of $[a, b]$ by $\xi_{-|n|}, \ldots, \xi_0$. By using Lemma 10.3.2, we obtain

$$(b - a)\mathbb{E}U_n \leq \mathbb{E}(\xi_0 - a)^+.$$

Similarly to the reasoning in Theorem 10.3.3 this yields the existence of a limit $\xi_{-\infty} = \lim\limits_{n\to-\infty} \xi_n$ a.e. However, in the present case, the sequence $\{\xi_n\}$ is at once uniformly integrable, since $\xi_n = \mathbb{E}^{\mathcal{F}_n}\xi_0$ for all $n = 0, -1, \ldots$. This ensures mean convergence. It is clear that the function $\xi_{-\infty}$ is measurable with respect to $\mathcal{F}_{-\infty}$. Given $A \in \mathcal{F}_{-\infty}$, we have $\mathbb{E}(I_A\xi_0) = \mathbb{E}(I_A\xi_n) \to \mathbb{E}(I_A\xi_{-\infty})$, whence we obtain the last assertion. □

10.3.17. Corollary. *Suppose that $(X, \mathcal{F}, \mu)$ is a probability space and that $\{\mathcal{F}_n\}_{n\in\{0,-1,\ldots\}}$ is a sequence of sub-σ-algebras in $\mathcal{F}$ with $\mathcal{F}_{n-1} \subset \mathcal{F}_n$ for all n. Set*

$$\mathcal{F}_{-\infty} = \bigcap_{n\leq 0} \mathcal{F}_n.$$

Then, for every function $f \in \mathcal{L}^1(\mu)$, one has $\mathbb{E}^{\mathcal{F}_n} f \to \mathbb{E}^{\mathcal{F}_{-\infty}} f$ a.e. and in $L^1(\mu)$.

PROOF. The sequence $\xi_n = \mathbb{E}^{\mathcal{F}_n} f$ is a reversed martingale. As shown above, it converges a.e. and in $L^1(\mu)$ to $\xi_{-\infty} = \mathbb{E}^{\mathcal{F}_{-\infty}}\mathbb{E}^{\mathcal{F}_0} f = \mathbb{E}^{\mathcal{F}_{-\infty}} f$. □

Reversed martingales can be applied to convergence of the Riemann sums.

10.3.18. Example. Let f be a function integrable on $[0,1)$ and defined on the whole real line periodically with a period 1. For every $n \in \mathbb{N}$, we define a function F_n by

$$F_n(x) = 2^{-n} \sum_{j=0}^{2^n-1} f(j2^{-n} + x).$$

Then, for almost all $x \in [0,1]$, we have

$$\lim_{n\to\infty} F_n(x) = \int_0^1 f(t)\, dt.$$

PROOF. By Example 10.1.4, F_k is the conditional expectation of f with respect to the σ-algebra $\mathcal{B}_k$ generated by 2^{-k}-periodic functions. Clearly, $\mathcal{B}_{k+1} \subset \mathcal{B}_k$. According to Exercise 5.8.109 only constants are measurable with respect to the σ-algebra $\bigcap_{k\ge 1} \mathcal{B}_k$. It remains to apply the above corollary with $\mathcal{F}_n = \mathcal{B}_{-n}$. $\square$

Finally, let us mention the following interesting fact.

10.3.19. Proposition. *Let $\{\xi_n\}$, $n \in \mathbb{N}$, be a supermartingale with respect to an increasing sequence of σ-algebras $\mathcal{F}_n$. Then one can find a martingale $\{\eta_n\}$ and an increasing process $\{\zeta_n\}$ such that $\xi_n = \eta_n - \zeta_n$.*

PROOF. Let $\alpha_k := \mathbb{E}(\xi_k - \xi_{k+1}|\mathcal{F}_k)$. Since $\{\xi_n\}$ is a supermartingale, one has $\alpha_k \ge 0$. Let $\zeta_n := \sum_{k=1}^{n-1} \alpha_k$. Then $\zeta_{n+1} \ge \zeta_n$. It is easy to verify that the sequence $\xi_n + \zeta_n$ is a martingale. $\square$

The decomposition obtained above (called the Doob decomposition) is a special case of the Doob–Meyer decomposition for supermartingales $\{\xi_t\}$, $t \ge 0$, satisfying certain mild assumptions (see Dellacherie [**424**, Ch. IV]).

10.4. Regular conditional measures

We have already encountered the concept of conditional measure in §10.1. We have discussed there the following situation. Let μ be a measure on a measurable space $(X,\mathcal{A})$ and let $\mathcal{B}$ be a sub-σ-algebra in $\mathcal{A}$. We may assume that $\mathcal{B}$ is generated by a measurable mapping π from X to some measurable space $(Y,\mathcal{E})$. One can take $(Y,\mathcal{E}) = (X,\mathcal{B})$ with the identity embedding π.

As we know, in the case of a nonnegative measure μ, for every $A \in \mathcal{A}$, there exists a $\mathcal{B}$-measurable function $\mu(A, \cdot\,)$ such that

$$\mu(A\cap B) = \int_B \mu(A,x)\, \mu(dx), \quad A \in \mathcal{A}, B \in \mathcal{B}.$$

By the $\mathcal{B}$-measurability of the function $x \mapsto \mu(A,x)$ and Theorem 2.12.3, there exists an $\mathcal{E}$-measurable function $y \mapsto \mu^y(A)$ on Y with $\mu(A,x) = \mu^{\pi(x)}(A)$. Letting $\nu := \mu\circ\pi^{-1}$, this formula can be written as follows:

$$\mu\bigl(A\cap \pi^{-1}(E)\bigr) = \int_E \mu^y(A)\, \nu(dy), \quad E \in \mathcal{E}.$$

In particular, letting $E = Y$ we obtain

$$\mu(A) = \int_Y \mu^y(A)\,\nu(dy), \quad A \in \mathcal{A}.$$

Thus, if μ^y is a measure on $\pi^{-1}(y)$ for every $y \in Y$, then the previous equality is a generalized Fubini-type theorem: in order to find the measure of A, one has to compute the conditional measures of A on the level sets $\pi^{-1}(y)$ and then integrate in y with respect to the measure ν.

However, as we shall see below, it is not always the case that for μ-almost all x the set function $\mu(A, x)$ (or $\mu^y(A)$ for ν-almost all y) is a countably additive measure. Nevertheless, this becomes possible under some additional conditions of set-theoretic or topological character.

10.4.1. Definition. *Suppose we are given a σ-algebra $\mathcal{A}$, its sub-σ-algebra $\mathcal{B}$, and a measure μ on $\mathcal{A}$. We shall say that a function*

$$\mu^{\mathcal{B}}(\,\cdot\,,\,\cdot\,)\colon\ \mathcal{A}\times X \to \mathbb{R}^1$$

is a regular conditional measure on $\mathcal{A}$ with respect to $\mathcal{B}$ if:

(1) *for every x, the function $A \mapsto \mu^{\mathcal{B}}(A, x)$ is a measure on $\mathcal{A}$;*

(2) *for every $A \in \mathcal{A}$, the function $x \mapsto \mu^{\mathcal{B}}(A, x)$ is measurable with respect to $\mathcal{B}$ and $|\mu|$-integrable;*

(3) *one has*

$$\mu(A \cap B) = \int_B \mu^{\mathcal{B}}(A, x)\,|\mu|(dx), \quad \forall\ A \in \mathcal{A},\ B \in \mathcal{B}. \tag{10.4.1}$$

In the cases where there is no risk of ambiguity the shortened notation $\mu(A, x)$ is used. An alternative notation for the same objects: $\mu^{\mathcal{B}}(A|x)$ and $\mu(A|x)$. The measures $A \mapsto \mu^{\mathcal{B}}(A, x)$ also are called regular conditional measures (to distinguish the individual measures $\mu^{\mathcal{B}}(\,\cdot\,, x)$ and the whole function $\mu^{\mathcal{B}}(\,\cdot\,,\,\cdot\,)$, the latter is sometimes called a system of conditional measures).

The term "regular conditional measure" is used in order to avoid confusion with the conditional probabilities in the sense of conditional expectations (which are not always countably additive). However, in the cases where there is no risk of confusion we shall omit the word "regular" for brevity.

If $x \mapsto \|\mu^{\mathcal{B}}(\,\cdot\,, x)\|$ is $|\mu|$-integrable (which is not always the case), equality (10.4.1) can be written in the following integral form: for every bounded $\mathcal{A}$-measurable function f and every $B \in \mathcal{B}$, one has

$$\int_B f(x)\,\mu(dx) = \int_B \int_X f(y)\,\mu^{\mathcal{B}}(dy, x)\,|\mu|(dx). \tag{10.4.2}$$

Indeed, for the indicators of sets in $\mathcal{A}$ this coincides with (10.4.1). Hence equality (10.4.2) holds for simple functions, which by means of uniform approximations enables us to extend it to all bounded $\mathcal{A}$-measurable functions. If the measures μ and $\mu^{\mathcal{B}}(\,\cdot\,, x)$ are nonnegative, then (10.4.2) extends to all $\mathcal{A}$-measurable μ-integrable functions f. Indeed, for nonnegative f, we consider the functions $f_n = \min(f, n)$. By the previous step, the integrals of the functions

$$x \mapsto \int_X f_n(y)\, \mu^{\mathcal{B}}(dy, x)$$

are uniformly bounded. By Fatou's theorem, for μ-a.e. x, the function f is integrable against $\mu^{\mathcal{B}}(dy, x)$. It remains to apply the monotone convergence theorem. In the general case, we consider separately f^+ and f^-.

In the situation where the σ-algebra $\mathcal{B}$ is generated by a measurable mapping $\pi\colon (X, \mathcal{A}) \to (Y, \mathcal{E})$, it is more convenient to parameterize conditional measures by points of the space Y.

10.4.2. Definition. *A system of regular conditional measures μ^y, $y \in Y$, generated by an $(\mathcal{A}, \mathcal{E})$-measurable mapping $\pi\colon X \to Y$ is defined as a function $(A, y) \mapsto \mu^y(A)$ on $\mathcal{A}\times Y$ such that, for every fixed y, it is a measure on $\mathcal{A}$, for every fixed $A \in \mathcal{A}$ is measurable with respect to $\mathcal{E}$ and $|\mu| \circ \pi^{-1}$-integrable, and for all $A \in \mathcal{A}$ and $E \in \mathcal{E}$ satisfies the equality*

$$\mu\Bigl(A \cap \pi^{-1}(E)\Bigr) = \int_E \mu^y(A)\, |\mu| \circ \pi^{-1}(dy). \tag{10.4.3}$$

If for $|\mu| \circ \pi^{-1}$-almost every point $y \in Y$ we have $\pi^{-1}(y) \in \mathcal{A}$ and the measure μ^y is concentrated on $\pi^{-1}(y)$, then we shall call μ^y proper conditional measures.

Sometimes the following more general definition of conditional measures is useful. Let $\mathcal{A}_0$ be a sub-σ-algebra in $\mathcal{A}$ (not necessarily containing $\mathcal{B}$). Then the conditional measures $\mu^{\mathcal{B}}_{\mathcal{A}_0}(A, x)$ on $\mathcal{A}_0$ with respect to $\mathcal{B}$ are defined as above, but with $\mathcal{A}_0$ in place of $\mathcal{A}$ in conditions (1)–(3). In particular, now in place of (10.4.1) we require the equality

$$\mu(A \cap B) = \int_B \mu^{\mathcal{B}}_{\mathcal{A}_0}(A, x)\, |\mu|(dx), \quad \forall\ A \in \mathcal{A}_0,\ B \in \mathcal{B}. \tag{10.4.4}$$

In a similar manner one defines regular conditional measures $\mu^y_{\mathcal{A}_0}$ on $\mathcal{A}_0$ in the case where $\mathcal{B}$ is generated by a mapping π.

10.4.3. Lemma. *Let $\mathcal{A}$ be countably generated. Then regular conditional measures are essentially unique: given two regular conditional measures $\mu^{\mathcal{B}}_1(\,\cdot\,,\,\cdot\,)$ and $\mu^{\mathcal{B}}_2(\,\cdot\,,\,\cdot\,)$ on $\mathcal{A}$, there exists a set $Z \in \mathcal{B}$ with $|\mu|(Z) = 0$ such that $\mu^{\mathcal{B}}_1(A, x) = \mu^{\mathcal{B}}_2(A, x)$ for all $A \in \mathcal{A}$ and $x \in X\backslash Z$. Similarly, the measures $\mu^{\mathcal{B}}_{\mathcal{A}_0}(\,\cdot\,, x)$ on $\mathcal{A}_0$ are essentially unique if $\mathcal{A}_0$ is countably generated (the whole σ-algebra $\mathcal{A}$ need not be countably generated in this case).*

If, in addition, μ is a probability measure, then $\mu^{\mathcal{B}}(\,\cdot\,, x)$ is a probability measure for μ-a.e. x.

Proof. There is a countable algebra $\mathcal{R} = \{A_n\}$ generating $\mathcal{A}$. By equality (10.4.1), for every $A_n \in \mathcal{R}$, there is a set $Z_n \in \mathcal{B}$ such that $|\mu|(Z_n) = 0$ and $\mu^{\mathcal{B}}_1(A_n, x) = \mu^{\mathcal{B}}_2(A_n, x)$ for all $x \in X\backslash Z_n$. Now we take $Z := \bigcup_{n=1}^{\infty} Z_n$. The case of $\mathcal{A}_0$ is similar.

If μ is a probability measure, then, for every n, the function $\mu^{\mathcal{B}}(A_n, x)$ is nonnegative μ-a.e. because its integral over every $B \in \mathcal{B}$ is nonnegative. Similarly, $\mu^{\mathcal{B}}(X, x) = 1$ for μ-a.e. x. Hence for μ-a.e. x, the measure $\mu^{\mathcal{B}}(\,\cdot\,, x)$ is nonnegative on $\mathcal{R}$ and $\mu^{\mathcal{B}}(X, x) = 1$, which yields that $\mu^{\mathcal{B}}(\,\cdot\,, x)$ is a probability measure for such x. □

For an arbitrary σ-algebra $\mathcal{A}$, both assertions may be false even if μ is separable (see Exercise 10.10.44 for a simple counterexample).

10.4.4. Remark. (i) We observe that if a signed measure μ possesses regular conditional measures $\mu^{\mathcal{B}}(\,\cdot\,, x)$ and $X = X^+ \cup X^-$ is the Hahn decomposition for μ, then the measures $|\mu|^{\mathcal{B}}(\,\cdot\,, x) := \mu^{\mathcal{B}}(\,\cdot \cap X^+, x) - \mu^{\mathcal{B}}(\,\cdot \cap X^-, x)$ serve as regular conditional measures for $|\mu|$. Conversely, given regular conditional measures $|\mu|^{\mathcal{B}}(\,\cdot\,, x)$ (possibly, signed) for $|\mu|$, we obtain regular conditional measures $|\mu|^{\mathcal{B}}(\,\cdot \cap X^+, x)$ and $|\mu|^{\mathcal{B}}(\,\cdot \cap X^-, x)$ for μ^+ and μ^-, respectively. Hence μ has regular conditional measures $|\mu|^{\mathcal{B}}(\,\cdot \cap X^+, x) - |\mu|^{\mathcal{B}}(\,\cdot \cap X^-, x)$.

(ii) Let μ be a probability measure such that there exist probability measures $A \mapsto \nu(A, x)$, $x \in X$, on $\mathcal{A}$ and a probability measure σ on $\mathcal{B}$ satisfying the equality

$$\mu(A \cap B) = \int_B \nu(A, x)\, \sigma(dx)$$

for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$, where the functions $x \mapsto \nu(A, x)$ are measurable with respect to $\mathcal{B}$. Letting $A = X$, we see that σ coincides with the restriction of μ to $\mathcal{B}$, i.e., we obtain regular conditional measures. If the measures $\nu(\,\cdot\,, x)$ and σ are nonnegative, but not necessarily normalized, then the function $\theta(x) = \nu(X, x)$ is $\mathcal{B}$-measurable, and the measure $\sigma_0 := \theta \cdot \sigma$ is probabilistic. Replacing $\nu(\,\cdot\,, x)$ by the probability measure $\nu_0(\,\cdot\,, x) = \theta(x)^{-1}\nu(\,\cdot\,, x)$ for all x with $\theta(x) > 0$, we arrive at the previous case.

10.4.5. Theorem. (i) *Suppose that the σ-algebra $\mathcal{A}$ is countably generated and that μ has a compact approximating class in $\mathcal{A}$. Then, for every sub-σ-algebra $\mathcal{B} \subset \mathcal{A}$, there exists a regular conditional measure $\mu^{\mathcal{B}}$ on $\mathcal{A}$.*

(ii) *More generally, let $\mathcal{A}_0$ be a sub-σ-algebra in the σ-algebra $\mathcal{A}$ such that there exists a countable algebra $\mathcal{U}$ generating $\mathcal{A}_0$. Assume, additionally, that there is a compact class $\mathcal{K}$ such that for every $A \in \mathcal{U}$ and $\varepsilon > 0$, there exist $K_\varepsilon \in \mathcal{K}$ and $A_\varepsilon \in \mathcal{A}$ with $A_\varepsilon \subset K_\varepsilon \subset A$ and $|\mu|(A \backslash A_\varepsilon) < \varepsilon$. Then, for every sub-$\sigma$-algebra $\mathcal{B} \subset \mathcal{A}$, there exists a regular conditional measure $\mu^{\mathcal{B}}_{\mathcal{A}_0}$ on $\mathcal{A}_0$ (a probability if μ is nonnegative).*

In addition, for every $\mathcal{A}_0$-measurable μ-integrable function f, one has

$$\int_X f\, d\mu = \int_X \int_X f(y)\, \mu^{\mathcal{B}}_{\mathcal{A}_0}(dy, x)\, |\mu|(dx). \tag{10.4.5}$$

PROOF. Let us consider first the case of a probability measure.

(1) We shall prove the more general second assertion. Let $\mathcal{U}$ consist of countably many sets A_n. For every n, we find sets $C_{n,k} \in \mathcal{K}$ and $A_{n,k} \in \mathcal{A}$, $k \in \mathbb{N}$, such that

$$A_{n,k} \subset C_{n,k} \subset A_n \quad \text{and} \quad \mu(A_n \backslash A_{n,k}) < 1/k. \tag{10.4.6}$$

The sets $A_{n,k}$ along with the sets A_n generate a countable algebra $\mathcal{U}_0 \subset \mathcal{A}$. By the Radon–Nikodym theorem, for every set $A \in \mathcal{U}_0$, there exists a nonnegative $\mathcal{B}$-measurable function $x \mapsto p_0(A,x)$ such that $p_0(X,x) = 1$, $p_0(\varnothing,x) = 0$ for all x, and

$$\mu(A \cap B) = \int_B p_0(A,x)\,\mu(dx), \quad \forall\, B \in \mathcal{B}. \tag{10.4.7}$$

We observe that there exists a measure zero set $N_0 \in \mathcal{B}$ such that for all $x \in X \backslash N_0$, the function $A \mapsto p_0(A,x)$ is additive on $\mathcal{U}_0$. Indeed, it follows by (10.4.7) that $p_0(A \cup B, x) = p_0(A,x) + p_0(B,x)$ μ-a.e. whenever $A, B \in \mathcal{U}_0$ and $A \cap B = \varnothing$. Since the set of pairs (A,B), where $A, B \in \mathcal{U}_0$, is countable, the union of all sets on which the indicated equality fails for some pair of sets in $\mathcal{U}_0$ has measure zero.

(2) We now prove that for a.e. x one has

$$p_0(A_n, x) = \sup_k p_0(A_{n,k}, x), \quad \forall\, n \in \mathbb{N}. \tag{10.4.8}$$

In particular, for such x, the set function $p_0(\,\cdot\,,x)$ is approximated on the algebra $\mathcal{U}$ by the class $\mathcal{K}$ with respect to the algebra $\mathcal{U}_0$ (see Remark 1.4.7). We denote the right-hand side of (10.4.8) by $q_n(x)$. It is clear that the function q_n is measurable with respect to $\mathcal{B}$. The inclusion $A_{n,k} \subset A_n$ yields that there exist measure zero sets $N_{n,k} \in \mathcal{B}$ such that $p_0(A_{n,k},x) \le p_0(A_n,x)$ for all $x \notin N_{n,k}$. Hence

$$q_n(x) \le p_0(A_n,x), \quad \forall\, x \notin N_0 := \textstyle\bigcup_{n,k=1}^\infty N_{n,k}.$$

On the other hand, the obvious inequality $p_0(A_{n,k},x) \le q_n(x)$ yields that

$$\mu(A_{n,k}) = \int_X p_0(A_{n,k},x)\,\mu(dx) \le \int_X q_n(x)\,\mu(dx),$$

whence on account of the equality $\mu\Bigl(\bigcup_{n,k=1}^\infty N_{n,k}\Bigr) = 0$ we obtain

$$\sup_k \mu(A_{n,k}) \le \int_X q_n(x)\,\mu(dx) \le \int_X p_0(A_n,x)\,\mu(dx) = \mu(A_n).$$

Since the left-hand side equals $\mu(A_n)$, we have $q_n(x) = p_0(A_n,x)$ everywhere, with the exception of some measure zero set $N_1 \in \mathcal{B}$.

(3) According to steps (1) and (2), for all $x \notin N := N_0 \cup N_1$ the additive set function $p_0(\,\cdot\,,x)$ on the algebra $\mathcal{U}_0$ has the property that the compact class $\mathcal{K}$ approximates $p_0(\,\cdot\,,x)$ on $\mathcal{U}$ with respect to $\mathcal{U}_0$. By Remark 1.4.7, this set function is countably additive on $\mathcal{U}$ and extends uniquely to a countably additive measure on $\mathcal{A}_0$, which we take for $\mu(\,\cdot\,,x) = \mu^{\mathcal{B}}_{\mathcal{A}_0}(\,\cdot\,,x)$. It is clear that we obtain a probability measure. Finally, for all $x \in N$ let $\mu(\,\cdot\,,x) = \mu$.

(4) Let us verify that we have obtained the required conditional measures. Indeed, if $A = A_n$, then the function $x \mapsto \mu(A,x)$ is measurable with respect to $\mathcal{B}$. The class of all sets $A \in \sigma(\mathcal{U})$ for which this is true is monotone. Hence it coincides with $\sigma(\mathcal{U})$. Further, if $B \in \mathcal{B}$ and $A = A_n$, then by construction one has (10.4.7). Let $B \in \mathcal{B}$ be fixed. The class $\mathcal{E}$ of all sets $A \in \sigma(\mathcal{U})$

such that (10.4.7) holds is monotone: if sets $E_j \in \mathcal{E}$ are increasing to E, then $\lim\limits_{j\to\infty} \mu(E_j, x) = \mu(E, x)$, whence by the dominated convergence theorem we obtain the inclusion $E \in \mathcal{E}$. Therefore, $\mathcal{E} = \sigma(\mathcal{U})$. Since $\sigma(\mathcal{U}) = \mathcal{A}_0$ by hypothesis, we arrive at (10.4.4).

(5) It suffices to obtain equality (10.4.5) for the indicators of sets in $\mathcal{A}_0$, but in this case it is true by definition.

If μ is nonnegative, but is not probabilistic, the conditional probability measures for $\mu/\|\mu\|$ are conditional measures for μ as well (if $\mu = 0$ and X is not empty, then one can take a fixed Dirac measure for conditional measures). Finally, conditional measures for a signed measure μ are constructed as the differences of the conditional measures for μ^+ and μ^- in the following way. For the measure μ^+ we take probability conditional measures $\mu_1(\,\cdot\,, x)$, $x \in X^+$, concentrated on X^+; for the measure μ^- we take probability conditional measures $\mu_2(\,\cdot\,, x)$, $x \in X^-$, concentrated on X^-. Let $\mu_1(\,\cdot\,, x) = 0$ if $x \in X^-$, $\mu_2(\,\cdot\,, x) = 0$ if $x \in X^+$. Then the measures $\mu(\,\cdot\,, x) := \mu_1(\,\cdot\,, x) - \mu_2(\,\cdot\,, x)$ are conditional for μ and one has $\|\mu(\,\cdot\,, x)\| = 1$ (moreover, either $\mu(\cdot,, x)$ is a probability measure or $-\mu(\cdot,, x)$ is a probability measure). □

Let us note that by construction we have $\|\mu^{\mathcal{B}}_{\mathcal{A}_0}(\,\cdot\,, x)\| = 1$.

We now give the major special case for applications.

10.4.6. Corollary. *Let μ be a Borel measure on a Souslin space X. Then, for every sub-σ-algebra $\mathcal{B} \subset \mathcal{B}(X)$, there exists a regular conditional measure $\mu^{\mathcal{B}}$ on $\mathcal{B}(X)$.*

PROOF. It suffices to recall that the measure μ is Radon and $\mathcal{B}(X)$ is countably generated. □

10.4.7. Corollary. *Let X be a Hausdorff space and let μ be a Radon measure on X. Then, for every sub-σ-algebra $\mathcal{B} \subset \mathcal{B}(X)$ and every countably generated sub-σ-algebra $\mathcal{A}_0 \subset \mathcal{B}(X)$, there exists a regular conditional measure $\mu^{\mathcal{B}}_{\mathcal{A}_0}$ on $\mathcal{A}_0$.*

PROOF. Let $\mathcal{A} = \mathcal{B}(X)$ and take for $\mathcal{K}$ the class of all compact sets in the space X. □

Let us represent the obtained results in terms of a mapping π generating the σ-algebra $\mathcal{B}$.

10.4.8. Theorem. *Let μ be a measure (possibly signed) on a measurable space $(X, \mathcal{A})$, let $(Y, \mathcal{E})$ be a measurable space, and let $\pi\colon (X, \mathcal{A}) \to (Y, \mathcal{E})$ be a mapping measurable with respect to $(\mathcal{A}_\mu, \mathcal{E})$. Suppose that $\pi(X) \in \mathcal{E}_{|\mu|\circ\pi^{-1}}$, where $\mathcal{E}_{|\mu|\circ\pi^{-1}}$ is the completion of $\mathcal{E}$ with respect to the measure $|\mu| \circ \pi^{-1}$. Assume also that $\mathcal{A}$ is countably generated and that the measure μ has a compact approximating class. Then, there exist regular conditional measures μ^y, $y \in Y$, generated by π on $\mathcal{A}$ (probabilities if $\mu \geq 0$).*

More generally, π generates regular conditional measures $\mu^y_{\mathcal{A}_0}$, $y \in Y$, on every countably generated σ-algebra $\mathcal{A}_0 \subset \mathcal{A}$ on which μ possesses a compact approximating class.

PROOF. Set $\mathcal{B} := \pi^{-1}(\mathcal{E})$. By Theorem 10.4.5 (assertion (ii) is applicable with $\mathcal{A}_0 = \mathcal{A}$ and $\mathcal{A} = \mathcal{A}_\mu$), there is a conditional measure $\mu^{\mathcal{B}}(A, x)$ on $\mathcal{A}$ such that the function $\mu^{\mathcal{B}}(A, x)$ is measurable with respect to $\mathcal{B}$. This means that for every $A \in \mathcal{A}$, one has an $\mathcal{E}$-measurable function $\eta(A, y)$ such that $\mu^{\mathcal{B}}(A, x) = \eta\bigl(A, \pi(x)\bigr)$. By hypothesis, there exists $Y_0 \in \mathcal{E}$ with $Y_0 \subset \pi(X)$ and $|\mu| \circ \pi^{-1}(Y \backslash Y_0) = 0$. It is clear that for each $y \in Y_0$, $\eta(A, y)$ is a measure as a function of A. We take this measure for μ^y. If $y \notin Y_0$, then let $\mu^y = \mu$. For every $A \in \mathcal{A}$, the function $\mu^y(A)$ is $\mathcal{E}$-measurable, since $Y \backslash Y_0 \in \mathcal{E}$. The same reasoning proves the last assertion. □

10.4.9. Remark. Suppose that in the situation of Theorem 10.4.8 we have $\mathcal{A}_0 := \xi^{-1}(\mathcal{F})$, where ξ is a mapping from X to a measurable space $(Z, \mathcal{F})$, $\nu = |\mu| \circ \xi^{-1}$, $\xi(X) \in \mathcal{F}_\nu$. Then for the existence of regular conditional measures $\mu^y_{\mathcal{A}_0}$ generated by π on $\mathcal{A}_0$ the following conditions are sufficient: $\mathcal{F}$ is countably generated and the measure ν on $\mathcal{F}$ (or on $\mathcal{F}_\nu$) has a compact approximating class. This follows from the fact that the σ-algebra $\xi^{-1}(\mathcal{F})$ is countably generated and the measure μ on $\mathcal{A}_0$ has a compact approximating class according to Exercise 9.12.40.

The use of the measure $|\mu| \circ \pi^{-1}$ in the case of a signed measure μ is absolutely natural because the measure $\mu \circ \pi^{-1}$ may be identically zero for a nonzero measure μ. We note that the measures μ^y constructed above may not be concentrated on the sets $\pi^{-1}(y)$ (which may not be even measurable). Let us give a sufficient condition of the existence of proper conditional measures.

10.4.10. Corollary. *Suppose that in Theorem* 10.4.8 *the σ-algebra $\mathcal{E}$ is countably generated and contains all singletons. Then, there exist regular conditional measures μ^y on the σ-algebra $\mathcal{A}'$ generated by $\mathcal{A}$ and $\pi^{-1}(\mathcal{E})$ such that, for $|\mu| \circ \pi^{-1}$-a.e. y, the measure μ^y is concentrated on the set $\pi^{-1}(y)$. If π has an $(\mathcal{A}, \mathcal{E})$-measurable version $\widetilde{\pi}$ such that $\widetilde{\pi}(\mathcal{A}) \subset \mathcal{E}_{|\mu| \circ f^{-1}}$, then $\pi^{-1}(y) \in \mathcal{A}_{\mu^y}$ for $|\mu| \circ \pi^{-1}$-a.e. y.*

PROOF. It suffices to consider only probability measures. Let $\nu = \mu \circ \pi^{-1}$. Under our assumptions one has $\pi^{-1}(y) \in \mathcal{A}'$. There exists a countable algebra of sets $\mathcal{E}_0 = \{E_n\}$ generating $\mathcal{E}$. It is clear that $\mathcal{A}'$ is countably generated as well. We know that there exist regular conditional measures μ^y, $y \in Y$, on $\mathcal{A}'$. Let us fix $E_n \in \mathcal{E}_0$. For every $E \in \mathcal{E}$ one has

$$\int_E \mu^y\bigl(\pi^{-1}(E_n)\bigr)\,\nu(dy) = \mu\bigl(\pi^{-1}(E) \cap \pi^{-1}(E_n)\bigr) = \mu\bigl(\pi^{-1}(E \cap E_n)\bigr)$$
$$= \nu(E \cap E_n) = \int_E I_{E_n}(y)\,\nu(dy),$$

whence $\mu^y\bigl(\pi^{-1}(E_n)\bigr) = I_{E_n}(y)$ ν-a.e. Therefore, there exists a set Y_0 of full ν-measure such that $\mu^y\bigl(\pi^{-1}(E_n)\bigr) = I_{E_n}(y)$ for all $y \in Y_0$ and all n. This yields the relationship

$$\mu^y\bigl(\pi^{-1}(E)\bigr) = I_E(y), \quad \forall\, y \in Y_0,\ E \in \mathcal{E}.$$

Indeed, for every fixed $y \in Y_0$ both sides of this equality are measures as functions of E and coincide on $\mathcal{E}_0$. In particular, we obtain $\mu^y\bigl(\pi^{-1}(y)\bigr) = 1$. If there is a modification $\widetilde{\pi}$ with the properties listed in the formulation, then there exists a set $X_0 \in \mathcal{A}$ of full μ-measure on which π coincides with $\widetilde{\pi}$ and is $(\mathcal{A}, \mathcal{E})$-measurable. Then $Y_0 := \pi(X_0) = \widetilde{\pi}(X_0) \in \mathcal{E}_\nu$. The measure μ on X_0 possesses regular conditional measures μ^y, $y \in Y_0$, on $\mathcal{A}_{X_0} = \mathcal{A} \cap X_0$, such that $\mu^y\bigl(X_0 \cap \pi^{-1}(y)\bigr) = 1$ for ν-a.e. $y \in Y_0$. We extend the measures μ^y to $\mathcal{A}$ by setting $\mu^y(X\backslash X_0) = 0$. Since $X_0 \cap \pi^{-1}(y)$ belongs to $\mathcal{A}$ and is contained in $\pi^{-1}(y)$, the last assertion is proved. $\square$

We recall that if $\mathcal{E}$ is countably generated and countably separated, a mapping $f\colon X \to E$ is measurable with respect to $(\mathcal{A}_\mu, \mathcal{E})$, and $|\mu|$ is perfect, then the set $f(X)$ is $|\mu| \circ f^{-1}$-measurable.

10.4.11. Example. Let X and Y be Souslin spaces, μ a measure on $\mathcal{A} = \mathcal{B}(X)$, $\mathcal{E} = \mathcal{B}(Y)$, and let $\pi\colon X \to Y$ be measurable with respect to μ. Then there exist regular conditional measures μ^y, $y \in Y$, on $\mathcal{B}(X)$ such that $|\mu^y|\bigl(X\backslash\pi^{-1}(y)\bigr) = 0$ for $|\mu| \circ \pi^{-1}$-a.e. y.

Proof. If π is Borel measurable, then the above results apply, since $\mathcal{B}(X)$ and $\mathcal{B}(Y)$ are countably generated and separate the points. In the general case, there is a set $X_0 \in \mathcal{B}(X)$ with $|\mu|(X) = |\mu|(X_0)$ on which π is Borel. Then $Y_0 := \pi(X_0)$ is a Souslin set, hence there exists a Borel set $E \subset Y_0$ with $|\mu| \circ \pi^{-1}(Y_0\backslash E) = 0$. If $y \in E$, then we take measures μ^y constructed for $\pi|_{X_0}$; if $y \notin E$, then we let $\mu^y = \mu$. $\square$

In general, one cannot combine the $\mathcal{E}$-measurability of all functions $\mu^y(A)$, $A \in \mathcal{A}$, and the equality $|\mu^y|\bigl(X\backslash\pi^{-1}(y)\bigr) = 0$ for all $y \in \pi(X)$. Counterexamples exist even for continuous functions on a Borel subset of the interval (see Exercise 10.10.48). At the expense of the $\mathcal{E}$-measurability of all functions $\mu^y(A)$, $A \in \mathcal{A}$, but requiring their $|\mu| \circ \pi^{-1}$-measurability, the measures μ^y in Example 10.4.11 can be chosen in such a way that for every $y \in \pi(X)$ the measure μ^y will be concentrated on $\pi^{-1}(y)$. To this end, for all $y \in \pi(X)\backslash Y_0$ we take for μ^y a measure concentrated at an arbitrary point in $\pi^{-1}(y)$.

In the case of a Borel mapping one can find a $\sigma(\mathcal{S}_Y)$-measurable version of proper conditional measures, where $\mathcal{S}_Y$ is the class of Souslin sets in Y.

10.4.12. Proposition. *Let X and Y be Souslin spaces, let μ be a Borel probability measure on X, and let $f\colon X \to Y$ be a Borel mapping. Then there exist Borel probability measures $\mu(\,\cdot\,, y)$, $y \in Y$, on X such that:*

(i) *the functions $y \mapsto \mu(B, y)$, $B \in \mathcal{B}(X)$, are $\sigma(\mathcal{S}_Y)$-measurable,*

(ii) *one has $\mu\bigl(f^{-1}(y), y\bigr) = 1$ for every $y \in f(X)$,*

(iii) *for all $B \in \mathcal{B}(X)$ and $E \in \mathcal{B}(Y)$ one has*

$$\mu\bigl(B \cap f^{-1}(E)\bigr) = \int_E \mu(B, y)\, \mu \circ f^{-1}(dy).$$

PROOF. We know that there exist regular conditional measures μ^y on X such that (i) (even with the Borel measurability in place of $\sigma(\mathcal{S}_Y)$-measurability) and (iii) hold, and (ii) holds for $\mu\circ f^{-1}$-a.e. y. In order to obtain (ii) for all $y\in f(X)$, we redefine μ^y as follows. There is a Borel set $Y_0\subset Y$ such that $\mu\circ f^{-1}(Y_0)=1$ and (ii) holds for all $y\in Y_0$. In addition, there is a $\bigl(\sigma(\mathcal{S}_{f(X)}),\mathcal{B}(X)\bigr)$-measurable mapping $g\colon f(X)\to X$ such that $f\bigl(g(y)\bigr)=y$ for all $y\in f(X)$. Now let $\mu(\,\cdot\,,y)=\mu^y$ if $y\in Y_0$ and $\mu(\,\cdot\,,y)=\delta_{g(y)}$ if $y\in f(X)\backslash Y_0$. It is readily seen that we obtain desired measures. □

The constructed measures are also called conditional, although property (i) is weaker than the corresponding requirement in Definition 10.4.2. Such measures give a disintegration in the sense of §10.6.

10.4.13. Corollary. *Let X be a Souslin space and let $\mathcal{A}\subset\mathcal{B}(X)$ be a countably generated sub-σ-algebra. Then, for every Borel probability measure μ on X, there exist Borel probability measures $\mu(\,\cdot\,,x)$, $x\in X$, such that*:

(i) *the functions $x\mapsto\mu(B,x)$, $B\in\mathcal{B}(X)$, are $\sigma\bigl(S(\mathcal{A})\bigr)$-measurable,*

(ii) *$\mu(A,x)=1$ for all $A\in\mathcal{A}$ and all $x\in A$,*

(iii) *for all $B\in\mathcal{B}(X)$ and $A\in\mathcal{A}$ one has*

$$\mu(A\cap B)=\int_A\mu(B,x)\,\mu(dx).$$

PROOF. There is a Borel function $f\colon X\to[0,1]$ such that $\mathcal{A}=f^{-1}(\mathcal{B})$, $\mathcal{B}=\mathcal{B}([0,1])$. Let us take measures $\mu(\,\cdot\,,y)$ according to the proposition and set $\mu_0(B,x):=\mu\bigl(B,f(x)\bigr)$. Then we have (i), as $f^{-1}(\mathcal{S}_{[0,1]})\subset S(\mathcal{A})$, and (iii). In order to verify (ii) we observe that $A=f^{-1}(E)$, where $E\in\mathcal{B}$. Hence $y=f(x)\in E$ and $\mu_0(A,x)=\mu\bigl(f^{-1}(E),y\bigr)\geq\mu\bigl(f^{-1}(y),y\bigr)=1$. □

Clearly, both results extend to signed measures in the same spirit as above.

Now we consider the following important special case: $\Omega=X\times Y$, where $(X,\mathcal{A}_X)$ and $(Y,\mathcal{A}_Y)$ are two measurable spaces, $\mathcal{A}=\mathcal{A}_X\otimes\mathcal{A}_Y$. Let $\mathcal{B}_X$ and $\mathcal{B}_Y$ be the sub-σ-algebras in $\mathcal{A}$ formed, respectively, by the sets $A\times Y$ with $A\in\mathcal{A}_X$ and sets $X\times B$ with $B\in\mathcal{A}_Y$. Let μ_X and μ_Y be the images of μ under the natural projections to X and Y, and let $|\mu|_X$ and $|\mu|_Y$ denote the projections of the measure $|\mu|$. As above, let

$$A^y=\bigl\{A\cap(X\times\{y\})\bigr\}\quad\text{and}\quad A_y=\bigl\{x\colon\ (x,y)\in A\bigr\}.$$

10.4.14. Theorem. *Suppose that $\mathcal{A}_X$ is countably generated and that $|\mu|_X$ on $\mathcal{A}_X$ has a compact approximating class. Then, for every $y\in Y$, there exist a measure $\mu(\,\cdot\,,y)$ on $\mathcal{A}$ and a measure μ_y on $\mathcal{A}_X$ (probabilistic if so is μ) such that the function $y\mapsto\mu(A,y)=\mu_y(A_y)$ is measurable with respect to $|\mu|_Y$ for every $A\in\mathcal{A}$ and for all $B\in\mathcal{A}_Y$ one has*

$$\mu\bigl(A\cap(X\times B)\bigr)=\int_B\mu(A,y)\,|\mu|_Y(dy)=\int_B\mu(A,y)\,|\mu|_Y(dy),\qquad(10.4.9)$$

where $\mu(A,y)=\mu(A^y,y)$ *if* $\mathcal{A}_Y$ *contains the singletons. In addition, for every* $\mathcal{A}$*-measurable* μ*-integrable function* f *one has*

$$\int_\Omega f(x,y)\,\mu\bigl(d(x,y)\bigr)=\int_Y\int_X f(x,y)\,\eta^y(dx)\,|\mu|_Y(dy). \tag{10.4.10}$$

Finally, for any other families of measures $\mu'(\,\cdot\,,y)$ *and* μ'_y *with the stated properties one has* $\mu'(\,\cdot\,,y)=\mu(\,\cdot\,,y)$ *and* $\mu'_y=\mu_y$ *for* $|\mu|_Y$*-a.e.* y.

PROOF. It suffices to consider nonnegative measures by taking the Jordan–Hahn decomposition. The sets in $\mathcal{B}_X$ have the form $E\times Y$, $E\in\mathcal{A}_X$. According to Remark 10.4.9 applied to $\pi=\pi_Y$ and $\xi=\pi_X$, there exist probability measures $\mu^y_{\mathcal{B}_X}$, $y\in Y$, on $\mathcal{B}_X$ such that for all $E\in\mathcal{A}_X$, $B\in\mathcal{A}_Y$, one has

$$\mu\bigl((E\times Y)\cap(X\times B)\bigr)=\int_B \mu^y_{\mathcal{B}_X}(E\times Y)\,\mu_Y(dy), \tag{10.4.11}$$

where the integrand is $\mathcal{A}_Y$-measurable. We define probability measures μ_y on $\mathcal{A}_X$ by $\mu_y(E)=\mu^y_{\mathcal{B}_X}(E\times Y)$. Set

$$\mu(A,y)=\mu_y(A_y),\quad A\in\mathcal{A}.$$

By using that $A_y\in\mathcal{A}_X$, it is readily verified that $\mu(\,\cdot\,,y)$ is a probability measure on $\mathcal{A}$ for every $y\in Y$. Then, given $A=(E\times Y)\cap(X\times B)$, in view of the relationship $\pi_X\bigl((E\times Y)^y\bigr)=E$, equality (10.4.11) is written in the form

$$\mu(A)=\int_Y \mu(A^y,y)\,\mu_Y(dy), \tag{10.4.12}$$

since the section A^y is empty if $y\notin B$. It is clear that the class of all sets $A\in\mathcal{A}$ for which the function $y\mapsto\mu(A^y,y)$ is measurable with respect to $\mathcal{A}_Y$ and (10.4.12) holds, is monotone. By the above this class contains any finite unions of measurable rectangles, hence it coincides with $\mathcal{A}$. Then (10.4.9) holds as well, as $\bigl(A\cap(X\times B)\bigr)^y=A^y$ if $y\in B$, and if $y\notin B$, then this set is empty. If $\mathcal{A}_Y$ contains all singletons, then $A^y\in\mathcal{A}$ for all $A\in\mathcal{A}$ and hence $\mu(A,y)=\mu(A^y,y)$ because both sides equal $\mu_y(A_y)$. Formula (10.4.10) follows from what we have proved. The uniqueness statement follows by Lemma 10.4.3. □

If $\mathcal{A}_Y$ does not contain all singletons, then $A^y\notin\mathcal{A}$ and one has to employ the measures μ_y. Certainly, from the very beginning we could deal with the sets A_y, i.e., the projections of the geometric sections, as is done in Fubini's theorem. However, it is often more convenient to assume that the conditional measures are defined on $X\times\{y\}$. If $A^y\notin\mathcal{A}$, then this can be achieved by defining the measures μ^y by the equality $\mu^y(A):=\mu_y(A_y)$ on the distinct σ-algebras $\mathcal{A}_X\times\{y\}=\mathcal{A}\cap X\times\{y\}$ on $X\times\{y\}$ (which yields a disintegration in the sense of §10.6). There is no principal difference here, one should only remember that this is a question of conventions, in which one has to be consistent.

It is clear that the conclusion of the above theorem is true in the case where the whole σ-algebra $\mathcal{A}$ is countably generated and the measure μ on $\mathcal{A}$ has a compact approximating class (this follows from Theorem 10.4.5, but also is a corollary of the above theorem because one can verify that $\mathcal{A}_X$ is countably generated and has a compact approximating class). The next result demonstrates the advantages of our more general formulation.

10.4.15. Corollary. *Let $\Omega = X\times Y$, where X is a Souslin space with its Borel σ-algebra $\mathcal{A}_X = \mathcal{B}(X)$ and $(Y,\mathcal{A}_Y)$ is a measurable space, and let μ be a measure on $\mathcal{A} = \mathcal{A}_X\otimes\mathcal{A}_Y$. Then, for all $y\in Y$, there exist Radon measures μ^y on the spaces $X\times\{y\}$ and Radon measures μ_y on X such that for every set $A\in\mathcal{A}_X\otimes\mathcal{A}_Y$, the function $y\mapsto\mu_y(A_y) = \mu^y(A^y)$ is $\mathcal{A}_Y$-measurable and one has the equalities*

$$\mu(A) = \int_Y \mu^y(A^y)\,|\mu|_Y(dy) = \int_Y \mu_y(A_y)\,|\mu|_Y(dy).$$

For every other collection of measures μ'_y with the same properties one has $\mu'_y = \mu_y$ for $|\mu|_Y$-a.e. y and similarly for μ^y.

10.4.16. Example. Let X be a Souslin space, let Y be a Hausdorff space, and let μ be a Radon measure on $X\times Y$. Then, there exist Radon measures μ^y on the spaces $X\times\{y\}$, $y\in Y$, such that for every set $A\in\mathcal{B}(X\times Y)$ the function $y\mapsto\mu^y\bigl(A\cap(X\times\{y\})\bigr)$ is $|\mu|_Y$-measurable and one has the equality

$$\mu(A) = \int_Y \mu^y\bigl(A\cap(X\times\{y\})\bigr)\,|\mu|_Y(dy).$$

The measures μ^y are defined uniquely up to a set of $|\mu|_Y$-measure zero since $\mathcal{B}(X\times Y)\subset\sigma\bigl(S\bigl(\mathcal{B}(X)\bigr)\bigr)\otimes\mathcal{B}(Y)$ by Lemma 6.4.2 and Theorem 6.9.1.

As noted above, equality (10.4.15) (or (10.4.12)) is equivalent to (10.4.9), hence implies the essential uniqueness of measures μ_y. Certainly, the equality

$$\mu(A) = \int_Y \mu(A,y)\,|\mu|_Y(dy)$$

uniquely determines the measures $\mu(\,\cdot\,,y)$ for $|\mu|_Y$-a.e. y only if we have the equality $\mu(A,y) = \mu_y(A_y)$.

The next result follows easily from Theorem 10.4.14 (consider first the indicators of rectangles).

10.4.17. Corollary. *Let $(X,\mathcal{A})$ be the product of measurable spaces $(X_i,\mathcal{A}_i)$, $i = 1,\ldots,n$. Suppose that $\mathcal{A}$ is countably generated and that a probability measure μ on $\mathcal{A}$ has a compact approximating class. Denote by $\mu(dx_{k+1},x_1,\ldots,x_k)$ a regular conditional probability on $\mathcal{A}_{k+1}$ with respect to $\bigotimes_{i=1}^k\mathcal{A}_i$. Then the integral of any $\mathcal{A}$-measurable function $f\in\mathcal{L}^1(\mu)$ with respect to the measure μ equals*

$$\int_{X_1}\cdots\int_{X_n} f(x_1,\ldots,x_n)\,\mu(dx_n,x_1,\ldots,x_{n-1})\cdots\mu(dx_2,x_1)\,\mu_1(dx_1),$$

where μ_1 is the projection of μ on X_1.

Now we consider how to construct conditional expectations by means of regular conditional measures; a justification is clear from (10.4.5).

10.4.18. Proposition. *In the situation of Theorem* 10.4.5*, for every function* $f \in \mathcal{L}^1(\mu)$ *one has*

$$\mathbb{E}^{\mathcal{B}} f(x) = \int_X f(y)\,\mu(dy,x). \tag{10.4.13}$$

In the situation of Theorem 10.4.8*, one has*

$$\mathbb{E}^{\mathcal{B}} f(x) = \int_X f(z)\,\mu^{\pi(x)}(dz).$$

Now we give an example where there is no regular conditional measure even for a countably generated σ-algebra. Thus, the existence of a compact approximating class is an essential condition. Let us take two disjoint sets S_1 and S_2 in the interval $[0,1]$ such that both have inner measure 0, outer measure 1, and $S_1 \cup S_2 = [0,1]$ (see Example 1.12.13). Let $\mathcal{B}$ be the Borel σ-algebra of the interval and let $\mathcal{A}$ be the σ-algebra generated by $\mathcal{B}$ and the set S_1. It is clear that both σ-algebras are countably generated. Every set $A \in \mathcal{A}$ has the form

$$A = (B_1 \cap S_1) \cup (B_2 \cap S_2), \quad B_1, B_2 \in \mathcal{B}([0,1]).$$

Let λ be Lebesgue measure on $[0,1]$. It has been shown in Theorem 1.12.14 that the formula $\mu(A) = \bigl(\lambda(B_1) + \lambda(B_2)\bigr)/2$ defines a measure on $\mathcal{A}$ that coincides with λ on $\mathcal{B}$.

10.4.19. Example. On $\mathcal{A}$, there are no regular conditional measures with respect to $\mathcal{B}$.

PROOF. It is clear that the identity mapping of $([0,1],\mathcal{A})$ to $([0,1],\mathcal{B})$ is measurable. The image of the measure μ under this mapping is λ. It is seen from the proof of Corollary 10.4.10 that for λ-a.e. y, regular conditional measures must be Dirac measures: $\mu^y(A) = \delta_y(A)$. In particular, $\mu^y(S_1) = \delta_y(S_1)$ for all y outside some set Z of Lebesgue measure zero. Obviously, this contradicts the requirement of the λ-measurability of the function $\mu^y(S_1)$, which equals 1 on a set that differs from the nonmeasurable set S_1 only in a set of Lebesgue measure zero. □

See also Example 10.6.4 and Example 10.6.5 in §10.6. Taking Lebesgue measure λ on $\mathcal{B}([0,1])$ and the mapping $\pi(x) = x$ to the interval $[0,1]$ equipped with the σ-algebra $\mathcal{E}$ generated by the singletons, we obtain an example where there exist regular conditional measures $\lambda(A,y) \equiv \lambda(A)$, but there are no proper conditional measures, since such measures would coincide with δ_y, whereas the function $\mapsto \delta_y([0,1/2])$ is not $\mathcal{E}$-measurable.

Now we consider some examples of computation of conditional measures.

10.4.20. Example. Let μ be a Borel probability measure on the square $[0,1]^2$ defined by a density f with respect to Lebesgue measure. Then, regular conditional measures with respect to the projection to the first coordinate axis have the form

$$\mu^x(B) = \int_{\{y\colon\ (x,y)\in B\}} \frac{f(x,y)}{f_1(x)}\, dy, \quad x \in [0,1], \tag{10.4.14}$$

where

$$f_1(x) = \int_0^1 f(x,y)\, dy,$$

and we set $f(x,y)/f_1(x) = 0$ if $f_1(x) = 0$. In other words, the measure μ^x is concentrated on the vertical interval $\{x\}\times[0,1]$ and is given by the density $y \mapsto f(x,y)/f_1(x)$ with respect to the natural Lebesgue measure on this interval.

PROOF. According to Exercise 9.12.48 the image of the measure μ under the projection to the first coordinate axis (which we denote by ν) is given by the density f_1. By Fubini's theorem, the function defined by the right-hand side of (10.4.14) is finite for almost all x and ν-integrable. In addition, integrating this function against the measure ν, we obtain the integral of fI_B against Lebesgue measure. Since this function depends only on x, our assertion is proved. $\square$

10.4.21. Example. Suppose that for a probability measure μ on a measurable space $(X,\mathcal{A})$ we know regular conditional measures with respect to a sub-σ-algebra $\mathcal{B}\subset\mathcal{A}$. Then, for every measure η with a density ϱ with respect to the measure μ, regular conditional measures are given by the formula

$$\eta(A,x) = \frac{1}{\varrho_{\mathcal{B}}(x)} \int_A \varrho(y)\, \mu(dy,x), \tag{10.4.15}$$

where $\varrho_{\mathcal{B}}$ is the Radon–Nikodym density of the restriction of η to $\mathcal{B}$ with respect to the restriction of μ to $\mathcal{B}$, $|\eta|(\{\varrho_{\mathcal{B}}=0\})=0$, and we set $\eta(A,x)=0$ if $\varrho_{\mathcal{B}}(x)=0$ (one can also set $\eta(A,x)=\mu(A)$).

PROOF. Let $Z := \{x\colon\ \varrho_{\mathcal{B}}(x)=0\}$. Then $|\eta|(Z)=0$ because for every bounded $\mathcal{B}$-measurable function φ we have

$$\int_Z \varphi\, d\eta = \int_X I_Z\varphi\varrho\, d\mu = \int_X I_Z\varphi\varrho_{\mathcal{B}}\, d\mu = 0.$$

It follows by (10.4.2) that the function defined by the right-hand side of (10.4.15) is finite η-a.e. It is readily verified that this function is measurable with respect to $\mathcal{B}$ (it suffices to approximate the function ϱ by simple functions). Finally, according to (10.4.2) one has

$$\begin{aligned}\int_B \eta(A,x)\,\eta(dx) &= \int_B\int_A \varrho(y)\,\mu(dy,x)\,\mu(dx)\\ &= \int_B I_A(x)\varrho(x)\,\mu(dx) = \eta(A\cap B)\end{aligned}$$

for all $B\in\mathcal{B}$. $\square$

In a similar manner the next example is justified.

10.4.22. Example. (i) Let $(X_1, \mathcal{B}_1, \mu_1)$ and $(X_2, \mathcal{B}_2, \mu_2)$ be two spaces with probability measures. Then, for the measure $\mu := \mu_1 \otimes \mu_2$ on $\mathcal{B}_1 \otimes \mathcal{B}_2$, the conditional measures with respect to the σ-algebra generated by the projection to X_1 have the form

$$\mu(B, x_1, x_2) = \mu_2\bigl(y_2 \in X_2\colon\ (x_1, y_2) \in B\bigr).$$

In other words, $\mu(\,\cdot\,, x_1, x_2) = \delta_{x_1} \otimes \mu_2$. In terms of conditional measures generated by the indicated projection this can be written as $\mu^{x_1} = \delta_{x_1} \otimes \mu_2$.

(ii) Let ν be a probability measure on $\mathcal{B}_1 \otimes \mathcal{B}_2$ absolutely continuous with respect to the measure $\mu = \mu_1 \otimes \mu_2$ in (i) and let $\varrho = d\nu/d\mu$. Then, the conditional measures for ν with respect to the σ-algebra generated by the projection to X_1 have the following form:

$$\nu(B, x_1, x_2) = \biggl(\int_{X_2} \varrho(x_1, y_2)\, \mu_2(dy_2)\biggr)^{-1} \int_X I_B(x_1, y_2)\varrho(x_1, y_2)\, \mu_2(dy_2).$$

Now we consider convergence of conditional measures in variation.

10.4.23. Proposition. *Suppose that measures μ_n on a measurable space $(X, \mathcal{A})$ converge in variation to a measure μ. Let $\mathcal{B}$ be a sub-σ-algebra in $\mathcal{A}$ such that the measure $\nu := \sum_{n=1}^{\infty} 2^{-n}|\mu_n|$ on $\mathcal{A}$ has a regular conditional probability measure $\nu(\,\cdot\,,\,\cdot\,)$ with respect to $\mathcal{B}$ (which is the case if X is a Souslin space and $\mathcal{A} = \mathcal{B}(X)$). Then one can choose a subsequence $\{n_i\}$ and regular with respect to $\mathcal{B}$ conditional measures $\mu_{n_i}(\,\cdot\,,\,\cdot\,)$ and $\mu(\,\cdot\,,\,\cdot\,)$ for μ_{n_i} and μ such that for $|\mu|$-a.e. x, the measures $\mu_{n_i}(\,\cdot\,, x)$ converge in variation to $\mu(\,\cdot\,, x)$.*

Proof. It is clear that $\mu_n \ll \nu$ and $\mu \ll \nu$. Let us set $f_n := d\mu_n/d\nu$, $f := d\mu/d\nu$, where we choose $\mathcal{A}$-measurable versions, and let g_n and g be the conditional expectations of f_n and f with respect to the σ-algebra $\mathcal{B}$ and the measure ν. In view of Example 10.4.21 one has

$$\mu_n(\,\cdot\,, x) = g_n(x)^{-1} f_n(\,\cdot\,) \cdot \nu(\,\cdot\,, x)$$

for μ_n-a.e. x (where $|\mu_n|(\{g_n = 0\}) = 0$) and

$$\mu(\,\cdot\,, x) = g(x)^{-1} f(\,\cdot\,) \cdot \nu(\,\cdot\,, x)$$

for μ-a.e. x (where $|\mu|(\{g = 0\}) = 0$). If $g_n(x) = 0$ or $g(x) = 0$, then we set respectively $\mu_n(\,\cdot\,, x) = \nu(\,\cdot\,, x)$ or $\mu(\,\cdot\,, x) = \nu(\,\cdot\,, x)$. We have

$$\|\mu_n - \mu\| = \int_X |f_n - f|\, d\nu = \int_X \int_X |f_n(y) - f(y)|\, \nu(dy, x)\, \nu(dx).$$

Since the measures μ_n converge to μ in variation, there is a subsequence $\{n_i\}$ such that for ν-a.e. x, the sequence of functions f_{n_i} converges to the function f in $L^1\bigl(\nu(\,\cdot\,, x)\bigr)$, i.e., the measures $f_{n_i} \cdot \nu(\,\cdot\,, x)$ converge in variation to the measure $f \cdot \nu(\,\cdot\,, x)$. The functions g_n converge to g in $L^1(\nu)$, which gives convergence almost everywhere if we choose a suitable subsequence in $\{n_i\}$

denoted by the same symbol. Since $g(x) \neq 0$ for $|\mu|$-a.e. x, we obtain a desired subsequence. $\square$

Let us note that in the case of probability measures the obtained result gives the μ-a.e. convergence of the conditional expectations $\mathbb{E}^{\mathcal{B}}_{\mu_{n_i}} f \to \mathbb{E}^{\mathcal{B}}_{\mu} f$ for any bounded $\mathcal{A}$-measurable function f.

If we require the ν-a.e. convergence $f_n \to f$ and $g_n \to g$ (as in Gänssler, Pfanzagl [**655**]), then we obtain the $|\mu|$-a.e. convergence of the conditional measures.

The following example shows that in the considered situation, there might be no convergence (even in the weak topology!) of the whole sequence, so that it is indeed necessary to select a subsequence.

10.4.24. Example. There is a sequence of Borel probability measures μ_n on $[0,1]\times[0,1]$ with densities $\varrho_n > 0$ that converges in variation to Lebesgue measure λ, but, for every fixed $x \in [0,1]$, the conditional measures μ_n^x do not converge weakly on $[0,1]$, in particular, do not converge in variation.

PROOF. Let $\varrho_n(x,y) = 1 + \varphi_n(x)\psi(y)$, where $\psi_n(y) = 1$ if $y \in [0,1/2]$, $\psi_n(y) = -1$ if $y \in (1/2,1]$, $0 \le \varphi_n \le 1$, and $\{\varphi_n\}$ converges to 0 in measure but at no point. Then $|\varrho_n(x,y) - 1| \le \varphi_n(x)$, which yields convergence in $L^1([0,1]\times[0,1])$. The conditional measure μ_n^x is given by the density $y \mapsto \varrho_n(x,y)$, which is a probability density. Clearly, there is no weak convergence of these conditional measures. Indeed, the integral of $\varrho_n(x,y)$ in y over $[0,t]$ equals $t + t\varphi_n(x)$ if $t \le 1/2$. $\square$

The following result on convergence of conditional measures is proved in Blackwell, Dubins [**181**].

10.4.25. Proposition. *Suppose we are given a sequence of measurable spaces $(X_i, \mathcal{A}_i)$ and two probability measures μ and ν on their product $(X, \mathcal{A})$ with $\nu \ll \mu$. Assume that for every n, the measure μ has regular conditional probability measures $\mu_{x_1,\ldots,x_n}$, $(x_1,\ldots,x_n) \in \prod_{i=1}^n X_i$, on the σ-algebra $\mathcal{B}_{n+1} := \mathcal{A}_{n+1} \bigotimes \mathcal{A}_{n+2} \bigotimes \cdots$ in the space $Z_{n+1} := X_{n+1} \times X_{n+2} \times \cdots$. Then the measure ν also has regular conditional probability measures $\nu_{x_1,\ldots,x_n}$ on $\mathcal{B}_{n+1}$, and for ν-a.e. $(x_1, x_2, \ldots) \in X$ one has*

$$\lim_{n\to\infty} \|\mu_{x_1,\ldots,x_n} - \nu_{x_1,\ldots,x_n}\| = 0.$$

PROOF. Let us fix an $\mathcal{A}$-measurable version ϱ of the Radon–Nikodym density $d\nu/d\mu$. Let

$$\varrho_n(x_1,\ldots,x_n) = \int_{Z_{n+1}} \varrho(x_1,x_2,\ldots)\, \mu_{x_1,\ldots,x_n}\bigl(d(x_{n+1},x_{n+2},\ldots)\bigr),$$

$$\xi_n(x_1,\ldots,x_n,x_{n+1},\ldots) = \frac{\varrho(x_1,x_2,\ldots)}{\varrho_n(x_1,\ldots,x_n)}$$

whenever $\varrho_n(x_1,\ldots,x_n)\neq 0$ and $\xi_n(x_1,\ldots,x_n,x_{n+1},\ldots)=0$ otherwise. Let us introduce functions

$$\psi_{x_1,\ldots,x_n}(x_{n+1},\ldots)=\xi_n(x_1,\ldots,x_n,x_{n+1},\ldots)$$

on Z_{n+1}. Then the measures

$$\nu_{x_1,\ldots,x_n}:=\psi_{x_1,\ldots,x_n}\cdot\mu_{x_1,\ldots,x_n}$$

serve as regular conditional probability measures for ν. For every $\varepsilon>0$, we have

$$\begin{aligned}\|\mu_{x_1,\ldots,x_n}-\nu_{x_1,\ldots,x_n}\| &= \int_{Z_{n+1}}|1-\psi_{x_1,\ldots,x_n}|\,d\mu_{x_1,\ldots,x_n}\\ &= 2\int_{\{\psi_{x_1,\ldots,x_n}>1\}}[\psi_{x_1,\ldots,x_n}-1]\,d\mu_{x_1,\ldots,x_n}\\ &\le 2\varepsilon+2\int_{\{\psi_{x_1,\ldots,x_n}>1+\varepsilon\}}[\psi_{x_1,\ldots,x_n}-1]\,d\mu_{x_1,\ldots,x_n}\\ &\le 2\varepsilon+2\nu_{x_1,\ldots,x_n}\bigl(\{\psi_{x_1,\ldots,x_n}>1+\varepsilon\}\bigr).\end{aligned}$$

Let us observe that

$$\nu_{x_1,\ldots,x_n}\bigl(\{\psi_{x_1,\ldots,x_n}>1+\varepsilon\}\bigr)=\mathbb{E}_\nu^{\mathcal{F}_n}I_{\{\xi_n>1+\varepsilon\}}(x_1,\ldots,x_n),$$

where $\mathcal{F}_n:=\mathcal{A}_1\bigotimes\cdots\bigotimes\mathcal{A}_n$. By the martingale convergence theorem, for ν-a.e. $(x_1,x_2,\ldots)$ the sequence $\varrho_n(x_1,\ldots,x_n)$ converges to $\varrho(x_1,x_2,\ldots)$, which yields convergence of the sequence $\xi_n(x_1,\ldots,x_n,x_{n+1},\ldots)$ to 1. Hence one has $I_{\{\xi_n>1+\varepsilon\}}\to 0$. Therefore, $\mathbb{E}_\nu^{\mathcal{F}_n}I_{\{\xi_n>1+\varepsilon\}}(x_1,\ldots,x_n)\to 0$ for ν-a.e. point $(x_1,x_2,\ldots)$ according to Exercise 10.10.39, which completes the proof. $\square$

10.5. Liftings and conditional measures

In this section, we consider another approach to constructing conditional measures that is based on the concept of lifting, which definitely deserves a discussion in its own right. This concept arises in fact right after introducing classes of equivalent functions in the sense of equality almost everywhere. Is it possible to pick in every equivalence class in the set of all bounded measurable functions exactly one representative in such a way that the algebraic relationships (sums and products) that hold for classes be fulfilled pointwise for these representatives? Such a choice is called a lifting. Let us give precise definitions.

10.5.1. Definition. *Let $(X,\mathcal{A},\mu)$ be a measurable space with a nonnegative measure μ (possibly with values in $[0,+\infty]$) and let $\mathcal{L}^\infty_{\mathcal{A}}$ be the space of all bounded $\mathcal{A}$-measurable functions. A lifting on $\mathcal{L}^\infty_{\mathcal{A}}$ is a mapping $L\colon\ \mathcal{L}^\infty_{\mathcal{A}}\to\mathcal{L}^\infty_{\mathcal{A}}$ satisfying the following conditions*:

(i) $L(f)=f\ \ \mu$*-a.e.*;

(ii) $L(f)(x)=L(g)(x)$ *for all* $x\in X$ *if* $f=g\ \ \mu$*-a.e.*;

(iii) $L(f)(x)=1$ *for all* $x\in X$ *if* $f=1$ μ*-a.e.*;

(iv) $L(\alpha f+\beta g)(x)=\alpha L(f)(x)+\beta L(g)(x)$ *for all* $x\in X$, $f,g\in\mathcal{L}^\infty_{\mathcal{A}}$ *and* $\alpha,\beta\in\mathbb{R}^1$;

(v) $L(fg)(x)=L(f)(x)L(g)(x)$ *for all* $x\in X$, $f,g\in\mathcal{L}^\infty_{\mathcal{A}}$.

We observe that if L is a lifting, then for all $A\in\mathcal{A}$ we have $L(I_A)=L(I_A^2)=L(I_A)^2$, i.e., the function $L(I_A)$ takes values in $\{0,1\}$ and hence is the indicator of some set $\widetilde{A}\in\mathcal{A}$. This enables us to define the mapping $L\colon\mathcal{A}\to\mathcal{A}$, $L(A):=\widetilde{A}$. By the properties of liftings, this mapping satisfies the following conditions:

(1) $\mu\bigl(L(A)\bigtriangleup A\bigr)=0$,

(2) $L(A)=L(B)$ if $\mu(A\bigtriangleup B)=0$,

(3) $L(X)=X$, $L(\varnothing)=\varnothing$,

(4) $L(A\cup B)=L(A)\cup L(B)$,

5) $L(A\cap B)=L(A)\cap L(B)$.

The mapping $L\colon\mathcal{A}\to\mathcal{A}$ is called a lifting of the σ-algebra $\mathcal{A}$. It is clear that every lifting of the σ-algebra $\mathcal{A}$ uniquely defines a lifting on $\mathcal{L}^\infty_{\mathcal{A}}$ by the formula $L(I_A):=I_{L(A)}$, extended by linearity to all simple functions and then by means of uniform approximations to all of the space $\mathcal{L}^\infty_{\mathcal{A}}$. Thus, liftings correspond one-to-one to liftings of the σ-algebra.

For every lifting we have

(iv') $L(f)\ge 0$ if $f\ge 0$ μ-a.e.

Indeed, $L(f)=L(\sqrt{f})L(\sqrt{f})$.

The most important is the case when $\mathcal{A}$ is $\mathcal{A}_\mu$. This is exactly the above-mentioned problem of selecting in every equivalence class in $\mathcal{L}^\infty(\mu)$ a representative with pointwise preservation of the algebraic operations. It is clear that liftings of $\mathcal{L}^\infty(\mu)$ can be identified with homomorphisms from the algebra $L^\infty(\mu)$ to the algebra $\mathcal{L}^\infty(\mu)$ such that any equivalence class is sent to its representative. For this reason, L is also called a lifting of $L^\infty(\mu)$.

A weaker concept than a lifting is a *linear lifting*. This is a mapping L with properties (i)–(iv) and (iv'). The following result enables one to reduce the construction of a lifting to finding a linear lifting, which is somewhat simpler, as we shall see below.

10.5.2. Lemma. *Let L_0 be a linear lifting for a complete probability measure μ on a measurable space $(X,\mathcal{A})$. For each $A\in\mathcal{A}$ let*

$$E(A):=\bigl\{x\colon\ L_0(I_A)(x)=1\bigr\},\quad P(A):=\bigl\{x\colon\ L_0(I_A)(x)>0\bigr\}.$$

Then, there exists a lifting L such that $I_{E(A)}\le L(I_A)\le I_{P(A)}$ for all $A\in\mathcal{A}$.

PROOF. We consider the set Λ of all linear liftings l such that

$$I_{E(A)}\le l(I_A)\le I_{P(A)},\quad A\in\mathcal{A}.$$

Then $L_0\in\Lambda$, which follows by the definition of $E(A)$ and $P(A)$, since we have $0\le L_0(I_A)\le 1$. The set Λ is convex if it is regarded as a subset in the product $\mathbb{R}^\Omega$, $\Omega=\mathcal{L}^\infty_{\mathcal{A}}\times X$, by means of the natural embedding

$$l\mapsto\bigl(l(f)(x)\bigr)_{(f,x)\in\Omega}.$$

It is clear that Λ is contained in the product of compact intervals, since $|l(f)(x)| \le \sup_{y\in X} |f(y)|$ for all $l \in \Lambda$ by property (iv') of a linear lifting.

The set Λ is closed in $\mathbb{R}^\Omega$ in the product topology. Indeed, let an element $\xi\colon (f,x) \mapsto \xi(f,x) \in \mathbb{R}^\Omega$ be the limit of a net of elements $l_\alpha \in \Lambda$, i.e.,

$$\xi(f,x) = \lim_\alpha l_\alpha(f)(x) \quad \text{for all } f \in \mathcal{L}^\infty_{\mathcal{A}} \text{ and } x \in X.$$

We set $l(f)(x) := \xi(f,x)$ and show that $l \in \Lambda$. It is obvious that conditions (ii)–(iv) and (iv') from the definition of a linear lifting are satisfied and one has the estimate

$$I_{E(A)} \le l(I_A) \le I_{P(A)} \quad \text{for all } A \in \mathcal{A},$$

since all these relationships are pointwise. However, we have to verify the equality $l(f) = f$ a.e. (because we deal with a possibly uncountable net). Let $f = I_A$, where $A \in \mathcal{A}$. Then one has a.e.

$$I_{E(A)} \le l(I_A) \le I_{P(A)} \quad \text{and} \quad I_{E(A)}(x) = I_{P(A)}(x) = I_A(x)$$

which by the completeness of the measure μ on $\mathcal{A}$ yields the $\mathcal{A}$-measurability of $l(f)$ and the equality $l(f)(x) = f(x)$ a.e. This equality extends to finite linear combinations of indicators of sets in $\mathcal{A}$. An arbitrary function $f \in \mathcal{L}^\infty_{\mathcal{A}}$ is the uniform limit of a sequence of simple functions $f_j \in \mathcal{L}^\infty_{\mathcal{A}}$, for which the equality $l(f_j) = f_j$ a.e. is already established. Since the functions $l(f_j)$ converge uniformly to $l(f)$ by property (iv'), we obtain $l(f) = f$ a.e.

Since the product of compact intervals is compact, the set Λ is convex and compact. By the Krein–Milman theorem (see Dunford, Schwartz [**503**, Ch. V, §8]) Λ has extreme points, i.e., points that are not representable as a convex combination $tl' + (1-t)l''$ with $t \in (0,1)$, $l', l'' \in \Lambda$. Let L be such an extreme point. We show that L is a required lifting. In fact, we have to verify that $L(fg) = L(f)L(g)$. Suppose that this is not true, i.e., there exist $f, g \in \mathcal{L}^\infty_{\mathcal{A}}$ and $a \in X$ such that $L(fg)(a) \ne L(f)(a)L(g)(a)$. Then we observe that one can take g with $0 \le g \le 1$ (the validity of the above equality for all g with this restriction yields its validity for all g). Let $L_1(\varphi) = L(\varphi) + L(g\varphi) - L(g)L(\varphi)$, $L_2(\varphi) = L(\varphi) - L(g\varphi) + L(g)L(\varphi)$. It is clear that $L = (L_1 + L_2)/2$ and $L_1 \ne L_2$ because $L_1(f)(a) \ne L_2(f)(a)$. Let us verify that $L_1, L_2 \in \Lambda$. The functionals L_1 and L_2 are linear and $L_1(1) = L_2(1) = 1$. If $\varphi \ge 0$, then $L_1(\varphi) = \bigl(1 - L(g)\bigr)L(\varphi) + L(g\varphi) \ge 0$, since $L(g) \le 1$, $L(\varphi) \ge 0$ and $L(g\varphi) \ge 0$. Similarly, $L_2(\varphi) \ge 0$. Therefore, $0 \le L_i(\psi) \le 1$, $i = 1, 2$, whenever $0 \le \psi \le 1$. It is clear that $L_1(\varphi) = L_2(\varphi) = \varphi$ a.e. Finally, we have $I_{E(A)} \le L_1(I_A) \le I_{P(A)}$ and $I_{E(A)} \le L_2(I_A) \le I_{P(A)}$, since these inequalities hold for $L = (L_1 + L_2)/2$ and $0 \le L_i(I_A) \le 1$. Hence we obtain a contradiction with the fact that L is an extreme point. □

Linear liftings are easier to construct. We shall consider a special case – Lebesgue measure on an interval, which makes transparent the proof in the general case.

10.5.3. Example. All functions on $[0,1]$ will be extended by zero outside $[0,1]$. We know that for every bounded measurable function f the limit of the quantities

$$E_n f(x) := n \int_x^{x+n^{-1}} f(y)\,dy$$

a.e. equals $f(x)$. On the space m of all bounded sequences with the sup-norm, there exists a generalized limit, i.e., a continuous linear functional Λ that is nonnegative on nonnegative sequences and coincides with the usual limit on all convergent sequences (see Exercise 2.12.100). Set

$$L(f)(x) := \Lambda\Bigl(\bigl(E_n f(x)\bigr)_{n=1}^{\infty}\Bigr).$$

Then $L(f)(x) = f(x)$ at all points x where $f(x) = \lim\limits_{n\to\infty} E_n f(x)$, i.e., almost everywhere. The linearity and nonnegativity of L are obvious. Thus, L is a linear lifting and $L(f) = f$ for all continuous f. By Lemma 10.5.2 we obtain the existence of a lifting on $[0,1]$ with Lebesgue measure.

10.5.4. Theorem. *For every complete probability measure μ, there exists a lifting on $\mathcal{L}^\infty(\mu)$.*

Proof. We consider the set $\mathcal{M}$ consisting of all pairs $(\mathcal{E}, L)$, where $\mathcal{E}$ is a sub-σ-algebra in $\mathcal{A}$ containing the σ-algebra $\mathcal{A}_0$ generated by all measure zero sets and L is a lifting on $\mathcal{E}$. The set $\mathcal{M}$ is not empty, since on $\mathcal{A}_0$ one has the lifting L_0 defined as follows: $L_0(I_A) = 0$ if $\mu(A) = 0$, $L_0(I_A) = 1$ if $\mu(A) = 1$. The set $\mathcal{M}$ is equipped with the following order: $(\mathcal{E}_1, L_1) \le (\mathcal{E}_2, L_2)$ if $\mathcal{E}_1 \subset \mathcal{E}_2$ and $L_2|_{\mathcal{E}_1} = L_1$. We show that $\mathcal{M}$ contains a maximal element. By Zorn's lemma it suffices to verify that every linearly ordered part $\{(\mathcal{E}_\alpha, L_\alpha)\}$ in $\mathcal{M}$ has an upper bound. Let $\mathcal{E}$ be the σ-algebra generated by all $\mathcal{E}_\alpha$. We shall construct on $\mathcal{E}$ a lifting L whose restriction to $\mathcal{E}_\alpha$ is L_α for all α. Then the pair $(\mathcal{E}, L)$ will be an upper bound.

Suppose first that for every sequence $(\mathcal{E}_{\alpha_n}, L_{\alpha_n})$ in $\mathcal{M}$ one has an upper bound $(\mathcal{E}_\beta, L_\beta)$ in $\mathcal{M}$. Then L can be defined in the following way. It is easy to see that for every $\mathcal{E}$-measurable bounded function f, there exists a countable collection of indices α_n such that f is measurable with respect to the σ-algebra generated by all $\mathcal{E}_{\alpha_n}$. We take β with $(\mathcal{E}_{\alpha_n}, L_{\alpha_n}) \le (\mathcal{E}_\beta, L_\beta)$ for all n and set $L(f) = L_\beta(f)$. It is readily verified that due to the linear ordering of the considered collection, L is well-defined. It is clear that L is a lifting. Suppose now that our assumption is false for some sequence $(\mathcal{E}_{\alpha_n}, L_{\alpha_n})$. It is clear that this sequence can be taken as increasing. Since for every α, there exists n with $(\mathcal{E}_\alpha, L_\alpha) \le (\mathcal{E}_{\alpha_n}, L_{\alpha_n})$ because otherwise $\alpha_n \le \alpha$ for all n, we obtain that $\mathcal{E}$ is generated by an increasing sequence of σ-algebras $\mathcal{E}_{\alpha_n}$. By Theorem 10.2.3, for every bounded $\mathcal{E}$-measurable function f, almost everywhere there exists a limit $\lim\limits_{n\to\infty} \mathbb{E}^{\mathcal{E}_n} f(x)$ and this limit equals $f(x)$ a.e. By means of this limit we define a linear lifting L_0. To this end, as in the case of the interval, we take a generalized limit Λ on the space m of all bounded sequences and set

$$L_0(f)(x) := \Lambda\Big(\big(L_{\alpha_n}\big[\mathbb{E}^{\mathcal{E}_n} f\big](x)\big)_{n=1}^{\infty}\Big).$$

Then L_0 is a linear mapping, $0 \le L_0(f) \le 1$ whenever $0 \le f \le 1$, and one has $L_0(f) = L_{\alpha_n}(f)$ if the function f is measurable with respect to $\mathcal{E}_{\alpha_n}$. Indeed, in this case for all $k \ge n$ we have $L_{\alpha_k}\big[\mathbb{E}^{\mathcal{E}_k} f\big] = L_{\alpha_n}\big[\mathbb{E}^{\mathcal{E}_n} f\big] = L_{\alpha_n}(f)$, whence the indicated equality follows. For every $\mathcal{E}$-measurable bounded function f we have $L_0(f) = f$ a.e., since we have $f(x) = \lim\limits_{n\to\infty} \mathbb{E}^{\mathcal{E}_n} f(x)$ a.e. and $\mathbb{E}^{\mathcal{E}_n} f(x) = L_{\alpha_n}\big[\mathbb{E}^{\mathcal{E}_n} f\big](x)$ a.e. Let L be a lifting on $\mathcal{E}$ given by Lemma 10.5.2. The restriction of L to $\mathcal{E}_\alpha$ is L_α for every α. Indeed, there is n with $\alpha \le \alpha_n$, whence for any $A \in \mathcal{E}_\alpha$ we obtain $L_0(I_A) = L_{\alpha_n}(I_A)$. Hence $I_{E(A)} = I_{P(A)} = L_{\alpha_n}(I_A)$. Therefore, $L(I_A) = L_{\alpha_n}(I_A)$, which yields the equality on all bounded $\mathcal{E}_\alpha$-measurable functions.

Thus, $\mathcal{M}$ has at least one maximal element $(\mathcal{E}, L)$. We show that $\mathcal{E} = \mathcal{A}$, which will bring our proof to an end. As we have explained earlier, it suffices to prove that if there is a measurable set $E_0 \notin \mathcal{E}$, then there exists a lifting of the σ-algebra $\mathcal{E}_0$ generated by $\mathcal{E}$ and E_0 that extends the given lifting L on $\mathcal{E}$. The elements of $\mathcal{E}_0$ are all sets of the form

$$C = (E_0 \cap A) \cup [(X\backslash E_0) \cap B], \quad A, B \in \mathcal{E}. \tag{10.5.1}$$

For every set S, let $\mathcal{Z}(S)$ denote the collection of all sets $E \in \mathcal{E}$ such that $\mu(E \cap S) = 0$. Let Ω_1 denote the union of the sets $L(D)$ over all $D \in \mathcal{Z}(E_0)$, and let Ω_2 be the union of the sets $L(F)$ over all $F \in \mathcal{Z}(X\backslash E_0)$. Let

$$E_0' := \big(E_0 \cap (X\backslash\Omega_1)\big) \cup \big((X\backslash E_0) \cap \Omega_2\big).$$

There exist sets $E_n \in \mathcal{Z}(E_0)$ such that $E_n \subset E_{n+1}$ and

$$\lim_{n\to\infty} \mu(E_n) = \sup\{\mu(E)\colon\ E \in \mathcal{Z}(E_0)\}.$$

It is clear that $E_\infty := \bigcup_{n=1}^\infty E_n \in \mathcal{Z}(E_0)$. We observe that $L(E_\infty) = \Omega_1$. Indeed, $L(E_\infty) \subset \Omega_1$. On the other hand, for every set $D \in \mathcal{Z}(E_0)$, we have $\mu(D\backslash E_\infty) = 0$ by the construction of E_∞, whence one has the inclusion $L(D) \subset L(E_\infty)$. Thus, $\Omega_1 \in \mathcal{Z}(E_0) \subset \mathcal{E}$. Similarly, we prove the existence of a set $D_\infty \in \mathcal{Z}(X\backslash E_0)$ such that $\Omega_2 = L(D_\infty) \in \mathcal{Z}(X\backslash E_0)$. Therefore, $E_0' \in \mathcal{E}_0$. One can readily verify the equality $\mu(E_0 \triangle E_0') = 0$. Further,

$$\Omega_1 \cap \Omega_2 = L(E_\infty) \cap L(D_\infty) = L(E_\infty \cap D_\infty) = \varnothing.$$

Now for every $A \in \mathcal{Z}(E_0)$, we obtain $E_0' \cap L(A) \subset E_0' \cap \Omega_1 \subset \Omega_2 \cap \Omega_1 = \varnothing$. Similarly, for all $B \in \mathcal{Z}(X\backslash E_0)$, we have $(X\backslash E_0') \cap L(B) = \varnothing$. For sets of the form (10.5.1), we let

$$L_0(C) := \big(E_0' \cap L(A)\big) \cup \big[(X\backslash E_0') \cap L(B)\big].$$

By using the above relationships it is easily verified that L_0 is a lifting on $\mathcal{E}_0$ such that $L_0|_{\mathcal{E}} = L$ and $L_0(E_0) = E_0'$. $\square$

It is clear that a lifting exists for any complete nonnegative σ-finite measure (even for any decomposable measure, and the converse is true, see Exercise 10.10.52).

It remains an open question how essential the completeness of the measure μ is (which has been used in Lemma 10.5.2). For example, the question arises whether in the lifting theorem one can choose Borel representatives in the equivalence classes in the case of Lebesgue measure on the real line. It was shown in von Neumann, Stone [**1367**] under the continuum hypothesis that a Borel lifting exists in the case of Lebesgue measure. However, according to Shelah [**1695**], it is consistent with set theory ZFC that there is no such lifting.

There are no linear liftings on the spaces $L^p[0,1]$ with $1 \le p < \infty$ (Exercise 10.10.53).

Now we employ liftings to prove one more result on the existence of regular conditional measures. We begin with an auxiliary lemma. Let $(X, \mathcal{A}, \mu)$ be a measurable space with a finite nonnegative measure, let L be a lifting on the space $\mathcal{L}^\infty(\mu)$, and let $\mathcal{L} := L\bigl(\mathcal{L}^\infty(\mu)\bigr)$. Then $\mathcal{L}$ turns out to be a complete vector lattice (see Chapter 4, §4.7(i)). Due to property (ii) of liftings the order relation in $\mathcal{L}$ is the pointwise inequality $f(x) \le g(x)$ (unlike the a.e. inequality in $\mathcal{L}^\infty(\mu)$). Let M be a subset of $\mathcal{L}$ bounded from above. Denote by $\vee(M)$ the lattice supremum of M (which exists, since $\mathcal{L}$ is complete) and set

$$\sup(M)(x) := \sup\{f(x)\colon\ f \in M\}.$$

It turns out that the function $\sup(M)$ is measurable. Certainly, this is due to a special structure of the set $\mathcal{L}$: it is easy to give an example of a family of uniformly bounded measurable functions whose supremum is not measurable.

10.5.5. Lemma. (i) *Suppose that M is a subset of $\mathcal{L}$ bounded from above. Then* $\sup(M)$ *is a μ-measurable function,* $\sup(M) = \vee(M)$ *a.e. and* $\sup(M) \le \vee(M)$ *everywhere.*

(ii) *Let $\{f_\alpha\}$ be a bounded increasing net in $\mathcal{L}$. Then*

$$\int_X \sup_\alpha f_\alpha(x)\,\mu(dx) = \sup_\alpha \int_X f_\alpha(x)\,\mu(dx).$$

In particular, if $\{A_\alpha\}$ is an increasing net of measurable sets, then the set $\bigcup_\alpha L(A_\alpha)$, where L also denotes the lifting of $\mathcal{A}_\mu$, is measurable and its measure is $\sup_\alpha \mu(A_\alpha)$.

Proof. (i) We have the pointwise inequality $\sup(M)(x) \le \vee(M)(x)$, since $f(x) \le \vee(M)(x)$ for all x (we recall again that the order relation in $\mathcal{L}$ is the pointwise inequality). By Corollary 4.7.2, there exists a sequence $\{f_n\} \subset M$ such that $\vee(M) = \vee\{f_n\}$. Let $f = \sup_n f_n$. Then the function f is measurable with respect to μ and $f \le \sup(M) \le \vee(M)$ everywhere. On the other hand, $f \ge f_n$ for every n, hence by the definition of a lifting $Lf \ge f_n$ everywhere. Therefore, $Lf \ge \vee\{f_n\} = \vee(M)$, whence $f \ge \vee(M)$ a.e.

(ii) Set $M = \{f_\alpha\}$ and choose a sequence $\{f_n\}$ as above. One can assume that $\{f_n\}$ is increasing because due to the increasing of $\{f_\alpha\}$ one can pass to the sequence $\max_{i=1}^n f_i$. Then $\sup(M) = \sup_n f_n = \lim\limits_{n\to\infty} f_n$ a.e., so one has

$$\int_X \sup(M)(x)\,\mu(dx) = \lim_{n\to\infty}\int_X f_n(x)\,\mu(dx),$$

which is majorized by

$$\sup_\alpha \int_X f_\alpha(x)\,\mu(dx).$$

The reverse inequality is trivial. $\square$

It should be noted that the equality in (ii) does not extend to arbitrary increasing bounded nets, i.e., the membership in the range of a lifting is essential. For example, one can take a net of functions on $[0,1]$ with finite supports on which these functions equal 1, such that the supremum of this net equals 1 at every point, but the integrals are all zero.

10.5.6. Theorem. *Let μ be a Radon measure on a topological space X and let π be a μ-measurable mapping from X to a measurable space $(Y,\mathcal{E})$. Then, there exist Radon conditional measures on X, i.e., there exists a mapping $(B,y)\mapsto \mu(B,y)$, $\mathcal{B}(X)\times Y\to \mathbb{R}^1$, with the following properties*:

(1) *for every $y\in Y$, the set function $B\mapsto \mu(B,y)$ is a Radon measure on X;*

(2) *for every $B\in\mathcal{B}(X)$, the function $y\mapsto\mu(B,y)$ is measurable with respect to the measure $\nu := |\mu|\circ\pi^{-1}$;*

(3) *for all $B\in\mathcal{B}(X)$ and $E\in\mathcal{E}$, one has*

$$\int_E \mu(B,y)\,\nu(dy) = \mu\bigl(B\cap\pi^{-1}(E)\bigr). \tag{10.5.2}$$

Proof. Suppose first that μ is a probability measure and X is compact. For every $\varphi\in C(X)$, let

$$\mu_\varphi(E) = \int_{\pi^{-1}(E)} \varphi(x)\,\mu(dx), \quad E\in\mathcal{E}.$$

The measure μ_φ is absolutely continuous with respect to ν, the mapping $\varphi\mapsto\mu_\varphi$ is linear, and one has the estimate

$$|\mu_\varphi|(E)\le \|\varphi\|_\infty \nu(E).$$

Denote by $p(\varphi,\cdot)$ the Radon–Nikodym density of μ_φ with respect to ν. By the above estimate, the norm of $p(\varphi,\cdot)$ in $L^\infty(\nu)$ is majorized by $\|\varphi\|_\infty$. According to Theorem 10.5.4, there exists a lifting L of the space $L^\infty(\nu)$. Therefore, one can set

$$r(\varphi,\cdot) := L\bigl(p(\varphi,\cdot)\bigr).$$

By the definition of the Radon–Nikodym density and properties of liftings we obtain that for every $y\in Y$ the mapping $\varphi\mapsto r(\varphi,y)$ is a positive linear functional on the space $C(X)$, $r(1,y)=1$ and $|r(\varphi,y)|\le \sup_x|\varphi(x)|$. According to the Riesz theorem, there exist Radon probability measures $\mu(\,\cdot\,,y)$ on the compact space X such that

$$\int_X \varphi(x)\,\mu(dx,y) = r(\varphi,y).$$

We recall that the function $r(\varphi,\cdot)$ represents the equivalence class of the density of the measure μ_φ with respect to ν.

We verify that the family of measures $\mu(\cdot,y)$ has the required properties. Let $\mathcal{F}$ denote the class of all bounded Borel functions φ on X for which the function

$$y \mapsto \int_X \varphi(x)\,\mu(dx,y)$$

on Y is measurable with respect to the Lebesgue completion of ν and for every $E\in\mathcal{E}$ one has the equality

$$\int_E\int_X \varphi(x)\,\mu(dx,y)\,\nu(dy) = \int_{\pi^{-1}(E)} \varphi(x)\,\mu(dx). \tag{10.5.3}$$

By construction, this class contains $C(X)$. In addition, it is a linear space that is closed with respect to pointwise convergence of uniformly bounded sequences, i.e., if $\varphi_n\in\mathcal{F}$, $|\varphi_n|\le C$, $\varphi(x)=\lim\limits_{n\to\infty}\varphi_n(x)$, then $\varphi\in\mathcal{F}$. Let us verify that the indicator functions of open sets belong to $\mathcal{F}$. Let U be open in X. Set

$$\Psi=\bigl\{\psi\in C(X)\colon\ 0\le\psi\le I_U\bigr\},\quad \Psi^*=\bigl\{r(\psi,\,\cdot\,)\colon\ \psi\in\Psi\bigr\}.$$

The subset Ψ^* in the lattice $\mathcal{L}=L\bigl(L^\infty(\nu)\bigr)$ is bounded from above by the unit function. We observe that for every $y\in Y$, in view of the Radon property of $\mu(\,\cdot\,,y)$ one has

$$\mu(U,y)=\sup\bigl\{r(\psi,y)\colon\ \psi\in\Psi\bigr\}. \tag{10.5.4}$$

Indeed, given $\varepsilon>0$, there exists a compact set K in U with $\mu(U\backslash K,y)<\varepsilon$. Since X is compact, there exists a continuous function $\psi\colon X\to[0,1]$ that equals 1 on K and 0 outside U. By the definition of $\mu(\,\cdot\,,\,\cdot\,)$, we have

$$r(\psi,y)=\int_X\psi(x)\,\mu(dx,y)\ge\mu(K,y)\ge\mu(U,y)-\varepsilon.$$

In view of the inequality $r(\psi,y)\le\mu(U,y)$, we arrive at (10.5.4). By Lemma 10.5.5 (or Lemma 7.2.6), the function $y\mapsto\mu(U,y)$ is measurable with respect to the Lebesgue completion of ν. Let us fix a set $E\in\mathcal{E}$ and verify the validity of formula (10.5.3). Since the measure $I_{\pi^{-1}(E)}\cdot\mu$ is Radon, one has

$$\mu\bigl(U\cap\pi^{-1}(E)\bigr)=\sup\biggl\{\int_X I_E\bigl(\pi(x)\bigr)\psi(x)\,\mu(dx)\colon\ \psi\in\Psi\biggr\},$$

which equals

$$\sup\biggl\{\int_Y I_E(y)r(\psi,y)\,\nu(dy)\colon\ \psi\in\Psi\biggr\}$$

because $(\psi\cdot\mu)\circ\pi^{-1} = r(\psi,\,\cdot\,)\cdot\nu$. On the other hand, applying Lemma 10.5.5 to the family of functions $\{r(\psi,\,\cdot\,)\colon\ \psi\in\Psi\}$ on the space Y with the measure $I_E\cdot\nu$, we obtain

$$\int_Y I_E(y)\,\mu(U,y)\,\nu(dy) = \sup\biggl\{\int_Y I_E(y)r(\psi,y)\,\nu(dy)\colon\ \psi\in\Psi\biggr\}.$$

Thus, (10.5.3) is verified. By Theorem 2.12.9, the class $\mathcal{F}$ coincides with the collection of all bounded Borel functions. In particular, for every $B\in\mathcal{B}(X)$, the function $y\mapsto\mu(B,y)$ is measurable with respect to ν. In addition, one has (10.5.2).

We observe that if μ is a nonnegative (but not probability) measure, then applying the above construction to the corresponding normalized measure, we obtain the required representation, where all conditional measures are probabilities.

Now we consider the case where μ is still a probability measure, but the space X is arbitrary. We choose an increasing sequence of compact sets K_n with $\mu(K_n)\to 1$ and let $S_n = K_n\backslash K_{n-1}$, $S_1 = K_1$. Let $\mu_n = I_{S_n}\cdot\mu$ and let ϱ_n be the Radon–Nikodym density of the measure $\mu_n\circ\pi^{-1}$ with respect to $\nu=\mu\circ\pi^{-1}$. Let us apply the case considered to every measure μ_n considered on the compact set K_n. We denote the corresponding conditional measures on K_n by $\mu_n(\,\cdot\,,y)$. We observe that $\sum_{n=1}^\infty\mu_n=\mu$ and $\sum_{n=1}^\infty\varrho_n = 1$ ν-a.e. Letting $\mu(B,y) = \sum_{n=1}^\infty\varrho_n(y)\mu_n(B\cap S_n,y)$, we obtain (10.5.2). To this end, it suffices to observe that this equality is true if $B\subset S_n$. Indeed,

$$\begin{aligned}\int_E\mu(B,y)\,\nu(dy) = \int_E\varrho_n(y)\mu_n(B,y)\,\nu(dy) &= \int_E\mu_n(B,y)\,\mu_n\circ\pi^{-1}(dy)\\ &= \mu_n\bigl(\pi^{-1}(E)\cap B\bigr) = \mu\bigl(\pi^{-1}(E)\cap B\bigr).\end{aligned}$$

In the general case, we apply the already-proven assertions to the measures μ^+ and μ^- that yield two families of conditional probability measures $\mu'(\,\cdot\,,\,\cdot\,)$ and $\mu''(\,\cdot\,,\,\cdot\,)$, respectively. Let ϱ^+ and ϱ^- be the Radon–Nikodym densities of the measures $\mu^+\circ\pi^{-1}$ and $\mu^-\circ\pi^{-1}$ with respect to $\nu=|\mu|\circ\pi^{-1}$. Letting

$$\mu(B,y) = \varrho^+(y)\mu'(B,y) - \varrho^-(y)\mu''(B,y),$$

we arrive at the desired representation. $\square$

10.5.7. Corollary. *Suppose that under the hypotheses of Theorem* 10.5.6 *the graph of* π *belongs to* $\mathcal{B}(X)\otimes\mathcal{E}$. *Then, the conditional probability* $\mu(\,\cdot\,,\,\cdot\,)$ *has the following property: for* ν*-almost every* $y\in Y$, *the measure* $\mu(\,\cdot\,,y)$ *is concentrated on the set* $\pi^{-1}(y)$ (*and all such sets are Borel*).

Proof. There exist $\{B_n\}\subset\mathcal{B}(X)$ and $\{E_n\}\subset\mathcal{E}$ such that the graph Γ_π of π belongs to the σ-algebra generated by the sets $B_n\times E_n$. The sets $\pi^{-1}(y)$ belong to $\sigma(\{B_n\})$ because $I_{\Gamma_\pi} = \varphi(I_{B_1}I_{E_1}, I_{B_2}I_{E_2},\ldots)$, where φ is a Borel function on $\mathbb{R}^\infty$. Hence the reasoning from Corollary 10.4.10 is applicable. $\square$

This result yields easily the already-known assertion from Example 10.4.11 on conditional measures in the case of measurable mappings of Souslin spaces.

10.6. Disintegrations of measures

In this section, we discuss certain generalizations of conditional measures. The principal difference as compared to our previous setting is that now conditional measures will be defined on different σ-algebras.

Let $(X, \mathfrak{F}, \mu)$ be a probability space, let $\mathfrak{B} \subset \mathfrak{F}$ be a sub-σ-algebra, and let $\mathfrak{B} \cap E = \mathfrak{B}_E$ denote the restriction of $\mathfrak{B}$ to $E \subset X$.

10.6.1. Definition. *Suppose that for each $x \in X$ we are given a sub-σ-algebra $\mathfrak{F}_x \subset \mathfrak{F}$ and a measure $\mu(\,\cdot\,, x)$ on $\mathfrak{F}_x$ satisfying the following conditions*:

(i) *for every $A \in \mathfrak{F}$, there exists a set $N_A \in \mathfrak{B}$ such that $\mu(N_A) = 0$ and $A \in \mathfrak{F}_x$ for all $x \notin N_A$, and the function $x \mapsto \mu(A, x)$ on $X \backslash N_A$ is measurable with respect to $\mathfrak{B} \cap (X \backslash N_A)$ and μ-integrable*;

(ii) *for all $A \in \mathfrak{F}$ and $B \in \mathfrak{B}$ one has*

$$\int_B \mu(A, x)\, \mu(dx) = \mu(A \cap B).$$

Then we shall say that the measures $\mu(\,\cdot\,, x)$ give a disintegration of the measure μ with respect to $\mathfrak{B}$ and call these measures conditional measures.

It is clear that if there exist regular conditional measures $\mu(\,\cdot\,, x)$ with respect to $\mathfrak{B}$, then they give a disintegration, and one can let $\mathfrak{F}_x = \mathfrak{F}$ for all x. The difference between disintegrations and regular conditional measures is that, in the first place, the measures $\mu(\,\cdot\,, x)$ may be defined on different σ-algebras, and, secondly, the condition of $\mathfrak{B}$-measurability of $\mu(A, x)$ is weakened at the expense of admitting sets N_A of measure zero. As we shall see below, these distinctions lead indeed to a more general object. However, we shall show first that in the case of a countably generated σ-algebra $\mathfrak{F}$, the existence of a disintegration is equivalent to the existence of conditional measures with respect to $\mathfrak{B}$ (we recall that conditional measures do not always exist even for countably generated σ-algebras, see Example 10.4.19). Somewhat different disintegrations are considered below in §10.10(ii).

10.6.2. Proposition. *Suppose that $\mathfrak{F}$ is a countably generated σ-algebra. Then, the existence of a disintegration with respect to a σ-algebra $\mathfrak{B} \subset \mathfrak{F}$ is equivalent to the existence of a regular conditional measure with respect to $\mathfrak{B}$.*

Proof. If we have a regular conditional measure, then we have a disintegration. Let us show the converse. Suppose that measures $\mu(\,\cdot\,, x)$ on $\mathfrak{F}_x \subset \mathfrak{F}$ give a disintegration of the measure μ on $\mathfrak{F}$ with respect to $\mathfrak{B}$ and construct a new disintegration $\mu_1(\,\cdot\,, x)$ such that all conditional measures are defined on $\mathfrak{F}$. Let $\mathfrak{R}$ be a countable algebra generating $\mathfrak{F}$. For every $A_i \in \mathfrak{R}$, we find a measure zero set $N_{A_i} \subset \mathfrak{B}$ with the properties from Definition 10.6.1. We may assume that $\mu(A_i, x) \geq 0$ and $\mu(X, x) = 1$ if $x \notin N_{A_i}$ since $\mu(A_i, x) \geq 0$ a.e. and $\mu(X, x) = 1$ a.e. by the identity in (ii) in the definition. Let $N = \bigcup_{i=1}^\infty N_{A_i}$. It is clear that $N \in \mathfrak{B}$ and $\mu(N) = 0$. By Lemma 10.4.3 we may assume that $\mu(\,\cdot\,, x)$ is a probability measure for each

$x \in X\backslash N$. Let us consider the class of all sets $E \in \mathfrak{F}$ such that $E \in \mathfrak{F}_x$ for all $x \in X\backslash N$ and the function $x \mapsto \mu(E, x)$ on $X\backslash N$ is measurable with respect to $\mathfrak{B} \cap (X\backslash N)$. It is clear that this class contains $\mathfrak{R}$ and is monotone. Therefore, it coincides with $\mathfrak{F}$. Now we let $\mu_1(A, x) = \mu(A, x)$ if $x \notin N$ and $A \in \mathfrak{F}$. This is possible because $A \in \mathfrak{F}_x$ if $x \notin N$. If $x \in N$, then we set $\mu_1(A, x) = \mu(A)$. Since $N \in \mathfrak{B}$, it follows that for every $A \in \mathfrak{F}$, the function $x \mapsto \mu_1(A, x)$ is measurable with respect to $\mathfrak{B}$. Finally, if $B \in \mathfrak{B}$, then by the equality $\mu(N) = 0$, the integral of $\mu_1(A, x)$ over B equals the integral of $\mu(A, x)$ over B, hence equals $\mu(A \cap B)$. □

10.6.3. Remark. There is yet another definition of conditional measures that is intermediate between regular conditional measures and disintegrations. We shall say that a family of measures $\mu(\,\cdot\,, x)$ on $\mathfrak{F}$ gives for the measure μ conditional measures with respect to $\mathfrak{B} \subset \mathfrak{F}$ in the sense of Doob if in Definition 10.4.1 of regular conditional measures in place of the $\mathfrak{B}$-measurability of functions $x \mapsto \mu(A, x)$ with $A \in \mathfrak{F}$, we require only their $\mathfrak{B}_\mu$-measurability. The connection between conditional measures in the sense of Doob and disintegrations consists in the following: the existence of conditional measures with respect to $\mathfrak{B}$ in the sense of Doob is equivalent to the existence of a disintegration $\mu(\,\cdot\,, x)$ with $\mathfrak{F}_x = \mathfrak{F}$ for all $x \in X$ (Exercise 10.10.49).

Now we give an example (using the continuum hypothesis) that shows that the existence of a disintegration does not guarantee the existence of regular conditional measures.

10.6.4. Example. Let $X = [0,1]^2 \times [0,1]$, $\mathfrak{F} = \mathfrak{L}_2 \otimes \mathfrak{L}_1$, where $\mathfrak{L}_2$ and $\mathfrak{L}_1$ are the σ-algebras of Lebesgue measurable sets in $[0,1]^2$ and $[0,1]$, respectively. Let

$$\mu(A) = \lambda_2\bigl((x_1, x_2)\colon\ (x_1, x_2, x_2) \in A\bigr), \quad A \in \mathfrak{F},$$

where λ_2 is Lebesgue measure on $[0,1]^2$. Let us take for $\mathfrak{B}$ the σ-algebra of all cylinders $B = [0,1]^2 \times B_0$, $B_0 \in \mathfrak{L}_1$. Then the set of all compact sets in X is a compact approximating class for μ on $\mathfrak{F}$ and there exists a disintegration of μ with respect to $\mathfrak{B}$. However, under the continuum hypothesis, one cannot choose conditional measures $\mu(\,\cdot\,, x)$ such that for μ-a.e. x, the measure $\mu(\,\cdot\,, x)$ be defined on $\mathfrak{F}$.

PROOF. We observe that the measure μ is well-defined on $\mathfrak{F}$, since the mapping $\psi\colon (x_1, x_2) \mapsto (x_1, x_2, x_2)$ is measurable with respect to the σ-algebras $\mathfrak{L}_2$ and $\mathfrak{L}_2 \otimes \mathfrak{L}_1$ because $\psi^{-1}(A_2 \times A_1) = A_2 \cap ([0,1] \times A_1) \in \mathfrak{L}_2$ for all $A_2 \in \mathfrak{L}_2$ and $A_1 \in \mathfrak{L}_1$. According to our definition, $\mu = \lambda_2 \circ \psi^{-1}$. It is readily seen that the measure μ is approximated by the class of all compact sets. As shown in Theorem 10.6.6 below, this implies the existence of a disintegration with respect to $\mathfrak{B}$. Let us show that one cannot have almost all conditional measures defined on $\mathfrak{F}$. Suppose the contrary. We show that there exists a set $M \in \mathfrak{B}$ such that $\mu(M) = 0$ and, for all $x \notin M$ and all Borel sets $E \subset [0,1]^2$, one has the equality

$$\mu(E \times [0,1], x) = \lambda_1(E_{x_3}),\ x = (x_1, x_2, x_3),\ E_{x_3} = \{t\colon\ (t, x_3) \in E\}. \quad (10.6.1)$$

Since $\mathcal{B}([0,1]^2)$ is a countably generated σ-algebra and both sides of (10.6.1) are measures as set functions on $E \in \mathcal{B}([0,1]^2)$, it suffices to verify that there exists a set $M \in \mathfrak{B}$ of μ-measure zero such that (10.6.1) is true for all $x \notin M$ and every set E in some countable algebra generating $\mathcal{B}([0,1]^2)$. Hence it suffices to show that for any fixed set E, equality (10.6.1) is true $\mu|_{\mathfrak{B}}$-a.e. In turn, it suffices to show that the integrals of both sides of (10.6.1) over every set $B \in \mathfrak{B}$ coincide. Since $B = [0,1]^2 \times B_0$, the integral over B of the left-hand side is

$$\mu\bigl(B \cap (E \times [0,1])\bigr) = \lambda_2\bigl(E \cap ([0,1] \times B_0)\bigr) = \int_{B_0} \lambda_1(E_{x_3})\, \lambda_1(dx_3)$$

by the definition of a disintegration and Fubini's theorem. It remains to observe that

$$\int_{B_0} \lambda_1(E_{x_3})\, \lambda_1(dx_3) = \int_{[0,1]^2 \times B_0} \lambda_1(E_{x_3})\, \mu(dx),$$

since $\lambda_1(E_{x_3})$ does not depend on (x_1, x_2) and the image of μ under the mapping $(x_1, x_2, x_3) \mapsto x_3$ is λ_1 (the latter is easily verified). Thus, we obtain a required set M. Now let $x = (x_1, x_2, x_3) \notin M$. We observe that for an arbitrary set $C \subset [0,1]$, the set $C \times \{x_3\} \times [0,1]$ belongs to $\mathfrak{F}$, since we have $\lambda_2(C \times \{x_3\}) = 0$. Due to our assumption that almost all conditional measures are defined on $\mathfrak{F}$, there exists at least one point $x \notin M$ for which the set function $\overline{\lambda}(C) = \mu(C \times \{x_3\} \times [0,1], x)$ is defined on the class of all sets $C \subset [0,1]$. It is clear that $\overline{\lambda}$ is a countably additive measure that vanishes on all singletons by (10.6.1). According to Corollary 1.12.41 we have $\overline{\lambda} = 0$, a contradiction. □

Let us consider one more close example, in which, however, the σ-algebra $\mathfrak{B}$ is not complete with respect to μ.

10.6.5. Example. Assume the continuum hypothesis. Let $X = [0,1]^2$, let $\mathfrak{F}$ be the σ-algebra of all Lebesgue measurable sets in $[0,1]^2$, let μ be Lebesgue measure on $[0,1]^2$, and let $\mathfrak{B}$ be the σ-algebra generated by the projection to the first coordinate, i.e., the collection of all sets of the form $B = B_0 \times [0,1]$, $B_0 \in \mathcal{B}([0,1])$. Then, one cannot choose conditional measures $\mu(\,\cdot\,, x)$ in such a way that for each x, the measure $\mu(\,\cdot\,, x)$ be defined on $\mathfrak{F}$.

PROOF. Suppose that such conditional measures exist. The functions $\mu(A, x)$ depend only on the first coordinate x_1 of the point $x \in [0,1]^2$. Hence we may denote them by $\lambda(A, x_1)$. Similarly to Corollary 10.4.10 one verifies that for almost all x_1 one has the equality $\lambda(\{x_1\} \times [0,1], x_1) = 1$. Let $\{B_n\}$ be the set of all rational intervals in $[0,1]$ and λ Lebesgue measure on $[0,1]$. For every $B \in \mathcal{B}([0,1])$, we have

$$\lambda(B_n)\lambda(B) = \mu\Bigl(\bigl([0,1] \times B_n\bigr) \cap \bigl(B \times [0,1]\bigr)\Bigr) = \int_B \lambda\bigl([0,1] \times B_n, x_1\bigr)\, \lambda(dx_1),$$

whence it follows that $\lambda([0,1] \times B_n, x_1) = \lambda(B_n)$ for almost all x_1. This means that for almost all x_1 the measure $\lambda(\,\cdot\,, x_1)$ on the class of Borel sets is the natural Lebesgue measure on the interval $\{x_1\}\times[0,1]$. Let x_1 be such a value. Then $\lambda(\,\cdot\,,x_1)$ gives a countably additive extension of Lebesgue measure to all subsets of $\{x_1\}\times[0,1]$, since such sets have zero measure in the square and hence belong to $\mathfrak{F}$. This contradicts the continuum hypothesis. □

10.6.6. Theorem. *Let $(X,\mathfrak{F},\mu)$ be a probability space and let $\mathfrak{B}$ be a sub-σ-algebra in $\mathfrak{F}$ such that the measure $\mu|_{\mathfrak{B}}$ is complete. Suppose that there exists a compact class $\mathcal{K}\subset\mathfrak{F}$ that is closed with respect to finite unions and countable intersections, contains $\varnothing$ and approximates μ. Then, there is a disintegration $\{\mathfrak{F}_x,\mu(\,\cdot\,,x)\}_{x\in X}$ with respect to $\mathfrak{B}$ such that for every $x\in X$, $\mu(\,\cdot\,,x)$ is a probability measure, and the class $\mathcal{K}$ belongs to $\mathfrak{F}_x$ and approximates $\mu(\,\cdot\,,x)$ on $\mathfrak{F}_x$.*

PROOF. Let L be a lifting on $(X,\mathfrak{B},\mu|_{\mathfrak{B}})$. For every $A\in\mathfrak{F}$, we fix some version of the conditional expectation $\mathbb{E}^{\mathfrak{B}}I_A$ with respect to $\mathfrak{B}$. Set

$$\beta_x(K)=L(\mathbb{E}^{\mathfrak{B}}I_K)(x),\quad K\in\mathcal{K}.$$

It follows by the properties of conditional expectations and liftings that β_x is a monotone modular function with $\beta_x(X)=1$. According to Lemma 1.12.38, for every x, there exists a monotone modular function ζ_x on $\mathcal{K}$ with $\zeta_x\ge\beta_x$, $\zeta_x(X)=1$, and $\zeta_x(K)+(\zeta_x)_*(X\backslash K)=1$, $\forall\, K\in\mathcal{K}$. Let

$$\mathfrak{F}_x=\Bigl\{E\in\mathfrak{F}\colon\ (\zeta_x)_*(E)+(\zeta_x)_*(X\backslash E)=1\Bigr\}.$$

Denote by $\mu(\,\cdot\,,x)$ the restriction of $(\zeta_x)_*$ to $\mathfrak{F}_x$. By Corollary 1.12.39, $\mathfrak{F}_x$ is a σ-algebra and $\mu(\,\cdot\,,x)$ is a countably additive measure on $\mathfrak{F}_x$, in addition, the class $\mathcal{K}$ is contained in $\mathfrak{F}_x$ and approximates the measure $\mu(\,\cdot\,,x)$. It remains to verify that we have obtained a disintegration. Let $A\in\mathfrak{F}$. We find two increasing sequences $\{K_n\}$, $\{L_n\}$ in $\mathcal{K}$ such that $K_n\subset A$, $L_n\subset X\backslash A$, $\mu(K_n)\to\mu(A)$, and $\mu(L_n)\to\mu(X\backslash A)$. For every $B\in\mathfrak{B}$, we have

$$\begin{aligned}\mu(B\cap A)&=\lim_{n\to\infty}\mu(B\cap K_n)=\lim_{n\to\infty}\int_B\mathbb{E}^{\mathfrak{B}}I_{K_n}(x)\,\mu(dx)\\&=\lim_{n\to\infty}\int_B\beta_x(K_n)\,\mu(dx)=\int_B\lim_{n\to\infty}\beta_x(K_n)\,\mu(dx).\end{aligned}\tag{10.6.2}$$

Similarly, we verify that

$$\mu(B\cap A)=\mu(B)-\mu\bigl(B\cap(X\backslash A)\bigr)=\int_B\lim_{n\to\infty}\bigl[1-\beta_x(L_n)\bigr]\,\mu(dx).\tag{10.6.3}$$

We observe that for every x, one has the inequalities

$$\lim_{n\to\infty}\beta_x(K_n)\le(\zeta_x)_*(A)\le1-(\zeta_x)_*(X\backslash A)\le1-\lim_{n\to\infty}\beta_x(L_n).$$

Hence (10.6.2) and (10.6.3) yield that for $\mu|_{\mathfrak{B}}$-a.e. x, one has the equalities

$$(\zeta_x)_*(A)=1-(\zeta_x)_*(X\backslash A)\quad\text{and}\quad\int_B(\zeta_x)_*(A)\,\mu(dx)=\mu(B\cap A).$$

Thus, for $\mu|_{\mathfrak{B}}$-a.e. x, we obtain that $A \in \mathfrak{F}_x$ and

$$\mu(A, x) = (\zeta_x)_*(A) = \lim_{n\to\infty} \beta_x(K_n).$$

In particular, the function $\mu(A, x)$ is measurable with respect to the measure $\mu|_{\mathfrak{B}}$. Finally, one has equality (10.6.2). $\square$

10.6.7. Corollary. *Let $(X, \mathfrak{F}, \mu)$ be a probability space such that μ has a compact approximating class. Then, for every sub-σ-algebra $\mathfrak{B} \subset \mathfrak{F}$, there exists a disintegration $\bigl\{\mathfrak{F}_x, \mu(\,\cdot\,, x)\bigr\}_{x\in X}$ with respect to $\mathfrak{B}$ such that $\mu(\,\cdot\,, x)$ is a probability measure with a compact approximating class in $\mathfrak{F}_x$ for every x.*

PROOF. According to Proposition 1.12.4, there is a compact class $\mathcal{K} \subset \mathfrak{F}$ that approximates μ and is closed with respect to finite unions and countable intersections. The class $\mathcal{K}$ is approximating for the completed σ-algebra $\mathfrak{F}_\mu$ as well. Let $\mathfrak{B}_\mu$ be the completion of $\mathfrak{B}$ with respect to $\mu|_{\mathfrak{B}}$. By the above theorem the measure μ on $\mathfrak{F}_\mu$ has a disintegration $\bigl\{\overline{\mathfrak{F}}_x, \overline{\mu}(\,\cdot\,, x)\bigr\}_{x\in X}$ with respect to $\mathfrak{B}_\mu$ such that $\mathcal{K}$ is contained in $\overline{\mathfrak{F}}_x \subset \mathfrak{F}_\mu$ and approximates $\overline{\mu}(x, \,\cdot\,)$ for all x. Let $\mathfrak{F}_x = \overline{\mathfrak{F}}_x \cap \mathfrak{F}$ and $\mu(\,\cdot\,, x) = \overline{\mu}(\,\cdot\,, x)|_{\mathfrak{F}_x}$. We verify that this is a required disintegration. Let $A \in \mathfrak{F}$. Let us take a set $N \in \mathfrak{B}_\mu$ of μ-measure zero such that for each $x \notin N$ the set A belongs to $\overline{\mathfrak{F}}_x$ and the function $\overline{\mu}(A, x)$ on $X\backslash N$ is measurable with respect to $\mathfrak{B}_\mu \cap (X\backslash N)$. Next we find a set $M \in \mathfrak{B}$ containing N and having μ-measure zero such that the function $\overline{\mu}(A, x)$ on $X\backslash M$ is measurable with respect to $\mathfrak{B} \cap (X\backslash M)$. Thus, for each $x \notin M$, we have $A \in \mathfrak{F}_x$ and the function $\overline{\mu}(A, x)$ on $X\backslash M$ is measurable with respect to $\mathfrak{B} \cap (X\backslash M)$. In addition, for all x, the class $\mathcal{K}$ is contained in $\mathfrak{F}_x$ and approximates $\mu(\,\cdot\,, x)$ on $\mathfrak{F}_x$. Finally, it is clear that for any $B \in \mathfrak{B}$, the integral of $\mu(A, x)$ over B coincides with the integral of $\overline{\mu}(A, x)$ and equals $\mu(B \cap A)$. $\square$

10.7. Transition measures

Conditional measures provide an example of transition measures, which we discuss in greater detail in this section.

10.7.1. Definition. *Let $(X_1, \mathcal{B}_1)$ and $(X_2, \mathcal{B}_2)$ be a pair of measurable spaces. A transition measure for this pair is a function $P(\,\cdot\,|\,\cdot\,)\colon\ X_1 \times \mathcal{B}_2 \to \mathbb{R}^1$ with the following properties*:

(i) *for every fixed $x \in X_1$, the function $B \mapsto P(x|B)$ is a measure on $\mathcal{B}_2$*;

(ii) *for every fixed $B \in \mathcal{B}_2$, the function $x \mapsto P(x|B)$ is measurable with respect to $\mathcal{B}_1$.*

In the case where transition measures are probabilities in the second argument, they are called *transition probabilities.*

10.7.2. Theorem. *Let $P(\,\cdot\,|\,\cdot\,)$ be a transition probability for spaces $(X_1, \mathcal{B}_1)$ and $(X_2, \mathcal{B}_2)$ and let ν be a probability measure on $\mathcal{B}_1$. Then, there* *exists a unique probability measure* μ *on* $(X_1 \times X_2, \mathcal{B}_1 \otimes \mathcal{B}_2)$ *with*

$$\mu(B_1 \times B_2) = \int_{B_1} P(x|B_2)\,\nu(dx), \quad \forall\, B_1 \in \mathcal{B}_1,\ B_2 \in \mathcal{B}_2. \tag{10.7.1}$$

In addition, given any function $f \in \mathcal{L}^1(\mu)$, *for* ν*-a.e.* $x_1 \in X_1$, *the function* $x_2 \mapsto f(x_1, x_2)$ *on* X_2 *is measurable with respect to the completed* σ*-algebra* $(\mathcal{B}_2)_{P(x_1|\,\cdot\,)}$ *and* $P(x_1|\,\cdot\,)$*-integrable, the function*

$$x_1 \mapsto \int_{X_2} f(x_1, x_2)\, P(x_1|dx_2)$$

is measurable with respect to $(\mathcal{B}_1)_\nu$ *and* ν*-integrable, and one has*

$$\int_{X_1 \times X_2} f(x_1, x_2)\, \mu\bigl(d(x_1, x_2)\bigr) = \int_{X_1} \int_{X_2} f(x_1, x_2)\, P(x_1|dx_2)\, \nu(dx_1). \tag{10.7.2}$$

Proof. In order to prove the first assertion, it suffices to show that the nonnegative set function μ defined by the right-hand side of (10.7.1) on the semialgebra of rectangles is countably additive. Let $A \times B = \bigcup_{n=1}^\infty A_n \times B_n$, where $A, A_n \in \mathcal{B}_1$, $B, B_n \in \mathcal{B}_2$, and $A_n \times B_n$ are pairwise disjoint. This means that $I_A(x_1) I_B(x_2) = \sum_{n=1}^\infty I_{A_n}(x_1) I_{B_n}(x_2)$. By using the countable additivity of $P(x_1|\,\cdot\,)$ and interchanging the summation and integration we obtain

$$I_A(x_1) P(x_1|B) = \sum_{n=1}^\infty I_{A_n}(x_1) P(x_1|B_n).$$

Integrating against the measure ν, we obtain $\mu(A \times B) = \sum_{n=1}^\infty \mu(A_n \times B_n)$, as required. Now we prove that for every set $E \in \mathcal{B}_1 \otimes \mathcal{B}_2$, the function $x_1 \mapsto P(x_1|E_{x_1})$, where $E_{x_1} = \{x_2\colon (x_1, x_2) \in E\}$, is measurable with respect to $\mathcal{B}_1$. To this end, we observe that $E_{x_1} \in \mathcal{B}_2$ according to Proposition 3.3.2 and that the class $\mathcal{E}$ of all sets $E \in \mathcal{B}_1 \otimes \mathcal{B}_2$ with the property to be proven is an algebra and by definition contains all rectangles. It is clear that the class $\mathcal{E}$ is closed with respect to formation of unions of increasing sequences. Therefore, $\mathcal{E}$ is a σ-algebra that coincides with $\mathcal{B}_1 \otimes \mathcal{B}_2$.

It follows that for every bounded $\mathcal{B}_1 \otimes \mathcal{B}_2$-measurable function f, the function

$$\widehat{f}\colon\ x_1 \mapsto \int_{X_2} f(x_1, x_2)\, P(x_1|dx_2)$$

is measurable with respect to $\mathcal{B}_1$ and one has (10.7.2).

Now let f be a nonnegative $\mathcal{B}_1 \otimes \mathcal{B}_2$-measurable function that is integrable with respect to μ. We consider the functions $f_n = \min(f, n)$ and obtain that the corresponding functions $\widehat{f_n}$ are measurable with respect to $\mathcal{B}_1$ and equality (10.7.2) is fulfilled for them. By the monotone convergence theorem the function $\widehat{f}$ is μ-integrable as well and satisfies (10.7.2). Thus, the second assertion of the theorem is true for all $\mathcal{B}_1 \otimes \mathcal{B}_2$-measurable μ-integrable functions.

Finally, we extend the result to all functions $f \in \mathcal{L}^1(\mu)$. As in the previous step, it suffices to do this for bounded functions. In turn, it suffices to consider the indicator of a μ-measurable set E. The set E is the union of a set E_0 in $\mathcal{B}_1 \otimes \mathcal{B}_2$ and some set C with $\mu(C) = 0$. It remains to observe that for ν-a.e. x_1, the set $C_{x_1} = \{x_2\colon (x_1, x_2) \in C\}$ has $P(x_1|\,\cdot\,)$-measure zero. This follows from the fact that there exists a set $D \in \mathcal{B}_1 \otimes \mathcal{B}_2$ such that $C \subset D$ and $\mu(D) = 0$. Indeed, $C_{x_1} \subset D_{x_1}$ for all $x_1 \in X_1$ and $P(x_1|D_{x_1}) = 0$ for ν-a.e. x_1 according to (10.7.2). □

It is clear that this theorem extends to signed measures if the function $\|P(x|\,\cdot\,)\|$ is integrable with respect to $|\nu|$.

Now we prove the following theorem of Ionescu Tulcea, which is useful in the theory of random processes.

10.7.3. Theorem. *Let $(X_n, \mathcal{B}_n)$, $n = 0, 1, \ldots$, be measurable spaces such that for every $n = 0, 1, \ldots$, we are given a transition probability $P_{n+1}^{0,\ldots,n}$ for the pair of spaces*

$$\Bigl(\prod_{k=0}^{n} X_k, \bigotimes_{k=0}^{n} \mathcal{B}_k\Bigr) \quad \textit{and} \quad (X_{n+1}, \mathcal{B}_{n+1}).$$

Then, for every $x_0 \in X_0$, there exists a unique probability measure P_{x_0} on the measurable space $(X, \mathcal{B}) = \Bigl(\prod_{n=0}^{\infty} X_n, \bigotimes_{n=0}^{\infty} \mathcal{B}_n\Bigr)$ such that for all $B_k \in \mathcal{B}_k$

$$P_{x_0}\Bigl(\prod_{k=0}^{n} B_k\Bigr) = \int_{B_1} \cdots \int_{B_n} P_n^{0,\ldots,n-1}(x_0, \ldots, x_{n-1}|dx_n) \qquad (10.7.3)$$
$$\cdots P_2^{0,1}(x_0, x_1|dx_2)\, P_1^{0}(x_0|dx_1)\, I_{B_0}(x_0).$$

Proof. Suppose first that we are given a finite sequence of spaces X_k and transition probabilities $P_{k+1}^{0,\ldots,k}$, $k = 0, 1, \ldots, N$. We define probabilities $P_{x_0,\ldots,x_k}$ on $\prod_{j=0}^{N} X_j$ by the recursive formulas (in the order of decreasing indices k)

$$P_{x_0,\ldots,x_N}(A) = I_A(x_0, \ldots, x_N),$$
$$P_{x_0,\ldots,x_k}(A) = \int_{X_{k+1}} P_{x_0,\ldots,x_k,x_{k+1}}(A) P_{k+1}^{0,\ldots,k}(x_0, \ldots, x_k|dx_{k+1}).$$

It is easy to see that $P_{x_0,\ldots,x_k}$ is a probability measure on $\bigotimes_{j\le N} \mathcal{B}_j$, and for every set $A \in \bigotimes_{j\le N} \mathcal{B}_j$, the function $P_{x_0,\ldots,x_k}(A)$ is $\bigotimes_{j\le k} \mathcal{B}_j$-measurable with respect to $(x_0, \ldots, x_k)$, and for any fixed $(x_0, \ldots, x_{k-1})$, it is $\mathcal{B}_k$-measurable with respect to x_k. It is clear that for every nonnegative $\bigotimes_{j\le N} \mathcal{B}_j$-measurable function ζ and all $k \le N$, one has the equality

$$\int \zeta\, dP_{x_0,\ldots,x_k} = \int_{X_N} \cdots \int_{X_{k+1}} \zeta(x_0, \ldots, x_N) P_N^{0,\ldots,N-1}(x_0, \ldots, x_{N-1}|dx_N)$$
$$\cdots P_{k+1}^{0,\ldots,k}(x_0, \ldots, x_k|dx_{k+1}).$$

We proceed to the infinite sequence case. As above, we shall construct probabilities $P_{x_0,\ldots,x_k}$. For every N, the construction of the previous step gives probabilities $P_{x_0,\ldots,x_k}^{(N)}$ on $\bigotimes_{j\le N} \mathcal{B}_j$. It is seen from our construction that these probabilities are consistent for different N. Thus, on the algebra $\mathfrak{A}$ obtained as the union of $\bigotimes_{j\le N}\mathcal{B}_j$, we are given set functions $P_{x_0,\ldots,x_N}$ whose restrictions to $\bigotimes_{j\le N}\mathcal{B}_j$ coincide with $P^{(N)}_{x_0,\ldots,x_k}$. It is clear that these functions are additive. If we show that they are countably additive on $\mathfrak{A}$, then their countably additive extensions to $\mathcal{B}$ will be the required probabilities. In particular, we shall have probability measures P_{x_0}. Let sets $A_n\in\mathfrak{A}$ be decreasing to $\varnothing$. Suppose that $\lim\limits_{n\to\infty} P_{y_0,\ldots,y_N}(A_n)>0$ for some N and $y_0,\ldots,y_N$. Then

$$\int_{X_{N+1}} \lim_{n\to\infty} P_{y_0,\ldots,y_N,x_{N+1}}(A_n)\, P^{0,\ldots,N}_{N+1}(y_0,\ldots,y_N|dx_{N+1}) = \lim_{n\to\infty} P_{y_0,\ldots,y_N}(A_n)>0.$$

Therefore, there exists y_{N+1} such that $\lim\limits_{n\to\infty} P_{y_0,\ldots,y_{N+1}}(A_n)>0$. By induction we find a sequence $y=(y_0,y_1,\ldots)$ with

$$\lim_{n\to\infty} P_{y_0,\ldots,y_k}(A_n)>0 \quad \text{for all } k\ge N.$$

On the other hand, for every fixed n, we have $A_n\in\bigotimes_{j\le m}\mathcal{B}_j$ for all sufficiently large m, whence one has $P_{y_0,\ldots,y_m}(A_n)=I_{A_n}(y_0,\ldots,y_m)$. Therefore, $y\in A_n$ for all n, which contradicts the fact that the intersection of the sets A_n is empty. Thus, we have established the countable additivity of the measures $P_{x_0,\ldots,x_n}$. $\square$

It follows by this theorem that for every bounded $\bigotimes_{j\le n}\mathcal{B}_j$-measurable function ζ, one has

$$\int_{X_0}\zeta(x)\,P_{x_0}(dx)=\int_{X_1}\cdots\int_{X_n}\zeta(x_0,x_1,\cdots,x_n)\,P^{0,\ldots,n-1}_{n}(x_0,\ldots,x_{n-1}|dx_n) \cdots P^{0,1}_2(x_0,x_1|dx_2)\,P^0_1(x_0|dx_1).$$

10.7.4. Corollary. *Let $(X_n,\mathcal{B}_n)$, $n=0,1,\ldots,$ be measurable spaces such that for every n, we are given a transition probability $P^{0,\ldots,n}_{n+1}$ for the spaces*

$$\Bigl(\prod_{k=0}^n X_k,\bigotimes_{k=0}^n\mathcal{B}_k\Bigr) \quad \textit{and} \quad (X_{n+1},\mathcal{B}_{n+1}).$$

Let P_0 be a probability measure on $(X_0,\mathcal{B}_0)$. Then, there exists a unique probability measure P on the space $(X,\mathcal{B}):=\Bigl(\prod_{n=0}^\infty X_n,\bigotimes_{n=0}^\infty\mathcal{B}_n\Bigr)$ satisfying

$$P\Bigl(\prod_{k=0}^n B_k\Bigr)=\int_{B_0}\int_{B_1}\cdots\int_{B_n} P^{0,\ldots,n-1}_n(x_0,\ldots,x_{n-1}|dx_n) \cdots P^{0,1}_2(x_0,x_1|dx_2)\,P^0_1(x_0|dx_1)\,P_0(dx_0).$$

As already noted above, transition measures can be constructed by using conditional measures.

Let us consider the following example of application of Theorem 10.4.14.

10.7.5. Example. Suppose that $(\Omega, \mathcal{A}, P)$ is a probability space, X is a Souslin space, $(Y, \mathcal{A}_Y)$ is a measurable space, $\xi\colon (\Omega, \mathcal{A}) \to \bigl(X, \mathcal{B}(X)\bigr)$ and $\eta\colon (\Omega, \mathcal{A}) \to (Y, \mathcal{A}_Y)$ are measurable mappings. Then there is a transition probability $(y, B) \mapsto \mu(y|B)$ on $Y \times \mathcal{B}(X)$ such that for every $B \in \mathcal{B}(X)$, one has $P(\xi \in B|\eta) = \mu(\eta|B)$ a.e. The family of measures $\mu(y|\,\cdot\,)$ is uniquely determined up to a redefinition on a set of $P\circ\eta^{-1}$-measure zero.

Proof. Let μ be the image of P under the mapping (ξ, η) with values in $X\times Y$. Set $\mu(y|B) := \mu(B, y)$, where the measures $\mu(\,\cdot\,, y)$ are constructed in the cited theorem. For any fixed $B \in \mathcal{B}(X)$ and every $E \in \mathcal{A}_Y$ we have

$$\mathbb{E}[I_E \circ \eta P(\xi \in B|\eta)] = \mathbb{E}[I_E \circ \eta I_B \circ \xi] = \int_{X\times Y} I_E(y) I_B(x)\, \mu\bigl(d(x, y)\bigr)$$
$$= \int_E \mu(B, y)\, \mu_Y(dy) = \mathbb{E}[I_E \circ \eta \mu(B, \eta)],$$

whence we obtain $P(\xi \in B|\eta) = \mu(B, \eta)$ a.e. The uniqueness assertion can be easily derived from the fact that $\mathcal{B}(X)$ is countably generated. $\square$

The following result enables one to obtain transition probabilities as distributions of random elements.

10.7.6. Proposition. *Let $(X, \mathcal{A})$ be a measurable space, let T be a Souslin space, and let $(x, B) \mapsto \mu(x|B)$ be a transition probability on $X\times\mathcal{B}(T)$. Then there exists an $\bigl(\mathcal{A}\otimes\mathcal{B}([0,1]), \mathcal{B}(T)\bigr)$-measurable mapping $f\colon\ X\times[0,1] \to T$ such that for every random variable ξ with the uniform distribution in $[0,1]$, the mapping $f(x, \xi)$ has the distribution $\mu(x|\,\cdot\,)$ for all $x \in X$.*

Proof. We may assume that $T \subset [0,1]$. Set

$$f(x, t) = \sup\bigl\{r \in [0,1]\colon\ \mu(x|[0,r]) < t\bigr\}.$$

The function f is measurable, since the indicated supremum can be taken over all rational numbers r, and the function $(x, t) \mapsto \mu(x|[0,r]) - t$ is measurable with respect to $\mathcal{A}\otimes\mathcal{B}([0,1])$. For every random variable ξ on $(\Omega, \mathcal{F}, P)$ uniformly distributed in $[0,1]$ we have

$$P\bigl(f(x,\xi) \le s\bigr) = P\bigl(\xi \le \mu(x|[0,s])\bigr) = \mu(x|[0,s])$$

for all $s \in [0,1]$. Hence the mapping $f(x, \xi)$ has the distribution $\mu(x|\,\cdot\,)$. $\square$

10.7.7. Corollary. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(S, \mathcal{S})$ a measurable space, T a Souslin space, and let*

$$\xi, \xi'\colon\ (\Omega, \mathcal{A}) \to (S, \mathcal{S}) \quad \text{and} \quad \eta\colon\ (\Omega, \mathcal{A}) \to \bigl(T, \mathcal{B}(T)\bigr)$$

be measurable mappings such that ξ and ξ' have a common distribution. Suppose there exists a random variable θ uniformly distributed in $[0,1]$ such that θ and ξ' are independent. Then there exists a measurable mapping $\eta'\colon\ \Omega \to T$ such that the mappings (ξ, η) and (ξ', η') have a common distribution.

Moreover, η' can be taken in the form $\eta' = f(\xi', \theta)$ with some measurable mapping $f\colon\ S\times[0,1] \to T$.

Proof. We know that there exist probability measures $\mu(\,\cdot\,, s)$, $s \in S$, on $\mathcal{B}(T)$ such that the functions $s \mapsto \mu(B, s)$ are measurable with respect to $\mathcal{S}$ and $\mu(B,\xi) = P(\eta \in B|\xi)$ a.e. (see Example 10.7.5). According to the above proposition, there exists a measurable mapping $f\colon\ S\times[0,1] \to T$ such that the random element $f(s,\theta)$ has the distribution $\mu(\,\cdot\,, s)$ for each $s \in S$. Let $\eta' = f(\xi', \theta)$. For every bounded $\mathcal{S}\otimes\mathcal{B}(T)$-measurable function g on $S\times T$ we obtain

$$\mathbb{E}g(\xi',\eta') = \mathbb{E}g\bigl(\xi', f(\xi',\theta)\bigr) = \mathbb{E}\int_0^1 g\bigl(\xi, f(\xi,u)\bigr)\,du$$
$$= \mathbb{E}\int_T g\bigl(\xi, f(\xi,t)\bigr)\,\mu(dt,\xi) = \mathbb{E}g(\xi,\eta),$$

which gives the equality of the distributions of (ξ',η') and (ξ,η). $\square$

10.8. Measurable partitions

A partition of a measure space $(M,\mathcal{M},\mu)$ is a representation of M in the form of the union of pairwise disjoint measurable sets ζ_α, where the index α runs through some nonempty set T. Let $\zeta = (\zeta_\alpha)_{\alpha\in T}$. A basic example is the partition into preimages of points under a measurable function.

Arbitrary unions of elements of a partition ζ will be called ζ-sets. For example, if ζ is the partition of the square $[0,1]^2$ into intervals parallel to the ordinate axis, then the ζ-sets are sets of the form $A\times[0,1]$, where $A\subset[0,1]$.

Suppose we are given a countable family of measurable sets $S = (S_n)$. For every sequence $\omega = (\omega_n) \in \{0,1\}^\infty$, let $S_n(\omega_n) = S_n$ if $\omega_n = 1$ and $S_n(\omega_n) = M\backslash S_n$ if $\omega_n = 0$. Let us consider the set $\bigcap_{n=1}^\infty S_n(\omega_n)$. It is clear that the obtained sets (we take into account only nonempty ones) form a partition, which is denoted by $\zeta(S)$. The family S is called a basis of the partition.

10.8.1. Definition. *A partition ζ is called measurable if it has the form $\zeta = \zeta(S)$ for some at most countable collection S of measurable sets.*

We have the following characterization of measurable partitions.

10.8.2. Lemma. *A partition is measurable if and only if it has the form $\zeta = \bigl(f^{-1}(c)\bigr)_{c\in[0,1]}$ for some measurable function $f\colon\ M\to[0,1]$.*

Proof. The partition into preimages of points is measurable, since it has a basis $f^{-1}(I_n)$, where $\{I_n\}$ are all intervals with rational endpoints. Conversely, let $S = (S_n)$ be a basis of a measurable partition ζ. The mapping

$$g\colon\ M \to \{0,1\}^\infty, \quad g(x) = \bigl(I_{S_n}(x)\bigr)_{n=1}^\infty,$$

is measurable if $\{0,1\}^\infty$ is equipped with its standard Borel σ-algebra. It is clear that ζ coincides with the partition into preimages of points under the mapping g. It remains to take an injective Borel function $\varphi\colon\ \{0,1\}^\infty \to [0,1]$ and set $f = \varphi\circ g$. $\square$

It is clear that any partition into preimages of points under a measurable mapping to a space with a countably generated and countably separated σ-algebra is measurable. Since the elements of a measurable partition have the form $f^{-1}(c)$, $c \in \mathbb{R}^1$, according to §10.4 one obtains regular conditional measures on them.

We shall say that two partitions ζ and ζ' are identical mod0 if there exists a set M_0 of full μ-measure such that the partitions of the set M_0 that are induced by ζ and ζ' are equal.

The set of partitions has the following natural order: $\zeta \le \zeta'$ if every element of the partition ζ is constituted of some collection of elements of the partition ζ'. In this case, ζ' is called a finer partition (respectively, ζ is called a coarser partition).

For every sequence of measurable partitions ζ_n, there is the coarsest partition ζ that is finer than every ζ_n. This partition is denoted by $\bigvee_{n=1}^{\infty} \zeta_n$ and can be defined as the partition into preimages of points under the mapping $x \mapsto \bigl(f_n(x)\bigr)$, $M \to [0,1]^\infty$, where f_n generates the partition ζ_n according to the above lemma and $[0,1]^\infty$ is equipped with its natural Borel σ-algebra.

Let μ be a probability measure. Two measurable partitions ζ and η are called *independent* if they are generated by functions f and g that are independent random variables on $(M, \mathcal{M}, \mu)$, i.e., one has

$$\mu\bigl(x\colon\ f(x) < a, g(x) < b\bigr) = \mu\bigl(x\colon\ f(x) < a\bigr)\mu\bigl(x\colon\ g(x) < b\bigr)$$

for all $a, b \in \mathbb{R}^1$ (see §10.10(i)). According to Exercise 10.10.50, this is equivalent to saying that for every measurable ζ-set A and every measurable η-set B one has the equality

$$\mu(A \cap B) = \mu(A)\mu(B).$$

Two measurable partitions ζ and η are called *mutually complementary* if $\zeta \bigvee \eta$ is identical mod0 to the partition into single points. Thus, if ζ is generated by a function f and η is generated by a function g, then it is required that the mapping $(f, g)\colon\ M \to \mathbb{R}^2$ be injective on a set of full measure.

Mutually complementary independent partitions are called *independent complements* of each other.

10.8.3. Theorem. *Suppose that ζ is a measurable partition of a Lebesgue–Rohlin space $(M, \mathcal{M}, \mu)$, where μ is a probability measure, such that almost all conditional measures on the elements of the partition have no atoms. Then ζ possesses an independent complement.*

PROOF. In terms of random variables we have to prove the following. Let a measurable function $f\colon\ M \to [0,1]$ be such that for $\mu \circ f^{-1}$-a.e. y the conditional measure μ^y on $f^{-1}(y)$ has no atoms. Then, there exists a measurable function g on M with values in $[0,1]$ such that the mapping $(f, g)\colon\ M \to [0,1]^2$ is injective on a set of full measure and transforms μ into a measure $\nu \otimes \nu_0$, where $\nu = \mu \circ f^{-1}$ and ν_0 is some probability measure. By the isomorphism theorem we may assume that μ is a Borel measure on $[0,1]$ and f is a Borel function. Let

$$g(x) = \mu^{f(x)}\bigl([0,x)\bigr).$$

We observe that the function $(x,t) \mapsto \mu^{f(x)}\bigl([0,t)\bigr)$ is Borel measurable. This follows by Lemma 6.4.6 since the function $\mu^y([0,t])$ is Borel in y for any fixed t and left continuous in t for any fixed y. Then g is a Borel function. The mapping (f,g) is injective on the set Ω of all $x \in M$ such that the measure $\mu^{f(x)}$ has no atoms and $g(x) < g(x+n^{-1})$ for all $n \in \mathbb{N}$. Indeed, if $x_1, x_2 \in \Omega$ and $x_1 < x_2$, then either $f(x_1) \neq f(x_2)$ or $f(x_1) = f(x_2) = y$ and $g(x_1) < g(x_2)$ because $x_2 > x_1 + n^{-1}$ for some n. One has $\mu(\Omega) = 1$ since Ω contains the intersection Ω_0 of the set $\{x\colon\, g(x) < g(x+n^{-1})\ \forall\, n \in \mathbb{N}\}$ and the set $f^{-1}(E)$, where E is a Borel set such that $\nu(E) = 1$ and the conditional measures μ^y have no atoms. Indeed, the set Ω_0 is μ-measurable, and $\mu^y(\Omega_0) = 1$ for all $y \in E$, which is clear from the following observation: for every atomless Borel probability measure σ on $[0,1]$ with the distribution function F_σ, for σ-a.e. t, one has $F_\sigma(t) < F_\sigma(t+n^{-1})$ for all n (the topological support of σ has the form $[0,1]\backslash\bigcup_{k=1}^\infty (a_k,b_k)$, so every point $t \notin \bigcup_{k=1}^\infty [a_k,b_k]$ has the aforementioned property). We show that the measure μ is transformed by the mapping (f,g) to the product of the measure ν and Lebesgue measure λ on $[0,1]$. To this end, it suffices to show that whenever $a<b$, $c<d$, one has the equality

$$\mu\bigl((f,g)^{-1}([a,b]\times[c,d])\bigr) = \nu([a,b])\lambda([c,d]).$$

Since $\mu^y\bigl(f^{-1}(y)\bigr) = 1$, the left-hand side of this equality is

$$\int_{[a,b]} \mu^y\bigl((f,g)^{-1}([a,b]\times[c,d])\bigr)\,\nu(dy) = \int_{[a,b]} \mu^y\bigl(f^{-1}(y)\cap g^{-1}([c,d])\bigr)\,\nu(dy).$$

It remains to observe that on the set $f^{-1}(y)$ the function g coincides with the distribution function of the measure μ^y. Since the measure μ^y is concentrated on $f^{-1}(y)$, it follows by Example 3.6.2 that for all $y \in E$ we have $\mu^y\bigl(f^{-1}(y)\cap g^{-1}([c,d])\bigr) = \lambda([c,d])$, which yields the assertion. $\square$

10.9. Ergodic theorems

In this section, we prove several principal theorems of ergodic theory — an intensively developing field of mathematics on the border of measure theory, the theory of dynamical systems, mathematical physics, and probability theory. In these theorems, one is concerned with a family of measure-preserving transformations T_t, where the parameter t takes values in $\mathbb{N}$ or $[0,+\infty)$, and the problem is the study of the asymptotic behavior of these transformations for large t. Certainly, in this introductory discussion, it is impossible even to mention all interesting problems of measure theory arising in the described situation. The interested reader is referred to the books Arnold, Avez [**71**], Billingsley [**168**], Cornfeld, Sinaĭ, Fomin [**376**], Garsia [**671**], Halmos [**780**], Krengel [**1058**], Petersen [**1437**], Sinai [**1730**].

One of the first results of ergodic theory was the following Poincaré recurrence theorem.

10.9.1. Theorem. *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space and let $T\colon \Omega \to \Omega$ be a $(\mathcal{B}_\mu, \mathcal{B})$-measurable mapping such that $\mu\circ T^{-1} = \mu$. If A is a μ-measurable set, then for μ-almost every $x \in A$, there exists an infinite sequence of indices n_i such that $T^{n_i}x \in A$. In particular, if $\mu(A) > 0$, then there exists a point $x \in A$ such that $T^n x \in A$ for infinitely many n.*

PROOF. If $\mu(A) = 0$, then our claim is true in the trivial way. We assume further that $\mu(A) > 0$. We prove first a weaker assertion that for almost every $x \in A$, there exists $n \in \mathbb{N}$ such that $T^n x \in A$. Points with such a property are called recurrent. Denote by E the set of all points $x \in A$ such that $T^n x \notin A$ for all $n \geq 1$. It is easy to see that the set E is measurable. In order to show that $\mu(E) = 0$, it suffices to verify that the sets E, $T^{-1}(E)$, $T^{-1}\bigl(T^{-1}(E)\bigr)$ and so on are pairwise disjoint, since by hypothesis they have equal measures. These sets will be denoted by E_k: $E_{k+1} := T^{-1}(E_k)$, $E_0 := E$. Suppose that $x \in E_m \cap E_p$, where $m > p$. Then

$$T^p x \in E \cap T^p E_m = E \cap E_{m-p}.$$

Therefore, letting $y = T^p x \in E$ we obtain $T^{m-p} y \in E \subset A$ contrary to the definition of E.

Now the initial assertion follows by the considered partial case. Indeed, for every $k \in \mathbb{N}$, the measurable mapping T^k transforms the measure μ into μ. As we have proved, almost all points in A are recurrent for T^k. Therefore, almost all points in A are recurrent simultaneously for all T^k, which completes the proof. □

We shall now see that the Poincaré theorem admits a substantial reinforcement. The so-called individual ergodic theorem (the Birkhoff–Khinchin theorem) proven below is one of the key results of ergodic theory. Given a measurable transformation T of a probability space $(\Omega, \mathcal{B}, \mu)$, we denote by $\mathcal{T}$ the σ-algebra of all sets $B \in \mathcal{B}$ with $B = T^{-1}(B)$. The conditional expectation with respect to $\mathcal{T}$ will be denoted by $\mathbb{E}^{\mathcal{T}}$.

We observe that if $T\colon \Omega \to \Omega$ is a $(\mathcal{B}_\mu, \mathcal{B})$-measurable mapping that preserves the measure μ, i.e., $\mu = \mu \circ T^{-1}$, then $\mu\bigl(T^{-1}(Z)\bigr) = 0$ for every set Z of μ-measure zero. Hence, for any $f \in \mathcal{L}^1(\mu)$, the function $f \circ T$ is a.e. defined and μ-integrable.

10.9.2. Lemma. *Let T be a measure-preserving transformation of a probability space $(\Omega, \mathcal{B}, \mu)$, $f \in \mathcal{L}^1(\mu)$, $k \in \mathbb{N}$, and let*

$$f_k(x) = f(T^k x), \quad S_k = f_0 + \cdots + f_{k-1}, \quad M_k = \max(0, S_1, \ldots, S_k).$$

Then

$$\int_{\{M_k > 0\}} f\, d\mu \geq 0.$$

PROOF. For all $j \le k$ we have $M_k(Tx) \ge S_j(Tx)$, whence

$$M_k(Tx) + f(x) \ge S_j(Tx) + f(x) = S_{j+1}(x),$$

i.e., we have the inequality $f(x) \ge S_{j+1}(x) - M_k(Tx)$, $j = 1, \ldots, k$. In addition, we have $f(x) = S_1(x) \ge S_1(x) - M_k(Tx)$. Hence

$$\int_{\{M_k>0\}} f\,d\mu \ge \int_{\{M_k>0\}} \bigl[\max(S_1, \ldots, S_k) - M_k \circ T\bigr]\,d\mu$$
$$= \int_{\{M_k>0\}} [M_k - M_k \circ T]\,d\mu \ge 0,$$

since the integral of $M_k - M_k \circ T$ over Ω vanishes, whereas on the complement of $\{M_k > 0\}$ we have $M_k = 0$ and $M_k \circ T \ge 0$. □

10.9.3. Corollary. *In the situation of the above lemma one has*

$$\mu\bigl(\max(S_1, S_2/2, \ldots, S_k/k) > r\bigr) \le r^{-1} \int_\Omega |f|\,d\mu, \quad \forall\, r > 0.$$

PROOF. Let us set $B = \bigl\{\max(S_1, S_2/2, \ldots, S_k/k) > r\bigr\}$ and

$$g = f - r,\ \widetilde{S}_k = g + \cdots + g \circ T^{k-1},\ \widetilde{M}_k = \max(0, g, \ldots, \widetilde{S}_k).$$

By the lemma the integral of g over $\{\widetilde{M}_k > 0\}$ is nonnegative. We observe that $B = \{\widetilde{M}_k > 0\}$. Indeed, $\widetilde{S}_j = S_j - jr$, hence the inequalities $\widetilde{S}_j > 0$ and $S_j/j > r$ are equivalent. Therefore, $r\mu(B)$ does not exceed the integral of f over B . Since the integral of f is majorized by that of $|f|$, the claim follows. □

Now we can prove the Birkhoff–Khinchin theorem.

10.9.4. Theorem. *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space and let f be a μ-integrable function. Suppose that $T\colon\ \Omega \to \Omega$ is a $(\mathcal{B}_\mu, \mathcal{B})$-measurable mapping such that $\mu \circ T^{-1} = \mu$. Then for μ-a.e. x, there exists a limit*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) := \overline{f}(x).$$

In addition, $\overline{f}$ is in $L^1(\mu)$, coincides a.e. with $\mathbb{E}^{\mathcal{T}} f$ and

$$\int_\Omega f\,d\mu = \int_\Omega \overline{f}\,d\mu.$$

PROOF. Since T preserves μ, we may assume that f is defined everywhere (its redefinition on a measure zero set does not affect our assertion). We observe that $(\mathbb{E}^{\mathcal{T}} f) \circ T = \mathbb{E}^{\mathcal{T}} f$. Indeed, $I_B \circ T = I_B$ for all $B \in \mathcal{T}$, hence for every bounded $\mathcal{T}$-measurable function ψ we have $\psi \circ T = \psi$, which yields the same equality for every $\mathcal{T}$-measurable function. Therefore, one can pass to $f - \mathbb{E}^{\mathcal{T}} f$ and assume further that $\mathbb{E}^{\mathcal{T}} f = 0$. Let $S_k = f + f \circ T + \cdots + f \circ T^{k-1}$,

$g = \limsup_{n\to\infty} S_n/n$, $\varepsilon > 0$, and $E := \{g > \varepsilon\}$. We show that $\mu(E) = 0$. Let

$$f^\varepsilon = (f-\varepsilon)I_E,\ S_k^\varepsilon = f^\varepsilon + \cdots + f^\varepsilon \circ T^{k-1},\ M_k^\varepsilon = \max(0, S_1^\varepsilon, \ldots, S_k^\varepsilon).$$

It is clear that $E \in \mathcal{T}$, since $g \circ T = g$. In addition, the sequence of functions M_k^ε is increasing and $E = \bigcup_{k=1}^\infty \{M_k^\varepsilon > 0\}$. This is easily seen from the equality $S_k^\varepsilon = (S_k - k\varepsilon)I_E$. Therefore, by Lemma 10.9.2 and the monotone convergence theorem we obtain

$$0 \le \int_{\{M_k^\varepsilon > 0\}} f^\varepsilon\, d\mu \to \int_E f^\varepsilon\, d\mu.$$

By virtue of the equality $\mathbb{E}^{\mathcal{T}} f = 0$ and the inclusion $E \in \mathcal{T}$, we have according to the definition of conditional expectation

$$\int_E f\, d\mu = \int_E \mathbb{E}^{\mathcal{T}} f\, d\mu = 0.$$

Thus, the above estimate can be written in the form $-\varepsilon\mu(E) \ge 0$, i.e., one has $\mu(E) = 0$. Hence $S_n/n \to 0$ a.e.

Now we prove mean convergence. For any fixed $N \in \mathbb{N}$ let us set $\psi_N = fI_{\{|f|\le N\}}$, $\varphi_N = f - \psi_N$. Then $|\psi_N| \le N$ and by the previous step the functions $n^{-1}\sum_{k=0}^{n-1} \psi_N \circ T^k$ converge to $\mathbb{E}^{\mathcal{T}}\psi_N$ in $L^1(\mu)$ as $n \to \infty$. We observe that by the invariance of μ with respect to T and the estimate $\|\mathbb{E}^{\mathcal{T}}\varphi_N\|_{L^1(\mu)} \le \|\varphi_N\|_{L^1(\mu)}$, one has the inequality

$$\int_\Omega \Bigl| n^{-1}\sum_{k=0}^{n-1} \varphi_N \circ T^k - \mathbb{E}^{\mathcal{T}}\varphi_N \Bigr|\, d\mu$$
$$\le n^{-1}\sum_{k=0}^{n-1} \int_\Omega |\varphi_N \circ T^k|\, d\mu + \int_\Omega |\mathbb{E}^{\mathcal{T}}\varphi_N|\, d\mu \le 2\int_\Omega |\varphi_N|\, d\mu.$$

Since the right-hand side of this inequality tends to zero as $N \to \infty$ and $f = \psi_N + \varphi_N$, the theorem is proven. □

Let us consider continuous time systems.

10.9.5. Corollary. *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space and let $(T_t)_{t\ge 0}$ be a semigroup of measure-preserving transformations, i.e., $T_0 = I$, $T_{s+t} = T_s \circ T_t$, the mappings T_t are $(\mathcal{B}_\mu, \mathcal{B})$-measurable, and $\mu \circ T_t^{-1} = \mu$. Suppose $f \in \mathcal{L}^1(\mu)$ is such that $(x,t) \mapsto f\bigl(T_t(x)\bigr)$ is $\mathcal{B}_\mu \otimes \mathcal{B}\bigl([0,+\infty)\bigr)$-measurable. Then μ-a.e. and in $L^1(\mu)$ there exists a limit*

$$\overline{f}(x) := \lim_{t\to+\infty} t^{-1}\int_0^t f\bigl(T_s(x)\bigr)\, ds$$

and $\overline{f} = \mathbb{E}^{\mathcal{T}_\infty} f$ a.e., where $\mathcal{T}_\infty$ is the σ-algebra generated by all μ-measurable functions φ such that, for every $\tau > 0$, one has $\varphi\bigl(T_\tau(x)\bigr) = \varphi(x)$ a.e.

PROOF. Let us apply the ergodic theorem to the function g defined as follows: $g(x)$ is the Lebesgue integral of $f(T_s(x))$ in s over $[0,1]$. We observe that the function g is measurable and that the equality

$$\sum_{k=0}^{n-1} g(T_1^k x) = S_n(x) := \int_0^n f\bigl(T_s(x)\bigr)\,ds$$

holds. Hence a.e. there exists a limit $h(x) := \lim\limits_{n\to\infty} n^{-1}S_n(x)$. It suffices to consider the case $f \geq 0$. This gives at once the existence of the limit indicated in the theorem almost everywhere and its coincidence with $h(x)$ because $n^{-1}\bigl(S_{n+1}(x) - S_n(x)\bigr) \to 0$ a.e. In order to prove convergence in $L^1(\mu)$ it suffices to consider bounded functions f since the L^1-norm of the function

$$S_t(x) = t^{-1}\int_0^t f\bigl(T_s(x)\bigr)\,ds$$

does not exceed the norm of f. For bounded f, the equality $\lim\limits_{n\to\infty} \|S_t - \overline{f}\|_1 = 0$ is obvious from the already-established facts. It is clear that $\overline{f}\bigl(T_\tau(x)\bigr) = \overline{f}(x)$ a.e. for each $\tau > 0$. For any $\mathcal{T}_\infty$-measurable bounded function φ we have

$$\begin{aligned}\int_\Omega f(x)\varphi(x)\,\mu(dx) &= \int_\Omega f\bigl(T_s(x)\bigr)\varphi\bigl(T_s(x)\bigr)\,\mu(dx)\\ &= \int_\Omega f\bigl(T_s(x)\bigr)\varphi(x)\,\mu(dx),\end{aligned}$$

which yields the equality $\overline{f} = \mathbb{E}^{\mathcal{T}_\infty} f$. □

One can find a version $\overline{f}$ with values in $[-\infty,+\infty]$ such that $\overline{f}\bigl(T_t(x)\bigr) = \overline{f}(x)$ for all $x \in \Omega$, $t \geq 0$. To this end, for nonnegative functions f, we set

$$\overline{f}(x) := \lim_{r\to+\infty}\limsup_{n\to\infty} n^{-1}\int_r^{r+n} f\bigl(T_s(x)\bigr)\,ds.$$

10.9.6. Example. Let $\Omega = [0,1)$ be equipped with Lebesgue measure λ and let $T(x) = x + \theta(\mathrm{mod}\,1)$, where $\theta \in \mathbb{R}^1$ is a fixed number. Then T preserves the measure λ. If θ is irrational, then for every Borel set B, one has

$$n^{-1}\sum_{k=0}^{n-1} I_B \circ T^k \to \lambda(B) \quad \text{a.e.}$$

This follows by the ergodic theorem taking into account that the σ-algebra $\mathcal{T}$ is trivial: every $\mathcal{T}$-measurable function a.e. equals some constant, since by the irrationality of θ it has arbitrarily small periods (see Exercise 5.8.109).

Kozlov and Treschev [**1054**] discovered the following very interesting averaging property in the case of continuous time.

10.9.7. Theorem. *Suppose that in the situation of Corollary* 10.9.5 *the function f is bounded. Let ϱ be a probability density on* $[0,+\infty)$*. Then, the* *function* $(x,t,s)\mapsto f\bigl(T_{st}(x)\bigr)$ *is* $\mathcal{B}_\mu\otimes\mathcal{B}\bigl([0,+\infty)\bigr)\otimes\mathcal{B}\bigl([0,+\infty)\bigr)$*-measurable and* μ*-a.e. one has*

$$\overline{f}(x)=\lim_{t\to+\infty}\int_0^\infty f\bigl(T_{st}(x)\bigr)\varrho(s)\,ds.$$

PROOF. Let us approximate ϱ in $L^1(\mathbb{R}^1)$ by a sequence of compactly supported probability densities that assume finitely many values and are piecewise constant. The claim for such densities follows by Corollary 10.9.5. It remains to observe that the difference between the considered integrals for ϱ and ϱ_n does not exceed $\|\varrho_n-\varrho\|_{L^1}\sup_x|f(x)|$. Additional results in this direction can be found in Bogachev, Korolev [**219**]. $\square$

In connection with the ergodic theorem several interesting concepts arise, of which we only mention the ergodicity and mixing.

10.9.8. Definition. *Suppose that* $(\Omega,\mathcal{B},\mu)$ *is a probability space and* T *is a transformation preserving the measure* μ*. Then* T *is called ergodic if every set in* $\mathcal{T}$ *has measure either* 0 *or* 1*.*

If for every $A,B\in\mathcal{B}$ *we have*

$$\lim_{n\to\infty}\mu\bigl(A\cap T^{-n}(B)\bigr)=\mu(A)\mu(B),\tag{10.9.1}$$

then T *is called mixing.*

Ergodicity is equivalent to the property that the space of all $\mathcal{T}$-measurable functions in $L^1(\mu)$ consists of constants. In turn, this is equivalent to the property that $\mathbb{E}^{\mathcal{T}}$ coincides with the usual expectation. Hence for any ergodic measure, the averages indicated in the ergodic theorem converge to the integral of the function over the space. In other words, the time averages coincide with the space averages, which has an important physical sense.

It is clear that the mixing implies the ergodicity, since we have the equality $\mu(A)=\mu(A)^2$ whenever $A=B\in\mathcal{T}$. On the other hand, the ergodicity is equivalent to a somewhat weaker relationship than (10.9.1), namely, to the following property:

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\mu\bigl(A\cap T^{-k}(B)\bigr)=\mu(A)\mu(B),\quad A,B\in\mathcal{B}.\tag{10.9.2}$$

Indeed, by the ergodic theorem, for any ergodic T we have a.e.

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}I_B\circ T^k=\mu(B),$$

which after integration over A yields (10.9.2). If (10.9.2) is fulfilled, then on account of the relationship $n^{-1}\sum_{k=0}^{n-1}I_B\circ T^k\to\mathbb{E}^{\mathcal{T}}I_B$ a.e., we obtain

$$\int_A\mathbb{E}^{\mathcal{T}}I_B\,d\mu=\mu(A)\mu(B).$$

This means that $\mathbb{E}^{\mathcal{T}}I_B=\mu(B)$ a.e., hence $\mathcal{T}$ is trivial.

10.9.9. Example. (i) A transformation T with irrational θ in Example 10.9.6 is ergodic, but not mixing. Indeed, its ergodicity has been explained in Example 10.9.6. In order to see that it is not mixing we observe that, by the irrationality of θ, there exists a sequence of natural numbers n_k with $n_k\theta (\mathrm{mod}\, 1) \to 1/2$. Let $A = B = [0, 1/4)$. Then, for large k, the sets A and $T^{-n_k}(B)$ do not meet, so that (10.9.1) is impossible.

(ii) Let $(X, \mathcal{A}, \mu)$ be a probability space and let $\Omega = X^{\mathbb{Z}}$ be equipped with a measure P that is the product of countably many copies of μ. Then the transformation $T\colon (x_n) \mapsto (x_{n+1})$ preserves P and is mixing. Indeed, for cylindrical sets A and B, for all sufficiently large n we have the equality $P\bigl(A \cap T^{-n}(B)\bigr) = P(A)P(B)$, which yields (10.9.1) for all measurable sets.

Bourgain [**244**] proved that if T is an ergodic measure-preserving transformation of a probability space $(\Omega, \mathcal{B}, \mu)$, then for all $f, g \in \mathcal{L}^\infty(\mu)$ and all natural numbers p and q, the limit $\lim\limits_{n\to\infty} n^{-1}\sum_{k=1}^n f(T^{pk}x)g(T^{qk}x)$ exists a.e.

We close this section with some results from the recent paper Ivanov [**872**], where very interesting connections between certain ergodic type limit theorems and elementary properties of increasing functions have been discovered.

Let S be a measurable set of finite measure on the real line and let F be an increasing function on S. We fix two numbers α and β with $0 < \alpha < \beta$. A screen of the point $x \in S$ is any interval $(y, z) \subset S$ such that $x < y$ and

$$F(y + 0) - F(x) \geq \beta(y - x), \quad F(z - 0) - F(x) \leq \alpha(z - x).$$

Let S^* denote the set of all points in S possessing screens (with these α and β). V.V. Ivanov [**871**], [**872**] discovered the following surprising inequality.

10.9.10. Theorem. *Under the assumptions made above, one has the estimate* $\lambda(S^*) \leq \frac{\alpha}{\beta}\lambda(S)$.

10.9.11. Corollary. *Let* $I = [a, b]$ *and let* F *be an increasing function on* I. *Given* $0 < \alpha < \beta$ *and* $k \in \mathbb{N}$, *let* I_k *denote the set of all points* $x \in I$ *for which there exists a chain* $x < y_1 < z_1 < \cdots < y_k < z_k \leq b$ *such that*

$$[F(y_i) - F(x)]/(y_i - x) \geq \beta \quad \textit{and} \quad [F(z_i) - F(x)]/(z_i - x) \leq \alpha$$

for all $i = 1, \ldots, k$. *Then* $\lambda(I_k) \leq (\alpha/\beta)^k\lambda(I)$.

The remarkable inequality of Ivanov has already found applications, one of which is discussed below. For these applications, it suffices to be able to prove Ivanov's inequality in the simplest case where S is a closed interval and the function F is piece-wise constant and assumes only finitely many values. Surprisingly enough, even in this partial case, the proof, although completely elementary, is rather involved (in fact, in [**872**], the general case is reduced to this partial case whose accurate justification takes about two pages).

Now we consider a probability space $(\Omega, \mathcal{B}, \mu)$ and a semigroup $\{T_t\}_{t\geq 0}$ of mappings $T_t\colon \Omega \to \Omega$ preserving the measure μ. We shall assume that the mapping $T_t(\omega)$ is measurable in (t,ω). Then, for every integrable function f on Ω, one obtains μ-integrable functions

$$\sigma_t(\omega) := \frac{1}{t}\int_0^t f\bigl(T_s(\omega)\bigr)\,ds.$$

In the case where the transformations T_n are defined only for $n \in \mathbb{N}$, i.e., $T_n = T^n$, where T is a measure-preserving transformation, we set

$$\sigma_n(\omega) := n^{-1}\sum_{k=1}^{n} f\bigl(T_k(\omega)\bigr).$$

For any fixed $0 < \alpha < \beta$ and $k \in \mathbb{N}$, we denote by $\Omega_k(\alpha,\beta)$ the set of all points $\omega \in \Omega$ such that there exists a chain $0 < s_1 < t_1 < \cdots < s_k < t_k$ for which $\sigma_{s_i}(\omega) \geq \beta$ and $\sigma_{t_i}(\omega) \leq \alpha$ for all $i = 1,\ldots,k$. Thus, the trajectory of the point ω up-crosses at least k times the strip between the levels α and β. Analogous sets are defined in the discrete time case. According to Exercise 10.10.70, the sets $\Omega_k(\alpha,\beta)$ are measurable. By using Theorem 10.9.10 the following remarkable estimate is derived in [**872**].

10.9.12. Theorem. *Let $f \geq 0$. Then $\mu\bigl(\Omega_k(\alpha,\beta)\bigr) \leq (\alpha/\beta)^k$.*

It is clear from the proof of the individual ergodic theorem that this estimate not only implies the ergodic theorem, but also gives a universal estimate of fluctuations of the averages. In the continuous time case, Ivanov's estimate gives an alternative proof of the existence of a limit $f^* = \lim\limits_{t\to\infty} \sigma_t$. For a bounded function f, by the dominated convergence theorem and invariance of μ we obtain that the integrals of f^* and f are equal, which yields easily that the same is true for all integrable functions.

10.10. Supplements and exercises

10.10(i). Independence

In this subsection we briefly discuss the concept of independence, which is crucial for probability theory, and is often of use and importance in measure theory.

10.10.1. Definition. *Let $(X,\mathcal{A},\mu)$ be a probability space and let*

$$\xi\colon\, X \to E_1 \quad \textit{and} \quad \eta\colon\, X \to E_2$$

be measurable mappings to measurable spaces $(E_1,\mathcal{E}_1)$ and $(E_2,\mathcal{E}_2)$. The mappings ξ and η are called independent (or stochastically independent) if

$$\mu(\xi \in A_1, \eta \in A_2) = \mu(\xi \in A_1)\mu(\eta \in A_2) \quad \textit{for all } A_1 \in \mathcal{E}_1, A_2 \in \mathcal{E}_2.$$

It is clear that if a measurable mapping η is constant, then, for any measurable mapping ξ, the mappings ξ and η are independent. In addition, if ξ and η are independent and $\psi_1\colon E_1 \to E_1$ and $\psi_2\colon E_2 \to E_2$ are measurable mappings, then $\psi_1 \circ \xi$ and $\psi_2 \circ \eta$ are independent. If $E_1 = E_2 = \mathbb{R}^1$ and $\mathcal{E}_1 = \mathcal{E}_2 = \mathcal{B}(\mathbb{R}^1)$, then the independence of ξ and η is equivalent to the equality $\mu(\xi < a, \eta < b) = \mu(\xi < a)\mu(\eta < b)$ for all a and b. This follows from the fact that $\mu(\xi \in A_1, \eta \in A_2)$ and $\mu(\xi \in A_1)\mu(\eta \in A_2)$ are measures as functions of A_1 and A_2, and any Borel measure on the real line is uniquely determined by its values on rays.

It is seen from the definition that the concept of independence is related not only to the mappings and measure, but also to the σ-algebras $\mathcal{E}_i$. The most important for applications is the case where $E_1 = E_2 = \mathbb{R}^1$ and $\mathcal{E}_1 = \mathcal{E}_2 = \mathcal{B}(\mathbb{R}^1)$. In that case, it suffices to take for A_1 and A_2 only intervals. We remark that one can introduce a stronger concept of independence (independence in the sense of Kolmogorov) by requiring the equality

$$\mu(\xi \in A_1, \eta \in A_2) = \mu(\xi \in A_1)\mu(\eta \in A_2)$$

for all $A_i \subset E_i$ such that $\xi^{-1}(A_1) \in \mathcal{A}$, $\eta^{-1}(A_2) \in \mathcal{A}$. Even in the case $E_1 = E_2 = \mathbb{R}^1$ and $\mathcal{E}_1 = \mathcal{E}_2 = \mathcal{B}(\mathbb{R}^1)$, this definition is strictly stronger (Exercise 10.10.73). However, if $E_1 = E_2 = \mathbb{R}^1$, $\mathcal{E}_1 = \mathcal{E}_2 = \mathcal{B}(\mathbb{R}^1)$, and the measure μ is perfect, then both definitions are obviously equivalent (see Ramachandran [**1519**] on other cases of equivalence).

It is clear that measurable mappings ξ and η with values in $(E_1, \mathcal{E}_1)$ and $(E_2, \mathcal{E}_2)$ are independent precisely if

$$\mu \circ (\xi, \eta)^{-1} = (\mu \circ \xi^{-1}) \otimes (\mu \circ \eta^{-1})$$

on $(E_1 \times E_2, \mathcal{E}_1 \otimes \mathcal{E}_2)$. By analogy one defines independence of families of measurable mappings. Namely, given a sequence (finite or countable) of measurable mappings ξ_n on X with values in measurable spaces $(E_n, \mathcal{E}_n)$, we call it a sequence of independent random elements if the image of μ under the mapping $(\xi_1, \xi_2, \ldots)$ to $\prod_{n=1}^\infty E_n$ coincides with the countable product of the measures $\mu \circ \xi_n^{-1}$. Obviously, this is equivalent to the equality

$$\mu\bigl(x\colon\ \xi_1(x) \in A_1, \ldots, \xi_n(x) \in A_n\bigr) = \mu\bigl(x\colon\ \xi_1(x) \in A_1\bigr) \cdots \mu\bigl(x\colon\ \xi_n(x) \in A_n\bigr)$$

for all $n \in \mathbb{N}$ and $A_i \in \mathcal{E}_i$. More generally, given a family of measurable mappings ξ_t with values in measurable spaces $(E_t, \mathcal{E}_t)$, we call it independent random elements if every finite subfamily is independent in the above sense. It should be noted that this independence is stronger than the pairwise independence of ξ_t (Exercise 10.10.80).

Two measurable sets A and B in a probability space $(X, \mathcal{A}, \mu)$ are called (stochastically) independent if their indicators I_A and I_B are independent. This is equivalent to the equality $\mu(A \cap B) = \mu(A)\mu(B)$. More generally, a family of measurable sets A_t in a probability space $(X, \mathcal{A}, \mu)$ is called (stochastically) independent if the family of functions I_{A_t} is independent. An equivalent condition: $\mu(A_{t_1} \cap \cdots A_{t_n}) = \mu(A_{t_1}) \cdots \mu(A_{t_n})$ for all distinct

$t_1, \ldots, t_n$. Two family of sets $\mathcal{A}$ and $\mathcal{B}$ are called independent if $\mu(A \cap B) = \mu(A)\mu(B)$ for all $A \in \mathcal{A}$, $B \in \mathcal{B}$. Finally, families $\mathcal{A}_t$ of measurable sets are called independent if the sets A_t are independent whenever $A_t \in \mathcal{A}_t$. All these properties refer to an a priori given probability measure.

10.10.2. Lemma. *If two functions ξ and η on $(X, \mathcal{A}, \mu)$ are independent and integrable, then the function $\xi\eta$ is integrable as well and one has*

$$\int_X \xi\eta \, d\mu = \int_X \xi \, d\mu \int_X \eta \, d\mu.$$

PROOF. Let ξ assume finitely many values a_i on disjoint sets X_i, $i = 1, \ldots, n$, and let η assume finitely many values b_j on disjoint sets Y_j, $j = 1, \ldots, m$. Then the integral of $\xi\eta$ equals $\sum_{i,j} a_i b_j \mu(X_i \cap Y_j)$, which coincides with the product of the integrals of ξ and η, since $\mu(X_i \cap Y_j) = \mu(X_i)\mu(Y_j)$ due to independence. Let ξ and η be bounded and take values in $(-M, M)$. For every $k \in \mathbb{N}$, we partition $[-M, M]$ into k disjoint intervals $I_i = (a_i, b_i]$ of the same length and set $\xi_k(x) = b_i$ if $\xi(x) \in I_i$. Similarly, we define the functions η_k. The functions ξ_k and η_k are independent for any fixed k, since $\xi_k = \varphi_k \circ \xi$, $\eta_k = \varphi_k \circ \eta$, where φ_k is a Borel function defined by the equality $\varphi_k(t) = b_i$ whenever $t \in I_i$. Since the equality to be proven is true for ξ_k and η_k, it remains valid for ξ and η. When ξ and η are not bounded, we consider the functions $\min(k, |\xi|)$ and $\min(k, |\eta|)$ and by the monotone convergence theorem obtain the desired equality for $|\xi|$ and $|\eta|$. This shows the integrability of $\xi\eta$. Now the same reasoning completes the proof. $\square$

Let us give two interesting results (due to Banach and Marczewski) related to independence. A class of sets $\mathfrak{E}$ in a space X is called independent if for every sequence of distinct sets $E_i \in \mathfrak{E}$ we have $\bigcap_{i=1}^{\infty} D_i \neq \varnothing$, where D_i is either E_i or $X \backslash E_i$. Note that this concept involves no measures. Marczewski [**1250**] (see also the papers [**1817**], [**1251**] by the same author) obtained the following result.

10.10.3. Theorem. *Let $\mathfrak{E}$ be an independent class of subsets of a space X and let ν be a function on $\mathfrak{E}$ with values in $[0, 1]$. Then, on the σ-algebra $\sigma(\mathfrak{E})$ generated by the class $\mathfrak{E}$, there exists a probability measure μ such that*

$$\mu(E) = \nu(E) \quad \textit{for all } E \in \mathfrak{E},$$

and the sets in $\mathfrak{E}$ are stochastically independent with respect to μ.

Suppose we are given a family $\mathcal{A}_t$ of σ-algebras in a space X, where $t \in T$. This family is called countably independent if for every countable collection of nonempty sets $A_i \in \mathcal{A}_{t_i}$ with distinct t_i we have $\bigcap_{i=1}^{\infty} A_i \neq \varnothing$. Banach [**107**] proved the following theorem, which substantially generalizes the previous one (the proof below is due to Sherman [**1696**]; it is considerably shorter than the original one). The previous theorem corresponds to the case where each $\mathcal{A}_t$ is generated by a single set A_t.

10.10.4. Theorem. *Suppose we are given a countably independent family of σ-algebras $\mathcal{A}_t$, $t \in T$, in a space X such that every $\mathcal{A}_t$ is equipped with a probability measure μ_t. Then, on the σ-algebra $\mathcal{A}$ generated by all $\mathcal{A}_t$, there exists a probability measure μ such that $\mu(A) = \mu_t(A)$ for all $A \in \mathcal{A}_t$ and $\mu\bigl(\bigcap_{i=1}^{\infty} A_i\bigr) = \prod_{i=1}^{\infty} \mu_{t_i}(A_i)$ for all $A_i \in \mathcal{A}_{t_i}$, where $t_i \neq t_j$ if $i \neq j$, i.e., the σ-algebras $\mathcal{A}_t$ are stochastically independent with respect to μ.*

PROOF. The measure $\bigotimes_{t\in T} \mu_t$ on the σ-algebra $\mathcal{B} := \bigotimes_{t\in T} \mathcal{A}_t$ will be denoted by ν. Let us consider the mapping $\varphi\colon X \to X^T$ defined by the formula $\varphi(x) = (x_t)_{t\in T}$, where $x_t = x$ for all $t \in T$. Let D be the image of φ. We define μ by the equality $\mu\bigl(\varphi^{-1}(B)\bigr) := \nu(B)$, $B \in \mathcal{B}$. The theorem will be proven once we establish that the mapping $\varphi^{-1}\colon \mathcal{B} \to \mathcal{A}$ is a σ-isomorphism. It is clear that φ^{-1} takes complements to complements and countable unions (or intersections) to countable unions (respectively, intersections). For every fixed $\tau \in T$ and any $E \in \mathcal{A}_\tau$, the image of the set $B = \{(x_t)_{t\in T}\colon\, x_\tau \in E\}$ is the set E. Together with the aforementioned properties this means that $\varphi^{-1}(\mathcal{B}) = \mathcal{A}$. It remains to verify the injectivity of φ^{-1}. It suffices to show that if $B \in \mathcal{B}$ and $B \cap D = \varnothing$, then $B = \varnothing$. It is at this stage that we need the countable independence of $\mathcal{A}_t$.

Suppose first that B has the form $B = \{(x_t)_{t\in T}\colon\, x_{t_i} \in B_i\}$, where $\{t_i\}$ is a finite or countable set and $B_i \in \mathcal{A}_{t_i}$. Sets of such a form will be called blocks. If B is nonempty, then all B_i are nonempty. By hypothesis, there exists a point $x \in \bigcap_{i=1}^{\infty} B_i$, which gives a point in $B \cap D$. In order to complete the proof we show that every set in $\mathcal{B}$ is a union (possibly, uncountable) of a family of blocks. Denote by $\mathcal{B}_0$ the subclass in $\mathcal{B}$ consisting of all sets for which this is true. Since $\mathcal{B}_0$ contains all blocks, for the proof of the equality $\mathcal{B}_0 = \mathcal{B}$ it suffices to show that $\mathcal{B}_0$ is a monotone class. Obviously, $\mathcal{B}_0$ admits arbitrary unions. Let $B_n \in \mathcal{B}_0$ and $B_{n+1} \subset B_n$ for all n. For every point $x \in B := \bigcap_{n=1}^{\infty} B_n$ and every n, there is a block $C_n(x) \subset B_n$ that contains x. The sets $C(x) := \bigcap_{n=1}^{\infty} C_n(x)$ are blocks and their union over $x \in B$ is B because $C(x) \subset B$. Thus, $\mathcal{B}_0$ is a monotone class, hence we obtain $\mathcal{B}_0 = \mathcal{B}$. $\square$

We note that the independent σ-algebras $\mathcal{A}_t$ can have in common only the empty set and the whole space X (otherwise $A \cap (X \backslash A)$ would be nonempty). Hence the measures μ_t yield at once a well-defined single set function on all $\mathcal{A}_t$ (as was assumed from the very beginning in Banach's paper). However, the existence of a further extension is not obvious.

For independent random variables one has the so-called zero–one laws, discussed in §10.10(iv), and laws of large numbers, discussed in §10.10(v).

Let us briefly discuss the concept of conditional independence, which is useful for the study of many probabilistic problems, in particular, related to limit theorems, Markov processes, and Gibbs measures.

Let $(\Omega, \mathcal{A}, P)$ be a probability space and let $\mathcal{F}_1, \ldots, \mathcal{F}_n, \mathcal{G} \subset \mathcal{A}$ be sub-σ-algebras. We shall say that $\mathcal{F}_1, \ldots, \mathcal{F}_n$ are conditionally independent with respect to $\mathcal{G}$ (or given $\mathcal{G}$) if for all $B_k \in \mathcal{F}_k$, $k = 1, \ldots, n$, we have

$$P^{\mathcal{G}}(B_1 \cap \cdots \cap B_n) = \prod_{k=1}^{n} P^{\mathcal{G}}(B_k) \text{ a.e.}$$

For an infinite family of σ-algebras $\mathcal{F}_t$, $t \in T$, conditional independence with respect to $\mathcal{G}$ is defined as conditional independence for every finite collection $\mathcal{F}_{t_i}$ with distinct t_i. The concept of conditional independence is transferred to random elements. Random elements ξ and η are called conditionally independent with respect to a random element ζ if $\sigma(\xi)$ and $\sigma(\eta)$ are conditionally independent with respect to $\sigma(\zeta)$.

It is clear that the σ-algebras $\mathcal{F}$ and $\mathcal{G}$ are conditionally independent given $\mathcal{G}$. Independence of the σ-algebras $\mathcal{F}_1$ and $\mathcal{F}_2$ does not imply their conditional independence given $\mathcal{G}$. For example, the coordinate functions on $[-1/2, 1/2]^2$ with Lebesgue measure are independent, but are not conditionally independent given the function $\zeta(x_1, x_2) = x_1 x_2$ because their conditional expectations with respect to $\sigma(\zeta)$ vanish (which is seen from the fact that the integral of $x_1(x_1x_2)^n$ vanishes for all $n = 0, 1, \ldots$).

As we shall now see, conditional independence means that enlarging $\mathcal{G}$ by $\mathcal{F}$ does not change the corresponding conditional expectations.

10.10.5. Proposition. *Sub-σ-algebras $\mathcal{F}$ and $\mathcal{E}$ are conditionally independent with respect to a sub-σ-algebra $\mathcal{G}$ if and only if for every $E \in \mathcal{E}$*

$$P^{\sigma(\mathcal{F}\cup\mathcal{G})}(E) = P^{\mathcal{G}}(E) \quad \textit{a.e.}$$

PROOF. Conditional independence yields that for any $F \in \mathcal{F}$, $G \in \mathcal{G}$, $E \in \mathcal{E}$ we have

$$\begin{aligned}\int_{F\cap G} P^{\mathcal{G}}(E)\,dP &= \int_{\Omega} P^{\mathcal{G}}(F)P^{\mathcal{G}}(G)P^{\mathcal{G}}(E)\,dP \\ &= \int_{G} P^{\mathcal{G}}(F\cap E)\,dP = P(G\cap F\cap E).\end{aligned}$$

By the monotone class theorem we conclude that for every $A \in \sigma(\mathcal{F}\cup\mathcal{G})$, the integral of $I_A P^{\mathcal{G}}(E)$ equals $P(A\cap E)$, which gives the indicated equality. If this equality holds, then for all $F \in \mathcal{F}$ and $E \in \mathcal{E}$ we have

$$\mathbb{E}^{\mathcal{G}}(I_F I_E) = \mathbb{E}^{\mathcal{G}}\mathbb{E}^{\sigma(\mathcal{F}\cup\mathcal{G})}(I_F I_E) = \mathbb{E}^{\mathcal{G}}(I_F \mathbb{E}^{\sigma(\mathcal{F}\cup\mathcal{G})} I_E) = \mathbb{E}^{\mathcal{G}} I_E \mathbb{E}^{\mathcal{G}} I_F,$$

which shows conditional independence. □

10.10.6. Proposition. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, let X be a Souslin space, let $(Y, \mathcal{B})$ and $(Z, \mathcal{E})$ be measurable spaces, and let mappings*

$$\xi\colon (\Omega, \mathcal{A}) \to \bigl(X, \mathcal{B}(X)\bigr), \quad \eta\colon (\Omega, \mathcal{A}) \to (Y, \mathcal{B}), \quad \zeta\colon (\Omega, \mathcal{A}) \to (Z, \mathcal{E})$$

be measurable. Suppose that there exists a random variable θ on Ω uniformly distributed in $[0,1]$ such that θ and (η, ζ) are independent. Then conditional independence of ξ and ζ with respect to η is equivalent to the existence of a measurable mapping $f\colon Y\times[0,1] \to X$ and a random variable $\widetilde{\theta}$ uniformly distributed in $[0,1]$ such that $\widetilde{\theta}$ and (η, ζ) are independent and $\xi = f(\eta, \widetilde{\theta})$ a.e.

Proof. We may assume that $X \subset [0,1]$. If such a function f exists, then it suffices to use conditional independence of (η,θ) and ζ with respect to η, which follows by independence of θ and (η,ζ) and Proposition 10.10.5. If we are given conditional independence, then by Corollary 10.7.7 there exists a measurable function $f\colon\ Y\times[0,1]\to X$ such that the random element $\widetilde{\xi}=f(\eta,\theta)$ has the same distribution as ξ, and (ξ,η) and $(\widetilde{\xi},\eta)$ also have a common distribution. As shown above, $\widetilde{\xi}$ and ζ are conditionally independent with respect to η. According to Proposition 10.10.5 and the equality of the distributions of (ξ,η) and $(\widetilde{\xi},\eta)$ we obtain

$$P\bigl(\widetilde{\xi}\in B|(\eta,\zeta)\bigr)=P(\widetilde{\xi}\in B|\eta)=P(\xi\in B|\eta)=P\bigl(\xi\in B|(\eta,\zeta)\bigr),$$

which yields the equality of the distributions of $(\widetilde{\xi},\eta,\zeta)$ and (ξ,η,ζ). By Corollary 10.7.7, there exists a random variable $\widetilde{\theta}$ uniformly distributed in $[0,1]$ such that the random element $(\xi,\eta,\zeta,\widetilde{\theta})$ has the same distribution as $(\widetilde{\xi},\eta,\zeta,\theta)$. Then $\widetilde{\theta}$ and (η,ζ) are independent, and the random elements $\bigl(\xi,f(\eta,\widetilde{\theta})\bigr)$ and $\bigl(\widetilde{\xi},f(\eta,\theta)\bigr)$ have equal distributions. Since $\widetilde{\xi}-f(\eta,\theta)=0$ a.e., one has $\xi-f(\eta,\widetilde{\theta})=0$ a.e. □

Under very broad assumptions on a probability space $(\Omega,\mathcal{F},P)$ and a measurable space E, for any random element π on Ω with values in E, one can find a random element π' with the same distribution as π and a random variable θ uniformly distributed in $[0,1]$ such that π' and θ are independent. For example, it suffices that Ω and E be Souslin spaces equipped with their Borel σ-algebras and that the measure P be Borel and atomless. This follows from the fact that, given a Borel function $\pi\colon\ [0,1]\to[0,1]$, one can transform Lebesgue measure λ on $[0,1]$ into the measure $\lambda\otimes(\lambda\circ\pi^{-1})$ on $[0,1]^2$. Certainly, one cannot always take $\pi'=\pi$. For example, if $P=\lambda$ and $\pi(t)=t$ on $[0,1]$, then there is no Borel function θ such that $\lambda\circ(\pi,\theta)^{-1}=\lambda\otimes\lambda$.

10.10(ii). Disintegrations

This subsection contains additional information about disintegrations.

10.10.7. Lemma. *Let $(Y,\mathfrak{B},\nu)$ be a probability space such that ν possesses a compact approximating class, let $X\subset Y$ be a set with $\nu^*(X)=1$, let $\mathfrak{F}=\mathfrak{B}_X$, and let $\mu=\nu|_X$ (see Chapter 1 about restrictions of measures). Let $\widehat{\mathfrak{B}}$ denote the σ-algebra generated by $\mathfrak{B}$ and X and let $\widehat{\nu}$ denote the measure on $\widehat{\mathfrak{B}}$ defined by the formula $\widehat{\nu}(A)=\mu(A\cap X)$, $A\in\widehat{\mathfrak{B}}$. Suppose that the measure $\widehat{\nu}$ on $(Y,\widehat{\mathfrak{B}})$ has a disintegration with respect to $\mathfrak{B}$. Then the measure μ has a compact approximating class.*

Proof. We know from §1.12(ii) that one can find a compact class $\mathcal{L}\subset\mathfrak{B}$ that approximates the measure ν and is closed with respect to countable intersections. By hypothesis, the measure $\widehat{\nu}$ has a disintegration $\bigl\{\mathfrak{F}_y,\widehat{\nu}(\,\cdot\,,y)\bigr\}_{y\in Y}$ with respect to $\mathfrak{B}$. Let

$$\mathcal{K}=\bigl\{K\in\mathfrak{F}|\,\exists\,L\in\mathcal{L}\colon\ K=L\cap X\in\mathfrak{F}_y,\widehat{\nu}(K,y)=1,\forall\,y\in L\bigr\}.$$

It is clear that the class $\mathcal{K}$ is closed with respect to countable intersections. We show that $\mathcal{K}$ is a compact class. Suppose that sets $K_n \in \mathcal{K}$ are decreasing and $\bigcap_{n=1}^{\infty} K_n = \varnothing$. For every n we find $L_n \in \mathcal{L}$ such that $L_n \cap X = K_n$, $K_n \in \mathfrak{F}_y$ and $\widehat{\nu}(K_n, y) = 1$ for all $y \in L_n$. We may assume that the sets L_n are decreasing, passing to $\bigcap_{i=1}^{n} L_i$ and using that the sets K_n are decreasing and $\mathcal{L}$ is closed with respect to intersections. Then one has $\bigcap_{n=1}^{\infty} L_n = \varnothing$. Indeed, if $y \in \bigcap_{n=1}^{\infty} L_n$, then by the definition of L_n we arrive at the following contradiction:

$$1 = \lim_{n\to\infty} \widehat{\nu}(K_n, y) = \widehat{\nu}\Bigl(\bigcap_{n=1}^{\infty} K_n, y\Bigr) = \nu(\varnothing, y) = 0.$$

Therefore, there exists m such that $L_m = \varnothing$, whence $K_m = \varnothing$. Thus, $\mathcal{K}$ is a compact class.

Now we show that $\mathcal{K}$ approximates μ. Let $A \in \mathfrak{F}$ and $\varepsilon > 0$. We can find $B_1 \in \mathfrak{B}$ with $B_1 \cap X = A$. Let us choose $L_1 \in \mathcal{L}$ with $L_1 \subset B_1$ such that $\nu(B_1 \backslash L_1) < \varepsilon/2$. By definition we have $L_1 \cap X \in \mathfrak{F}_y$ for $\widehat{\nu}$-a.e. $y \in L_1$ and

$$\int_{L_1} \widehat{\nu}(L_1 \cap X, y)\, \nu(dy) = \mu(L_1 \cap X) = \nu(L_1).$$

Hence there exists a set $B_2 \in \mathfrak{B}$ with $B_2 \subset L_1$ and $\nu(L_1 \backslash B_2) = 0$ such that $L_1 \cap X \in \mathfrak{F}_y$ and $\widehat{\nu}(L_1 \cap X, y) = 1$ for all $y \in B_2$. Next we find a set $L_2 \in \mathcal{L}$ with $L_2 \subset B_2$ such that $\nu(B_2 \backslash L_2) < \varepsilon/4$, and a set $B_3 \in \mathfrak{B}$ such that

$$B_3 \subset L_2,\ \nu(L_2 \backslash B_3) = 0,\ L_2 \cap X \in \mathfrak{F}_y \quad \text{and} \quad \widehat{\nu}(L_2 \cap X, y) = 1 \quad \text{for all } y \in B_3.$$

Continuing our construction by induction we obtain two sequences of sets $B_n \in \mathfrak{B}$ and $L_n \in \mathcal{L}$ such that

$$B_{n+1} \subset L_n \subset B_n,\ \nu(L_n \backslash B_{n+1}) = 0,\ \nu(B_n \backslash L_n) < \varepsilon 2^{-n},\ L_n \cap X \in \mathfrak{F}_y,$$

$$\widehat{\nu}(L_n \cap X, y) = 1 \quad \text{for all } y \in B_{n+1}.$$

Set $L = \bigcap_{n=1}^{\infty} L_n = \bigcap_{n=1}^{\infty} B_n$ and $K = L \cap X$. Then $K \in \mathfrak{F}$ and $K \subset A$. For all $y \in L$ we have $K \in \mathfrak{F}_y$ and $\widehat{\nu}(K, y) = \lim\limits_{n\to\infty} \widehat{\nu}(L_n \cap X, y) = 1$. Hence $K \in \mathcal{K}$. Finally, one has

$$\mu(A \backslash K) = \nu(B_1 \backslash L) = \sum_{n=1}^{\infty} \nu(B_n \backslash B_{n+1}) < \varepsilon.$$

The lemma is proven. □

The following deep result has been obtained in Pachl [**1414**]. The question on its validity remained open for a long time in spite of its very elementary formulation.

10.10.8. Theorem. *Suppose that $(X, \mathfrak{F}, \mu)$ is a probability space such that $\mathfrak{F}$ contains a compact class approximating μ. Let $\mathfrak{F}^*$ be a sub-σ-algebra in $\mathfrak{F}$. Then $\mathfrak{F}^*$ also contains a compact class that approximates $\mu|_{\mathfrak{F}^*}$.*

PROOF. Let μ_0 be the restriction of μ to $\mathfrak{F}^*$. By the existence of an approximating compact class the measure μ has a disintegration $\{\mathfrak{F}_x, \mu(\,\cdot\,,x)\}_{x\in X}$ with respect to $\mathfrak{F}^*$. For every $x \in X$ let

$$\mathfrak{F}_x^* = \mathfrak{F}_x \cap \mathfrak{F}^*.$$

It is readily verified that $\{\mathfrak{F}_x^*, \mu(\,\cdot\,,x)\}_{x\in X}$ is a disintegration of μ with respect to $\mathfrak{F}^*$. Now one can take $Y = X$, $\nu = \mu$, $\mathfrak{B} = \mathfrak{F}^*$ and apply the foregoing lemma, according to which the measure μ_0 on $\mathfrak{F}^*$ has a compact approximating class. □

The role of the compactness condition in the problem of the existence of disintegrations in the case of product spaces has been investigated in Pachl [**1414**], where somewhat different disintegrations have been considered (see also Edgar [**511**], Valadier [**1911**]).

Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be two probability spaces and let λ be a probability measure on $\mathcal{A}\otimes\mathcal{B}$ such that $\lambda\circ\pi_X^{-1} = \mu$ and $\lambda\circ\pi_Y^{-1} = \nu$, where π_X and π_Y are, respectively, the projection operators from $X\times Y$ to X and Y. A family $\{\mathcal{A}_y, \mu_y\}$, $y \in Y$, is called a ν-disintegration of the measure λ if:

(1) for every $y \in Y$, the class $\mathcal{A}_y$ is a σ-algebra in X and μ_y is a probability measure on $\mathcal{A}_y$;

(2) for every $A \subset \mathcal{A}$, there exists a set $Z \subset \mathcal{B}$ such that $\nu(Z) = 0$, $A \in \mathcal{A}_y$ for all $y \in Y\backslash Z$, and the function $y \mapsto \mu_y(A)$ on $\bigl(Y\backslash Z, \mathcal{B}\cap(Y\backslash Z)\bigr)$ is measurable;

(3) for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$, one has

$$\int_B \mu_y(A)\,\nu(dy) = \lambda(A\times B).$$

10.10.9. Remark. Suppose that $Y = X$ and $\mathcal{B} \subset \mathcal{A}$. Let ν be the restriction of μ to $\mathcal{B}$. Let us take for λ the image of the measure μ under the mapping $x \mapsto (x, x)$. Then a disintegration $\{\mathcal{A}_x, \mu(\,\cdot\,,x)\}_{x\in X}$ of μ with respect to $\mathcal{B}$ in the sense of Definition 10.6.1 with probability conditional measures exists precisely when there exists a ν-disintegration $\{\mathfrak{A}_y, \mu(\,\cdot\,,y)\}_{y\in Y}$ of the measure λ (Exercise 10.10.66).

The following result (see the proof in [**1414**, Theorem 3.5]) reinforces Theorem 10.4.14.

10.10.10. Theorem. *Suppose that in the situation described above the measure space $(Y, \mathcal{B}, \nu)$ is complete and that μ has a compact approximating class $\mathcal{K} \subset \mathcal{A}$. Then, the measure λ has a ν-disintegration $\{\mathcal{A}_y, \mu_y\}$, $y \in Y$, such that $\mathcal{K} \subset \mathcal{A}_y$ for all y. If the class $\mathcal{K}$ is closed with respect to finite unions and finite intersections, then such a disintegration can be found with the additional property that $\mathcal{K}$ approximates μ_y for each y.*

According to the following important result from [**1414**], the existence of a compact approximating class is necessary for the existence of disintegrations for all possible λ.

10.10.11. Theorem. *Suppose that a probability space $(X, \mathcal{A}, \mu)$ has the following property: for every complete probability space $(Y, \mathcal{B}, \nu)$ and every probability measure λ on $\mathcal{A}\otimes\mathcal{B}$ with $\lambda\circ\pi_X^{-1} = \mu$ and $\lambda\circ\pi_Y^{-1} = \nu$, there exists a ν-disintegration. Then μ has a compact approximating class $\mathcal{K} \subset \mathcal{A}$.*

This theorem along with the results in §10.6 yields that the class of probability measures μ possessing a ν-disintegration for every probability measure ν coincides with the class of probability measures μ that have disintegrations in the sense of Definition 10.6.1 with probability conditional measures (since in both cases one obtains the class of compact measures). A direct proof of the coincidence of these two classes has been given in Remy [**1548**].

According to Sazonov [**1656**, Theorem 7], analogous results are valid for perfect measures.

10.10.12. Theorem. *Let P be a perfect probability measure on a space $(X, \mathcal{S})$ and let $\mathcal{S}_1$, $\mathcal{S}_2$ be two σ-algebras of measurable sets such that $\mathcal{S}_1$ is countably generated. Then, there exists a function $p(\,\cdot\,,\,\cdot\,)\colon\ \mathcal{S}_1\times X \to [0,1]$ such that:*

(i) *the function $x\mapsto p(E,x)$ is $\mathcal{S}_2$-measurable for every $E\in\mathcal{S}_1$;*

(ii) *$E\mapsto p(E,x)$ is a perfect probability measure on S_1 for every $x\in X$;*

(iii) *for all $E\in\mathcal{S}_1$ and $B\in\mathcal{S}_2$, one has*

$$P(E\cap B) = \int_B p(E,x)\,P(dx).$$

PROOF. By Theorem 7.5.6 any perfect measure has a compact approximating class on every countably generated sub-σ-algebra. □

10.10(iii). Strong liftings

In many special cases (for example, for the interval with Lebesgue measure), there exist liftings with stronger properties.

10.10.13. Definition. *Let X be a topological space and let μ be a Borel (or Baire) measure on X that is positive on nonempty open sets. We shall say that L is a strong lifting on $\mathcal{L}^\infty(\mu)$ if L is a lifting with the following property: $L(f) = f$ for all $f\in C_b(X)$.*

10.10.14. Theorem. *A strong lifting exists in the case of Lebesgue measure on an interval.*

PROOF. Follows by Example 10.5.3 and an obvious modification of the reasoning in Lemma 10.5.2. □

The existence of a strong lifting on a space implies the existence of measurable selections of some special form for mappings to this space. It is known that a strong lifting exists if X is a compact metric space (see A. & C. Ionescu Tulcea [**867**]). It was unknown for quite a long time whether one can omit the assumption of metrizability. It turned out that the answer is negative: Losert [**1190**] constructed his celebrated counter-example.

10.10.15. Theorem. *There exists a Radon probability measure on a compact space of the form* $X = \{0,1\}^\tau$ *such that it is positive on all nonempty open sets and has no strong lifting.*

There exist strong liftings that are not Borel liftings (see Johnson [**915**]).

The next result (see A. & C. Ionescu Tulcea [**867**, Theorem 3, p. 138] establishes a close connection between strong liftings and proper regular conditional measures.

10.10.16. Theorem. *Let* T *be a compact space and let* μ *be a positive Radon measure on* T *with* $\operatorname{supp}\mu = T$*. The following assertions are equivalent*:

(i) *there exists a strong lifting for* μ;

(ii) *for every triple* $\{S,\nu,\pi\}$*, where* S *is a compact space with a positive Radon measure* ν *and* $\pi\colon S\to T$ *is a continuous mapping of* S *onto* T *with* $\mu=\nu\circ\pi^{-1}$*, there exists a mapping* $t\mapsto\lambda_t$ *of the space* T *to the space* $\mathcal{P}_r(S)$ *of Radon probability measures such that the functions* $t\mapsto\lambda_t(E)$, $E\in\mathcal{B}(S)$*, are* μ*-measurable and one has* $\operatorname{supp}\lambda_t\subset\pi^{-1}(t)$ *for every* $t\in T$ *and*

$$\nu(E)=\int_T \lambda_t(E)\,\mu(dt),\quad E\in\mathcal{B}(S).$$

10.10(iv). Zero–one laws

Zero–one laws (0-1 laws) are assertions of the sort that under certain conditions every set in some class has probability either 0 or 1. Let consider some examples. The most important of them is the following 0-1 law of Kolmogorov. Suppose we are given measurable spaces $(X_i,\mathcal{A}_i)$, $i\in\mathbb{N}$. Their product $X=\prod_{i=1}^\infty X_i$ is equipped with the σ-algebra $\mathcal{A}=\bigotimes_{i=1}^\infty\mathcal{A}_i$. Let $\mathcal{X}_n:=\bigotimes_{i=n+1}^\infty\mathcal{A}_i$ and $\mathcal{X}:=\bigcap_{n=1}^\infty\mathcal{X}_n$, where sets from $\mathcal{X}_n$ are naturally identified with subsets of X. The following terms are used for $\mathcal{X}$: the tail σ-algebra, the asymptotic σ-algebra. The class $\mathcal{X}$ contains sets that are unchanged under all transformations of the space X which alter only finitely many coordinates. Typical examples of sets in $\mathcal{X}$ are

$$L:=\Bigl\{x\in\mathbb{R}^\infty\colon\ \exists\lim_{n\to\infty}x_n\Bigr\},\quad S:=\Bigl\{x\in\mathbb{R}^\infty\colon\ \limsup_{n\to\infty}x_n<\infty\Bigr\}.$$

10.10.17. Theorem. *Let* μ_i *be probability measures on* $(X_i,\mathcal{A}_i)$ *and let* $\mu=\bigotimes_{i=1}^\infty\mu_i$*. Then, for every* $E\in\mathcal{X}$*, we have either* $\mu(E)=1$ *or* $\mu(E)=0$*. In particular, every* $\mathcal{X}$*-measurable function a.e. equals some constant.*

PROOF. By Corollary 10.2.4 the functions

$$\int I_E(x_1,\ldots,x_n,x_{n+1},\ldots)\bigotimes_{k=n+1}^\infty\mu_k\bigl(d(x_{n+1},x_{n+2},\ldots)\bigr)$$

converge to I_E a.e. and in $L^1(\mu)$. If $E\in\mathcal{X}$, then these functions are constant, hence I_E a.e. coincides with some constant. It is clear that such a constant can be only 0 or 1. $\square$

As an application of this theorem we note that in the case where $\mathbb{R}^\infty$ is equipped with a measure μ that is the countable product of probability measures on the real line, given a sequence of numbers $c_n > 0$, one has that either $\lim\limits_{n\to\infty} c_n x_n$ exists for a.e. x or there is no limit for a.e. x. Certainly, the theorem does not tell us which of the two cases occurs, but sometimes it is useful to know that no other case is possible. In probabilistic terms, this means that for any sequence of independent random variables x_n, the above limit either exists almost surely or does not exist almost surely. In general, diverse asymptotic properties of sequences of independent random variables are a typical object of applications of zero–one laws.

In the case where all $(X_n, \mathcal{A}_n)$ coincide with the space $(X_1, \mathcal{A}_1)$, one can consider yet another interesting σ-algebra, called the symmetric σ-algebra and defined by the equality $\mathcal{S} := \bigcap_{n=1}^\infty \mathcal{S}_n$, where $\mathcal{S}_n$ is the σ-algebra generated by all $\mathcal{A}$-measurable functions that are invariant with respect to permutations of $x_1, \ldots, x_n$, i.e., functions f such that

$$f(x_1, \ldots, x_n, x_{n+1}, \ldots) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}, x_{n+1}, \ldots)$$

for every permutation σ of the set $\{1, \ldots, n\}$. It is clear that $\mathcal{X}_n \subset \mathcal{S}_n$ and hence $\mathcal{X} \subset \mathcal{S}$. This inclusion, however, may be strict. Indeed, let us consider the set

$$E = \Bigl\{x \in \mathbb{R}^\infty \colon \ \sum_{i=1}^n x_i = 0 \quad \text{for infinitely many } n\Bigr\}.$$

Then $E \in \mathcal{S}$, but $E \notin \mathcal{X}$, since the point $(-1, 1, 0, 0, \ldots)$ belongs to E, but the point $(0, 1, 0, 0, \ldots)$, which differs only in the first coordinate, does not. It turns out that for some classes of measures, the classes $\mathcal{S}$ and $\mathcal{X}$ coincide up to sets of measure zero.

A measure μ on $\mathcal{A}$ will be called invariant with respect to permutations or symmetric if it is invariant with respect to all transformations of X of the form $(x_i) \mapsto (x_{\sigma(i)})$, where σ is a permutation of $\mathbb{N}$ that replaces only finitely many elements. An example of such a measure is the product of identical measures μ_n on $(X_n, \mathcal{A}_n)$.

The following result is the zero–one law of Hewitt and Savage.

10.10.18. Theorem. *Let μ be a probability measure on $\mathcal{A}$ that is invariant with respect to permutations. Then $\mathbb{E}^{\mathcal{X}} = \mathbb{E}^{\mathcal{S}}$ on the space $L^1(\mu)$.*

In particular, if μ is the product of identical measures μ_n, then for all $E \in \mathcal{S}$ we have either $\mu(E) = 1$ or $\mu(E) = 0$.

PROOF. It suffices to verify that $\mathbb{E}^{\mathcal{S}} f = \mathbb{E}^{\mathcal{X}} f$ a.e. for every bounded measurable function f that depends on the coordinates x_i, $i \le n$, since the set of such functions is dense in $L^1(\mu)$ and the operators $\mathbb{E}^{\mathcal{X}}$ and $\mathbb{E}^{\mathcal{S}}$ are continuous on $L^1(\mu)$. Whenever $k > n$ we set $f_k(x) = f(x_{1+k}, \ldots, x_{n+k})$. We observe that $f_k(x) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$, where σ is the permutation of the set $\{1, \ldots, n+k\}$ that interchanges i and $i+k$, $i = 1, \ldots, n$, and leaves unchanged the elements $n+1, \ldots, k$. Since the sequence $\{f_k\}$ is uniformly bounded, there exists a subsequence $\{f_{k_j}\}$ that converges to some function $g \in L^2(\mu)$ in the weak topology of $L^2(\mu)$. Then $\mathbb{E}^{\mathcal{S}} f_{k_j} \to \mathbb{E}^{\mathcal{S}} g$ in the weak topology. The function f_k does not depend on $x_1, \ldots, x_k$, i.e., is measurable with respect to $\mathcal{X}_k$. Hence the function g a.e. equals some $\mathcal{X}$-measurable function h. Since $\mathcal{X} \subset \mathcal{S}$, we obtain that $\mathbb{E}^{\mathcal{S}} g = h = g$ a.e. On the other hand, $\mathbb{E}^{\mathcal{S}} f_k = \mathbb{E}^{\mathcal{S}} f$ a.e. by the invariance of μ. Thus, $\mathbb{E}^{\mathcal{S}} f = \mathbb{E}^{\mathcal{S}} g = g = h$ a.e. The inclusion $\mathcal{X} \subset \mathcal{S}$ and the $\mathcal{X}$-measurability of h yield that $\mathbb{E}^{\mathcal{S}} f = \mathbb{E}^{\mathcal{X}} f$ a.e. The last claim follows by the Kolmogorov zero–one law. $\square$

This theorem means that for every set $E \in \mathcal{S}$, there is a set $E' \in \mathcal{X}$ with $\mu(E \triangle E') = 0$. Indeed, $I_E(x) = \mathbb{E}^{\mathcal{X}} I_E(x)$ a.e. and for E' one can take the set $E' = \bigl\{x\colon\ \mathbb{E}^{\mathcal{X}} I_E(x) = 1\bigr\}$.

The next theorem proved in Ressel [**1557**] generalizes a classical result of de Finetti (see de Finetti [**419**]) and a number of its subsequent improvements (see Hewitt, Savage [**826**], Aldous [**22**], Diaconis, Freedman [**440**]). According to this theorem, any probability measure invariant with respect to permutations is a mixture of product measures.

10.10.19. Theorem. *Let $X = T^\infty$, where T is a completely regular space. Then, for every Radon probability measure μ on X that is invariant with respect to permutations, there exists a Radon probability measure Π on the space $\mathcal{P}_r(T)$ equipped with the weak topology such that*

$$\mu(B) = \int_{\mathcal{P}_r(T)} m^\infty(B)\, \Pi(dm), \quad B \in \mathcal{B}(X),$$

where for any measure $m \in \mathcal{P}_r(T)$, the symbol m^∞ denotes the Radon extension of the countable power of m.

If we are given a sequence of independent random variables ξ_n on a probability space $(\Omega, \mathcal{F}, P)$, then the series $\sum_{n=1}^\infty \xi_n$ either converges a.e. or diverges a.e. The following "Kolmogorov three series theorem" determines which of the two cases occurs. Its proof can be read in Shiryaev [**1700**]. Let $\xi_n^{(c)}(\omega) = \xi_n(\omega)$ if $|\xi_n(\omega)| \le c$ and $\xi_n^{(c)}(\omega) = 0$ if $|\xi_n(\omega)| > c$. Let $\mathbb{E}\xi$ denote the expectation (integral) of a random variable ξ.

10.10.20. Theorem. *Let $\{\xi_n\}$ be a sequence of independent random variables on a probability space $(\Omega, \mathcal{F}, P)$. The series $\sum_{n=1}^\infty \xi_n$ converges a.e. precisely when for every $c > 0$, one has convergence of the series*

$$\sum_{n=1}^\infty P(|\xi_n| \ge c), \quad \sum_{n=1}^\infty \mathbb{E}\xi_n^{(c)}, \quad \sum_{n=1}^\infty \mathbb{E}\bigl(\xi_n^{(c)} - \mathbb{E}\xi_n^{(c)}\bigr)^2.$$

Moreover, convergence of these series for some $c > 0$ is sufficient.

10.10.21. Example. Let random variables ξ_n be independent.

(i) If $|\xi_n| \le c$ for some $c > 0$, then a necessary and sufficient condition of a.e. convergence of the series $\sum_{n=1}^\infty \xi_n$ is convergence of the two series with the terms $\mathbb{E}\xi_n$ and $\mathbb{E}(\xi_n-\mathbb{E}\xi_n)^2$. If, in addition, $\mathbb{E}\xi_n=0$, then only convergence of the series of $\mathbb{E}\xi_n^2$ is required.

(ii) If $\mathbb{E}\xi_n=0$ and the series of $\mathbb{E}\xi_n^2$ converges, then the series $\sum_{n=1}^\infty \xi_n$ converges a.e. Indeed, the Chebyshev inequality yields convergence of the series of $P(|\xi_n|\geq 1)$. The series of $\mathbb{E}|\xi_n^{(1)}|^2$ converges as well, which by the Cauchy–Bunyakowsky inequality yields convergence of the series of $|\mathbb{E}\xi_n^{(1)}|^2$. Hence the series of $\mathbb{E}\bigl(\xi_n^{(1)}-\mathbb{E}\xi_n^{(1)}\bigr)^2$ converges. We note that this partial case is usually proved before Kolmogorov's theorem and is used in its proof.

(iii) One has $\sum_{n=1}^\infty \xi_n^2$ a.e. precisely when $\sum_{n=1}^\infty \mathbb{E}\bigl(\xi_n^2/(1+\xi_n^2)\bigr)<\infty$. Indeed, convergence of the latter series yields a.e. convergence of the series of $\xi_n^2/(1+\xi_n^2)$, which, as one can easily see, is equivalent to convergence of the series of ξ_n^2. If the series of ξ_n^2 converges a.e., then the series of uniformly bounded variables $\xi_n^2/(1+\xi_n^2)$ converges a.e. as well, which gives convergence of their expectations according to (i).

For various special classes of measures and sets, there are other 0-1 laws based on specific features of the involved objects. See Bogachev [**208**], Buczolich [**271**], Dudley, Kanter [**497**], Fernique [**564**], Hoffmann-Jørgensen [**846**], Janssen [**884**], Smolyanov [**1752**], Takahashi, Okazaki [**1825**], Zinn [**2032**], and Exercise 10.10.76.

10.10(v). Laws of large numbers

A law of large numbers is an assertion about convergence of the normalized sums $(\xi_1+\cdots+\xi_n)/n$ for a given sequence of random variables. Results of this kind constitute an important branch in probability theory (see Bauer [**136**], Loève [**1179**], Petrov [**1439**], [**1440**], Révész [**1558**], Shiryaev [**1700**], and references therein). As an example we mention the following theorem due to Kolmogorov.

10.10.22. Theorem. *Suppose that random variables ξ_n are independent, equally distributed and integrable. Then the sequence $(\xi_1+\cdots+\xi_n)/n$ converges a.e. to the expectation of ξ_1.*

We prove a law of large numbers in another case that will be used in the proof of the Komlós theorem stated in Chapter 4.

10.10.23. Theorem. *Let (Ω,P) be a probability space, let $\{\xi_n\}\subset L^2(P)$, and let $\mathbb{E}(\xi_n|\xi_1,\ldots,\xi_{n-1})$ be the conditional expectation of ξ_n with respect to the σ-algebra generated by $\xi_1,\ldots,\xi_{n-1}$. Let us set $\zeta_1:=\xi_1-\mathbb{E}\xi_1$ and $\zeta_n:=\xi_n-\mathbb{E}(\xi_n|\xi_1,\ldots,\xi_{n-1})$ if $n\geq 2$. Then:*

(i) *for all $\varepsilon>0$, $m=0,1,\ldots$ and $n\in\mathbb{N}$, we have*

$$P\Bigl(\max_{1\leq k\leq n}\Bigl|\sum_{j=1+m}^{k+m}\zeta_j\Bigr|\geq\varepsilon\Bigr)\leq\frac{1}{\varepsilon^2}\sum_{k=m+1}^{m+n}\mathbb{E}\zeta_k^2; \qquad (10.10.1)$$

(ii) *if $\sum_{k=1}^\infty \mathbb{E}(\xi_k-\mathbb{E}\xi_k)^2<\infty$, then the series $\sum_{k=1}^\infty\zeta_k$ converges a.e.;*
(iii) *if $\sum_{k=1}^\infty k^{-2}\mathbb{E}(\xi_k-\mathbb{E}\xi_k)^2<\infty$, then $\lim\limits_{n\to\infty} n^{-1}\sum_{k=1}^n\zeta_k=0$ a.e.*

Proof. (i) Let $\varepsilon > 0$, $m \in \{0, 1, \ldots\}$, and $n \in \mathbb{N}$ be fixed. We set

$$A := \Bigl\{x\colon\ \max_{1\le k\le n}\Bigl|\sum_{j=m+1}^{m+k}\zeta_j\Bigr| \ge \varepsilon\Bigr\}, \quad \eta_k := \zeta_{m+1} + \cdots + \zeta_{m+k},$$

$$A_k := \Bigl\{x\colon\ |\eta_1(x)| < \varepsilon, \cdots, |\eta_{k-1}(x)| < \varepsilon, |\eta_k(x)| \ge \varepsilon\Bigr\}.$$

Then $A_i \cap A_j = \varnothing$ if $i \ne j$ and $A = \bigcup_{k=1}^n A_k$. We observe that the functions ζ_i are mutually orthogonal in $L^2(P)$. Moreover, it is readily verified that for any $i < j$ and every set B in the σ-algebra generated by $\xi_1, \ldots, \xi_{j-1}$, one has $(I_B\zeta_i, \zeta_j)_{L^2(P)} = 0$. In particular, for every $k \le n$, one has

$$(\eta_n - \eta_k, I_{A_k}\eta_k)_{L^2(P)} = 0.$$

Hence

$$\begin{aligned}\int_{A_k}\eta_n^2\,dP &= \int_{A_k}\eta_k^2\,dP + \int_{A_k}(\eta_n-\eta_k)^2\,dP + 2\int_{A_k}(\eta_n-\eta_k)\eta_k\,dP\\ &= \int_{A_k}\eta_k^2\,dP + \int_{A_k}(\eta_n-\eta_k)^2\,dP \ge \int_{A_k}\eta_k^2\,dP \ge \varepsilon^2 P(A_k),\end{aligned}$$

whence we obtain

$$\varepsilon^2 P(A) \le \sum_{k=1}^n \int_{A_k}\eta_n^2\,dP \le \sum_{k=m+1}^{m+n}\mathbb{E}\zeta_k^2.$$

(ii) Let $S_k = \zeta_1 + \cdots + \zeta_k$, $\alpha_m(x) := \sup_k |S_{m+k}(x) - S_m(x)|$ and $\alpha(x) := \inf_m \alpha_m(x)$. If $\alpha(x) = 0$, then $\lim\limits_{k\to\infty} S_k(x)$ exists and is finite. Hence it suffices to show that $\alpha(x) = 0$ a.e. According to (10.10.1), for any $m < n$ we have

$$P\Bigl(x\colon\ \sup_{1\le k\le n}|S_{m+k}(x) - S_m(x)| \ge \varepsilon\Bigr) \le \frac{1}{\varepsilon^2}\sum_{k=m+1}^{m+n}\mathbb{E}\zeta_k^2.$$

Therefore, for all m we obtain

$$P\bigl(x\colon\ \alpha(x) \ge \varepsilon\bigr) \le P\bigl(x\colon\ \alpha_m(x) \ge \varepsilon\bigr) \le \frac{1}{\varepsilon^2}\sum_{k=m+1}^{\infty}\mathbb{E}\zeta_k^2.$$

We observe that $\mathbb{E}\zeta_k^2 \le \mathbb{E}(\xi_k - \mathbb{E}\xi_k)^2$, since ζ_k is the orthogonal projection of ξ_k to the closed linear subspace in $L^2(P)$ formed by the functions that are measurable with respect to the σ-algebra generated by $\xi_1, \ldots, \xi_{k-1}$.

(iii) Applying (ii) to the functions ξ_k/k we obtain a.e. convergence of the series $\sum_{k=1}^\infty k^{-1}\zeta_k$, which by the well-known Kronecker lemma yields our claim. □

10.10.24. Corollary. *Let*

$$\sum_{n=1}^\infty n^{-2}\int_\Omega \xi_n^2\,dP < \infty, \quad \lim_{n\to\infty}\mathbb{E}(\xi_n|\xi_1, \ldots, \xi_{n-1}) = 0 \ \textit{a.e.}$$

Then $\lim_{n\to\infty} n^{-1}(\xi_1+\cdots+\xi_n)=0$ *a.e. In particular, this is true if* $\xi_k\in L^2(P)$ *are independent and have zero means.*

Now we are in a position to prove the Komlós theorem.

10.10.25. Theorem. *Let* μ *be a probability measure and let a sequence* $\{\xi_n\}$ *be bounded in* $L^1(\mu)$*. Then, there exist a subsequence* $\{\eta_n\}$ *in* $\{\xi_n\}$ *and a function* $\eta\in L^1(\mu)$ *such that for every subsequence* $\{\eta'_n\}$ *in* $\{\eta_n\}$*, one has almost everywhere* $\lim_{n\to\infty} n^{-1}\bigl(\eta'_1(x)+\cdots+\eta'_n(x)\bigr)=\eta(x)$*.*

PROOF. The main idea of the proof is to achieve a situation where the hypotheses of Theorem 10.10.23 are satisfied. First we show how to pick a subsequence $\{\eta_n\}$ in $\{\xi_n\}$ with the convergent arithmetic means, and then the necessary changes will be described in order to cover all subsequences in $\{\eta_n\}$ as well. One can assume from the very beginning (passing to a subsequence) that

$$\sum_{n=1}^{\infty}\mu\bigl(|\xi_n|\geq n\bigr)<\infty. \tag{10.10.2}$$

For every k, the sequence $\xi_{n,k}:=\xi_n I_{[-k,k]}\circ\xi_n$ is bounded in $L^2(\mu)$ and hence has a weakly convergent subsequence. By the standard diagonal procedure we pick a subsequence $\{\xi'_n\}$ in $\{\xi_n\}$ such that, for every fixed k, the sequence $\xi'_{n,k}=\xi'_n I_{[-k,k]}\circ\xi'_n$ converges weakly in $L^2(\mu)$ to some function β_k as $n\to\infty$. By Proposition 4.7.31, there exists a function $\eta\in L^1(\mu)$ such that

$$\lim_{k\to\infty}\beta_k(x)=\eta(x)\ \text{ a.e. and }\ \lim_{k\to\infty}\|\beta_k-\eta\|_{L^1(\mu)}=0. \tag{10.10.3}$$

One can pick in $\{\xi'_n\}$ a further subsequence $\bigl\{\xi_n^{(1)}\bigr\}$ such that for some number $p_1\in[0,1]$ one has

$$\lim_{n\to\infty}\mu\Bigl(0\leq\bigl|\xi_n^{(1)}\bigr|<1\Bigr)=p_1,\quad \frac{p_1}{2}\leq\mu\Bigl(0\leq\bigl|\xi_n^{(1)}\bigr|<1\Bigr)<p_1+1,\ \forall\, n\in\mathbb{N}.$$

By induction, for every $k\in\mathbb{N}$, we construct a sequence $\bigl\{\xi_n^{(k)}\bigr\}\subset\bigl\{\xi_n^{(k-1)}\bigr\}$ such that for all $n\in\mathbb{N}$, one has

$$\lim_{n\to\infty}\mu\Bigl(k-1\leq\bigl|\xi_n^{(k)}\bigr|<k\Bigr)=p_k,\quad \frac{p_k}{2}\leq\mu\Bigl(k-1\leq\bigl|\xi_n^{(k)}\bigr|<k\Bigr)<p_k+\frac{1}{k^3},$$

where $0\leq p_k\leq 1$. Set $\zeta_n=\xi_n^{(n^2)}$. Then, for the sequence $\{\zeta_n\}$ and each of its subsequences, we have

$$\lim_{n\to\infty}\mu\bigl(k-1\leq|\zeta_n|<k\bigr)=p_k,\quad \forall\, k\in\mathbb{N}, \tag{10.10.4}$$

$$\frac{p_k}{2}\leq\mu\bigl(k-1\leq|\zeta_n|<k\bigr)<p_k+\frac{1}{k^3},\quad \forall\, n\in\mathbb{N},\ k=1,\ldots,n^2. \tag{10.10.5}$$

The last inequality yields

$$\sum_{k=1}^{n^2}kp_k\leq 2\sum_{k=1}^{n^2}k\mu\bigl(k-1\leq|\zeta_n|<k\bigr)\leq 2\bigl(\|\zeta_n\|_{L^1(\mu)}+1\bigr)\leq 2C+2,$$

where $C := \sup_n \|\xi_n\|_{L^1(\mu)}$, whence we obtain

$$\sum_{k=1}^{\infty} kp_k \le 2C + 2. \tag{10.10.6}$$

Now let $\{\eta_n\}$ be an arbitrary subsequence in $\{\zeta_n\}$ and $\overline{\eta}_n := \eta_n I_{[-n,n]} \circ \eta_n$. We show that

$$\sum_{n=1}^{\infty} n^{-2} \|\overline{\eta}_n\|_2^2 \le 4C + 8. \tag{10.10.7}$$

Indeed, by (10.10.5) we have

$$\|\overline{\eta}_n\|_2^2 \le \sum_{k=1}^{n} k^2 \mu\bigl(k-1 \le |\eta_n| < k\bigr) < \sum_{k=1}^{n} k^2 (p_k + k^{-3}).$$

In view of (10.10.6) and the estimate $\sum_{n=k}^{\infty} n^{-2} \le 2k^{-1}$ this yields

$$\sum_{n=1}^{\infty} n^{-2} \|\overline{\eta}_n\|_2^2 < \sum_{n=1}^{\infty} \sum_{k=1}^{n} n^{-2} k^2 (p_k + k^{-3}) = \sum_{k=1}^{\infty} k^2 (p_k + k^{-3}) \sum_{n=k}^{\infty} n^{-2}$$
$$\le 2 \sum_{k=1}^{\infty} k(p_k + k^{-3}) \le 4C + 8.$$

Similarly to (10.10.7) one proves the estimate

$$\sum_{n=1}^{\infty} n^{-2} \|\beta_n\|_2^2 \le 4C + 8. \tag{10.10.8}$$

Indeed, let $\zeta_{n,k} := \zeta_n I_{[-k,k]} \circ \zeta_n$. For any $m \ge n$ we have by (10.10.5)

$$\|\zeta_{m,n}\|_2^2 < \sum_{k=1}^{n} k^2 \mu\bigl(k-1 \le |\zeta_m| < k\bigr) < \sum_{k=1}^{n} k^2 (p_k + k^{-3}).$$

Hence $\|\beta_n\|_2^2 \le \sum_{k=1}^{n} k^2 (p_k + k^{-3})$, since $\zeta_{m,n} \to \beta_n$ weakly as $m \to \infty$. As above, we arrive at (10.10.8). By the inequality $\mu(\eta_n \ne \overline{\eta}_n) = \mu(|\eta_n| > n)$ and (10.10.2) we have

$$\sum_{n=1}^{\infty} \mu(\eta_n \ne \overline{\eta}_n) < \infty.$$

By the Borel–Cantelli lemma (see Exercise 1.12.89), for almost every x we obtain $\eta_n(x) = \overline{\eta}_n(x)$ for all $n > n(x)$. Hence the equalities

$$\mu\Bigl(\lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} \overline{\eta}_k = \eta\Bigr) = 1 \quad \text{and} \quad \mu\Bigl(\lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} \eta_k = \eta\Bigr) = 1$$

are equivalent. In view of (10.10.3) it suffices to achieve a situation where, letting $\gamma_k := \overline{\eta}_k - \beta_k$, one has

$$\mu\Bigl(\lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} \gamma_k = 0\Bigr) = 1. \tag{10.10.9}$$

To this end, we pick in $\{\zeta_n\}$ a suitable subsequence $\{\eta_n\}$ as follows. For $a > 0$, we set

$$G_a(t) = ak \quad \text{if } ak \le t < ak + a,\ k \in \mathbb{Z}.$$

Let $\eta_1 = \zeta_1$, $\gamma_1 = \eta_1 I_{[-1,1]} \circ \eta_1 - \beta_1$, $\eta_1' = G_{1/2} \circ \gamma_1$. The function γ_1 is bounded, hence the function η_1' assumes only finitely many values and the σ-algebra generated by η_1' is finite. Let $A_{1,1},\ldots,A_{1,N_1}$ be all sets of positive measure in this σ-algebra. Let $\varepsilon_1 = \min_{1\le k\le N_1} \mu(A_{1,k})$. As $m \to \infty$ the sequence $\{\zeta_{m,2}\}$ converges weakly in $L^2(\mu)$ to β_2, since $\zeta_{m,k} = \xi_{m,k}$ whenever $m \ge k^2$ by our choice of η_n. There is m_2 such that

$$\biggl|\int_{A_{1,k}} (\zeta_{m,2} - \beta_2)\, d\mu\biggr| \le \frac{\varepsilon_1}{2}, \quad \forall\, k = 1,\ldots,N_1,\ \forall\, m \ge m_2.$$

Let $\eta_2 = \zeta_{m_2}$, $\gamma_2 = \eta_2 - I_{[-2,2]} \circ \eta_2 - \beta_2$, and $\eta_2' = G_{\varepsilon_1/4} \circ \gamma_2$. Since the functions η_1' and η_2' assume only finitely many values, they generate a finite σ-algebra. Let $A_{2,1},\ldots,A_{2,N_2}$ be all sets of positive measure in this σ-algebra. Let $\varepsilon_2 = \min_{1\le k\le N_2} \mu(A_{2,k})$. As above, the sequence of functions $\zeta_{m,3} - \beta_3$ converges weakly to zero in $L^2(\mu)$ and hence there exists $m_3 > m_2$ with

$$\biggl|\int_{A_{2,k}} (\zeta_{m,3} - \beta_3)\, d\mu\biggr| \le \frac{\varepsilon_2}{3}, \quad \forall\, k = 1,\ldots,N_2,\ \forall\, m \ge m_3.$$

We set $\eta_3 = \zeta_{m_3}$, $\gamma_3 = \eta_3 I_{[-3,3]} \circ \eta_3 - \beta_3$, $\eta_3' = G_{\varepsilon_2/8} \circ \gamma_3$ and continue our construction inductively. Let

$$\eta_n := \zeta_{m_n},\ \gamma_n := \eta_n I_{[-n,n]} \circ \eta_n - \beta_n,\ \eta_n' := G_{\varepsilon_{n-1}/2^n} \circ \gamma_n,$$

and let $\mathcal{E}_n$ be the finite σ-algebra generated by the functions $\eta_1',\ldots,\eta_{n-1}'$. Thus, we obtain numbers $m_n > m_{n-1}$ such that for all $m \ge m_n$ one has

$$\biggl|\int_{A_{n-1,k}} (\zeta_{m,n} - \beta_n)\, d\mu\biggr| \le \frac{\varepsilon_{n-1}}{n}, \quad \forall\, k = 1,\ldots,N_{n-1}, \tag{10.10.10}$$

where $A_{n-1,k}$ are all sets of positive measure in $\mathcal{E}_n$ and

$$\varepsilon_{n-1} = \min_{1\le k\le N_{n-1}} \mu(A_{n-1,k}).$$

We show that (10.10.9) is fulfilled. It follows by the definition of η_n' and $G_{\varepsilon_{n-1}/2^n}$ that

$$0 \le \gamma_n - \eta_n' \le \varepsilon_{n-1} 2^{-n} \le 2^{-n}.$$

Hence

$$0 \le \frac{\gamma_1 + \cdots + \gamma_n}{n} - \frac{\eta_1' + \cdots + \eta_n'}{n} \le \frac{1}{n}.$$

Thus, it suffices to establish that $(\eta_1' + \cdots + \eta_n')/n \to 0$ a.e. This will be done by using Theorem 10.10.23. According to (10.10.7) and (10.10.8) we have

$$\sum_{n=1}^{\infty} n^{-2} \|\gamma_n\|_2^2 < \infty.$$

Hence $\sum_{n=1}^{\infty} n^{-2}\|\eta'_n\|_2^2 < \infty$. It remains to verify that for the conditional expectation with respect to the σ-algebra $\mathcal{A}_{n-1}$ generated by $\eta'_1, \ldots, \eta'_{n-1}$ we have

$$\lim_{n\to\infty} \mathbb{E}^{\mathcal{A}_{n-1}}\eta'_n = 0 \quad \text{a.e.}$$

To this end, by virtue of (10.10.10) we obtain almost everywhere

$$\begin{aligned}|\mathbb{E}^{\mathcal{A}_{n-1}}\eta'_n| &\le \max_{1\le k\le N_{n-1}} \mu(A_{n-1,k})^{-1}\biggl|\int_{A_{n-1,k}} \eta'_n\, d\mu\biggr| \\ &\le \varepsilon_{n-1}^{-1}\biggl|\int_{A_{n-1,k}} \gamma_n\, d\mu\biggr| + \varepsilon_{n-1}^{-1}\biggl|\int_{A_{n-1,k}} [\eta'_n - \gamma_n]\, d\mu\biggr| \\ &\le \varepsilon_{n-1}n^{-1}\varepsilon_{n-1}^{-1} + \varepsilon_{n-1}^{-1}\|\eta'_n - \gamma_n\|_1 \le n^{-1} + 2^{-n}.\end{aligned}$$

Now it is clear how to modify our reasoning in order to have convergence of the arithmetic means of every subsequence in $\{\eta_n\}$ and not only of the sequence itself. In the inductive construction of η_n we shall find positive numbers φ_n as follows. Let $\varphi_1 = 1/2$. Instead of a single function η'_{n-1} we shall consider all possible collections F_{n-1} of functions $G_{\varphi_l/2^{il}}\circ(\eta_{l,n-1}-\beta_{n-1})$, $1 \le l \le n-1$. The finite σ-algebra generated by the functions in the collections $F_1, \ldots, F_{n-1}$ is denoted by $\mathcal{A}_{n-1}$ and the minimum of measures of all sets of positive measure in $\mathcal{A}_{n-1}$ is denoted by φ_n. Then we find numbers $m_{n,k}$ such that for every set $A \in \mathcal{A}_{n-1}$ one has the inequality

$$\biggl|\int_A (\zeta_{m,k} - \beta_k)\, d\mu\biggr| \le \varphi_n k^{-1}, \quad \forall\, m \ge m_{n,k}.$$

Finally, let $m_n = \max_{1\le k\le n} m_{n,k}$ and $\eta_n = \zeta_{m_n}$. As above, one verifies that $\{\eta_n\}$ is a required sequence. □

Let us briefly comment on further generalizations of the Komlós theorem. A sequence of numbers s_n is called Cesàro summable to $s \in [0, +\infty]$ if $\dfrac{s_1 + \cdots + s_n}{n} \to s$. Berkes [**157**] has shown that a subsequence in the Komlós theorem can be found in such a way that all its permutations will also be Cesàro summable. von Weizsäcker [**1970**] investigated the role of the condition that the functions ξ_n are integrable and their norms are uniformly bounded. Simple examples show that one cannot completely drop this condition. However, some generalizations in this direction are possible. For example, it is obvious that it suffices to have the above condition with respect to some measure equivalent to the measure μ. This simple observation enlarges considerably the range of admissible sequences. Surprisingly enough, for nonnegative ξ_n this is the best possible extension of the Komlós theorem if one admits only finite functions. We state the corresponding result from von Weizsäcker [**1970**].

10.10.26. Theorem. *Let $\{\xi_n\}$ be a sequence of nonnegative measurable functions on a probability space $(\Omega, \mathcal{F}, P)$. Then, there exist a measurable function ξ with values in $[0, +\infty]$ and a subsequence $\{\xi_{n_k}\}$ such that* *every permutation of* $\{\xi_{n_k}\}$ *is a.e. Cesàro summable to* ξ*, and the sequence* $\{I_{\{\xi<\infty\}}\xi_{n_k}\}$ *is bounded in* $L^1(Q)$ *for some probability measure* Q *equivalent to the measure* P.

Talagrand [**1835**] considered "stable classes" of functions on a probability space $(X, \mathcal{A}, \mu)$. One of the equivalent descriptions of a stable class S of uniformly bounded measurable functions is this:

$$\lim_{k\to\infty} \sup_{f\in S} \Big| \frac{1}{k} \sum_{i=1}^{k} f(x_i) - \int_X f\,d\mu \Big| = 0$$

for a.e. $(x_i) \in X^\infty$ with respect to the countable power of μ.

10.10(vi). Gibbs measures

The fundamental Kolmogorov theorem enables us to construct measures on infinite-dimensional spaces from their finite-dimensional projections. Here we consider a dual (in a certain sense) problem of constructing measures from their conditional measures on finite-dimensional subspaces. Suppose that we are given an infinite index set S (usually in applications S is a countable set like the integer lattice $\mathbb{Z}^d$) and that for every $s \in S$, a measurable space $(X_s, \mathcal{B}_s)$ is given. In typical applications X_s is a set in $\mathbb{R}^n$ or in some manifold. As usual, X^S will denote the space of all collections $x = (x_s)_{s\in S}$, where $x_s \in X_s$. If all X_s coincide with one and the same space X, then X^S is the usual power. For every subset $\Lambda \subset S$, let X^Λ denote the class of all collections $x = (x_s)_{s\in\Lambda}$ with $x_s \in X_s$. The space X^Λ is equipped with the σ-algebra $\mathcal{B}_\Lambda$ generated by the coordinate mappings $\pi_{s_i}\colon (x_s)_{s\in\Lambda} \mapsto x_{s_i}$, $s_i \in \Lambda$, from X^Λ to $(X_{s_i}, \mathcal{B}_{s_i})$. Let π_E denote the natural projection of X^S to X^E for every set $E \subset S$.

10.10.27. Definition. *Suppose that for every finite set* $\Lambda \subset S$ *we are given a transition probability* $P^\Lambda(\,\cdot\,,\,\cdot\,)$ *on* $\mathcal{B}_\Lambda \times X^{S\setminus\Lambda}$. *We shall say that a probability measure* P *on* $\mathcal{B}^S$ *is Gibbs with the conditional distributions* P^Λ *if for every finite set* $\Lambda \subset S$ *one has the equality*

$$P(B) = \int_{X^{S\setminus\Lambda}} P^\Lambda(B, y)\, P\circ\pi_{S\setminus\Lambda}^{-1}(dy), \quad B \in \mathcal{B}^S. \tag{10.10.11}$$

The first questions arising in connection with this definition are whether Gibbs measures exist and whether they are unique. Certainly, in the theory of Gibbs measures there are many other questions. It is worth noting that this theory, which has been created relatively recently and in which unsolved problems are in abundance, is a very promising field of applications of measure theory. We shall briefly consider only the "finite-dimensional" case, i.e., the problem of recovering a measure on a finite product from its conditional measures.

10.10.28. Example. Let $(X_1, \mathcal{B}_1, \lambda_1)$ and $(X_2, \mathcal{B}_2, \lambda_2)$ be two probability spaces and let μ be a measure on $X_1\times X_2$ given by a positive density f with respect to the measure $\lambda_1 \otimes \lambda_2$. Let $\mathcal{B}_1$ and $\mathcal{B}_2$ contain all singletons. Then the measure μ is uniquely determined by the conditional measures $\mu^1(\,\cdot\,, x_2)$ on $X_1 \times \{x_2\}$ and conditional measures $\mu^2(\,\cdot\,, x_1)$ on $\{x_1\} \times X_2$.

PROOF. According to Exercise 9.12.48, the projection of μ on X_1 is given by the density

$$\varrho_1(x_1) = \int_{X_2} f(x_1, x_2)\, \lambda_2(dx_2)$$

with respect to λ_1. The projection of μ on X_2 is given by the density

$$\varrho_2(x_2) = \int_{X_1} f(x_1, x_2)\, \lambda_1(dx_1)$$

with respect to λ_2. In addition, the conditional measures $\mu^1(\,\cdot\,, x_2)$ on the sections $X_1 \times \{x_2\}$ are given by the densities

$$\psi_1(x_1, x_2) := f(x_1, x_2)/\varrho_2(x_2)$$

with respect to the measures $\lambda_1 \otimes \delta_{x_2}$ and the conditional measures $\mu^2(\,\cdot\,, x_1)$ on $\{x_1\} \times X_2$ are given by the densities

$$\psi_2(x_1, x_2) := f(x_1, x_2)/\varrho_1(x_1)$$

with respect to $\delta_{x_1} \otimes \lambda_2$. Thus, we have to recover μ knowing a pair of positive functions ψ_1 and ψ_2. Let us integrate the function $\psi_1(x_1, x_2)/\psi_2(x_1, x_2)$ in x_1 against the measure λ_1. Then we obtain $1/\varrho_2(x_2)$. Thus, the function ϱ_2 is uniquely recovered from the functions ψ_1 and ψ_2. Now we can uniquely recover the measure μ itself: we have found its projection on X_2 and we know the conditional measures for every fixed x_2. $\square$

Note that we have actually used the positivity of the densities ϱ_1 and ϱ_2 and the positivity of conditional densities. For infinite products, however, this is not enough (Exercise 10.10.72). The positivity of conditional densities is essential even in the case of finite products.

10.10.29. Example. Let $E_1 = [0, 1/2] \times [0, 1/2]$, $E_2 = (1/2, 1] \times (1/2, 1]$. Let $f_1(x_1, x_2) = 2$ if $(x_1, x_2) \in E_1 \cup E_2$, $f_1 = 0$ outside $E_1 \cup E_2$, $f_2(x_1, x_2) = 3$ if $(x_1, x_2) \in E_1$, $f_2(x_1, x_2) = 1$ if $(x_1, x_2) \in E_2$. Then f_1 and f_2 are distinct probability densities on $[0, 1]^2$ and their projections to the sides of the square have strictly positive densities with respect to Lebesgue measure. However, both measures have equal conditional measures on the horizontal and vertical lines. For example, on the section of the square with the ordinate $x_2 \in [0, 1/2]$, the corresponding common conditional density equals $2I_{[0,1/2]}(x_1)$, and on the section with the ordinate $x_2 \in (1/2, 1]$, it equals $2I_{(1/2,1]}(x_1)$.

10.10(vii). Triangular mappings

In this subsection we consider an interesting class of measure transformations on product-spaces. Suppose we are given a finite or countable family of measurable spaces $(X_n, \mathcal{A}_n)$. Let $(X, \mathcal{A})$ denote their product. A mapping

$T = (T_1, T_2, \ldots)\colon X \to X$ is called *triangular* if, for every n, the component T_n depends only on $x_1, \ldots, x_n$. In the case where all the spaces X_n coincide with the real line (or are subsets of the real line), the mapping T is called increasing if, for every n, the function $x_n \mapsto T_n(x_1, \ldots, x_{n-1}, x_n)$ is increasing for all fixed $x_1, \ldots, x_{n-1}$. The same terminology is used for mappings defined on subsets of product-spaces. The term "triangular" is explained by the fact that the derivative of a differentiable triangular mapping on $\mathbb{R}^n$ is given by a triangular matrix. Triangular transformations arise naturally in many problems, see, e.g., Knothe [**1013**], Talagrand [**1837**], and the recent papers Aleksandrova [**25**], Bogachev, Kolesnikov [**212**], [**213**], Bogachev, Kolesnikov, Medvedev [**217**], [**218**], on which our exposition is based. In spite of their rather special form, triangular mappings provide us with a powerful tool for transforming measures. Say, the countable product of Lebesgue measures on $[0,1]$ can be transformed by a Borel increasing triangular transformation into an arbitrary Borel probability measure on $[0,1]^\infty$.

We recall that every Borel probability measure μ on the product of two Souslin spaces X_1 and X_2 possesses conditional Borel probability measures μ_{x_1}, $x_1 \in X_1$, on X_2 such that, for every Borel set B in $X_1 \times X_2$, the function $x_1 \mapsto \mu_{x_1}(B^{x_1})$, where $B^{x_1} = \{x_2 \in X_2\colon (x_1, x_2) \in B\}$, is Borel on X_1 and one has

$$\mu(B) = \int_{X_1} \mu_{x_1}(B^{x_1})\, \mu_1(dx_1),$$

where μ_1 is the projection of μ on X_1. Note that given a Borel measure μ on the product of three Souslin spaces X_1, X_2, and X_3, its conditional measures on X_2 serve as conditional measures for its projection on the space $X_1 \times X_2$.

10.10.30. Theorem. (i) *Let X_1 and X_2 be Souslin spaces and let μ and ν be two Borel probability measures on $X_1 \times X_2$. Suppose that the projection of μ onto the first factor and the conditional measures μ_{x_1}, $x_1 \in X_1$, on X_2 have no atoms. Then there exists a Borel triangular mapping $T\colon X_1 \times X_2 \to X_1 \times X_2$ such that $\mu \circ T^{-1} = \nu$.*

(ii) *Let $X = \prod_{n=1}^\infty X_n$, where every X_n is a Souslin space. Let μ be a Borel probability measure on X such that its projection on $\prod_{j=1}^n X_j$ and the conditional measures on X_n have no atoms for all n. Then, for every Borel probability measure ν on X, there exists a triangular Borel mapping $T\colon X \to X$ such that $\mu \circ T^{-1} = \nu$.*

Proof. (i) First we consider the case $X_1 = X_2 = [0,1]$. Let μ_1 and ν_1 denote the projections on the first factor. There exists a monotone function T_1 such that $\mu_1 \circ T_1^{-1} = \nu_1$. We shall use the canonical version of this function defined by the formula $T_1(t) = G_{\nu_1}\bigl(F_{\mu_1}(t)\bigr)$, where

$$G_{\nu_1}(t) := \inf\bigl\{s \in [0,1]\colon\ F_{\nu_1}(s) \geq t\bigr\},$$

and F_{μ_1} and F_{ν_1} are the distribution functions of the measures μ_1 and ν_1, respectively, i.e., $F_{\mu_1}(t) = \mu_1\bigl([0,t)\bigr)$. It is readily seen that G_{ν_1} is increasing. In addition, it is left-continuous. Indeed, if a sequence $\{t_i\}$ increases to t and $G_{\nu_1}(t_i) \le G_{\nu_1}(t) - \varepsilon$ for some $\varepsilon > 0$, then $F_{\nu_1}\bigl(G_{\nu_1}(t) - \varepsilon\bigr) \ge t_i$, hence $F_{\nu_1}\bigl(G_{\nu_1}(t) - \varepsilon\bigr) \ge t$, which is impossible. Therefore, T_1 is increasing and left-continuous as well. The function F_{μ_1} transforms μ_1 into Lebesgue measure (see Example 3.6.2) and the function G_{ν_1} transforms Lebesgue measure into ν_1. In Theorem 8.5.4, for transforming Lebesgue measure we used the function $\xi_{\nu_1}(t) = \sup\bigl\{s \in [0,1]\colon\ F_{\nu_1}(s) \le t\bigr\}$, but this function may differ from G_{ν_1} only at countably many points. For every $x_1 \in [0,1]$, we take the above-defined canonical increasing function $x_2 \mapsto T_2(x_1, x_2)$ that takes μ_{x_1} to $\nu_{T_1(x_1)}$. The function T_2 is Borel. Indeed, it is increasing and left-continuous in x_2. Hence its Borel measurability follows by its Borel measurability in x_1 for every fixed x_2 (see Lemma 6.4.6). In order to verify the Borel measurability in x_1 we recall that

$$T_2(x_1, x_2) = G^{T_1(x_1)}\Bigl(\mu_{x_1}\bigl([0, x_2)\bigr)\Bigr),$$

where the function $x_1 \mapsto \mu_{x_1}\bigl([0, x_2)\bigr)$ on $[0,1]$ is Borel and

$$G^z(t) := \inf\bigl\{s \in [0,1]\colon\ \nu_z\bigl([0,s)\bigr) \ge t\bigr\}, \quad t \in [0,1].$$

Therefore, it is sufficient to verify the Borel measurability of the function $g(z) := G^z(t)$ with respect to z for every fixed t, since $x_1 \mapsto T_1(x_1)$ is a Borel function. Thus, we consider the function

$$g(z) = \inf\bigl\{s\colon\ \nu_z\bigl([0,s)\bigr) \ge t\bigr\}.$$

According to our choice of conditional measures, the Borel measurability of g follows by Exercise 6.10.85. Let us verify that $\nu = \mu \circ T^{-1}$. Let $E = A \times B$, where A and B are Borel sets. Then one has

$$\begin{aligned}
\mu \circ T^{-1}(E) &= \int_0^1 \int_0^1 I_E\bigl(T(x)\bigr)\, \mu_{x_1}(dx_2)\, \mu_1(dx_1) \\
&= \int_0^1 I_A\bigl(T_1(x_1)\bigr) \int_0^1 I_B\bigl(T_2(x_1, x_2)\bigr)\, \mu_{x_1}(dx_2)\, \mu_1(dx_1) \\
&= \int_0^1 I_A\bigl(T_1(x_1)\bigr) \int_0^1 I_B(y_2)\, \nu_{T_1(x_1)}(dy_2)\, \mu_1(dx_1) \\
&= \int_0^1 \int_0^1 I_A(y_1) I_B(y_2)\, \nu_{y_1}(dy_2)\, \nu_1(dy_1) = \nu(E).
\end{aligned}$$

In the general case there exist injective Borel functions $h_i\colon X_i \to [0,1]$. Hence we may assume that the spaces X_i are Souslin subsets of the interval $[0,1]$. Extending both measures to $[0,1]^2$ we find the mapping T constructed above. The set X_1 contains a Borel subset Y_1 of full measure with respect to μ_1 such that $T_1(Y_1) \subset X_1$. Outside Y_1 we redefine T_1 by some constant value from X_1. This gives a Borel function $\widetilde{T}_1$ on X_1 with values in X_1 that μ_1-a.e. equals T_1. Finally, one can find a Borel function $\widetilde{T}_2$ on $X_1 \times X_2$ with values in X_2 such that $\widetilde{T}_2(x_1, x_2) = T_2(x_1, x_2)$ for μ-a.e. (x_1, x_2). To this end, we observe that $\mu\bigl((x_1, x_2) \in X_1 \times X_2\colon\ T_2(x_1, x_2) \in X_2\bigr) = 1$. Indeed, the indicated set is Souslin. For μ_1-almost every fixed x_1, the conditional measure μ_{x_1} is concentrated on X_2, and also for ν_1-almost every fixed y_1, the conditional measure ν_{y_1} is concentrated on X_2. Hence for μ_1-a.e. x_1, the conditional measure $\nu_{T_1(x_1)}$ is concentrated on X_2, i.e., one has the inclusion $T_2(x_1, x_2) \in X_2$ for μ-a.e. (x_1, x_2).

(ii) Induction on n proves our assertion for every finite product of the spaces X_j. Denoting by μ_n and ν_n the projections of μ and ν on $\prod_{j=1}^n X_j$ and using the finite product case we obtain Borel mappings T_n from $\prod_{j=1}^n X_j$ to X_n such that $\mu_n \circ (T_1, \ldots, T_n)^{-1} = \nu_n$ for all n. Then $\mu \circ T^{-1} = \nu$, where $T = (T_n)_{n=1}^\infty$. □

In the case where the spaces X_n coincide with the interval $[0, 1]$, the Borel triangular mappings constructed above have the property that the functions $x_k \mapsto T(x_1, \ldots, x_k)$ are increasing and left continuous. We shall call these increasing Borel triangular mappings *canonical triangular mappings*. A canonical triangular transformation of a measure μ to a measure ν will be denoted by $T_{\mu,\nu}$. In the case where the measures μ and ν are defined on all of $\mathbb{R}^n$, an analogous construction yields a triangular increasing Borel mapping $T_{\mu,\nu} = (T_1, \ldots, T_n)$ with values in $\mathbb{R}^n$ defined on some Borel set $\Omega \subset \mathbb{R}^n$ of full μ-measure. Moreover, every function T_k as a function of the variables $x_1, \ldots, x_k$ is defined on some Borel set in $\mathbb{R}^k$ whose intersections with the straight lines parallel to the kth coordinate line are intervals. This is obvious from our inductive construction and the one-dimensional case, in which the composition $G_{\nu_1} \circ F_{\mu_1}$ is defined either on the whole real line or on a ray or on an interval (if the function G_{ν_1} has no finite limits at the points 0 and 1 and the measure μ_1 is concentrated on a bounded interval). For example, if μ is Lebesgue measure on $[0, 1]$ considered on the whole real line and ν is the standard Gaussian measure, then the mapping $T_{\mu,\nu}$ is defined on the interval $(0, 1)$, but has no increasing extension to the whole real line. If the measure ν on $\mathbb{R}^n$ has a bounded support, then the mapping $T_{\mu,\nu}$ is defined on all of $\mathbb{R}^n$. The same is true for any measure ν if the projection of μ on the first coordinate line and its conditional measures on the other coordinate lines are not concentrated on bounded sets. For example, this is the case if the measure μ is equivalent to Lebesgue measure because one can take a strictly positive Borel version of its density. We observe that the case of $\mathbb{R}^n$ reduces to that of $[0, 1]^n$. To this end, by using the mapping $(x_1, \ldots, x_n) \mapsto (\mathrm{arctg} x_1, \ldots, \mathrm{arctg} x_n)$ and its inverse we pass from $\mathbb{R}^n$ to $(0, 1)^n$ (this preserves the class of increasing triangular Borel mappings). Given two measures μ and ν on $(0, 1)^n$, we take the mapping $T_{\mu,\nu}$ on the cube $[0, 1]^n$ corresponding to their extensions to this cube and let $\Omega = T_{\mu,\nu}^{-1}\bigl((0, 1)^n\bigr)$.

Since conditional measures are uniquely determined up to sets of measure zero, canonical triangular mappings are defined up to modifications, too. However, we shall now see that the uniqueness of a canonical mapping holds in a broader class of transformations.

10.10.31. Lemma. *Let μ and ν be two Borel probability measures on $\mathbb{R}^n$ possessing atomless projections on the first coordinate line and atomless* *conditional measures on the other coordinate lines. Then the mapping $T_{\mu,\nu}$ is injective on a Borel set of full μ-measure. The same is true for measures on $\mathbb{R}^\infty$.*

Proof. It suffices to consider the case of $\mathbb{R}^n$ because in the case of an infinite product we obtain the injectivity on the set of full μ-measure that is the intersection of the sets $E_n\times\mathbb{R}^1\times\mathbb{R}^1\times\cdots$ of full μ-measure, where E_n is a Borel set in $\mathbb{R}^n$ of full measure with respect to the projection of μ such that the mapping $(T_1,\ldots,T_n)$ is injective on E_n. The conditional measures on the first n coordinate lines for the projection of μ on $\mathbb{R}^n$ are atomless, since they coincide with the corresponding conditional measures of the measure μ. In the case $n=1$ the mapping $T_{\mu,\nu}$ is strictly increasing on the set $\mathbb{R}^1\backslash\bigcup_{k=1}^\infty[a_k,b_k]$, where $\mathbb{R}^1\backslash\bigcup_{k=1}^\infty(a_k,b_k)$ is the topological support of μ. The multidimensional case is justified by induction. To this end, we take a set $E\subset\mathbb{R}^{n-1}$ with $\mu_{n-1}(E)=1$ on which the mapping $(T_1,\ldots,T_{n-1})$ is injective. The set $E\times\mathbb{R}^1$ contains a set of full μ-measure on which $T_{\mu,\nu}$ is injective, since for every $y=(x_1,\ldots,x_{n-1})\in E$, the function $t\mapsto T_n(x_1,\ldots,x_{n-1},t)$ is injective on a set of full μ_y-measure. □

10.10.32. Lemma. *Let μ be a Borel probability measure on $\mathbb{R}^\infty$. Suppose we are given two increasing triangular Borel mappings $T=(T_n)_{n=1}^\infty$ and $S=(S_n)_{n=1}^\infty$ such that $\mu\circ T^{-1}=\mu\circ S^{-1}$ and, for every n, the mapping $(T_1,\ldots,T_n)$ is injective on a Borel set of full measure with respect to the projection of μ on $\mathbb{R}^n$. Then $T(x)=S(x)$ for μ-a.e. x.*

In particular, if the projections of the measures μ and ν on the spaces $\mathbb{R}^n$ are absolutely continuous, then there exists a canonical triangular mapping $T_{\mu,\nu}$, and it is unique up to μ-equivalence in the class of increasing Borel triangular mappings transforming μ into ν.

Proof. Clearly, the assertion reduces to the case of $\mathbb{R}^n$. Let us prove it by induction on n. Let $n=1$. Suppose that a point x_0 belongs to the topological support of μ. If $T(x_0)<S(x_0)$, then x_0 cannot be an atom of μ, since $\mu\bigl(x\colon T(x)<t\bigr)=\mu\bigl(x\colon S(x)<t\bigr)$ for all t, and one can take $t=\bigl(T(x_0)+S(x_0)\bigr)/2$. Now we may assume that both functions T and S are continuous at x_0, since the sets of their discontinuity points are at most countable. By the continuity of both functions at x_0, there exists a point $x_1>x_0$ that is not an atom of μ such that the functions T and S are continuous at x_1 and $T(x_1)<S(x_0)$. Taking $t=T(x_1)$ we obtain that there exists a point $y<x_0$ such that $\mu\bigl((y,x_1)\bigr)=0$, contrary to the fact that x_0 belongs to the topological support of μ.

Suppose our assertion is already proven for some $n\geq 1$. Let us consider the case of $\mathbb{R}^{n+1}$. Set $\nu:=\mu\circ T^{-1}=\mu\circ S^{-1}$. Denote by μ_n and ν_n the projections of μ and ν on $\mathbb{R}^n$. On the last coordinate axis we fix conditional measures μ_y and ν_y, $y\in\mathbb{R}^n$. By the inductive assumption, whenever $i\leq n$, we have $T_i(x)=S_i(x)$ for μ-a.e. x. Indeed, the images of the measure μ_n under the mappings $T_0:=(T_1,\ldots,T_n)$ and $S_0:=(S_1,\ldots,S_n)$ are equal (they coincide with ν_n). This gives $T_0 = S_0$ μ_n-a.e., which is equivalent to the equality of these mappings μ-a.e., since they depend only on $y := (x_1, \dots, x_n)$. Now let us show that for μ_n-a.e. $y = (x_1, \dots, x_n)$, we have the equality $T_{n+1}(x_1, \dots, x_n, x_{n+1}) = S_{n+1}(x_1, \dots, x_n, x_{n+1})$ for μ_y-a.e. x_{n+1}. To this end, by the one-dimensional case it suffices to verify the equality μ_n-a.e. of the measures $\mu_y \circ F_y^{-1}$ and $\mu_y \circ G_y^{-1}$, where

$$F_y(t) = T_{n+1}(x_1, \dots, x_n, t), \quad G_y(t) = S_{n+1}(x_1, \dots, x_n, t).$$

By hypothesis, there exists a Borel set $E \subset \mathbb{R}^n$ with $\mu_n(E) = 1$ such that the mapping $T_0 = S_0$ is Borel and injective on E. One can find a Borel mapping J on $\mathbb{R}^n$ such that $J\bigl(T_0(y)\bigr) = J\bigl(S_0(y)\bigr) = y$ for all $y \in E$. Let us take a countable family of bounded Borel functions φ_i on $\mathbb{R}^n$ separating the Borel measures, and an analogous countable family of functions ψ_j on the real line. Set $\zeta_i = \varphi_i \circ J$. Then $\zeta_i\bigl(S_0(y)\bigr) = \zeta_i\bigl(T_0(y)\bigr) = \varphi_i(y)$ for all $y \in E$, i.e., μ_n-a.e. For all i and j, one has the equality

$$\begin{aligned}\int_{\mathbb{R}^{n+1}} \zeta_i(y)\psi_j(t)\, \nu(dydt) &= \int_{\mathbb{R}^{n+1}} \zeta_i\bigl(S_0(y)\bigr)\psi_j\bigl(S_{n+1}(y,t)\bigr)\, \mu(dydt) \\ &= \int_{\mathbb{R}^n} \Bigl(\int_{\mathbb{R}^1} \psi_j\bigl(S_{n+1}(y,t)\bigr)\, \mu_y(dt)\Bigr)\varphi_i(y)\, \mu_n(dy) \\ &= \int_{\mathbb{R}^n} \Bigl(\int_{\mathbb{R}^1} \psi_j(t)\, \mu_y \circ G_y^{-1}(dt)\Bigr)\varphi_i(y)\, \mu_n(dy).\end{aligned}$$

The same equality is fulfilled for the measures $\mu_y \circ F_y^{-1}$ in place of $\mu_y \circ G_y^{-1}$. According to our choice of the functions φ_i and ψ_j we obtain the equality $\mu_y \circ G_y^{-1} = \mu_y \circ F_y^{-1}$ for μ_n-a.e. y. $\square$

The assumption that ν possesses atomless conditional measures on the coordinate lines is essential for the uniqueness statement. Indeed, let μ be Lebesgue measure on $[0,1]^2$ and let $T_1(x_1) = S_1(x_1) = 0$, $T_2(x_1, x_2) = x_2$, $S_2(x_1, x_2) = (x_2 + 1)/2$ if $0 \le x_1 \le 1/2$, and $S_2(x_1, x_2) = (x_2 - 1)/2$ if $1/2 < x_1 \le 1$. Then T and S transform μ into Lebesgue measure on the unit interval of the second coordinate line.

10.10.33. Theorem. *Let $\{\mu_j\}$ and $\{\nu_j\}$ be two sequences of Borel probability measures on $\mathbb{R}^\infty$ convergent in variation to measures μ and ν, respectively. Suppose that the measures μ_j and μ satisfy the hypotheses of Theorem* 10.10.30. *Then the canonical triangular mappings T_{μ_j,ν_j}, extended in an arbitrary way to Borel mappings of the whole space outside their initial domains, converge in measure μ to the mapping $T_{\mu,\nu}$.*

PROOF. If follows from our previous considerations that it suffices to consider the case of measures on $[0,1]^n$. Moreover, it suffices to show that every subsequence in the given sequence of mappings has a further subsequence that converges almost everywhere.

First we consider the case when all the measures μ_j coincide with μ. In fact, we need the case where μ is Lebesgue measure. Let $n = 1$. Then

$\lim_{j\to\infty} T_{\mu,\nu_j}(t) = T_{\mu,\nu}(t)$ for almost every t, since μ has no atoms and $\lim_{j\to\infty} G_{\nu_j}(u) = G_\nu(u)$ for all points $u \in [0,1]$ at which the function G_ν is continuous, i.e., with the exception of an at most countable set (in the case of Lebesgue measure $T_{\mu,\nu_j} = G_{\nu_j}$). Suppose the theorem is proved for some $n \geq 1$ and we are given probability measures ν_j convergent in variation to a measure ν on $I_{n+1} := [0,1]^{n+1}$. It suffices to verify that every subsequence in $\{T_{\mu,\nu_j}\}$ contains a subsequence convergent μ-a.e.

Denote by π_n the projection on $I_n = [0,1]^n$ and let $\mu_0 := \mu \circ \pi_n^{-1}$, $\nu_0 = \nu \circ \pi_n^{-1}$, $T_{\mu,\nu_j} = (T_1^j, \ldots, T_{n+1}^j)$, $T_{\mu,\nu} = (T_1, \ldots, T_{n+1})$. Let ν_y and ν_y^j, $y \in I_n$, denote the conditional measures for ν and ν^j corresponding to the factorization $I_{n+1} = I_n \times [0,1]$. By the inductive assumption and Riesz's theorem we may assume that the mappings $(T_1^j, \ldots, T_n^j)$ converge μ_0-a.e. to the mapping $(T_1, \ldots, T_n)$, since by our construction they coincide with the canonical mappings $S_j := T_{\mu_0, \nu_j \circ \pi_n^{-1}}$ and $S := T_{\mu_0,\nu_0}$ on I_n. It follows by the above inductive construction of the components of canonical mappings and the considered one-dimensional case that in order to have convergence of the functions T_{n+1}^j to T_{n+1} it suffices to obtain weak convergence of the one-dimensional conditional measures $\nu_{S_j(y)}^j$ to the conditional measure $\nu_{S(y)}$ for μ_0-almost all $y \in I_n$. In turn, for every fixed $k \in \mathbb{N}$ letting

$$\psi_j(y) := \int_0^1 t^k\, \nu_y^j(dt), \quad \psi(y) := \int_0^1 t^k\, \nu_y(dt),$$

it suffices to have convergence μ_0-a.e. of the numbers $\psi_j\bigl(S_j(y)\bigr)$ to $\psi\bigl(S(y)\bigr)$. Moreover, as observed above, it suffices to ensure this for some subsequence of indices j. According to Proposition 10.4.23, passing to a subsequence, we may assume that the measures ν_z^j converge in variation to the measure ν_z for ν_0-a.e. z. Then the functions $z \mapsto \psi_j(z)$ converge ν_0-almost everywhere to the function $z \mapsto \psi(z)$. By convergence of the measures $\mu_0 \circ S_j^{-1}$ to the measure $\mu_0 \circ S^{-1}$ in variation and Corollary 9.9.11 we obtain convergence of the functions $\psi_j\bigl(S_j(y)\bigr)$ to $\psi\bigl(S(y)\bigr)$ in measure μ_0. Passing to a subsequence once again we obtain convergence almost everywhere.

Now let us consider another special case where a sequence of measures μ_j convergent in variation is transformed into Lebesgue measure λ on $[0,1]^n$. In this case all the components of our canonical triangular mappings transform the conditional measures (or one-dimensional projections) into Lebesgue measure, i.e., are the distribution functions of the corresponding measures. Therefore, arguing by induction, it suffices to pass to a subsequence of measures for which one has convergence in variation for the conditional measures.

Finally, in the general case we have $T_{\mu_j,\nu_j} = T_{\lambda,\nu_j} \circ T_{\mu_j,\lambda}$. In view of the two cases considered above the sequences of mappings $T_{\mu_j,\lambda}$ and T_{λ,ν_j} converge in measure with respect to the measures μ and λ, correspondingly. Since the measures $\mu \circ T_{\mu_j,\lambda}^{-1}$ converge in variation to the measure λ (this follows by the fact that $\mu_j \circ T_{\mu_j,\lambda}^{-1} = \lambda$ and $\|\mu_j - \mu\| \to 0$), Corollary 9.9.11 used above yields the desired convergence. We recall that if the projections of μ and μ_j to all subspaces $\mathbb{R}^n$ are equivalent to Lebesgue measure, then the canonical triangular mappings are defined on the whole space from the very beginning. $\square$

It follows from the theorem that some subsequence of mappings T_{μ_j,ν_j} converges to $T_{\mu,\nu}$ almost everywhere with respect to μ. In such a formulation, the theorem extends to countable products of arbitrary Souslin spaces (see Aleksandrova [**25**]), and if the factors are metrizable, then convergence of the whole sequence in measure μ remains valid. As Example 10.4.24 shows, there might be no almost everywhere convergence of the whole sequence T_{μ_j,ν_j}.

We have the following change of variables formula for increasing triangular mappings.

10.10.34. Lemma. *Let* $T = (T_1, \dots, T_n)\colon \mathbb{R}^n \to \mathbb{R}^n$ *be an increasing Borel triangular mapping. Suppose that the functions*

$$x_i \mapsto T_i(x_1, \dots, x_i)$$

are absolutely continuous on bounded intervals for a.e. $(x_1, \dots, x_{i-1}) \in \mathbb{R}^{i-1}$. *Let us set by definition* $\det DT := \prod_{i=1}^n \partial_{x_i} T_i$. *Then for every Borel function* φ *that is integrable on the set* $T(\mathbb{R}^n)$, *the function* $\varphi \circ T \det DT$ *is integrable over* $\mathbb{R}^n$ *and one has*

$$\int_{T(\mathbb{R}^n)} \varphi(y)\, dy = \int_{\mathbb{R}^n} \varphi\bigl(T(x)\bigr) \det DT(x)\, dx. \tag{10.10.12}$$

If the mapping T *is defined only on a Borel set* $\Omega \subset \mathbb{R}^n$ *and every function* T_i *is defined on a Borel set in* $\mathbb{R}^i$ *whose sections by the straight lines parallel to the* i*th coordinate line are intervals and the indicated condition is fulfilled for the compact intervals in those sections, then the same assertion is true with* Ω *in place of* $\mathbb{R}^n$.

PROOF. For $n = 1$ our assertion coincides with the classic change of variables formula for absolutely continuous functions. Next we apply induction on n and assume the assertion to be true in the case of dimension $n-1$. We make the function φ zero outside the Souslin set $T(\mathbb{R}^n)$. Let $S = (T_1, \dots, T_{n-1})$. Then for almost every $y_n \in \mathbb{R}^1$, the function $(y_1, \dots, y_{n-1}) \mapsto \varphi(y_1, \dots, y_n)$ is integrable over $\mathbb{R}^{n-1}$, hence by the inductive assumption and the fact that the mapping S on $\mathbb{R}^{n-1}$ satisfies our hypotheses, we obtain

$$\int_{T(\mathbb{R}^n)} \varphi(y)\, dy = \int_{\mathbb{R}^n} \varphi(y)\, dy = \int_{-\infty}^{+\infty} \int_{\mathbb{R}^{n-1}} \varphi\bigl(S(z), y_n\bigr) \det DS(z)\, dz\, dy_n,$$

which after interchanging the limits of integration and the change of variable $y_n = T_n(z, x_n)$ for fixed $z \in \mathbb{R}^{n-1}$ leads to (10.10.12) by the equality $\det DT = (\det DS)\partial_{x_n} T_n$. A similar reasoning applies to the second case mentioned in the formulation, when T is defined on Ω. $\square$

Let us give a simple sufficient condition on the measures μ and ν ensuring the absolute continuity of the ith component of $T_{\mu,\nu}$ with respect to the variable x_i.

10.10.35. Lemma. *A canonical triangular mapping $T_{\mu,\nu}$ on $\mathbb{R}^n$ that transforms an absolutely continuous probability measure μ to a probability measure ν equivalent to Lebesgue measure satisfies the hypothesis of the preceding lemma.*

PROOF. It suffices to observe that in the one-dimensional case the function $T_{\mu,\nu}$ is absolutely continuous on the intervals, since $T_{\mu,\nu} = G_\nu \circ F_\mu$, where both functions are increasing and absolutely continuous on the intervals. The absolute continuity of F_μ is obvious and the absolute continuity (on every bounded interval) of the function G_ν that is inverse to the absolutely continuous function F_ν follows by the fact that it is continuous, increasing and has Lusin's property (N) (see Exercise 5.8.51). Property (N) follows by the condition $F_\nu' > 0$ a.e. (see Lemma 5.8.13). $\square$

If the measure ν is not equivalent to Lebesgue measure, then the ith component of the canonical triangular mapping may be discontinuous. For example, the canonical mapping of Lebesgue measure on $[0,1]$ to the measure ν with density 2 on $[0,1/4]\cup[3/4,1]$ and 0 on $(1/4,3/4)$ has a jump. Nevertheless, the change of variables formula proven above remains valid without assumption on the absolute continuity made in the lemma if T is a canonical mapping of absolutely continuous measures (certainly, not every increasing Borel triangular mapping has this property).

10.10.36. Proposition. *Let μ and ν be probability measures on $\mathbb{R}^n$ with densities ϱ_μ and ϱ_ν with respect to Lebesgue measure. Then, for the canonical triangular mapping $T_{\mu,\nu} = (T_1,\ldots,T_n)$, we have the equality*

$$\varrho_\mu(x) = \varrho_\nu\bigl(T_{\mu,\nu}(x)\bigr)\det DT_{\mu,\nu}(x) \quad \textit{for } \mu\textit{-a.e. } x, \tag{10.10.13}$$

where $\det DT_{\mu,\nu} := \prod_{i=1}^n \partial_{x_i} T_i$ *exists almost everywhere by the monotonicity of T_i in x_i.*

PROOF. Let us consider first the one-dimensional case. Then $T_{\mu,\nu} = S\circ T$, where T is the canonical mapping of the measure μ to Lebesgue measure λ on $(0,1)$, i.e., the distribution function of the measure μ, and S is the canonical mapping of the measure λ to the measure ν, i.e., the inverse function to the distribution function F_ν of the measure ν. By differentiating the identity $F_\nu\bigl(S(y)\bigr) = y$ we obtain $\varrho_\nu\bigl(S(y)\bigr)S'(y) = 1$ a.e. Indeed, it suffices to observe that if Z is a Lebesgue measure zero set on which the derivative of F_ν does not exist or differs from ϱ_ν, then $S^{-1}(Z)$ has Lebesgue measure zero. This is a direct consequence of the equality $\lambda\circ S^{-1} = \nu$ and the absolute continuity of ν. Now we observe that

$$\varrho_\nu\Bigl(S\bigl(T(x)\bigr)\Bigr)S'\bigl(T(x)\bigr) = 1 \quad \text{for } \mu\text{-a.e. } x.$$

This is clear from the equality $\mu\circ T^{-1} = \lambda$. By using this equality we conclude as above that

$$T'_{\mu,\nu}(x) = S'\bigl(T(x)\bigr)T'(x) \quad \text{for } \mu\text{-a.e. } x.$$

Thus, for μ-a.e. x we obtain

$$\varrho_\nu\bigl(T_{\mu,\nu}(x)\bigr)T'_{\mu,\nu}(x) = \varrho_\nu\bigl(T_{\mu,\nu}(x)\bigr)S'\bigl(T(x)\bigr)T'(x) = T'(x) = \varrho_\mu(x).$$

Next we use induction on n and assume that our assertion is true in dimension $n-1$. We write the points of $\mathbb{R}^n$ in the form (x, x_n), $x \in \mathbb{R}^{n-1}$. Set $\widetilde{T}(x) = \bigl(T_1(x), \ldots, T_{n-1}(x)\bigr)$. The projections of the measures μ and ν on $\mathbb{R}^{n-1}$ are denoted by μ' and ν', and their densities with respect to Lebesgue measure on $\mathbb{R}^{n-1}$ are denoted by $\varrho_{\mu'}$ and $\varrho_{\nu'}$, respectively. We observe that $\widetilde{T}$ coincides with $T_{\mu',\nu'}$. By the inductive assumption one has

$$\varrho_{\mu'}(x) = \varrho_{\nu'}\bigl(\widetilde{T}(x)\bigr)\det D\widetilde{T}(x) \quad \mu'\text{-a.e.} \tag{10.10.14}$$

For μ'-a.e. fixed $x \in \mathbb{R}^{n-1}$, the function $t \mapsto T_n(x,t)$ transforms the one-dimensional conditional density $\varrho_\mu^x(x_n) = \varrho_\mu(x,x_n)/\varrho_{\mu'}(x)$ of the measure μ to the conditional density

$$\varrho_\nu^{\widetilde{T}(x)}(x_n) = \varrho_\nu\bigl(\widetilde{T}(x), x_n\bigr)/\varrho_{\nu'}\bigl(\widetilde{T}(x)\bigr)$$

of the measure ν. According to the one-dimensional case we obtain

$$\frac{\varrho_\mu(x,x_n)}{\varrho_{\mu'}(x)} = \frac{\varrho_\nu\bigl(\widetilde{T}(x), T_n(x,x_n)\bigr)}{\varrho_{\nu'}\bigl(\widetilde{T}(x)\bigr)}\partial_{x_n}T_n(x,x_n) \quad \text{for } \mu_x\text{-a.e. } x_n.$$

By using the equality $\det DT(x,x_n) = \partial_{x_n}T_n(x,x_n)\det D\widetilde{T}(x)$ and relation (10.10.14) we complete the proof. □

We emphasize once again that the partial derivative in the formulation is an almost everywhere existing usual partial derivative, not the one in the sense of distributions (which has a singular component in the case of a function that is not absolutely continuous).

We shall say that a Borel probability measure μ with a twice differentiable density $\exp(-\Phi_n)$ on $\mathbb{R}^n$ is uniformly convex with constant $C > 0$ if Φ_n is a convex function and $D^2\Phi_n(x) \geq C \cdot I$, i.e., $\partial_e^2\Phi_n(x) \geq C$ for every unit vector $e \in \mathbb{R}^n$. A Borel probability measure μ on $\mathbb{R}^\infty$ is called uniformly convex with constant $C > 0$ if its projections on the spaces $\mathbb{R}^n$ are uniformly convex with constant C.

The following result is proved in Bogachev, Kolesnikov, Medvedev [**217**], [**218**]. This result generalizes the inequality obtained by Talagrand [**1837**] in the case of a Gaussian measure.

10.10.37. Theorem. *Suppose that a probability measure μ on $\mathbb{R}^n$ is uniformly convex with constant C (for example, let μ be the standard Gaussian measure). Let ν be an absolutely continuous probability measure on $\mathbb{R}^n$ such* *that for $f := d\nu/d\mu$ one has $f\log f \in L^1(\mu)$. Then, there exists a Borel increasing triangular mapping T such that $\nu = \mu\circ T^{-1}$ and*

$$\int_{\mathbb{R}^n} |x-T(x)|^2\,\mu(dx) \le \frac{2}{C}\int_{\mathbb{R}^n} f(x)\log f(x)\,\mu(dx).$$

In the case of the standard Gaussian measure, one has $C = 1$.

Let $H := l^2$ and $|h|_H := \bigl(\sum_{n=1}^\infty h_n^2\bigr)^{1/2}$. The following theorem is proved in Bogachev, Kolesnikov [**213**].

10.10.38. Theorem. *Suppose that a Borel probability measure μ on $X := \mathbb{R}^\infty$ is uniformly convex with constant $C > 0$. Let $\nu \ll \mu$ be a probability measure and let $f := d\nu/d\mu$.*

(i) *If $f\log f \in L^1(\mu)$, then the canonical triangular mapping $T_{\mu,\nu}$ has the property that*

$$\int_X |T_{\mu,\nu}(x)-x|_H^2\,\mu(dx) \le \frac{2}{C}\int_X f\log f\,d\mu.$$

(ii) *If μ has the form $\mu_1\otimes\mu'$, where μ' is a measure on the product of the remaining real lines, then there exists a Borel triangular mapping T of the form $T(x) = x + F(x)$ with $F\colon X\to H$ such that $\nu = \mu\circ T^{-1}$.*

(iii) *If μ is equivalent to the measure $\mu_{e_1}\colon B\mapsto \mu(B-e_1)$, where $e_1 = (1,0,0,\ldots)$, then there exists a Borel mapping T of the form $T(x) = x+F(x)$ with $F\colon X\to H$ such that $\nu = \mu\circ T^{-1}$.*

The assumptions (ii) and (iii) are fulfilled for the countable power of any uniformly convex measure on the real line. In particular, this theorem applies to the countable power of the standard Gaussian measure on the real line. Consequently, the conclusion is true for every Radon Gaussian measure.

Exercises

10.10.39.° Let $(\Omega,\mathcal{F},P)$ be a probability space, let $\mathcal{F}_n$ be an increasing sequence of σ-algebras generating $\mathcal{F}$, and let $|\xi_n| \le \eta$, where ξ_n and η are integrable. Suppose that $\xi_n\to\xi$ a.e. Prove that $\mathbb{E}^{\mathcal{F}_n}\xi_n\to\xi$ a.e.

Hint: as $\mathbb{E}^{\mathcal{F}_n}\xi\to\xi$ a.e. by the martingale convergence theorem, the assertion reduces to the case $\xi = 0$. Given $\varepsilon > 0$, one can find a set E with $P(E) < \varepsilon$ such that $|\xi_n|\le\varepsilon$ outside E for all $n\ge n_\varepsilon$. Then for all $n\ge n_\varepsilon$ we have

$$\mathbb{E}^{\mathcal{F}_n}|\xi_n| \le \varepsilon + \mathbb{E}^{\mathcal{F}_n}(\eta I_E)\quad\text{a.e.}$$

It remains to observe that $\mathbb{E}^{\mathcal{F}_n}(\eta I_E)\to\eta I_E$ a.e. and ηI_E vanishes outside E.

10.10.40.° (Moy [**1339**], Rota [**1614**]) Show that if μ is a probability measure on a space $(X,\mathcal{F})$ and $T\colon L^p(\mu)\to L^p(\mu)$ is a linear operator for some $p\in[1,\infty)$ such that $\|T\| = 1$, $T1 = 1$ and $T(gTf) = TgTf$ for all $g\in L^\infty(\mu)$, $f\in L^p(\mu)$, then there exists a sub-σ-algebra in $\mathcal{E}\subset\mathcal{F}$ such that $Tf = \mathbb{E}^{\mathcal{E}}f$.

Hint: let T^* be the adjoint operator on $L^q(\mu)$, $q = p(p-1)^{-1}$. It follows that $T^*1 = 1$, as the integral of T^*1 equals 1 by the equality $T1 = 1$ and the estimate $\|T^*1\|_q\le 1$. Note also that $Tf\in L^\infty(\mu)$ if $f\in L^\infty(\mu)$. Indeed, we have

$(Tf)^2 = T(fTf) \in L^p(\mu)$. By induction one has $(Tf)^n = T\bigl(f(Tf)^{n-1}\bigr) \in L^p(\mu)$ for all n. By the equality $\|T\| = 1$ and Hölder's inequality we find

$$\|Tf\|_{np}^{np} \le \|f(Tf)^{n-1}\|_p^p \le \|f\|_{np}^p \|Tf\|_{np}^{np-p},$$

whence it follows that $\|Tf\|_{np} \le \|f\|_\infty$ for all n, hence $Tf \in L^\infty(\mu)$. Let us consider the class Φ of all bounded $\mathcal{F}$-measurable functions φ with $T\varphi = \varphi$ a.e. and denote by $\mathcal{E}$ the σ-algebra generated by Φ. By induction one obtains $T(\varphi)^n = \varphi^n$ for all $n \in \mathbb{N}$ and $\varphi \in \Phi$. Hence $T[\psi(\varphi)] = \psi(\varphi)$ for all polynomials ψ, which gives $T[\psi(\varphi)] = \psi(\varphi)$ for any bounded Borel function ψ. Therefore, $Tg = g$ for every bounded $\mathcal{E}$-measurable function g. Let $f \in \mathcal{L}^\infty(\mu)$. Then $T(Tf) = Tf$, so Tf has a version $\varphi \in \Phi$. Finally, for any $E \in \mathcal{E}$ the integral of the function

$$I_E Tf = TfTI_E = T(fTI_E) = T(fI_E)$$

equals the integral of fI_E, i.e., $\varphi = \mathbb{E}^{\mathcal{E}} f$.

10.10.41.° (Šidák [**1705**]) Let μ be a probability measure on a space $(X, \mathcal{F})$ and let M be a closed linear subspace in $L^2(\mu)$. Show that the following conditions are equivalent:

(i) $1 \in M$ and $\max(f, g) \in M$ for all $f, g \in M$,

(ii) there exists a sub-σ-algebra $\mathcal{E} \subset \mathcal{F}$ such that $M = \mathbb{E}^{\mathcal{E}}(L^2(\mu))$.

Hint: (i) yields that $\mathcal{E} := \{E \in \mathcal{F} \colon I_E \in M\}$ is a σ-algebra. Let L be the closed linear subspace in $L^2(\mu)$ generated by the functions I_E, $E \in \mathcal{E}$. Then $L \subset M$. Note that $\min(f, g) \in M$ if $f, g \in M$. If $-1 \le f \le 0$, then $\max(nf, -1) \to -I_{\{f<0\}}$, which gives $\{f < 0\} \in \mathcal{E}$. It follows that $\{f < c\} \in \mathcal{E}$ for all $f \in M$ and $c \in \mathbb{R}$. Hence $f \in L$, i.e., one has $M = L$. It is readily verified that (ii) implies (i).

10.10.42. (Zięba [**2029**]) Let $(X, \mathcal{A}, \mu)$ be a probability space and let $\mathcal{B} \subset \mathcal{A}$ be a sub-σ-algebra. A sequence $\{\eta_n\}$ of measurable functions is called uniformly $\mathcal{B}$-integrable if for every $\mathcal{B}$-measurable a.e. positive function α, there exists a $\mathcal{B}$-measurable function β such that $\beta(x) > 0$ a.e. and $\sup_n \mathbb{E}^{\mathcal{B}}\bigl(|\eta_n| I_{\{|\eta_n|>\beta\}}\bigr) < \alpha$ a.e.

(i) Prove that if a sequence $\{\xi_n\}$ of integrable functions is such that the sequence of functions ξ_n^+ is uniformly $\mathcal{B}$-integrable, then $\limsup\limits_{n\to\infty} \mathbb{E}^{\mathcal{B}} \xi_n \le \mathbb{E}^{\mathcal{B}} \limsup\limits_{n\to\infty} \xi_n$ a.e. If the sequence $\{\xi_n\}$ is uniformly $\mathcal{B}$-integrable and converges a.e. to ξ, then we have $\lim\limits_{n\to\infty} \mathbb{E}^{\mathcal{B}} \xi_n = \mathbb{E}^{\mathcal{B}} \xi$ a.e.

(ii) Construct an example showing that the usual uniform integrability of ξ_n^+ is not sufficient for the conclusion in (i).

10.10.43. (Blackwell, Dubins [**182**]) Show that if functions $f_n \ge 0$ are integrable with respect to a probability measure μ and converge a.e. to a function $f \in L^1(\mu)$ such that the function $g := \sup_n f_n$ is not integrable, then one can find a probability space $(\Omega, \mathcal{F}, P)$, functions $\varphi_n, \varphi \in L^1(P)$, and a sub-$\sigma$-algebra $\mathcal{E} \subset \mathcal{F}$ such that the sequence $(\varphi, \varphi_1, \varphi_2, \ldots)$ has the same distribution as $(f, f_1, f_2, \ldots)$ (i.e., both sequences induce one and the same measure on $\mathbb{R}^\infty$) and $P\bigl(\omega\colon \lim\limits_{n\to\infty} \mathbb{E}^{\mathcal{E}} \varphi_n(\omega) = \mathbb{E}^{\mathcal{E}} \varphi(\omega)\bigr) = 0$.

10.10.44.° Let $X = [-1/2, 1/2]$ be equipped with the σ-algebra $\mathcal{A}$ of all sets that are either at most countable or have at most countable complements, let $\mathcal{B} = \mathcal{A}$, and let λ be Lebesgue measure. Show that Dirac's measures δ_x serve as regular conditional measures $\lambda^{\mathcal{B}}(\,\cdot\,, x)$. Show that the probability measures $\lambda^x := \delta_{-x}$ as well as the signed measures $\lambda_x := 2\delta_x - \delta_{-x}$ also serve as regular conditional measures for λ. Hence there is no essential uniqueness of regular conditional measures even in the class of probability conditional measures, although μ is separable; in addition, a probability measure may have signed regular conditional measures. Finally, letting $\lambda_x := [x^{-1}+1]\delta_x - x^{-1}\delta_{-x}$ if $x \neq 0$, we get regular conditional measures with non-integrable $\|\lambda_x\|$.

HINT: the first claim is trivial. The second claim follows from the fact that for any countable set A, the functions $\lambda^x(A) = I_{-A}(x)$ and $\lambda_x(A) = 2I_A(x) - I_{-A}(x)$ are $\mathcal{B}$-measurable and their Lebesgue integrals vanish; if $A = X$, then both functions equal 1.

10.10.45° (cf. Krylov [**1066**]) Let E be a Borel (or coanalytic) set in a complete separable metric space M and let $D(E)$ be the space of all mappings $x\colon [0,+\infty) \to E$ that are right-continuous and have left limits. Let $\mathcal{A}$ denote the smallest σ-algebra in $D(E)$ making measurable all mappings $x \mapsto x(t)$, $t \geq 0$. Prove that for every probability measure μ on $\mathcal{A}$ and every sub-σ-algebra $\mathcal{B} \subset \mathcal{A}$, there exists a regular with respect to $\mathcal{B}$ conditional probability on $\mathcal{A}$.

HINT: use that $D(E)$ is a coanalytic set in the Polish space $D(M)$ (see Theorem 6.10.19) and that $\mathcal{A}$ is generated by countably many mappings $x \mapsto x(t)$, $t \in \mathbb{Q}$.

10.10.46° Let $(X,\mathcal{A})$ and $(Y,\mathcal{B})$ be measurable spaces. Suppose that for every $x \in X$, we are given a probability measure μ^x on $\mathcal{B}$ such that the function $x \mapsto \mu^x(B)$ is measurable with respect to $\mathcal{A}$ for all $B \in \mathcal{B}$. Show that for every $E \in \mathcal{A}\otimes\mathcal{B}$, the function $x \mapsto \mu^x(E_x)$, where $E_x := \{y \in Y\colon\ (x,y) \in E\}$, is measurable with respect to $\mathcal{A}$.

HINT: the class $\mathcal{E}$ of all sets $E \in \mathcal{A}\otimes\mathcal{B}$ with the required property is σ-additive and contains the class of all products $A \times B$, where $A \in \mathcal{A}$, $B \in \mathcal{B}$, which is closed with respect to intersections. Hence $\mathcal{E} = \mathcal{A}\otimes\mathcal{B}$ (see §1.9).

10.10.47. (Blackwell, Ryll-Nardzewski [**186**]) Let X and Y be Borel sets in Polish spaces and let $\mathcal{A}$ be a countably generated sub-σ-algebra in $\mathcal{B}(X)$. Suppose we are given a set $S \in \mathcal{A}\otimes\mathcal{B}(Y)$ and a mapping $x \mapsto \mu^x$ from X to $\mathcal{P}(Y)$ such that for all $B \in \mathcal{B}(Y)$, the function $\mu^x(B)$ is measurable with respect to $\mathcal{A}$.

(i) Show that for every $\theta \in [0,1)$, there exists a set $E \in \mathcal{A}\otimes\mathcal{B}(Y)$ such that $E \subset S$, all sections $E_x := \{y\colon\ (x,y) \in E\}$ are closed and $\mu^x(E_x) \geq \theta\mu^x(S_x)$ for all $x \in X$.

(ii) Let $\mu^x(S_x) > 0$ for all $x \in X$. Prove that S contains the graph of some $\bigl(\mathcal{A},\mathcal{B}(Y)\bigr)$-measurable mapping $f\colon\ X \to Y$.

HINT: (i) the class $\mathcal{F}$ of all sets in $\mathcal{A}\otimes\mathcal{B}(Y)$ with the required property admits finite unions, countable unions of increasing sets and countable intersections of decreasing sets. For example, let $F = \bigcap_{n=1}^\infty F^n$, $F^n \in \mathcal{F}$, $F^{n+1} \subset F^n$ and $\theta \in (0,1)$. By the previous exercise the function $\psi\colon\ x \mapsto \mu^x(F_x)$ is measurable with respect to $\mathcal{E}$. Let $X_k := \{(k+1)^{-1} \leq \psi < k^{-1}\}$. Then $\mathcal{A}\otimes\mathcal{B}$ contains a set $E^n \subset F^n$ with closed sections such that $\mu^x(F^n_x\backslash E^n_x) \leq (k+1)^{-1}(1-\theta)2^{-n}$ for all $x \in X_k$. Let $E := \bigcap_{n=1}^\infty E^n$. Then $E \subset F$ and whenever $\mu^x(F_x) > 0$, we have

$$\mu^x(E_x) \geq \mu^x(F_x)\Bigl[1 - \sum_{n=1}^\infty \mu^x(F_x\backslash E^n_x)/\mu^x(F_x)\Bigr] \geq \theta\mu^x(F_x),$$

since $\mu^x(F_x\backslash E^n_x) \leq \mu^x(F^n_x\backslash E^n_x) \leq (k+1)^{-1}(1-\theta)2^{-n} \leq (1-\theta)2^{-n}\mu^x(F_x)$. It is clear that $\mathcal{F}$ contains all sets of the form $A\times B$, where $A \in \mathcal{A}$ and $B \subset Y$ is closed. The same is true for the complements of such sets, since any open set in Y is the union of a sequence of increasing closed sets. (ii) We may assume that Y is a complete separable metric space. By using (i) one can find sets $S_n \in \mathcal{A}\otimes\mathcal{B}$ with $S_{n+1} \subset S_n \subset S$ such that their sections are closed, nonempty and have diameters at most $1/n$ in the metric Y. Hence $\bigcap_{n=1}^{\infty} S_n$ is the graph of a mapping f, and this mapping is $\mathcal{A}$-measurable; see Blackwell, Ryll-Nardzewski [**186**], another proof is given in Kechris [**968**, Corollary 18.7].

10.10.48. (Blackwell, Ryll-Nardzewski [**186**]) (i) Let μ be a Borel probability measure on a Borel set X in a Polish space and let f be a Borel function on X. Let $\sigma(f)$ be the σ-algebra generated by f. Prove that the existence of regular conditional probabilities μ^y, $y \in \mathbb{R}^1$, that for all $y \in f(X)$ are concentrated on $f^{-1}(y)$ and for which all functions $y \mapsto \mu^y(A)$, $A \in \mathcal{B}(X)$, are Borel measurable, is equivalent to the existence of a mapping $F\colon X \to X$ that is $\bigl(\sigma(f), \mathcal{B}(X)\bigr)$-measurable and satisfies the condition $f\bigl(F(x)\bigr) = f(x)$.

(ii) Show that a necessary condition for the existence of a mapping F as in (i) is the Borel measurability of the set $f(X)$. In particular, there exists a continuous (even smooth) mapping f on a Borel set in $[0,1]$ for which there are no conditional measures with the properties mentioned in (i).

HINT: in the case where the indicated conditional measures exist we apply Exercise 10.10.47 to $X = Y$, $\mathcal{A} = \sigma(f)$ and the mapping $x \mapsto \mu^{f(x)}$, taking for S the set of all (x, y) with $f(x) = f(y)$. Then S contains the graph of some $\bigl(\sigma(f), \mathcal{B}(X)\bigr)$-measurable mapping F and $f\bigl(F(x)\bigr) = f(x)$. There is a Borel mapping $g\colon \mathbb{R}^1 \to X$ with $F(x) = g\bigl(f(x)\bigr)$. Hence the image of f is the Borel set $\bigl\{t\colon f\bigl(g(t)\bigr) = t\bigr\}$. Conversely, if F with the listed properties exists, then we take regular with respect to $\sigma(f)$ conditional measures $B \mapsto \mu(B, x)$, $x \in X$, and set $\mu^{f(x)}(B) := \mu\bigl(B, F(x)\bigr)$ for all $x \in F^{-1}(B)$, $\mu^{f(x)}(B) := 0$ for all $x \notin F^{-1}(B)$. If $y \notin f(X)$, then $\mu^y := \delta_0$.

10.10.49. Show that the existence of conditional measures in the sense of Doob with respect to $\mathfrak{B}$ (see Remark 10.6.3) is equivalent to the existence of a disintegration $\mu(\,\cdot\,, x)$ with $\mathfrak{F}_x = \mathfrak{F}$ for all $x \in X$.

HINT: if one has conditional measures in the sense of Doob, then for every set $A \in \mathfrak{F}$, there is a measure zero set $N_A \in \mathfrak{B}$ on the complement to which the function $\mu(A, x)$ is $\mathfrak{B}$-measurable. The converse is obvious.

10.10.50. Let $(M, \mathcal{M}, \mu)$ be a probability space. Prove that measurable partitions ζ and η are independent precisely when for every measurable ζ-set A and every measurable η-set B, one has the equality $\mu(A \cap B) = \mu(A)\mu(B)$.

10.10.51. (Dieudonné [**447**]) Let $X = [0,1]^\infty$ be equipped with the measure μ that is the countable product of Lebesgue measures on $[0,1]$. For every μ-integrable function f and every finite set $J \subset \mathbb{N}$, we let $J' := \mathbb{N}\backslash J$ and

$$f_J(x) := \int_{[0,1]^{J'}} f(x_J, x_{J'})\, \mu_{J'}(dx_{J'}),$$

where $x_J := (x_n)_{n\in J}$ and $\mu_{J'}$ is the projection of μ on $[0,1]^{J'}$, i.e., the sub-product of the copies of Lebesgue measure corresponding to J'. Given an increasing sequence J_n of finite parts of $\mathbb{N}$ with the union $\mathbb{N}$, we obtain by the martingale convergence theorem that $f_{J_n}(x) \to f(x)$ a.e. Show that this assertion may fail for nets, by constructing a measurable set E of positive μ-measure whose indicator $f = I_E$ has the following property: the net $\{f_J\}$ indexed by all finite sets $J \subset \mathbb{N}$ does not converge to f, i.e., *it is not true* that for μ-a.e. $x \in X$ and every $\varepsilon > 0$, there exists a finite set $J_0 \subset \mathbb{N}$ such that $|f(x) - f_J(x)| < \varepsilon$ for every finite set J that contains J_0.

10.10.52. Let μ be a measure with values in $[0+\infty]$ on a σ-algebra $\mathcal{A}$. Prove that the existence of a lifting on $\mathcal{L}^\infty_{\mathcal{A}}$ is equivalent to that μ is decomposable.

HINT: see, e.g., A. & C. Ionescu Tulcea [**867**, p. 48], Levin [**1164**, Ch. 3, §3].

10.10.53°**.** Show that there are no linear liftings on the spaces $L^p[0,1]$ in the case $1 \le p < \infty$.

HINT: if L is a linear lifting on $L^p[0,1]$, $1 \le p < \infty$, then for every t, the functional $l_t(f) = L(f)(t)$ on $L^p[0,1]$ is linear and nonnegative on nonnegative functions, which by Exercise 4.7.88 yields its continuity. Hence the functional l_t is represented by a function g_t in $L^q[0,1]$, $q = p/(p-1)$. For every n, we partition $[0,1]$ into n intervals $J_{n,1}, \ldots, J_{n,n}$ by the points k/n. Let $E_{n,k} := \bigl\{x\colon L(I_{J_{n,k}})(x) = 1\bigr\}$ and $E_n := \bigcup_{k=1}^n E_{n,k}$. Then $\lambda(E_n) = 1$ by the properties of liftings. There exists a point $t \in \bigcap_{n=1}^\infty E_n$. For every n, there is $j(n)$ with $t \in E_{n,j(n)}$, i.e., $L(I_{J_{n,j(n)}})(t) = 1$. Since $L(I_{J_{n,k}}) = I_{J_{n,k}}$ a.e., for all k we have

$$L(I_{J_{n,k}})(t) = \int_0^1 I_{J_{n,k}}(s) g_t(s)\, ds \le n^{-1/p} \|g_t\|_{L^q},$$

which leads to a contradiction. The same reasoning applies to any continuous measure, see A. & C. Ionescu Tulcea [**867**].

10.10.54°**.** Let $(X, \mathcal{A}, \mu)$ be a probability space and let $T\colon X \to X$ be a transformation that preserves the measure μ and is ergodic. Suppose that f is a μ-measurable nonnegative function such that μ-a.e. $I(x) := \lim\limits_{n\to\infty} n^{-1} \sum_{k=1}^n f(T^k x)$ exists and is finite. Prove that the function f is integrable.

HINT: let $f_N = \min(f, N)$, then for any fixed N the analogous limit exists and equals the integral of f_N for a.e. x. Hence the integral of f_N is majorized by $I(x)$ a.e. for every N, which yields the boundedness of the sequence of integrals of f_N, since it suffices to find a common point x for all N.

10.10.55. Let $n \in \mathbb{N}$ and let f_n be the transformation of the interval $[0,1]$ into itself taking x to the fractional part of nx.

(i) Prove that $\lambda \circ f_n^{-1} = \lambda$, where λ is Lebesgue measure.

(ii) Prove that for every set $E \subset [0,1]$ of positive measure, almost every point $x \in [0,1]$ has the property that $f_n(x) \in E$ for infinitely many n.

HINT: see Billingsley [**168**, Ch. 1, §3].

10.10.56. Let T be the transformation of the space $[0,1)$ into itself that takes $x > 0$ to the fractional part of $1/x$, $T(0) = 0$. Let us consider the following Gauss measure: $\mu := (\ln 2)^{-1}(x+1)^{-1}\, dx$. (i) Prove that $\mu \circ T^{-1} = \mu$. (ii) Prove that T is ergodic on $[0,1)$ with the measure μ and hence for every integrable function f on $[0,1)$ for a.e. x one has

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) = \frac{1}{\ln 2} \int_0^1 \frac{f(y)}{1+y}\, dy.$$

HINT: see Billingsley [**168**, Ch. 1, §4].

10.10.57. (Khinchin [**997**]) Let f be a positive continuous function on $(0,+\infty)$ such that $xf(x)$ is a decreasing function. Prove that if the integral of f over $[0,\infty)$

is infinite, then for almost all α the inequality $|\alpha - p/q| < f(q)/q$ has infinitely many solutions in integer numbers p and q ($q > 0$), and if this integral is finite, then for almost all α the indicated inequality has finitely many solutions. Apply this to $f(x) = x^{-1}(\log x)^{-1}$ and $f(x) = x^{-1}(\log x)^{-2}$.

HINT: see Billingsley [**168**, Ch. 1, §4], Khinchin [**997**, §14].

10.10.58.$^\circ$ (i) Let $\{\xi_n\}$ be a martingale with respect to $\{\mathcal{F}_n\}$ and let ψ be a convex function such that the functions $\psi(\xi_n)$ are integrable. Prove that $\{\psi(\xi_n)\}$ is a submartingale with respect to $\{\mathcal{F}_n\}$. In particular, $\{|\xi_n|^p\}$ is a submartingale if the functions $|\xi_n|^p$ are integrable. (ii) Prove that the conclusion in (i) remains true if $\{\xi_n\}$ is a submartingale and ψ is an increasing convex function. In particular, the functions $\max(\xi_n - c, 0)$ form a submartingale for all c.

10.10.59.$^\circ$ Let $\{\xi_n\}$, $n = 0, 1, \ldots$, be a submartingale and let a bounded nonnegative function g_n be measurable with respect to $\mathcal{F}_{n-1}$ for each $n \geq 1$. Prove that the sequence $[g, \xi]_n := \sum_{m=1}^n g_m(\xi_m - \xi_{m-1})$, $[g, \xi]_0 := 0$, is a submartingale.

10.10.60. Construct an example of a martingale $\{\xi_n\}$ that converges to zero in measure, but not a.e. and an example of a martingale $\{\xi_n\}$ that tends to $+\infty$ a.e.

10.10.61.$^\circ$ Let $\{\xi_n\}$ be a supermartingale with respect to $\{\mathcal{F}_n\}$ and let τ be a stopping time. Prove that $\{\xi_{\min(\tau,n)}\}$ is a supermartingale.

10.10.62. Let $\{\xi_n\}$ be a martingale with respect to $\{\mathcal{F}_n\}$ and let ν be the corresponding additive set function on the algebra $\mathcal{R} = \bigcup_{n=1}^\infty \mathcal{F}_n$ defined in Remark 10.3.7. Show that ν is countably additive if and only if $\mathbb{E}\xi_\tau = \mathbb{E}\xi_1$ for all finite stopping times τ. In this case $\xi_\infty = \lim\limits_{n\to\infty} \xi_n$ is the Radon–Nikodym density of the absolutely continuous component of ν with respect to P.

HINT: see Neveu [**1369**, Proposition III-1].

10.10.63. (Gilat [**686**]) Let $\{\xi_n\}$ be a nonnegative submartingale on a probability space $(\Omega, \mathcal{F}, P)$. Prove that there exists a martingale $\{\eta_n\}$ on some probability space $(\Omega', \mathcal{F}', P')$ such that the image of the measure P under the mapping $\xi = (\xi_n)\colon \Omega \to \mathbb{R}^\infty$ coincides with the image of the measure P' under the mapping $\eta = (|\eta_n|)\colon \Omega' \to \mathbb{R}^\infty$, i.e., the sequences $\{\xi_n\}$ and $\{|\eta_n|\}$ have the same distribution.

10.10.64. (i) Deduce Corollary 10.3.10 from Proposition 10.3.9.

(ii) Deduce from Corollary 10.3.10 the following inequality of Kolmogorov: if ξ_n are independent square integrable random variables with the zero mean, then

$$P\Bigl(\max_{1\leq k\leq n} |\xi_1 + \cdots + \xi_n| \geq r\Bigr) \leq r^{-2}\mathbb{E}|\xi_1 + \cdots + \xi_n|^2, \quad \forall\, r > 0.$$

10.10.65. (i) Show that the boundedness of the sequence $\{\|\xi_n^+\|_{L^1(P)}\}$ does not imply the boundedness of $\{\|X_n\|_{L^1(P)}\}$ in the situation of Corollary 10.3.11.

(ii) Prove that in the situation of Corollary 10.3.11 one has

$$\mathbb{E}X_n \leq \frac{e}{e-1}\Bigl(1 + \mathbb{E}\xi_n^+ \max(\log \xi_n^+, 0)\Bigr).$$

HINT: see Example 10.3.8 and Durrett [**505**, §4.4, Exercises 4.2, 4.7].

10.10.66. Prove the claim in Remark 10.10.9.

10.10.67. (Gaposhkin [**658**]) Let μ be a probability measure and let a sequence of functions f_n converge to zero in the weak topology of $L^p(\mu)$ for some $p \in [1, \infty)$. Prove that there exist a subsequence $\{f_{n_k}\}$ and a sequence of functions $g_k \in L^p(\mu)$ such that

$$\sum_{k=1}^{\infty} \|f_{n_k} - g_k\|_{L^p(\mu)} < \infty \quad \text{and} \quad \mathbb{E}(g_k|g_1, \ldots, g_{k-1}) = 0, \ \forall\, k \in \mathbb{N}.$$

10.10.68. (Oxtoby, Ulam [**1411**]) Show that the set of all points x in $(0, 1)$ for which the number of units among the first n coefficients in the expansion in negative powers of 2 divided by n tends to $1/2$ is a first category set (i.e., the law of large numbers fails for category in place of measure).

10.10.69. (Bryc, Kwapień [**268**]) Let $(\Omega, \mathcal{F}, P)$ be a probability space, let $\mathcal{F}_i$ be a sequence of mutually independent sub-σ-algebras in $\mathcal{F}$, and let $\xi_i \in L^1(\Omega, \mathcal{F}_i, P)$ be such that the integral of ξ_i is zero. Prove that the following conditions are equivalent: (a) there exists $\xi \in L^1(\Omega, \mathcal{F}, P)$ with $\xi_i = \mathbb{E}^{\mathcal{F}_i}\xi$ for all i, (b) $\lim\limits_{i\to\infty} \|\xi_i\|_{L^1(P)} = 0$.

10.10.70.° Let $(\Omega, \mathcal{A}, \mu)$ be a probability space and let $f(t, \omega)$ be a measurable function on $[0, 1] \times \Omega$, continuous in t. Denote by Ω_k the set of all ω for which there exists a chain $0 < s_1 < t_1 < \cdots < s_k < t_k \le 1$ such that $f(s_i, \omega) \ge 1$ and $f(t_i, \omega) \le 0$ for all $i = 1, \ldots, k$. Show that Ω_k is measurable.

Hint: for any fixed $\varepsilon > 0$, consider the set $\Omega_{k,\varepsilon}$ that is defined analogously to Ω_k with the inequalities $f(s_i, \omega) > 1 - \varepsilon$ and $f(t_i, \omega) < \varepsilon$. By the continuity of f in t, one can pass to rational s_i and t_i, which gives measurability of $\Omega_{k,\varepsilon}$. One has $\Omega_k = \bigcap_{j=1}^{\infty} \Omega_{k,1/j}$ by the continuity of f in t.

10.10.71. (Bellow [**145**]) Suppose that $(\Omega, \mathcal{F}, \mu)$ is a complete probability space and Λ is a lifting on $L^\infty(\mu)$. Let K be a compact space and let a mapping $g\colon\ \Omega \to K$ be $\bigl(\mathcal{F}, \mathcal{B}a(K)\bigr)$-measurable. For every $\omega \in \Omega$, consider the function $\psi \mapsto \Lambda(\psi \circ g)(\omega)$ on $C_b(K)$. (i) Show that there exists a unique element $\Lambda_K(g)(\omega) \in K$ such that the equality $\psi\bigl(\Lambda_K(g)(\omega)\bigr) = \Lambda(\psi \circ g)(\omega)$ holds for all $\psi \in C_b(K)$. (ii) Prove that the mapping $\Lambda_K(g)\colon\ \Omega \to K$ is Borel measurable. (iii) Prove that the image of the measure μ with respect to $\Lambda_K(g)$ is a Radon measure on K.

10.10.72. Construct two distinct centered Gaussian measures on $\mathbb{R}^\infty$ that for all n have equal conditional measures on all lines $y + \mathbb{R}^1 e_n$, $y \in \Pi_n$, where Π_n is the hyperplane $\{x \in \mathbb{R}^n\colon\ x_n = 0\}$, $e_n = (e_n^j)$, $e_n^n = 1$ and $e_n^j = 0$ if $j \ne n$.

Hint: see Bogachev [**208**, Theorem 7.3.7] or Bogachev, Röckner [**223**].

10.10.73. (Jessen [**897**], Doob [**466**]) Construct an example of a probability measure μ on a space Ω and two independent measurable functions ξ and η that are not independent in the sense of Kolmogorov (see remark after Definition 10.10.1).

10.10.74. (Stroock [**1796**], Kallianpur, Ramachandran [**941**]) Let X be a nonempty set with two σ-algebras $\mathcal{A}$ and $\mathcal{B}$. Let μ be a probability measure on $\mathcal{A}$ and ν a probability measure on $\mathcal{B}$. A probability measure η on the σ-algebra $\sigma(\mathcal{A}\cup\mathcal{B})$ is called a splicing of the measures μ and ν if $\eta(A \cap B) = \mu(A)\nu(B)$ for all $A \in \mathcal{A}$, $B \in \mathcal{B}$. Thus, $\eta = \mu$ on $\mathcal{A}$, $\eta = \nu$ on $\mathcal{B}$, and $\mathcal{A}$ and $\mathcal{B}$ are independent with respect to η. Prove that a splicing of measures μ and ν exists precisely when $\sum_{n=1}^{\infty} \mu(A_n)\nu(B_n) \ge 1$ for all sequences of sets $A_n \in \mathcal{A}$ and $B_n \in \mathcal{B}$ such that $X = \bigcup_{n=1}^{\infty} (A_n \cap B_n)$.

10.10.75. (i) (Lipchius [**1174**]) Let $(\Omega, \mathcal{F}, P)$ be a probability space, $(X, \mathcal{A})$ a measurable space with a countably generated and countably separated σ-algebra $\mathcal{A}$. Suppose that two mappings $f, g\colon \Omega \to X$ are measurable and independent and that g satisfies the following two conditions: 1) $g(E) \in \mathcal{A}_{P\circ g^{-1}}$ for all $E \in \mathcal{F}$, 2) for every sequence of pairwise disjoint sets $A_k \in \mathcal{F}$ such that $P(A_k) > 0$ and $\lim\limits_{k\to\infty} P(A_k) = 0$, there exists n such that $P \circ g^{-1}\bigl(g(A_n)\bigr) < 1$.

Prove that f coincides a.e. with a finitely many valued mapping.

(ii) (Ottaviani [**1407**]) Let g be an absolutely continuous function on $[0,1]$ that is not a constant. Suppose that a measurable function f on $[0,1]$ is such that f and g are independent random variables on $[0,1]$ with Lebesgue measure. Prove that f coincides a.e. with a function that assumes only finitely many values. Note that (i) implies (ii).

10.10.76. (Borell [**236**]) Let μ be a convex Radon probability measure on a locally convex space X and let G be an additive subgroup in X. Prove that either $\mu_*(G) = 0$ or $\mu_*(G) = 1$.

10.10.77.° Let μ be a probability measure. Prove that two μ-measurable functions f and g are independent precisely when for all t and s one has the equality

$$\int \exp(itf + isg)\, d\mu = \int \exp(itf)\, d\mu \int \exp(isg)\, d\mu.$$

Hint: if this equality holds, then for any function ψ that is a finite linear combination of the functions of the form $\exp(itx)$, the integral of $\psi(f)\psi(g)$ equals the product of the integrals of $\psi(f)$ and $\psi(g)$. It is clear by the Weierstrass theorem that this remains true for all $\psi \in C_0(\mathbb{R})$, hence for all bounded Borel functions.

10.10.78. (Rüschendorf, Thomsen [**1628**]) Suppose that $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ are measurable spaces. Let μ be a probability measure on $(X \times Y, \mathcal{A} \otimes \mathcal{B})$, let μ_X be the projection of μ on X, and let μ_Y be the projection of μ on Y. Set

$$S := \Bigl\{ f \in L^0(\mu)\colon\ f(x,y) = \varphi(x) + \psi(y), \varphi \in L^0(\mu_X), \psi \in L^0(\mu_Y) \Bigr\}.$$

(i) Let g be a positive finite μ-measurable function. Prove that the set

$$\{ f \in S\colon\ |f(x,y)| \le g(x,y) \text{ a.e.}\}$$

is closed in $L^0(\mu)$.

(ii) Give an example showing that S may not be closed.

10.10.79. (Jacobs [**875**]) Let Ω be a Polish space, μ a Borel probability measure on Ω, and $T\colon \Omega \to \Omega$ a continuous transformation. Suppose that there is an increasing sequence of integers $k_n \to \infty$ such that the measures $\mu\circ (T^{k_n})^{-1}$ converge weakly to μ. Prove the following extension of the Poincaré recurrence theorem: for μ-a.e. x, there is a sequence of integers $p_n \to \infty$ such that $T^{p_n} x \to x$.

Hint: let U be open; the set $G = \bigcup_{n\ge 0} T^{-n}(U)$ is open, $T^{-1}(G) \subset G$, $G\backslash T^{-1}(G) = U\backslash U_1$, where U_1 is the set of all points in U that return to U. It suffices to show that $\mu(G) = \mu\bigl(T^{-1}(G)\bigr)$. Let $\varepsilon > 0$ and let $f \in C_b(\Omega)$ be such that $0 \le f \le 1$,

$$\int f\, d\mu \ge \mu(G) - \varepsilon.$$

By weak convergence, there is n such that

$$\int f \circ T^n \, d\mu \geq \mu(G) - 2\varepsilon.$$

Hence $\mu(G) \geq \mu\bigl(T^{-1}(G)\bigr) \geq \mu\bigl(T^{-n}(G)\bigr) \geq \mu(G) - 2\varepsilon$, whence the claim follows.

10.10.80. Construct three random variables on a probability space that are pairwise independent, but are not independent.

10.10.81. (i) Let X be a Souslin space with a Borel measure μ, $\mathcal{A}$ a sub-σ-algebra in $\mathcal{B}(X)$, and let $\mu(\,\cdot\,,\,\cdot\,)$ be a regular conditional measure with respect to $\mathcal{A}$. Suppose that the measures $\mu(\,\cdot\,,x)$ are absolutely continuous with respect to a nonnegative measure ν on $\mathcal{A}$. Prove that there exists an $\mathcal{A}\otimes\mathcal{B}(X)$-measurable function ϱ on X^2 such that $d\mu(\,\cdot\,,x_1)/d\nu(x_2) = \varrho(x_1,x_2)$.

(ii) Let X and Y be Polish spaces, μ a Borel measure on $X\times Y$, μ_Y the projection of μ on Y, and let ν be a Borel probability measure on X such that μ_Y-a.e. conditional measures μ_y on X are absolutely continuous with respect to ν. Prove that there exists a Borel function ϱ on $X\times Y$ such that $d\mu_y/d\nu(x) = \varrho(x,y)$.

Hint: use Exercise 6.10.72.

10.10.82. Suppose that the distribution P_ξ of a random vector $\xi = (\xi_1,\ldots,\xi_n)$ in $\mathbb{R}^n$ is invariant with respect to permutations of coordinates and a Borel function φ on $\mathbb{R}^n$ is invariant with respect to permutations of coordinates. Let $\mathcal{B}$ denote the σ-algebra generated by the random variable $\varphi(\xi_1,\ldots,\xi_n)$. Show that if the variables ξ_i are integrable, then $\mathbb{E}^{\mathcal{B}}\xi_1 = \mathbb{E}^{\mathcal{B}}\xi_i$ for all $i \leq n$. In particular, if $\varphi(x) = x_1+\cdots+x_n$, then the equality $\mathbb{E}^{\mathcal{B}}\xi_i = (\xi_1+\cdots+\xi_n)/n$ holds.

Hint: for every bounded Borel function ψ on the real line, the integral of the function $(x_1 - x_k)\psi\circ\varphi(x_1,\ldots,x_n)$ with respect to P_ξ vanishes because the transformation that interchanges the first and the kth coordinates leaves this integral unchanged, but at the same time transforms it into the opposite number.

10.10.83. (i) (Burkholder [**289**]) Let ξ be an integrable random variable and $\xi_1,\xi_2,\ldots$ independent random variables each with the same distribution as ξ. Show that the following statements are equivalent:

(a) $|\xi|\log^+|\xi|$ is not integrable, where $\log^+ x = \log x$ if $x > 1$ and $\log^+ x = 0$ otherwise,

(b) $\sup_n |\xi_n|/n$ is not integrable,

(c) $\sup_n[|\xi_1|+\cdots+|\xi_n|]/n$ is not integrable.

(ii) (Blackwell, Dubins [**182**]) Show that if ξ is a nonnegative integrable random variable such that $\xi\log^+\xi$ is not integrable, then there exist a probability space $(\Omega,\mathcal{F},P)$, a decreasing sequence of sub-σ-fields $\mathcal{F}_n\subset\mathcal{F}$, and a random variable ξ_1 on $(\Omega,\mathcal{F},P)$ with the same distribution as ξ such that $\sup_n \mathbb{E}^{\mathcal{F}_n}\xi$ is not integrable.

Hint: (ii) let $\xi_1,\xi_2,\ldots$ be independent and have the same distribution as ξ and let $\mathcal{F}_n$ be generated by $\xi_1+\cdots+\xi_k$, $k\geq n$; observe that $\mathbb{E}^{\mathcal{F}_n}\xi_1 = \mathbb{E}^{\mathcal{F}_n}\xi_k$ for each $k\leq n$, hence $\mathbb{E}^{\mathcal{F}_n}\xi_1 = (\xi_1+\cdots+\xi_n)/n$.

10.10.84.° Let $(\Omega,\mathcal{A},P)$ be a probability space. Prove that the following conditions on sub-σ-algebras $\mathcal{F},\mathcal{G}\subset\mathcal{A}$ are equivalent: (i) $\mathbb{E}^{\mathcal{F}}\xi = \mathbb{E}^{\mathcal{G}}\xi$ a.e. for every integrable function ξ, (ii) for every $F\in\mathcal{F}$, there exists a set $G\in\mathcal{G}$ with $P(F\triangle G)=0$, and for every $G\in\mathcal{G}$, there exists a set $F\in\mathcal{F}$ with $P(F\triangle G)=0$.

HINT: if we have (i) and $F \in \mathcal{F}$, then $I_F = \mathbb{E}^{\mathcal{F}} I_F = \mathbb{E}^{\mathcal{G}} I_F$ a.e. and one can take $G = \{\mathbb{E}^{\mathcal{G}} I_F = 1\}$. If we have (ii), then every $\mathcal{F}$-measurable function equals a.e. some $\mathcal{G}$-measurable function and conversely.

10.10.85.° Let $(\Omega, \mathcal{A}, P)$ be a probability space, let $\mathcal{F}, \mathcal{G} \subset \mathcal{A}$ be sub-σ-algebras, and let $\xi, \eta \in \mathcal{L}^1(P)$. Suppose that a set $A \in \mathcal{F} \cap \mathcal{G}$ is such that $\xi = \eta$ a.e. on A and $\{A \cap G\colon G \in \mathcal{G}\} = \{A \cap F\colon F \in \mathcal{F}\}$. Show that $\mathbb{E}^{\mathcal{F}}\xi = \mathbb{E}^{\mathcal{G}}\eta$ a.e. on A.

HINT: let $E := A \cap \{\mathbb{E}^{\mathcal{F}}\xi > \mathbb{E}^{\mathcal{G}}\eta\}$. Observe that $E \in \mathcal{F} \cap \mathcal{G}$ and show that $P(E) = 0$ by verifying that the integral of $\mathbb{E}^{\mathcal{F}}\xi - \mathbb{E}^{\mathcal{G}}\eta$ over E vanishes.

10.10.86.° Let $\xi \geq 0$ be an integrable random variable on a probability space $(\Omega, \mathcal{A}, P)$ and let $\mathcal{B} \subset \mathcal{A}$ be a sub-σ-algebra. Show that if $\xi > 0$ on a set of positive measure, then $\mathbb{E}^{\mathcal{B}}\xi > 0$ on a set of positive measure.

HINT: if $\mathbb{E}^{\mathcal{B}}\xi = 0$ a.e., then $\mathbb{E}\xi = 0$.

10.10.87.° Suppose we are given a probability space $(\Omega, \mathcal{A}, P)$, a sequence of integrable functions $\xi_n \geq 0$, and a sequence of sub-σ-algebras $\mathcal{A}_n \subset \mathcal{A}$. Let $\mathbb{E}^{\mathcal{A}_n}\xi_n \to 0$ in probability. Prove that $\xi_n \to 0$ in probability.

HINT: observe that $\mathbb{E}^{\mathcal{A}_n}\xi_n(\xi_n + 1)^{-1} \to 0$ in probability, which yields convergence $\xi_n(\xi_n + 1)^{-1} \to 0$ in L^1-norm.

10.10.88.° Let f be an integrable function on a probability space $(X, \mathcal{A}, \mu)$ and let $\mathcal{B} \subset \mathcal{A}$ be a sub-σ-algebra. Let V be a strictly convex function on the real line, i.e., $V(x) - V(y) > V'_+(y)(x - y)$ whenever $x \neq y$, and let the function $V \circ f$ be integrable. Suppose that $\mathbb{E}^{\mathcal{B}}(V \circ f) = V \circ \mathbb{E}^{\mathcal{B}} f$ a.e. Prove that $f = \mathbb{E}^{\mathcal{B}} f$ a.e.

HINT: letting $g := \mathbb{E}^{\mathcal{B}} f$ we have $h := V(f) - V(g) - V'_+(g)(f - g) \geq 0$ a.e. If $\mu(\{f \neq g\}) > 0$, then $\mu(\{h > 0\}) > 0$, whence $\mu(\{\mathbb{E}^{\mathcal{B}} h > 0\}) > 0$. It remains to observe that $\mathbb{E}^{\mathcal{B}}[V'_+(g)(f - g)] = \mathbb{E}^{\mathcal{B}} V'_+(g)\mathbb{E}^{\mathcal{B}}(f - g) = 0$ a.e.

10.10.89. Let f be an integrable function on a probability space $(X, \mathcal{A}, \mu)$ and let $\mathcal{B} \subset \mathcal{A}$ be a sub-σ-algebra. Show that if $\mathbb{E}^{\mathcal{B}} f$ and f have equal distributions, then we have $f = \mathbb{E}^{\mathcal{B}} f$ a.e.

HINT: there exists a strictly increasing convex function V such that the function $V(f)$ is integrable. One has $\mathbb{E}^{\mathcal{B}} V(f) \geq V(\mathbb{E}^{\mathcal{B}} f)$ by Jensen's inequality, and the integrals of both sides are equal, since $\mathbb{E}^{\mathcal{B}} f$ and f are equally distributed. This is possible only if $\mathbb{E}^{\mathcal{B}} V(f) = V(\mathbb{E}^{\mathcal{B}} f)$ a.e., which gives $f = \mathbb{E}^{\mathcal{B}} f$ a.e. by Exercise 10.10.88.

10.10.90. Suppose that on a probability space we are given integrable random variables ξ, ξ' and random variables η, η' such that (ξ, η) and (ξ', η') have the same distribution. Prove that $\mathbb{E}(\xi|\eta)$ and $\mathbb{E}(\xi'|\eta')$ have a common distribution.

HINT: $\mathbb{E}(\xi|\eta) = f(\eta)$ for some Borel function f, whence for every bounded Borel function φ we obtain

$$\mathbb{E}[\varphi\circ\eta'\mathbb{E}(\xi'|\eta')] = \mathbb{E}[\xi'(\varphi\circ\eta')] = \mathbb{E}[\xi(\varphi\circ\eta)] = \mathbb{E}[(f\circ\eta)(\varphi\circ\eta)] = \mathbb{E}[(f\circ\eta')(\varphi\circ\eta')],$$

which gives $\mathbb{E}(\xi'|\eta') = f \circ \eta'$ a.e.

10.10.91. Let random elements ξ and η on a probability space $(\Omega, \mathcal{A}, P)$ take values in a Souslin space S. Suppose that a Borel mapping $F\colon S \to S$ is such that the random elements $(\xi, F \circ \eta)$ and (ξ, η) have one and the same distribution. Prove that:

(i) $P^{\sigma(\eta)}(A) = P^{\sigma(F\circ\eta)}(A)$ for all $A \in \sigma(\xi)$,

(ii) the random elements ξ and η are conditionally independent with respect to $F\circ\eta$.

HINT: (i) the function I_A has the form $\psi\circ\xi$, the function $\mathbb{E}^{\sigma(\eta)}I_A$ has the form $\theta\circ\eta$ and is a unique (up to equivalence) function of η on which the minimum of the distances from $\psi\circ\xi$ to elements of the subspace of $\sigma(\eta)$-measurable functions is attained. Since the function $\theta(F\circ\eta)$ is $\sigma(\eta)$-measurable and the function $\psi\circ\xi-\theta(F\circ\eta)$ has the same L^2-norm as $\psi\circ\xi-\theta\circ\eta$, we obtain $\theta\circ\eta=\theta(F\circ\eta)$ a.e., hence $\theta\circ\eta$ has a $\sigma(F\circ\eta)$-measurable modification. (ii) By (i) for all $A\in\sigma(\xi)$ and $B\in\sigma(\eta)$ we have

$$\begin{aligned}\mathbb{E}^{\sigma(F\circ\eta)}(I_AI_B)&=\mathbb{E}^{\sigma(F\circ\eta)}\mathbb{E}^{\sigma(\eta)}(I_AI_B)=\mathbb{E}^{\sigma(F\circ\eta)}(I_B\mathbb{E}^{\sigma(\eta)}I_A)\\&=\mathbb{E}^{\sigma(F\circ\eta)}(I_B\mathbb{E}^{\sigma(F\circ\eta)}I_A)=\mathbb{E}^{\sigma(F\circ\eta)}I_B\mathbb{E}^{\sigma(F\circ\eta)}I_A.\end{aligned}$$

10.10.92.° Let random variables ξ,η,ζ be such that the vector (ξ,ζ) and η are independent. Show that ξ and η are conditionally independent given ζ.

HINT: let bounded functions f, g, and h be measurable with respect to $\sigma(\xi)$, $\sigma(\eta)$, and $\sigma(\zeta)$, respectively. Then $\mathbb{E}(fgh)=\mathbb{E}g\mathbb{E}(fh)$ and

$$\mathbb{E}[h\mathbb{E}(f|\zeta)\mathbb{E}(g|\zeta)]=\mathbb{E}[hg\mathbb{E}(f|\zeta)]=\mathbb{E}g\mathbb{E}[h\mathbb{E}(f|\zeta)]=\mathbb{E}g\mathbb{E}(fh),$$

which gives the equality $\mathbb{E}(fg|\zeta)=\mathbb{E}(f|\zeta)\mathbb{E}(g|\zeta)$.

10.10.93. 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 $(Z,\mathcal{E})$. Prove that $\nu_{\sigma,T}:=(\nu\otimes\sigma)\circ T^{-1}\ll\mu_{\sigma,T}:=(\mu\otimes\sigma)\circ T^{-1}$ and that

$$\int_Z V\Bigl(\frac{d\nu_{\sigma,T}}{d\mu_{\sigma,T}}\Bigr)\,d\mu_{\sigma,T}\le\int_X V\Bigl(\frac{d\nu}{d\mu}\Bigr)\,d\mu$$

for any convex function V such that $V(d\nu/d\mu)\in L^1(\mu)$.

HINT: 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 let $\mathcal{F}$ be the σ-algebra generated by T. It is readily verified that $g\circ T=\mathbb{E}^{\mathcal{F}}_{\mu\otimes\sigma}f$. It remains to apply Jensen's inequality for conditional expectations.

10.10.94. Let X and Y be Polish spaces and let a Borel probability measure μ on $X\times Y$ be such that its projection μ_X on X has no atoms. Prove that there exists a sequence of Borel mappings $\varphi_n\colon\ X\to Y$ such that the measures $\mu_n:=\mu_X\circ F_n^{-1}$, where $F_n(x)=\bigl(x,\varphi_n(x)\bigr)$, converge weakly to μ.

HINT: let μ_x, $x\in X$, be conditional probabilities on Y for the measure μ. Since the weak topology on $\mathcal{P}(X\times Y)$ is metrizable and the mapping $x\mapsto\mu_x$ from X to $\mathcal{P}(Y)$ is measurable, it suffices to prove the assertion in the case where the mapping $x\mapsto\mu_x$ is simple, i.e., the space X is partitioned into finitely many Borel parts B_i such that $\mu_x=\mu_i$ for every $x\in B_i$, $\mu_i\in\mathcal{P}(Y)$. Clearly, this case reduces to the case where $\mu=\mu_X\otimes\nu$ with some $\nu\in\mathcal{P}(Y)$. We can approximate $\mu_X\otimes\nu$ by a sequence of measures of the form $\mu_X\otimes\nu_n$, where ν_n has a finite support. Hence we may assume that $\nu=\sum_{i=1}^p c_i\delta_{y_i}$, $y_i\in Y$, $0<c_i\le 1$, $\sum_{i=1}^p c_i=1$. Now we proceed as in Example 8.3.3: given n, we partition X in Borel sets B_j of positive μ_X-measure and diameter less than $1/n$; each B_j is partitioned into p Borel parts $B_{j,i}$ with $\mu_X(B_{j,i})=c_i\mu_X(B_j)$. Finally, let φ_n be defined as follows: $\varphi_n(x)=y_i$ if $x\in B_{j,i}$. Let $f\in\mathrm{Lip}_1(X\times Y)$. The difference between the integrals of f against μ

and $\mu_X \circ F_n^{-1}$ does not exceed $2/n$. Indeed, pick $x_j \in B_j$. Then

$$\biggl|\int_{X\times Y} f\,d\mu - \sum_{i=1}^{p}\sum_{j=1}^{\infty} c_i f(x_j, y_i)\mu_X(B_j)\biggr| \le 1/n$$

because $|f(x,y) - f(x_j,y)| \le 1/n$ whenever $x \in B_j$. Similarly,

$$\biggl|\int_{X} f\circ F_n\,d\mu_X - \sum_{j=1}^{\infty}\sum_{i=1}^{p} f(x_j, y_i)\mu_X(B_{j,i})\biggr| \le 1/n.$$

It remains to recall that $\mu_X(B_{j,i}) = c_i\mu_X(B_j)$.

10.10.95. Suppose a sequence of Borel probability measures μ_n on $[0,1]^2$ converges weakly to Lebesgue measure. Is it possible that, for all n and $x \in [0,1]$, the conditional measures μ_n^x on the vertical line are Dirac measures at some points?

HINT: yes, it is: see the previous exercise.

10.10.96. Bogachev, Korolev [**219**]) Show that Theorem 10.9.7 may fail for unbounded functions f. More specifically, show that in the case of the group of rotations of the unit circle with Lebesgue measure there exist an unbounded Borel function f on the unit circle and a probability density ϱ on $[0,1]$ for which Theorem 10.9.7 fails.

Bibliographical and Historical Comments

> Upon superficial observation mathematics appears to be a fruit of many thousands of scarcely related individuals scattered through the continents, centuries and millenniums. But the internal logic of its development looks much more like the work of a single intellect that is developing his thought continuously and systematically, using as a tool only the variety of human personalities. As in an orchestra performing a symphony by some composer, a theme is passing from one instrument to another, and when a performer has to finish his part, another one is continuing it as if playing from music.
>
> *I.R. Shafarevich.* On some tendencies of the development of mathematics.

> Unfortunately, it is in the very nature of such a systematic exposition that newly obtained knowledge merges with the old one, so that the historical development becomes unrecognizable.
>
> *C. Carathéodory.* Vorlesungen über reelle Funktionen.

Chapter 6.

§§6.1–6.8. In this chapter, along with some topological concepts we present the basic facts of the so-called descriptive set theory which are necessary for applications in measure theory. This theory arose simultaneously with measure theory, to a large extent under the influence of the latter (let us mention Lebesgue's work [**1123**]). Considerable contributions to its creation are due to E. Borel, R. Baire, H. Lebesgue, N.N. Lusin, F. Hausdorff, M.Ya. Souslin, W. Sierpiński, P.S. Alexandroff, P.S. Novikoff, A.A. Lyapunov, and other researchers; see comments to §1.10 in Volume 1 concerning the history of discovery of Souslin sets and Arsenin, Lyapunov [**72**], Hausdorff [**797**] Kanovei [**947**], Kuratowski [**1082**], Lyapunov [**1217**], Novikov [**1385**], and comments in [**216**], [**1209**], [**1211**]. The Souslin sets (A-sets or analytic sets in the terminology of that time; the term "Souslin sets" was introduced by Hausdorff in his book [**797**]) were first considered by Souslin, Lusin, Sierpiński, and other researchers in the space $\mathbb{R}^n$ and its subspaces, but already then the special role of the space of irrational numbers (or the space of all sequences) was realized. So the step to a study of Souslin sets in topological spaces was natural; see, e.g., Shneider [**1701**]. Among later works note Bressler, Sion [**253**], Choban [**341**], Choquet [**350**], Frolík [**642**], Hoffmann-Jørgensen [**841**], Jayne [**886**], [**887**], Rao, Rao [**1532**], Sion [**1731**], [**1732**], Topsøe [**1881**], and Topsøe, Hoffmann-Jørgensen [**1882**], where one can find additional references. A more detailed exposition of this direction can be found in Dellacherie [**425**], Kechris [**968**], Rogers, Jayne [**1589**], Srivastava [**1772**]. Dellacherie [**424**] discusses descriptive set theory in relation to the theory of capacities and certain measurability problems in the theory of random processes. In the 1920–1930s a whole direction arose and was intensively developing at the intersection of measure theory, descriptive set theory, general topology and partly mathematical logic; this direction can be called set-theoretic measure theory. Considerable contributions to this direction are due to Banach [**108**], Sierpiński [**1721**], [**1723**], Szpilrajn-Marczewski [**1819**], [**1256**], Ulam [**1898**].

Proposition 6.5.4 was obtained in Hoffmann-Jørgensen [**841**] for Souslin spaces; for separable Banach spaces it was also noted in Afanas'eva, Petunin [**12**] and Perlman [**1432**].

In order to describe the σ-algebra generated by a sequence of sets E_n and construct isomorphisms of measurable spaces Szpilrajn [**1815**], [**1816**] employed "the characteristic function of a sequence of sets", i.e., the function f defined by $f(x) = 2\sum_{n=1}^{\infty} 3^{-n} I_{E_n}(x)$; it was noted in [**1815**] that a compact form of representation of such a function had been suggested by Kuratowski.

The absence of a countable collection of generators of the σ-algebra $\mathcal{S}$ generated by Souslin sets was established in Rao [**1529**] (whence we borrowed the reasoning in Example 6.5.9) and Mansfield [**1247**]; see also Rao [**1530**]. Rao [**1528**] proved that under the continuum hypothesis there exists a countably generated σ-algebra of subsets of the interval $[0,1]$ containing all Souslin sets (the question about this as well as the problem of the existence of countably many generators of $\mathcal{S}$ was raised by S. Ulam, see *Fund. Math.*, 1938, V. 30, p. 365). In the same work [**1528**], the following more general fact was established: if X is a set of cardinality κ equal to the first uncountable cardinal, then for every collection of sets $X_\alpha \subset X$ that has cardinality κ, there exists a countably generated σ-algebra containing all singletons in X and all sets X_α.

A simple description of the Borel isomorphic types of Borel sets leads to the analogous problem for Souslin sets. However, here the situation is more complicated, and one cannot give an answer without additional set-theoretic axioms. It is consistent with the standard axioms that every two non-Borel Souslin sets on the real line are Borel isomorphic. On the other hand, one can add an axiom which ensures the existence of a non-Borel Souslin set A that is not Borel isomorphic to A^2 and $A \times [0,1]$. For example, if there exists a non-Borel coanalytic set $C \subset [0,1]$ without perfect subsets, then one can take $A = [0,1]\backslash C$. See details in Cenzer, Mauldin [**321**], Maitra, Ryll-Nardzewski [**1239**], Mauldin [**1276**].

§6.9. Measurable selection theorems go back to Lusin (see [**1209**], [**1208**]) and Novikoff (see [**1383**], [**1385**]) in respect of fundamental ideas and general approach, but the first explicit result of the type of Theorem 6.9.1 was obtained by Jankoff [**882**]. Some authors call this theorem the Lusin–Jankoff (Yankov) theorem, see Arsenin, Lyapunov [**72**]; it was shown in Lusin [**1208**] that every Borel set B in the plane is uniformizable by a coanalytic set C (a set M_1 is said to be uniformizable by a set $M_2 \subset M_1$ if M_2 is the graph of a function defined on the projection of M_1 to the axis of abscissas), and Jankoff observed that one can take for C the graph of a measurable function, which yields a measurable selection. This approach is described in detail in [**72**]. The measurable selection theorem was later proved independently by von Neumann [**1363**]. For this reason, the discussed theorem is also called the Jankoff–von Neumann theorem. It appears that this terminology is justified and that, on the other hand, the name "the measurable selection theorem" has an advantage in being informative and a disadvantage in being applicable to too many results in this area. There are comments in Wagner [**1956**] with some information that von Neumann could have proved the result even before World War II, but since no analogous investigation with respect to the other authors was done, we refer only to the published works.

Theorem 6.9.3 was discovered by Rohlin [**1596**] and later by Kuratowski and Ryll-Nardzewski [**1084**]. Wagner [**1956**] detects a gap in the proof in [**1596**], but also indicates a simple and sufficiently obvious way to correct it, keeping the main idea; independently of the way of correcting that gap, it is obvious that the very fact of announcing such an important theorem had a principal significance. Regarding measurable selections, see also Castaing, Valadier [**319**], Graf [**721**], Graf, Mauldin [**723**], Levin [**1164**], Saint-Raymond [**1639**], Wagner [**1956**], [**1957**]; related questions (such as measurable modifications) are discussed in Cohn [**361**], Mauldin [**1277**].

§6.10. The idea of applying compact classes to the characterization of abstract Souslin sets as projections goes back to the work Marczewski, Ryll-Nardzewski [**1258**]. It should be noted that many results of this chapter on Souslin spaces are valid in a more abstract setting, where no topologies are employed and the main role is played by compact classes, see Hoffmann-Jørgensen [**841**].

Interesting results related to the Borel structure can be found in Christensen [**355**]. Various problems connected with measurability in functional spaces (in particular, with Borel or Souslin sets) arise in the theory of random processes and mathematical statistics, see Dellacherie [**424**], Dynkin [**507**], Chentsov [**335**], [**336**], [**337**], [**338**], Ma, Röckner [**1219**], Dellacherie, Meyer [**427**], Rao [**1539**], Thorisson [**1854**].

The assertion of Exercise 6.10.53 is found in Kuratowski, Szpilrajn [**1085**] with attribution to M-lle Braun.

Chapter 7.

§§7.1–7.4. Measure theory on topological spaces began to develop in the 1930s under the influence of descriptive set theory and general topology as well as in connection with problems of functional analysis, dynamical systems, and other fields. In particular, this development was considerably influenced by the discovery of Haar measures on locally compact topological groups. This influence was so strong that until recently the chapters on measures on topological spaces in measure theory textbooks (in those advanced treatises where such chapters were included) dealt almost exclusively with locally compact spaces. Among the works of the 1930–1950s that played a particularly significant role in the development of measure theory on topological spaces we note the following: Alexandroff[1] [**30**], Bogoliouboff, Kryloff [**227**], Choquet [**349**], Gnedenko, Kolmogorov [**700**], Haar [**758**], Hopf [**854**], Marczewski [**1254**], Oxtoby, Ulam [**1412**], Prohorov [**1496**], [**1497**], Rohlin [**1595**], Stone [**1788**], [**1789**], [**1790**], Weil [**1965**], as well as Halmos's book [**779**] and the first edition of Bourbaki [**242**]. It should be added that Radon [**1514**] had already worked out the key ideas of topological measure theory in the case of the space $\mathbb{R}^n$. Certainly, an important role was played by research on the border of measure theory and descriptive set theory (Lusin, Sierpiński, Szpilrajn-Marczewski, and others). Finally, topological measure theory was obviously influenced by the investigations of Wiener, Kolmogorov, Doob, and Jessen on integration in infinite-dimensional spaces and the distributions of random processes; this influence became especially significant in the subsequent decades.

The first thorough and very general investigation of measures on topological spaces was accomplished in a series of papers (of book size) by A.D. Alexandroff [**30**], after which it became possible to speak of a new branch of measure theory. In this fundamental work, under very general assumptions on the considered spaces (even more general than topological, although in many statements one was concerned with normal topological spaces), regular additive set functions of bounded variation (called charges) were investigated. A.D. Alexandroff introduced and studied the concept of a τ-additive signed measure (he called such measures "real"), considered tight measures (measures concentrated on countable unions of compact sets; the term "tight" was later coined by Le Cam), established the correspondence between charges and functionals on the space of bounded continuous functions, in particular, the correspondence between τ-additive measures and τ-smooth functionals, and obtained the decomposition of a τ-additive measure into the difference of two nonnegative τ-additive measures, and many other results, which along with later generalizations form the basis of our exposition. In addition, in the same work, the investigation of weak convergence of measures on topological spaces was initiated, which is the subject of Chapter 8. Varadarajan

[1]An alternative spelling used in the translations of some later works is Aleksandrov.

[**1918**] wrote a survey of the main directions in topological measure theory, based principally on the works by A.D. Alexandroff and Yu.V. Prohorov, with a number of important generalizations and simplifications. The books by Bourbaki [**242**], Parthasarathy [**1424**], Topsøe [**1873**], Schwartz [**1681**], and Vakhania, Tarieladze, Chobanyan [**1910**] have become standard references in measure theory on metric or topological spaces. A very informative survey of measures on topological spaces is included in Tortrat [**1887**]. Schwartz's book [**1681**] has played an important role in the development and popularization of the theory of Radon measures on general topological spaces. Recently, an extensive treatise by Fremlin [**635**] has been published, a large portion of which is devoted to measures on topological spaces and related set-theoretic problems. Detailed surveys covering many special directions were published by Gardner [**660**], Gardner, Pfeffer [**666**], Wheeler [**1979**], and the author [**207**]. These surveys contain many additional results and references. Note also that Gardner [**660**], Gardner, Pfeffer [**666**], and Fremlin [**635**] contain a lot of information on infinite Borel measures, which is outside the scope of this book (except for a few occasional remarks).

S. Ulam (see [**1899**], [**1411**]) was one of the first to notice the property of tightness of Borel measures on complete separable metric spaces. As already mentioned in the comments to Volume 1, for $\mathbb{R}^n$ this property had already been found by Radon. A bit later this property was independently established by A.D. Alexandroff. It seems that at the end of the 1930s several other mathematicians observed this simple, but very important property, namely Kolmogorov, von Neumann, and Rohlin; however, in published form it appeared only in their later works. After A.D. Alexandroff, the property of τ-additivity was considered by many authors, see Amemiya, Okada, Okazaki [**46**], Gardner [**660**], Gardner, Pfeffer [**666**], and Tortrat [**1889**], [**1890**], where one can find additional references.

The concept of a universally measurable set was first considered, apparently, by Marczewski (see Marczewski [**1256**, p. 168]).

Some authors call the set S_μ defined in §7.2 the support of μ if $|\mu|(S_\mu) > 0$ (but S_μ does not necessarily have full measure); then measures concentrated on S_μ are called support concentrated.

Among many papers devoted to extensions of measures on topological spaces we especially note the classical works by A.D. Alexandroff [**30**] and Marczewski [**1254**] that revealed the role of compact approximations, and the subsequent works in this circle of ideas by Choksi [**344**], Erohin [**537**], Henry [**812**], Kisyński [**1007**], Mallory [**1245**], Topsøe [**1878**], [**1879**], [**1880**]. Very important for applications, Theorem 7.3.2 goes back to Prohorov [**1498**]. The formulation in the text along with the proof is borrowed from Vakhania, Tarieladze, Chobanyan [**1910**]. We note that the regularity of the space in (ii) is essential (see a counter-example in Fremlin [**635**, §419H]). There are many papers on extensions of measures with values in more general spaces (see, e.g., Lipecki [**1177**]), but here we are only concerned with real measures.

In the classical book by Halmos [**779**], the Baire sets are defined as sets in the σ-algebra generated by compact G_δ-sets, whereas the Borel sets are elements of the σ-ring generated by compact sets in a locally compact space; this differs from the modern terminology.

Measures on Souslin spaces (first for subspaces of the real line, then in the abstract setting) became a very popular object of study starting from old works by Lusin and Sierpiński (see comments to §1.10). Such spaces turned out to be very convenient in applications, since they include most of the spaces actually encountered and enable one to construct various necessary objects of measure theory (conditional measures, measurable selections, etc.). In this connection we note the paper Mackey [**1223**]. The fact that any Borel measure on a Souslin space is Radon can be deduced from the properties of capacities (which was pointed out by G. Choquet).

It is known that it is consistent to assume that there exists a Souslin set on the plane such that the projection of its complement is not Lebesgue measurable. This result was noted by K. Gödel and proved by P.S. Novikov [**1384**].

§7.5. Perfect measures were introduced in the classical book by Gnedenko and Kolmogorov [**700**]; for injective functions the main determining property was considered by Halmos and von Neumann [**781**] among other properties characterizing their "normal measures". Perfect measures were thoroughly investigated by Ryll-Nardzewski [**1631**] who characterized them in terms of quasi-compactness and by Sazonov [**1656**]. Compact measures introduced by Marczewski [**1254**] turned out to be closely connected with perfect measures. Vinokurov [**1929**] noted the existence of a perfect but not compact measure. The first example of such a measure was given in Vinokurov, Mahkamov [**1930**]; another example was constructed in Musiał [**1346**]. The relative intricacy of these examples also shows that both properties are very close. Dekiert [**422**] established the existence of a perfect probability measure without a monocompact, in the sense of Theorem 1.12.5, approximating class (actually, it was proved that so is the measure from Musiał [**1346**]). Fremlin [**634**] constructed a probability measure that possesses a monocompact approximating class but has no compact approximating classes. Our exposition of the fundamentals of the theory of perfect measures follows mainly the paper [**1656**] and the book Hennequin, Tortrat [**811**], although it contains a lot of additional results. Perfect measures and related objects are also discussed in Adamski [**8**], Darst [**406**], van Dulst [**498**], Frolík, Pachl [**643**], Koumoullis [**1043**], [**1045**], Koumoullis, Prikry [**1050**], Musiał [**1345**], [**1347**], Ramachandran [**1521**], Remy [**1548**].

§7.6–7.7. Products of measures on topological spaces, in particular, products of Radon measures are investigated in Bledsoe, Morse [**188**], Bledsoe, Wilks [**189**], Elliott [**527**], Godfrey, Sion [**703**], Grekas [**734**], Grekas, Gryllakis [**737**], [**738**], Gryllakis, Grekas [**749**], Johnson [**907**], [**908**], [**909**], [**910**], [**911**], [**912**], Johnson, Wajch, Wilczyński [**913**], Plebanek [**1466**]. It is proved in de Leeuw [**423**] that the function $\int h(x,y)\,\mu(dy)$ is Borel measurable provided that μ is a Radon measure on a compact space K and h is a bounded

Borel function on K^2. Concerning measurability of functions on product spaces, see also Grande [**726**], [**727**].

For probability distributions on the countable product of real lines, Daniell [**402**] obtained a result close to the Kolmogorov theorem (which appeared later), but presented it in a less convenient form in terms of the distribution functions of infinitely many variables (functions of bounded variation and positive type according to Daniell's terminology), i.e., Daniell characterized functions of the form $F(x_1, x_2, \ldots) = \mu\bigl(\prod_{n=1}^{\infty}(-\infty, x_n)\bigr)$, where μ is a probability measure on $\mathbb{R}^{\infty}$. In order to derive the Kolmogorov theorem from this result, given consistent finite-dimensional distributions, one has to construct the corresponding function on $\mathbb{R}^{\infty}$. By using compact classes, Marczewski [**1254**] obtained an important generalization of Kolmogorov's theorem on consistent probability distributions. Later this direction was developing in the framework of projective systems of measures (see §9.12(i)). Its relations to transition probabilities and conditional probabilities are discussed in Dinculeanu [**451**], Lamb [**1101**].

§7.8. Daniell's construction [**399**], [**400**], [**403**] turned out to be very efficient in the theory of integration on locally compact spaces. It enabled one to construct the integral without prior constructing measures, which is convenient when the corresponding measures are not σ-finite. This was manifested especially by the theory of Haar measures. In that case, it turned out to be preferable to regard measures as functionals on spaces of continuous functions. Daniell's construction was substantially developed by Stone [**1790**]; let us also mention the work of Goldstine [**710**] that preceded Stone's series of papers and was concerned with the representation of functionals as integrals in Daniell's spirit. Certain constructions close to Daniell's approach had been earlier developed by Young (see [**2010**], [**2013**], [**2015**]). It should be noted that also in the real analysis, F. Riesz proposed a scheme of integration avoiding prior construction of measure theory and leading to a somewhat more economical presentation of the fundamentals of the theory of integration (see Riesz [**1571**], [**1572**] and the textbooks mentioned below). In the middle of the 20th century there was a very widespread point of view in favor of presentation of the theory of integration following Daniell's approach, and some authors even declared the traditional presentation to be "obsolete". Apart the above-mentioned conveniences in the consideration of measures on locally compact spaces, an advantage of such an approach for pedagogical purposes seemed to be that it "leads to the goal much faster, avoiding auxiliary constructions and subtleties of measure theory". In Wiener, Paley [**1987**, p. 145], one even finds the following statement: "In an ideal course on Lebesgue integration, all theorems would be developed from the point of view of the Daniell integral". But fashions pass, and now it is perfectly clear that the way of presentation in which the integral precedes measure can be considered as no more than equivalent to the traditional one. This is caused by a number of reasons. First of all, we note that the economy of Daniell's scheme can be seen only in considerations of the very elementary properties of the Lebesgue integral (this may be important if perhaps in the course of the theory of representations of groups one has to explain briefly the concept of the integral), but in any advanced presentation of the theory this initial economy turns out to be imaginary. Secondly, the consideration of measure theory (and not only the integral) is indispensable for most applications (in many of which measures are the principal object), so in Daniell's approach sooner or later one has to prove the same theorems on measures, and they do not come as simple corollaries of the theory of the integral. It appears that even if there are problems whose investigation requires no measure theory, but involves the Lebesgue integral, then it is very likely that most of them can also be managed without the latter.

It should be added that in order to define the integral in the traditional way one needs very few facts about measures (they can be explained in a couple of lectures), so that the fears of "subtleties of measure theory" necessary for the usual definition of integral are considerably exaggerated. Also from the methodological point of view, the preliminary acquaintance with the basic concepts of measure theory is very useful for the true understanding of the role of different conditions encountered in any definition of the integral (for example, the monotone convergence). In addition, it must be said that the use of the concept of a measure zero set without definition of measure (which is practised in a number of approaches to the integral) seems to be highly unnatural independently of possible technical advantages of such constructions. Finally, it should be remarked that the approach based on Daniell's scheme turned out to be of little efficiency in the construction and investigation of measures on infinite-dimensional spaces, although consideration of measures as functionals (which was a source of Daniell's method and which should not be identified with the latter) is used here very extensively. Taking into account all these circumstances, one can conclude that application of Daniell's method in a university course on measure and integration is justified chiefly by a desire to diversify the course, to provide a stronger functional-analytic trend and minimize the set-theoretic considerations. Lebesgue [**1133**, p. 320] remarked in this connection: "S'il ne s'agit que d'une question d'ordre de paragraphes, peu m'importe, mais je crois qu'il serait mauvais de se passer de la théorie des ensembles". Certainly, for the researchers in measure theory and functional analysis, acquaintance with Daniell's method is necessary for broadening the technical arsenal. Among many books offering a systematic presentation of Daniell's approach we mention Bichteler [**166**], Cotlar, Cignoli [**377**], Filter, Weber [**586**], Hildebrandt [**831**], Hirsch, Lacombe [**834**], Janssen, van der Steen [**885**], Klambauer [**1009**], Nielsen [**1371**], Pfeffer [**1445**], Riesz, Sz.-Nagy [**1578**], Shilov, Gurevich [**1699**], and Zaanen [**2020**].

§§7.9–7.10. F. Riesz [**1568**] proved his famous representation theorem in the case $X = [a, b]$; Radon [**1514**] extended it to compact sets in $\mathbb{R}^n$. For metrizable compact spaces this result was proved by Banach and Saks (see Banach [**104**], Saks [**1642**]). Markov [**1268**] obtained related results for more general normal spaces by using finitely-additive measures, and for general compact spaces Theorem 7.10.4 was stated explicitly and proven in Kakutani [**932**]. A thorough investigation of such problems was undertaken by A.D. Alexandroff [**30**] and continued by Varadarajan [**1918**]. Theorem 7.10.6 is found in Bourbaki [**242**, Ch. IX, §5.2]. It can be extracted from the results in [**1918**]. For additional comments, see Batt [**131**], Dunford, Schwartz [**503**, Chapter IV].

It is worth noting that in [**30**] (see §2, 3^o, Definition 6, p. 326; §10, 2^o, Definition 2, p. 596), in the definition of a convergent net of functions f_α, the following condition is forgotten: for every pair of indices α and β, there exists an index γ such that $\alpha \le \gamma$, $\beta \le \gamma$ and $f_\alpha \ge f_\gamma$, $f_\beta \ge f_\gamma$. It is obvious from the proofs that this condition is implicitly included, and without it many assertions are obviously false. The main results of [**30**] on the correspondence between measures and functionals (with the aforementioned condition, of course) are equivalent to the results established in §§7.9,7.10 in terms of monotone nets. To this end, it suffices to observe that if we are given a net of functions f_α satisfying the above condition, then one can take a new directed index set $\widetilde{\Lambda}$ which consists of finite subsets of the initial index set Λ partially ordered by inclusion. For every $\lambda = (\alpha_1, \ldots, \alpha_n) \in \widetilde{\Lambda}$ we let $g_\lambda := \min(f_{\alpha_1}, \ldots, f_{\alpha_n})$. Our new net $\{g_\lambda\}_{\lambda\in\widetilde{\Lambda}}$ is decreasing. Moreover, for every $\alpha \in \Lambda$ and $\lambda \in \widetilde{\Lambda}$, there exist $\alpha' \in \Lambda$ and $\lambda' \in \widetilde{\Lambda}$ such that $\lambda \le \lambda'$, $g_{\lambda'} \le f_\alpha$, $\alpha \le \alpha'$, and $f_{\alpha'} \le g_\lambda$. Indeed, under our assumptions one can find an index α' such that $\alpha_i \le \alpha'$ and $f_{\alpha'} \le f_{\alpha_i}$ whenever $i = 1, \ldots, n$.

Various results connected to integral representations of linear functionals on function spaces and related topologies on spaces of functions and measures, in particular, generalizations of the Riesz theorem, are discussed in Anger, Portenier [**53**], Collins [**364**], Fremlin [**619**], Garling [**668**], Hewitt [**824**], Lorch [**1183**], Mosiman, Wheeler [**1336**], Pollard, Topsøe [**1480**], Topsøe [**1876**], Zakharov, Mikhalev [**2024**]. The number of related publications is very high. It should be noted, though, that in this direction there are many rather artificial settings of problems that are far removed from any applications.

§7.11. Measure theory on locally compact spaces is presented in many books, including Bourbaki [**242**], Dinculeanu [**453**]. For this reason, in this book we give minimal attention to this question, although we include the principal results.

§7.12. The investigation of general probability measures on Banach and more general linear spaces was initiated by Kolmogorov [**1026**], Fréchet (see [**615**], [**616**], [**618**]), Fortet, Mourier [**600**], Mourier [**1338**], Bochner [**202**], Prohorov [**1497**]. An important motivation was the construction of the Wiener measure [**1984**], [**1986**]. Later, measures on linear spaces were studied in Badrikian [**91**], Badrikian, Chevet [**92**], Chevet [**339**], Da Prato, Zabczyk [**392**], Gelfand, Vilenkin [**677**], Grenander [**739**], Hoffmann-Jørgensen [**845**], Kuo [**1080**], Ledoux, Talagrand [**1140**], Schwartz [**1683**], [**1685**], Skorokhod

[**1741**], Słowikowski [**1742**], Umemura [**1901**], Vakhania [**1907**], Vershik, Sudakov [**1926**], Xia [**1999**], Yamasaki [**2000**]. The most complete exposition of the linear theory is given in the book Vakhania, Tarieladze, Chobanyan [**1910**], which has become a standard reference in the field. Sudakov [**1803**] developed an interesting direction in measure theory on linear spaces, connected with geometry and approximation theory.

For the theory of random processes, it is important to consider measures in sufficiently general function spaces. In those cases where such a space is not Polish or Souslin (like, e.g., the space of all functions on the interval with the topology of pointwise convergence), there arise various problems with measurability, partly described in the text. Such problems were investigated in Ambrose [**41**], Doob [**467**], [**463**], [**465**], Chentsov [**335**], [**336**], [**337**], [**338**], Kakutani [**933**], Nelson [**1359**]. The main motif of these works is an extension of a measure μ on the σ-algebra generated by cylinders in the spaces $[0,1]^T$ or $\mathbb{R}^T$ to a measure on larger σ-algebras. Such a question arose naturally after the appearance of Kolmogorov's theorem. One of the observations in Kakutani [**933**] (see also Nelson [**1359**]) is that if in place of $\mathbb{R}^T$ one considers the compact space $\overline{\mathbb{R}}^T$, where $\overline{\mathbb{R}}$ is the one-point compactification of the real line, then a Baire measure μ on this compact space can be extended to a Radon measure, which makes measurable many more sets than in the usual construction of Kolmogorov. However, Bourbaki and N.N. Chentsov observed independently that anyway, many natural and effectively described sets remain nonmeasurable (see Exercises 7.14.157, 7.14.158); a result of this kind is found in Hewitt, Ross [**825**, §16.13(f)]. Related aspects are discussed in Kendall [**981**], Talagrand [**1833**].

Kuelbs [**1073**] showed that a Radon measure on a Banach space X is concentrated on a compactly embedded Banach space E, and the constructed space E was a dual space (not necessarily separable). Ostrovskiĭ [**1406**] showed in a different way that E can be taken to be a dual space, and Buldygin [**274**] proved that E can be chosen to be separable reflexive. In Bogachev [**205**], this fact was extended to Fréchet spaces by means of a short reasoning combining some ideas from [**1073**] and [**274**] (it is given in Theorem 7.12.4).

Concerning moments of measures, see Vakhania, Tarieladze, Chobanyan [**1910**], Kruglov [**1063**], Graf, Luschgy [**722**], Ledoux, Talagrand [**1140**], Kwapień, Woyczyński [**1096**].

Convergence of random series and other limit theorems in infinite-dimensional spaces are considered in Buldygin [**273**], Vakhania [**1907**], Vakhania, Tarieladze, Chobanyan [**1910**], Buldygin, Solntsev [**276**], Kwapień, Woyczyński [**1096**].

Differential properties of measures on infinite-dimensional spaces are investigated in Bogachev [**206**], Bogachev, Smolyanov [**225**], Dalecky, Fomin [**394**], and Uglanov [**1896**], which contain extensive bibliographies.

§7.13. Characteristic functionals of measures on infinite-dimensional spaces were introduced by Kolmogorov [**1027**]. Later they were considered by many other authors (see, e.g., Le Cam [**1137**], Prohorov [**1497**], [**1498**], Prohorov, Sazonov [**1499**]). Important ideas related to characteristic functionals and developed later in other works were proposed in Prohorov [**1497**]. As observed by Kolmogorov [**1031**], the work [**1497**] contained the main inequality on which are based the celebrated theorems of Minlos and Sazonov on the description of characteristic functionals of measures on the duals to nuclear spaces and Hilbert spaces. It should be noted that in spite of the subsequent intensive studies in this field and numerous generalizations of these two theorems, in applications one uses these original results. Extensive information on characteristic functionals of measures on locally convex spaces is presented in the books Vakhania, Tarieladze, Chobanyan [**1910**] and Mushtari [**1348**]. See also the papers Gross [**743**], Kwapień, Tarieladze [**1095**], Mouchtari [**1337**], Mushtari, Chuprunov [**1349**], Smolyanov [**1754**], Smolyanov, Fomin [**1755**], Tarieladze [**1839**]. There is an extensive literature (see the works cited above) devoted to the so-called sufficient topologies on locally convex spaces (i.e., topologies τ on X^* such that the τ-continuity of the Fourier transform of a nonnegative cylindrical quasi-measure ν on X implies the tightness of ν) and necessary topologies (respectively, the topologies τ on X^* in which are continuous the characteristic functionals of all tight nonnegative cylindrical quasi-measures on X). An important result due to Tarieladze [**1840**], [**1841**] states that any sufficient topology is sufficient for signed measures as well in the following sense: let τ be a sufficient topology on X^* and let φ be the τ-continuous Fourier transform of a signed cylindrical quasi-measure μ of bounded variation on $\sigma(X^*)$; then μ is countably additive and tight (the question about this was raised by O.G. Smolyanov in the 1970s and in some special cases was answered positively by E.T. Shavgulidze). However, in this assertion one cannot replace the boundedness of variation of μ by the boundedness of $|\varphi|$ (Exercise 7.14.135). Smolyanov, Shavgulidze [**1756**] simplified the proof of the Tarieladze theorem. Related to this circle of problems is the concept of measurable seminorm (not in the sense of measurability with respect to a measure), which is discussed in Dudley, Feldman, Le Cam [**496**], Maeda [**1225**], Maeda, Harai, Hagihara [**1226**], Smolyanov [**1754**].

§7.14. An interesting example connected with measurability on products is constructed in Dudley [**492**], [**493**].

The term "completion regular" was used in Halmos [**779**]. Moran [**1330**] introduced the property of measure-compactness. Related properties were also considered in Gardner [**660**], Gardner, Pfeffer [**666**], Okada, Okazaki [**1396**].

The separability of Radon measures on compact spaces was investigated in Dzamonja, Kunen [**509**], Kunen, van Mill [**1078**], and Plebanek [**1467**], where one can find additional references. In particular, it was shown that the question of the existence of a first countable Corson compact space that is the support of a nonseparable Radon measure is undecidable in ZFC (with an extra set-theoretic assumption such a space is constructed in [**1078**], and the non-existence result is established in [**1467**] under the negation of that additional assumption).

Theorem 7.14.3 goes back to a result of Kakutani [**933**] who proved that if Ω_γ, $\gamma \in \Gamma$, are compact metric spaces equipped with Borel probability measures μ_γ that are positive on nonempty open sets, then the Lebesgue completion of the product measure $\bigotimes_{\gamma\in\Gamma} \mu_\gamma$ coincides with the Radon measure μ constructed from the measure $\bigotimes_{\gamma\in\Gamma} \mu_\gamma$ by means of the Riesz theorem; in other words, all Borel sets belong to the Lebesgue completion of $\bigotimes_{\gamma\in\Gamma} \mathcal{B}(\Omega_\gamma)$. Concerning other results connected with completion regular measures, see also Babiker, Graf [**86**], Babiker, Knowles [**87**], Gryllakis [**748**]. Wheeler [**1979**] raised the question whether any finite τ-additive Baire measure μ on a completely regular space X has a Lindelöf subset of full μ-outer measure. If such a set exists, then (X, μ) is said to have property L. Aldaz [**18**] investigated from this point of view the Sorgenfrey plane X with Lebesgue measure λ. He proved that (i) there exists a model of the set theory ZF in which (X, λ) has no property L, (ii) (X, λ) has property L in ZFC+CH, (iii) the existence of a τ-additive measure without property L is consistent with ZFC. Finally, Plebanek [**1469**]) constructed an example (in ZFC) of a τ-additive Baire measure without Lindelöf subspaces of full measure.

Interesting examples of compact spaces without strictly positive measures (i.e., positive on nonempty open sets) are constructed in Argyros [**65**]. A discussion of connections between strictly positive measures on a compact space X, strictly convex renormings of $C(X)$, and the chain condition can be found in Comfort, Negrepontis [**366**, Ch. VI]. Connections between nonmeasurable cardinals and existence of separable supports of measures on metric spaces are studied in Marczewski, Sikorski [**1260**]. For additional information about supports of measures, see Adamski [**6**], van Casteren [**320**], Gardner [**660**], Gardner, Pfeffer [**666**], Hebert, Lacey [**805**], Kharazishvili [**988**], Okada [**1395**], Plebanek [**1468**], Sato [**1651**], Seidel [**1690**].

Generalizations of Lusin's theorem were considered by many authors. For example, Schaerf [**1662**] gave a generalization in the case of mappings from topological spaces to second countable spaces. Sometimes the measurability is defined as Lusin's C-property (see Bourbaki [**242**]).

Approximations of analytical sets by compact sets for some outer measures were also constructed in Glivenko [**698**], Kelley [**977**]. The paper Mattila, Mauldin [**1273**] deals with the measurability of functions of the form $K \mapsto h(K)$ on the space of compact sets in a Polish space equipped with the Hausdorff distance, where h is some set function, for example, a Hausdorff measure.

The foundations of the abstract theory of capacities were laid by Choquet [**349**], [**350**], [**351**], but certain assertions had been known earlier. For example, Korovkin [**1041**] proved an analog of Egoroff's theorem for capacities.

As shown by Alexandroff [**30**] and Glicksberg [**696**], a Hausdorff space X is pseudocompact if and only if every additive regular set function on X is countably additive on $\mathcal{B}a(X)$.

Vakhania, Tarieladze, Chobanyan [**1910**, §I.5] give a more direct (but longer) proof of Corollary 7.14.59.

There are examples where two distinct Borel probability measures on a compact metric space coincide on all balls, see Davies [**412**], [**415**], Darst [**408**]. According to Preiss, Tišer [**1491**], two Radon probability measures on a Banach space that agree on all balls are equal. The problem of to what extent a measure is determined by its values on balls is discussed in Riss [**1582**], [**1583**]. For related results, see Gorin, Koldobskiĭ [**714**], Mejlbro, Preiss, Tišer [**1298**], Preiss [**1487**], Preiss, Tišer [**1490**].

Connections between measure and category had already been examined in the 1930s, see, e.g., Sierpiński [**1718**], Szpilrajn [**1813**], Marczewski, Sikorski [**1261**]; as a few later works we mention Oxtoby [**1409**], Ayerbe-Toledano [**82**], Gardner [**660**].

Concerning the theory of infinitely divisible and stable measures we refer to the books Hazod, Siebert [**804**], Kruglov [**1063**], Linde [**1172**] and the papers Acosta [**1**], Acosta, Samur [**2**], Bogachev [**204**], Dudley, Kanter [**497**], Fernique [**564**], Kanter [**949**], Linde [**1172**], Sztencel [**1820**], Tortrat [**1888**].

Convex measures are studied in Bobkov [**193**], Bogachev, Kolesnikov [**213**], [**214**], Borell [**236**], [**238**], [**239**], Krugova [**1064**].

The theory of Gaussian measures is presented in detail in the recent books Bogachev [**208**], Fernique [**570**], and Lifshits [**1171**], where one can find an extensive bibliography.

The notion of a measurable linear function is connected with that of the linear kernel of a measure μ (i.e., the topological dual to the space X^* equipped with the topology of convergence in measure μ), which is not discussed here; see Chevet [**339**], [**340**], Khafizov [**984**], Kwapień, Tarieladze [**1095**], Smoleński [**1747**], [**1748**], [**1749**], [**1750**], Takahashi [**1824**], Tien, Tarieladze [**1855**], Urbanik [**1902**] and the references therein. Measurable polylinear functions are considered in Smolyanov [**1751**].

Measures on groups and related concepts are studied in Armstrong [**69**], Becker, Kechris [**141**], Berg, Christensen, Ressel [**152**], Bloom, Heyer [**191**], Csiszár [**389**], Edwards [**519**], Fox [**601**], Grekas [**735**], [**736**], Hazod, Siebert [**804**], Hewitt, Ross [**825**], Heyer [**828**], [**829**], Högnäs, Mukherjea [**849**], Panzone, Segovia [**1421**], Peterson [**1438**], Pier [**1454**], Sazonov, Tutubalin [**1658**], and Wijsman [**1988**], where one can find a more complete bibliography.

Various regularity properties of measures are discussed in Adamski [**7**], [**10**], Anger, Portenier [**53**], Babiker [**84**], Babiker, Graf [**86**], Bachman, Sultan [**89**], Berezanskiĭ [**150**], Cooper, Schachermayer [**375**], Dixmier [**458**], Flachsmeyer, Lotz [**589**], Fremlin [**626**], Gardner [**660**], [**666**], Gould, Mahowald [**715**], Katětov [**960**], Kharazishvili [**988**], [**990**], Kubokawa [**1068**], Lotz [**1193**], de Maria, Rodriguez-Salinas [**1265**], Métivier [**1308**], Plebanek

[**1470**], [**1471**], Prinz [**1495**], Rao [**1541**], Sondermann [**1766**], Topsøe [**1873**], [**1878**], [**1879**], [**1880**].

Radon measures are considered in many papers and books, in particular, in Anger, Portenier [**53**], Bogachev [**208**], Bourbaki [**242**], Schwartz [**1681**], Semadeni [**1691**], Tjur [**1861**], Vakhania, Tarieladze, Chobanyan [**1910**].

Assertion (i) in Example 7.14.60 goes back to Ionescu Tulcea [**862**], [**863**]; Tortrat [**1890**] extended it to metrizable locally convex spaces (the proof is similar; this result is called the Tortrat theorem). The existence of Radon extensions with respect to the norm topology for weakly Radon measures goes back to Phillips [**1452**] where a result of this sort (called the Phillips theorem) is obtained in the form of the strong measurability of weakly measurable mappings; an analogous assertion was also obtained by A. Grothendieck.

Measures on Banach spaces with the weak topology are discussed in many works, see, e.g., de Maria, Rodriguez-Salinas [**1266**], Jayne, Rogers [**888**], Rybakov [**1630**], Schachermayer [**1659**], Talagrand [**1834**].

In addition to the works cited in §7.14(xviii), infinite Borel measures are studied in Jimenez-Guerra, Rodriguez-Salinas [**901**], Novoa [**1386**], Rodriguez-Salinas [**1585**]. Products of infinite measures are considered in Elliott [**527**], Elliott, Morse [**528**], Hahn [**772**], and Luther [**1213**], where one can find additional references.

Certain special properties of compact sets related to measures are studied in Dzamonja, Kunen [**508**], [**509**], Fremlin [**632**], Kunen, van Mill [**1078**].

Chapter 8.

§§8.1–8.4. A large portion of the results in this chapter is taken from the outstanding works of A.D. Alexandroff [**30**] and Yu.V. Prohorov [**1497**] who laid the foundations of the modern theory. As pointed by A.D. Alexandroff himself, a source of his abstract work in general measure theory was his research [**29**] (see Alexandrov [**32**]) in geometry of convex bodies. Among important earlier works we note Helly [**809**], Radon [**1514**], Bray [**250**], and a series of works of Lévy, including his book [**1167**] containing results on convergence of the distribution functions. Close to them in the sense of ideas are the paper Gâteaux [**672**] and Lévy's book [**1166**] on averaging on functional spaces. Let us also mention Glivenko [**699**]. The subsequent development of this area was considerably influenced by the works of Skorohod [**1739**], [**1740**], Le Cam [**1138**], and Varadarajan [**1918**]. It had already been shown by Radon [**1514**] that every bounded sequence of signed measures on a compact set in $\mathbb{R}^n$ contains a weakly convergent subsequence; earlier in the one-dimensional case the result had been obtained by Helly [**809**] in terms of functions of bounded variation. The term "schwach konvergent" — weakly convergent — was used by Radon in [**1516**]. The space of measures and weak convergence were employed by Radon in the study of the operators adjoint to linear operators on spaces of continuous functions and in potential theory. Bogolubov and Krylov [**227**] (in the paper spelled as Bogoliouboff and Kryloff) showed that a complete separable metric space X is compact precisely when the space of probability measures on X is compact in the weak topology. In the same work, they proved the uniform tightness of any weakly compact set of probability measures on a metric space whose balls are compact. The space of probability measures with the weak topology was also investigated in Blau [**187**] (who considered the A-topology). It should be noted that in many works Alexandroff's theorem on weak convergence (Theorem 8.2.3) is called the "portmanteau theorem". The English word "portmanteau" (originally a French word, meaning a coat-hanger) has the archaic meaning of a large traveling bag and may also denote multi-purpose or multi-function objects or concepts. I do not know who invented such a nonsensical name for Alexandroff's theorem. It seems there is no need to attach a meaningless label without any mnemonic content to a result with obvious and generally recognized authorship, rather than just calling it by the inventor's name.

The continuity sets of measures on $\mathbb{R}^n$ were considered in Gunther [**752**, p. 13], Jessen, Wintner [**900**], Cramér, Wold [**381**]. Romanovsky [**1603**] studied locally uniform convergence of multivariate characteristic functions. Multivariate distribution functions and their weak convergence were also considered in Haviland [**799**].

Beginning from the 1950s, in the theory of weak convergence of measures, apart from a purely probabilistic direction related to the study of asymptotic behavior of random variables, there has been intensive development of the direction laid by the above-mentioned works by A.D. Alexandroff and Yu.V.Prohorov and belonging rather to measure theory and functional analysis but in many respects furnishing the foundations for the first direction. Naturally, in our book only this second direction is discussed.

The fundamentals of the theory of weak convergence of measures on metric spaces are presented in the books Billingsley [**169**] and Gikhman, Skorokhod [**685**]. See also Bergström [**155**], [**156**], Dalecky, Fomin [**394**], Dudley [**495**], Ethier, Kurtz [**543**], Gänssler [**654**], Gänssler, Stute [**656**], Hennequin, Tortrat [**811**], Hoffmann-Jørgensen [**847**], Kruglov [**1063**], Pollard [**1478**], Shiryaev [**1700**], Stroock [**1797**], Stroock, Varadhan [**1799**], Vakhania, Tarieladze, Chobanyan [**1910**]. Weak convergence and weak compactness are investigated in an important series of works by Topsøe (see [**1873**] and [**1870**], [**1871**], [**1872**], [**1874**], [**1875**], [**1877**]).

Proposition 8.2.8 was obtained in Prohorov [**1497**] in the case of complete separable metric spaces, but extensions to more general cases meet no difficulties (this concerns Theorem 8.2.13 and Theorem 8.2.17 as well).

The Kantorovich–Rubinshtein metric goes back to Kantorovich's work [**951**]. Later this metric was used in Fortet, Mourier [**599**] in the study of convergence of empirical distributions. In relation to some extremal problems, the Kantorovich–Rubinshtein metric was considered in Kantorovich, Rubinshtein [**953**], [**954**] in the case of compact metric spaces (in a somewhat different form); see also Kantorovich, Akilov [**952**, Ch. VIII, §4] and comments in Vershik [**1925**]. In form (8.10.5) this metric was also defined in Vasershtein

[**1919**] (sometimes $W(\mu,\nu)$ is also called the Wasserstein metric, see, e.g., Dobrushin [**460**], although there is no author with this name). An extensive bibliography on related problems can be found in Rachev [**1506**], [**1507**]. Some comments given below in relation to metrics on spaces of probability measures also concern the Kantorovich–Rubinshtein metric. For a study of geometry of metric spaces of measures, see Ambrosio [**45**] and Sturm [**1800**].

§8.5. Additional results on the Skorohod representation and parameterization of weakly convergent sequences of measures or the set of all probability measures can be found in Banakh, Bogachev, Kolesnikov [**114**], [**115**], [**116**], [**117**], Bogachev, Kolesnikov [**211**], Choban [**342**], Cuesta-Albertos, Matrán-Bea [**391**], Jakubowski [**879**], Letta, Pratelli [**1160**], Schief [**1671**], Tuero [**1894**], Wichura [**1981**]. An interesting approach to parameterization of measures on $\mathbb{R}^n$ has been suggested by Krylov [**1067**] who obtained a parameterization with certain differentiability properties. This method is also connected with the Monge–Kantorovich problem (see, e.g., Bogachev, Kolesnikov [**214**, Example 2.1]) and certain extremal problems for measures with given marginals, which is briefly discussed in §9.12(vii). It is to be noted that in Blackwell, Dubins [**184**], there is a very short sketch of the proof of Theorem 8.5.4, but a detailed proof on this way with the verification of all details is not that short (see Fernique [**566**] and Lebedev [**1117**, Ch. 5]).

§§8.6–8.9. Investigations of weak compactness in spaces of measures and conditions of tightness were considerably influenced by the already-mentioned Prohorov work [**1497**], the ideas, methods, and concrete results of which are now presented in textbooks and have for half a century been successfully applied by many researchers. It is worth noting that in this work the fundamental Prohorov theorem was proved for probability measures on complete separable metric spaces, but the term "Prohorov theorem" is traditionally applied to numerous later generalizations of the whole theorem or only its direct or inverse assertions. This is explained by the exceptional importance of the phenomenon discovered in the theorem, whose value in the theory and applications even in the case of the simplest spaces is not overshadowed by deep and non-trivial extensions. A.D. Alexandroff [**30**] established the "absence of eluding load" (his own terminology) for weakly convergent sequences of measures (see Proposition 8.1.10), which yields directly certain partial cases of the Prohorov theorem. The idea to apply weak convergence in l^1 to weak convergence of measures is also due to A.D. Alexandroff [**30**]. Dieudonné [**449**] established the uniform tightness of any weakly convergent sequence of Radon measures on a paracompact locally compact space and constructed an example showing that the local compactness alone is not enough. Le Cam [**1138**] proved that in the case of a locally compact σ-compact space X, a family of measures is relatively compact in $\mathcal{M}_t(X)$ with the weak topology precisely when it is uniformly tight. He also observed that this assertion follows from Dieudonné [**448**]. The fact that the uniform tightness of a family of measures implies the compactness of its closure in the case of general completely regular spaces was observed by several researchers (L. Le Cam, P.-A. Meyer,

L. Schwartz) soon after the appearance of Prohorov's work and under its influence. The proof of this fact is quite simple, unlike the less obvious inverse assertion and the sequential compactness which hold for more narrow classes of spaces. Certainly, the consideration of signed measures brings additional difficulties. Example 8.6.9 is borrowed from Varadarajan [**1918**]. Compactness conditions for capacities are considered in O'Brien, Watson [**1388**].

The important Theorem 8.7.1 was established by A.D. Alexandroff [**30**] for Borel measures on perfectly normal spaces, but an analogous proof applies to Baire measures on arbitrary spaces. The proof given in the text is due to Le Cam [**1138**].

Theorem 8.9.4 is due to Varadarajan [**1918**] (see also Granirer [**729**] for another proof).

It was proved in Varadarajan [**1917**], Hoffmann-Jørgensen [**841**], Schwartz [**1681**], and Oppel [**1401**], [**1402**] that the spaces of measures on a space X are Souslin or Lusin in the weak topology under appropriate conditions on X. The fact that the space of signed measures of unit variation norm on a Polish space is Polish in the weak topology was established in Oppel [**1402**]. Additional results and references concerning properties of spaces of measures and connections with general topology can be found in Banakh [**113**], Banakh, Cauty [**118**], Banakh, Radul [**119**], [**120**], Brow, Cox [**261**], Constantinescu [**367**], [**368**], [**369**], [**370**], Fedorchuk [**557**], [**559**], [**558**], Flachsmeyer, Terpe [**590**], Frankiewicz, Plebanek, Ryll-Nardzewski [**602**], Kirk [**1005**], [**1006**], Koumoullis [**1044**], Talagrand [**1830**].

A number of authors investigated locally convex topologies on the space $C_b(X)$ for which the dual spaces are spaces of measures; these investigations are also connected with consideration of tight or weakly compact families of measures, see Conway [**373**], Hoffmann-Jørgensen [**843**], Mosiman, Wheeler [**1336**], Sentilles [**1692**], and the survey Wheeler [**1979**].

It is shown in Mohapl [**1325**] that if X is a complete metric space, then the space $\mathcal{M}_r(X)$ of Radon measures coincides with the space of all bounded linear functionals l on the space of bounded Lipschitzian functions on X such that the restriction of l to the unit ball in the sup-norm is continuous in the topology of uniform convergence on compact sets.

§§8.10. Prohorov's work [**1497**] had a decisive influence on the development of the theory of weak convergence, and the appearance of the concept of a "Prohorov space" illustrates this. It is worth noting that in the literature one can find several different notions of a "Prohorov space". Indeed, for generalizations of the Prohorov theorem one has at least the following possibilities: (1) to consider compact families of tight nonnegative Baire measures (as in Definition 8.10.8); (2) to consider compact families of not necessarily tight nonnegative Baire measures; (3) to consider weakly convergent sequences of tight nonnegative Baire measures with tight limits; (4) to consider countably compact families of type (1) or (2); (5) to consider in (1)–(4) completely bounded (i.e., precompact) families instead of compact; (6) to deal with signed measures in place of nonnegative ones. Certainly, there exist other reasonable possibilities. The situation with signed measures is less studied.

Prohorov spaces are investigated in Banakh, Bogachev, Kolesnikov [**114**], [**115**], Choban [**342**], Cox [**379**], Koumoullis [**1047**], [**1048**], Mosiman, Wheeler [**1336**], Smolyanov [**1753**]. Saint-Raymond [**1638**] gives a simpler proof that $\mathbb{Q}$ is not a Prohorov space.

The last claim of Example 8.10.14 (borrowed from Hoffmann-Jørgensen [**844**]) was stated in Smolyanov, Fomin [**1755**] for signed measures (and reproduced in Daletskii, Smolyanov [**394**]); however, it is not clear whether it remains true for signed measures because its proof was based on the erroneous Lemma 3 in [**1755**] (see also [**394**, Lemma 2.1, Ch. III] and [**395**]) asserting that for any disjoint sequence of compact sets K_n with disjoint open neighborhoods and any weakly convergent sequence $\{\mu_n\}$ of Radon measures one has $\lim\limits_{n\to\infty} \sup_i |\mu_i|(K_n) = 0$. Clearly, this is false if K_n is the point $1/n$ in $[0,1]$ and μ_n is Dirac's measure at this point. Example 8.10.25 is taken from Fremlin, Garling, Haydon [**636**] (its special case can also be found in [**1755**, §5, Theorem 3], but the proof contains the above-mentioned gap). In their spirit and ideas, these assertions are close to the results of A.D. Alexandroff in §8.1 on the "absence of eluding load".

Concerning weak convergence of measures on nonseparable metric spaces, see Dudley [**488**], [**490**], van der Vaart, Wellner [**1915**].

In addition to the already-mentioned works, the weak topology and weak convergence of measures are the main subjects in Adamski [**5**], Baushev [**137**], Borovkov [**240**], Conway [**374**], Crauel [**382**], De Giorgi, Letta [**420**], Dudley [**489**], [**491**], Fernique [**563**], [**567**], [**568**], Kallianpur [**940**], Léger, Soury [**1144**], Mohapl [**1324**], Nakanishi [**1354**], Pollard [**1475**], [**1477**], Prigarin [**1494**], Wilson [**1992**].

On weak compactness in spaces of measures, see also Adamski, Gänssler, Kaiser [**11**], Fernique [**567**], [**568**], Gerard [**681**], [**682**], Haydon [**801**], Pollard [**1476**]. Uniformity in weak convergence is studied in Billingsley, Topsøe [**171**]. Some properties of the weak topology on the space of measures on a compact space and averaging operators are considered in Bade [**90**].

Young measures are called after L.C. Young (who used them in the calculus of variations, see [**2004**]), a son of W.H. Young and G.C. Young.

Metrics on certain subspaces of the space of measures (mainly on the subspace of probability measures) were studied in Dudley [**491**], [**494**], [**495**], Givens, Shortt [**692**], Kakosyan, Klebanov, Rachev [**931**], Rachev, Rüschendorf [**1508**], Zolotarev [**2034**], [**2035**], where one can find additional references. Theorem 8.10.45 was proved in Kantorovich, Rubinshtein [**954**]. Other proofs were proposed by a number of authors, see Fernique [**565**], Szulga [**1821**]. A metric analogous to the L^p-metric of the Kantorovich–Rubinshtein type was considered in Kusuoka, Nakayama [**1091**] on the set of pairs (μ, ξ), where μ is a probability measure and ξ is a mapping. The Kantorovich–Rubinshtein norm on the space of signed measures was considered in Fedorchuk, Sadovnichiĭ [**560**], Hanin [**784**], and Sadovnichii [**1635**] (note that in [**784**, Proposition 4] it is mistakenly claimed that convergence with respect to the Kantorovich–Rubinshtein norm is equivalent to weak convergence for uniformly bounded sequences of signed measures; see Exercise 8.10.138). The property of the Kantorovich–Rubinshtein norm $\|\cdot\|_0^*$ described in Exercise 8.10.143 was discovered by Kantorovich and Rubinshtein [**954**]. This property means that the space of Lipschitzian functions on a bounded metric space vanishing at a fixed point is the dual space to the space M_0 of signed measures of total zero mass equipped with norm $\|\cdot\|_0^*$. This gives another proof of the fact that in nontrivial cases the weak topology on the whole space M_0 does not coincide with the topology generated by $\|\cdot\|_0^*$.

Convergence classes for probability measures in the sense of Theorem 8.10.56 have been investigated by several authors. It has been established that (i) the class $\mathcal{G}$ of all open sets is a convergence class for τ-additive measures on regular spaces; (ii) the class $\mathcal{G}_0$ of all functionally open sets is a convergence class for Baire measures on Hausdorff spaces, for τ-additive measures on completely regular spaces, and for regular Borel measures on normal spaces; (iii) the class $\mathcal{G}_r$ of all regular open sets is a convergence class for τ-additive measures on regular spaces and for regular Borel measures on normal spaces. Proofs of these facts and additional references can be found in Adamski, Gänssler, Kaiser [**11**].

In some problems, one has to consider spaces of locally finite measures on a locally compact space M with the topology of duality with $C_0(M)$. For example, the configuration space Γ_M is the set of all measures of the form $\gamma = \sum_{n=1}^{\infty} k_n \delta_{x_n}$, where k_n are nonnegative integer numbers and $\{x_n\} \subset M$ has no limit points. The compactness conditions in Γ_M are obtained in Bogachev, Pugachev, Röckner [**222**], where one can find additional references.

Chapter 9.

§§9.1–9.2. Some results on nonlinear transformations of measures were known in the early years of the theory of integration. For example, Riesz [**1569**, p. 497] noted without proof that every measurable set in $\mathbb{R}^n$ of measure m can be mapped by means of a measure-preserving one-to-one function onto an interval of length m, and Radon [**1514**, p. 1342] considered an isomorphism between a square with the two-dimensional Lebesgue measure and an interval with the linear Lebesgue measure (these observations were not forgotten and were later noted, for example, in Bochner, von Neumann [**203**]). Intensive investigations of transformations of measures began in the 1930s, when problems related to transformations of measures arose not only in measure theory, but also in such fields as the theory of dynamical systems, functional analysis, and probability theory. Steinhaus [**1784**] constructed a mapping $\theta\colon (0,1) \to (0,1)^\infty$ that is one-to-one on a set of full measure and transforms Lebesgue measure λ into λ^∞ (see Exercise 9.12.50). The goal of his work was to study random series. This goal was shared by a series of works by Wiener, Paley, and Zygmund (see references and comments in the book Wiener, Paley [**1987**]). In particular, the Wiener measure on the infinite-dimensional space of continuous functions was represented as the image of Lebesgue measure under some measurable mapping. The theory of dynamical systems was also an important impetus in the development of the theory of nonlinear transformations of measures. In this connection one has to mention the works Birkhoff [**174**], Bogoliouboff, Kryloff [**227**], Hopf [**854**], von Neumann [**1362**], [**1361**] (see also Halmos, Neumann [**781**]), and Oxtoby, Ulam [**1411**], [**1412**]. Finally, an important role was played by works on invariant measures on groups.

Application of measurable selection theorems to the proof of the existence of preimages of measures, as in Theorem 9.1.3, is standard and was employed by many authors (see, e.g., Varadarajan [**1917**, Lemma 2.2], Mackey [**1223**]). In Bourbaki [**242**, Ch. IX, §2.4], the existence of a preimage of a measure under a surjection of Souslin spaces is deduced from Theorem 9.1.9 and certain properties of capacities. A result analogous to Theorem 9.1.9 was proved in Fremlin, Garling, Haydon [**636**]. Lembcke [**1149**], [**1150**], [**1152**], introduced the following terminology: a Borel mapping $f\colon X \to Y$ between topological spaces is called conservative if every nonnegative Radon measure μ on Y such that $\mu^*(C \cap f(X)) = \mu(C)$ for every compact set $C \subset Y$, has a Radon preimage in X (in these works, unbounded measures are considered as well). Such a mapping is called strongly conservative if a preimage exists provided that the set $Y \backslash f(X)$ is μ-zero. According to [**1152**, Theorem 3.3], a continuous mapping f is strongly conservative if $f^{-1}(C)$ is contained in a $\mathcal{K}$-analytic subset of X for every compact set $C \subset Y$, and f is conservative if the same is true for all compact sets $C \subset f(X)$. Preimages of measures were also studied in Bauer [**133**], [**134**].

Proposition 9.1.7 was proved in Federer, Morse [**556**] by using an analogous result for continuous f obtained earlier by Banach [**100**] (this result was presented in Saks [**1640**, p. 282, Ch. IX, §7, Lemma 7.1] and found independently also by Kolmogorov [**1025**]).

An analog of Proposition 9.1.12 for infinite Baire measures is obtained in Kellerer [**976**], which gives a necessary and sufficient condition for the existence of a continuous transformation of an infinite Baire measure into Lebesgue measure on a half-line or on the whole real line.

The existence of simultaneous preimages for a family of measures μ_α on spaces X_α and mappings $f_\alpha\colon X \to X_\alpha$ was investigated in Lembcke [**1149**], [**1150**], [**1152**] and in the works cited therein. Related problems were considered by Ershov [**538**], [**539**], [**540**], [**542**] who developed a general approach to stochastic equations as the problem of finding preimages of measures under measurable mappings. On a related problem of finding measures with given marginal projections, see §9.12(vii).

§§9.3–9.5. Kolmogorov [**1022**] defined an isometry between two measures as an isometry between the corresponding measure algebras and singled out the separable case, noting that in that case there is an isometry with a measure on an interval. Szpilrajn [**1818**] showed that for a probability measure μ on $(X, \mathcal{A})$, the space $\mathcal{A}/\mu$ is isometric to the space $\mathcal{L}/\lambda$, where λ is Lebesgue measure on $[0, 1]$ and $\mathcal{L}$ is the class of all measurable sets, precisely when μ is separable and has no atoms. A finer classification of separable measure spaces was proposed independently by Halmos and von Neumann [**781**] and Rohlin [**1595**]. Maharam [**1228**], [**1229**], [**1230**] obtained fundamental results on the structure of general measure spaces. We remark that V.A. Rohlin announced his results before World War II, but their publication was considerably delayed: Rohlin participated in the war as a volunteer, was captured and spent several years in the concentration camps, then in special filtration camps for former prisoners of war, and in the subsequent years had to overcome a lot of obstacles on his way back to science (see [**1601**]). The spaces called "Lebesgue spaces" by Rohlin deserve the name "Lebesgue–Rohlin spaces", and we follow this terminology. This class of spaces coincides with the class introduced by Halmos and von Neumann, but Rohlin's axiomatics turned out to be more convenient, and, what is most principal, Rohlin developed a deep structural theory of such spaces (see [**1593**], [**1594**], [**1596**], [**1597**], [**1598**], [**1599**], [**1600**], [**1601**]), which influenced the subsequent applications in the theory of dynamical systems. Lebesgue–Rohlin spaces and related objects are studied in Haezendonck [**764**], Ramachandran [**1520**], [**1522**], Rudolph [**1626**], de La Rue [**1627**], Vinokurov [**1929**]. The books Samorodnitskiĭ [**1645**], [**1646**] develop a theory of nonseparable analogs of Lebesgue–Rohlin spaces.

There are interesting problems of classification of measure spaces with additional structures (for example, metric, linear or differential-geometric) with the preservation of a given structure. For example, one can consider isometries of metric spaces with measures that preserve measure (see Gromov [**742**], Vershik [**1924**]).

§§9.6–9.7. Theorem 9.6.3 for compact metric spaces had been earlier proved by Bourbaki (see Bourbaki [**242**, Ch. V, §6, Exercise 8c]). On measure-preserving homeomorphisms, see Alpern, Prasad [**38**], Katok, Stepin [**961**]. The problem of description of continuous images of Lebesgue measure was raised by P.V. Paramonov as part of a more general problem of characterization of images of Lebesgue measure (on an interval or a cube) under mappings of the class C^k. This general problem is open (see also Exercise 9.12.81).

§9.8. Example 9.8.1 is borrowed from Maitra, Rao, Rao [**1238**], where it is attributed to E. Marczewski. The example from Exercise 9.12.63 was constructed by Ershov [**539**]; the example from Exercise 9.12.49 is borrowed from Fremlin [**635**, §439].

§9.9. Theorem 9.9.3 goes back to a theorem from Lusin [**1205**, §47] according to which a continuous function without property (N) takes some perfect set of measure zero to a set of positive measure. The necessity part of Theorem 9.9.3 was obtained by Rademacher [**1509**, Satz VII, p. 196] who also proved the sufficiency part for continuous functions (see Satz VIII in p. 200 of the cited work). In view of Lusin's theorem, an analogous reasoning applies to any measurable functions and yields the general result that was explicitly given in Ellis [**529**] (the proof for continuous functions given in Natanson [**1356**, §3 Ch. IX] also applies to measurable functions in view of Lusin's theorem). The proofs given in the cited works are quite simple and follow, essentially, by the measurability of images of Borel sets under Borel mappings combined with the elementary fact that every set of positive Lebesgue measure contains a nonmeasurable subset. Moreover, these proofs apply to much more general cases (in particular, yield the results from Wiśniewski [**1994**]). Some problems related to transformations of measures on $\mathbb{R}^n$ are considered in Radó, Reichelderfer [**1513**].

Nonlinear transformations of general measures arise in the study of transformations of various special measures, for example, Gaussian, see Bogachev [**208**], Üstünel, Zakai [**1905**].

§9.10. Transformations of measures generated by shifts along trajectories of dynamical systems, in particular, along integral curves of differential equations, were considered by Liouville, Poincaré, Birkhoff, Kolmogorov, von Neumann, Bogolubov and Krylov, and other classics. This problematic remains an important source of new problems in measure theory as well as a field of application of new results and methods. The study of infinite-dimensional systems appears to be a promising direction. Additional results and references can be found in Ambrosio [**43**], Ambrosio, Gigli, Savaré [**45**], Bogachev, Mayer-Wolf [**220**], Cruzeiro [**386**], DiPerna, Lions [**456**], and Peters [**1436**].

§9.11. Haar [**758**] gave the first general construction of the measures that now bear his name. Simplified constructions were given by von Neumann, H. Cartan, Weyl, and other researchers (see Banach [**103**], Cartan [**315**], Weyl [**1965**], Johnson [**906**]). Haar measures are discussed in many works, see, e.g., Bourbaki [**242**], Hewitt, Ross [**825**], Nachbin [**1352**], Naimark [**1353**], Weyl [**1965**]; in particular, in several courses on measure theory, see, e.g., Federer [**555**], Halmos [**779**], Royden [**1618**]. The books Greenleaf [**733**] and Paterson [**1426**] deal with more general invariant means on groups.

§9.12. Projective systems of measures appeared under the influence of the Kolmogorov theorem and were introduced in a more abstract setting by Bochner; they are studied in Bourbaki [**242**], Choksi [**343**], Mallory [**1244**], [**1245**], Mallory, Sion [**1246**], Métivier [**1307**], Rao, Sazonov [**1543**].

Let λ^∞ be the countable power of Lebesgue measure on $[0,1]$. Let $[0,1]^\infty$ be equipped with the following metric d: $d(x,y)^2 = \sum_{n=1}^\infty a_n(x_n - y_n)^2$, where $a_n > 0$ and $\sum_{n=1}^\infty a_n < \infty$. S. Ulam raised the question about the equality $\lambda^\infty(A) = \lambda^\infty(B)$ for isometric sets A and B in $\bigl([0,1]^\infty, d\bigr)$ (it is not assumed that the isometry extends to the whole space). Mycielski [**1351**] gave a partial answer to this question: isometric open sets have equal measures. In the same paper, he constructed metrics on $[0,1]^\infty$ that define the same topology and have the property that λ^∞ is invariant with respect to all isometries. The results of Mycielski [**1350**] yield that on any nonempty compact metric space, there is a Borel probability measure such that isometric open sets have equal measures (the paper contains a more general assertion).

In relation to §9.12(vii), see Dudley [**495**], Jacobs [**876**], Kellerer [**972**], [**973**], [**975**], Ramachandran, Röschendorf [**1524**], [**1525**], Sazonov [**1657**], Skala [**1738**], Strassen [**1791**], Sudakov [**1803**]. Some historical comments on measures with given marginals are given in Dall'Aglio [**397**]. This subsection is closely related to the Monge–Kantorovich problem of optimal measure transport, on which there is extensive literature; see the works cited in §8.10(viii) and the recent work Léonard [**1153**], where one can find many references.

In addition to his well-known theorem on representation of Boolean algebras given in the text, Stone [**1788**], [**1789**] obtained many other results on the structure of Boolean algebras. The Stone theorem can be extended to semifinite measures (the corresponding space will be locally compact), see Fremlin [**635**, §343B].

Chapter 10.

§§10.1–10.3. The concept of conditional expectation was introduced by Kolmogorov [**1026**]; an important role was played by the abstract Radon–Nikodym theorem just discovered by Nikodym. Later this concept was investigated by B. Jessen, P. Lévy, J. Doob, and many other authors (see [**895**], [**1167**], [**467**]). Certainly, one should have in mind that the heuristic concept of conditional probability had existed long before the cited works: we speak here of rigorous constructions in the framework of general measure theory. The first attempts to construct sufficiently general countably additive conditional probabilities (i.e., the regular conditional probabilities discussed in §10.4) were made in Doob [**463**] and Halmos [**777**], but Andersen and Jessen (see [**49**]) and Dieudonné (see [**446**]) constructed disproving counterexamples; see also Halmos [**778**]. Below we return to this question.

In addition to the characterization of conditional expectations as orthogonal projections or other operators with certain special properties, there is their description by means of L^1-valued measures, see Olson [**1400**].

Fundamental theorems on convergence of conditional expectations and more general martingale convergence theorems were obtained by Jessen [**895**], P. Lévy [**1167**, p. 129], Doob [**464**], [**467**], and Andersen and Jessen [**48**], [**49**], [**50**] (Kolmogorov was interested in this question too, see, e.g., his note [**1030**]), and then they became the subject of intensive studies by many authors, see the books Hall, Heyde [**776**], Hayes, Pauc [**803**], Woyczynski [**1998**], and the papers Chatterji [**326**], [**329**] which emphasize the functional-analytic aspects. There is an extensive probabilistic literature on the theory of martingales and their applications (see, e.g., Bass [**129**], Bauer [**136**], Durrett [**504**], [**505**], Edgar, Sucheston [**517**], Letta [**1157**], Neveu [**1369**], Rao [**1540**], and Shiryaev [**1700**], where one can find further references).

Interesting results on the equivalence of product measures are obtained in Fernique [**569**].

Remarks related to Example 10.3.18 are given in the comments to Chapter 4.

§§10.4–10.6. Regular conditional measures in the case of product measures were explicitly indicated by Jessen. When Doob addressed the problem of their existence in more general cases, and the above-mentioned examples by Andersen, Jessen, and Dieudonné were found, it became clear that one has to impose additional conditions of the topological character. The first general results on regular conditional measures were obtained by Dieudonné [**446**], Rohlin [**1595**], Jiřina [**903**], [**904**], Sazonov [**1656**]. In this chapter, they are presented in the modern form accumulating the contributions of many authors. Conditional measures and disintegrations are discussed in Blackwell, Dubins [**183**], Blackwell, Maitra [**185**], Blackwell, Ryll-Nardzewski [**186**], Calbrix [**302**], Chatterji [**325**], Császár [**387**], Dubins, Heath [**476**], Graf, Mauldin [**724**], Hennequin, Tortrat [**811**], Kulakova [**1075**], Ma [**1218**], Maitra, Ramakrishnan [**1237**], Metivier [**1306**], [**1307**], Musial [**1345**], Pachl [**1414**], [**1415**], Pellaumail [**1431**], Pfanzagl [**1443**], Ramachandran [**1520**], [**1521**], [**1522**], [**1523**], Rao [**1538**], [**1539**], [**1540**], [**1542**], Remy [**1548**], Rényi [**1549**], [**1550**], Saint-Pierre [**1637**], Schwartz [**1682**], [**1684**], Sokal [**1763**], Tjur [**1860**].

A number of authors, starting with A. Ionescu Tulcea and C. Ionescu Tulcea [**865**], [**866**], constructed conditional measures by using liftings; our exposition is close to Hoffmann-Jørgensen [**842**].

Concerning proper conditional measures, see Blackwell, Dubins [**183**], Blackwell, Ryll-Nardzewski [**186**], Faden [**547**], Musiał [**1345**], Sokal [**1763**].

An important role in the study of disintegrations and conditional measures was played by Pachl's work [**1414**]. One of its fascinating results was the proof of the fact that the restriction of any compact measure to a sub-σ-algebra is compact. This work, as well as Ramachandran's work [**1522**], became a basis of our exposition of part of the results in §10.5. Ramachandran [**1523**] observed that Example 10.6.5, constructed in [**1414**], solves a problem raised by Sazonov in [**1656**], i.e., shows that there exist a perfect probability space and a σ-algebra for which there are no regular conditional probabilities in the sense of Doob.

Schwartz [**1682**], Valadier [**1911**], and Edgar [**511**] considered disintegrations on product spaces. In Dieudonné [**446**], as well as in [**511**], [**1682**], [**1684**], the investigation of disintegrations is based on vector measures and the Radon–Nikodym theorem for such measures (instead of liftings). Disintegrations for unbounded measures are studied in Saint-Pierre [**1637**]. Adamski [**8**] gave a characterization of perfect measures by means of conditional measures.

The existence of a lifting for Lebesgue measure on the interval was proved by von Neumann [**1360**]. Maharam [**1231**] gave a proof in the general case, considerably more difficult than the case of Lebesgue measure (she noted that earlier von Neumann had presented orally his proof for the general case which was never written down and the details of which are unknown). Shortly after that a different proof was given by A.&C. Ionescu Tulcea (see [**864**], [**867**]). A somewhat more elementary proof was proposed in Traynor [**1892**]. The theory of liftings is thoroughly discussed in the book A. Ionescu Tulcea, C. Ionescu Tulcea [**867**]. Extensive information is presented in the books Fremlin [**635**], Levin [**1164**]. In the literature, one can find different proofs of the existence of liftings; in addition to the already-mentioned works, see Dinculeanu [**452**], Jacobs [**876**], Sion [**1736**]. On the theory of liftings, in particular, on liftings with certain additional properties (e.g., consistent with products of spaces), see also Burke [**286**], [**287**], Edgar, Sucheston [**517**], Grekas, Gryllakis [**737**], [**738**], Losert [**1191**], [**1192**], Macheras, Strauss [**1220**], [**1221**], [**1222**], Sapounakis [**1649**], Talagrand [**1832**], [**1834**]. Measurability problems related to liftings are considered in Cohn [**360**], [**361**], Talagrand [**1836**]. A recent survey is Strauss, Macheras, Musiał [**1792**].

§10.7. The Ionescu Tulcea theorem on transition probabilities (obtained in [**868**]) was generalized by several authors, see, e.g., Jacobs [**876**], Ershov [**541**]. This theorem is presented in many books, our exposition follows Neveu [**1368**].

In relation to conditional and transition measures, Burgess, Mauldin [**283**], Gardner [**661**], Maharam [**1234**], Mauldin, Preiss, von Weizsäcker [**1278**], and Preiss, Rataj [**1489**] studied families of measures possessing diverse disjointness properties (for example, pairwise mutually singular). It is shown in Fremlin, Plebanek [**638**] that under Martin's axiom, there exists a compact space X of cardinality of the continuum $\mathfrak{c}$ such that one can find $2^{\mathfrak{c}}$ mutually singular Radon measures on X.

§10.8. Measurable partitions play an important role in ergodic theory, in particular, in the classification of dynamical systems; see the books on ergodic theory cited at the beginning of §10.9 and the work Vershik [**1923**].

§10.9. The Poincaré recurrence theorem was discovered by him in connection with considerations of systems of the classical mechanics (see [**1472**, pp. 67–72] or p. 314 in V. 7 of his works), but his reasoning with obvious changes is applicable in the general case as well, which was observed by Carathéodory [**309**] (see V. 4 in [**311**]). Theorem 10.9.4, called the Birkhoff or Birkhoff–Khinchin theorem, was obtained in Birkhoff [**175**] in a somewhat less general form and was soon generalized (with certain simplification and clarification of the proof and keeping the main idea) in Khinchin [**996**]. In subsequent years many interesting applications and generalizations of this theorem were found (see Dunford, Schwartz [**503**, Ch. VIII]); we only mention a couple of old works by Hartman, Marczewski, Ryll-Nardzewski [**791**] and Riesz [**1576**], where, in particular, transformations of the interval with Lebesgue measure were considered; the modern bibliography can be found in the books cited in §10.9. A survey of estimates of the rate of convergence in ergodic theorems is given in Kachurovskiĭ [**924**]. Important works in this direction are Ivanov [**871**], [**872**] and Bishop [**177**].

§10.10. The concept of independence (of functions, sets, σ-algebras) is one of the central ones in probability theory; it is important in measure theory as well. Diverse problems of measure theory related to this concept have been studied in many works. Among many old functional-analytic works we mention Banach [**106**], [**107**], Fichtenholz, Kantorovitch [**584**], Kac [**922**], Kac, Steinhaus [**923**], Marczewski [**1250**], [**1251**], [**1253**]; one can hardly estimate the number of works of probabilistic nature. See Chaumont, Yor [**330**] for exercises on conditional independence.

Fremlin [**633**] gave a different proof of Theorem 10.10.8, also using disintegrations. Theorem 10.10.18 was obtained in Hewitt, Savage [**826**]; the presented proof is borrowed from Letta [**1158**]. See Novikoff, Barone [**1382**] for some historical remarks.

Several results close to the Komlós theorem are obtained in Chatterji [**324**], [**327**], [**328**], Gaposhkin [**658**]. Interesting and very broad generalizations of this theorem are found in Aldous [**21**], Berkes, Péter [**158**], Péter [**1435**].

Gibbs measures are a very popular object in the literature on probability theory and statistical physics; they originated in the works by Dobrushin [**460**], [**461**] and Lanford and Ruelle [**1104**] and have been investigated by many authors. The books Georgii [**680**], Preston [**1492**], Prum, Fort [**1500**], Sinai [**1729**], [**1730**] are devoted to this direction.

Triangular transformations of measures is a very interesting and sufficiently new object of study requiring modest background. In spite of the fact that such transformations are almost as universal as general isomorphisms of measures, their advantageous distinction is an effective method of construction and a simple character of dependence of the components on the coordinates. Triangular mappings have been employed in Bogachev, Kolesnikov, Medvedev [**218**] to give a positive answer to a long-standing question on the possibility of transforming a Gaussian measure μ into every probability measure ν that is absolutely continuous with respect to μ by a mapping of the form $T(x) = x + F(x)$, where F takes on values in the Cameron–Martin space of the measure μ (this result follows from assertion (ii) in Theorem 10.10.38). It remains unknown whether in assertions (ii) and (iii) in Theorem 10.10.38 one can take for T the canonical triangular mappings $T_{\mu,\nu}$. It is of interest to continue the study of the continuity and differentiability properties of canonical triangular mappings.

References

[1] Acosta A. de. *Stable measures and seminorms.*[1] Ann. Probab. 1975. V. 3. P. 865–875. [*451*][2]

[2] de Acosta A., Samur J.D. *Infinitely divisible probability measures and the converse Kolmogorov inequality in Banach spaces.* Studia Math. 1979. V. 66. P. 143–160. [*451*]

[3]* Adams M., Guillemin V. Measure theory and probability. Birkhäuser, Boston, 1996; xvi+205 pp. [3]

[4]* Adams R.A. Sobolev spaces. Academic Press, New York, 1975; 268 pp.

[5] Adamski W. *An abstract approach to weak topologies in spaces of measures.* Bull. Soc. Math. Grèce (N.S.). 1977. V. 18, №1. P. 28–68. [*456*]

[6] Adamski W. *Note on support-concentrated Borel measures.* J. Austral. Math. Soc. Ser. A. 1980. V. 29. P. 310–315. [*156, 450*]

[7] Adamski W. *Extensions of tight set functions with applications in topological measure theory.* Trans. Amer. Math. Soc. 1984. V. 283, №1. P. 353–368. [*131, 451*]

[8] Adamski W. *Factorization of measures and perfection.* Proc. Amer. Math. Soc. 1986. V. 97, №1. P. 30–32. [*444, 462*]

[9] Adamski W. *On the existence of invariant probability measures.* Israel J. Math. 1989. V. 65, №1. P. 79–95. [*336*]

[10] Adamski W. *On extremal extensions of regular contents and measures.* Proc. Amer. Math. Soc. 1994. V. 121, №4. P. 1159–1164. [*451*]

[11] Adamski W., Gänssler P., Kaiser S. *On compactness and convergence in spaces of measures.* Math. Ann. 1976. B. 220. S. 193–210. [*244, 456*]

[12] Afanas'eva L.G., Petunin Yu.G. *σ-algebras generated by comparable normed topologies.* Trudy Inst. Matem. Voronezh. Univ. 1971. №1. P. 3–11 (in Russian). [*440*]

[13]* 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.

[14]* Akhiezer (Achieser) N.I. Theory of approximation. Translated from the Russian. Frederick Ungar Publ., New York, 1956; x+307 pp. (Russian ed.: Moscow, 1947).

[15]* 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).

[16]* 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).

[17] Akin E. *Measures on Cantor space.* Topology Proc. 1999. V. 24, №2. P. 1–34. [*288*]

[1] The article titles are printed in italics to distinguish them from the book titles.

[2] In square brackets we indicate in italics all page numbers where the corresponding work is cited; for the works cited in both volumes, the labels **I** and **II** indicate the volume; if a work is cited only in vol. 2, then all the page numbers refer to this volume.

[3] The asterisk marks the works cited only in vol. 1, where the pages of citing are indicated.

[18] Aldaz J.M. *On tau-smooth measure spaces without thick Lindelöf subsets*. Real Anal. Exchange. 1991–1992. V. 17, №1. P. 379–385. [*450*]

[19] Aldaz J.M. *Borel extensions of Baire measures*. Fund. Math. 1997. V. 154. P. 275–293. [*131*]

[20] Aldaz J.M., Render H. *Borel measure extensions of measures defined on sub-σ-algebras*. Adv. Math. 2000. V. 150. P. 233–263. [*166*]

[21] Aldous D.J. *Limit theorems for subsequences of arbitrarily-dependent sequences of random variables*. Z. Wahrscheinlichkeitstheorie verw. Geb. 1977. B. 40, №1. S. 59–82. [*464*]

[22] Aldous D.J. Exchangeability and related topics. Lecture Notes in Math. 1985. V. 1117. P. 1–198. [*409*]

[23]* 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).

[24]* 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).

[25] Aleksandrova D.E. *Convergence of triangular transformations of measures*. Theory Probab. Appl. 2005. V. 50, №1. P. 145–150. [*418, 424*]

[26]* Aleksandrova D.E., Bogachev V.I., Pilipenko A.Yu. *On the convergence of induced measures in variation*. Mat. Sbornik. 1999. V. 190, №9. P. 3–20 (in Russian); English transl.: Sb. Math. 1999. V. 190, №9–10. P. 1229–1245.

[27]* Aleksjuk V.N. Set functions. Leningrad. Gos. Pedagog. Inst., Leningrad, 1982; 78 pp. (in Russian)

[28]* Alexander R. *Random compact sets related to the Kakeya problem*. Proc. Amer. Math. Soc. 1975. V. 53, №2. P. 415–419.

[29] 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 [**32**]. [**I**: *431;* **II**: *452*]

[30] 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. [**I**: *422, 429, 431;* **II**: *64, 108, 113, 135, 179, 184, 250, 442, 443, 447, 451, 452, 454*]

[31]* 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.

[32] Alexandrov A.D. Selected works. Part 1: Selected scientific papers. Gordon and Breach, Amsterdam, 1996; x+322 pp. [**I**: *409;* **II**: *452*]

[33]* Alexandroff P. *Sur la puissance des ensembles mesurables B*. C. R. Acad. Sci. Paris. 1916. T. 162. P. 323–325.

[34]* 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).

[35] Alfsen E.M. Compact convex sets and boundary integrals. Springer-Verlag, New York – Heidelberg, 1971; 210 pp. [*146*]

[36]* Aliprantis Ch.D., Burkinshaw O. Principles of real analysis. 3d ed. Academic Press, San Diego, California, 1998; xii+415 pp.

[37]* Aliprantis Ch.D., Burkinshaw O. Problems in real analysis. A workbook with solutions. 2nd ed. Academic Press, San Diego, California, 1999; viii+403 pp.

[38] Alpern S., Prasad V.S. Typical dynamics of volume preserving homeomorphisms. Cambridge University Press, Cambridge, 2000; xx+216 pp. [*288, 459*]

[39]* Alt H.W. Lineare Funktionalanalysis: eine anwendungsorientierte Einführung. 3e Aufl. Springer, Berlin, 1999; xiii+415 S.

[40]* Amann H., Escher J. Analysis III. Birkhäuser, Basel – Boston – Berlin, 2001; 480 S.

[41] Ambrose W. *On measurable stochastic processes*. Trans. Amer. Math. Soc. 1940. V. 47. P. 66–79. [*448*]

[42] Ambrosio L. *Lecture notes on optimal transport problems*. Lecture Notes in Math. 2003. V. 1812. P. 1–52. [*236*]

[43] Ambrosio L. *Transport equation and Cauchy problem for BV vector fields*. Invent. Math. 2004. V. 158, №2. P. 227–260. [*460*]

[44]* 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.

[45] Ambrosio L., Gigli N., Savaré G. Gradient flows in metric spaces and in the space of probability measures. Birkhäuser, Basel, 2005; vii+333 pp. [*454, 460*]

[46] Amemiya I., Okada S., Okazaki Y. *Pre-Radon measures on topological spaces*. Kodai Math. J. 1978. V. 1, №1. P. 101–132. [*156, 443*]

[47]* 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.

[48] Andersen E.S., Jessen B. *Some limit theorems on integrals in an abstract set*. Danske Vid. Selsk. Math.-Fys. Medd. 1946. V. 22, №14. P. 1–29. [*461*]

[49] Andersen E.S., Jessen B. *On the introduction of measures in infinite product sets*. Danske Vid. Selsk. Math.-Fys. Medd. 1948. V. 25, №4. P. 1–7. [*461*]

[50] Andersen E.S., Jessen B. *Some limit theorems on set-functions*. Danske Vid. Selsk. Math.-Fys. Medd. 1948. V. 25, №5. P. 1–8. [*461*]

[51]* 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.

[52]* Anger B., Bauer H. Mehrdimensionale Integration. Eine Einführung in die Lebesguesche Theorie. Walter de Gruyter, Berlin – New York, 1976; 188 S.

[53] Anger B., Portenier C. Radon integrals. Birkhauser Boston, Boston, 1992; 332 pp. [**I**: *415;* **II**: *447, 451*]

[54] Aniszczyk B. *Two-valued measure need not be purely $\aleph_0$-compact*. Comment. Math. Univ. Carolin. 1982. V. 23, №1. P. 167–171. [*173*]

[55] Anosov D.V. *The additive functional homology equation that is connected with an ergodic rotation of the circles*. Izv. Akad. Nauk SSSR Ser. Mat. 1973. V. 37. P. 1259–1274 (in Russian); English transl.: Math. USSR Izv. 1973. V. 7. P. 1257–1271. [*335*]

[56]* Ansel J.-P., Ducel Y. Exercices corrigés en théorie de la mesure et de l'intégration. Ellipses, Paris, 1995; 126 pp.

[57]* 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).

[58]* 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).

[59]* Areshkin G.Ya. *On measure theory of B-sets*. Soobšč. Akad. Nauk Gruzin. SSR. 1949. V. 7, №8. P. 505–506 (in Russian).

[60]* 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).

[61]* 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).

[62]* 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).

[63]* 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).

[64]* 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.

[65] Argyros S. *On compact spaces without strictly positive measures.* Pacif. J. Math. 1983. V. 105, №2. P. 257–272. [*450*]

[66]* Arias de Reyna J. Pointwise convergence of Fourier series. Lecture Notes in Math. V. 1785. Springer-Verlag, Berlin, 2002; xviii+175 pp.

[67]* Arino O., Delode C., Genet J. Mesure et intégration. Exercices et problèmes avec solutions. Maîtrises de mathématiques. Vuibert, Paris, 1976; 239 pp.

[68] Arkhangel'skiĭ A.V., Ponomarev V.I. Fundamentals of general topology. Problems and exercises. Translated from the Russian. Reidel, Dordrecht, 1984; xvi+415 pp. [*9, 64*]

[69] Armstrong T. *Borel measures on compact groups are meager.* Illinois J. Math. 1981. V. 25, №4. P. 667–672. [*451*]

[70]* Arnaudies J.-M. L'intégrale de Lebesgue sur la droite. Vuibert, Paris, 2001; 320 pp.

[71] Arnold V.I., Avez A. Problèmes ergodiques de la mécanique classique. Gauthier-Villars, Paris, 1967; ii+243 pp. [*391*]

[72] Arsenin V.Ya., Lyapunov A.A. *The theory of A-sets.* Uspehi Mat. Nauk. 1950. V. 5. P. 45–108 (in Russian); German transl.: Ljapunow A.A., Stschegolkow E.A., Arsenin, W.J. Arbeiten zur deskriptiven Mengenlehre. Mathematische Forschungsberichte. 1. VEB Deutscher Verlag Wiss., Berlin, 1955; 108 S. [*37, 439, 441*]

[73]* Artémiadis N.K. Real analysis. Southern Illinois University Press, Carbondale; Feffer & Simons, London – Amsterdam, 1976; xii+581 pp.

[74]* Ascherl A., Lehn J. *Two principles for extending probability measures.* Manuscr. Math. 1977. V. 21. P. 43–50.

[75]* Ash R.B. Real analysis and probability. Academic Press, New York – London, 1972; xv+476 pp.

[76]* Ash R.B. Measure, integration, and functional analysis. Academic Press, New York – London, 1972; xiii+284 pp.

[77]* Asplund E., Bungart L. A first course in integration. Holt, Rinehart and Winston, New York – Toronto – London, 1966; xiii+489 pp.

[78]* Aumann G. Reelle Funktionen. Springer-Verlag, Berlin – Göttingen – Heidelberg, 1954; 2e Aufl., 1969; 418 S.

[79]* Aumann G., Haupt O. Einführung in die reelle Analysis. III. 3e Aufl. Walter de Gruyter, Berlin – New York, 1983; 297 S.

[80] Aumann R.J. *Measurable utility and measurable choice theorem.* Proc. Colloque Int. CNRS "La Décision" (Aix-en-Provence, 1967), pp. 15–26. CNRS, Paris, 1969. [*40*]

[81] Averna D. *Lusin type theorems for multifunctions, Scorza Dragoni's property and Carathéodory selections.* Boll. Un. Mat. Ital. A (7). 1994. V. 8, №2. P. 193–202. [*138*]

[82] Ayerbe-Toledano J.-M. *Category measures on Baire spaces.* Publ. Mat. 1990. V. 34, №2. P. 299–305. [*451*]

[83] Babiker A.G. *On almost discrete spaces.* Mathematika. 1971. V. 18. P. 163–167. [*136*]

[84] Babiker A.G. *On uniformly regular topological measure spaces.* Duke Math. J. 1976. V. 43, №4. P. 773–789. [*451*]

[85] Babiker A.G. *Lebesgue measures on topological spaces.* Mathematika. 1977. V. 24. P. 52–59. [*288*]

[86] Babiker A.G., Graf S. *Homomorphism-compact spaces.* Canad. J. Math. 1983. V. 35, №3. P. 558–576. [*450, 451*]

[87] Babiker A.G., Knowles J. *An example concerning completion regular measures, images of measurable sets and measurable selections.* Mathematika. 1978. V. 25. P. 120–124. [*334, 450*]

[88] Babiker A.G., Knowles J.D. *Functions and measures on product spaces.* Mathematika. 1985. V. 32, №1. P. 60–67. [*163*]

[89] Bachman G., Sultan A. *On regular extensions of measures.* Pacif. J. Math. 1980. V. 86, №2. P. 389–395. [*131, 451*]

[90] Bade W.G. The Banach space $C(S)$. Lecture Notes Series, №26. Matematisk Inst., Aarhus Universitet, Aarhus, 1971; ii+154 pp. [*456*]

[91] Badrikian A. Séminaire sur les fonctions aléatoire linéaires et les mesures cylindriques. Lecture Notes in Math. V. 139. Springer-Verlag, Berlin – New York, 1970; vii+221 pp. [*447*]
[92] Badrikian A., Chevet S. Mesures cylindriques, espaces de Wiener et fonctions aléatoires gaussiennes. Lecture Notes in Math. V. 379. Springer-Verlag, Berlin – New York, 1974; x+383 pp. [*447*]
[93] Baire R. Leçons sur les fonctions discontinues. Gauthier-Villars, Paris, 1904; viii+128 pp. [**I**: *148;* **II**: *12*]
[94] Baire R. Œuvres scientifiques. Gauthier-Villars, Paris, 1990. [**I**: *409;* **II**: *12*]
[95] Baker R. *"Lebesgue measure" on* $\mathbb{R}^{\infty}$. Proc. Amer. Math. Soc. 1991. V. 113. P. 1023–1029; *II*: ibid. 2004. V. 132, №9. P. 2577–2591. [*335*]
[96] Balder E.J. *On ws-convergence of product measures.* Math. Oper. Res. 2001. V. 26, №3. P. 494–518. [*249*]
[97]* Ball J.M., Murat F. *Remarks on Chacon's biting lemma.* Proc. Amer. Math. Soc. 1989. V. 107, №3. P. 655–663.
[98]* Banach S. *Sur le problème de la mesure.* Fund. Math. 1923. V. 4. P. 7–23.
[99]* Banach S. *Sur les lignes rectifiables et les surfaces dont l'aire est finie.* Fund. Math. 1925. V. 7. P. 225–236.
[100] Banach S. *Sur une classe de fonctions continues.* Fund. Math. 1926. V. 8. P. 225–237. [**I**: *392;* **II**: *458*]
[101]* Banach S. *Über additive Massfunktionen in abstrakten Mengen.* Fund. Math. 1930. V. 15. P. 97–101.
[102]* Banach S. Théorie des opérations linéaires. Warszawa, 1932; vii+254 pp.
[103] Banach S. *On Haar's measure.* Note I in [**1640**] (2nd ed.). P. 314–319. [*460*]
[104] Banach S. *The Lebesgue integral in abstract spaces.* Note II in [**1640**] (2nd ed.). P. 320–330. [**I**: *424;* **II**: *446*]
[105]* Banach S. *Sur les suites d'ensembles excluant l'existence d'une mesure.* Colloq. Math. 1947–1948. V. 1. P. 103–108.
[106] Banach S. *Sur la représentation des fonctions indépendantes à l'aide des fonctions de variables distinctes.* Colloq. Math. 1947–1948. V. 1. P. 109–121. [*464*]
[107] Banach S. *On measures in independent fields.* Studia Math. 1948. V. 10. P. 159–177. [*400, 464*]
[108] Banach S. Œuvres, V. I, II. PWN-Éditions Scientifiques de Pologne, Warszawa, 1967, 1979. [**I**: *409;* **II**: *440*]
[109]* Banach S., Kuratowski K. *Sur une généralisation du problème de la mesure.* Fund. Math. 1929. V. 14. P. 127–131.
[110]* Banach S., Saks S. *Sur les fonctions absolument continues des fonctions absolument continues.* Fund. Math. 1928. V. 11. P. 113–116.
[111]* Banach S., Saks S. *Sur la convergence forte dans les champs* L^p. Studia Math. 1930. V. 2. P. 51–57.
[112]* Banach S., Tarski A. *Sur la décomposition des ensembles de points en parties respectivement congruentes.* Fund. Math. 1924. V. 6. P. 244–277.
[113] Banakh T.O. *Topology of spaces of probability measures. I, II.* Matem. Studii (Lviv). 1995. №5. P. 65–87, 88–106 (in Russian). [*228, 455*]
[114] Banakh T.O., Bogachev V.I., Kolesnikov A.V. *Topological spaces with the strong Skorokhod property.* Georgian Math. J. 2001. V. 8, №2. P. 201–220. [*202, 225, 454, 456*]
[115] Banakh T.O., Bogachev V.I., Kolesnikov A.V. *On topological spaces with the Prokhorov and Skorokhod properties.* Dokl. Ros. Akad. Nauk. 2001. V. 380, №6. P. 727–730 (in Russian); English transl.: Dokl. Math. 2001. V. 64, №2. P. 244–247. [*202, 454, 456*]
[116] Banakh T.O., Bogachev V.I., Kolesnikov A.V. *Topological spaces with the strong Skorokhod property, II.* In: Functional Analysis and its Applications, V. Kadets and W. Zelazko eds. (Proceedings of the international conference on functional analysis

and its applications dedicated to the 110th anniversary of Stefan Banach, May 28–31, 2002, Lviv), pp. 23–47. Elsevier, Amsterdam, 2004. [*202, 454*]

[117] Banakh T.O., Bogachev V.I., Kolesnikov A.V. *Topological spaces with the Skorohod representation property*. Ukrainian Math. J. 2005. V. 57, №9. P. 1171–1186. [*454*]

[118] Banakh T.O., Cauty R. *Topological classification of spaces of probability measures on co-analytic sets*. Matem. Zametki. 1994. V. 55, №1. P. 9–19 (in Russian); English transl.: Math. Notes. 1994. V. 55, №1–2. P. 8–13. [*455*]

[119] Banakh T.O., Radul T.N. *Topology of spaces of probability measures*. Mat. Sbornik. 1997. V. 188, №7. P. 23–46 (in Russian); English transl.: Sbornik Math. 1997. V. 188, №7. P. 973–995. [*455*]

[120] Banakh T.O., Radul T.N. *Geometry of mappings of spaces of probability measures*. Matem. Studii (Lviv). 1999. V. 11, №1. P. 17–30 (in Russian). [*228, 455*]

[121]* Barner M., Flohr F. Analysis. II. 2e Aufl., Walter de Gruyter, Berlin, 1989; 449 S.

[122]* 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.

[123]* Barra G. de. Measure theory and integration. Ellis Horwood, Chichester; John Wiley & Sons, New York, 1981; 239 pp. (2nd ed.: 2003).

[124]* Bartle R.G. The elements of integration and Lebesgue measure. John Wiley & Sons, New York, 1995; xii+179 pp. (1st ed.: 1966).

[125]* Bartle R.G. A modern theory of integration. Amer. Math. Soc., Providence, Rhode Island, 2001; xiv+458 pp.

[126]* 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).

[127]* 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.

[128]* Bass J. Cours de mathématiques. T. III: Topologie, intégration, distributions, équations intégrales, analyse harmonique. Masson et Cie, Paris, 1971; 405 pp.

[129] Bass R.F. Probabilistic techniques in analysis. Springer, Berlin – New York, 1995; 392 pp. [*461*]

[130]* Basu A.K. Measure theory and probability. Prentice-Hall of India, New Delhi, 2003; 236 pp.

[131] Batt J. *Die Verallgemeinerungen des Darstellungssatzes von F. Riesz und ihre Anwendungen*. Jahresber. Deutsch. Math.-Verein. 1973. B. 74, №4. S. 147–181. [*447*]

[132] Bauer H. *Sur l'équivalence des théories de l'intégration selon N. Bourbaki et selon M.H. Stone*. Bull. Soc. Math. France. 1957. V. 85. P. 51–75. [*155*]

[133] Bauer H. *Konservative Abbildungen lokal-kompakter Räume*. Math. Ann. 1959. B. 138. S. 398–427. [*458*]

[134] Bauer H. *Mesures avec une image donnée*. Rev. Roumaine Math. Pures Appl. 1966. V. 11. P. 747–752. [*458*]

[135]* 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.;

[136] Bauer H. Probability theory. Walter de Gruyter, Berlin, 2001; xvi+230 pp.; (translation from the 4th German ed.) German 5th ed.: Wahrscheinlichkeitstheorie. Walter de Gruyter, Berlin, 2002; xix+520 S. [*410, 461*]

[137] Baushev A.N. *On the weak convergence of probability measures in a Banach space*. Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. 1999. V. 260. Veroyatn. i Statist. 3. P. 17–30 (in Russian); English transl.: J. Math. Sci. (New York). 2002. V. 109, №6. P. 2037–2046. [*456*]

[138]* Beals R. Analysis. An introduction. Cambridge University Press, Cambridge, 2004; x+261 pp.

[139]* Bear H.S. A primer of Lebesgue integration. Academic Press, San Diego, California, 1995; xii+163 pp.

[140] Beck A., Corson H.H., Simon A.B. *The interior points of the product of two subsets of a locally compact group.* Proc. Amer. Math. Soc. 1958. V. 9. P. 648–652. [*333*]

[141] Becker H., Kechris A. *Borel actions of Polish groups.* Bull. Amer. Math. Soc. 1993. V. 28, №2. P. 334–341. [*451*]

[142]* Behrends E. Maß- und Integrationstheorie. Springer-Verlag, Berlin – New York, 1987; 260 S.

[143]* Belkner H., Brehmer S. Lebesguesche Integrale. VEB Deutscher Verlag der Wissenschaften, Berlin, 1984; 151 S.

[144]* Bellach J., Franken P., Warmuth E., Warmuth W. Maß, Integral und bedingter Erwartungswert. Vieweg, Braunschweig, 1978; 145 S.

[145] Bellow A. *Lifting compact spaces.* Lecture Notes in Math. 1980. V. 794. P. 233–253. [*433*]

[146]* Benedetto J.J. Real variable and integration. With historical notes. B.G. Teubner, Stuttgart, 1976; 278 pp.

[147]* Benoist J., Salinier A. Exercices de calcul intégral: avec rappels de cours. Masson, Paris, 1997; ix+212 pp.

[148]* Berberian S.K. Measure theory and integration. Macmillan, New York, 1965; xviii+312 pp.

[149]* Berberian S.K. Fundamentals of real analysis. Springer-Verlag, New York, 1999; xii+479 pp.

[150] Berezanskiĭ I.A. *Measures on uniform spaces and molecular measures.* Trudy Moskov. Mat. Ob. 1968. V. 19. P. 3–40 (in Russian); English transl.: Trans. Moscow Math. Soc. 1968. V. 19. P. 1–40. [*451*]

[151]* 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).

[152] Berg C., Christensen J.P.R., Ressel P. Harmonic analysis on semigroups. Springer, Berlin – New York, 1984; 289 pp. [*451*]

[153]* Bergh J., Löfström J. Interpolation spaces. An introduction. Springer-Verlag, Berlin – New York, 1976; x+207 pp.

[154] Bergin J. *On the continuity of correspondences on sets of measures with restricted marginals.* Econom. Theory. 1999. V. 13, №2. P. 471–481. [*266*]

[155] Bergström H. Limit theorems for convolutions. Almqvist & Wiksell, Stockholm – Göteborg – Uppsala; John Wiley & Sons, New York – London, 1963; 347 pp. [*453*]

[156] Bergström H. Weak convergence of measures. Academic Press, New York – London, 1982; 245 pp. [*453*]

[157] Berkes I. *An extension of the Komlós subsequence theorem.* Acta Math. Hungar. 1990. V. 55. P. 103–110. [*415*]

[158] Berkes I., Péter E. *Exchangeable random variables and the subsequence principle.* Probab. Theory Related Fields. 1986. V. 73, №3. P. 395–413. [*464*]

[159] Berliocchi H., Lasry J.-M. *Intégrandes normales et mesures paramétrées en calcul des variations.* Bull. Soc. Math. France. 1973. V. 101. P. 129–184. [*137*]

[160]* Bertin E.M.J., Cuculescu I., Theodorescu R. Unimodality of probability measures. Kluwer Academic Publ., Dordrecht, 1997; xiv+251 pp.

[161]* Besicovitch A.S. *On Kakeya's problem and a similar one.* Math. Z. 1928. B. 27. S. 312–320.

[162]* Besicovitch A.S. *On a problem concerning Lebesgue integrals.* Math. Ann. 1938. B. 115. S. 613–618.

[163]* 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.

[164]* 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).

[165]* 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.

[166] Bichteler K. Integration – a functional approach. Birkhäuser, Boston – Berlin, 1998; viii+193 pp. [**I**: *413;* **II**: *446*]

[167]* Bierlein D. *Über die Fortsetzung von Wahrscheinlichkeitsfeldern.* Z. Wahrscheinlichkeitstheorie verw. Geb. 1962. B. 1. S. 28–46.

[168] Billingsley P. Ergodic theory and information. John Wiley & Sons, New York – London – Sydney, 1965; xiii+195 pp. [*391, 431*]

[169] Billingsley P. Convergence of probability measures. Wiley, New York, 1968; xii+253 pp. [*53, 453*]

[170]* Billingsley P. Probability and measure. 3d ed. Wiley, New York, 1995; 593 pp.

[171] Billingsley P., Topsøe F. *Uniformity in weak convergence.* Z. Wahrscheinlichkeitstheorie verw. Geb. 1967. B. 7. S. 1–16. [*456*]

[172]* 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.

[173]* Bingham N.H. *Doob: a half-century on.* J. Appl. Probab. 2005. V. 42, №1. P. 257–266.

[174] Birkhoff G.D. Dynamical systems. New York, 1927. Rev. ed.: Amer. Math. Soc., Providence, 1966; xii+305 pp. [*458*]

[175] Birkhoff G.D. *Proof of the ergodic theorem.* Proc. Nat. Acad. Sci. U.S.A. 1931. V. 17. P. 656–660. [*463*]

[176]* Birkhoff G. Lattice theory. Corrected reprint of the 1967 third edition. Amer. Math. Soc., Providence, Rhode Island, 1979; vi+418 pp.

[177] Bishop E. *An upcrossing inequality with applications.* Michigan Math. J. 1966. V. 13. P. 1–13. [*463*]

[178]* Bishop E. Foundations of constructive analysis. McGraw-Hill Book Co., New York – Toronto, 1967; xiii+370 pp.

[179] Blackwell D. *On a theorem of Lyapunov.* Ann. Math. Statist. 1951. V. 22. P. 112–114. [*338*]

[180] Blackwell D.H. *On a class of probability spaces.* Proc. of the Third Berkeley Sympos. on Math. Statist. and Probab. (Berkeley, 1954/55). P. 1–6. Univ. California Press, Berkeley, 1956. [*50*]

[181] Blackwell D., Dubins L. *Merging of opinions with increasing information.* Ann. Math. Statist. 1962. V. 33. P. 882–886. [*370*]

[182] Blackwell D., Dubins L.E. *A converse to the dominated convergence theorem.* Illinois J. Math. 1963. V. 7, №3. P. 508–514. [*428*]

[183] Blackwell D., Dubins L.E. *On existence and nonexistence of proper, regular, conditional distributions.* Ann. Probab. 1975. V. 3, №5. P. 741–752. [*462*]

[184] Blackwell D., Dubins L.E. *An extension of Skorohod's almost sure representation theorem.* Proc. Amer. Math. Soc. 1983. V. 89, №4. P. 691–692. [*199, 454*]

[185] Blackwell D., Maitra A. *Factorization of probability measures and absolutely measurable sets.* Proc. Amer. Math. Soc. 1984. V. 92, №2. P. 251–254. [*462*]

[186] Blackwell D., Ryll-Nardzewski C. *Non-existence of everywhere proper conditional distributions.* Ann. Math. Statist. 1963. V. 34. P. 223–225. [*429, 462*]

[187] Blau J.H. *The space of measures on a given set.* Fund. Math. 1951. V. 38. P. 23–34. [*453*]

[188] Bledsoe W.W., Morse A.P. *Product measures.* Trans. Amer. Math. Soc. 1955. V. 79. P. 173–215. [*444*]

[189] Bledsoe W.W., Wilks C.E. *On Borel product measures*. Pacific J. Math. 1972. V. 42. P. 569–579. [*444*]

[190]* Bliss G.A. *Integrals of Lebesgue*. Bull. Amer. Math. Soc. 1917. V. 24. P. 1–47.

[191] Bloom W.R., Heyer H. Harmonic analysis of probability measures on hypergroups. Walter de Gruyter, Berlin, 1995; vi+601 pp. [*451*]

[192]* Blumberg H. *The measurable boundaries of an arbitrary function*. Acta Math. 1935. V. 65. P. 263–282.

[193] Bobkov S.G. *Isoperimetric and analytic inequalities for log-concave probability measures*. Ann. Probab. 1999. V. 27, №4. P. 1903–1921. [**I**: *431;* **II**: *451*]

[194] Bobkov S.G. *Remarks on the growth of L^p-norms of polynomials*. Lecture Notes in Math. 2000. V. 1745. P. 27–35. [*150*]

[195]* 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.

[196]* 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.

[197]* 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).

[198]* Boccara N. Functional analysis. An introduction for physicists. Academic Press, Boston, 1990; xiv+327 pp.

[199]* Boccara N. Intégration. Ellipses, Paris, 1995; 120 pp.

[200]* 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.

[201]* Bochner S. *Monotone Funktionen, Stieltjessche Integrale und harmonische Analyse*. Math. Ann. 1933. B. 108. S. 378–410 (see [**200**]).

[202] Bochner S. Harmonic analysis and the theory of probability. University of California Press, Berkeley, 1955; 176 pp. [*447*]

[203] Bochner S., Neumann J. von. *On compact solutions of operational-differential equations. I*. Ann. Math. 1935. V. 36, №1. P. 255–291. [*457*]

[204] Bogachev V.I. *Indices of asymmetry of stable measures*. Matem. Zametki. 1986. V. 40, №1. P. 127–138 (in Russian); English transl.: Math. Notes. 1986. V. 40. P. 569–575. [*451*]

[205] Bogachev V.I. *Locally convex spaces with the CLT property and supports of measures*. Vestnik Moskovsk. Univ. 1986. №6. P. 16–20 (in Russian); English transl.: Moscow Univ. Math. Bull. 1986. V. 41, №6. P. 19–23. [*448*]

[206] Bogachev V.I. *Differentiable measures and the Malliavin calculus*. J. Math. Sci. (New York), 1997. V. 87, №5. P. 3577–3731. [*305, 448*]

[207] Bogachev V.I. *Measures on topological spaces*. J. Math. Sci. (New York). 1998. V. 91, №4. P. 3033–3156. [*53, 443*]

[208] Bogachev V.I. Gaussian measures, Amer. Math. Soc., Providence, Rhode Island, 1998; 433 pp. [**I**: *198, 431;* **II**: *98, 142, 144, 410, 433, 451, 452, 460*]

[209] Bogachev V.I. *Average approximations and moments of measures*. J. Complexity. 2000. V. 16. P. 390–410. [*170*]

[210]* 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).

[211] Bogachev V.I., Kolesnikov A.V. *Open mappings of probability measures and the Skorohod representation theorem*. Teor. Verojatn. i Primen. 2001. V. 46, №1. P. 3–27 (in Russian); English transl.: Theory Probab. Appl. 2001. V. 46, №1. P. 20–38. [*199, 202, 225, 228, 229, 454*]

[212] Bogachev V.I., Kolesnikov A.V. *Nonlinear transformations of convex measures and entropy of the Radon–Nikodym densities*. Russian Math. Dokl. 2004. V. 397, №2. P. 155–159 (in Russian); English transl.: Dokl. Math. 2004. V. 70, №1. P. 524–528. [*418*]

[213] Bogachev V.I., Kolesnikov A.V. *Nonlinear transformations of convex measures.* Teor. Verojatn. i Primen. 2005. V. 50, №1. P. 27–51. (in Russian); English transl.: Theory Probab. Appl. 2005. V. 50, №1. [*418, 427, 451*]

[214] Bogachev V.I., Kolesnikov A.V. *Integrability of absolutely continuous transformations of measures and applications to optimal mass transportation.* Theory Probab. Appl. 2005. V. 50, №3. P. 433–456 (in Russian); English transl.: Theory Probab. Appl. 2005. V. 50, №3. [*236, 451, 454*]

[215]* 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.

[216] 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. 2005. №10(45). P. 300–307 (in Russian). [**I**: *420;* **II**: *439*]

[217] Bogachev V.I., Kolesnikov A.V., Medvedev K.V. *On triangular transformations of measures.* Russian Math. Dokl. 2004. V. 396, №6. P. 727–732 (in Russian); English transl.: Dokl. Math. 2004. V. 69, №3. P. 438–442. [*418, 426*]

[218] Bogachev V.I., Kolesnikov A.V., Medvedev K.V. *Triangular transformations of measures.* Matem. Sbornik. 2005. V. 196, №3. P. 3–30 (in Russian); English transl.: Sbornik Math. 2005. V. 196, №3. [*418, 426, 464*]

[219] Bogachev V.I., Korolev A.V. *On the ergodic theorem in the Kozlov–Treshchev form.* Doklady Math. 2006. [*396*]

[220] Bogachev V.I., Mayer-Wolf E. *Absolutely continuous flows generated by Sobolev class vector fields in finite and infinite dimensions.* J. Funct. Anal. 1999. V. 167, №1. P. 1–68. [*301, 302, 460*]

[221] Bogachev V.I., Prostov Yu.I. *A polynomial diffeomorphism of a ball without invariant measures.* Funkt. Anal. i Prilozh. 1989. V. 23, №4. P. 75–76 (in Russian); English transl.: Funct. Anal. Appl. 1989. V. 23, №4. P. 317–319. [*319*]

[222] Bogachev V.I., Pugachev O.V., Röckner M. *Surface measures and tightness of (r,p)-capacities on Poisson space.* J. Funct. Anal. 2002. V. 196. P. 61–86. [*457*]

[223] Bogachev V.I., Röckner M. *Regularity of invariant measures on finite and infinite dimensional spaces and applications.* J. Funct. Anal. 1995. V. 133. P. 168–223. [*433*]

[224] Bogachev V.I., Sadovnichiĭ Yu.V., Fedorchuk V.V. *On the uniqueness of a preimage of a measure.* Dokl. Russian Acad. Sci. 2006. V. 408, №2 (in Russian); English transl.: Dokl. Math. 2006. V. 73, №3. P. 344–348. [*311*]

[225] Bogachev V.I., Smolyanov O.G. *Analytic properties of infinite dimensional distributions.* Uspehi Mat. Nauk. 1990. V. 45, №3. P. 3–83 (in Russian); English transl.: Russian Math. Surveys. 1990. V. 45, №3. P. 1–104. [*167, 448*]

[226] Böge W., Krickeberg K., Papangelou F. *Über die dem Lebesgueschen Maß isomorphen topologischen Maßräume.* Manuscr. Math. 1969. V. 1. P. 59–77. [*323*]

[227] Bogoliouboff N.N., Kryloff N.M. *La théorie générale de la mesure dans son application à l'étude de systèmes dynamiques de la mécanique non-linéaire.* Ann. Math. 1937. B. 38. S. 65–113. [*318, 442, 452, 458*]

[228]* Bogoljubov A.N. Mathematicians. Mechanicists. (Biographical reference book). Naukova Dumka, Kiev, 1983 (in Russian).

[229]* Boman J. *An example of nonuniqueness for a generalized Radon transform.* J. Anal. Math. 1993. V. 61. P. 395–401.

[230] Borel E. Leçons sur la théorie des fonctions. Gauthier-Villars, Paris, 1898 (4e éd.: 1950; xii+295 pp.) [**I**: *417;* **II**: *10*]

[231]* Borel E. *Un théorème sur les ensembles mesurables.* C. R. Acad. Sci. Paris. 1903. T. 137. P. 966–967.

[232]* Borel E. *Les probabilités dénombrables et leurs applications arithmétiques.* Rend. Circ. Mat. Palermo. 1909. T. 27. P. 247–271.

[233] Borel E. Introduction géométrique à quelques théories physiques. Gauthier-Villars, Paris, 1914; vii+137 pp. [*254*]

[234] Borel E. Œuvres de Émile Borel. T. I–IV. Éditions du Centre National de la Recherche Scientifique, Paris, 1972. [**I**: *409;* **II**: *10*]

[235]* 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.

[236] Borell C. *Convex measures on locally convex spaces.* Ark. Math. 1974. V. 12, №2. P. 239–252. [**I**: *226;* **II**: *150, 434, 451*]

[237]* Borell C. *Convex set functions in d-space.* Period. Math. Hungar. 1975. V. 6, №2. P. 111–136.

[238] Borell C. *A note on conditional probabilities of a convex measure.* Lecture Notes in Math. 1978. V. 644. P. 68–72. [*150, 451*]

[239] Borell C. *Convexity of measures in certain convex cones in vector space σ-algebras.* Math. Scand. 1983. V. 53, №1. P. 125–144. [*451*]

[240] Borovkov A.A. *Convergence of measures and random processes.* Uspehi Mat. Nauk. 1976. V. 31, №2. P. 3–68 (in Russian); English transl.: Russian Math. Surveys. 1976. V. 31, №2. P. 1–69. [*456*]

[241]* Borovkov A.A. Probability theory. Translated from the Russian. Gordon and Breach, Amsterdam, 1998; x+474 pp. (Russian ed.: Moscow, 1986).

[242] 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). [**I**: *412;* **II**: *59, 125, 172, 442, 443, 447, 450, 452, 458, 459, 460*]

[243]* 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.

[244] Bourgain J. *Double recurrence and almost sure convergence.* J. Reine Angew. Math. 1990. B. 404. S. 140–161. [*397*]

[245]* Bouyssel M. Intégrale de Lebesgue. Mesure et intégration. Exercices avec solutions et rappels de cours. Cépaduès-éditions, Toulouse, 1997; 383 pp.

[246] Bouziad A. *Coincidence of the upper Kuratowski topology with the co-compact topology on compact sets, and the Prohorov property.* Topology Appl. 2002. V. 120. P. 283-299. [*225*]

[247] Bouziad A. *Luzin measurability of Carathéodory type mappings.* Topology Appl. 2006. [*138*]

[248]* Bouziad A., Calbrix J. Théorie de la mesure et de l'intégration. Publ. de l'Univ. de Rouen, Mont–Saint-Aignan, 1993; 254 pp.

[249]* 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.

[250] Bray H.E. *Elementary properties of the Stieltjes integral.* Ann. Math. (2). 1918–1919. V. 20. P. 177–186. [*452*]

[251]* Brehmer S. Einführung in die Maßtheorie. Akademie-Verlag, Berlin, 1975; 186 S.

[252] Brenier Y. *Polar factorization and monotone rearrangement of vector–valued functions.* Comm. Pure Appl. Math. 1991. V. 44. P. 375–417. [**I**: *382;* **II**: *236*]

[253] Bressler D.W., Sion M. *The current theory of analytic sets.* Canad. J. Math. 1964. V. 16. P. 207–230. [*440*]

[254]* Brézis H. *How to recognize constant functions. Connections with Sobolev spaces.* Russian Math. Surveys. 2002. V. 54, №4. P. 693–708.

[255]* 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.

[256]* Briane M., Pagès G. Théorie de l'intégration. 2e éd. Vuibert, Paris, 2000; 304 pp.

[257]* Bridges D.S. Foundations of real and abstract analysis. Springer-Verlag, New York, 1998; xiv+322 pp.

[258]* 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).

[259]* Brooks J.K., Chacon R.V. *Continuity and compactness of measures.* Adv. Math. 1980. V. 37, №1. P. 16–26.

[260]* Broughton A., Huff B.W. *A comment on unions of sigma-fields.* Amer. Math. Monthly. 1977. V. 84, №7. P. 553–554.

[261] Brow J.B., Cox G.V. *Baire category in spaces of probability measures II.* Fund. Math. 1984. V. 121, №2. P. 143–148. [*455*]

[262]* Browder A. Mathematical analysis. An introduction. Springer-Verlag, New York, 1996; xiv+333 pp.

[263]* 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.

[264]* 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.)

[265]* Bruckner A.M., Bruckner J.B., Thomson B.S. Real analysis. Prentice-Hall, 1997; 713 pp.

[266]* Brudno A.L. The theory of functions of a real variable. Nauka, Moscow, 1971; 120 pp. (in Russian).

[267] Bruijn N.G. de, Post K.A. *A remark on uniformly distributed sequences and Riemann integrability.* Indag. Math. 1968. V. 30. P. 149–150. [*257*]

[268] Bryc W., Kwapień S. *On the conditional expectation with respect to a sequence of independent σ-fields.* Z. Wahrscheinlichkeitstheorie verw. Geb. 1978–1979. B. 46, №2. S. 221–225. [*433*]

[269]* Brzuchowski J., Cichon J., Grzegorek E., Ryll-Nardzewski C. *On the existence of nonmeasurable unions.* Bull. Acad. Polon. Sci. 1979. V. 27, №6. P. 447–448.

[270]* Buchwalter H. Le calcul intégral en licence de mathématiques. Ellipses, Paris, 1991; iv+145 pp.

[271] Buczolich Z. *Birkhoff sums and zero–one law.* Real Anal. Exchange. Denton Symp. 2000. P. 13–21. [**I**: *172;* **II**: *410*]

[272]* Bukovský L. *Any partition into Lebesgue measure zero sets produces a nonmeasurable set.* Bull. Acad. Polon. Sci. 1979. V. 27, №6. P. 431–435.

[273] Buldygin V.V. Convergence of random elements in topological spaces. Naukova Dumka, Kiev, 1980 (in Russian). [*448*]

[274] Buldygin V.V. *Supports of probability measures in separable Banach spaces.* Teor. Verojatn. i Primen. 1984. V. 29, №3. P. 528–532 (in Russian); English transl.: Theory Probab. Appl. 1984. V. 29. P. 546–549. [*448*]

[275]* 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).

[276] Buldygin V.V., Solntsev S.A. Asymptotic behaviour of linearly transformed sums of random variables. Kluwer Academic Publ., Dordrecht, 1997; xvi+500 pp. [*448*]

[277]* Burago D.M., Zalgaller V.A. Geometric inequalities. Translated from the Russian. Springer-Verlag, Berlin, 1988; xiv+331 pp. (Russian ed.: Leningrad, 1980).

[278]* 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.

[279]* 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.

[280] Burgess J.P. *A measurable selection theorem.* Fund. Math. 1980. V. 110, №2. P. 91–100. [*43*]

[281] Burgess J.P. *Sélections mesurables pour relations d'équivalence analytiques à classes* G_δ. Bull. Sci. Math. (2). 1980. V. 104, №4. P. 435–440. [*43*]
[282] Burgess J.P. *Careful choices – a last word on Borel selectors.* Notre Dame J. Formal Logic. 1981. V. 22, №3. P. 219–226. [*37*]
[283] Burgess J.P., Mauldin R.D. *Conditional distributions and orthogonal measures.* Ann. Probab. 1981. V. 9, №5. P. 902–906. [*463*]
[284]* Burk F. Lebesgue measure and integration. An introduction. John Wiley & Sons, New York, 1998; xvi+292 pp.
[285] Burke D.K., Pol R. *Note on separate continuity and the Namioka property.* Topology Appl. 2005. V. 152, P. 258–268. [*129*]
[286] Burke M.R. *Liftings for Lebesgue measure.* Set theory of the reals (Ramat Gan, 1991), pp. 119–150. Israel Math. Conf. Proc. 6, Bar–Ilan Univ., Ramat Gan, 1993. [*463*]
[287] Burke M.R. *Liftings and the property of Baire in locally compact groups.* Proc. Amer. Math. Soc. 1993. V. 117, №4. P. 1075–1082. [*463*]
[288] Burke M.R., Fremlin D. *A note on measurability and almost continuity.* Proc. Amer. Math. Soc. 1988. V. 102, №3. P. 611–612. [*137*]
[289] Burkholder D.L. *Successive conditional expectations of an integrable function.* Ann. Math. Statist. 1962. V. 33. P. 887–893. [*435*]
[290]* Burkill J.C. *Functions of intervals.* Proc. London Math. Soc. 1924. V. 22. P. 275–310.
[291]* Burkill J.C. The Lebesgue integral. Cambridge University Press, Cambridge, 1951; viii+87 pp.
[292]* Burkill J.C. *Charles-Joseph de la Vallée Poussin.* J. London Math. Soc. 1964. V. 39. P. 165–175.
[293]* Burrill C.W. Measure, integration, and probability. McGraw-Hill Book Co., New York, 1972; xiv+464 pp.
[294]* Burrill C.W., Knudsen J.R. Real variables. Holt, Rinehart and Winston, New York – Montreal – London, 1969; xii+419 pp.
[295]* Burstin C. *Über eine spezielle Klasse reeller periodischer Funktionen.* Monatsh. Math. Phys. 1915. B. 26. S. 229–262.
[296]* Buseman H., Feller W. *Zur Differentiation des Lebesgueschen Integrale.* Fund. Math. 1934. V. 22. P. 226–256.
[297]* Busemann H., Petty C.M. *Problems on convex bodies.* Math. Scand. 1956. V. 4. P. 88–94.
[298]* Caccioppoli R. *Sull'integrazione delle funzioni discontinue.* Rend. Circ. Mat. Palermo. 1928. T. 52. P. 1–29.
[299]* Caccioppoli R. *Integrali impropri di Stieltjes. Estensione del teorema di Vitali.* Rend. Acc. Sc. Fis. Mat. Napoli (3). 1929. T. 35. P. 34–40.
[300] Caffarelli L. *Monotonicity properties of optimal transportations and the FKG and related inequalities.* Comm. Math. Phys. 2000. V. 214, №3. P. 547–563. [**I**: *382;* **II**: *236*]
[301]* Cafiero F. Misura e integrazione. Edizioni Cremonese, Roma, 1959; 451 pp.
[302] Calbrix J. *Mesures non σ-finies: désintégration et quelques autres proprietés.* Ann. Inst. H. Poincaré, Sect. B. 1981. V. 17, №1. P. 75–95. [*462*]
[303]* Calderón A.P., Zygmund A. *On the existence of certain singular integrals.* Acta Math. 1952. V. 88. P. 85–139.
[304]* Cantelli F.P. *La tendenza ad un limite nel senso del calcolo delle probabilità.* Rend. Circ. Mat. Palermo. 1916. T. 41. P. 191–201.
[305]* Capiński M., Kopp E. Measure, integral and probability. 2nd ed. Springer-Verlag London, London, 2004; xvi+311 pp.
[306]* Capiński M., Zastawniak T. Probability through problems. Springer-Verlag, New York, 2001; viii+257 pp.

[307]* Carathéodory C. *Über das lineare Maß von Punktmengen – eine Verallgemeinerung des Längenbegriffs.* Nachr. Ges. Wiss. Göttingen, Math.-phys. Kl. 1914. S. 404–426.

[308] Carathéodory C. Vorlesungen über reelle Funktionen. B.G. Teubner, Leipzig – Berlin, 1918; 704 S. (2e Aufl. 1927; 718 S.) [**I**: *100, 410, 418, 419, 420;* **II**: *164*]

[309] Carathéodory C. *Über den Wiederkehrsatz von Poincaré.* Sitz. Preuß. Akad. Wiss., Math.-phys. Kl. 1919. S. 580–584. [*463*]

[310]* 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.

[311] Carathéodory C. Gesammelte mathematische Schriften. C.H. Beck, München, 1954. [**I**: *409;* **II**: *463*]

[312]* 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.

[313]* Carlson T. *Extending Lebesgue measure by infinitely many sets.* Pacif. J. Math. 1984. V. 115, №1. P. 33–45.

[314]* Carothers N.L. Real analysis. Cambridge University Press, Cambridge, 2000; xiv+401 pp.

[315] Cartan H. *Sur la mesure de Haar.* C. R. Acad. Sci. Paris. 1940. T. 211. P. 759–762. [*460*]

[316]* Carter M., Brunt B. van. The Lebesgue–Stieltjes integral. A practical introduction. Springer-Verlag, New York, 2000; x+228 pp.

[317] Castaing C. *Une nouvelle extension du théorème de Dragoni–Scorza.* C. R. Acad. Sci. Paris, sér. A-B. 1970. T. 271. P. A396–A398. [*137*]

[318] Castaing C., Raynaud de Fitte P., Valadier M. Young measures on topological spaces. With applications in control theory and probability theory. Kluwer Acad. Publ., Dordrecht, 2004; xi+320 pp. [*231, 249*]

[319] Castaing C., Valadier M. Convex analysis and measurable multifunctions. Lecture Notes in Math. V. 580. Springer-Verlag, Berlin – New York, 1977; 278 pp. [*39, 441*]

[320] Casteren J.A. van. *Strictly positive Radon measures.* J. London Math. Soc. (2). 1994. V. 49, №1. P. 109–123. [*450*]

[321] Cenzer D., Mauldin R.R *Borel equivalence and isomorphism of coanalytic sets.* Dissert. Math. 1984. V. 228. 28 pp. [*440*]

[322]* Chae S.B. Lebesgue integration. 2nd ed. Springer-Verlag, New York, 1995; xiv+264 pp.

[323]* Chandrasekharan K. A course on integration theory. Hindustan Book Agency, New Delhi, 1996; viii+118 pp.

[324] Chatterji S.D. *A general strong law.* Invent. Math. 1969–1970. V. 9. P. 235–245. [*464*]

[325] Chatterji S.D. *Disintegration of measures and lifting.* In: Vector and operator valued measures and applications (Proc. Sympos., Snowbird Resort, Alta, Utah, 1972), pp. 69–83. Academic Press, New York, 1973. [*462*]

[326] Chatterji S.D. *Les martingales et leurs applications analytiques.* Lecture Notes in Math. 1973. V. 307. P. 27–164. [*461*]

[327] Chatterji S.D. *A principle of subsequences in probability theory: the central limit theorem.* Adv. Math. 1974. V. 13. P. 31–54; correction, ibid. 1974. V. 14. P. 266–269. [*464*]

[328] Chatterji S.D. *A subsequence principle in probability theory.* Jahresber. Deutsch. Math.-Verein. 1985. B. 87, №3. S. 91–107. [*464*]

[329] Chatterji S.D. *Martingale theory: an analytical formulation with some applications in analysis.* Lecture Notes in Math. 1986. V. 1206. P. 109–166. [*461*]

[330] Chaumont L., Yor M. Exercises in probability. A guided tour from measure theory to random processes, via conditioning. Cambridge University Press, Cambridge, 2003; xv+236 pp. [*464*]

[331]* Chavel I. Isoperimetric inequalities. Differential geometric and analytic perspectives. Cambridge University Press, Cambridge, 2001; xii+268 pp.

[332]* 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).

[333]* Cheney W. Analysis for applied mathematics. Springer-Verlag, New York, 2001; viii+444 pp.

[334]* 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).

[335] Chentsov (Čencov) N.N. *Topological measures and the theory of random functions*. Trudy V All Union meeting on probability theory and mathematical statistics (Erevan, 1958), Erevan, Akad. Nauk ArmSSR, 1960, pp. 83–87 (see [**338**]). [*59, 172, 441, 448*]

[336] Chentsov (Čencov) N.N. Statistical decision rules and optimal inference. Translated from the Russian. Amer. Math. Soc., Providence, Rhode Island, 1982; viii+499 pp. (Russian ed.: Moscow, 1972). [*441, 448*]

[337] Chentsov (Čencov) N.N. *Nonmeasurable subsets of the Tychonoff cube*. Problems of mathematical physics and numerical mathematics. Nauka, Moscow, 1977, pp. 316–320 (in Russian; see [**338**]). [*172, 441, 448*]

[338] Chentsov (Čencov) N.N. Selected works: Mathematics. Fizmatlit, Moscow, 2001; 400 pp. (in Russian). [*441, 448*]

[339] Chevet S. *Quelques nouveaux résultats sur les mesures cylindriques*. Lecture Notes in Math. 1978. V. 644. P. 125–158. [*447, 451*]

[340] Chevet S. *Kernel associated with a cylindrical measure*. Lecture Notes in Math. 1980. V. 860. P. 51–84. [*451*]

[341] Choban M.M. *Descriptive set theory and topology*. General topology, III, pp. 157–219. Encyclopaedia Math. Sci., 51, Springer, Berlin, 1995. [*440*]

[342] Choban M.M. *Spaces, mappings and compact subsets*. Bul. Acad. Sti. Rep. Moldova. 2001. №2(36). P. 3–52. [*225, 454, 456*]

[343] Choksi J.R. *Inverse limits of measure spaces*. Proc. London Math. Soc. (3). 1958. V. 8. P. 321–342. [*460*]

[344] Choksi J.R. *On compact contents*. J. London Math. Soc. 1958. V. 33. P. 387–398. [*443*]

[345] Choksi J.R. *Automorphisms of Baire measures on generalized cubes. II*. Z. Wahrscheinlichkeitstheorie verw. Geb. 1972. B. 23, №2. S. 97–102. [*320*]

[346] Choksi J.R. *Measurable transformations on compact groups*. Trans. Amer. Math. Soc. 1973. V. 184. P. 101–124. [*320*]

[347] Choksi J.R., Fremlin D.H. *Completion regular measures on product spaces*. Math. Ann. 1979. B. 241, №2. S. 113–128. [*320*]

[348]* 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.

[349] Choquet G. *Theory of capacities*. Ann. Inst. Fourier (Grenoble). 1955. V. 5. P. 131–295. [*142, 442, 450*]

[350] Choquet G. *Ensembles $\mathcal{K}$-analytiques et $\mathcal{K}$-sousliniens. Cas général et cas métrique*. Ann. Inst. Fourier (Grenoble). 1959. V. 9. P. 75–81. [*440, 450*]

[351] Choquet G. *Forme abstraite du théorème de capacitabilité*. Ann. Inst. Fourier (Grenoble). 1959. V. 9. P. 83–89. [*450*]

[352]* Choquet G. Lectures on analysis. V. I: Integration and topological vector spaces. W.A. Benjamin, New York – Amsterdam, 1969; xx+360+xxi pp.

[353] Choquet G. *Sur les ensembles uniformément négligéables*. Séminaire Choquet, 9e année. 1970. №6. 15 pp. [*224, 255, 261*]

[354]* Chow Y.S., Teicher H. Probability theory. Independence, interchangeability, martingales. 3d ed. Springer-Verlag, New York, 1997; xxii+488 pp.

[355] Christensen J.P.R. Topology and Borel structure. North-Holland, Amsterdam – London, Amer. Elsevier, New York, 1974; 133 pp. [*168, 441*]

[356]* Ciesielski K. *Isometrically invariant extensions of Lebesgue measure.* Proc. Amer. Math. Soc. 1990. V. 110. P. 799–801.

[357]* Ciesielski K., Fejzić H., Freiling C. *Measure zero sets with non-measurable sum.* Real Anal. Exchange. 2001–2002. V. 27, №2. P. 783–793.

[358]* Ciesielski K., Pelc A. *Extensions of invariant measures on Euclidean spaces.* Fund. Math. 1985. V. 125. P. 1–10.

[359]* Clarkson J.A. *Uniformly convex spaces.* Trans. Amer. Math. Soc. 1936. V. 40, №3. P. 396–414.

[360] Cohn D.L. *Measurable choice of limit points and the existence of separable and measurable processes.* Z. Wahrscheinlichkeitstheorie verw. Geb. 1972. B. 22, №2. S. 161–165. [*463*]

[361] Cohn D.L. *Liftings and the construction of stochastic processes.* Trans. Amer. Math. Soc. 1978. V. 246. P. 429–438. [*463*]

[362]* Cohn D.L. Measure theory. Birhäuser, Boston – Basel – Stuttgart, 1980; ix+373 pp.

[363]* Coifman R.R., Feffermann C. *Weighted norm inequalities for maximal functions and singular integrals.* Studia Math. 1974. V. 51. P. 241–250.

[364] Collins H.S. *Strict topologies in measure theory.* Proc. Conf. on Integration, Topology, and Geometry in Linear Spaces (Univ. North Carolina, Chapel Hill, N.C., 1979), pp. 1–13. Contemp. Math., 2, Amer. Math. Soc., Providence, Rhode Island, 1980. [*447*]

[365] Comfort W., Negrepontis S. Continuous pseudometrics. Marcel Dekker, New York, 1975; 126 pp. [*44*]

[366] Comfort W.W., Negrepontis S. Chain condition in topology. Cambridge University Press, Cambridge, 1982; xiii+300 pp. [*450*]

[367] Constantinescu C. Duality in measure theory. Lecture Notes in Math. V. 796. Springer-Verlag, Berlin – New York, 1980; 197 pp. [*455*]

[368] Constantinescu C. *Spaces of measures on topological spaces.* Hokkaido Math. J. 1981. V. 10. P. 89–156. [*455*]

[369] Constantinescu C. Spaces of measures. de Gruyter, Berlin – New York, 1984; 444 pp. [*455*]

[370] Constantinescu C. Some properties of spaces of measures. Suppl. V. 35 Degli Atti Del Semin. Mat. Fis. Univ. Modena. 1989. [*455*]

[371]* Constantinescu C., Filter W., Weber K. Advanced integration theory. With the collaboration of Alexia Sontag. Kluwer Academic Publ., Dordrecht, 1998; x+861 pp.

[372]* 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.

[373] Conway J. *The strict topology and compactness in the space of measures.* Trans. Amer. Math. Soc. 1967. V. 126. P. 474–486. [*455*]

[374] Conway J. *A theorem on sequential convergence of measures and some applications.* Pacific J. Math. 1969. V. 28. P. 53–60. [*456*]

[375] Cooper J., Schachermayer W. *Uniform measures and coSaks spaces.* Lecture Notes in Math. 1981. V. 843. P. 217–246. [*451*]

[376] Cornfeld I.P., Fomin S.V., Sinaĭ Ya.G. Ergodic theory. Translated from the Russian. Springer-Verlag, New York, 1982; x+486 pp. (Russian ed.: Moscow, 1980). [*391*]

[377] Cotlar M., Cignoli R. An introduction to functional analysis. North-Holland, Amsterdam–London; American Elsevier, New York, 1974; xiv+585 pp. [**I**: *413;* **II**: *446*]

[378]* Courrège P. Théorie de la mesure. Les cours de Sorbonne. 2e éd. Centre de Documentation Universitaire, Paris, 1965; 233 pp.

[379] Cox G.V. *On Prohorov spaces.* Fund. Math. 1983. V. 116, №1. P. 67–72. [*225, 456*]

[380]* Cramer H. Mathematical methods of statistics. Princeton University Press, Princeton, 1946; xvi+575 pp.

[381] Cramér H., Wold H. *Some theorems on distribution functions.* J. London Math. Soc. 1936. V. 11. P. 290–294. [*453*]

[382] Crauel H. Random probability measures on Polish spaces. Taylor&Francis, London, 2002; xvi+118 pp. [*456*]

[383]* Craven B.D. Lebesgue measure & integral. Pitman, Boston – London, 1982; ix+221 pp.

[384]* 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.

[385]* Crum M.M. *On positive definite functions.* Proc. London Math. Soc. (3). 1956. V. 6. P. 548–560.

[386] Cruzeiro A.-B. *Équations différentielles ordinaires: non explosion et mesures quasi-invariantes.* J. Funct. Anal. 1983. V. 54. P. 193–205. [*460*]

[387] Császár Á. *Sur la structure des espaces de probabilité conditionnelle.* Acta Math. Acad. Sci. Hung. 1955. V. 6. P. 337–361. [*462*]

[388]* Csiszár I. *Information-type measures of difference of probability distributions and indirect observations.* Studia Sci. Math. Hung. 1967. V. 2. P. 299–318.

[389] Csiszár I. *On the weak* continuity of convolution in a convolution algebra over an arbitrary topological group.* Stidia Sci. Math. Hung. 1971. V. 6. P. 27–40. [*451*]

[390]* Csőrnyei M. *How to make Davies' theorem visible.* Bull. London Math. Soc. 2001. V. 33, №1. P. 59–66.

[391] Cuesta-Albertos J.A., Matrán-Bea C. *Stochastic convergence through Skorohod representation theorems and Wasserstein distances.* First International Conference on Stochastic Geometry, Convex Bodies and Empirical Measures (Palermo, 1993). Rend. Circ. Mat. Palermo (2). Suppl. No. 35, pp. 89–113, 1994. [*454*]

[392] Da Prato G., Zabczyk J. Stochastic equations in infinite dimensions. Cambridge University Press, Cambridge, 1992; xviii+454 pp. [*447*]

[393]* Dalen D. van, Monna A.F. Sets and integration: an outline of the development. Wolters – Hoordhoff, Groningen, 1972; 162 pp.

[394] Dalecky Yu.L., Fomin S.V. Measures and differential equations in infinite-dimensional space. Translated from the Russian. Kluwer Academic Publ., Dordrecht, 1991; xvi+337 pp. (Russian ed.: Moscow, 1983). [*125, 448, 453, 456*]

[395] Daletskii Yu.L., Smolyanov O.G. *On the weak sequential completeness of the spaces of Radon measures.* Teor. Verojatn. i Primen. 1984. V. 29, №1. P. 141–147 (in Russian). English transl.: Theor. Probab. Appl. 1985. V. 29, №1. P. 142–147. [*456*]

[396] Dall'Aglio G. *Sugli estremi dei momenti delle funzioni di ripartizioni doppia.* Ann. Scuola Norm. Super. (3). 1956. V. 10. P. 33–74. [*263*]

[397] Dall'Aglio G. *Fréchet classes: the beginnings.* Advances in probability distributions with given marginals (Rome, 1990), pp. 1–12, Math. Appl., 67, Kluwer Academic Publ., Dordrecht, 1991. [*461*]

[398]* 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.

[399] Daniell P.J. *A general form of integral.* Ann. Math. (2). 1917–1918. V. 19. P. 279–294. [**I**: *419, 423;* **II**: *445*]

[400] Daniell P.J. *Integrals in an infinite number of dimensions.* Ann. Math. (2). 1918. V. 20. P. 281–288. [**I**: *419, 423, 429;* **II**: *445*]

[401]* Daniell P.J. *Stieltjes derivatives.* Bull. Amer. Math. Soc. 1919. V. 26. P. 444–448.

[402] Daniell P.J. *Functions of limited variation in an infinite number of dimensions.* Ann. Math. (2). 1919–1920. V. 21. P. 30–38. [**I**: *419;* **II**: *445*]

[403] Daniell P.J. *Further properties of the general integral.* Ann. Math. (2). 1920. V. 21. P. 203–220. [**I**: *419, 423;* **II**: *445*]

[404] D'Aristotile A., Diaconis P., Freedman D. *On merging of probabilities.* Sankhya. Ser. A. 1988. V. 50, №3. P. 363–380. [*237*]

[405]* 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.

[406] Darst R.B. *On universal measurability and perfect probability*. Ann. Math. Statist. 1971. V. 42. P. 352–354. [*444*]

[407]* Darst R.B. C^∞*-functions need not be bimeasurable*. Proc. Amer. Math. Soc. 1971. V. 27. P. 128–132.

[408] Darst R.B. *Two singular measures can agree on balls*. Mathematika. 1973. V. 20, №2. P. 224–225. [*451*]

[409]* David G., Semmes S. Analysis of and on uniformly rectifiable sets. Amer. Math. Soc., Providence, Rhode Island, 1993; xii+356 pp.

[410] Davies R.O. *Subsets of finite measure in analytic sets*. Indag. Math. 1952. V. 14. P. 488–489. [*140*]

[411]* 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.

[412] Davies R.O. *Measures not approximable or specifiable by means of balls*. Mathematika. 1971. V. 18. P. 157–160. [*171, 451*]

[413] Davies R.O. *A non-Prokhorov space*. Bull. London Math. Soc. 1971. V. 3. P. 341–342; Addendum ibid. 1972. V. 4. P. 310. [*224*]

[414]* Davies R.O. *Separate approximate continuity implies measurability*. Math. Proc. Camb. Phil. Soc. 1973. V. 73. P. 461–465.

[415] Davies R.O. *Some counterexamples in measure theory*. Proc. Conf. Topology and Measure III. Part 1, 2. J. Flachsmeyer, Z. Frolík, Ju.M. Smirnov, F. Topsøe, and F. Terpe eds., pp. 49–55. Ernst-Moritz-Arndt-Universität, Greifswald, 1982. [*451*]

[416]* 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.

[417] Davies R.O., Rogers C.A. *The problem of subsets of finite positive measure*. Bull. London Math. Soc. 1969. V. 1. P. 47–54. [*140*]

[418] Davies R.O., Schuss Z. *A proof that Henstock's integral includes Lebesgue's*. J. London Math. Soc. (2). 1970. V. 2. P. 561–562. [*160*]

[419] De Finetti B. *La prévision: ses lois logiques, ses sources subjectives*. Ann. Inst. H. Poincaré. 1937. V. 7. P. 1–68. [*409*]

[420] De Giorgi E., Letta G. *Une notion générale de convergence faible pour des fonctions croissantes d'ensemble*. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4). 1977. V. 4, №1. P. 61–99. [*456*]

[421]* Deheuvels P. L'intégrale. Presses Universitaires de France, Paris, 1980; 229 pp.

[422] Dekiert M. *Two results for monocompact measures*. Manuscripta Math. 1993. V. 80, №4. P. 339–346. [*444*]

[423] de Leeuw K. *The Fubini theorem and convolution formula for regular measures*. Math. Scand. 1962. V. 11. P. 117–122. [*444*]

[424] Dellacherie C. Capacités et processus stochastiques. Springer-Verlag, Berlin – New York, 1972; ix+155 pp. [*142, 356, 440, 441*]

[425] Dellacherie C. Un cours sur les ensembles analytiques. In: Analytic sets, pp. 184–316. Academic Press, New York, 1980. [*142, 261, 440*]

[426] Dellacherie C. *Mesurabilité des débuts et théorèmes de section: le lot à la portée de toutes les bourses*. Lecture Notes in Math. 1981. V. 850. P. 351–370. [*73*]

[427] Dellacherie C., Meyer P.-A. Probabilités et potentiel. Chap. I–IV. Hermann, Paris, 1975; x+291 pp. [*441*]

[428] Dembski W.A. *Uniform probability*. J. Theoret. Probab. 1990. V. 3, №4. P. 611–626. [*255*]

[429]* 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).

[430]* Denjoy A. *Une extension de l'intégrale de M. Lebesgue*. C. R. Acad. Sci. Paris. 1912. T. 154. P. 859–862.

[431]* Denjoy A. *Mémoire sur les nombres derivées des fonctions continues.* J. Math. Pures Appl. (7). 1915. V. 1. P. 105–240.

[432]* Denjoy A. *Sur les fonctions derivées sommables.* Bull. Soc. Math. France. 1915. V. 43. P. 161–248.

[433]* Denjoy A. *Sur une propriété des fonctions derivées.* Enseign. Math. 1916. T. 18. P. 320–328.

[434]* Denjoy A. Articles et mémoires. II. Le champ réel. Notes. II. Le champ réel. Gauthier-Villars, Paris, 1955, 1957.

[435]* Denkowski Z., Migórski S., Papageorgiou N.S. An introduction to nonlinear analysis: theory. Kluwer Academic Publ., Boston, 2003; xvi+689 pp.

[436]* Denneberg D. Non-additive measure and integral. Kluwer Academic Publ., Dordrecht, 1994; x+178 pp.

[437]* DePree J., Swartz C. Introduction to real analysis. John Wiley & Sons, New York, 1988; xii+355 pp.

[438]* Descombes R. Intégration. Hermann, Paris, 1972; 208 pp.

[439]* Dharmadhikari S., Joag-Dev K. Unimodality, convexity, and applications. Academic Press, Boston, 1988; xiv+278 pp.

[440] Diaconis P., Freedman D. *A dozen de Finetti-style results in search of a theory.* Ann. Inst. H. Poincaré. Probab. Statist. 1987. V. 23, №2, suppl. P. 397–423. [*409*]

[441]* DiBenedetto E. Real analysis. Birkhäuser Boston, Boston, 2002; xxiv+485 pp.

[442] Diestel J. Geometry of Banach spaces – selected topics. Lecture Notes in Math. V. 485. Springer-Verlag, Berlin – New York, 1975; xi+282 pp. [**I**: *282, 285, 319, 433;* **II**: *120*]

[443]* Diestel J. Sequences and series in Banach spaces. Springer, Berlin – New York, 1984; xi+261 pp.

[444] Diestel J., Uhl J.J. Vector measures. Amer. Math. Soc., Providence, 1977; xiii+322 pp. [**I**: *423;* **II**: *329*]

[445] Dieudonné J. *Un exemple d'un espace normal non susceptible d'une structure uniforme d'espace complet.* C. R. Acad. Sci. Paris. 1939. V. 209. P. 145–147. [*68*]

[446] Dieudonné J. *Sur le théorème de Lebesgue–Nikodým. III.* Ann. Inst. Fourier (Grenoble) 1947–1948. V. 23. P. 25–53. [*462, 462*]

[447] Dieudonné J. *Sur un théorème de Jessen.* Fund. Math. 1950. V. 37. P. 242–248. [*430*]

[448] Dieudonné J. *Sur la convergence des suites des mesures de Radon.* An. Acad. Brasil. Cienc. 1951. V. 23. P. 21–38; Addition: ibid., P. 277–282. [*241, 454*]

[449] Dieudonné J. *Sur le produit de composition.* Compositio Math. 1954. V. 12. P. 17–34. [*454*]

[450]* Dieudonné J. Éléments d'analyse. Tome II: Chapitres XII à XV. Gauthier-Villars, Paris, 1968; x+408 pp.

[451] Dinculeanu N. *Remarques sur les mesures dans les espaces produits.* Colloq. Math. 1957. V. 5. P. 51–54. [*445*]

[452] Dinculeanu N. Vector measures. Pergamon Press, Oxford – New York – Toronto, VEB Deutscher Verlag der Wissenschaften, Berlin, 1967; x+432 pp. [**I**: *423;* **II**: *463*]

[453] Dinculeanu N. Integration on locally compact spaces. Noordhoff, Leiden, 1974; xv+626 pp. [*447*]

[454]* Dinculeanu N. Vector integration and stochastic integration in Banach spaces. Wiley–Interscience, New York, 2000; xvi+424 pp.

[455]* 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.

[456] DiPerna R.J., Lions P.L. *Ordinary differential equations, transport theory and Sobolev spaces.* Invent. Math. 1989. V. 98, №3. P. 511–547. [*460*]

[457] Ditor S., Eifler L. *Some open mapping theorems for measures.* Trans. Amer. Math. Soc. 1972. V. 164. P. 287–293. [*228*]

[458] Dixmier J. *Sur certains espaces considérés par M.H. Stone*. Summa Brasil Math. 1951. V. 2. P. 151–182. [*451*]

[459]* Dixmier J. L'intégrale de Lebesgue. "Les cours de Sorbonne", Mathématique I. Centre de Documentation Universitaire, Paris, 1964; 93 pp.

[460] Dobrushin R.L. *The description of a random field by means of conditional probabilities and conditions of its regularity.* Teor. Verojatn. i Primen. 1968. V. 13. P. 201–229 (in Russian); English transl.: Theory Probab. Appl. 1968. V. 13, №2. P. 197–224. [*454, 464*]

[461] Dobrushin R.L. *Prescribing a system of random variables by conditional distributions.* Teor. Verojat. i Primen. 1970. V. 15, №3. P. 469–497 (in Russian); English transl.: Theory Probab. Appl. 1970. V. 15. P. 458–486. [*464*]

[462]* 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.

[463] Doob J.L. *Stochastic processes with an integral-valued parameter.* Trans. Amer. Math. Soc. 1938. V. 44, №1. P. 87–150. [*448, 461*]

[464] Doob J.L. *Regularity properties of certain families of chance variables.* Trans. Amer. Math. Soc. 1940. V. 47. P. 455–486. [*461*]

[465] Doob J.L. *Probability in function space.* Bull. Amer. Math. Soc. 1947. V. 53. P. 15–30. [*448*]

[466] Doob J.L. *On a problem of Marczewski.* Colloq. Math. 1948. V. 1. P. 216–217. [*433*]

[467] Doob J.L. Stochastic processes. John Wiley & Sons, New York, 1953; viii+654 pp. [**I**: *412;* **II**: *99, 448, 461*]

[468]* Doob J.L. Measure theory. Springer-Verlag, New York, 1994; xii+210 pp.

[469]* Dorogovtsev A.Ya. Mathematical analysis. Collection of exercises. Vischa Shkola, Kiev, 1987; 408 pp. (in Russian).

[470]* Dorogovtsev A.Ya. Elements of the general theory of measure and integral. Vischa Shkola, Kiev, 1989; 152 pp. (in Russian).

[471]* Douglas R.G. *On extremal measures and subspace density.* Michigan Math. J. 1964. V. 11. P. 243–246.

[472]* Drewnowski L. *Topological rings of sets, continuous set functions. Integration I, II.* Bull. Acad. Polon. Sci. 1972. V. 20. P. 269–286.

[473]* Drewnowski L. *Equivalence of Brooks–Jewett, Vitali–Hahn–Saks, and Nikodym theorems.* Bull. Acad. Polon. Sci. 1972. V. 20. P. 725–731.

[474]* 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.

[475]* 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.

[476] Dubins L.E., Heath D. *With respect to tail sigma fields, standard measures possess measurable disintegrations.* Proc. Amer. Math. Soc. 1983. V. 88, №3. P. 416–418. [*462*]

[477]* 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.

[478]* 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).

[479]* 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).

[480]* 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).

[481]* Dubrovskiĭ V.M. *On properties of absolute continuity and equi-continuity*. Dokl. Akad. Nauk SSSR. 1948. V. 63. P. 483–486 (in Russian).

[482]* Dubrovskiĭ V.M. *On certain conditions of compactness*. Izv. Akad. Nauk SSSR. 1948. V. 12. P. 397–410 (in Russian).

[483]* 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).

[484]* 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).

[485]* 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).

[486]* Dubrovskiĭ V.M. *On a property of a formula of Nikodym*. Dokl. Akad. Nauk SSSR. 1952. V. 85. P. 693–696 (in Russian).

[487]* 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).

[488] Dudley R.M. *Weak convergence of probabilities on nonseparable metric spaces and empirical measures on Euclidean spaces*. Illinois J. Math. 1966. V. 10. P. 109–126. [*456*]

[489] Dudley R. *Convergence of Baire measures*. Studia Math. 1966. V. 27. P. 251–268; Addendum ibid. 1974. V. 51. P. 275. [*456*]

[490] Dudley R.M. *Measures on nonseparable metric spaces*. Illinois J. Math. 1967. V. 11. P. 449–453. [*456*]

[491] Dudley R.M. *Distances of probability measures and random variables*. Ann. Math. Statist. 1968. V. 39, №5. P. 1563–1572. [*456*]

[492] Dudley R.M. *On measurability over product spaces*. Bull. Amer. Math. Soc. 1971. V. 77. P. 271–274. [*449*]

[493] Dudley R.M. *A counter-example on measurable processes*. Proceedings of the sixth Berkeley symposium on mathematical statistics and probability (Berkeley, 1970/71). V. II. P. 57–66. Univ. California Press, Berkeley, 1972; correction: Ann. Probab. 1973. V. 1. P. 191–192. [*449*]

[494] Dudley R.M. Probability and metrics. Lect. Notes Aarhus Univ., 1976. [*456*]

[495] Dudley R.M. Real analysis and probability. Wadsworth & Brooks, Pacific Grove, California, 1989; xii+436 pp. [**I**: *62, 228, 413, 415;* **II**: *166, 236, 453, 456, 461*]

[496] Dudley R.M., Feldman J., Le Cam L. *On seminorms and probabilities, and abstract Wiener spaces*. Ann. Math. (2). 1971. V. 93. P. 390–408. Corrections: ibid. 1976. V. 104, №2. P. 391. [*449*]

[497] Dudley R.M., Kanter M. *Zero–one laws for stable measures*. Proc. Amer. Math. Soc. 1974. V. 45, №2. P. 245–252; Correction: ibid. 1983. V. 88, №4. P. 689–690. [*410, 451*]

[498] Dulst D. van. Characterizations of Banach spaces not containing l^1. Centrum voor Wiskunde en Informatica, Amsterdam, 1989; iv+163 pp. [*444*]

[499]* Dunford N. *Integration in general analysis*. Trans. Amer. Math. Soc. 1935. V. 37. P. 441–453.

[500]* Dunford N. *Uniformity in linear spaces*. Trans. Amer. Math. Soc. 1938. V. 44. P. 305–356.

[501]* Dunford N. *A mean ergodic theorem*. Duke Math. J. 1939. V. 5. P. 635–646.

[502]* Dunford N., Pettis B.J. *Linear operations on summable functions*. Trans. Amer. Math. Soc. 1940. V. 47. P. 323–392.

[503] Dunford N., Schwartz J.T. Linear operators, I. General Theory. Interscience, New York, 1958; xiv+858 pp. [**I**: *240, 282, 283, 321, 413, 415, 421, 423, 424, 434, 435;* **II**: *113, 264, 326, 373, 447, 463*]

[504] Durrett R. Brownian motion and martingales in analysis. Wadsworth, Belmont, 1984; xi+328 pp. [*461*]

[505] Durrett R. Probability. Theory and Examples. 2nd ed., Duxbury Press, Wadsworth Publ., Belmont, 1996; xiii+503 pp. [**I**: *413;* **II**: *432, 461*]

[506]* D'yachenko M.I., Ulyanov P.L. Measure and integral. Faktorial, Moscow, 1998; 160 pp. (in Russian); Spanish transl.: Addison-Wesley Iberoamericana, 2000.

[507] Dynkin E.B. Markov processes. GIFML, Moscow, 1963 (in Russian); English transl.: V. I,II. Academic Press, New York; Springer-Verlag, Berlin – Heidelberg, 1965; xii+365 pp., viii+274 pp. [*441*]

[508] Dzamonja M., Kunen K. *Measures on compact HS spaces.* Fund. Math. 1993. V. 143, №1. P. 41–54. [*452*]

[509] Dzamonja M., Kunen K. *Properties of the class of measure separable compact spaces.* Fund. Math. 1995. V. 147. P. 261–277. [*449, 452*]

[510]* Eaton M.L. Lectures on topics in probability inequalities. Centrum voor Wiskunde en Informatica, Amsterdam, 1987; vi+197 pp.

[511] Edgar G.A. *Disintegration of measures and the vector-valued Radon–Nikodym theorem.* Duke Math. J. 1975. V. 42, №3. P. 447–450. [*405, 462*]

[512] Edgar G.A. *Measurable weak sections.* Illinois J. Math. 1976. V. 20, №4. P. 630–646. [*321, 322*]

[513] Edgar G.A. *Measurability in a Banach space. I.* Indiana Math. J. 1977. V. 26, №4. P. 663–677. [*45, 52*]

[514] Edgar G.A. *Measurability in a Banach space. II.* Indiana Math. J. 1979. V. 28, №4. P. 559–579. [*45, 52*]

[515] Edgar G.A. *On pointwise-compact sets of measurable functions.* Lecture Notes in Math. 1982. V. 945. P. 24–28. [*151*]

[516]* Edgar G.A. Integral, probability, and fractal measures. Springer-Verlag, New York, 1998; 286 pp.

[517] Edgar G.A., Sucheston L. Stopping times and directed processes. Cambridge University Press, Cambridge, 1992; 428 pp. [**I**: *435, 438;* **II**: *461, 463*]

[518] Edwards R.E. Functional analysis. Theory and applications. Holt, Rinehart and Winston, New York – London, 1965; xiii+781 pp. [**I**: *423;* **II**: *119, 146, 319*]

[519] Edwards R.E. Integration and harmonic analysis on compact groups. Cambridge University Press, London – New York, 1972; vi+184 pp. [*451*]

[520]* 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.

[521]* Eggleston H.G. *Two measure properties of Cartesian product sets.* Quart. J. Math., Oxford, Ser. (2). 1954. V. 5. P. 108–115.

[522]* Egoroff D.-Th. *Sur les suites de fonctions mesurables.* C. R. Acad. Sci. Paris. 1911. T. 152. P. 244–246.

[523]* Egoroff D.-Th. *Sur l'intégration des fonctions mesurables.* C. R. Acad. Sci. Paris. 1912. T. 155. P. 1474–1475.

[524] Eifler L.Q. *Open mapping theorems for probability measures on metric spaces.* Pacif. J. Math. 1976. V. 66. P. 89–97. [*228*]

[525] Eisele K.-Th. *On the uniqueness of preimages of measures.* Lecture Notes in Math. 1976. V. 541. P. 5–12. [*311*]

[526]* Eisen M. Introduction to mathematical probability theory. Prentice Hall, Englewood Cliffs, New Jersey, 1969; xxi+535 pp.

[527] Elliott E.O. *Measures on product spaces.* Trans. Amer. Math. Soc. 1967. V. 128. P. 379–388; corr.: ibid., 1969. V. 139. P. 511. [*444, 452*]

[528] Elliott E.O., Morse A.P. *General product measures.* Trans. Amer. Math. Soc. 1964. V. 110. P. 245–283; corr.: ibid., 1971. V. 157. P. 505–506. [*452*]

[529] Ellis H.W. *Darboux properties and applications to non-absolutely convergent integrals.* Canad. J. Math. 1951. V. 3. P. 471–485. [*460*]

[530] Elstrodt J. Mass- und Integrationstheorie. 4e Aufl. Springer-Verlag, Berlin – New York, 2005; xvi+434 S. [**I**: *413, 415;* **II**: *61*]

[531]* Ene V. Real functions: current topics. Lecture Notes in Math. V. 1603. Springer-Verlag, Berlin – New York, 1995; xi+310 pp.

[532] Engelking P. General topology. Polish Sci. Publ., Warszawa, 1977; 626 pp. [*1, 6, 7, 8, 9, 13, 45, 54, 55, 58, 62, 75, 77, 83, 111, 114, 166, 173, 201, 244, 289*]

[533]* Erdős P., Kestelman H., Rogers C.A. *An intersection property of sets with positive measure*. Colloq. Math. 1963. V. 11. P. 75–80.

[534]* 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.

[535] Erdős P., Stone A.H. *On the sum of two Borel sets*. Proc. Amer. Math. Soc. 1970. V. 25. P. 304–306. [*60*]

[536]* Erdős P., Taylor S.J. *The Hausdorff measure of the intersection of sets of positive Lebesgue measure*. Mathematika. 1963. V. 10. P. 1–9.

[537] Erohin V.D. *A note on measure theory*. Uspehi Mat. Nauk. 1961. V. 16, №3. P. 175–180 (in Russian). [*173, 443*]

[538] Ershov M.P. *Extensions of measures. Stochastic equations*. Lecture Notes in Math. 1973. V. 330. P. 516–526. [*458*]

[539] Ershov M.P. *Extension of measures and stochastic equations*. Teor. Verojatn. i Primen. 1974. V. 19, №3. P. 457–471 (in Russian); English transl.: Theory Probab. Appl. 1974. V. 19. P. 431–444. [*311, 458, 459*]

[540] Ershov M.P. *The Choquet theorem and stochastic equations*. Anal. Math. 1975. V. 1, №4. P. 259–271. [*458*]

[541] Ershov M.P. *On a generalization of the Ionescu Tulcea construction of a measure by transition kernels*. Lecture Notes in Math. 1982. V. 945. P. 29–33. [*463*]

[542] Ershov M.P. (Jerschow M.) *Causal selections and solutions of stochastic equations*. Stochast. and Stoch. Reports. 1994. V. 50. P. 161–173. [*458*]

[543] Ethier S.N., Kurtz T.G. Markov processes. Characterization and convergence. John Wiley & Sons, New York, 1986; x+534 pp. [*453*]

[544]* Evans C., Gariepy R.F. Measure theory and fine properties of functions. CRC Press, Boca Raton – London, 1992; viii+268 pp.

[545] Evstigneev I.V. *Measurable selection theorems and probabilistic models of control in general topological spaces*. Mat. Sbornik. 1986. V. 131, №1. P. 27–39 (in Russian); English transl.: Math. USSR Sb. 1988. V. 59, №1. P. 25–37. [*41*]

[546]* Faber V., Mycielski J. *The vanishing of certain integrals*. Houston J. Math. 1990. V. 16, №3. P. 313–316.

[547] Faden A.M. *The existence of regular conditional probabilities: necessary and sufficient conditions*. Ann. Probab. 1985. V. 13, №1. P. 288–298. [*462*]

[548]* 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.

[549]* Falconer K.J. *Sections of sets of zero Lebesgue measure*. Mathematika. 1980. V. 27, №1. P. 90–96.

[550]* Falconer K. Fractal geometry. John Wiley & Sons, New York, 1990; xxii+288 pp.

[551]* Falconer K.J., Mauldin R.D. *Fubini-type theorems for general measure constructions*. Mathematika. 2000. V. 47, №1–2. P. 251–265.

[552]* Farrell R.H. *Dense algebras of functions in L^p*. Proc. Amer. Math. Soc. 1962. V. 13. P. 324–328.

[553]* Fatou P. *Séries trigonométriques et séries de Taylor*. Acta Math. 1906. V. 30. P. 335–400.

[554]* Federer H. *Surface area. II*. Trans. Amer. Math. Soc. 1944. V. 55. P. 438–456.

[555] Federer H. Geometric measure theory. Springer-Verlag, New York, 1969; xiv+676 pp. [**I**: *79, 243, 312, 373, 381, 413, 430, 437;* **II**: *460*]

[556] Federer H., Morse A.P. *Some properties of measurable functions*. Bull. Amer. Math. Soc. 1943. V. 49. P. 270–277. [*331*]

[557] Fedorchuk V.V. *Probability measures in topology.* Uspehi Mat. Nauk. 1991. V. 46, №1. P. 41–80 (in Russian); English transl.: Russian Math. Surveys. 1991. V. 46. P. 45–93. [*455*]
[558] Fedorchuk V.V. *Functors of probability measures in topological categories.* J. Math. Sci. (New York). 1998. V. 91, №4. P. 3157–3204. [*455*]
[559] Fedorchuk V.V. *On the topological completeness of spaces of measures.* Izv. Ross. Akad. Nauk Ser. Mat. 1999. V. 63, №4. P. 207–223 (in Russian); English transl.: Izv. Math. 1999. V. 63, №4. P. 827–843. [*455*]
[560] Fedorchuk V.V., Sadovnichiĭ Yu.V. *On some topological and categorical properties of measures with alternating signs.* Fundam. Prikl. Mat. 1999. V. 5, №2. P. 597–618 (in Russian). [*457*]
[561] Fedorchuk V.V., Filippov V.V. General topology. Basic constructions. Moskov. Gos. Univ., Mocsow, 1988; 252 pp. (in Russian). [*201, 245*]
[562]* Fernandez P.J. Medida e integracao. Instituto de Matematica Pura e Aplicada, Rio de Janeiro, 1976; viii+198 pp.
[563] Fernique X. *Processus linéaires, processus généralisés.* Ann. Inst. Fourier (Grenoble). 1967. V. 17. P. 1–92. [*224, 456*]
[564] Fernique X. *Une démonstration simple du théorème de R.M. Dudley et M. Kanter sur les lois zero–un pour les mesures stables.* Lecture Notes in Math. 1974. V. 381. P. 78–79. [*410, 451*]
[565] Fernique X. *Sur le théorème de Kantorovitch–Rubinstein dans les espaces polonais.* Lecture Notes in Math. 1981. V. 850. P. 6–10. [*456*]
[566] Fernique X. *Un modèle presque sûr pour la convergence en loi.* C. R. Acad. Sci. Paris, Sér. 1. 1988. T. 306. P. 335–338. [*199, 454*]
[567] Fernique X. *Convergence en loi de fonctions aléatoires continues ou cadlag, propriétés de compacité des lois.* Lecture Notes in Math. 1991. V. 1485. P. 178–195. [*456*]
[568] Fernique X. *Convergence en loi de variables aléatoires et de fonctions aléatoires, propriétés de compacité des lois. II.* Lecture Notes in Math. 1993. V. 1557. P. 216–232. [*456*]
[569] Fernique X. *Sur l'équivalence de certaines mesures produits.* Probab. Theory Related Fields. 1994. V. 98, №1. P. 77–90. [*462*]
[570] Fernique X. Fonctions aléatoires gaussiennes, vecteurs aléatoires gaussiens. Université de Montréal, Centre de Recherches Mathématiques, Montréal, 1997; iv+217 pp. [*451*]
[571] Feyel D., Üstünel A.S. *Monge–Kantorovitch measure transportation and Monge–Ampère equation on Wiener space.* Probab. Theor. Related Fields. 2004. V. 128, №3. P. 347–385. [*236*]
[572]* Fichera G. Lezioni sulle trasformazioni lineari. V. I. Introduzione all'analisi lineare. Istituto Matematico, Università, Trieste, 1954; xvii+502+iv pp.
[573]* Fichtenholz G. *Un théorème sur l'intégration sous le signe intégrale.* Rend. Circ. Mat. Palermo. 1913. T. 36. P. 111–114.
[574]* Fichtenholz (Fihtengol'c) G.M. *Definite integrals depending on a parameter.* Mat. Sbornik. 1913. V. 29. P. 53–66 (in Russian).
[575] Fichtenholz G. *Notes sur les limites des fonctions représentées par des intégrales définies.* Rend. Circ. Mat. Palermo. 1915. T. 40. P. 153–166. [**I**: *428, 433;* **II**: *188*]
[576] Fichtenholz G.M. Theory of simple integrals dependent on a parameter. Petrograd, 1918; vii+333 pp. (in Russian). [**I**: *276, 411, 428;* **II**: *241*]
[577]* Fichtenholz G. *Note sur les fonctions absolument continues.* Bull. Acad. Belgique (5). 1922. V. 8. P. 430–443.
[578] Fichtenholz G. *Sur les suites convergentes des intégrales définies.* Bulletin 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. [**I**: *276, 433;* **II**: *241*]

[579]* Fichtenholz G. *Sur une fonction de deux variables sans intégrale double.* Fund. Math. 1924. V. 6. P. 30–36.

[580]* Fichtenholz G.M. *On absolutely continuous functions.* Mat. Sbornik. 1924. V. 31. P. 286–296 (in Russian).

[581]* Fichtenholz G. *Sur un problème de M. Banach.* Fund. Math. 1927. V. 10. P. 302–304.

[582]* Fichtenholz G. *Sur l'intégration des suites des fonctions sommables.* C. R. Acad. Sci. Paris. 1927. T. 184. P. 436–438.

[583]* Fichtenholz G.M. Differential- und Integralrechnung. B. I–III. Translated from the Russian. VEB Deutscher Verlag der Wissenschaften, Berlin, 1986, 1987.

[584] Fichtenholz G., Kantorovitch L. *Sur les opérations linéaires dans l'espace de fonctions bornées.* Studia Math. 1936. V. 5. P. 68–98. [**I**: *435;* **II**: *464*]

[585] Filippov V.V. *On a question of E.A. Michael.* Comm. Math. Univ. Carol. 2004. V. 45, №4. P. 735–738. [*229*]

[586] Filter W., Weber K. Integration theory. Chapman & Hall, London, 1997; xii+294 pp. [**I**: *413, 422;* **II**: *446*]

[587]* Fischer E. *Sur la convergence en moyenne.* C. R. Acad. Sci. Paris. 1907. T. 144. P. 1022–1024.

[588]* Fischer E. *Applications d'un théorème sur la convergence en moyenne.* C. R. Acad. Sci. Paris. 1907. T. 144. P. 1148–1151.

[589] Flachsmeyer J., Lotz S. *A survey on hyperdiffuse measures.* Proc. Conf. Topology and Measure, I (Zinnowitz, 1974), Part 1, pp. 87–128. Ernst-Moritz-Arndt Univ., Greifswald, 1978. [*451*]

[590] Flachsmeyer J., Terpe F. *Some applications of extension theory for topological spaces and measure theory.* Uspehi Mat. Nauk. 1977. V. 32, №5. P. 125–162 (in Russian); English transl.: Russian Math. Surveys. 1977. V. 32, №5. P. 133–171. [*455*]

[591]* Fleming W. Functions of several variables. 2nd ed. Springer-Verlag, Berlin – New York, 1977; xi+411 pp.

[592]* Floret K. Maß- und Integrationstheorie. B.G. Teubner, Stuttgart, 1981; 360 S.

[593]* Folland G.B. Real analysis: modern techniques and their applications. 2nd ed. Wiley, New York, 1999; xiv+386 pp.

[594]* 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.

[595]* Fonda A. Lezioni sulla teoria dell'integrale. Edizioni goliardiche, Udine, 2001; 206 pp.

[596] Fonf V.P., Johnson W.B., Pisier G., Preiss D. *Stochastic approximation properties in Banach spaces.* Studia Math. 2003. V. 159, №1. P. 103–119. [*120, 145*]

[597]* Foran J. Fundamentals of real analysis. Marcel Dekker, New York, 1991; xiv+473 pp.

[598]* Forster O. Analysis. B. 3. Integralrechnung im $\mathbb{R}^n$ mit Anwendungen. Vieweg, Braunschweig, 1999; vii+285 S.

[599] Fortet R., Mourier E. *Convergence de la répartition empirique vers la répartition théorique.* Ann. Sci. École Norm. Sup. (3). 1953. V. 70. P. 267–285. [*453*]

[600] Fortet R., Mourier E. *Les fonctions aléatoires comme éléments aléatoires dans les espaces de Banach.* Studia Math. 1955. V. 15. P. 62–73. [*447*]

[601] Fox G. Intégration dans les groupes topologiques. Séminaire de Mathématiques Supérieures, No. 12 (Été, 1964). Les Presses de l'Université de Montréal, Montréal, 1966; 357 pp. [*451*]

[602] Frankiewicz R., Plebanek G., Ryll-Nardzewski C. *Between Lindelöf property and countable tightness.* Proc. Amer. Math. Soc. 2001. V. 129, №1. P. 97–103. [*455*]

[603]* Fréchet M. *Sur les ensembles de fonctions et les opérations linéaires.* C. R. Acad. Sci. Paris. 1907. T. 144. P. 1414–1415.

[604]* Fréchet M. *Définition de l'intégrale sur un ensemble abstrait.* C. R. Acad. Sci. Paris. 1915. T. 161. P. 839–840.

[605]* Fréchet M. *Sur l'intégrale d'une fonctionnelle étendue a un ensemble abstrait*. Bull. Sci. Math. France. 1915. V. 43. P. 249–265.

[606]* 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.

[607]* Fréchet M. *Familles additives et fonctions additives d'ensembles abstraits*. Enseignement Math. 1922. V. 22. P. 113–129.

[608]* Fréchet M. *Sur la distance de deux ensembles*. C. R. Acad. Sci. Paris. 1923. T. 176. P. 1123–1124.

[609]* Fréchet M. *Des familles et fonctions additives d'ensembles abstraits*. Fund. Math. 1923. V. 4. P. 329–365.

[610]* Fréchet M. *Des familles et fonctions additives d'ensembles abstraits. Suite*. Fund. Math. 1924. V. 5. P. 206–251.

[611]* Fréchet M. *Sur la distance de deux ensembles*. Bull. Calcutta Math. Soc. 1924. V. 15. P. 1–8.

[612]* Fréchet M. *Sur les ensembles compacts de fonctions mesurables*. Fund. Math. 1927. V. 9. P. 25–32.

[613]* Fréchet M. *Sur la convergence en probabilité*. Metron. 1930. V. 8. P. 3–50.

[614]* Fréchet M. *Sur les ensembles compacts de fonctions de carrés intégrables*. Acta Sci. Math. (Szeged). 1937. V. 8. P. 116–126.

[615] Fréchet M. *Intégrale abstraite d'une fonction abstraite d'une variable abstraite et son application à la moyenne d'un élément aléatoire de nature quelconque*. Revue Sci. (Rev. Rose Illus.) 1944. V. 82. P. 483–512. [*447*]

[616] Fréchet M. *Les éléments aléatoires de nature quelconque dans un espace distancié*. Ann. Inst. H. Poincaré. 1948. V. 10. P. 215–310. [**I**: *426;* **II**: *447*]

[617]* Fréchet M. Généralités sur les probabilités. Éléments aléatoires. 2d ed. Gauthier-Villars, Paris, 1950; xvi+355 pp.

[618] Fréchet M. Pages choisies d'analyse générale. Gauthier-Villars, Paris, 1953; 213 pp. [**I**: *409, 426;* **II**: *447*]

[619] Fremlin D.H. Topological Riesz spaces and measure theory. Cambridge University Press, London, 1974; 266 pp. [**I**: *312;* **II**: *104, 447*]

[620] Fremlin D.H. *Topological measure theory: two counter-examples*. Math. Proc. Cambridge Phil. Soc. 1975. V. 78. P. 95–106. [*155*]

[621] Fremlin D.H. *Pointwise compact sets of measurable functions*. Manuscr. Math. 1975. V. 15. P. 219–242. [*129, 151, 153, 162*]

[622] Fremlin D.H. *Products of Radon measures: a counter-example*. Canad. Math. Bull. 1976. V. 19, №3. P. 285–289. [*127*]

[623] Fremlin D.H. *Uncountable powers of* **R** *can be almost Lindelöf*. Manuscr. Math. 1977. V. 22. P. 77–85. [*131*]

[624] Fremlin D.H. *Borel sets in non-separable Banach spaces*. Hokkaido Math. J. 1980. V. 9. P. 179–183. [*46*]

[625] Fremlin D.H. *Measurable functions and almost continuous functions*. Manuscr. Math. 1981. V. 33, №3–4. P. 387–405. [*137, 157, 322*]

[626] Fremlin D.H. *On the additivity and cofinality of Radon measures*. Mathematika. 1984. V. 31, №2. P. 323–335. [*451*]

[627] Fremlin D.H. Consequences of Martin's Axiom. Cambridge University Press, Cambridge, 1985; xii+325 pp. [**I**: *78, 80;* **II**: *136*]

[628] Fremlin D.H. *Measure-additive coverings and measurable selectors*. Dissertationes Math. (Rozprawy Mat.). 1987. V. 260. 116 pp. [*137*]

[629] Fremlin D. *Measure algebras*. Handbook of Boolean Algebras, J.D. Monk ed., V. 3, pp. 877–980. North-Holland, 1989. [*280*]

[630] Fremlin D.H. *On the average of inner and outer measures*. Fund. Math. 1991. V. 139, №1. P. 9–15. [**I**: *100;* **II**: *166*]

[631]* 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.

[632] Fremlin D.H. *On compact spaces carrying Radon measures of uncountable Maharam type.* Fund. Math. 1997. V. 154, №3. P. 295–304. [*452*]

[633] Fremlin D.H. *Compact measure spaces.* Mathematika. 1999. V. 46, №2. P. 331–336. [*464*]

[634] Fremlin D.H. *Weakly α-favourable measure spaces.* Fund. Math. 2000. V. 165, №1. P. 67–94. [**I**: *53;* **II**: *444*]

[635] Fremlin D. Measure theory. V. 1–5. University of Essex, Colchester, 2000–2003. [**I**: *74, 98, 235, 237, 325, 413, 421, 434;* **II**: *104, 131, 135, 151, 153, 154, 155, 159, 164, 166, 166, 171, 280, 308, 320, 337, 443, 459, 461, 463*]

[636] Fremlin D., Garling D., Haydon R. *Bounded measures on topological spaces.* Proc. London Math. Soc. 1972. V. 25. P. 115–136. [*224, 255, 309, 337, 456, 458*]

[637] Fremlin D., Grekas S. *Products of completion regular measures.* Fund. Math. 1995. V. 147, №1. P. 27–37. [*134*]

[638] Fremlin D., Plebanek G. *Large families of mutually singular Radon measures.* Bull. Polish Acad. Sci. 2003. V. 51, №2. P. 169–174. [*463*]

[639] Fremlin D., Talagrand M. *On the representation of increasing linear functionals on Riesz spaces by measures.* Mathematika. 1978. V. 25. P. 213–215. [*104*]

[640]* Friedman H. *A consistent Fubini–Tonelli theorem for nonmeasurable functions.* Illinois J. Math. 1980. V. 24. P. 390–395.

[641]* Fristedt B., Gray L. A modern approach to probability theory. Birkhäuser Boston, Boston, 1997; xx+756 pp.

[642] Frolík Z. *A survey of separable descriptive theory of sets and spaces.* Czech. Math. J. 1970. V. 20 (95). P. 406–467. [*440*]

[643] Frolík Z., Pachl J. *Pure measures.* Comment. Math. Univ. Carol. 1973. V. 14. P. 279–293. [*173, 444*]

[644]* 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).

[645]* 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.

[646]* 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.

[647]* Fubini G. Opere scelte. V. 1–3. Edizioni Cremonese, Roma, 1957, 1958, 1962.

[648]* Fukuda R. *Exponential integrability of sub-Gaussian vectors.* Probab. Theory Relat. Fields. 1990. V. 85. P. 505–521.

[649]* Galambos J. Advanced probability theory. Marcel Dekker, New York, 1988; viii+405 pp.

[650]* Galambos J., Simonelli I. Bonferroni-type inequalities with applications. Springer, New York, 1996; 296 pp.

[651] Gale S.L. *Measure-compact spaces.* Topology Appl. 1992. V. 45, №2. P. 103–118. [*131*]

[652] Gänssler P. *A convergence theorem for measures in regular Hausdorff spaces.* Math. Scand. 1971. V. 29. P. 237–244. [*244*]

[653] Gänssler P. *Compactness and sequential compactness in spaces of measures.* Z. Wahrscheinlichkeitstheorie verw. Geb. 1971. B. 17. S. 124–146. [*244*]

[654] Gänssler P. Empirical processes. Inst. Math. Statist., Hayward, California, 1983, ix+179 pp. [*453*]

[655] Gänssler P., Pfanzagl J. *Convergence of conditional expectations.* Ann. Math. Statist. 1971. V. 42, №1. P. 315–324. [*370*]

[656] Gänssler P., Stute W. Wahrscheinlichkeitstheorie. Springer-Verlag, Berlin – New York, 1977; xii+418 pp. [**I**: *413;* **II**: *453*]

[657]* 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.

[658] 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. [**I**: *434;* **II**: *433, 464*]

[659]* García-Cuerva J., Rubio de Francia J.L. Weighted norm inequalities and related topics. North-Holland, Amsterdam, 1985; x+604 p.

[660] Gardner R.J. *The regularity of Borel measures.* Lecture Notes in Math. 1982. V. 945. P. 42–100. [*127, 134, 135, 165, 225, 443, 449, 451, 451*]

[661] Gardner R.J. *A note on conditional distributions and orthogonal measures.* Ann. Probab. 1982. V. 10, №3. P. 877–878. [*463*]

[662] 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*]

[663]* Gardner R.J. *The Brunn–Minkowski inequality.* Bull. Amer. Math. Soc. 2002. V. 39, №3. P. 355–405.

[664] Gardner R.J., Gruenhage G. *Finite Borel measures on spaces of cardinality less than c.* Proc. Amer. Math. Soc. 1981. V. 81. P. 624–628. [*131*]

[665]* 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.

[666] Gardner R.J., Pfeffer W.F. *Borel measures.* Handbook of set-theoretic topology, pp. 961–1043. North-Holland, Amsterdam – New York, 1984. [*155, 443, 449, 450, 451*]

[667] Gardner R.J., Pfeffer W.F. *Decomposability of Radon measures.* Trans. Amer. Math. Soc. 1984. V. 283, №1. P. 283–293. [*155*]

[668] Garling D. *A "short" proof of the Riesz representation theorem.* Proc. Camb. Phil. Soc. 1973. V. 73, №3. P. 459–460. [*447*]

[669]* Garnir H.G. Fonctions de variables réelles. T. 2. Librairie Universitaire, Louvain, Gauthier-Villars, Paris, 1965; viii+560 pp.

[670]* 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.

[671] Garsia A.M. Topics in almost everywhere convergence. Markham Publ., Chicago, 1970; x+154 pp. [**I**: *261;* **II**: *391*]

[672] Gâteaux R. *Sur la notion d'intégrale dans le domaine fonctionnel et sur la théorie du potentiel.* Bull. Soc. Math. France. 1919. V. 47. P. 48–67. [*254, 452*]

[673]* Gaughan E. Introduction to analysis. Brooks/Coole, Belmont, 1968; viii+310 pp.

[674] Gelbaum B.R. *Cantor sets in metric measure spaces.* Proc. Amer. Math. Soc. 1970. V. 24. P. 341–343. [*330*]

[675]* Gelbaum B. Problems in real and complex analysis. Springer, New York, 1992; x+488 pp.

[676]* Gelbaum B.R., Olmsted J.M.H. Counterexamples in analysis. Holden-Day, San Francisco – London – Amsterdam, 1964; xxiv+194 pp.

[677] Gel'fand I.M., Vilenkin N.Ya. Generalized functions. V. 4. Applications of harmonic analysis. Translated from the Russian. Academic Press, New York – London, 1964; xiv+384 pp. (Russian ed.: Moscow, 1961). [*447*]

[678]* Genet J. Mesure et intégration. Théorie élémentaire. Cours et exercices resolus. Vuibert, Paris, 1976; 322 pp.

[679]* George C. Exercises in integration. Springer-Verlag, Berlin – New York, 1984; 550 pp.

[680] Georgii H.-O. Gibbs measures and phase transitions. Walter de Gruyter, Berlin, 1988; xiv+525 pp. [*464*]

[681] Gerard P. *Suites de Cauchy et compacité dans les espaces de mesures.* Bull. Soc. Roy. Sci. Liège. 1973. V. 42, №1–2. P. 41–49. [*456*]

[682] Gerard P. *Un critère de compacité dans l'espace* $M_t^+(E)$. Bull. Soc. Roy. Sci. Liège. 1973. V. 42, №5–6. P. 179–182. [*456*]

[683] 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. [**I**: *379;* **II**: *231, 252*]

[684]* 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).

[685] Gikhman I.I., Skorokhod A.V. The theory of stochastic processes. V. 1. Translated from the Russian. Springer-Verlag, New York – Heidelberg, 1974; viii+570 pp. (Russian ed.: Moscow, 1971). [*98, 453*]

[686] Gilat D. *Every nonnegative submartingale is the absolute value of a martingale.* Ann. Probab. 1977. V. 5. P. 475–481. [*432*]

[687]* Gillis J. *Note on a property of measurable sets.* J. London Math. Soc. 1936. V. 11. P. 139–141.

[688]* Gillis J. *Some combinatorial properties of measurable sets.* Quart. J. Math. 1936. V. 7. P. 191–108.

[689]* Girardi M. *Compactness in L_1, Dunford-Pettis operators, geometry of Banach spaces.* Proc. Amer. Math. Soc. 1991. V. 111, №3. P. 767–777.

[690]* Girardi M. *Weak vs. norm compactness in L_1: the Bocce criterion.* Studia Math. 1991. V. 98. P. 95–97.

[691]* Giustu E. Minimal surfaces and functions of bounded variation. Birkhäuser Verlag, Basel, 1984; xii+240 pp.

[692] Givens C.R., Shortt R.M. *A class of Wasserstein metrics for probability distributions.* Michigan Math. J. 1984. V. 31, №2. P. 231–240. [*456*]

[693]* Gładysz S., Marczewski E., Ryll-Nardzewski C. *Concerning distances of sets and distances of functions.* Colloq. Math. 1961. V. 8. P. 71–75.

[694]* 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.

[695]* Gleason A.M. Fundamentals of abstract analysis. Reading, Addison-Wesley, Massachusetts, 1966; xi+404 pp.

[696] Glicksberg I. *The representation of functionals by integrals.* Duke Math. J. 1952. V. 19. P. 253–261. [*451*]

[697] Glicksberg I. *Weak compactness and separate continuity.* Pacif. J. Math. 1961. V. 11. P. 205–214. [*130*]

[698] Glivenko E.V. *On measures of the Hausdorff type.* Mat. Sbornik. 1956. V. 39, №4. P. 423–432 (in Russian). [*450*]

[699] Glivenko V.I. The Stieltjes integral. ONTI, Moscow – Leningrad, 1936; 216 pp. (in Russian) [**I**: *425;* **II**: *264, 265, 452*]

[700] 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. [**I**: *412;* **II**: *442, 444*]

[701]* Gneiting T. *Curiosities of characteristic functions.* Exposition. Math. 2001. V. 19, №4. P. 359–363.

[702]* Godement R. Analyse mathématique. IV. Springer-Verlag, Berlin, 2003; xii+599 pp.

[703] Godfrey M.C., Sion M. *On products of Radon measures.* Canad. Math. Bull. 1969. V. 12. P. 427–444. [*127, 444*]

[704]* Goffman C. Real functions. Prindle, Weber & Schmidt, Boston, 1953; xii+263 pp.

[705]* Goffman C., Pedrick G. First course in functional analysis. Prentice-Hall, Englewood Cliffs, 1965; xi+282 pp.

[706]* Goguadze D.F. On the Kolmogorov integrals and some their applications. Metsniereba, Tbilisi, 1979; 288 pp. (in Russian).

[707]* Gohman E.H. The Stieltjes integral and its applications. Fizmatgiz, Moscow, 1958; 192 pp. (in Russian).

[708]* Goldberg R.R. Methods of real analysis. 2nd ed. John Wiley & Sons, New York – London – Sydney, 1976; x+402 pp.

[709] 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). [**I**: *379;* **II**: *142*]

[710] Goldstine H.H. *Linear functionals and integrals in abstract spaces.* Bull. Amer. Math. Soc. 1941. V. 47. P. 615–620. [*445*]

[711]* Gomes R.L. Intégral de Riemann. Junta de Investigaọ Matemática, Porto, 1949; viii+309+i pp.

[712]* Gordon R.A. The integrals of Lebesgue, Denjoy, Perron, and Henstock. Amer. Math. Soc., Providence, Rhode Island, 1994; 395 pp.

[713]* Gordon R.A. *Some comments on the McShane and Henstock integrals.* Real Anal. Exchange. 1997–1998. V. 23, №1. P. 329–342.

[714] Gorin E.A., Koldobskiĭ A.L. *On potentials of measures in Banach spaces.* Sibirsk. Mat. Zh. 1987. V. 28, №1. P. 65–80 (in Russian); English transl.: Sib. Math. J. 1987. V. 28, №1. P. 46–58. [*451*]

[715] Gould G., Mahowald M. *Measures on completely regular spaces.* J. London Math. Soc. 1962. V. 37. P. 103–111. [*451*]

[716]* Gouyon R. Intégration et distributions. Vuibert, Paris, 1967; 170 pp.

[717]* Gowurin M.K. *On sequences of indefinite integrals.* Bull. Amer. Math. Soc. 1936. V. 42, №12. P. 930–936.

[718] Graf S. *A measurable selection theorem for compact-valued maps.* Manuscr. Math. 1979. V. 27. P. 341–352. [*41, 64*]

[719] Graf S. *Induced σ-homomorphisms and a parametrization of measurable sections via extremal preimage measures.* Math. Ann. 1980. B. 247, №1. S. 67–80. [*310, 321*]

[720] Graf S. *Measurable weak selections.* Lecture Notes in Math. 1980. V. 794. P. 117–140. [*311*]

[721] Graf S. *Selected results on measurable selections.* Proceedings of the 10th Winter School on Abstract Analysis (Srní, 1982). Rend. Circ. Mat. Palermo (2). 1982. P. 87–122. [*441*]

[722] Graf S., Luschgy H. Foundations of quantization for probability distributions. Lecture Notes in Math. V. 1730. Springer-Verlag, Berlin – New York, 2000; x+230 pp. [*448*]

[723] Graf S., Mauldin R.D. *Measurable one-to-one selections and transition kernels.* Amer. J. Math. 1985. V. 107, №2. P. 407–425. [*441*]

[724] Graf S., Mauldin R.D. *A classification of disintegrations of measures.* Measure and measurable dynamics (Rochester, NY, 1987), pp. 147–158. Contemp. Math., 94, Amer. Math. Soc., Providence, Rhode Island, 1989. [*462*]

[725]* Gramain A. Intégration. Hermann, Paris, 1988; x+229 pp.

[726] Grande Z. *La mesurabilité des fonctions de deux variables et de superposition* $F(x, f(x))$. Dissert. Math. 1978. V. 159. P. 1–50. [*445*]

[727] Grande Z. *On the measurability of functions defined on the product of two topological spaces.* Real Anal. Exchange. 1997–1998. V. 23 (1). P. 203–210. [*445*]

[728] Grande Z., Lipiński J.S. *Un exemple d'une fonction sup-mesurable qui n'est pas mesurable.* Colloq. Math. 1978. V. 39, №1. P. 77–79. [*164*]

[729] Granirer E.E. *On Baire measures on D-topological spaces.* Fund. Math. 1967. V. 60. P. 1–22. [*455*]

[730]* Grauert H., Lieb I. Differential- und Integralrechnung. III. Integrationstheorie. Kurven- und Flächenintegrale. Vektoranalysis. 2e Aufl. Springer-Verlag, Berlin – New York, 1977; xiv+210 S.

[731]* Gravé D. *Sur les lignes composées de parties rectilignes.* C. R. Acad. Sci. Paris. 1898. T. 127. P. 1005–1007.

[732]* Graves L.M. The theory of functions of real variables. McGraw-Hill, New York, 1946 (2nd ed.: 1956; xii+375 pp.).

[733] Greenleaf F.P. Invariant means on topological groups and their applications. Van Nostrand, New York, 1969; ix+113 pp. [*333, 460*]

[734] Grekas S. *On products of completion regular measures*. J. Math. Anal. Appl. 1992. V. 171, №1. P. 101–110. [*444*]

[735] Grekas S. *Isomorphic measures on compact groups*. Math. Proc. Cambridge Philos. Soc. 1992. V. 112, №2. P. 349–360; corr.: ibid. 1994. V. 115, №2. P. 377. [*451*]

[736] Grekas S. *Structural properties of compact groups with measure-theoretic applications*. Israel J. Math. 1994. V. 87, №1–3. P. 89–95. [*451*]

[737] Grekas S., Gryllakis C. *Completion regular measures on product spaces with application to the existence of Baire strong liftings*. Illinois J. Math. 1991. V. 35, №2. P. 260–268. [*444, 463*]

[738] Grekas S., Gryllakis C. *Measures on product spaces and the existence of strong Baire liftings*. Monatsh. Math. 1992. V. 114, №1. P. 63–76. [*444, 463*]

[739] Grenander U. Probabilities on algebraic structures. John Wiley & Sons, New York – London; Almqvist & Wiksell, Stockholm – Göteborg – Uppsala, 1963; 218 pp. [*447*]

[740] Grömig W. *On a weakly closed subset of the space of τ-smooth measures*. Proc. Amer. Math. Soc. 1974. V. 43. P. 397–401. [*256*]

[741]* Gromov M. *Monotonicity of the volume of intersection of balls*. Lecture Notes in Math. 1987. V. 1267. P. 1–4.

[742] Gromov M. Metric structures for Riemannian and non-Riemannian spaces. Birkhäuser, Boston – Berlin, 1999; xix+585 pp. [*459*]

[743] Gross L. *Harmonic analysis on Hilbert space*. Mem. Amer. Math. Soc. 1963. V. 46. P. 1–62. [*449*]

[744] Grothendieck A. *Sur les applications linéaires faiblement compactes d'espaces du type $C(K)$*. Canad. J. Math. 1953. V. 5. P. 129–173. [*241, 244*]

[745] Grothendieck A. Topological vector spaces. Gordon and Breach, New York, 1973; 245 pp. [*262*]

[746]* 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.

[747] Gruenhage G., Pfeffer W.F. *When inner regularity of Borel measures implies regularity*. J. London Math. Soc (2). 1978. V. 17. P. 165–171. [*155*]

[748] Gryllakis C. *Products of completion regular measures*. Proc. Amer. Math. Soc. 1988. V. 103, №2. P. 563–568. [*450*]

[749] Gryllakis C., Grekas S. *On products of Radon measures*. Fund. Math. 1999. V. 159, №1. P. 71–84. [*444*]

[750] Gryllakis C., Koumoullis G. *Completion regularity and τ-additivity of measures on product spaces*. Compos. Math. 1990. V. 73. P. 329–344. [*134*]

[751] Grzegorek E. *Solution to a problem of Banach on σ-fields without continuous measures*. Bull. Acad. Polon. Sci. 1980. V. 28. P. 7–10. [*133*]

[752] 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). [**I**: *425;* **II**: *453*]

[753]* Günzler H. Integration. Bibliogr. Inst., Mannheim, 1985; 392 S.

[754]* Gut A. Probability: a graduate course. Springer, Berlin – New York, 2005; xxiv+603 pp.

[755]* 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.

[756]* Guzmán M. de. Real variable methods in Fourier analysis. North-Holland, Amsterdam, 1981; 392 pp.

[757]* Guzmán M. de, Rubio B. Integración: teoría y técnicas. Editorial Alhambra, Madrid, 1979; x+171 pp.

[758] Haar A. *Der Massbegriff in der Theorie der kontinuierlichen Gruppen*. Ann. Math. (2). 1933. V. 34. P. 147–169. [*442, 460*]

[759]* Haaser N.B., Sullivan J.A. Real analysis. Dover Publications, New York, 1991; x+341 pp.

[760]* 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).

[761] Hackenbroch W. *Dilated sections and extremal preimage measures.* Arch. Math. 1984. V. 43, №5. P. 434–439. [*311*]

[762]* Hackenbroch W. Integrationstheorie. Eine Einführung in die Integrationstheorie und ihre Anwendungen. Teubner, Stuttgart, 1987; 144 S.

[763]* Hadwiger H. Vorlesungen über Inhalt, Oberfläche und Isoperimetrie. Springer-Verlag, Berlin – Heidelberg, 1957; xiii+312 S.

[764] Haezendonck J. *Abstract Lebesgue–Rokhlin spaces.* Bull. Soc. Math. Belg. 1973. V. 25, №3. P. 243–258. [*459*]

[765]* Hahn H. *Über den Fundamentalsatz der Integralrechnung.* Monatsh. Math. Phys. 1905. B. 16. S. 161–166.

[766]* 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.

[767]* 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.

[768]* Hahn H. *Über eine Verallgemeinerung der Riemannschen Integraldefinition.* Monatsh. Math. Phys. 1915. B. 26. S. 3–18.

[769]* 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.

[770] Hahn H. Theorie der reellen Funktionen. B. I. Springer, Berlin, 1921; vii+614 S. [**I**: *411, 418, 421, 429;* **II**: *160*]

[771]* Hahn H. *Über Folgen linearer Operationen.* Monatsh. Math. Phys. 1922. B. 32. S. 3–88.

[772] Hahn H. *Über die Multiplikation total-additiver Mengenfunktionen.* Annali Scuola Norm. Super. Pisa (2). 1933. V. 2. P. 429–452. [**I**: *419;* **II**: *452*]

[773]* 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.

[774]* Hahn H., Rosenthal A. Set functions. Univ. New Mexico Press, Albuquerque, New Mexico, 1948; 324 pp.

[775]* Hajłasz P. *Change of variables formula under minimal assumptions.* Colloq. Math. 1993. V. 64, №1. P. 93–101.

[776] Hall P., Heyde C.C. Martingale limit theory and its applications. Academic Press, New York, 1980; 308 pp. [*461*]

[777] Halmos P. *The decomposition of measures.* Duke Math. J. 1941. V. 8. P. 386–392. [*461*]

[778] Halmos P.R. *On a theorem of Dieudonné.* Proc. Nat. Acad. Sci. U.S.A. 1949. V. 35, №1. P. 38–42. [*461*]

[779] Halmos P. Measure theory. Van Nostrand, New York, 1950; xi+304 pp. [**I**: *v, 180, 412;* **II**: *44, 308, 442, 444, 449, 460*]

[780] Halmos P. Lectures on ergodic theory. Chelsea, New York, 1960; vii+101 pp. [*391*]

[781] Halmos P.R., Neumann J. von. *Operator methods in classical mechanics. II.* Ann. Math. 1942. V. 43, №2. P. 332–350. [*444, 459*]

[782]* 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.

[783] Hammersley J.M. *An extension of the Slutzky–Fréchet theorem.* Acta Math. 1952. V. 87. P. 243–257. [*199*]

[784] Hanin L.G. *An extension of the Kantorovich norm.* Monge–Ampère equations: applications to geometry and optimization, pp. 113–130. Contemp. Math., 226, Amer. Math. Soc., Providence, Rhode Island, 1999. [*457*]

[785]* Hanisch H., Hirsch W.M., Renyi A. *Measures in denumerable spaces.* Amer. Math. Monthly. 1969. V. 76, №5. P. 494–502.

[786]* Hanner O. *On the uniform convexity of L^p and l^p.* Ark. Math. 1956. B. 3. S. 239–244.

[787]* Hardy G.H., Littlewood J.E., Pólya G. Inequalities. 2d ed. Cambridge, at the University Press, 1952; xii+324 pp.

[788]* Hardy G.H., Rogosinski W.W. Fourier series. Cambridge University Press, Cambridge, 1944; 100 pp.

[789] Hart J.E., Kunen K. *Orthogonal continuous functions.* Real Anal. Exchange. 1999–2000. V. 25, №2. P. 653–659. [*158*]

[790] Hartman S., Marczewski E. *On the convergence in measure.* Acta Sci. Math. Szeged. 1950. V. 12. P. 125–131. [*254*]

[791] Hartman S., Marczewski E., Ryll-Nardzewski C. *Théorèmes ergodiques et leurs applications.* Colloq. Math. 1950. V. 2. P. 109–123. [*463*]

[792]* Hartman S., Mikusiński J. The theory of Lebesgue measure and integration. Pergamon Press, New York, 1961; 176 pp.

[793]* 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.)

[794]* Hausdorff F. Grundzüge der Mengenlehre. Leipzig, 1914; 476 S.

[795]* Hausdorff F. *Die Mächtigkeit der Borelschen Mengen.* Math. Ann. 1916. B. 77, №3. S. 430–437.

[796]* Hausdorff F. *Dimension und äusseres Mass.* Math. Ann. 1919. B. 79. S. 157–179.

[797] Hausdorff F. Mengenlehre. 3e Aufl. Walter de Gruyter, Berlin, 1935; 307 S. English transl.: Set theory. 2nd ed. Chelsea, New York, 1962; 352 pp. [*439*]

[798]* Hausdorff F. Gesammelte Werke. B. 1–8. Springer, Berlin, 2001–2005.

[799] Haviland E.K. *On the theory of absolutely additive distribution functions.* Amer. J. Math. 1934. V. 56. P. 625–658. [*453*]

[800]* Hawkins T. Lebesgue's theory of integration. Its origins and development. University of Wisconsin Press, Madison – London, 1970; xv+227 pp.

[801] Haydon R. *On compactness in spaces of measures and measure compact spaces.* Proc. London Math. Soc. 1974. V. 29, №1. P. 1–16. [*256, 456*]

[802] Haydon R. *On dual L^1-spaces and injective bidual Banach spaces.* Israel J. Math. 1978. V. 31, №2. P. 142–152. [*136*]

[803] Hayes C.A., Pauc C.Y. Derivation and martingales. Springer, Berlin – New York, 1970; vii+203 pp. [**I**: *438;* **II**: *461*]

[804] Hazod W., Siebert E. Stable probability measures on Euclidean spaces and on locally compact groups. Structural properties and limit theorems. Kluwer Academic Publ., Dordrecht, 2001; xviii+612 p. [*451*]

[805] Hebert D.J., Lacey H.E. *On supports of regular Borel measures.* Pacif. J. Math. 1968. V. 27. P. 101–118. [*136, 450*]

[806]* Heinonen J. Lectures on analysis on metric spaces. Springer, New York, 2001; x+140 pp.

[807]* Helgason S. The Radon transform. 2nd ed. Birkhäuser, Boston – Basel, 1999; xiv+188 pp.

[808]* Hellinger E. *Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen.* J. Reine Angew. Math. 1909. B. 136. S. 210–271.

[809] Helly E. *Über lineare Funktionaloperationen.* Sitz. Akad. Wiss. Wien, Math.-naturwiss. Kl. IIa. 1912. B. 12. S. 265–297. [*452*]

[810] Hengartner W., Theodorescu R. Concentration functions. Academic Press, New York – London, 1973; xii+139 pp. [*257*]
[811] Hennequin P.-L., Tortrat A. Théorie des probabilités et quelques applications. Masson et Gie, Paris, 1965; viii+457 pp. [**I**: *413;* **II**: *444, 453, 462*]
[812] Henry J.P. *Prolongements de mesures de Radon.* Ann. Inst. Fourier (Grenoble). 1969. V. 19, №1. P. 237–247. [*84, 85, 443*]
[813]* Henstock R. Theory of integration. Butterworths, London, 1963; ix+168 pp.
[814]* Henstock R. *A Riemann–type integral of Lebesgue power.* Canad. J. Math. 1968. V. 20. P. 79–87.
[815]* Henstock R. Lectures on the theory of integration. World Sci., Singapore, 1988; xiv+206 pp.
[816]* Henstock R. *A short history of integration theory.* Southeast Asian Bull. Math. 1988. V. 12, №2. P. 75–95.
[817]* Henstock R. The general theory of integration. Clarendon Press, Oxford University Press, New York, 1991; xii+262 pp.
[818]* Henze E. Einführung in die Maßtheorie. 2e Aufl. Bibliogr. Inst., Mannheim, 1985; 235 S.
[819] Herer W. *Stochastic bases in Fréchet spaces.* Demonstr. Math. 1981. V. 14. P. 719–724. [*120*]
[820]* Herglotz G. *Über Potenzreichen mit positiven reellen Teil im Einheitskreis.* Leipziger Berichte. 1911. B. 63. S. 501–511.
[821] Herz C.S. *The spectral theory of bounded functions.* Trans. Amer. Math. Soc. 1960. V. 94. P. 181–232. [*332*]
[822]* Hesse C. Angewandte Wahrscheinlichkeitstheorie. Ein fundierte Einführung mit über 500 realitätsnahen Beispielen und Aufgaben. Vieweg, Wiesbaden, 2003; x+505 S.
[823]* Heuser H. Lehrbuch der Analysis. B. 2. 11e Aufl., Teubner, Stuttgart, 2000; 737 S.
[824] Hewitt E. *Linear functionals on spaces of continuous functions.* Fund. Math. 1950. V. 37. P. 161–189. [*447*]
[825] Hewitt E., Ross K.A. Abstract harmonic analysis. V. 1. Academic Press, New York; Springer-Verlag, Berlin – Heidelberg, 1963; viii+519 pp. [*306, 308, 320, 448, 451, 460*]
[826] Hewitt E., Savage L.J. *Symmetric measures on Cartesian products.* Trans. Amer. Math. Soc. 1955. V. 80. P. 470–501. [*464*]
[827]* Hewitt E., Stromberg K. Real and abstract analysis. Englewood Springer, Berlin – New York, 1975; x+476 pp.
[828] Heyer H. Probability measures on locally compact groups. Springer-Verlag, Berlin – New York, 1977; x+531 pp. [*451*]
[829] Heyer H. Structural aspects in the theory of probability. World Sci., Singapore, 2004; 400 pp. [*451*]
[830]* 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.
[831] Hildebrandt T.H. Introduction to the theory of integration. Academic Press, New York – London, 1963; ix+385 pp. [**I**: *414;* **II**: *446*]
[832]* Hille E. Methods in classical and functional analysis. Addison-Wesley, Reading – London, 1972; ix+486 pp.
[833]* Hinderer K. Grundbegriffe der Wahrscheinlichkeitstheorie. 3e Aufl. Springer, Berlin, 1985; vii+247 S.
[834] Hirsch F., Lacombe G. Éléments d'analyse fonctionnelle. Masson, Paris, 1997; x+340 pp. [*446*]
[835] Hlawka E. *Folgen auf kompakten Räumen.* Abh. Math. Sem. Univ. Hamburg. 1956. B. 20. S. 223–241. [*237, 258*]
[836] Hlawka E. Theorie der Gleichverteilung. Bibliogr. Inst., Mannheim, 1979; x+142 S. [*237*]

[837]* 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.
[838]* 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.
[839]* Hoffman K. Analysis in Euclidean space. Prentice-Hall, Englewood Cliffs, 1975; 432 pp.
[840]* Hoffmann D., Schäfke F.W. Integrale. Bibliogr. Inst., Mannheim, 1992; xiv+333 S.
[841] Hoffmann-Jørgensen J. The theory of analytic spaces. Aarhus Various Publ. Series. №10. 1970; vi+314 pp. [*27, 29, 46, 440, 441, 455*]
[842] Hoffmann-Jørgensen J. *Existence of conditional probabilities.* Math. Scand. 1971. V. 28, №2. P. 257–264. [*462*]
[843] Hoffmann-Jørgensen J. *A generalization of the strict topology.* Math. Scand. 1972. V. 30, №2. P. 313–323. [*220, 455*]
[844] Hoffmann-Jørgensen J. *Weak compactness and tightness of subsets of $M(X)$.* Math. Scand. 1972. V. 31, №1. P. 127–150. [*215, 217, 254, 456*]
[845] Hoffmann-Jørgensen J. Probability in Banach spaces. Lecture Notes in Math. 1976. V. 598. P. 1–186. [*46, 447*]
[846] Hoffmann-Jørgensen J. *Integrability of seminorms, the $0-1$ law and the affine kernel for product measures.* Studia Math. 1977. V. 61, №2. P. 137–159. [*410*]
[847] Hoffmann-Jørgensen J. Stochastic processes on Polish spaces. Aarhus Universitet, Matematisk Institut, Aarhus, 1991; ii+278 pp. [*46, 453*]
[848] Hoffmann-Jørgensen J. Probability with a view toward statistics. V. I, II. Chapman & Hall, New York, 1994; 589 pp., 533 pp. [**I**: *95, 414, 421;* **II**: *56*]
[849] Högnäs G., Mukherjea A. Probability measures on semigroups. Convolution products, random walks, and random matrices. Plenum Press, New York, 1995; xii+388 pp. [*451*]
[850]* Holdgrün H.S. Analysis. B 2. Leins-Verl., Göttingen, 2001; x+485 S.
[851] Holický P., Kalenda O. *Descriptive properties of spaces of measures.* Bull. Polish Acad. Sci. Math. 1999. V. 47, №1. P. 37–51. [*227*]
[852] Holický P., Ponomarev S.P., Zajíček L., Zelený M. *Structure of the set of continuous functions with Luzin's property (N).* Real Anal. Exchange. 1998–1999. V. 24 (2). P. 635–656. [*335*]
[853]* Hopf E. *On causality, statistics and probability.* J. Math. Phys. 1934. V. 13. P. 51–102.
[854] Hopf E. Ergodentheorie. Springer, Berlin, 1937; iv+83 S. [**I**: *419;* **II**: *442, 458*]
[855]* Howard E.J., Pfeffer W. *Luzin's theorem for charges.* Proc. Amer. Math. Soc. 2004. V. 132, №3. P. 857–863.
[856] Howroyd J.D. *On dimension and on the existence of sets of finite positive Hausdorff measure.* Proc. London Math. Soc. (3). 1995. V. 70. P. 581–604. [*140*]
[857]* Hu S. Elements of real analysis. Holden-Day, San Francisco, 1967; xii+365 pp.
[858]* Hulanicki A. *Invariant extensions of the Lebesgue measure.* Fund. Math. 1962. V. 51. P. 111–115.
[859]* 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.
[860]* Hunt G.A. Martingales et processus de Markov. Dunod, Paris, 1966; 143 pp.
[861]* Ingleton A.W. Notes on integration. Mathematical Institute, Oxford University, Oxford, 1972; i+90 pp.
[862] Ionescu Tulcea A. *On pointwise convergence, compactness, and equicontinuity in the lifting topology, I.* Z. Wahrscheinlichkeitstheorie verw. Geb. 1973. B. 26. S. 197–205. [*151, 452*]
[863] Ionescu Tulcea A. *On pointwise convergence, compactness, and equicontinuity, II.* Adv. Math. 1974. V. 12. P. 171–177. [*151, 452*]

[864] Ionescu Tulcea A., Ionescu Tulcea C. *On the lifting property. I.* J. Math. Anal. Appl. 1961. V. 3. P. 537–546. [*463*]

[865] Ionescu Tulcea A., Ionescu Tulcea C. *On the lifting property. II. Representation of linear operators on the spaces L_E^r, $1 \le r < \infty$.* J. Math. Mech. 1962. V. 11. P. 773–796. [*462*]

[866] Ionescu Tulcea A., Ionescu Tulcea C. *On the lifting property. IV, Disintegration of measures.* Ann. Inst. Fourier (Grenoble). 1964. V. 14. P. 445–472. [*462*]

[867] Ionescu Tulcea A., Ionescu Tulcea C. Topics in the theory of lifting. Springer, Berlin – New York, 1969; x+188 pp. [*407, 431, 463*]

[868] Ionescu Tulcea C. *Mesures dans les espaces produits.* Atti Accad. Naz. Lincei Rend. Cl. sci. fis., mat. e natur. Roma. (8). 1949. V. 7. P. 208–211. [*463*]

[869]* Ivanov L.D. Variations of sets and functions. Nauka, Moscow, 1975; 352 pp. (in Russian).

[870]* 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.

[871] Ivanov V.V. *Oscillations of averages in the ergodic theorem.* Dokl. Ros. Akad. Nauk. 1996. V. 347, №6. P. 736–738 (in Russian); English transl.: Dokl. Math. 1996. V. 53, №2. P. 263–265. [*463*]

[872] Ivanov V.V. *Geometric properties of monotone functions and the probabilities of random oscillations.* Sibirsk. Mat. Zh. 1996. V. 37, №1. P. 117–150 (in Russian); English transl.: Sib. Math. J. 1996. V. 37, №1. P. 102–129. [*397, 463*]

[873] Iwanik A. *Representation of operators on L^1-spaces by nondirect product measures.* Colloq. Math. 1982. V. 47, №1. P. 91–95. [*174*]

[874] Jackson S., Mauldin R.D. *On the σ-class generated by open balls.* Proc. Cambridge Phil. Soc. 1999. V. 127, №1. P. 99–108. [*61*]

[875] Jacobs K. *On Poincaré's recurrence theorem.* Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, 1965/66). V. II. Part 2. P. 375–404. Univ. California Press, Berkeley, 1967. [*434*]

[876] Jacobs K. Measure and integral. Academic Press, New York, 1978; 575 pp. [**I**: *414;* **II**: *461, 463, 463*]

[877] Jacod J., Mémin J. *Sur un type de convergence intermédiare entre la convergence en loi et la convergence en probabilité.* Lecture Notes in Math. 1981. V. 850. P. 529–546; Rectification: Lecture Notes in Math. 1983. V. 986. P. 509–511. [*249*]

[878] Jakubowski A. *On the Skorohod topology.* Ann. Inst. H. Poincaré Probab. Statist. 1986. V. 22, №3. P. 263–285. [*53*]

[879] Jakubowski A. *The almost sure Skorokhod representation for subsequences in nonmetric spaces.* Theory Probab. Appl. 1997. V. 42, №1. P. 167–174. [*454*]

[880]* Jain P.K., Gupta V.P. Lebesgue measure and integration. Wiley Eastern, New Delhi, 1986; viii+260 pp.

[881]* James R.C. Advanced calculus. Wadsworth, Belmont, 1966; x+646 pp.

[882] Jankoff W. (Yankov V.) *Sur l'uniformisation des ensembles A.* Dokl. Akad. Nauk SSSR. 1941. V. 30. P. 591–592 (in Russian); French transl.: C. R. (Doklady) Acad. Sci. URSS. 1941. V. 30. P. 597–598. [*441*]

[883] Janssen A. *Integration separat stetiger Funktionen.* Math. Scand. 1981. V. 48, №1. P. 68–78. [*130*]

[884] Janssen A. *A survey about zero-one laws for probability measures on linear spaces and locally compact groups.* Lecture Notes in Math. 1984. V. 1064. P. 551–563. [*410*]

[885]* 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. [**I**: *414;* **II**: *446*]

[886] Jayne J. *Structure of analytic Hausdorff spaces.* Mathematika. 1976. V. 23, №2. P. 208–211. [*49, 440*]

[887] Jayne J. *Generation of σ-algebras, Baire sets, and descriptive Borel sets.* Mathematika. 1977. V. 24, №2. P. 241–256. [**I**: *421;* **II**: *44, 46, 62, 440*]

[888] Jayne J., Rogers C.A. *Radon measures on Banach spaces with their weak topologies.* Serdica Math. J. 1995. V. 21, №4. P. 283–334. [*452*]

[889]* Jean R. Mesure et intégration. Les Presses de l'Université du Quebec, Montréal, 1975; 305 pp.

[890]* 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.

[891] Jech T. Set theory. Academic Press, New York, 1978; xi+621 pp. [**I**: *62, 79, 80;* **II**: *331*]

[892]* 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.

[893]* Jeffery R. The theory of functions of a real variable. 2nd ed., University of Toronto Press, 1953; xi+232 pp.

[894]* Jensen J.L.W.V. *Sur les fonctions convexes et les inégalités entres les valeurs moyennes.* Acta. Math. 1906. V. 30. P. 175–193.

[895] Jessen B. *The theory of integration in a space of an infinite number of dimensions.* Acta Math. 1934. V. 63. P. 249–323. [**I**: *419, 429, 435;* **II**: *461*]

[896]* Jessen B. *On the approximation of Lebesgue integrals by Riemann sums.* Ann. Math. (2). 1934. V. 35. P. 248–251.

[897] Jessen B. *On two notions of independent functions.* Colloq. Math. 1947–1948. V. 1. P. 214–215. [*433*]

[898]* Jessen B. Abstrakt Maal- og Integralkteori. Matematisk Forening i København, København, 1947; iii+108 pp.

[899]* Jessen B., Marcinkiewicz J., Zygmund A. *Note on the differentiability of multiple integrals.* Fund. Math. 1935. V. 25. P. 217–234.

[900] Jessen B., Wintner A. *Distribution functions and the Riemann zeta function.* Trans. Amer. Math. Soc. 1935. V. 38. P. 48–88. [**I**: *430;* **II**: *453*]

[901] Jimenez-Guerra P., Rodriguez-Salinas B. *Strictly localizable measures.* Nagoya Math. J. 1982. V. 85. P. 81–86. [*452*]

[902]* Jiménez Pozo M.A. Medida, integración y funcionales. Editorial Pueblo y Educación, Habana, 1989; 504 pp.

[903] Jiřina M. *Conditional probabilities on algebras with countable base.* Czech. Math. J. 1954. V. 4 (79). P. 372–380 (in Russian). [*462*]

[904] Jiřina M. *On regular conditional probabilities.* Czech. Math. J. 1959. V. 9. P. 445–451. [*462*]

[905] Johnson B.E. *Separate continuity and measurability.* Proc. Amer. Math. Soc. 1969. V. 20. P. 420–422. [*129, 163*]

[906] Johnson D.L. *A short proof of the uniqueness of Haar measure.* Proc. Amer. Math. Soc. 1976. V. 55, №1. P. 250–251. [*460*]

[907] Johnson Roy A. *On product measures and Fubini's theorem in locally compact spaces.* Trans. Amer. Math. Soc. 1966. V. 123. P. 112–129. [*127, 444*]

[908] Johnson Roy A. *Measurability of cross section measure of a product Borel set.* J. Austral. Math. Soc. Ser. A. 1979. V. 28. P. 346–352. [*127, 444*]

[909] Johnson Roy A. *Nearly Borel sets and product measures.* Pacif. J. Math. 1980. V. 87, №1. P. 97–109. [*444*]

[910] Johnson Roy A. *Extending the product of two regular Borel measures.* Illinois J. Math. 1980. V. 24, №4. P. 639–644. [*444*]

[911] Johnson Roy A. *Products of two Borel measures.* Trans. Amer. Math. Soc. 1982. V. 269, №2. P. 611–625. [*127, 444*]

[912] Johnson Roy A. *Absolute continuity, singularity and product measures.* Internat. J. Math. Math. Sci. 1982. V. 5, №4. P. 793–807. [*444*]

[913] Johnson Roy A., Wajch E., Wilczyński W. *Metric spaces and multiplication of Borel sets.* Rocky Mountain J. Math. 1992. V. 22, №4. P. 1341–1347. [*444*]

[914] Johnson Roy A., Wilczyński W. *Finite products of Borel measures.* In: Measure theory (Oberwolfach, 1990), B. Pettineo and P. Vetro eds.; Rend. Circ. Mat. Palermo

(2) Suppl. No. 28 (1992), pp. 141–148; Circolo Matematico di Palermo, Palermo, 1992. [*164*]

[915] Johnson Russell A. *Strong liftings that are not Borel liftings.* Proc. Amer. Math. Soc. 1980. V. 80. P. 234–236. [*407*]

[916]* Jones F.B. *Measure and other properties of a Hamel basis.* Bull. Amer. Math. Soc. 1942. V. 48. P. 472–481.

[917]* Jones F. Lebesgue integration on Euclidean space. Jones and Bartlett, Boston, 1993; xvi+588 pp.

[918]* 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.

[919]* Jordan C. Cours d'analyse de l'École Polytechnique. T. 1–3. 2e éd. Paris, 1893–1896; 612 pp., 627 pp., 542 pp.

[920]* Jost J. Postmodern analysis. Springer-Verlag, Berlin, 1998; xviii+353 pp.

[921] Juhász I. Cardinal functions in topology. Math. Cent. Tracts №34, Mathematisch Centrum, Amsterdam, 1971; xiii+149 pp. [*136*]

[922] Kac M. *Sur les fonctions indépendantes.* Studia Math. 1936. V. 6. P. 46–58. [*464*]

[923] Kac M., Steinhaus H. *Sur les fonctions indépendantes. II.* Studia Math. 1936. V. 6. P. 59–66. [*464*]

[924] Kachurovskiĭ A.G. *Rates of convergence in ergodic theorems.* Uspehi Mat. Nauk. 1996. V. 51, №4. P. 73–124 (in Russian); English transl.: Russian Math. Surveys. 1996. V. 51, №4. P. 653–703. [*463*]

[925]* Kaczmarz S., Nikliborc L. *Sur les suites des fonctions convergentes en moyenne.* Fund. Math. 1928. V. 11. P. 151–168.

[926]* Kaczor W.J., Nowak M.T. Problems in mathematical analysis III: integration. Amer. Math. Soc., Providence, Rhode Island, 2003.

[927]* Kadec M.I., Pełczyński A. *Bases, lacunary sequences and complemented subspaces in the spaces L_p.* Studia Math. 1962. T. 21. P. 161–176.

[928]* Kahane C.S. *Evaluating Lebesgue integrals as limits of Riemann sums.* Math. Japon. 1993. V. 38. P. 1073–1076.

[929]* 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.

[930]* Kahane J.-P. *Trois notes sur les ensembles parfaits linéaires.* Enseign. Math. (2). 1969. V. 15. P. 185–192.

[931] Kakosyan A.V., Klebanov L.V., Rachev S.T. Quantitative criteria for convergence of probability measures. Aĭastan, Erevan, 1988; 249 pp. (in Russian). [*456*]

[932] Kakutani S. *Concrete representation of abstract (M)-spaces. (A characterization of the space of continuous functions).* Ann. Math. 1941. V. 42. P. 994–1024. [*447*]

[933] Kakutani S. *Notes on infinite product measures, I, II.* Proc. Imperial Acad. Tokyo. 1943. V. 19. P. 148–151, 184–188. [**I**: *429;* **II**: *448, 450*]

[934]* Kakutani S. *Notes on divergent series and integrals.* Proc. Imper. Acad. Tokyo. 1944. V. 20. P. 74–76.

[935] Kakutani S. *On equivalence of infinite product measures.* Ann. Math. 1948. V. 49. P. 214–224. [*351*]

[936]* Kakutani S. Selected papers. V. 2. Birkhäuser, Boston, 1986.

[937] Kakutani S., Kodaira K. *Über das Haarsche Mass in der lokal bikompakten Gruppe.* Proc. Imper. Acad. Tokyo. 1944. V. 20. P. 444–450. [*308*]

[938]* 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.

[939] Kallenberg O. Foundations of modern probability. 2nd ed. Springer-Verlag, New York, 2002; xx+638 pp. [**I**: *414;* **II**: *262*]

[940] Kallianpur G. *The topology of weak convergence of probability measures.* J. Math. Mech. 1961. V. 10, №6. P. 947–969. [*456*]

[941] Kallianpur G., Ramachandran D. *On the splicing of measures.* Ann. Probab. 1983. V. 11, №3. P. 819–822. [*433*]

[942]* Kamke E. Das Lebesguesche Integral. Eine Einführung in die neuere Theorie der reellen Funktionen. Teubner, Leipzig, 1925; 151 S.

[943]* Kamke E. Das Lebesgue–Stieltjes–Integral. B.G. Teubner Verlagsgesellschaft, Leipzig, 1956; vi+226 S.

[944]* Kampen E.R. van. *Infinite product measures and infinite convolutions.* Amer. J. Math. 1940. V. 62, №2. P. 417–448.

[945]* Kannan R., Krueger C.K. Advanced analysis on the real line. Springer, New York, 1996; ix+259 p.

[946]* 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.

[947] Kanovei V.G. Supplement to the Russian translation of "Handbook of mathematical logic. Part II." Ed. by K. Jon Barwise. P. 273–364. Nauka, Moscow, 1982 (in Russian). [*439*]

[948]* Kanovei V.G. The axiom of choice and the axiom of determinacy. Nauka, Moscow, 1984; 65 pp. (in Russian).

[949] Kanter M. *Linear sample spaces and stable processes.* J. Funct. Anal. 1972. V. 9, №4. P. 441–459. [*149, 451*]

[950] Kanter M. *Random linear functionals and why we study them.* Lecture Notes in Math. 1978. V. 645. P. 114–123. [*149*]

[951] Kantorovitch L.V. *On the translocation of masses.* Dokl. Akad. Nauk SSSR. 1942. V. 37, №7–8. P. 227–229 (in Russian); English transl.: C. R. (Doklady) Acad. Sci. URSS. 1942. V. 37. P. 199–201. [*453*]

[952] Kantorovich L.V., Akilov G.P. Functional analysis. Translated from the Russian. 2nd ed. Pergamon Press, Oxford, 1982; xiv+589 pp. (Russian ed.: Nauka, Moscow, 1977). [*453*]

[953] Kantorovich L.V., Rubinshtein (Rubinštein) G.Sh. *On a functional space and certain extremum problems.* Dokl. Akad. Nauk SSSR. 1957. V. 115, №6. P. 1058–1061 (in Russian). [*453*]

[954] Kantorovich L.V., Rubinstein (Rubinšteĭn) G.Sh. *On a space of completely additive functions.* Vestnik Leningrad. Univ. 1958. №7(2). P. 52–59 (in Russian). [*453, 456, 457*]

[955]* Kantorovitz S. Introduction to modern analysis. Oxford University Press, New York, 2003; xii+434 pp.

[956]* Kappos D.A. Strukturtheorie der Wahrscheinlichkeitsfelder und -räume. Springer, Berlin, 1960; 136 S.

[957]* Kappos D.A. Probability algebras and stochastic spaces. Academic Press, New York, 1969; x+267 pp.

[958]* Karr A.F. Probability. Springer-Verlag, New York, 1993; xxii+282 pp.

[959]* 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).

[960] Katětov M. *Measures in fully normal spaces.* Fund. Math. 1951. V. 38. P. 73–84. [*451*]

[961] Katok A.B., Stepin A.M. *Metric properties of measure preserving homeomorphisms.* Uspehi Mat. Nauk. 1970. V. 25. P. 193–220 (in Russian); English transl.: Russian Math. Surveys. 1970. V. 25. P. 191–220. [*459*]

[962] Kats M.P. *Continuity of universally measurable linear mappings.* Sibirsk. Mat. Zh. 1982. V. 23, №3. P. 83–90; correction: ibid. 1983. V. 24, №3. P. 217 (in Russian); English transl.: Sib. Math. J. 1982. V. 23, №3. P. 358–364. [*168*]

[963]* Katznelson Y., Stromberg K. *Everywhere differentiable, nowhere monotone, functions.* Amer. Math. Monthly. 1974. V. 81. P. 349–354.

[964]* Kaufman R.P., Rickert N.W. *An inequality concerning measures.* Bull. Amer. Math. Soc. 1966. V. 72. P. 672–676.

[965]* Kaufman R., Wu J.-M. *Two problems on doubling measures.* Rev. Mat. Iberoamer. 1995. V. 11. P. 527–545.

[966] Kawabe J. *Convergence of compound probability measures on topological spaces.* Colloq. Math. 1994. V. 67, №2. P. 161–176. [*258*]

[967]* Kawata T. Fourier analysis in probability theory. Academic Press, New York, 1972; xii+668 pp.

[968] Kechris A.S. Classical descriptive set theory. Springer, Berlin – New York, 1995; xviii+402 pp. [*37, 262, 430, 440*]

[969]* Keleti T. *Density and covering properties of intervals in* $\mathbb{R}^n$. Mathematika. 2000. V. 47. P. 229–242.

[970] Keleti T., Preiss D. *The balls do not generate all Borel sets using complements and countable disjoint unions.* Math. Proc. Cambridge Philos. Soc. 2000. V. 128, №3. P. 539–547. [*61*]

[971] Keller O.H. *Die Homoiomorphie der kompakten konvexen Mengen in Hilbert'schen Raum.* Math. Ann. 1931. B. 105. S. 748–758. [*83*]

[972] Kellerer H.G. *Masstheoretische Marginalprobleme.* Math. Ann. 1964. B. 153. S. 168–198. [*461*]

[973] Kellerer H.G. *Schnittmass-Funktionen in mehrfachen Produkträumen.* Math. Ann. 1964. B. 155. S. 369–391. [*461*]

[974] Kellerer H.G. *Baire sets in product spaces.* Lecture Notes in Math. 1980. V. 794. P. 38–44. [*45*]

[975] Kellerer H.G. *Duality theorems for marginal problems.* Z. Wahrsch. verw. Geb. 1984. B. 67, №4. S. 399–432. [*461*]

[976] Kellerer H.G. *Continuous transformation of Baire measures into Lebesgue measure.* Proc. Amer. Math. Soc. 2002. V. 130, №8. P. 2305–2309. [*458*]

[977] Kelley J.D. *The increasing sets lemma, and the approximation of analytic sets from within by compact sets, for the measures generated by method III.* Proc. London Math. Soc. (2). 1974. V. 2. P. 29–43. [*450*]

[978]* Kelley J.L. *Measures on Boolean algebras.* Pacif. J. Math. 1959. V. 9, №4. P. 1165–1177.

[979]* Kelley J.L., Srinivasan T.P. Measure and integral. V. 1. Springer-Verlag, New York – Berlin, 1988; x+150 pp.

[980] Kemperman J.H.B., Maharam D. $\mathbf{R}^c$ *is not almost Lindelöf.* Proc. Amer. Math. Soc. 1970. V. 24, №4. P. 772–773. [*131*]

[981] Kendall D.G. *Separability and measurability for stochastic processes: a survey.* Stochastic analysis (a tribute to the memory of Rollo Davidson), pp. 415–443. Wiley, London, 1973. [*448*]

[982]* Kenyon H., Morse A.P. Web derivatives. Mem. Amer. Math. Soc., No. 132. Amer. Math. Soc., Providence, Rhode Island, 1973; xiii+177 pp.

[983]* Kestelman H. Modern theories of integration. Oxford, 1937. 2nd revised ed. Dover Publications, New York, 1960; x+309 pp.

[984] Khafizov M.U. *On the kernel of a Radon measure in a Banach space.* Vestnik Moskov. Univ. Ser. I Mat. Mekh. 1993. №1. P. 3–8 (in Russian); English transl.: Moscow Univ. Math. Bull. 1993. V. 48, №1. P. 1–5. [*451*]

[985]* 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).

[986]* Kharazishvili A.B. Problems of set theory and measure theory. Tbilis. Gos. Univ., Tbilisi, 1978; 178 pp. (in Russian).

[987]* Kharazishvili A.B. Invariant extensions of the Lebesgue measure. Tbilis. Gos. Univ., Tbilisi, 1983; 204 pp. (in Russian).

[988] Kharazishvili A.B. Topological aspects of measure theory. Naukova Dumka, Kiev, 1984; 120 pp. (in Russian). [**I**: *80;* **II**: *46, 450, 451*]

[989]* Kharazishvili A.B. Vitali's theorem and its generalizations. Tbilis. Gos. Univ., Tbilisi, 1991; 108 pp. (in Russian).

[990] Kharazishvili A.B. *Some problems in measure theory.* Colloq. Math. 1991. V. 62. P. 197–220. [*451*]

[991]* Kharazishvili A.B. Transformation groups and invariant measures. Set-theoretical aspects. World Sci., River Edge, 1998; viii+260 pp.

[992] Kharazishvili A.B. Strange functions in real analysis. Marcel Dekker, New York, 2000; viii+297 pp. [**I**: *80, 91, 211;* **II**: *60*]

[993]* Kharazishvili A.B. Nonmeasurable sets and functions. Elsevier, Amsterdam, 2004; xi+337 pp.

[994]* Khintchine A. *Sur la dérivation asymptotique.* C. R. Acad. Sci. Paris. 1917. T. 164. P. 142–144.

[995]* 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.

[996] Khinchin A.Ja. *Zur Birckhoffs Lösung des Ergodenproblem.* Math. Ann. 1933. B. 107. S. 485–488. [*463*]

[997] Khinchin A.Ya. Continued fractions. Translated from the Russian. Dover, Mineola, 1997; xii+95 pp. (Russian ed.: Moscow, 1978). [*431*]

[998] Kindler J. *A simple proof of the Daniell–Stone representation theorem.* Amer. Math. Monthly. 1983. V. 90, №6. P. 396–397. [*166*]

[999]* Kindler J. *A Mazur–Orlicz type theorem for submodular set functions.* J. Math. Anal. Appl. 1986. V. 120, №2. P. 533–546.

[1000]* Kindler J. *Supermodular and tight set functions.* Math. Nachr. 1987. B. 134. S. 131–147.

[1001]* Kindler J. *Sandwich theorems for set functions.* J. Math. Anal. Appl. 1988. V. 133, №2. P. 529–542.

[1002]* Kingman J.F.C., Taylor S.J. Introduction to measure and probability. Cambridge University Press, London – New York – Ibadan, 1966; x+401 pp.

[1003]* 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).

[1004] Kirk R.B. *Measures in topological spaces and B-compactness.* Indag. Math. 1969. V. 31. P. 172–183. [*131*]

[1005] Kirk R.B. *Topologies on spaces of Baire measures.* Bull. Amer. Math. Soc. 1973. V. 79, №3. P. 542–545. [*455*]

[1006] Kirk R.B. *Complete topologies on spaces of Baire measures.* Trans. Amer. Math. Soc. 1973. V. 184. P. 1–21. [*455*]

[1007] Kisyński J. *On the generation of tight measures.* Studia Math. 1968. V. 30. P. 141–151. [*443*]

[1008]* Kisyński J. *Remark on strongly additive set functions.* Fund. Math. 1968. V. 63. P. 327–332.

[1009]* Klambauer G. Real analysis. American Elsevier, New York, 1973; 436 pp. [**I**: *414;* **II**: *446*]

[1010]* Klei H.-A., Miyara M. *Une extension du lemme de Fatou.* Bull. Sci. Math. 1991. V. 115, №2. P. 211–221.

[1011]* Klimkin V.M. Introduction to the theory of set functions. Saratov. Gos. Univ., Kuibyshev, 1989; 210 pp. (in Russian).

[1012]* Kluvánek I., Knowles G. Vector measures and control systems. North-Holland, Amsterdam – Oxford; American Elsevier, New York, 1976; ix+180 pp.

[1013] Knothe H. *Contributions to the theory of convex bodies.* Michigan Math. J. 1957. V. 4. P. 39–52. [*418*]

[1014] Knowles J. *On the existence of non-atomic measures*. Mathematika. 1967. V. 14. P. 62–67. [*136, 317*]

[1015] Knowles J. *Measures on topological spaces*. Proc. London Math. Soc. 1967. V. 17. P. 139–156. [*113, 135*]

[1016]* Kodaira S., Kakutani S. *A nonseparable translation-invariant extension of the Lebesgue measure space*. Ann. Math. (2). 1950. V. 52. P. 574–579.

[1017] Kolesnikov A.V. *On the topological properties of the Skorokhod space*. Teor. Verojatn. i Primen. 1998. V. 43, №4. P. 781–786 (in Russian); English transl.: Theory Probab. Appl. 1999. V. 43, №4. P. 640–644. [*53*]

[1018] Kolesnikov A.V. *On continuous images of Lebesgue measure*. Matem. Zametki. 1999. V. 65, №5. P. 790–793 (in Russian); English transl.: Math. Notes. 1999. V. 65, №5-6. P. 665–668. [*289*]

[1019] Kolesnikov A.V. *Convexity inequalities and optimal transport of infinite-dimensional measures*. J. Math. Pures Appl. 2004. V. 83, №11. P. 1373–1404. [*236*]

[1020]* 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*, [**1032**, V. 1, p. 13–14].

[1021]* 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*, [**1032**, V. 1, p. 15–32].

[1022] Kolmogorov A.N. *General measure theory and probability calculus*. Trudy Kom. Akad. Matem. 1929. T. 1. P. 8–21 (in Russian); English transl.: [**1032**, V. 2, p. 48–59]. [**I**: *419;* **II**: *459*]

[1023]* Kolmogoroff A. *Untersuchungen über den Integralbegriff*. Math. Ann. 1930. B. 103. S. 654–696; English transl.: *Studies on the concept of integral*, [**1032**, V. 1, p. 100–143].

[1024]* 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*, [**1032**, V. 1, p. 147–150].

[1025] Kolmogoroff A. *Beiträge zur Maßtheorie*. Math. Ann. 1933. B. 107. S. 351–366; English transl.: *On measure theory*, [**1032**, V. 1, p. 161–180]. [**I**: *248;* **II**: *458*]

[1026] 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. [**I**: *411, 419, 429;* **II**: *447, 461*]

[1027] Kolmogoroff A. *La transformation de Laplace dans les espaces linéaires*. C. R. Acad. Sci. Paris. 1935. T. 200. P. 1717–1718; English transl.: *Laplace transformation in linear spaces*, [**1032**, V. 1, p. 194–195]. [*448*]

[1028]* Kolmogorov A.N. Lectures on a course of Analysis–III. Moscow Univ., Moscow, 1946–1947 (in Russian).

[1029]* 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).

[1030] Kolmogorov A.N. *A theorem on convergence of conditional expectations and some of its applications*. A lecture at the First Congress of the Hungarian Mathematicians, pp. 367–376. Akadémiai Kiado, Budapest, 1950 (in Russian). [*461*]

[1031] Kolmogorov A.N. *A note on the papers of R.A. Minlos and V. Sazonov*. Teor. Verojant. i Primen. 1959. V. 4, №2. P. 237–239 (in Russian); English transl.: Theory Probab. Appl. 1959. V. 4. P. 221–223. [*124, 449*]

[1032]* Kolmogorov A.N. Selected works. V. I,II. Kluwer Academic Publ., Dordrecht, 1991, 1992.

[1033]* 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. (Russian ed.: Elements of the theory of functions and functional analysis, Moscow, 1954, 1960).

[1034] Kolmogoroff A., Prochorow Ju.W. *Zufällige Funktionen und Grenzverteilugssätze.* Bericht über die Tagung Wahrscheinlichkeitsrechnung und mathematische Statistik, Berlin. 1956. S. 113–126 (English transl.: *Random functions and limit theorems*, [**1032**, V. 2, p. 442–458]). [*189*]

[1035]* Kölzow D. Differentiation von Maßen. Lecture Notes in Math. V. 65. Springer-Verlag, Berlin – New York, 1968; 102 S.

[1036]* Komlós J. *A generalization of a problem of Steinhaus.* Acta Math. Acad. Sci. Hungar. 1967. V. 18. P. 217–229.

[1037]* König H. Measure and integration. An advanced course in basic procedures and applications. Springer–Verlag, Berlin, 1997; xxii+260 pp.

[1038]* Königsberger K. Analysis 2. 4e Aufl. Springer, Berlin, 2002; xii+459 S.

[1039]* Korevaar J. Mathematical methods. V. I: linear algebra, normed spaces, distributions, integration. Academic Press, London, 1968; x+505 pp.

[1040]* Körner T.W. *Besicovitch via Baire.* Studia Math. 2003. V. 158, №1. P. 65–78.

[1041] Korovkin P.P. *A generalization of the D.Th. Egoroff theorem.* Dokl. Akad. Nauk SSSR. 1947. V. 58, №7. P. 1265–1267 (in Russian). [*450*]

[1042]* 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).

[1043] Koumoullis G. *On perfect measures.* Trans. Amer. Math. Soc. 1981. V. 264, №2. P. 521–537. [*444*]

[1044] Koumoullis G. *Some topological properties of spaces of measures.* Pacif. J. Math. 1981. V. 96, №2. P. 419–433. [*228, 230, 256, 455*]

[1045] Koumoullis G. *Perfect, u-additive measures and strict topologies.* Illinois J. Math. 1982. V. 26, №3. P. 466–478. [*444*]

[1046] Koumoullis G. *On the almost Lindelöf property in products of separable metric spaces.* Compos. Math. 1983. V. 48, №1. P. 89–100. [*131*]

[1047] Koumoullis G. *Cantor sets in Prohorov spaces.* Fund. Math. 1984. V. 124, №2. P. 155–161. [*456*]

[1048] Koumoullis G. *Topological spaces containing compact perfect sets and Prohorov spaces.* Topol. Appl. 1985. V. 21, №1. P. 59–71. [*456*]

[1049] Koumoullis G., Prikry K. *The Ramsey property and measurable selections.* J. London Math. Soc. (2). 1983. V. 28. P. 203–210. [*137*]

[1050] Koumoullis G., Prikry K. *Perfect measurable spaces.* Ann. Pure Appl. Logic. 1986. V. 30, №3. P. 219–248. [*444*]

[1051] Koumoullis G., Sapounakis A. *Two countability properties of sets of measures.* Mich. Math. J. 1984. V. 31, №1. P. 31–47. [*230, 231*]

[1052]* Kovan'ko A.S. The Lebesgue integral. Lvov. Univ., Lvov, 1951; 204 pp. (in Russian).

[1053]* Kowalsky H.-J. Vektoranalysis. II. Walter de Gruyter, Berlin – New York, 1976; 251 S.

[1054] Kozlov V.V., Treschev D.V. *On new forms of the ergodic theorem.* J. Dynam. Control Syst. 2003. V. 9, №3. P. 449–453. [*395*]

[1055] Krasnosel'skiĭ M.A. Topological methods in the theory of nonlinear integral equations. Translated from the Russian. Macmillan, New York, 1964; xi+395 pp. (Russian ed.: Moscow, 1956). [*137*]

[1056]* 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).

[1057]* Krée P. Intégration et théorie de la mesure. Une approche géométrique. Cours et exercices corrigés. Ellipses, Paris, 1997; x+211 pp.

[1058] Krengel U. Ergodic theorems. Walter de Gruyter, Berlin – New York, 1985; 357 pp. [*391*]

[1059] Krickeberg K. *Mischende Transformationen auf Mannigfaltigkeiten unendlichen Masses.* Z. Wahrscheinlichkeitstheorie verw. Geb. 1967. B. 7. S. 235–247. [*323*]

[1060] Krickeberg K. *Ein Isomorphiesatz über topologische Maßräume.* Math. Nachr. 1968. B. 37. S. 59–66. [*323*]

[1061]* Krieger H.A. Measure-theoretic probability. University Press of America, Lanham, 1980; xi+382 pp.

[1062]* Kripke B. Introduction to analysis. W.H. Freeman, San Francisco – London, 1968; viii+274 pp.

[1063] Kruglov V.M. Supplementary chapters of probability theory. Vysš. Škola, Moscow, 1984; 264 pp. (in Russian). [*448, 451, 453*]

[1064]* 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.

[1065] Krupa G., Zięba W. *Strong tightness as a condition of weak and almost sure convergence.* Comment. Math. Univ. Carol. 1996. V. 37, №3. P. 641–650. [*173*]

[1066] Krylov N.V. *On regularity of conditional probabilities for random processes.* Teor. Verojatn. i Primen. 1973. V. 18, №1. P. 151–155 (in Russian); English transl.: Theory Probab. Appl. 1973. V. 18. P. 150–153. [*429*]

[1067] Krylov N.V. *On SPDEs and superdiffusions.* Ann. Probab. 1997. V. 25. P. 1789–1809. [*454*]

[1068] Kubokawa Y. *Coverable Radon measures in topological spaces with covering properties.* Colloq. Math. 1996. V. 70, №1. P. 13–23. [*451*]

[1069] Kucia A. *Scorza Dragoni type theorems.* Fund. Math. 1991. V. 138, №3. P. 197–203. [*137*]

[1070]* 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).

[1071]* 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).

[1072]* 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).

[1073] Kuelbs J. *Some results for probability measures on linear topological vector spaces with an application to Strassen's LogLog Law.* J. Funct. Anal. 1973. V. 14, №1. P. 28–43. [*448*]

[1074] Kuipers L., Niederreiter H. Uniform distribution of sequences. Wiley, New York, 1974; xiv+390 pp. [*237*]

[1075] Kulakova V.G. *On regularity of conditional probabilities.* Vestnik Leningrad. Univ. 1976. №1(1). P. 16–20 (in Russian). [*462*]

[1076]* Kuller R.G. Topics in modern analysis. Prentice-Hall, Englewood Cliffs, 1969; vii+296 p.

[1077] Kunen K. *A compact L-space under CH.* Topol. Appl. 1981. V. 12. P. 283–287. [*136*]

[1078] Kunen K., Mill J. van. *Measures on Corson compact spaces.* Fund. Math. 1995. V. 147, №1. P. 61–72. [*449, 452*]

[1079] Kunugui K. *Sur un théorème d'existence dans la théorie des ensembles projectifs.* Fund. Math. 1937. V. 29. P. 167–181. [*66*]

[1080] Kuo H. Gaussian measures in Banach spaces. Lecture Notes in Math. V. 463. Springer-Verlag, Berlin – New York, 1975; vi+224 pp. [*447*]

[1081] Kupka J., Prikry K. *The measurability of uncountable unions.* Amer. Math. Monthly. 1984. V. 91. P. 85–97. [*137*]

[1082] Kuratowski K. Topology. V. 1,2. Academic Press, New York – London, Polish Sci. Publ., Warsaw, 1966, 1968; xx+560 pp., xiv+608 pp. [*1, 8, 12, 27, 56, 61, 288, 439*]

[1083] 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. [**I**: *78, 79;* **II**: *50*]

[1084] Kuratowski K., Ryll-Nardzewski C. *A general theorem on selectors*. Bull. Acad. Polon. Sci. 1965. V. 13. P. 397–403. [*441*]

[1085] Kuratowski C., Szpilrajn E. *Sur les cribles fermés et leurs applications*. Fund. Math. 1932. V. 18. P. 160–170. [*441*]

[1086]* 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.

[1087]* Kurzweil J. *Generalized ordinary differential equations and continuous dependence on a parameter*. Czech. Math. J. 1957. V. 7(82). P. 418–446.

[1088]* Kurzweil J. Nichtabsolut konvergente Integrale. B.G. Teubner Verlagsgesellschaft, Leipzig, 1980; 184 S.

[1089]* Kurzweil J. Integration between the Lebesgue integral and the Henstock–Kurzweil integral. Imperial College Press, London, 2002; 140 pp.

[1090]* 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).

[1091] Kusuoka S., Nakayama T. *On a certain metric on the space of pairs of a random variable and a probability measure*. J. Math. Sci. Univ. Tokyo. 2001. V. 8, №2. P. 343–356. [*456*]

[1092]* Kuttler K. Basic analysis. Rinton Press, Princeton, 2001; xiv+541 pp.

[1093] Kwapień S. *Complément au théorème de Sazonov–Minlos*. C. R. Acad. Sci. Paris Sér. A-B. 1968. T. 267. P. A698–A700. [*168*]

[1094] Kwapień S. *Linear functionals invariant under measure preserving transformations*. Math. Nachr. 1984. B. 119. S. 175–179. [*335*]

[1095] Kwapień S., Tarieladze V. *Mackey continuity of characteristic functionals*. Georg. Math. J. 2002. V. 9, №1. P. 83–112. [*123, 449, 451*]

[1096] Kwapień S., Woyczyński W.A. Random series and stochastic integrals: single and multiple. Birkhäuser, Boston, 1992; xvi+360 pp. [*448*]

[1097]* Laamri I.H. Mesures, intégration, convolution et transformée de Fourier des fonctions. Dunod, Paris, 2001; xii+338 pp.

[1098] Lacey H.E. The isometric theory of classical Banach spaces. Springer-Verlag, Berlin – New York, 1974; x+270 pp. [**I**: *421;* **II**: *326*]

[1099]* Lacey M.T. *Carleson's theorem: proof, complements, variations*. Publ. Mat. 2004. V. 48, №2. P. 251–307.

[1100]* Lahiri B.K., Roy K.C. Real analysis. World Press, Calcutta, 1988; xvi+399 pp.

[1101] Lamb C.W. *A comparison of methods for constructing probability measures on infinite product spaces*. Canad. Math. Bull. 1987. V. 30, №3. P. 282–285. [*445*]

[1102]* Lamperti J.W. Probability. A survey of the mathematical theory. 2nd ed. John Wiley & Sons, New York, 1996; xii+189 pp.

[1103] Landers D., Rogge L. *Cauchy convergent sequences of regular measures with values in a topological group*. Z. Wahrscheinlichkeitstheorie verw. Geb. 1972. B. 21. S. 188–196. [*244*]

[1104] Lanford O.E., Ruelle D. *Observables at infinity and states with short range correlations in statistical mechanics*. Comm. Math. Phys. 1969. V. 13. P. 194–215. [*464*]

[1105]* Lang S. Real analysis. 2nd ed. Addison-Wesley, Reading, Massachusetts, 1983; xv+533 pp.

[1106]* Lang S. Real and functional analysis. 3d ed. Springer, New York, 1993; xiv+580 pp.

[1107] Lange K. *Borel sets of probability measures*. Pacific J. Math. 1973. V. 48. P. 141–161. [*256*]

[1108]* Larman D.G. *On the exponent of convergence of a packing of spheres*. Mathematika. 1966. V. 13. P. 57–59.

[1109]* Larman D.G. *On packings of unequal spheres in* R^n. Canad. J. Math. 1968. V. 20. P. 967–969.

[1110]* 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.

[1111]* 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).

[1112]* la Vallée Poussin Ch.J. de. *Sur l'intégrale de Lebesgue*. Trans. Amer. Math. Soc. 1915. V. 16. P. 435–501.

[1113]* 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.).

[1114]* la Vallée Poussin Ch.J. de. Œuvres scientifiques. V. I–III. Académie Royale de Belgique, Brussels; Circolo Matematico di Palermo, Palermo, 2000–2004.

[1115]* Lax P. Theory of functions of a real variable. Courant Inst. Math. Sci., New York Univ., New York, 1959; xi+154 pp.

[1116]* Leader S. The Kurzweil–Henstock integral and its differentials. Marcel Dekker, New York – Basel, 2001; 355 pp.

[1117] Lebedev V.A. Martingales, convergence of probability measures, and stochastic equations. Izdat. MAI, Moscow, 1996; 348 pp. (in Russian). [*53, 249, 454*]

[1118]* Lebesgue H. *Sur une généralisation de l'intégrale définie*. C. R. Acad. Sci. Paris. 1901. T. 132. P. 1025–1028.

[1119]* Lebesgue H. *Intégrale, longueur, aire*. Ann. Mat. Pura Appl. (3). 1902. T. 7. P. 231–359.

[1120]* Lebesgue H. *Sur une propriété des fonctions*. C. R. Acad. Sci. 1903. T. 137, №26. P. 1228–1230.

[1121]* Lebesgue H. *Sur l'existence des dérivées*. C. R. Acad. Sci. Paris. 1903. T. 136. P. 659–661.

[1122]* Lebesgue H. Leçons sur l'intégration et la recherche des fonctions primitives. Paris, Gauthier-Villars, 1904; 2e éd., Paris, 1928; xii+340 pp.

[1123] Lebesgue H. *Sur les fonctions représentables analytiquement*. J. Math. Pures Appl. (6). 1905. V. 1. P. 139–216. [**I**: *420;* **II**: *439*]

[1124]* Lebesgue H. Leçons sur les séries trigonométriques. Paris, 1906; 138 pp.

[1125]* 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.

[1126]* 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.

[1127]* 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.

[1128]* 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.

[1129]* Lebesgue H. *Sur les intégrales singulières*. Ann. Fac. Sci. Univ. Toulouse (3). 1909. V. 1. P. 25–117.

[1130]* Lebesgue H. *Sur les suites de fonctions mesurables*. C. R. Acad. Sci. Paris. 1909. T. 149. P. 102–103.

[1131]* Lebesgue H. *Sur l'intégration des fonctions discontinues*. Ann. Sci. Ecole Norm. Sup. 1910. V. 27. P. 361–450.

[1132]* Lebesgue H. *Sur l'intégrale de Stieltjes et sur les opérations fonctionnelles linéaires*. C. R. Acad. Sci. Paris. 1910. T. 150. P. 86–88.

[1133] Lebesgue H. *Remarques sur les théories de la mesure et de l'intégration*. Ann. Sci. Ecole Norm. Sup. 1918. V. 35. P. 191–250. [**I**: *416, 422;* **II**: *446*]

[1134]* Lebesgue H. Œuvres scientifiques. V. 1–5. L'Enseignement Mathématique, Université de Genève, Genève, 1972.

[1135]* 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.

[1136]* Lebesgue H. Les lendemains de l'intégrale. Lettres à Émile Borel. Éd. par B. Bru B. et P. Dugac. Vuibert, Paris, 2004; 240 pp.

[1137] Le Cam L. *Un instrument d'étude des fonctions aléatoires: la fonctionnelle caractéristique*. C. R. Acad. Sci. Paris. 1947. T. 224. P. 710. [*449*]

[1138] Le Cam L. *Convergence in distribution of stochastic processes*. Univ. Calif. Publ. Statist. 2 (1957), 207–236. [*197, 452, 454*]
[1139]* Ledoux M. The concentration of measure phenomenon. Amer. Math. Soc., Providence, Rhode Island, 2001; x+181 pp.
[1140] Ledoux M., Talagrand M. Probability in Banach spaces. Isoperimetry and processes. Springer-Verlag, Berlin – New York, 1991; xii+480 pp. [*447, 448*]
[1141]* Lee J.R. Advanced calculus with linear analysis. Academic Press, New York – London, 1972; xiv+218 pp.
[1142]* Lee P.Y., Výborný R. Integral: an easy approach after Kurzweil and Henstock. Cambridge University Press, Cambridge, 2000; xii+311 pp.
[1143] Leese S.J. *Measurable selections and the uniformization of Souslin sets*. Amer. J. Math. 1978. V. 100, №1. P. 19–41. [*39*]
[1144] Léger C., Soury P. *Le convexe topologique des probabilités sur un espace topologique*. J. Math. Pures Appl. 1971. V. 50. P. 363–425. [*456*]
[1145]* Lehmann E.L. Testing statistical hypotheses. John Wiley & Sons, New York, Chapman & Hall, London, 1959; xiii+369 pp.
[1146] Lehn J., Mägerl G. *On the uniqueness of pre-image measures*. Z. Wahrsch. verw. Geb. 1977. B. 38. S. 333–357. [*311*]
[1147]* Leichtweiss K. Konvexe Mengen. Springer-Verlag, Berlin – New York, 1980; 330 S.
[1148]* Leinert M. Integration und Maß. Friedr. Vieweg & Sohn, Braunschweig, 1995; xii+208 S.
[1149] Lembcke J. *Konservative Abbildungen und Fortsetzung regulärer Masse*. Z. Wahrscheinlichkeitstheorie verw. Geb. 1970. B. 15. S. 57–96. [*458*]
[1150] Lembcke J. *Reguläre Masse mit einer gegebenen Familie von Bildmassen*. Sitz. Bayer. Akad. Wiss., Math.-naturwiss. Kl. 1976. S. 61–115. München, 1977. [*458*]
[1151]* Lembcke J. *On a measure extension theorem of Bierlein*. Lecture Notes in Math. 1980. V. 794. P. 45–48.
[1152] Lembcke J. *On simultaneous preimage measures on Hausdorff spaces*. Lecture Notes in Math. 1982. V. 945. P. 110–115. [*458*]
[1153] Léonard Ch. *Characterization of the optimal plans for the Monge–Kantorovich problem*. http://arxiv.org: math.OC/0607604. P. 1–39. [*461*]
[1154]* Leont'eva T.A., Panferov V.S., Serov V.S. Problems in the theory of functions of a real variable. Moskov. Gos. Univ., Moscow, 1997; 208 pp. (in Russian).
[1155]* Letac G. Exercises and solutions manual for “Integration and probability” by Paul Malliavin. Springer-Verlag, New York, 1995; viii+142 pp.
[1156]* Letta G. Teoria elementare dell'integrazione. Editrice Tecnico Scientifica, Pisa, 1973; vii+204 pp. (rev. ed.: Boringhieri, Torino, 1976; 289 pp.)
[1157] Letta G. Martingales et intégration stochastique. Scuola Normale Superiore, Pisa, 1984; 192 pp. [*461*]
[1158] Letta G. *Sur les théorèmes de Hewitt-Savage et de de Finetti*. Lecture Notes in Math. 1989. V. 1372. P. 531–535. [*464*]
[1159] Letta G. *Convergence stable et applications*. Atti Sem. Mat. Fis. Modena (suppl.). 1998. V. 48. P. 191–211. [*249*]
[1160] Letta G., Pratelli L. *Le théorème de Skorohod pour des lois de Radon sur un espace métrisable*. Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. (5). 1997. T. 21. P. 157–162. [*454*]
[1161]* Levi B. *Sopra l'integrazione delle serie*. Rend. Ist. Lombardo (2). 1906. T. 39. P. 775–780.
[1162]* 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.
[1163]* 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.

[1164] Levin V.L. Convex analysis in spaces of measurable functions and its applications in mathematics and economics. Nauka, Moscow, 1985; 352 pp. (in Russian). [*431, 441, 463*]

[1165] Levin V.L. *Measurable selections of multivalued mappings with a bi-analytic graph and σ-compact values*. Trudy Moskov. Mat. Obshch. 1992. V. 54. P. 3–28 (in Russian); English transl.: Trans. Moscow Math. Soc. 1993. P. 1–22. [*37*]

[1166] Lévy P. Problèmes concrets d'analyse fonctionnelle. Paris, 1922 (2me éd.: Gauthier-Villars, Paris, 1951; xiv+484 pp.). [*452*]

[1167] Lévy P. Théorie de l'addition des variables aléatoires. Gautier-Villars, Paris, 1937 (2e éd., 1954; 385 pp.). [**I**: *419;* **II**: *452, 461*]

[1168]* Lichtenstein L. *Ueber die Integration eines bestimmten Integrals in Bezug auf einen Parameter*. Nachr. Ges. Wiss. Göttingen, Math.-phys. Kl. 1910. S. 468–475.

[1169]* Lieb E.H., Loss M. Analysis. 2nd ed. Amer. Math. Soc., Providence, Rhode Island, 2001; 346 pp.

[1170]* Liese F., Vajda I. Convex statistical distances. BSB B.G. Teubner Verlag, Leipzig, 1987; 224 pp.

[1171] Lifshits M.A. Gaussian random functions. Translated from the Russian. Kluwer Academic Publ., Dordrecht, 1995; xii+333 pp. (Russian ed.: Kiev, 1995). [*451*]

[1172] Linde W. Probability in Banach spaces – stable and infinitely divisible distributions. Wiley, New York, 1986; 195 pp. [*451*]

[1173]* Lindenstrauss J., Tzafriri L. Classical Banach spaces. V. I,II. Springer, Berlin – New York, 1977, 1979; xiii+190 pp., x+243 pp.

[1174] Lipchius A.A. *A property of independent mappings*. Theory Probab. Appl. 2004. V. 49, №2. P. 362–365. [*434*]

[1175] Lipchius A.A. *A note on the equality in the problems of Monge and Kantorovich*. Theory Probab. Appl. 2005. V. 50, №4. P. 779–782. [*236*]

[1176]* Lipecki Z. *On strongly additive set functions*. Colloq. Math. 1971. V. 22. P. 255–256.

[1177] Lipecki Z. *Extensions of tight set functions with values in a topological group*. Bull. Acad. Polon. Sci. 1974. V. 22. P. 105–113. [*443*]

[1178]* Lipecki Z. *Unique extensions of positive set functions*. Arch. Math. 1983. B. 41. S. 71–79.

[1179] 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.) [**I**: *vi, 412;* **II**: *410*]

[1180]* Łojasiewicz S. An introduction to the theory of real functions. Wiley, New York, 1988; ix+230 pp.

[1181]* 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.

[1182] Loomis L.H. *On the representation of σ-complete Boolean algebras*. Bull. Amer. Math. Soc. 1947. V. 53. P. 757–760. [*326*]

[1183] Lorch E.R. *Compactification, Baire functions, and Daniell integration*. Acta Sci. Math. (Szeged). 1963. V. 24. P. 204–218. [*447*]

[1184]* Lorentz G.G. *Who discovered analytic sets?* Math. Intelligencer. 2001. V. 23, №4. P. 28–32.

[1185]* Łoś J., Marczewski E. *Extensions of measure*. Fund. Math. 1949. V. 36. P. 267–276.

[1186]* 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.

[1187] Losert V. *Uniformly distributed sequences on compact, separable, non-metrizable groups*. Acta Sci. Math. (Szeged). 1978. V. 40, №1-2. P. 107–110. [*241*]

[1188] Losert V. *On the existence of uniformly distributed sequences in compact topological spaces. I.* Trans. Amer. Math. Soc. 1978. V. 246. P. 463–471. [*241, 257*]

[1189] Losert V. *On the existence of uniformly distributed sequences in compact topological spaces. II.* Monatsh. Math. 1979. B. 87, №3. S. 247–260. [*241*]

[1190] Losert V. *A measure space without the strong lifting property.* Math. Ann. 1979. B. 239, №2. S. 119–128. [*406*]

[1191] Losert V. *A counterexample on measurable selections and strong lifting.* Lecture Notes in Math. 1980. V. 794. P. 153–159. [*463*]

[1192] Losert V. *Strong liftings for certain classes of compact spaces.* Lecture Notes in Math. 1982. V. 945. P. 170–179. [*463*]

[1193] Lotz S. *A survey on hyperdiffuse measures. IV.* In: Proceedings of the Conference Topology and Measure III. Part 1,2. J. Flachsmeyer, Z. Frolík, Yu. M. Smirnov, F. Topsøe, and F. Terpe eds., pp. 127–163. Vitte, 1980; Hiddensee, 1980, Univ. Greifswald, 1982. [*451*]

[1194]* 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.

[1195] Lozanovskiĭ G.Ya. *Normal measures in a product of bicompacta, and a certain problem of V.I. Ponomarev.* Izv. Vysš. Učebn. Zaved. Mat. 1975. №7. P. 114–116 (in Russian); English transl.: Soviet Math. (Iz. VUZ). 1976. V. 19, №7. P. 91–92. [*166*]

[1196]* Lozinskiĭ S.M. *On the indicatrix of Banach.* Vestnik Leningrad. Univ. 1958. V. 13, №7. P. 70–87 (in Russian).

[1197]* Lubotzky A. Discrete groups, expanding graphs and invariant measures. Birkhäuser, Basel – Boston, 1994; 195 pp.

[1198]* 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.

[1199]* Lucia P. de, Pap E. *Convergence theorems for set functions.* P. 125–178 in [**1423**].

[1200]* Lukacs E. Characteristic functions. 2nd ed. Hafner Publ., New York, 1970; x+350 pp.

[1201]* Lukacs E. Developments in characteristic function theory. Griffin, London, 1983; viii+182 pp.

[1202]* Lukes J., Malý J. Measure and integral. Matfyzpress, Prague, 1995; 178 pp.

[1203]* Lusin N. *Sur les propriétés des fonctions mesurables.* C. R. Acad. Sci. Paris. 1912. T. 154. P. 1688–1690.

[1204]* Lusin (Luzin) N.N. *On the fundamental theorem of the integral calculus.* Mat. Sbornik. 1912. V. 28. P. 266–294 (in Russian).

[1205] 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). [**I**: *332, 388, 400, 402, 410, 437;* **II**: *459*]

[1206]* Lusin N. *Sur la classification de M. Baire.* C. R. Acad. Sci. Paris. 1917. T. 164. P. 91–94.

[1207]* Lusin N. *Sur la notion de l'intégrale.* Ann. Mat. Pura Appl. (3). 1917. V. 26. P. 77–129.

[1208] Lusin N. *Sur le problème de M. Jacques Hadamard d'uniformisation des ensembles.* Mathematica (Cluj). 1930. V. 4. P. 54–66. [*441*]

[1209] Lusin N.N. Leçons sur les ensembles analytiques. Gauthier-Villars, Paris, 1930 (2nd. corr. ed.: Chelsea, New York, 1972; xi+326 pp.). [*38, 50, 60, 439, 441*]

[1210]* Lusin N.N. The theory of functions of a real variable. 2nd ed. Učpedgiz, Moscow, 1948; 319 pp. (in Russian).

[1211] Lusin N.N. Collected works. V. 1–3. Fizmatgiz, Moscow, 1953, 1958, 1959 (in Russian). [**I**: *409;* **II**: *439*]

[1212]* Lusin N., Sierpiński W. *Sur quelques propriétés des ensembles (A).* Bull. Acad. Sci. Cracovie, Sér. A. 1918. №4–5a. P. 35–48.

[1213] Luther N.Y. *Unique extension and product measure.* Canad. J. Math. 1967. V. 19. P. 757–763. [**I**: *99, 236;* **II**: *452*]

[1214]* Luther N.Y. *A decomposition of measures.* Canad. J. Math. 1968. V. 20. P. 953–959.

[1215]* Luukkainen J., Saksman E. *Every complete doubling metric space carries a doubling measure*. Proc. Amer. Math. Soc. 1998. V. 126, №2. P. 531–534.
[1216] Lyapunov A.A. (Liapounoff A.) *On completely additive vector functions. I.* Izv. Russian Acad. Sci. 1940. V. 4. P. 465–478 (in Russian; French summary). [*326, 328*]
[1217] Lyapunov A.A. Problems of set theory and the theory of functions. Nauka, Moscow, 1979; 264 pp. (in Russian). [*326, 439*]
[1218] Ma Z. *Some results on regular conditional probabilities*. Acta Math. Sinica, New Ser. 1985. V. 1, №4. P. 302–307. [*462*]
[1219] Ma Z.M., Röckner M. An introduction to the theory of (non-symmetric) Dirichlet forms. Springer, Berlin, 1992; 209 pp. [*441*]
[1220] Macheras N.D., Strauss W. *On the permanence of almost strong liftings*. J. Math. Anal. Appl. 1993. V. 174, №2. P. 566–572. [*463*]
[1221] Macheras N.D., Strauss W. *On strong liftings for projective limits*. Fund. Math. 1994. V. 144, №3. P. 209–229. [*463*]
[1222] Macheras N.D., Strauss W. *The product lifting for arbitrary products of complete probability spaces*. Atti Sem. Mat. Fis. Univ. Modena. 1996. V. 44. P. 485–496. [*463*]
[1223] Mackey G.W. *Borel structures in groups and their duals*. Trans. Amer. Math. Soc. 1957. V. 85. P. 134–165. [*444, 458*]
[1224]* MacNeille H.M. *A unified theory of integration*. Proc. Nat. Acad. Sci. U.S.A. 1941. V. 27. P. 71–76.
[1225] Maeda M. *Some examples of measurable norms*. J. Math. Anal. Appl. 1984. V. 98, №1. P. 158–165. [*449*]
[1226] Maeda M., Harai K., Hagihara R. *Some examples and the connection between cylindrical measures and measurable norms*. J. Math. Anal. Appl. 2003. V. 288, №2. P. 556–564. [*449*]
[1227]* Magyar Z. The Lebesgue integral. Akadémiai Kiadó, Budapest, 1997; x+106 pp.
[1228] Maharam D. *On homogeneous measure algebras*. Proc. Nat. Acad. Sci. U.S.A. 1942. V. 28. P. 108–111. [*280, 459*]
[1229] Maharam D. *On measure in abstract sets*. Trans. Amer. Math. Soc. 1942. V. 51. P. 413–433. [*459*]
[1230] Maharam D. *Decompositions of measure algebras and spaces*. Trans. Amer. Math. Soc. 1950. V. 69. P. 142–160. [*459*]
[1231] Maharam D. *On a theorem of von Neumann*. Proc. Amer. Math. Soc. 1958. V. 9. P. 987–994. [*462*]
[1232] Maharam D. *Automorphisms of products of measure spaces*. Proc. Amer. Math. Soc. 1958. V. 9. P. 702–707. [*320*]
[1233] Maharam D. *On smoothing compact measure spaces by multiplication*. Trans. Amer. Math. Soc. 1975. V. 204. P. 1–39. [*320*]
[1234] Maharam D. *Orthogonal measures: an example*. Ann. Probab. 1982. V. 10, №3. P. 879–880. [*463*]
[1235] Maitra A. *Coanalytic sets that are not Blackwell spaces*. Fund. Math. 1970. V. 67, №2. P. 251–254. [*62*]
[1236] Maitra A. *A note on bimeasurable functions*. Bull. Acad. Polon. Sci. 1975. V. 23, №2. P. 131–134. [*60*]
[1237] Maitra A., Ramakrishnan S. *Factorization of measures and normal conditional distributions*. Proc. Amer. Math. Soc. 1988. V. 103, №4. P. 1259–1267. [*462*]
[1238] Maitra A., Rao B.V., Bhaskara Rao K.P.S. *A problem in the extension of measures*. Illinois J. Math. 1979. V. 23, №2. P. 211–216. [*459*]
[1239] Maitra A., Ryll-Nardzewski C. *On the existence of two analytic non-Borel sets which are not isomorphic*. Bull. Acad. Polon. Sci. 1970. V. 18. P. 177–178. [*440*]
[1240]* 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).

[1241]* Malik S.C., Arora S. Mathematical analysis. 2nd ed. Wiley, New York, 1992; xii+903 pp.
[1242]* Malliavin P., Airault H., Kay L., Letac G. Integration and probability. Springer-Verlag, New York – Berlin, 1995; 322 pp.
[1243] Malliavin P. Stochastic analysis. Springer-Verlag, New York – Berlin, 1997; 343 pp. [*305*]
[1244] Mallory D. *Topological properties of limits of inverse systems of measures.* Canad. J. Math. 1974. V. 26. P. 1455–1465. [*460*]
[1245] Mallory D. *Extension of set functions to measures and applications to inverse limit measures.* Canad. Math. Bull. 1975. V. 18, №4. P. 547–553. [**I**: *52;* **II**: *443, 460*]
[1246] Mallory D., Sion M. *Limits of inverse systems of measures.* Ann. Inst. Fourier (Grenoble). 1971. V. 21. P. 25–57. [*460*]
[1247] Mansfield R. *The solution of one of Ulam's problems concerning analytic sets, II.* Proc. Amer. Math. Soc. 1970. V. 26. P. 539–540. [*440*]
[1248]* Marcinkiewicz J., Salem R. *Sur les sommes riemanniennes.* Compositio Math. 1940. T. 7. P. 376–389.
[1249]* Marcinkiewicz J., Zygmund A. *Mean values of trigonometrical polynomials.* Fund. Math. 1937. V. 28. P. 131–166.
[1250] Marczewski E. *Ensembles indépendants et leurs applications à la théorie de la mesure.* Fund. Math. 1948. V. 35. P. 13–28. [*400, 464*]
[1251] Marczewski E. *Indépendance d'ensembles et prolongement de mesures.* Colloq. Math. 1948. V. 1. P. 122–132. [*400, 464*]
[1252] Marczewski E. *Généralisations du théorème de Steinhaus sur l'ensemble des distances.* Colloq. Math. 1948. V. 1. P. 248–249. [*335*]
[1253] Marczewski E. *Measures in almost independent fields.* Fund. Math. 1951. V. 38. P. 217–229. [*464*]
[1254] Marczewski E. *On compact measures.* Fund. Math. 1953. V. 40. P. 113–124. [**I**: *419;* **II**: *335, 442, 443, 445*]
[1255]* Marczewski E. *Remarks on the convergence of measurable sets and measurable functions.* Colloq. Math. 1955. V. 3. P. 118–124.
[1256] Marczewski E. Collected mathematical papers. PAN, Warszawa, 1996. [**I**: *409;* **II**: *440, 443*]
[1257] Marczewski E., Ryll-Nardzewski C. *Sur la mesurabilité des fonctions de plusieurs variables.* Ann. Soc. Polon. Math. 1952. V. 25. P. 145–154. [*161*]
[1258] Marczewski E., Ryll-Nardzewski C. *Projections in abstract sets.* Fund. Math. 1953. V. 40. P. 160–164. [*441*]
[1259] Marczewski E., Ryll-Nardzewski C. *Remarks on the compactness and non direct products of measures.* Fund. Math. 1953. V. 40. P. 165–170. [*161*]
[1260] Marczewski E., Sikorski R. *Measures in non-separable metric spaces.* Colloq. Math. 1948. V. 1. P. 133–139. [*450*]
[1261] Marczewski E., Sikorski R. *Remarks on measure and category.* Colloq. Math. 1949. V. 2. P. 13–19. [*451*]
[1262]* Marczewski E., Steinhaus H. *On a certain distance of sets and the corresponding distance of functions.* Colloq. Math. 1958. V. 6. P. 319–327.
[1263]* Margulis G.A. *Some remarks on invariant measures.* Monatsh. Math. 1980. B. 90, №3. S. 233–235.
[1264]* Margulis G.A. *Finitely additive invariant measures on Euclidean spaces.* Ergodic Theory Dynam. Syst. 1982. V. 2. P. 383–396.
[1265] Maria J.L. de, Rodriguez-Salinas B. *The space* $(l_\infty/c_0,\ \text{weak})$ *is not a Radon space.* Proc. Amer. Math. Soc. 1991. V. 112, №4. P. 1095–1100. [*451*]
[1266] Maria J.L. de, Rodriguez-Salinas B. *Banach spaces which are Radon spaces with the weak topology.* Bull. London Math. Soc. 1993. V. 25, №6. P. 577–581. [*452*]
[1267] Mařík J. *The Baire and Borel measure.* Czech. Math. J. 1957. V. 7(82), №2. P. 248–253. [*130*]

[1268] Markov A.A. *On mean values and exterior densities.* Mat. Sbornik. 1938. V. 4 (46). P. 165–191. [*446*]
[1269]* Marle C.-M. Mesures et probabilités. Hermann, Paris, 1974; 474 pp.
[1270]* Martin D.A. *A simple proof that determinancy implies Lebesgue measurability.* Rend. Sem. Mat. Univ. Politec. Torino. 2003. V. 61, №4. P. 393–397.
[1271] Matsak I.K., Plichko A.N. *On the support of a measure in a Banach space and finite representability.* Teor. Veroyatn. i Primen. 1991. V. 36, №2. P. 363–367 (in Russian); English transl.: Theory Probab. Appl. 1992. V. 36, №2. P. 381–385. [*120*]
[1272]* Mattila P. Geometry of sets and measures in Euclidean spaces. Cambridge University Press, Cambridge, 1995; 343 pp.
[1273] Mattila P., Mauldin R.D. *Measure and dimension functions: measurability and densities.* Math. Proc. Cambridge Philos. Soc. 1997. V. 121, №1. P. 81–100. [*450*]
[1274] Mauldin R.D. *Baire functions, Borel sets, and ordinary function systems.* Adv. Math. 1974. V. 12, №4. P. 418–450. [*46*]
[1275]* Mauldin R.D. *Countably generated families.* Proc. Amer. Math. Soc. 1976. V. 54. P. 291–297.
[1276] Mauldin R.D. *On non isomorphic analytic sets.* Proc. Amer. Math. Soc. 1976. V. 58. P. 241–244. [*440*]
[1277] Mauldin R.D. *Borel parametrizations.* Trans. Amer. Math. Soc. 1979. V. 250. P. 223–234. [*46, 441*]
[1278] Mauldin R.D., Preiss D., Weizsäcker H. von. *Orthogonal transition kernels.* Ann. Probab. 1983. V. 11, №4. P. 970–988. [*463*]
[1279]* Maurin K. Analysis. Part II. Integration, distributions, holomorphic functions, tensor and harmonic analysis. Reidel, Dordrecht, 1980; xvii+829 pp.
[1280]* Mawhin J. Analyse. Fondements, techniques, évolution. De Boeck Université, Brussels, 1992; iv+808 pp.
[1281]* Mayrhofer K. Inhalt und Maß. Wien, 1952; vii+265 S.
[1282]* Maz'ja V.G. Sobolev spaces. Translated from the Russian. Springer-Verlag, Berlin, 1985; xix+486 pp. (Russian ed.: Leningrad, 1985).
[1283] Mazurkiewicz S. *Sur une propriété des ensembles $C(A)$.* Fund. Math. 1927. V. 10. P. 172–174. [*61*]
[1284]* Mazurkiewicz S. *Sur les fonctions qui satisfont à la condition (N).* Fund. Math. 1930. V. 16. P. 348–352.
[1285] McCann R.J. *Existence and uniqueness of monotone measure-preserving maps.* Duke Math. J. 1995. V. 80. P. 309–323. [**I**: *382;* **II**: *236*]
[1286]* 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.
[1287]* McLeod R.M. The generalized Riemann integral. Mathematical Association of America, Washington, D.C., 1980; xiii+275 pp.
[1288]* McShane E.J. Integration. Princeton, 1944; viii+394 pp.
[1289]* McShane E.J. Unified integration. Academic Press, New York – London, 1983; xiii+607 pp.
[1290]* McShane E.J., Botts T.A. Real analysis. Van Nostrand, Princeton, 1959; ix+272 pp.
[1291]* Medeiros L.A., Mello E.A. de. A integral de Lebesgue. Inst. matem., Univ. Rio de Janeiro, Rio de Janeiro, 1985; 157 pp.
[1292]* Medvedev F.A. Development of set theory in the nineteenth century. Nauka, Moscow, 1965; 232 pp. (in Russian).
[1293]* Medvedev F.A. Development of the concept of the integral. Nauka, Moscow, 1974; 423 pp. (in Russian).
[1294]* Medvedev F.A. Scenes from the history of real functions. Translated from the Russian. Birkhäuser Verlag, Basel, 1991; 265 pp. (Russian ed.: Moscow, 1975).
[1295]* 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).

[1296]* Medvedev F.A. *Letters of D.Th. Egoroff to N.N. Lusin.* Istor.–Matem. Issled. 1980. V. 25. P. 335–361 (in Russian).
[1297]* Medvedev F.A. The early history of the axiom of choice. Nauka, Moscow, 1982; 304 pp. (in Russian).
[1298] Mejlbro L., Preiss D., Tišer J. *Positivity principles in geometrical measure theory.* In: Measure theory, Oberwolfach, 1990, B. Pettineo and P. Vetro eds. Rend. Circ. Mat. Palermo (2) Suppl. No. 28 (1992), pp. 163–167. Circolo Matematico di Palermo, Palermo, 1992. [*451*]
[1299]* Mejlbro L., Topsøe F. *Vitali systems in* $\mathbf{R}^n$ *with irregular sets.* Math. Scand. 1996. V. 79, №1. P. 86–100.
[1300]* 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).
[1301]* Menchoff D. *Sur la représentation conforme des domaines plans.* Math. Ann. 1926. B. 95, №5. S. 641–670.
[1302]* 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).
[1303] Mercourakis S. *Some remarks on countably determined measures and uniform distribution of sequences.* Monatsh. Math. 1996. B. 121, №1-2. S. 79–111. [*241*]
[1304]* 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.
[1305]* 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.
[1306] Métivier M. *Sur la désintégration des mesures.* C. R. Acad. Sci. Paris. 1963. T. 256. P. 1062–1065. [*462*]
[1307] Métivier M. *Limites projectives de mesures. Martingales. Applications.* Ann. Mat. Pura Appl. (4). 1963. V. 63. P. 225–352. [*460, 462*]
[1308] Métivier M. *Génération de mesures dans un* σ*-anneau pseudotopologique.* Rend. Sem. Mat. Univ. Padova. 1964. T. 34. P. 242–304. [*451*]
[1309]* Métivier M. Notions fondamentales de la théorie des probabilités: maîtrises de mathématiques. 2e éd., Dunod, Paris, 1972; xi+334 pp.
[1310]* 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.
[1311] Meyer P.-A. Probability and potentials. Blaisdell, Waltham – Toronto – London, 1966; xiii+266 pp. [**I**: *415;* **II**: *50, 142, 146*]
[1312]* Miamee A.G. *The inclusion* $L^p(\mu) \subseteq L^q(\nu)$. Amer. Math. Monthly. 1991. V. 98, №4. P. 342–345.
[1313] Michael E. *A linear mapping between function spaces.* Proc. Amer. Math. Soc., 1964. V. 15. P. 407–409. [*229*]
[1314] Michael E. *A selection theorem.* Proc. Amer. Math. Soc. 1966. V. 17. P. 1404–1406. [*228*]
[1315]* Michel A. Constitution de la théorie moderne de l'intégration. Mathesis. Librairie Philosophique J. Vrin, Paris, 1992; 338 pp.
[1316]* Michel H. Maß- und Integrationstheorie. I. VEB Deutscher Verlag der Wissenschaften, Berlin, 1978; 204 S.
[1317]* Mikusiński J. *Sur une définition de l'intégrale de Lebesgue.* Bull. Acad. Polon. Sci. 1964. V. 12. P. 203–204.
[1318]* Mikusiński J. The Bochner integral. Birkhäuser Verlag, Basel – Stuttgart, 1978; xii+233 pp.
[1319]* Miller H.I. *A theorem connected with results of Steinhaus and Smítal.* J. Math. Anal. Appl. 1987. V. 124. P. 27–32.

[1320] Minlos R.A. *Generalized random processes and their extension to a measure.* Trudy Moskov. Mat. Obsc. 1959. V. 8. P. 497–518 (in Russian); English transl.: Math. Stat. Probab. 1959. V. 3. P. 291–314. [*124*]

[1321]* Misiewicz J.K., Scheffer C.L. *Pseudo isotropic measures.* Nieuw Arch. Wisk. (4). 1990. V. 8, №2. P. 111–152.

[1322] Mitoma I. *Tightness of probabilities on* $C([0,1];\mathcal{S}')$ *and* $D([0,1];\mathcal{S}')$. Ann. Probab. 1983. V. 11, №4. P. 989–999. [*53*]

[1323]* Mitrinović D.S., Pečarić J.E., Fink A.M. Classical and new inequalities in analysis. Kluwer Academic Publ., Dordrecht, 1993; xviii+740 pp.

[1324] Mohapl J. *On weakly convergent nets in spaces of nonnegative measures.* Czech. Math. J. 1990. V. 40(115), №3. P. 408–421. [*456*]

[1325] Mohapl J. *The Radon measures as functionals on Lipschitz functions.* Czech. Math. J. 1991. V. 41, №3. P. 446–453. [*455*]

[1326]* Monfort A. Cours de probabilités. Economica, Paris, 1980; 210 pp.

[1327]* Moore E.H., Smith H.L. *A general theory of limits.* Amer. J. Math. 1922. V. 44. P. 102–121.

[1328] Moran W. *The additivity of measures on completely regular spaces.* J. London Math. Soc. 1968. V. 43. P. 633–639. [*131, 134*]

[1329] Moran W. *Separate continuity and supports of measures.* J. London Math. Soc. 1969. V. 44. P. 320–324. [*129*]

[1330] Moran W. *Measures and mappings on topological spaces.* Proc. London Math. Soc. 1969. V. 19. P. 493–508. [*449*]

[1331] Moran W. *Measures on metacompact spaces.* Proc. London Math. Soc. 1970. V. 20. P. 507–524. [*226*]

[1332]* Morgan F. Geometric measure theory. A beginner's guide. 3d ed. Academic Press, San Diego, California, 2000; 226 pp.

[1333]* Morse A.P. *Convergence in variation and related topics.* Trans. Amer. Math. Soc. 1937. V. 41. P. 48–83.

[1334]* Morse A.P. *Perfect blankets.* Trans. Amer. Math. Soc. 1947. V. 61. P. 418–442.

[1335]* Moser J. *On the volume elements on a manifold.* Trans. Amer. Math. Soc. 1965. V. 120. P. 286–294.

[1336] Mosiman S.E., Wheeler R.F. *The strict topology in a completely regular setting: relations to topological measure theory.* Canad. J. Math. 1972. V. 24, №5. P. 873–890. [*447, 455, 456*]

[1337] Mouchtari D. *La topologie du type de Sazonov pour les Banach et les supports hilbertiens.* Ann. Univ. Clermont. 1976. V. 61. P. 77–87. [*120, 449*]

[1338] Mourier E. *Eléments aléatoires dans un espace de Banach.* Ann. Inst. H. Poincaré. 1953. T. 19. P. 161–244. [*447*]

[1339] Moy S.C. *Characterizations of conditional expectations as a transformation on function spaces.* Pacif. J. Math. 1954. V. 4. P. 47–64. [*427*]

[1340]* Mozzochi C.J. *On a Riemann sum construction of Rudin.* Proc. Amer. Math. Soc. 1969. V. 22. P. 718.

[1341]* 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.

[1342]* Mukherjea A., Pothoven K. Real and functional analysis. Plenum Press, New York, 1978; x+529 pp.

[1343]* Muldowney P. A general theory of integration in function spaces, including Wiener and Feynman integration. Longman Scientific, Wiley, New York, 1987; viii+115 pp.

[1344]* Munroe M.E. Measure and integration. 2nd ed. Addison–Wesley, Reading, Massachusetts – London – Don Mills, 1971; xii+290 pp.

[1345] Musiał K. *Existence of proper conditional probabilities.* Z. Wahrscheinlichkeitstheorie verw. Geb. 1972. B. 22. S. 8–12. [*444, 462*]

[1346] Musiał K. *Inheritness of compactness and perfectness of measures by thick subsets.* Lecture Notes in Math. 1976. V. 541. P. 31–42. [*89, 444*]

[1347] Musiał K. *Projective limits of perfect measure spaces.* Fund. Math. 1980. V. 110, №3. P. 163–189. [*444*]

[1348] Mushtari D.Kh. Probabilities and topologies on linear spaces. Kazan Mathematics Foundation, Kazan', 1996; xiv+233 pp. [*125, 449*]

[1349] Mushtari D.Kh., Chuprunov A.N. *Sufficient topologies and norms.* Teor. Verojatn. i Primen. 1983. V. 28, №4. P. 700–714 (in Russian); English transl.: Theory Probab. Appl. 1984. V. 28, №4. P. 736–751. [*449*]

[1350] Mycielski J. *Remarks on invariant measures in metric spaces.* Colloq. Math. 1974. V. 32. P. 105–112. [*460*]

[1351] Mycielski J. *A conjecture of Ulam on the invariance of measure in Hilbert's cube.* Studia Math. 1977. V. 60, №1. P. 1–10. [*460*]

[1352] Nachbin L. The Haar integral. Van Nostrand, Princeton, 1965; xii+156 pp. [*460*]

[1353] Naĭmark M.A. Normed algebras. Translated from the 2nd Russian ed. 3d ed. Wolters-Noordhoff, Groningen, 1972; xvi+598 pp. [*460*]

[1354] Nakanishi S. *Weak convergence of measures on the union of metric spaces. I.* Math. Jap. 1986. V. 31, №3. P. 429–447. [*456*]

[1355]* Natanson I.P. Foundations of the theory of functions of a real variable. Leningrad. Univ., Leningrad, 1941; 296 pp. (in Russian).

[1356] 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. [**I**: *vi, 62, 149, 400, 406, 412, 437;* **II**: *460*]

[1357]* Natterer F. The mathematics of computerized tomography. B.G. Teubner, Stuttgart; John Wiley & Sons, New York, 1986; x+222 pp.

[1358]* Nekrasov V.L. Structure and measure of linear point domains. Tomsk Technol. Inst., Tomsk, 1907; viii+254 pp. (in Russian).

[1359] Nelson E. *Regular probability measures on function space.* Ann. Math. (2). 1959. V. 69, №3. P. 630–643. [*448*]

[1360] Neumann J. von. *Algebraische Repräsentaten der Funktionen "bis auf eine Menge vom Masse Null".* J. Reine Ang. Math. 1931. B. 165. S. 109–115. [*462*]

[1361] Neumann J. von. *Einige Sätze über messbare Abbildungen.* Ann. Math. 1932. V. 33. P. 574–586. [*284, 458*]

[1362] Neumann J. von. *Zur Operatorenmethode in der klassischen Mechanik.* Ann. Math. 1932. V. 33. P. 587–642. [*458*]

[1363] Neumann J. von. *On rings of operators. Reduction theory.* Ann. Math. (2). 1949. V. 50. P. 401–485. [*441*]

[1364]* Neumann J. von. Functional Operators. I. Measures and Integrals. Annals of Mathematics Studies, no. 21. Princeton University Press, Princeton, 1950; vii+261 pp.

[1365]* Neumann J. von. Collected works. V. 1–4. Pergamon Press, New York, 1961.

[1366]* Neumann J. von. Invariant measures. Amer. Math. Soc., Providence, Rhode Island, 1999; xvi+134 pp.

[1367] Neumann J. von, Stone M.H. *The determination of representative elements in the residual classes of a Boolean algebra.* Fund. Math. 1935. V. 25. P. 353–378. [*376*]

[1368] 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. [**I**: *vi, 414;* **II**: *99, 463*]

[1369] Neveu J. Discrete-parameter martingales. Translated from the French. Revised ed. North-Holland, Amsterdam – Oxford, American Elsevier, New York, 1975; viii+236 pp. [*432, 461*]

[1370] Niederreiter H. *On the existence of uniformly distributed sequences in compact spaces.* Compositio Math. 1972. V. 25. P. 93–99. [*238*]

[1371] Nielsen O.A. An introduction to integration and measure theory. John Wiley & Sons, New York, 1997; xvi+473 pp. [**I**: *320, 414;* **II**: *446*]

[1372]* Nikodym O.[2] *Sur la mesure des ensembles plans dont tous les points sont rectilinéairement accessibles.* Fund. Math. 1927. V. 10. P. 116–168.
[1373]* Nikodym O. *Sur une propriété de la mesure généralisée des ensembles.* Prace Matem.–Fiz. 1928–1929. T. 36. P. 65–71.
[1374]* Nikodym O. *Sur les fonctions d'ensemble.* Comptes Rendus du I Congrès des Math. des Pays Slaves, Varsovie, 1929 (1930), pp. 304–313.
[1375]* Nikodym O. *Sur une généralisation des intégrales de M. J. Radon.* Fund. Math. 1930. V. 15. P. 131–179.
[1376]* 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.
[1377]* Nikodym O. *Sur les suites de fonctions parfaitement additives d'ensembles abstraits.* C. R. Acad. Sci. Paris. 1931. T. 192. P. 727.
[1378]* Nikodym O. *Sur les familles bornées de fonctions parfaitement additives d'ensemble abstrait.* Monatsh. Math. Phys. 1933. B. 40. S. 418–426.
[1379]* Nikodym O. *Sur les suites convergentes de fonctions parfaitement additives d'ensemble abstrait.* Monatsh. Math. Phys. 1933. B. 40. S. 427–432.
[1380]* 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.
[1381]* 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.
[1382] Novikoff A., Barone J. *The Borel law of normal numbers, the Borel zero-one law, and the work of Van Vleck.* Historia Math. 1977. V. 4. P. 43–65. [*464*]
[1383] Novikoff P. *Sur les fonctions implicites mesurables B.* Fund. Math. 1931. V. 17. P. 8–25. [*33, 38, 441*]
[1384] Novikov (Novikoff) P.S. *On the consistency of some propositions of the descriptive theory of sets.* Trudy Steklov Matem. Inst. AN SSSR. 1951. V. 38. P. 279–316 (in Russian); English transl.: Amer. Math. Soc. Transl. (2). 1963. V. 29. P. 51–89. [*331, 444*]
[1385] Novikov P.S. Selected works. The theory of sets and functions. Mathematical logic and algebra. Nauka, Moscow, 1979 (in Russian). [*439, 441*]
[1386] Novoa J.F. *Coverability of Radon measures.* Atti Sem. Mat. Fis. Univ. Modena. 2000. V. 48, №1. P. 151–162. [*452*]
[1387] Nussbaum A.E. *On a theorem by I. Glicksberg.* Proc. Amer. Math. Soc. 1962. V. 13. P. 645–646. [*163*]
[1388] O'Brien G.L., Watson S. *Relative compactness for capacities, measures, upper semicontinuous functions and closed sets.* J. Theor. Probab. 1998. V. 11, №3. P. 577–588. [*455*]
[1389] Ochakovskaya O.A. *On some classes of mappings preserving of measure.* In: Functional Analysis and its Applications, V. Kadets and W. Zelazko eds. (Proceedings of the international conference on functional analysis and its applications dedicated to the 110th anniversary of Stefan Banach, May 28–31, 2002, Lviv), pp. 205–208. Elsevier, Amsterdam, 2004. [*338*]
[1390]* Ochan Yu.S. Collection of problems and theorems in the theory of functions of a real variable. Prosveščenie, Moscow, 1965; 231 pp. (in Russian).
[1391]* Ochan Yu.S. The integral. Moskov. Pedagog. Inst., Moscow, 1973; 313 pp. (in Russian).
[1392]* Ochan Yu.S. Collection of problems in mathematical analysis. Prosveščenie, Moscow, 1981; 271 pp. (in Russian).
[1393]* Oden J.T., Demkowicz L.F. Applied functional analyis. CRC Press, Boca Raton, 1996; xiv+653 pp.

[2] Another (later) spelling: Nikodým O.M.

[1394] Ohta H., Tamano K. *Topological spaces whose Baire measure admits a regular Borel extension*. Trans. Amer. Math. Soc. 1990. V. 317, №1. P. 393–415. [*131, 156*]

[1395] Okada S. *Supports of Borel measures*. J. Austral. Math. Soc. 1979. V. 27. P. 221–231. [*450*]

[1396] Okada S., Okazaki Y. *On measure-compactness and Borel measure-compactness*. Osaka Math. J. 1978. V. 15. P. 183–191. [*449*]

[1397] Okazaki Y. *Stochastic basis in Fréchet space*. Math. Ann. 1986. B. 274. S. 379–383. [*120*]

[1398]* Okikiolu G.O. Aspects of the theory of bounded integral operators in L^p-spaces. Academic Press, London – New York, 1971; ix+522 pp.

[1399]* Olevskiĭ A.M. Fourier series with respect to general orthogonal systems. Springer-Verlag, New York – Heidelberg, 1975; viii+136 pp.

[1400] Olson M.P. *A characterization of conditional probability*. Pacif. J. Math. 1965. V. 15, №3. P. 971–983. [*461*]

[1401] Oppel U. *Zur charakterisierung Suslinscher und Lusinscher Räume*. Z. Wahrscheinlichkeitstheorie verw. Geb. 1976. B. 34, №3. S. 183–192. [*455*]

[1402] Oppel U. *Zur schwachen Topologie auf dem Vektorraum der Borel-Masse polnischer und Lusinscher Räume*. Math. Z. 1976. B. 147, №1. S. 97–99. [*455*]

[1403] Orkin M. *A Blackwell space which is not analytic*. Bull. Acad. Polon. Sci. 1972. V. 20. P. 437–438. [*50*]

[1404]* Os C.H. van. Moderne integraalrekening. Noordhoff, Groningen, 1925; 204 pp.

[1405]* Osserman R. *The isoperimetric inequality*. Bull. Amer. Math. Soc. 1978. V. 84, №6. P. 1182–1238.

[1406] Ostrovskiĭ E.I. *On the supports of probability measures in separable Banach spaces*. Dokl. Akad. Nauk SSSR. 1980. V. 225, №6. P. 1319–1320 (in Russian); English transl.: Soviet Math. Dokl. 1980. V. 22, №3. P. 836–838. [*170, 448*]

[1407] Ottaviani G. *Sulla indipendenza delle funzioni misurabili*. Atti Accad. Lincei. Rend. Cl. sci. fis., mat. e natur. Roma. (8). 1947. V. 2. P. 393–398. [*434*]

[1408] Oxtoby J.C. *Homeomorphic measures in metric spaces*. Proc. Amer. Math. Soc. 1970. V. 24, №3. P. 419–423. [*286, 330*]

[1409] 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. [**I**: *93, 414;* **II**: *451*]

[1410] Oxtoby J., Ulam S. *On the equivalence of any set of first category to a set of measure zero*. Fund. Math. 1938. V. 31. P. 201–206. [*336*]

[1411] Oxtoby J.C., Ulam S.M. *On the existence of a measure invariant under a transformation*. Ann. Math. 1939. V. 40. P. 560–566. [*336, 433, 443, 458*]

[1412] Oxtoby J.C., Ulam S. *Measure preserving homeomorphisms and metrical transitivity*. Ann. Math. (2). 1941. V. 42. P. 874–920. [*442, 458*]

[1413] Pachl J.K. *Every weakly compact probability is compact*. Bull. Acad. Polon. Sci. 1975. V. 23, №4. P. 401–405. [*173*]

[1414] Pachl J.K. *Disintegration and compact measures*. Math. Scand. 1978. V. 43, №1. P. 157–168. [*404, 405, 462*]

[1415] Pachl J.K. *Two classes of measures*. Colloq. Math. 1979. V. 42. P. 331–340. [*160, 462*]

[1416] Pachl J.K. *Measures as functionals on uniformly continuous functions*. Pacif. J. Math. 1979. V. 82, №2. P. 515–521. [*219, 256*]

[1417] Padmanabhan A.R. *Convergence in probability and allied results*. Math. Japon. 1970. V. 15. P. 111–117. [*266*]

[1418]* Pallu de la Barrière R. Intégration. Ellipses, Paris, 1997; 208 pp.

[1419]* 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.

[1420]* 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.

[1421] Panzone R., Segovia C. *Measurable transformations on compact spaces and o. n. systems on compact groups*. Rev. Un. Mat. Urgentina. 1964. V. 22. P. 83–102. [*320, 451*]

[1422]* Pap E. Null-additive set functions. Kluwer Academic Publ., Dordrecht, 1995; xii+315 pp.

[1423]* Pap E. (ed.) Handbook of measure theory. V. 1,2. North-Holland, Amsterdam, 2002.

[1424] Parthasarathy K.R. Probability measures on metric spaces. Academic Press, New York, 1967; xi+276 pp. (Repr. ed.: AMS Chelsea, Providence, Rhode Island, 2005). [*443*]

[1425]* 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.

[1426] Paterson A.L.T. Amenability. American Math. Soc., Providence, Rhode Island, 1988; xx+452 pp. [*460*]

[1427]* 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.

[1428]* Peano G. Applicazioni geometriche del calcolo infinitesimale. Torino, 1887; xii+334 pp.

[1429]* Pedersen G.K. Analysis now. Springer-Verlag, New York, 1989; xiv+277 pp.

[1430] Pełczyński A. Linear extensions, linear averagings, and their applications to linear topological classification of spaces of continuous functions. Dissert. Math. (Rozprawy Mat.) 1968. V. 58. 92 pp. [*201*]

[1431] Pellaumail J. *Application de l'existence d'un relèvement à un théorème sur la désintégration de mesures*. Ann. Inst. H. Poincare, Sect. B (N.S.). 1972. V. 8, №3. P. 211–215. [*462*]

[1432] Perlman M.D. *Characterizing measurability, distribution and weak convergence of random variables in a Banach space by total subsets of linear functionals*. J. Multivar. Anal. 1972. V. 2, №3. P. 174–188. [*440*]

[1433]* 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).

[1434]* Pesin Y.B. Dimension theory in dynamical systems. Contemporary views and applications. University of Chicago Press, Chicago, 1997; xii+304 pp.

[1435] Péter E. *An extension of the subsequence principle*. Studia Sci. Math. Hungar. 2000. V. 36. P. 335–345. [*464*]

[1436] Peters G. *Anticipating flows on the Wiener space generated by vector fields of low regularity*. J. Funct. Anal. 1996. V. 142. P. 129–192. [*460*]

[1437] Petersen K. Ergodic theory. Cambridge University Press, 1983; xi+329 pp. [*391*]

[1438] Peterson H.L. *Regular and irregular measures on groups and dyadic spaces*. Pacific J. Math. 1969. V. 28. P. 173–182. [*451*]

[1439] Petrov V.V. Sums of independent random variables. Springer, New York, 1975; x+346 pp. Translated from the Russian (Russian ed.: Moscow, 1972). [*410*]

[1440] Petrov V.V. Limit theorems of probability theory. Sequences of independent random variables. Clarendon Press, Oxford University Press, New York, 1995; xii+292 pp. (Russian ed.: Moscow, 1987). [*410*]

[1441]* Pettis J. *On the extension of measures*. Ann. Math. 1951. V. 54. P. 186–197.

[1442] Pfanzagl J. *Convergent sequences of regular measures*. Manuscr. Math. 1971. V. 4, №1. P. 91–98. [*241, 259*]

[1443] Pfanzagl J. *Conditional distributions as derivatives*. Ann. Probab. 1979. V. 7, №6. P. 1046–1050. [*462*]

[1444]* Pfanzagl J., Pierlo W. Compact systems of sets. Lecture Notes in Math. V. 16. Springer-Verlag, Berlin – New York, 1966; ii+48 pp.

[1445] Pfeffer W.F. Integrals and measures. Marcel Dekker, New York, 1977; ix+259 pp. [**I**: *414;* **II**: *446*]

[1446]* Pfeffer W.F. The Riemann approach to integration. Local geometric theory. Cambridge University Press, Cambridge, 1993; xvi+302 pp.

[1447]* Pfeffer W.F. Derivation and integration. Cambridge University Press, Cambridge, 2001; xvi+266 pp.

[1448] Phelps R.R. Lectures on Choquet's theorem. Van Nostrand, Princeton, 1966; v+130 pp.; 2nd ed.: Lecture Notes in Math. V. 1757. Springer-Verlag, Berlin – New York, 2001. [*146*]

[1449]* Phillips E.R. *Nicolai Nicolaevich Luzin and the Moscow school of the theory of functions.* Hist. Math. 1978. V. 5. P. 275–305.

[1450]* Phillips E.R. An introduction to analysis and integration theory. Corr. repr. of the 1971 ed. Dover Publications, New York, 1984; xxviii+452 pp.

[1451]* Phillips R.S. *On linear transformations.* Trans. Amer. Math. Soc. 1940. V. 48. P. 516–541.

[1452] Phillips R.S. *On weakly compact subsets of a Banach space.* Amer. J. Math. 1943. V. 65. P. 108–136. [*452*]

[1453]* Picone M., Viola T. Lezione sulla teoria moderna dell'integrazione. Einaudi, Torino, 1952; 404 pp.

[1454] Pier J.-P. Amenable locally compact groups. Wiley, New York, 1984; 418 pp. [*451*]

[1455]* 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.

[1456]* Pier J.-P. Histoire de l'intégration. Vingt-cinq siècles de mathématiques. Culture Scientifique. Masson, Paris, 1996; xii+307 pp.

[1457]* 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.

[1458]* 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).

[1459] Pintacuda N. *Sul lemma di misurabilità di Doob.* Boll. Un. Mat. Ital. A (7). 1989. V. 3, №2. P. 237–241. [*51*]

[1460]* Pisier G. The volume of convex bodies and Banach space geometry. Cambridge University Press, Cambridge, 1989; xvi+250 pp.

[1461]* Pitt H.R. Integration, measure and probability. Oliver & Boyd, Edinburgh – London, 1963; viii+110 pp.

[1462]* Pitt H.R. Measure and integration for use. Oxford University Press, London – New York, 1985; xii+143 pp.

[1463]* 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.

[1464]* 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.

[1465] Plebanek G. *Solution to a problem of Thiemann concerning existence of measures.* Bull. Polish Acad. Sci. Math. 1986. V. 34, №7-8. P. 401–403. [*160*]

[1466] Plebanek G. *Measures on two-dimensional products.* Mathematika. 1989. V. 36, №2. P. 253–258. [*335, 444*]

[1467] Plebanek G. *On Radon measures on first-countable spaces.* Fund. Math. 1995. V. 148. P. 159–164. [*449*]

[1468] Plebanek G. *A note on strictly positive Radon measures.* Colloq. Math. 1995. V. 69, №2. P. 187–192. [*450*]

[1469] Plebanek G. *A τ-additive measure without thick Lindelöf subset.* Questions Answers Gen. Topology. 1997. V. 15, №2. P. 241–243. [*166, 450*]

[1470] Plebanek G. *Nonseparable Radon measures and small compact spaces*. Fund. Math. 1997. V. 153, №1. P. 25–40. [*452*]

[1471] Plebanek G. *Approximating Radon measures on first-countable compact spaces*. Colloq. Math. 2000. V. 86, №1. P. 15–23. [*241, 452*]

[1472] Poincaré H. *Sur les équations de la dynamique et le problème des trois corps*. Acta Math. 1890. V. 13. P. 1–270 (Œuvres, V. 7. Paris, Gauthier-Villars, 1953). [*463*]

[1473] Pol R. *Note on the spaces $P(S)$ of regular probability measures whose topology is determined by countable subsets*. Pacif. J. Math. 1982. V. 100, №1. P. 185–201. [*230*]

[1474]* Polischuk E.M. Émile Borel (1871–1956). Nauka, Leningrad, 1980; 168 pp. (in Russian).

[1475] Pollard D. *Induced weak convergence and random measures*. Z. Wahrscheinlichkeitstheorie verw. Geb. 1977. B. 37, №4. S. 321–328. [*456*]

[1476] Pollard D. *Compact sets of tight measures*. Studia Math. 1976. V. 56, №1. P. 63–67. [*456*]

[1477] Pollard D. *Weak convergence on nonseparable metric spaces*. J. Austral. Math. Soc. Ser. A. 1979. V. 28, №2. P. 197–204. [*456*]

[1478] Pollard D. Convergence of stochastic processes. Springer, Berlin – New York, 1984; xiv+215 pp. [*453*]

[1479]* Pollard D. A user's guide to measure theoretic probability. Cambridge University Press, Cambridge, 2002; xiv+351 p.

[1480] Pollard D., Topsøe F. *A unified approach to Riesz type representation theorems*. Studia Math. 1975. V. 54, №2. P. 173–190. [*447*]

[1481]* 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.

[1482]* Poroshkin A.G. The theory of measure and integral. Syktyvkar. Gos. Univ., Syktyvkar, 1996; 172 pp. (in Russian).

[1483]* Possel R. de. *Sur la dérivation abstraite des fonctions d'ensembles*. J. Math. Pures Appl. 1936. V. 15. P. 391–409.

[1484] Pratelli L. *Sur le lemme de mesurabilité de Doob*. Lecture Notes in Math. 1990. V. 1426. P. 46–51. [*51*]

[1485]* Pratt J.W. *On interchanging limits and integrals*. Ann. Math. Statist. 1960. V. 31. P. 74–77.

[1486] Preiss D. *Metric spaces in which Prohorov's theorem is not valid*. Z. Wahrscheinlichkeitstheorie verw. Geb. 1973. B. 27. S. 109–116. [*224, 225*]

[1487] Preiss D. *Differentiation of measures in infinitely-dimensional spaces*. In: Proceedings of the Conference Topology and Measure III. Part 1,2. J. Flachsmeyer, Z. Frolík, Ju.M. Smirnov, F. Topsøe, and F. Terpe eds., pp. 201–207. Ernst-Moritz-Arndt-Universitat, Greifswald, 1982. [*451*]

[1488]* Preiss D. *Geometry of measures in R^n: distribution, rectifiability, and densities*. Ann. Math. (2). 1987. V. 125, №3. P. 537–643.

[1489] Preiss D., Rataj J. *Maximal sets of orthogonal measures are not analytic*. Proc. Amer. Math. Soc. 1985. V. 93, №3. P. 471–476. [*463*]

[1490] Preiss D., Tišer J. *Differentiation of measures on Hilbert spaces*. Lecture Notes in Math. 1982. V. 945. P. 194–207. [*451*]

[1491] Preiss D., Tišer J. *Measures in Banach spaces are determined by their values on balls*. Mathematika. 1991. V. 38. P. 391–397. [*451*]

[1492] Preston C.J. Random fields. Lecture Notes in Math. V. 534. Springer-Verlag, Berlin – New York, 1976; 200 pp. [*464*]

[1493]* Priestley H.A. Introduction to integration. Clarendon Press, Oxford University Press, New York, 1997; xii+306 pp.

[1494] Prigarin S.M. *Weak convergence of probability measures in spaces of continuously differentiable functions*. Sibirsk. Mat. Zh. 1993. V. 34, №1. P. 140–144 (in Russian); English transl.: Sib. Math. J. 1993. V. 34, №1. P. 123–127. [*456*]

[1495] Prinz P. *Regularity of Riesz measures.* Proc. Amer. Math. Soc. 1986. V. 96, №2. P. 330–334. [*452*]

[1496] Prohorov Yu.V. *Probability distributions in functional spaces.* Uspehi Mat. Nauk. 1953. V. 8, №3. P. 165–167 (in Russian). [*442*]

[1497] Prohorov Yu.V. *Convergence of random processes and limit theorems in probability theory.* Teor. Verojatn. i Primen. 1956. V. 1, №2. P. 177–238 (in Russian); English transl.: Theory Probab. Appl. 1956. V. 1. P. 157–214. [*188, 202, 309, 442, 447, 449, 452, 453, 454, 455*]

[1498] Prohorov Yu.V. *The method of characteristic functionals.* Proc. 4th Berkeley Symp. on Math. Statistics and Probability. V. 2. P. 403–419. University of California Press, Berkeley, 1960. [*443, 449*]

[1499] Prohorov Yu.V., Sazonov V.V. *Some results associated with Bochner's theorem.* Teor. Verojatn. i Primen. 1961. V. 6, №1. P. 87–93 (in Russian); English transl.: Theory Probab. Appl. 1961. V. 6. P. 82–87. [*449*]

[1500] Prum B., Fort J.-C. Stochastic processses on a lattice and Gibbs measures. Kluwer Academic Publ., Dordrecht, 1991; ix+220 p. [*464*]

[1501]* Pták P., Tkadlec J. *A note on determinacy of measures.* Časopis Pěst. Mat. 1988. V. 113, №4. P. 435–436.

[1502]* 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.

[1503]* 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).

[1504]* Pugh C.C. Real mathematical analysis. Springer-Verlag, New York, 2002; xii+437 pp.

[1505] Purves R. *Bimeasurable functions.* Fund. Math. 1966. V. 58. P. 149–157. [*60*]

[1506] Rachev S.T. *The Monge-Kantorovich problem on mass transfer and its applications in stochastics.* Teor. Verojatn. i Primen. 1984. V. 29, №4. P. 625–653 (in Russian); English transl.: Theory Probab. Appl. 1984. V. 29, №4. P. 647–676. [*454*]

[1507] Rachev S.T. Probability metrics and the stability of stochastic models. Wiley, Chichester, 1991; xiv+494 pp. [*454*]

[1508] Rachev S.T., Rüschendorf L. Mass transportation problems. V. 1. Springer, New York, 1998; 508 pp. [*236, 456*]

[1509] Rademacher H. *Eineindeutige Abbildungen und Meßbarkeit.* Monatsh. Math. Phys. 1916. B. 27. S. 183–290. [*459*]

[1510]* Rademacher H. *Über eine Eigenschaft von meßbaren Mengen positiven Maßes.* Jahr. Deutsch. Math. Verein. 1921. B. 30. S. 130–132.

[1511]* Radó T. Length and area. Amer. Math. Soc., New York, 1948; v+572 pp.

[1512]* Radó T., Reichelderfer P. *A theory of absolutely continuous transformations in the plane.* Trans. Amer. Math. Soc. 1941. V. 49. P. 258–307.

[1513] Radó T., Reichelderfer P.V. Continuous transformations in analysis. Springer–Verlag, Berlin, 1955; vi+441 pp. [*460*]

[1514] Radon J. *Theorie und Anwendungen der absolut additiven Mengenfunktionen.* Sitz. Akad. Wiss. Wien, Math.-naturwiss. Kl. IIa. 1913. B. 122. S. 1295–1438. [**I**: *418, 425, 429, 431, 434, 437;* **II**: *442, 446, 452, 457*]

[1515]* 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.

[1516] Radon J. *Über lineare Funktionaltransformationen und Funktionalgleichungen.* Sitz. Akad. Wiss. Wien, Math.-naturwiss. Kl. IIa. 1919. B. 128, H. 7. S. 1–39. [*452*]

[1517]* Radon J. Gesammelte Abhandlungen, V. 1,2. Birkhäuser, Basel – Boston, 1987.

[1518]* Ramachandran B. Advanced theory of characteristic functions. Statist. Publ. Soc., Calcutta, 1967; vii+208 pp.

[1519] Ramachandran D. *On the two definitions of independence*. Colloq. Math. 1975. V. 32, №2. P. 227–231. [*399*]

[1520] Ramachandran D. *Existence of independent complements in regular conditional probability spaces*. Ann. Probab. 1979. V. 7, №3. P. 433–443. [*459*]

[1521] Ramachandran D. *Perfect mixtures of perfect measures*. Ann. Probab. 1979. V. 7, №3. P. 444–452. [*444*]

[1522] Ramachandran D. Perfect measures. Parts I,II. ISI Lecture Notes, V. 5,7. Macmillan Co. of India, New Delhi, 1979. [*459, 462, 462*]

[1523] Ramachandran D. *A note on regular conditional probabilities in Doob's sense*. Ann. Probab. 1981. V. 9, №5. P. 907–908. [*462, 462*]

[1524] Ramachandran D., Rüschendorf L. *A general duality theorem for marginal problems*. Probab. Theory Related Fields. 1995. V. 101, №3. P. 311–319. [*325, 461*]

[1525] Ramachandran D., Rüschendorf L. *Duality and perfect probability spaces*. Proc. Amer. Math. Soc. 1996. V. 124, №7. P. 2223–2228. [*325, 461*]

[1526]* Rana I.K. An introduction to measure and integration. 2nd ed. Amer. Math. Soc., Rhode Island, Providence, 2002; 459 pp.

[1527]* Randolph J.F. Basic real and abstract analysis. Academic Press, New York, 1968; 504 pp.

[1528] Rao B.V. *On discrete Borel spaces and projective sets*. Bull. Amer. Math. Soc. 1969. V. 75. P. 614–617. [**I**: *211;* **II**: *440*]

[1529] Rao B.V. *Remarks on analytic sets*. Fund. Math. 1970. V. 66. P. 237–239. [*440*]

[1530] Rao B.V. *Non-existence of certain Borel structures*. Fund. Math. 1970. V. 69. P. 241–242. [*440*]

[1531] Rao B.V. *Remarks on vector sums of Borel sets*. Colloq. Math. 1972. V. 25. P. 103–104. [*60*]

[1532] Rao K.P.S. Bhaskara, Rao B.V. *Borel spaces*. Dissertationes Math. (Rozprawy Mat.). 1981. V. 190; 63 pp. [*50, 58, 440*]

[1533]* Rao K.P.S. Bhaskara, Rao B.V. *On extensions of measures*. Colloq. Math. 1987. V. 53, №1. P. 43–47.

[1534]* 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.

[1535] Rao K.P.S. Bhaskara, Shortt R.M. *Separation of points by families of intervals*. Real Anal. Exchange. 1990–1991. V. 16, №1. P. 177–186. [*61*]

[1536] Rao M. Bhaskara, Rao K.P.S. Bhaskara. *Borel σ-algebra on $[0,\Omega]$*. Manuscr. Math. 1971. V. 5. P. 195–198. [*161*]

[1537] Rao M.M. *Projective limits of probability spaces*. J. Multivariate Anal. 1971. V. 1, №1. P. 28–57. [*173*]

[1538] Rao M.M. *Conditional measures and operators*. J. Multivar. Anal. 1975. V. 5. P. 330–413. [*462*]

[1539] Rao M.M. Stochastic processes theory and integration. Sijthoff & Noordhoff, Aphen aan den Rijn, 1979; 456 pp. [*441, 462*]

[1540] Rao M.M. Probability theory with applications. Academic Press, New York, 1984; 495 pp. [**I**: *414;* **II**: *461*]

[1541] 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.). [**I**: *242, 312, 320, 397, 414, 423;* **II**: *452*]

[1542] Rao M.M. Conditional measures and applications. Marcel Dekker, New York, 1993; xiv+417 pp. (2nd ed.: Chapman and Hall/CRC, Boca Raton, Florida, 2005; xxiv+483 pp.) [*462*]

[1543] Rao M.M., Sazonov V.V. *A theorem on a projective limit of probability spaces and its applications*. Teor. Verojatn. i Primen. 1992. V. 38, №2. P. 345–355 (in Russian); English transl.: Theory Probab. Appl. 1993. V. 38, №2. P. 307–315. [*460*]

[1544] Rao R.R. *Relations between weak and uniform convergences of measures with applications*. Ann. Math. Statist. 1962. V. 33, №2. P. 659–680. [*190*]

[1545]* Ray W.O. Real analysis. Prentice Hall, Englewood Cliffs, New Jersey, 1988; xii+307 pp.

[1546] Raynaud de Fitte P. *Compactness criteria for the stable topology.* Bull. Polish Acad. Sci. Math. 2003. V. 51, №4. P. 343–363. [*248, 249*]

[1547] Reiter H. *On some properties of locally compact groups.* Indag. Math. 1965. V. 27. P. 687–701. [*333*]

[1548] Remy M. *Disintegration and perfectness of measure spaces.* Manuscr. Math. 1988. V. 62, №3. P. 277–296. [*406, 444, 462*]

[1549] Rényi A. *On a new axiomatic theory of probability.* Acta Math. Acad. Sci. Hung. 1955. V. 6. P. 285–335. [*462*]

[1550] Rényi A. *On conditional probability spaces generated by a dimensionally ordered set of measures.* Theory Probab. Appl. 1956. V. 1. P. 61–71. [*462*]

[1551] Rényi A. *On stable sequences of events.* Sankhyā. Ser. A. 1963. V. 25. P. 293–302. [*248*]

[1552] Repovš D., Semenov P.V. Continuous selections of multivalued mappings. Kluwer Academic Publ., Dordrecht, 1998; 356 pp. [*228*]

[1553] Reshetnyak Yu.G. *General theorems on semicontinuity and convergence with a functional.* Sibirsk. Mat. Zh. 1967. V. 8, №5. P. 1051–1069 (in Russian); English transl.: Sib. Math. J. 1967. V. 8. P. 801–816. [*252*]

[1554]* 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.

[1555] Ressel P. *Some continuity and measurability results on spaces of measures.* Math. Scand. 1977. V. 40, №1. P. 69–78. [*127, 156, 245*]

[1556] Ressel P. *A topological version of Slutsky's theorem.* Proc. Amer. Math. Soc. 1982. V. 85, №2. P. 272–274. [*261*]

[1557] Ressel P. *De Finetti-type theorems: an analytical approach.* Ann. Probab. 1985. V. 13, №3. P. 898–922. [*409*]

[1558] Révész P. The laws of large numbers. Academic Press, 1968; Académiai Kiadó, Budapest, 1970; 176 pp. [*410*]

[1559]* Revuz D. Mesure et intégration. Hermann, Paris, 1997.

[1560]* Rey Pastor J. Elementos de la teoria de funciones. 3 ed. Ibero-Americana, Madrid – Buenos Aires, 1953; 560 pp.

[1561]* Richard U. Teoria delle funzioni. Ist. Matematico Del Politecnico di Torino, Torino, 1951; viii+367 pp.

[1562]* Richter H. Wahrscheinlichkeitstheorie. 2e Aufl. Springer, Berlin, 1966; xii+462 S.

[1563]* Ridder J. *Integration in abstrakten Räumen.* Fund. Math. 1935. V. 24. P. 72–117.

[1564]* Riečan B., Neubrunn T. Integral, measure, and ordering. Kluwer Academic Publ., Dordrecht; Ister Science, Bratislava, 1997; xiv+378 pp.

[1565]* Riesz F. *Sur les systèmes orthogonaux de fonctions.* C. R. Acad. Sci. Paris. 1907. T. 144. P. 615–619.

[1566]* Riesz F. *Sur une espèce de géométrie analytique des systèmes de fonctions sommables.* C. R. Acad. Sci. Paris. 1907. T. 144. P. 1409–1411.

[1567]* Riesz F. *Sur les suites de fonctions mesurables.* C. R. Acad. Sci. Paris. 1909. T. 148. P. 1303–1305.

[1568] Riesz F. *Sur les opérations fonctionnelles linéaires.* C. R. Acad. Sci. Paris. 1909. T. 149. P. 974–977. [**I**: *425;* **II**: *446*]

[1569] Riesz F. *Untersuchungen über Systeme integrierbarer Funktionen.* Math. Ann. 1910. B. 69. S. 449–497. [**I**: *424, 431;* **II**: *457*]

[1570]* Riesz F. *Sur certains systèmes singuliers d'équations intégrales.* Ann. École Norm. Sup. (3). 1911. V. 28. P. 33–62.

[1571] Riesz F. *Sur quelques points de la théorie des fonctions sommables.* C. R. Acad. Sci. Paris. 1912. T. 154. P. 641–643. [**I**: *423;* **II**: *445*]

[1572] Riesz F. *Sur intégrale de Lebesgue*. Acta Math. 1920. V. 42. P. 191–205. [**I**: *423;* **II**: *445*]

[1573]* Riesz F. *Sur la convergence en moyenne, I,II.* Acta Sci. Math. 1928–29. T. 4. P. 58–64, 182–185.

[1574]* Riesz F. *Über Sätze von Stone und Bochner*. Acta Univ. Szeged. 1933. V. 6. P. 184–198.

[1575]* 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.

[1576] Riesz F. *Sur la théorie ergodique*. Comment. Math. Helv. 1944–1945. V. 17. P. 221–239. [*463*]

[1577]* Riesz F. Œuvres complètes, V. I,II. Acad. Sci., Budapest, 1960.

[1578] 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. [**I**: *163, 412;* **II**: *446*]

[1579]* Riesz M. *Sur les ensembles compacts de fonctions sommables.* Acta Sci. Math. Szeged. 1933. V. 6. P. 136–142.

[1580] Rinkewitz W. *Compact measures and measurable weak sections.* Math. Scand. 2002. V. 91. P. 150–160. [*311*]

[1581] Rinow W. *Über die Eindeutigkeit der Caratheodoryschen Erweiterung eines Maßes.* Math. Nachr. 1969. B. 42. S. 289–294. [*421*]

[1582] Riss E.A. *On measures that agree on balls.* Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. Problemy Teorii Veroyatnost. Raspred. XI. 1989. V. 177. P. 122–128 (in Russian); English transl.: J. Soviet Math. 1992. V. 61, №1. P. 1907–1911. [*451*]

[1583] Riss E.A. *The positivity principle for equivalent norms.* Algebra i Analiz. 2000. V. 12, №3. P. 146–172 (in Russian); English transl.: St. Petersburg Math. J. 2001. V. 12, №3. P. 451–469. [*451*]

[1584]* Rivière T., Ye D. *Resolutions of the prescribed volume from equation.* Nonlinear Diff. Eq. Appl. 1996. V. 3. P. 323–369.

[1585] Rodriguez-Salinas B. *Strictly localizable measures.* Rend. Circ. Mat. Palermo (2). 1992. V. 41, №2. P. 295–301. [*452*]

[1586]* Rogers C. Packing and covering. Cambridge University Press, New York, 1964; viii+111 pp.

[1587] Rogers C.A. Hausdorff measures. Cambridge University Press, London, 1970; viii+179 pp. [**I**: *430;* **II**: *140*]

[1588] Rogers C.A. *A linear Borel set whose difference set is not a Borel set.* Bull. London Math. Soc. 1970. V. 2, №1. P. 41–42. [*60*]

[1589] Rogers C.A., Jayne J.E. K-analytic sets. In: Analytic sets, pp. 1–181. Academic Press, New York, 1980. [*8, 49, 56, 61, 440*]

[1590]* Rogers C.A., Sion M. *On Hausdorff measures in topological spaces.* Monatsh. Math. 1963. B. 67. S. 269–278.

[1591] Rogge L. *The convergence determining class of regular open sets.* Proc. Amer. Math. Soc. 1973. V. 37, №2. P. 581–585. [*244*]

[1592]* Rogosinski W.W. Volume and integral. Oliver and Boyd, Edinburgh and London; Interscience Publishers, New York, 1952; x+160 pp.

[1593] Rohlin V.A. *On the classification of measurable decompositions.* Dokl. Akad. Nauk SSSR. 1947. V. 58. P. 29–32 (in Russian). [*459*]

[1594] Rohlin (Rokhlin) V.A. *On the problem of the classification of automorphisms of Lebesgue spaces.* Dokl. Akad. Nauk SSSR. 1947. V. 58. P. 189–191 (in Russian). [*459*]

[1595] Rohlin V.A. *On the fundamental ideas of measure theory.* Mat. Sbornik. 1949. V. 25. P. 107–150; English transl.: Amer. Math. Soc. Transl. 1952. №71; reissued in AMS Translations. Ser. 1. 1962. V. 10. P. 1–54. [*284, 442, 459, 462*]

[1596] Rohlin V.A. *Selected topics from the metric theory of dynamical systems.* Uspehi Mat. Nauk. 1949. V. 4, №2. P. 57–128 (in Russian). [*441, 459*]

[1597] Rohlin V.A. *Metric classification of measurable functions.* Uspehi Mat. Nauk. 1957. V. 12, №2. P. 169–174 (in Russian). [*459*]

[1598] Rohlin V.A. *New progress in the theory of transformations with invariant measure.* Uspehi Mat. Nauk. 1961. V. 15, №4. P. 3–26 (in Russian); English transl.: Russian Math. Surveys. 1960. V. 15, №4. P. 1–22. [*459*]

[1599] Rohlin V.A. *Exact endomorphisms of a Lebesgue space.* Izv. Akad. Nauk SSSR Ser. Mat. 1961. V. 25. P. 499–530 (in Russian). [*459*]

[1600] Rokhlin V.A. *Lectures on the entropy theory of measure-preserving transformations.* Uspehi Mat. Nauk. 1967. V. 22, №5. P. 3–56 (in Russian); English transl.: Russian Math. Surveys. 1967. V. 22, №5. P. 1–52. [*459*]

[1601] Rohlin V.A. Selected works. Moscow Centre of Continuous Mathematical Education, Moscow, 1999 (in Russian). [**I**: *409;* **II**: *459*]

[1602]* Romanovski P. *Intégrale de Denjoy dans les espaces abstraits.* Mat. Sbornik USSR. 1941. T. 9 (51), №1. P. 67–120.

[1603] Romanovsky V. *Sur un théorème limite du calcul des probabilités.* Mat. Sbornik (Recuil Math. Soc. Math. Moscou). 1929. V. 36. P. 36–64. [*453*]

[1604]* Romero J.L. *When is $L^p(\mu)$ contained in $L^q(\mu)$?* Amer. Math. Monthly. 1983. V. 90. P. 203–206.

[1605]* Rooij A.C.M. van, Schikhof W.H. A second course on real functions. Cambridge University Press, Cambridge, 1982; xiii+200 pp.

[1606]* Rosenblatt J. *Uniqueness of invariant means for measure-preserving transformations.* Trans. Amer. Math. Soc. 1981. V. 265, №2. P. 623–636.

[1607]* Rosenthal A. *Beiträge zu Carathéodory's Messbarkeitstheorie.* Nachr. Ges. Wiss. Göttingen. 1916. S. 305–321.

[1608]* 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.

[1609]* Rosenthal H.P. *On relatively disjoint families of measures, with some applications to Banach space theory.* Studia Math. 1970. V. 37. P. 13–36.

[1610]* Rosenthal J.S. A first look at rigorous probability theory. World Sci., Singapore, 2000; xiv+177 pp.

[1611] Rosiński J. *On the convolution of cylindrical measures.* Bull. Acad. Polon. Sci. 1982. V. 30, №7–8. P. 379–383. [*147*]

[1612] Ross K., Stromberg K. *Baire sets and Baire measures.* Ark. Mat. 1965. V. 6. P. 151–160. [*44*]

[1613]* Ross K.A., Stromberg K. *Jessen's theorem on Riemann sums for locally compact groups.* Pacific J. Math. 1967. V. 20. P. 135–147.

[1614] Rota G.C. *On the representation of averaging operators.* Rend. Semin. Nat. Padova. 1960. T. 30. P. 52–64. [*427*]

[1615]* Rotar V.I. Probability theory. Translated from the Russian. World Sci., River Edge, 1997; xviii+414 pp. (Russian ed.: Moscow, 1992).

[1616] Roussas G.G. Contiguity of probability measures: some applications in statistics. Cambridge University Press, London – New York, 1972; xiii+248 pp. [*257*]

[1617]* Roussas G.G. An introduction to measure-theoretic probability. Amsterdam, Elsevier, Academic Press, New York, 2004; xviii+443 pp.

[1618] Royden H.L. Real analysis. 3d ed., Prentice Hall, Englewood Cliffs, New Jersey, 1988; 444 pp. (1st ed.: Macmillan, 1963). [**I**: *vi, 414;* **II**: *460*]

[1619]* Rubel L.A. *A pathological Lebesgue-measurable function.* J. London Math. Soc. 1963. V. 38. P. 1–4.

[1620]* Ruch J.-J., Weber M. *Quelques résultats dans la théorie des sommes riemanniennes.* Exposition. Math. 1997. V. 15, №3. P. 279–288.

[1621]* Ruckle W.H. Modern analysis. Measure theory and functional analysis with applications. PWS-Kent Publ., Boston, 1991; xv+265 pp.

[1622]* Rudin W. *An arithmetic property of Riemann sums.* Proc. Amer. Math. Soc. 1964. V. 15. P. 321–324.

[1623]* Rudin W. Principles of mathematical analysis. 2nd ed. McGraw-Hill, New York, 1964; ix+270 pp.

[1624]* Rudin W. Real and complex analysis. 2nd ed., McGraw-Hill, New York, 1974; xii+452 pp.

[1625] Rudin W. *Lebesgue's first theorem.* Mathematical analysis and applications, Part B, pp. 741–747, Adv. in Math. Suppl. Stud., 7b, Academic Press, New York – London, 1981. [*58*]

[1626] Rudolph D. Fundamentals of measurable dynamics. Ergodic theory on Lebesgue spaces. Clarendon Press, Oxford University Press, New York, 1990; x+168 pp. [*459*]

[1627] Rue Th. de La. *Espaces de Lebesgue.* Lecture Notes in Math. 1993. V. 1557. P. 15–21. [*459*]

[1628] Rüschendorf L., Thomsen W. *Closedness of sum spaces and the generalized Schrödinger problem.* Theory Probab. Appl. 1997. V. 42, №3. P. 483–494. [*434*]

[1629]* Ruziewicz S. *Remarque à la note de M. Banach "Sur une classe de fonctions continues".* Fund. Math. 1926. V. 8. P. 173–174.

[1630] Rybakov V.I. *Some properties of measures on a normed space that has the* RN *property.* Matem. Zametki. 1977. V. 21, №1. P. 81–92 (in Russian); English transl.: Math. Notes. 1977. V. 1, №1–2. P. 45–50. [*452*]

[1631] Ryll-Nardzewski C. *On quasi-compact measures.* Fund. Math. 1953. V. 40. P. 125–130. [*335, 444*]

[1632]* 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.

[1633]* Sadovnichiĭ V.A. Theory of operators. Translated from the second Russian edition. Consultants Bureau, New York, 1991; xii+396 pp. (Russian ed.: Moscow, 1986).

[1634]* Sadovnichiĭ V.A., Grigor'yan A.A., Konyagin S.V. Problems from students mathematical competitions. Moskov. Gos. Univ., Moscow, 1987; 312 pp. (in Russian).

[1635] Sadovnichii Yu.V. *On the Kantorovich norm for alternating measures.* Dokl. Ros. Akad. Nauk. 1999. V. 368, №4. P. 459–461 (in Russian); English transl.: Dokl. Math. 1999. V. 60, №2. P. 223–225. [*457*]

[1636] Sainte-Beuve M.F. *On the extension of von Neumann–Aumann's theorem.* J. Funct. Anal. 1974. V. 17. P. 112–129. [*40*]

[1637] Saint-Pierre J. *Désintégration d'une mesure non borné.* Ann. Inst. H. Poincaré, Sect. B. 1975. V. 11, №3. P. 275–286. [*462*]

[1638] Saint-Raymond J. *Caractérisation d'espaces polonais.* Séminaire Choquet (Initiation à l'analyse). 11–12e années. 1971–1973. №5. 10 pp. [*456*]

[1639] Saint-Raymond J. *Riemann-measurable selections.* Set-Valued Anal. 1994. V. 2, №3. P. 481–485. [*441*]

[1640] 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. [**I**: *332, 372, 392, 411, 437;* **II**: *458*]

[1641]* Saks S. *On some functionals.* Trans. Amer. Math. Soc. 1933. V. 35, №2. P. 549–556; Addition: ibid., №4. P. 965–970.

[1642] Saks S. *Integration in abstract metric spaces.* Duke Math. J. 1938. V. 4. P. 408–411. [*446*]

[1643] Saks S., Sierpiński W. *Sur une propriété générale de fonctions.* Fund. Math. 1928. V. 11. P. 105–112. [*160*]

[1644]* Salem R., Zygmund A. *Some properties of trigonometric series whose terms have random signs.* Acta Math. 1954. V. 91. P. 245–301.

[1645] Samorodnitskiĭ A.A. Measure theory. Izdat. Leningrad. Univ., Leningrad, 1990; 268 pp. (in Russian). [*459*]

[1646] Samorodnitskiĭ A.A. Theory of Lebesgue–Rohlin spaces. Syktyvkar. Gos. Univ., Syktyvkar, 1997; 288 pp. (in Russian). [*459*]
[1647]* Samuélidès M., Touzillier L. Analyse fonctionnelle. Cépaduès Éditions, Toulouse, 1989; iv+289 pp.
[1648]* Sansone G., Merli L. Lezioni sulla teoria delle funzioni di variabili reali. Cedam, Padova, 1964; 236 pp.
[1649] Sapounakis A. *The existence of strong liftings for totally ordered measure spaces.* Pacif. J. Math. 1983. V. 106, №1. P. 145–151. [*463*]
[1650]* Sarason D. *Functions of vanishing mean oscillation.* Trans. Amer. Math. Soc. 1975. V. 207. P. 391–405.
[1651] Sato H. *Hilbertian support of a probability measure on a Banach space.* Lecture Notes in Math. 1979. V. 709. P. 195–205. [*120, 450*]
[1652]* Saxe K. Beginning functional analysis. Springer-Verlag, New York, 2002; xii+197 pp.
[1653]* Saxena S.Ch., Shah S.M. Introduction to real variable theory. 2nd ed. Prentice-Hall of India, New Delhi, 1980; xv+334 pp.
[1654] Sazhenkov A.N. *The principle of uniform boundedness for topological measures.* Matem. Zametki. 1982. V. 31, №2. P. 263–267 (in Russian); English transl.: Math. Notes. 1982. V. 31, №1–2. P. 135–137. [*244*]
[1655] Sazonov V.V. *On characteristic functionals.* Teor. Veroyatn. i Primen. 1958. V. 3, №2. P. 201–205 (in Russian); English transl.: Theory Probab. Appl. 1958. V. 3. P. 188–192. [*124*]
[1656] Sazonov V.V. *On perfect measures.* Izv. Akad. Nauk SSSR Ser. Mat. 1962. V. 26, №3. P. 391–414 (in Russian); English transl.: Amer. Math. Soc. Transl. (2). 1965. V. 48. P. 229–254. [*90, 159, 406, 444, 462, 462*]
[1657] Sazonov V.V. *Answer to a question by R.L. Dobrušin.* Teor. Verojatn. i Primen. 1964. V. 9, №1. P. 180–181 (in Russian); English transl.: Theory Probab. Appl. 1964. V. 9. P. 160–161. [*461*]
[1658] Sazonov V.V., Tutubalin V.N. *Probability distributions on topological groups.* Teor. Verojatn. i Primen. 1966. V. 11, №1. P. 3–55 (in Russian); English transl.: Theor. Probab. Appl. 1966. V. 11, №1. P. 1–45. [*451*]
[1659] Schachermayer W. *Mesures cylindriques sur les espaces de Banach, qui ont la propriété de Radon–Nikodym.* C. R. Acad. Sci. Paris, Sér. A. 1976. T. 282, №4. P. 227–229. [*452*]
[1660] Schachermayer W. *Eberlein-compacts et espaces de Radon.* C. R. Acad. Sci. Paris, Sér. A. 1977. T. 284. P. A405–A407. [*135*]
[1661] Schaefer H.H. Topological vector spaces. Springer-Verlag, Berlin – New York, 1971; xi+294 pp. [**I**: *281;* **II**: *119, 123, 208*]
[1662] Schaerf H.M. *On the continuity of measurable functions in neighborhood spaces.* Portugaliae Math. 1947. V. 6. P. 33–44; errata, P. 66. [*450*]
[1663] Schäl M. *On dynamic programming: compactness of the space of policies.* Stoch. Processes Appl. 1975. V. 3. P. 345–364. [*249*]
[1664]* Schauder J.P. *The theory of surface measure.* Fund. Math. 1926. V. 8. P. 1–48.
[1665]* 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.
[1666]* Scheffé H. *A useful convergence theorem for probability distributions.* Ann. Math. Statist. 1947. V. 18. P. 434–438.
[1667] Schief A. *The continuity of subtraction and the Hausdorff property in spaces of Borel measures.* Math. Scand. 1988. V. 63, №2. P. 215–219. [*261*]
[1668] Schief A. *Topological properties of the addition map in spaces of Borel measures.* Math. Ann. 1988. B. 282, №1. S. 23–31. [*261*]
[1669] Schief A. *On continuous image averaging of Borel measures.* Topol. Appl. 1989. V. 31, №3. P. 309–315. [*228*]
[1670] Schief A. *An open mapping theorem for measures.* Monatsh. Math. 1989. B. 108, №1. S. 59–70. [*228, 260*]

[1671] Schief A. *Almost surely convergent random variables with given laws.* Probab. Theory Related Fields. 1989. V. 81. P. 559–567. [*454*]
[1672]* Schilling R. Measures, integrals and martingales. Cambridge University Press, Cambridge, 2005; 380 pp.
[1673]* Schlesinger L., Plessner A. Lebesguesche Integrale und Fouriersche Reihen. De Gruyter, Berlin – Leipzig, 1926; 229 S.
[1674]* Schmetterer L. Einführung in die mathematische Statistik. 2e Aufl. Springer-Verlag, Wien – New York, 1966; ix+597 S.
[1675]* Schmitz N. Einführung in die Maß- und Integrationstheorie. Gesellschaft zur Förderung der Mathematischen Statistik, Münster, 1987; 144 S.
[1676]* Schmitz N., Plachky D. Vorlesungen über Wahrscheinlichkeitstheorie. I. Verlag Anton Hain, Meisenheim am Glan, 1976; vi+216 S.
[1677]* Schneider R. Convex bodies: the Brunn–Minkowski theory. Cambridge University Press, 1993; xiii+490 pp.
[1678]* Schönflies A. Entwicklung der Mengenlehre und ihrer Anwendungen. Leipzig, Berlin, 1913; 388 S.
[1679] Schwartz L. *Sur le théorème du graphe fermé.* C. R. Acad. Sci. Paris. 1966. T. 263. P. 602–605. [*168*]
[1680]* Schwartz L. Analyse mathématique. V. I,II. Hermann, Paris, 1967; xxix+830 pp., xxiv+571 pp.
[1681] Schwartz L. Radon measures on arbitrary topological spaces and cylindrical measures. Oxford University Press, London, 1973; xii+393 pp. [*443, 452, 455*]
[1682] Schwartz L. *Surmartingales régulières à valeurs mesures et désintégrations régulières d'une mesure.* J. Anal. Math. 1973. V. 26. P. 1–168. [*462*]
[1683] Schwartz L. *Certaines propriétés des mesures sur les espaces de Banach.* Seminaire Maurey–Schwartz (1975–1976), Exposé 23, Centre Math., École Polytech., Palaiseau, 1976. [*447*]
[1684] Schwartz L. Lectures on disintegration of measures. Tata Institute of Fundamental Research, Bombay, 1975; 132 pp. [*462*]
[1685] Schwartz L. Geometry and probability in Banach spaces. Lecture Notes in Math. V. 852. Springer-Verlag, Berlin – New York, 1981; 101 pp. [*447*]
[1686] Scorza Dragoni G. *Un teorema sulle funzioni continue rispetto ad una e misurabili rispetto ad un'altra variabile.* Rend. Sem. Mat. Univ. Padova. 1948. V. 17. P. 102–106. [*137*]
[1687]* 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.
[1688]* Segal I.E. *Equivalences of measure spaces.* Amer. J. Math. 1951. V. 73. P. 275–313.
[1689]* Segal I., Kunze R.A. Integrals and operators. Springer, Berlin – New York, 1978; xiv+371 pp.
[1690] Seidel W. *Supports of Borel measures.* Fund. Math. 1989. V. 133, №1. P. 67–80. [*450*]
[1691] Semadeni Z. Banach spaces of continuous functions. V. I. Polish Sci. Publ., Warszawa, 1971; 584 pp. [*452*]
[1692] Sentilles F.D. *Bounded continuous functions on a completely regular space.* Trans. Amer. Math. Soc. 1972. V. 168. P. 311–336. [*455*]
[1693]* Severini C. *Sopra gli sviluppi in serie di funzioni ortogonali.* Atti Acc. Gioenia di Catania (5). 1910. T. 3, Mem. XI. P. 1–7.
[1694]* Severini C. *Sulle successioni di funzioni ortogonali.* Atti Acc. Gioenia di Catania (5). 1910. T. 3, Mem. XIII. P. 1–10.
[1695] Shelah S. *Lifting of the measure algebra.* Israel J. Math. 1983. V. 45. P. 90–96. [*376*]
[1696] Sherman S. *On denumerably independent families of Borel fields.* Amer. J. Math. 1950. V. 72. P. 612–614. [*400*]

[1697]* 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).
[1698]* Shilov G.E. Mathematical analysis. Functions of several real variables. Nauka, Moscow, 1972; 624 pp. (in Russian).
[1699] 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). [**I**: *397, 414, 438;* **II**: *107, 446*]
[1700] Shiryaev A.N. Probability. Translated from the Russian. Springer-Verlag, New York, 1996; xvi+623 pp. (3nd Russian ed.: Moscow, 2004). [**I**: *vi, 414;* **II**: *409, 410, 453, 461*]
[1701] Shneider (Šneĭder) V.E. *The descriptive theory of sets in topological spaces.* Učen. Zap. Moskov. Gos. Univ. 1949. №135. Matematika. T. II. P. 37–85 (in Russian). [*440*]
[1702] Shortt R.M. *The extension of measurable functions.* Proc. Amer. Math. Soc. 1983. V. 87, №3. P. 444–446. [*60*]
[1703] Shortt R.M. *Universally measurable spaces: an invariance theorem and diverse characterizations.* Fund. Math. 1984. V. 121, №2. P. 169–176. [*159*]
[1704] Shortt R.M., Rao K.P.S. Bhaskara. *Generalized Lusin sets with the Blackwell property.* Fund. Math. 1987. V. 127, №1. P. 9–39. [*50*]
[1705] Šidák Z. *On relations between strict sense and wide sense conditional expectations.* Theory Probab. Appl. 1957. V. 2. P. 267–272. [*428*]
[1706]* Sierpiński W. *Sur une propriété générale des ensembles de points.* C. R. Acad. Sci. Paris. 1916. T. 162. P. 716–717.
[1707]* Sierpiński W. *Sur la question de la mesurabilité de la base de M. Hamel.* Fund. Math. 1920. V. 1. P. 105–111.
[1708]* Sierpiński W. *Sur les fonctions convexes mesurables.* Fund. Math. 1920. V. 1. P. 125–129.
[1709]* Sierpiński W. *Sur un problème concernant les ensembles mesurables superficiellement.* Fund. Math. 1920. V. 1. P. 112–115.
[1710]* 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.
[1711]* Sierpiński W. *Sur les fonctions d'ensembles additives et continues.* Fund. Math. 1922. V. 3. P. 240–246.
[1712]* Sierpiński W. *Sur l'ensemble de distances entre les points d'un ensemble.* Fund. Math. 1925. V. 7. P. 144–148.
[1713] Sierpiński W. *Sur une propriété charactéristique des ensembles analytiques.* Fund. Math. 1927. V. 10. P. 169–171. [*60*]
[1714] Sierpiński W. *Sur un problème de M. Hausdorff.* Fund. Math. 1927. V. 10. P. 427–430. [*28*]
[1715] Sierpiński W. *Sur les ensembles boreliens abstraits.* Ann. Soc. Polon. Math. 1927. T. 6. P. 50–53. [**I**: *419;* **II**: *57*]
[1716]* Sierpiński W. *Un théorème général sur les familles des ensembles.* Fund. Math. 1928. V. 12. P. 206–210.
[1717] Sierpiński W. *Sur la continuité des fonctions absolument additives d'ensembles.* Ann. Soc. Polon. Math. 1928. T. 7. P. 75–78. [*160*]
[1718] Sierpiński W. *Sur la dualité entre la première catégorie et la mesure nulle.* Fund. Math. 1934. V. 22. P. 276–280. [*451*]
[1719]* Sierpiński W. *Sur une propriété de fonctions quelconques d'une variable réelle.* Fund. Math. 1935. V. 25. P. 1–4.
[1720] Sierpiński W. *Sur une suite de fonctions continues dont toute fonction d'accumulation est non mesurable.* Publ. Inst. Math. Beograd. 1947. V. 1. P. 5–10. [*60*]
[1721] Sierpiński W. Hypothèse du continu. 2nd ed. Chelsea, New York, 1956; xvii+274 pp. [**I**: *78, 422;* **II**: *440*]

[1722]* Sierpiński W. On the congruence of sets and their equivalence by finite decomposition. New York, 1967; 115 pp.

[1723] Sierpiński W. Œuvres choisies. T. I–III. PWN-éditions Scientifiques de Pologne, Warsaw, 1974, 1975, 1976. [**I**: *409, 419, 422;* **II**: *440*]

[1724] Sikorski R. *On the representation of Boolean algebras as fields of sets*. Fund. Math. 1948. V. 35. P. 247–258. [*326*]

[1725] Sikorski R. Boolean algebras. Springer-Verlag, Berlin – Heidelberg, 1960; ix+176 pp. [**I**: *421;* **II**: *325*]

[1726]* Sikorski R. Advanced calculus: functions of several variables. Polish Sci. Publ., Warszawa, 1969; 460 pp.

[1727]* Simon L. Lectures on geometric measure theory. Centre for Mathematical Analysis, Australian National University, Canberra, 1983; vii+272 pp.

[1728]* Simonnet M. Measures and probabilities. Springer-Verlag, New York, 1996; xiv+510 pp.

[1729] Sinaĭ Ya.G. Theory of phase transitions: rigorous results. Translated from the Russian. Pergamon Press, Oxford-Elmsford, 1982; viii+150 pp. (Russian ed.: Moscow, 1980). [*464*]

[1730] Sinai Ya.G. Topics in ergodic theory. Princeton University Press, Princeton, 1994; viii+218 pp. [*391, 464*]

[1731] Sion M. *On uniformization of sets in topological spaces*. Trans. Amer. Math. Soc. 1960. V. 96. P. 237–245. [*440*]

[1732] Sion M. *Topological and measure-theoretic properties of analytic sets*. Proc. Amer. Math. Soc. 1960. V. 11. P. 769–776. [*440*]

[1733] Sion M. *Approximate continuity and differentiation*. Canad. J. Math. 1962. V. 14. P. 467–475. [*139*]

[1734] Sion M. *On capacitability and measurability*. Ann. Inst. Fourier (Grenoble). 1963. V. 18. P. 88–99. [*142*]

[1735]* Sion M. Introduction to the methods of real analysis. Holt, Rinehart and Winston, New York – Montreal – London, 1968; x+134 pp.

[1736] Sion M. A theory of semigroup valued measures. Lecture Notes in Math. V. 355. Springer-Verlag, Berlin – New York, 1973; iv+140 pp. [**I**: *423;* **II**: *463*]

[1737]* Sion M., Willmott R.C. *Hausdorff measures on abstract spaces*. Trans. Amer. Math. Soc. 1966. V. 123. P. 275–309.

[1738] Skala H.J. *The existence of probability measures with given marginals*. Ann. Probab. 1993. V. 21, №1. P. 136–142. [*324, 461*]

[1739] Skorohod A.V. *Limit theorems for stochastic processes*. Teor. Verojatn. i Primen. 1956. V. 1. P. 261–290 (in Russian); English transl.: Theory Probab. Appl. 1956. V. 1. P. 261–290. [*53, 199, 452*]

[1740] Skorokhod A.V. Studies in the theory of random processes. Translated from the Russian. Addison-Wesley, Reading, Mass., 1965; viii+199 pp. (Russian ed.: Kiev, 1961). [*199, 452*]

[1741] Skorohod A.V. Integration in Hilbert space. Springer-Verlag, Berlin – New York, 1974; xii+177 pp. [*448*]

[1742] Słowikowski W. *Pre-supports of linear probability measures and linear Lusin measurable functionals*. Dissert. Math. 1972. V. 93. P. 1–43. [*448*]

[1743] Slutsky E. *Über stochastische Asymptoten und Grenzwerte*. Metron. 1925. V. 5. P. 3–89. [**I**: *171, 426;* **II**: *261*]

[1744]* 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.

[1745]* 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. (1st Russian ed.: Moscow, 1947; 2nd ed.: Moscow, 1959).

[1746]* Smith H.J.S. *On the integration of discontinuous functions*. Proc. London Math. Soc. 1875. V. 6. P. 140–153.
[1747] Smoleński W. *Pre-supports and kernels of probability measures in Fréchet spaces*. Demonstr. Math. 1977. V. 10, №3–4. P. 751–762. [*451*]
[1748] Smoleński W. *An abstract form of a counterexample of Marek Kanter*. Lecture Notes in Math. 1984. V. 1080. P. 288–291. [*451*]
[1749] Smoleński W. *On the approximation of measurable linear functionals*. Statist. Probab. Lett. 1985. V. 3, №4. P. 205–207. [*451*]
[1750] Smoleński W. *On the kernel of probability measures on a Banach space*. Demonstr. Math. 1988. V. 21. P. 569–572. [*451*]
[1751] Smolyanov O.G. *Measurable polylinear and power functionals in certain linear spaces with measure*. Dokl. Akad. Nauk SSSR. 1966. V. 170. P. 526–529 (in Russian); English transl.: Soviet Math. Dokl. 1966. V. 7. P. 1242–1246. [*451*]
[1752] Smolyanov O.G. *Measurable linear varieties in products of linear spaces with measure*. Matem. Zametki. 1969. V. 5, №5. P. 623–634 (in Russian). English transl.: Math. Notes. 1969. V. 5. P. 374–379. [*410*]
[1753] Smolyanov O.G. Analysis on topological linear spaces and its applications. Moskov. Gos. Univ., Moscow, 1979; 86 pp. (in Russian). [*456*]
[1754] Smolyanov O.G. *The Gross–Sazonov theorem for sign-variable cylindrical measures*. Vestnik Moskovsk. Univ. 1983. №4. P. 4–12 (in Russian); English transl.: Moscow Univ. Math. Bull. 1983. V. 38. P. 1–9. [*449*]
[1755] Smolyanov O.G., Fomin S.V. *Measures on topological linear spaces*. Uspehi Mat. Nauk. 1976. V. 31, №4. P. 3–56 (in Russian); English transl.: Russian Math. Surveys. 1976. V. 31, №4. P. 1–53. [*125, 449, 456*]
[1756] Smolyanov O.G., Shavgulidze E.T. *A simple proof of Tarieladze's theorem on sufficiency of positively sufficient topologies*. Teor. Verojatn. i Primen. 1992. V. 37, №2. P. 421–424 (in Russian); English transl.: Theory Probab. Appl. 1992. V. 37, №2. P. 402–404. [*449*]
[1757]* Sobolev S.L. Introduction to the theory of cubature formulae. Nauka, Moscow, 1974; 808 pp. (in Russian).
[1758]* Sobolev V.I. Lectures on supplementary chapters of mathematical analysis. Nauka, Moscow, 1968; 288 pp. (in Russian).
[1759] Sodnomov B.S. *On arithmetic sums of sets*. Dokl. Akad. Nauk SSSR. 1951. V. 80, №2. P. 173–175 (in Russian). [*60*]
[1760] Sodnomov B.S. *Example of two sets of type G_δ whose arithmetic sum is non-B-measurable*. Dokl. Akad. Nauk SSSR. 1954. V. 99, №6. P. 507–510 (in Russian). [*60*]
[1761]* Sodnomov B. *On a property of sets of positive measure*. Fund. Math. 1967. V. 60. P. 187–190.
[1762]* Sohrab H.H. Basic real analysis. Birkhäuser Boston, Boston, 2003; xiv+559 pp.
[1763] Sokal A.D. *Existence of compatible families of proper regular conditional probabilities*. Z. Wahrscheinlichkeitstheorie verw. Geb. 1981. B. 56, №4. S. 537–548. [*462*]
[1764]* Solovay R. *A model of set theory in which every set of reals is Lebesgue measurable*. Ann. Math. 1970. V. 92. P. 1–56.
[1765]* 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.
[1766] Sondermann D. *Masse auf lokalbeschränkten Räumen*. Ann. Inst. Fourier (Grenoble). 1969. V. 19, №2. P. 33–113. [*452*]
[1767]* Souslin M. *Sur une définition des ensembles mesurables B sans nombres transfinis*. C. R. Acad. Sci. Paris. 1917. T. 164, №2. P. 89–91.
[1768]* 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.

[1769]* Sprecher D.A. Elements of real analysis. 2nd ed. Dover Publications, New York, 1987; viii+343 pp.
[1770]* Srinivasan T.P. *On measurable sets*. J. Indian Math. Soc. (N.S.) 1954. V. 18. P. 1–8.
[1771]* Srinivasan T.P. *On extensions of measures*. J. Indian Math. Soc. (N.S.) 1955. V. 19. P. 31–60.
[1772] Srivastava S.M. A course on Borel sets. Springer-Verlag, New York, 1998; xvi+261 pp. [*440*]
[1773]* 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.
[1774] Steen L., Seebach J. Counterexamples in topology. 2nd ed. Springer, New York, 1978; xi+244 pp. [*9, 64*]
[1775] Stegall Ch. *The topology of certain spaces of measures.* Topology Appl. 1991. V. 41. P. 73–112. [*167*]
[1776]* Stein E. Singular integrals and differentiability properties of functions. Princeton University Press, Princeton, 1970; xiv+290 pp.
[1777]* Stein E.M. Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals. Princeton University Press, Princeton, 1993; xiii+695 pp.
[1778]* Stein E.M., Shakarchi R. Real analysis. Measure theory, integration, and Hilbert spaces. Princeton University Press, Princeton, 2005; xx+402 pp.
[1779]* Stein E., Weiss G. Introduction to Fourier analysis on Euclidean spaces. Princeton University Press, Princeton, 1971; x+297 pp.
[1780] Stein J.D. *A uniform boundedness theorem for measures.* Michigan Math. J. 1972. V. 19, №2. P. 161–165. [*244*]
[1781]* Steinhaus H. *Additive und stetige Funktionaloperationen.* Math. Z. 1919. B. 5. S. 186–221.
[1782]* Steinhaus H. *Sur les distances des points des ensembles de mesure positive.* Fund. Math. 1920. V. 1. P. 93–104.
[1783]* Steinhaus H. *Les probabilités dénombrables et leur rapport à la théorie de la mesure.* Fund. Math. 1923. V. 4. P. 286–310.
[1784] Steinhaus H. *Sur la probabilité de la convergence de séries.* Studia Math. 1930. V. 2. P. 21–39. [*332, 457*]
[1785]* Stepanoff W. *Sur une propriété caractéristique des fonctions mesurables.* Mat. Sbornik. 1924. V. 31. P. 487–489.
[1786]* Stieltjes T.J. *Recherches sur les fractions continues.* Ann. Fac. Sci. Toulouse (1). 1894. V. 8. P. J.1–J.122.
[1787]* 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.
[1788] Stone M.H. *The theory of representations for Boolean algebras.* Trans. Amer. Math. Soc. 1936. V. 40, №1. P. 37–111. [*442, 461*]
[1789] Stone M.H. *Applications of the theory of Boolean rings to general topology.* Trans. Amer. Math. Soc. 1937. V. 41, №3. P. 375–481. [*442, 461*]
[1790] Stone M.H. *Notes on integration. I–IV.* Proc. Nat. Acad. Sci. U.S.A. 1948. V. 34. P. 336–342; ibid., P. 447–455; ibid., P. 483–490; ibid., 1949. V. 35. P. 50–58. [*442, 445*]
[1791] Strassen V. *The existence of probability measures with given marginals.* Ann. Math. Statist. 1965. V. 36. P. 423–439. [*236, 324, 461*]
[1792] Strauss W., Macheras N.D., Musiał K. *Liftings.* P. 1131–1184 in [**1423**]. [*463*]
[1793] Stricker C., Yor M. *Calcul stochastique dependant d'un paramétre.* Z. Wahrscheinlichkeitstheorie verw. Geb. 1978. V. 45, №2. P. 109–133. [*63*]
[1794]* Stromberg K. *The Banach–Tarskii paradox.* Amer. Math. Monthly. 1979. V. 86 (March). P. 151–161.
[1795]* Stromberg K. An introduction to classical real analysis. Wadsworth, Belmont, 1981; ix+575 pp.

[1796] Stroock D.W. *Some comments on independent σ-algebras.* Colloq. Math. 1976. V. 35, №1. P. 7–13. [*433*]
[1797] Stroock D.W. Probability theory: an analytic view. Cambridge University Press, 1993; xv+512 pp. [*453*]
[1798]* Stroock D.W. A concise introduction to the theory of integration. 3d ed. Birkhäuser Boston, Boston, 1999; xiv+253 pp.
[1799] Stroock D.W., Varadhan S.R.S. Multidimensional diffusion processes. Springer, Berlin – New York, 1979; xii+338 pp. [*453*]
[1800] Sturm K.-T. *On the geometry of metric measure spaces. I,II.* Acta Math. 2006. V. 196, №1. P. 65–131, 133–177. [*454*]
[1801]* Subramanian B. *On the inclusion $L^p(\mu) \subset L^q(\mu)$.* Amer. Math. Monthly. 1978. V. 85, №6. P. 479–481.
[1802]* Sudakov V.N. *Criteria of compactness in function spaces.* Uspehi Mat. Nauk. 1957. V. 12, №3. P. 221–224 (in Russian).
[1803] Sudakov V.N. *Geometric problems of the theory of infinite–dimensional probability distributions.* Trudy Mat. Inst. Steklov. 1976. V. 141. P. 1–190 (in Russian); English transl.: Proc. Steklov Inst. Math. 1979. №2. P. 1–178. [*236, 448, 461*]
[1804]* Suetin P.K. Classical orthogonal polynoms. Nauka, Moscow, 1976; 328 pp. (in Russian).
[1805]* 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.
[1806] Sun Y. *Isomorphisms for convergence structures.* Adv. Math. 1995. V. 116, №2. P. 322–355. [**I**: *237;* **II**: *241, 323*]
[1807] Sun Y. *Isomorphisms for convergence structures. II.* Borel spaces and infinite measures. Adv. Math. 1996. V. 124, №1. P. 77–93. [*241, 323*]
[1808]* Svetic R.E. *The Erdos similarity problem: a survey.* Real Anal. Exchange. 2000–2001. V. 26 (2). P. 525–539.
[1809]* Swartz Ch. Measure, integration and function spaces. World Sci., River Edge, 1994; xii+277 pp.
[1810]* Swartz Ch. Introduction to gauge integrals. World Sci., Singapore, 2001; x+157 pp.
[1811]* Sz.-Nagy B. Introduction to real functions and orthogonal expansions. Oxford University Press, New York, 1965; xi+447 pp.
[1812] Szpilrajn E. *Sur certains invariants de l'opération (A).* Fund. Math. 1933. V. 21. P. 229–235. [**I**: *420;* **II**: *61*]
[1813] Szpilrajn E. *Remarques sur les fonctions complètement additives d'ensemble et sur les ensembles jouissant de la propriété de Baire.* Fund. Math. 1934. V. 22. P. 303–311. [*451*]
[1814]* Szpilrajn E. *Sur l'extension de la mesure lebesguienne.* Fund. Math. 1935. V. 25. P. 551–558.
[1815] Szpilrajn E. *Sur l'équivalence des suites d'ensembles et l'équivalence des fonctions.* Fund. Math. 1936. V. 26. P. 302–326. [*440*]
[1816] Szpilrajn E. *The characteristic function of a sequence of sets and some of its applications.* Fund. Math. 1938. V. 31. P. 207–223. [*440*]
[1817] Szpilrajn E. *Ensembles indépendants et mesures non séparables.* C. R. Acad. Sci. Paris. 1938. T. 207. P. 768–770. [*400*]
[1818] Szpilrajn E. *On the space of measurable sets.* Ann. Soc. Polon. Math. 1938. V. 17. P. 120–121. [*459*]
[1819] Szpilrajn E. *On problems of the theory of measure.* Uspehi Mat. Nauk. 1946. V. 1, №2. P. 179–188 (in Russian); English transl.: see [**1256**]. [*440*]
[1820] Sztencel R. *On the lower tail of stable seminorm.* Bull. Acad. Polon. Sci. 1984. V. 32, №11–12. P. 715–719. [*149, 451*]
[1821] Szulga A. *On minimal metrics in the space of random variables.* Theory Probab. Appl. 1982. V. 27. P. 424–430. [*456*]

[1822]* Szymanski W. *Who was Otto Nikodym?* Math. Intell. 1990. V. 12, №2. P. 27–31.

[1823]* Tagamlickiĭ Ya.A. *On the integration of sequences of functions.* Dokl. Akad. Nauk SSSR. 1947. V. 57, №1. P. 17–19 (in Russian).

[1824] Takahashi Y. *On the relation between Radonifying mappings and kernels of probability measures on Banach spaces.* Hokkaido Math. J. 1985. V. 14, №1. P. 97–106. [*451*]

[1825] Takahashi Y., Okazaki Y. 0-1 *laws of a probability measure on a locally convex space.* Publ. Res. Inst. Math. Sci. 1986. V. 22, №1. P. 97–102. [*410*]

[1826] Talagrand M. *Les boules peuvent engendrer la tribu borélienne d'un espace métrizable non séparable?* Séminaire Choquet, 17e anée. 1977–1978. F. 5, №2, 2 pp. Paris, 1978. [*59*]

[1827] Talagrand M. *Comparaison des boreliens d'un espace de Banach pour les topologies fortes et faibles.* Indiana Univ. Math. J. 1978. V. 27, №6. P. 1001–1004. [*52*]

[1828] Talagrand M. *Est-ce que l^∞ est un espace mesurable?* Bull. Sci. Math. 1979. T. 103, №3. P. 255–258. [*52*]

[1829] Talagrand M. *Hyperplans universellement mesurables.* C. R. Acad. Sci. Paris, Sér. A. 1980. T. 291. P. A501–A502. [*168*]

[1830] Talagrand M. *Separabilité vague dans l'espace des mesures sur un compact.* Israel J. Math. 1980. V. 37, №1–2. P. 171–180. [*230, 455*]

[1831] Talagrand M. *Compacts de fonctions mesurables et filtres non mesurables.* Studia Math. 1980. V. 67. P. 13–43. [*151*]

[1832] Talagrand M. *Closed convex hull of set of measurable functions, Riemann-measurable functions and measurability of translations.* Ann. Inst. Fourier (Grenoble). 1982. V. 32, №1. P. 39–69. [*154, 463*]

[1833] Talagrand M. *Processus canoniquement mesurables (ou: Doob avait raison.* Lecture Notes in Math. 1983. V. 986. P. 502–507. [*448*]

[1834] Talagrand M. Pettis integral and measure theory. Mem. Amer. Math. Soc. 1984. V. 51, №307. P. 1–224. [**I**: *235;* **II**: *52, 153, 452, 463*]

[1835] Talagrand M. *The Glivenko–Cantelli problem.* Ann. Probab. 1987. V. 15. P. 837–870. [*416*]

[1836] Talagrand M. *On liftings and the regularization of stochastic processes.* Probab. Theory Related Fields. 1988. V. 78, №1. P. 127–134. [*463*]

[1837] Talagrand M. *Transportation cost for Gaussian and other product measures.* Geom. Funct. Anal. 1996. V. 6. P. 587–600. [*418, 426*]

[1838]* Talagrand M. *Maharam's problem.* http://arxiv.org: math.FA/0601689. P. 1–30.

[1839] Tarieladze V.I. *Characteristic functionals and cylindrical measures in* DS-*groups.* In: Probability theory and mathematical statistics, V. II (Vilnius, 1985), pp. 625–648. VNU Sci. Press, Utrecht, 1987. [*449*]

[1840] Tarieladze V.I. *On a topological description of characteristic functionals.* Dokl. Akad. Nauk SSSR. 1987. V. 295, №6. P. 1320–1323 (in Russian); English transl.: Soviet Math. Dokl. 1988. V. 36, №1. P. 205–208. [*449*]

[1841] Tarieladze V.I. *Topological description of characteristic functionals on some groups.* Teor. Veroyatn. i Primen. 1989. V. 34, №4. P. 719–730 (in Russian); English transl.: Theory Probab. Appl. 1990. V. 34, №4. P. 658–667. [*449*]

[1842]* Tarski A. *Une contribution à la théorie de la mesure.* Fund. Math. 1930. V. 15. P. 42–50.

[1843]* Taylor A.E. General theory of functions and integration. 2d ed. Dover Publ., New York, 1985; xvi+437 pp. (1st ed. Blaisdell, 1965).

[1844]* 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.

[1845]* Taylor A.E., Dugac P. *Quatre lettres de Lebesgue à Fréchet.* Rev. Histoire Sci. Appl. 1981. V. 34, №2. P. 149–169.

[1846]* Taylor J.C. An introduction to measure and probability. Springer-Verlag, New York, 1997; xviii+299 pp.

[1847]* Taylor S.J. Introduction to measure and integration. Cambridge University Press, London, 1973; 266 pp.

[1848]* Telyakovskiĭ S.A. A collection of problems in the theory of functions of a real variable. Nauka, Moscow, 1980; 112 pp. (in Russian).

[1849]* Temple G. The structure of Lebesgue integration theory. Clarendon Press, Oxford, 1971; 184 pp.

[1850]* Ter Horst H.J. *Riemann–Stieltjes and Lebesgue–Stieltjes integrability.* Amer. Math. Monthly. 1984. V. 91. P. 551–559.

[1851]* Thielman H. Theory of functions of a real variable. Prentice-Hall, Englewood Cliffs, New Jersey, 1953; 209 pp.

[1852]* Thomson B.S. Real functions. Lecture Notes in Math. V. 1170. Springer-Verlag, Berlin – New York, 1985; vii+229 pp.

[1853]* Thomson B.S. *Differentiation.* P. 179–247 in [**1423**].

[1854] Thorisson H. Coupling, stationarity, and regeneration. Springer-Verlag, New York, 2000; 517 pp. [*441*]

[1855] Tien N.D., Tarieladze V. *Probability measures with big kernels.* Georgian Math. J. 2001. V. 8, №2. P. 333-346. [*451*]

[1856]* Tikhomirov V.M. *The discovery of A-sets.* Istor.-Matem. Issled. 1993. №34. P. 129–139 (in Russian).

[1857]* Titchmarsh E.C. *The convergence of certain integrals.* Proc. London Math. Soc. (2). 1926. V. 24. P. 347–358.

[1858]* Titchmarsh E. The theory of functions. Clarendon Press, Oxford, 1932; x+454 pp.

[1859]* Titchmarsh E.C. Introduction to the theory of Fourier integrals. 3d ed. Chelsea, New York, 1986; x+394 pp.

[1860] Tjur T. Conditional probability distributions. Lecture Notes, No. 2. Institute of Mathematical Statistics, University of Copenhagen, Copenhagen, 1974; iv+370 pp. [*462*]

[1861] Tjur T. Probability based on Radon measures. John Wiley & Sons, Chichester, 1980; xi+232 pp. [*452*]

[1862]* Tkadlec J. *Construction of a finite Borel measure with σ-porous sets as null sets.* Real Anal. Exchange. 1986–1987. №1. P. 349–353.

[1863] 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. [**I**: *159;* **II**: *165*]

[1864]* 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).

[1865]* Tolstov G.P. *On change of variables in multiple integrals.* Uspehi Mat. Nauk. 1950. V. 5, №4. P. 162–169 (in Russian).

[1866]* Tolstow (Tolstov) G.P. Maß und Integral. Translated from the Russian. Akademie-Verlag, Berlin, 1981; 296 S. (Russian ed.: Moscow, 1976).

[1867]* Tonelli L. *Sull'integrazione per parti.* Atti Accad. Lincei Rend. Cl. sci. fis., mat. e natur. Roma. (5). 1909. V. 18. P. 246–253.

[1868]* Tonelli L. *Sulla nozione di integrale.* Ann. Mat. Pura Appl. (4). 1923. T. 1. P. 105–145.

[1869]* Tonelli L. Opere scelte. V. 1. Edizioni Cremonese, Roma, 1960.

[1870] Topsøe F. *Preservation of weak convergence under mappings.* Ann. Math. Statist. 1967. V. 38, №6. P. 1661–1665. [*453*]

[1871] Topsøe F. *A criterion for weak convergence of measures with an application to convergence of measures on $D[0,1]$.* Math. Scand. 1969. V. 25. P. 97–104. [*453*]

[1872] Topsøe F. *Compactness in spaces of measures.* Studia Math. 1970. V. 36, №3. P. 195–212. [*244, 453*]

[1873] Topsøe F. Topology and measure. Lecture Notes in Math. V. 133. Springer-Verlag, Berlin – New York, 1970; xiv+79 pp. [*192, 227, 443, 452, 453*]

[1874] Topsøe F. *Compactness and tightness in a space of measures with the topology of weak convergence*. Math. Scand. 1974. V. 34, №2. P. 187–210. [*217, 224, 453*]

[1875] Topsøe F. *Some special results on convergent sequences of Radon measures*. Manuscr. Math. 1976. V. 19. P. 1–14. [*453*]

[1876] Topsøe F. *Further results on integral representations*. Studia Math. 1976. V. 55, №3. P. 239–245. [*447*]

[1877] Topsøe F. *Uniformity in weak convergence with respect to balls in Banach spaces*. Math. Scand. 1976. V. 38, №1. P. 148–158. [*453*]

[1878] 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. [**I**: *421;* **II**: *443, 452*]

[1879] Topsøe F. *Approximating pavings and construction of measures*. Colloq. Math. 1979. V. 42. P. 377–385. [*443, 452*]

[1880] Topsøe F. *Radon measures, some basic constructions*. Lecture Notes in Math. 1983. V. 1033. P. 303–311. [*443, 452*]

[1881] Topsøe F. *The Souslin operation in topology and measure theory, selected topics*. In: Proceedings of the Conference Topology and Measure III. Part 1,2. J. Flachsmeyer, Z. Frolík, Ju.M. Smirnov, F. Topsøe, and F. Terpe eds., pp. 283–312. Ernst-Moritz-Arndt-Universitat, Greifswald, 1982. [*440*]

[1882] Topsøe F., Hoffmann-Jørgensen J. Analytic spaces and their applications. In: Analytic sets, pp. 317–403, Proc. Symp. London Math. Soc., Academic Press, New York, 1980. [*440*]

[1883]* Toralballa L.V. Theory of functions. Charles Merill Books, Ohio, 1963; 671 pp.

[1884]* Torchinsky A. Real variable methods in harmonic analysis. Academic Press, New York, 1986; 462 pp.

[1885]* Torchinsky A. Real variables. Addison-Wesley, New York, 1988; 403 pp.

[1886]* Tornier E. Wahrscheinlichkeitsrechnung und allgemeine Integrationstheorie. Teubner, Leipzig, 1936; 160 S.

[1887] Tortrat A. Calcul des probabilités et introduction aux processus aléatoires. Masson, Paris, 1971; xiv+303 pp. [**I**: *414;* **II**: *443*]

[1888] Tortrat A. *Lois* $e(\lambda)$ *dans les espaces vectoriels et lois stables*. Z. Wahrscheinlichkeitstheorie verw. Geb. 1976. B. 37, №2. S. 175–182. [*149, 451*]

[1889] Tortrat A. τ*-regularité des lois, séparation au sens de A. Tulcea et propriété de Radon–Nikodym*. Ann. Inst. H. Poincaré Sect. B (N.S.). 1976. V. 12, №2. P. 131–150; Addendum: ibid. 1977. V. 13. P. 43. [*443*]

[1890] Tortrat A. *Prolongements* τ*-réguliers. Applications aux probabilités gaussiennes*. Sympos. Math. Inst. Naz. Alta Mat. V. 21. P. 117–138. London – New York, 1977. [*443, 452*]

[1891]* Townsend E.J. Functions of real variables. Henry Holt, New York, 1928; ix+405 pp.

[1892] Traynor T. *An elementary proof of the lifting theorem*. Pacif. J. Math. 1974. V. 53, №1. P. 267–272. [*463*]

[1893]* Tricomi F.G. Istituzioni di analisi superiore (metodi matematici della fisica). 2 ed. CEDAM, Padua, 1970; ix+472 pp.

[1894] Tuero A. *On the stochastic convergence of representations based on Wasserstein metrics*. Ann. Probab. 1993. V. 21, №1. P. 72–85. [*454*]

[1895]* Tumakov I.M. Henri Leon Lebesgue (1875–1941). Nauka, Moscow, 1975; 119 pp. (in Russian).

[1896] Uglanov A.V. Integration on infinite-dimensional surfaces and its applications. Kluwer Academic Publ., Dordrecht, 2000; 262 pp. [*448*]

[1897]* Ulam S. *Concerning functions of sets*. Fund. Math. 1929. V. 14. P. 231–233.

[1898] Ulam S. *Zur Masstheorie in der allgemeinen Mengenlehre*. Fund. Math. 1930. V. 16. P. 140–150. [**I**: *77;* **II**: *77*]

[1899] Ulam S. *On the distribution of a general measure in any complete metric space*. Bull. Amer. Math. Soc. 1938. V. 44. P. 786. [*443*]

[1900]* 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).
[1901] Umemura Y. *Measures on infinite dimensional vector spaces.* Publ. RIMS Kyoto Univ. 1965. V. 1, №1. P. 1–47. [*448*]
[1902] Urbanik K. *Random linear functionals and random integrals.* Colloq. Math. 1975. V. 38, №2. P. 255–263. [*149, 451*]
[1903]* 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.
[1904] Ursell H.D. *Some methods of proving measurability.* Fund. Math. 1939. V. 32. P. 311–330. [*161*]
[1905] Üstünel A.S., Zakai M. Transformation of measure on Wiener space. Springer-Verlag, Berlin, 2000; xiv+296 pp. [*460*]
[1906]* Väisälä J. *Quasiconformal maps and positive boundary measure.* Analysis. 1989. V. 9. P. 205–216.
[1907] Vakhania N.N. Probability distributions on linear spaces. Tbilisi, 1971 (in Russian); English transl.: North-Holland, Amsterdam, 1981; xv+123 pp. [*448*]
[1908]* Vakhania N.N., Kvaratskhelia V.V. *Unconditional convergence of weakly subgaussian series in Banach spaces.* Theory Probab. Appl. 2006.
[1909] Vakhania N.N., Tarieladze V.I. *Covariance operators of probability measures in locally convex spaces.* Teor. Verojatn. i Primen. 1978. V. 23. P. 3–26 (in Russian); English transl.: Theor. Probab. Appl. 1978. V. 23. P. 1–21. [*143*]
[1910] Vakhania N.N., Tarieladze V.I., Chobanyan S.A. Probability distributions in Banach spaces. Translated from the Russian. Kluwer Academic Publ., 1991; xxvi+482 pp. (Russian ed.: Moscow, 1985). [*125, 144, 148, 167, 172, 443, 448, 451, 452, 453*]
[1911] Valadier M. *Désintégration d'une mesure sur un produit.* C. R. Acad. Sci. Paris A–B. 1973. T. 276. P. A33–A35. [*405, 462*]
[1912] Valadier M. *Young measures.* Lecture Notes in Math. 1990. V. 1446. P. 152–188. [*231, 231*]
[1913] Valadier M. *A course on Young measures.* Workshop di Teoria della Misura e Analisi Reale (Grado, 1993). Rend. Istit. Mat. Univ. Trieste. 1994. T. 26, suppl., pp. 349–394. [*231*]
[1914] Vallander S.S. *Calculations of the Vasseršteĭn distance between probability distributions on the line.* Teor. Verojatn. i Primen. 1973. V. 18. P. 824–827 (in Russian); English transl.: Probab. Theory Appl. 1973. V. 18. P. 784–786. [*263*]

Vallée Poussin Ch.J. de la: see la Vallée Poussin Ch.J. de [**1111**]–[**1114**]

[1915] Vaart A.W. van der, Wellner J.A. Weak convergence and empirical processes. With applications to statistics. Springer-Verlag, New York, 1996; xvi+508 pp. [*456*]
[1916]* 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.
[1917] Varadarajan V.S. *Groups of automorphisms of Borel spaces.* Trans. Amer. Math. Soc. 1963. V. 109. P. 191–220. [*455, 458*]
[1918] Varadarajan V.S. *Measures on topological spaces.* Mat. Sbornik. 1961. V. 55. P. 35–100 (in Russian); English transl.: Amer. Math. Soc. Transl. (2). 1965. V. 48. P. 161–228. [*166, 197, 250, 443, 447, 452, 455*]
[1919] Vasershtein L.N. *Markov processes over denumerable products of spaces describing large system of automata.* Probl. Peredači Inform. 1969. V. 5, №3. P. 64–72 (in Russian); English transl.: Problems Inform. Transmission. 1969. V. 5, №3. P. 47–52. [*454*]
[1920]* Väth M. Integration theory. A second course. World Sci., River Edge, 2002; viii+277 pp.
[1921]* Veress P. *Über Funktionenmengen.* Acta Sci. Math. Szeged. 1927. V. 3. P. 177–192.
[1922]* Verley J.-L. Théorie élémentaire de l'intégration. "Les cours de Sorbonne". 2e éd. Centre de docum. Université Paris-V, Paris, 1968; 185 pp.

[1923] Vershik A.M. *A theory of decreasing sequences of measurable partitions.* Algebra and Analysis. 1994. V. 6, №4. P. 1–68 (in Russian); English transl.: St. Petersburg Math. J. 1995. V. 6, №4. P. 705–761. [*463*]

[1924] Vershik A.M. *The universal Uryson space, Gromov metric triples and random metrics on the natural numbers.* Uspehi Matem. Nauk. 1998. V. 53, №5. P. 57–64 (in Russian); English transl.: Russian Math. Surveys. 1998. V. 53, №5. P. 921–928. [*459*]

[1925] Vershik A.M. *The Kantorovich metric: the initial history and little-known applications.* Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. 2004. T. 312. P. 69–85 (in Russian); English transl.: J. Math. Sci. (New York). 2006. V. 133, №4. P. 1410–1417. [*453*]

[1926] Vershik A.M., Sudakov V.N. *Probability measures in infinite-dimensional spaces.* Zap. Nauchn. Sem. LOMI. 1969. V. 12. P. 7–67 (in Russian); English transl.: Seminars in Math. V.A. Steklov Math. Inst., Leningrad. 1971. V. 12. P. 1–28. [*448*]

[1927]* Vestrup E.M. The theory of measures and integration. Wiley-Interscience, Hoboken, New Jersey, 2003; xviii+594 pp.

[1928] Villani C. Topics in optimal transportation. Amer. Math. Soc., Rhode Island, 2003; 370 pp. [*236*]

[1929] Vinokurov V.G. *Compact measures and products of Lebesgue spaces.* Mat. Sbornik. 1967. V. 74. P. 434–472 (in Russian); English transl.: Math. USSR Sb. 1967. V. 3. P. 405–433. [*320, 444, 459*]

[1930] Vinokurov V.G., Mahkamov B.M. *On spaces with a perfect, but not compact measure.* Nauchn. Zap. Tashkent Inst. Narodn. Hoz. 1973. №7. P. 97–103 (in Russian). [*89, 444*]

[1931]* Vinti C. Lezioni sulla teoria dell'integrazione. Galeno, Perugia, 1988; 206 pp.

[1932]* Visintin A. *Strong convergence results related to strict convexity.* Comm. Partial Differential Equations. 1984. V. 9, №5. P. 439–466.

[1933]* Vitali G. *Sulla integrabilità delle funzioni.* Rend. Ist. Lombardo (2). 1904. T. 37. P. 69–73.

[1934]* Vitali G. *Sui gruppi di punti.* Rend. Circ. Mat. Palermo. 1904. T. 18. P. 116–126.

[1935]* Vitali G. *Una proprietà delle funzioni misurabili.* Rend. R. Ist. Lombardo (2). 1905. T. 38. P. 600–603.

[1936]* Vitali G. *Sulle funzioni ad integrale nullo.* Rend. Circ. Mat. Palermo. 1905. T. 20. P. 136–141.

[1937]* Vitali G. *Sulle funzioni integrali.* Atti. R. Accad. sci. Torino. 1904–1905. T. 40. P. 1021–1034.

[1938]* Vitali G. *Sul problema della misura dei gruppi di punti di una retta.* Bologna, Tipogr. Gamberini e Parmeggiani, 1905; 5 pp.

[1939]* Vitali G. *Sull'integrazione per serie.* Rend. Circ. Mat. Palermo. 1907. T. 23. P. 137–155.

[1940]* Vitali G. *Sui gruppi di punti e sulle funzioni di variabili reali.* Atti. R. Accad. sci. Torino. 1908. T. 43. P. 75–92.

[1941]* Vitali G. *Sulle funzioni continue.* Fund. Math. 1926. V. 8. P. 175–188.

[1942]* 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.

[1943]* Vitali G. Geometria nello spazio Hilbertiano. N. Zanichelli, Bologna, 1929; 283 pp.

[1944]* Vitali G. Opere sull'analisi reale e complessa. Carteggio. Edizioni Cremonese, Bologna, 1984.

[1945]* 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)

[1946]* Vitushkin A.G. On multidimensional variations. GITTL, Moscow, 1955; 220 pp. (in Russian).

[1947] 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. [**I**: *421;* **II**: *280, 326*]
[1948]* Vogel W. Wahrscheinlichkeitstheorie. Vandenhoeck & Ruprecht, Göttingen, 1970; 385 S.
[1949]* Vo-Khac Kh. Mesure, intégration, convolution, & analyse de Fourier. Interprétation dans le langage des probabilités. Ellipses, Paris, 1984; 256 pp.
[1950]* 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.
[1951]* 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.
[1952]* Volcic A. Lezioni sull'integrale. 2a ed. Edizioni goliardiche, Udine, 1997; 156 pp.
[1953]* Volterra V. *Sui principii del calcolo integrale*. G. Battaglini. 1881. T. 19. P. 333–372.
[1954]* Vulikh B.Z. A brief course of the theory of functions of a real variable. Nauka, Moscow, 1973; 352 pp. (in Russian); English transl.: Mir, Moscow, 1976; 356 pp.
[1955] Wage M.L. *The product of Radon spaces*. Uspehi Mat. Nauk. 1980. V. 35, №3. P. 151–153 (in Russian); English transl.: Russian Math. Surveys. 1980. V. 35. P. 185–187. [*135, 171*]
[1956] Wagner D. *Survey of measurable selection theorems*. SIAM J. Control Optim. 1977. V. 15. P. 859–903. [*441*]
[1957] Wagner D. *Survey of measurable selection theorems: an update*. Lecture Notes in Math. 1980. V. 794. P. 119–184. [*441*]
[1958]* Wagon S. The Banach–Tarski paradox. Cambridge University Press, Cambridge, 1993; xviii+253 pp.
[1959]* Wagschal C. Dérivation, intégration. Hermann, Paris, 1999; viii+472 pp.
[1960]* Wagschal C. Intégration. Exercices et problèmes corrigés. Hermann, Paris, 1999; vi+130 pp.
[1961]* Walter W. Analysis 2. 5e Aufl. Springer, Berlin, 2002, xiv+408 S.
[1962]* Wang Z.Y., Klir G.J. Fuzzy measure theory. Plenum Press, New York, 1992; x+354 pp.
[1963]* Ważewski T. *Sur les ensembles mesurables*. C. R. Acad. Sci. Paris. 1923. T. 176. P. 69–70.
[1964]* Weber H. *Ein Fortsetzungssatz für gruppenwertige Maße*. Arch. Math. 1980. B. 34. S. 157–159.
[1965] Weil A. L'intégration dans les groupes topologiques et ses applications. Hermann et Cie, Paris, 1940; 158 pp. [*442, 460*]
[1966]* Weir A.J. Lebesgue integration and measure. Cambridge University Press, London – New York, 1973; xii+281 pp.
[1967]* Weir A.J. General integration and measure. Cambridge University Press, London – New York, 1974; xi+298 pp.
[1968] Weizsäcker H. von. *Der Satz von Choquet–Bishop-de Leeuw für konvexe nicht kompakte Mengen straffer Maße über beliebigen Grundräumen*. Math. Z. 1975. B. 142. S. 161–165. [*146*]
[1969] Weizsäcker H. von. *A note on infinite dimensional convex sets*. Math. Scand. 1976. V. 38, №2. P. 321–324. [*168*]
[1970] Weizsäcker H. von. *Can one drop* L^1*-boundedness in Komlós' subsequence theorem?* Amer. Math. Monthly. 2004. V. 111, №10. P. 900–903. [*415*]
[1971] Weizsäcker H. von, Winkler G. *Integral representation in the set of solutions of a generalized moment problem*. Math. Ann. 1979. B. 246. S. 23–32. [*146*]
[1972] Wells B.B. Jr. *Weak compactness of measures*. Proc. Amer. Math. Soc. 1969. V. 20, №3. P. 124–134. [*244*]

[1973] Wentzell A.D. A course in the theory of stochastic processes. Translated from the Russian. McGraw-Hill International Book, New York, 1981; x+304 pp. (Russian ed.: Moscow, 1975). [*98*]

[1974]* Wesler O. *An infinite packing theorem for spheres.* Proc. Amer. Math. Soc. 1960. V. 11. P. 324–326.

[1975]* Weyl H. *Über die Konvergenz von Reihen, die nach Orthogonalfunktionen fortschreiten.* Math. Ann. 1909. B. 67. S. 225–245.

[1976] Weyl H. *Über die Gleichverteilung von Zahlen mod. Eins.* Math. Ann. 1916. B. 77. S. 313–352. [*237, 257*]

[1977]* Wheeden R.L., Zygmund A. Measure and integral. Marcel Dekker, New York – Basel, 1977; 274 pp.

[1978] Wheeler R. *Extensions of a σ-additive measure to the projective cover.* Lecture Notes in Math. 1980. V. 794. P. 81–104. [*131, 156*]

[1979] Wheeler R.F. *A survey of Baire measures and strict topologies.* Expos. Math. 1983. V. 1, №2. P. 97–190. [*131, 156, 212, 443, 450, 455*]

[1980]* Whitney H. *Totally differentiable and smooth functions.* Pacif. J. Math. 1951. V. 1. P. 143–159.

[1981] Wichura M.J. *On the construction of almost uniformly convergent random variables with given weakly convergent image laws.* Ann. Math. Statist. 1970. V. 41, №1. P. 284–291. [*454*]

[1982] Wichura M.J. *A note on the weak convergence of stochastic processes.* Ann. Math. Statist. 1971. V. 42, №5. P. 1769–1772. [*251*]

[1983]* Widom H. Lectures on measure and integration. Van Nostrand Reinhold, New York, 1969; viii+166 pp.

[1984] Wiener N. *Differential space.* J. Math. Phys. 1923. V. 2. P. 131–174. [**I**: *419;* **II**: *447*]

[1985]* Wiener N. The Fourier integral and certain of its applications. Dover, New York, 1958; 201 pp.

[1986] Wiener N. Collected works. V. 1–3, The MIT Press, Cambridge, 1976–1981. [**I**: *409, 419;* **II**: *447*]

[1987] Wiener N., Paley R. Fourier transforms in the complex domain. Amer. Math. Soc., New York, 1934; viii+184 pp. [**I**: *430;* **II**: *445, 458*]

[1988] Wijsman R.A. Invariant measures on groups and their use in statistics. Institute of Mathematical Statistics, Hayward, California, 1990; viii+238 pp. [*451*]

[1989]* Wilcox H.J., Myers D.L. An introduction to Lebesgue integration and Fourier series. Dover Publications, New York, 1994; viii+159 pp.

[1990]* Williams D. Probability with martingales. Cambridge University Press, Cambridge, 1994; xv+251 pp.

[1991]* Williamson J.H. Lebesgue integration. Holt, Rinehart and Winston, New York, 1962; viii+117 pp.

[1992] Wilson R.J. *Weak convergence of probability measures in spaces of smooth functions.* Stoch. Proc. Appl. 1986. V. 23, №2. P. 333–337. [*456*]

[1993] Wise G.L., Hall E.B. Counterexamples in probability and real analysis. Oxford University Press, New York – Oxford, 1994; xii+211 pp. [**I**: *81, 228, 395, 414;* **II**: *59, 171*]

[1994] Wiśniewski A. *Theorem of Kuratowski–Suslin for measurable mappings.* Proc. Amer. Math. Soc. 1995. V. 123, №5. P. 1475–1479. [*460*]

[1995] Wójcicka M. *The space of probability measures on a Prohorov space is Prohorov.* Bull. Polish Acad. Sci., Math. 1987. V. 35, №11–12. P. 809–811. [*223*]

[1996]* Wolff J. *Sur les séries* $\sum \frac{A_k}{z-\alpha_k}$. C. R. Acad. Sci. Paris. 1921. T. 173. P. 1056–1057.

[1997]* Wolff T. *Recent work connected with the Kakeya problem.* Prospects in Mathematics, pp. 129–162, Amer. Math. Soc., Providence, Rhode Island, 1999.

[1998] Woyczynski W.A. Geometry and Martingales in Banach spaces. Part II : independent increaments. Probability in Banach Spaces, 4, pp. 267–507. Dekker, New York – Basel, 1978. [*461*]

[1999] Xia D.X. Measure and integration theory on infinite-dimensional spaces. Abstract harmonic analysis. Academic Press, New York – London, 1972; x+425 pp. [*448*]

[2000] Yamasaki Y. Measures on infinite-dimensional spaces. World Sci., Singapore, 1985; x+256 pp. [*448*]

[2001]* Yeh J. Lectures on real analysis. World Sci., River Edge, 2000; xvi+548 pp.

[2002]* Yosida K., Hewitt E. *Finitely additive measures.* Trans. Amer. Math. Soc. 1952. V. 72. P. 46–66.

[2003]* Young G.C., Young W.H. Selected papers. Presses Polytechniques et Universitaires Romandes, Lausanne, 2000; x+870 pp.

[2004] Young L.C. Lectures on the calculus of variations and optimal control theory. W.B. Saunders, Philadelphia – London – Toronto, 1969; 338 pp. [*456*]

[2005]* Young W.H. *The general theory of integration.* Proc. Royal Soc. London. 1904. V. 73. P. 445–449.

[2006]* Young W.H. *Open sets and the theory of content.* Proc. London Math. Soc. (2). 1905. V. 2. P. 16–51.

[2007]* Young W.H. *Upper and lower integration.* Proc. London Math. Soc. (2). 1905. V. 2. P. 52–66.

[2008]* Young W.H. *On the general theory of integration.* Phil. Trans. Royal Soc. London (A). 1905. V. 204. P. 221–252.

[2009]* Young W.H. *On parametric integration.* Monatsh. Math. Phys. 1910. B. 21. S. 125–149.

[2010] Young W.H. *On a new method in the theory of integration.* Proc. London Math. Soc. (2). 1911. V. 9. P. 15–50. [**I**: *423;* **II**: *445*]

[2011]* Young W.H. *On semi-integrals and oscillating successions of functions.* Proc. London Math. Soc. (2). 1911. V. 9. P. 286–324.

[2012]* Young W.H. *On successions of integrals and Fourier series.* Proc. London Math. Soc. 1913. V. 11. P. 43–95.

[2013] Young W.H. *On the new theory of integration.* Proc. Royal Soc., Ser. A. 1913. V. 88. P. 170–178. [**I**: *423;* **II**: *445*]

[2014]* Young W.H. *On derivatives and their primitive functions.* Proc. London Math. Soc. 1913. V. 12. P. 207–217.

[2015] Young W.H. *On integration with respect to a function of bounded variation.* Proc. London Math. Soc. 1914. V. 13. P. 109–150. [**I**: *425;* **II**: *445*]

[2016]* 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.

[2017]* Young W.H. *On successions with subsequences converging to an integral.* Proc. London Math. Soc. (2). 1926. V. 24. P. 1–20.

[2018]* Younovitch B. *Sur la derivation des fonctions absolument additives d'ensemble.* C. R. (Dokl.) Acad. Sci. URSS. 1941. V. 30. P. 112–114.

[2019]* 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.

[2020] Zaanen A.C. Integration. North-Holland, Amsterdam, 1967; xiii+604 pp. [**I**: *310, 312, 320, 414, 438;* **II**: *446*]

[2021]* Zahn P. Ein konstruktiver Weg zur Masstheorie und Funktionalanalysis. Wissenschaftliche Buchgesellschaft, Darmstadt, 1978; 350 S.

[2022]* Zahorski Z. *Sur la première dérivée.* Trans. Amer. Math. Soc. 1950. V. 69. P. 1–54.

[2023]* Zajíček L. *Sets of σ-porosity and sets of σ-porosity (q).* Časopis Pěst. Mat. 1976. V. 101. P. 350–359.

[2024] Zakharov V.K., Mikhalev A.V. *The problem of the general Radon representation for an arbitrary Hausdorff space.* Izv. Ross. Akad. Nauk Ser. Mat. 1999. V. 63, №5. P. 37–82; *II*: ibid., 2002. V. 66, №6. P. 3–18 (in Russian); English transl.: Izv. Math. 1999. V. 63, №5. P. 881–921; 2002. V. 66, №6. P. 1087–1101. [*447*]

[2025]* Zalcman L. *Uniqueness and nonuniqueness for the Radon transform*. Bull. London Math. Soc. 1982. V. 14, №3. P. 241–245.

[2026]* Zamansky M. Introduction à l'algèbre et l'analyse modernes. 3me éd. Dunod, Paris, 1967; xviii+435 pp.

[2027]* Zhang G.Y. *Intersection bodies and the Busemann–Petty inequalities in* R^4. Ann. Math. (2). 1994. V. 140, №2. P. 331–346.

[2028]* Zhang G. *A positive solution to the Busemann-Petty problem in* $\mathbf{R}^4$. Ann. Math. 1999. V. 149, №2. P. 535–543.

[2029] Zięba W. *A note on conditional Fatou lemma*. Probab. Theory Related Fields. 1988. V. 78. P. 73–74. [*428*]

[2030]* Ziemer W. Weakly differentiable functions. Springer-Verlag, New York – Berlin, 1989; xvi+308 pp.

[2031]* Zink R.E. *On the structure of measure spaces*. Acta Math. 1962. V. 107. P. 53–71. [**I**: *93;* **II**: *160*]

[2032] Zinn J. *Zero-one laws for non-Gaussian measures*. Proc. Amer. Math. Soc. 1974. V. 44, №1. P. 179–185. [*410*]

[2033] Zolotarev V.M. One-dimensional stable distributions. Translated from the Russian. Amer. Math. Soc., Providence, Rhode Island, 1986; x+284 pp. (Russian ed.: Moscow, 1983). [*149*]

[2034] Zolotarev V.M. *Probabilistic metrics*. Teor. Verojatn. i Primen. 1983. V. 28, №2. P. 264–287 (in Russian); English transl.: Theory Probab. Appl. 1984. V. 28, №2. P. 278–302. [*456*]

[2035] Zolotarev V.M. Modern theory of summation of random variables. Translated from the Russian. VSP, Utrecht, 1997; x+412 pp. (Russian ed.: Moscow, 1986). [*456*]

[2036]* 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).

[2037]* Zubieta Russi G. Teoría de la integral. Editorial Trillas, Mexico City, 1970; 222 pp.

[2038]* Zygmund A. Trigonometric series. V. I,II. 3d ed. Cambridge University Press, Cambridge, 2002; xiv+383 pp., viii+364 pp.

Author Index

[1]The labels **I** and **II** indicate the volume.

Subject Index

Notation:

[1]The labels **I** and **II** indicate the volume.