MEASURE THEORY

Volume 1

D.H.Fremlin

By the same author:
Topological Riesz Spaces and Measure Theory, Cambridge University Press, 1974.
Consequences of Martin's Axiom, Cambridge University Press, 1982.

Companions to the present volume:
Measure Theory, vol. 2, Torres Fremlin, 2001;
Measure Theory, vol. 3, Torres Fremlin, 2002;
Measure Theory, vol. 4, Torres Fremlin, 2003;
Measure Theory, vol. 5, Torres Fremlin, 2008.

First edition May 2000

Second edition January 2011

MEASURE THEORY

Volume 1

The Irreducible Minimum

D.H.Fremlin

Research Professor in Mathematics, University of Essex

Dedicated by the Author
to the Publisher

This book may be ordered from the printers, `http://www.lulu.com/buy`.

First published in 2000
by Torres Fremlin, 25 Ireton Road, Colchester CO3 3AT, England

Library of Congress classification QA312.F72

AMS 2010 classification 28-01

ISBN 978-0-9538129-8-1

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Printed by Lulu.com

Contents

General Introduction 6

Introduction to Volume 1 7

Chapter 11: Measure Spaces

Introduction 9

111 σ-algebras 10

Definition of σ-algebra; countable sets; σ-algebra generated by a family of sets; Borel σ-algebras.

112 Measure spaces 14

Definition of measure space; the use of ∞; elementary properties; negligible sets; point-supported measures; image measures.

113 Outer measures and Carathéodory's construction 19

Outer measures; Carathéodory's construction of a measure from an outer measure.

114 Lebesgue measure on $\mathbb{R}$ 23

Half-open intervals; Lebesgue outer measure; Lebesgue measure; Borel sets are measurable.

115 Lebesgue measure on $\mathbb{R}^r$ 28

Half-open intervals; Lebesgue outer measure; Lebesgue measure; Borel sets are measurable.

Chapter 12: Integration

Introduction 35

121 Measurable functions 35

Subspace σ-algebras; measurable real-valued functions; partially defined functions; Borel measurable functions; operations on measurable functions; generating Borel sets from half-spaces.

122 Definition of the integral 43

Simple functions; non-negative integrable functions; integrable real-valued functions; virtually measurable functions; linearity of the integral.

123 The convergence theorems 52

B.Levi's theorem; Fatou's lemma; Lebesgue's Dominated Convergence Theorem; differentiating through an integral.

Chapter 13: Complements

Introduction 56

131 Measurable subspaces 56

Subspace measures on measurable subsets; integration over measurable subsets.

132 Outer measures from measures 58

The outer measure associated with a measure; Lebesgue outer measure again; measurable envelopes.

133 Wider concepts of integration 61

∞ as a value of an integral; complex-valued functions; upper and lower integrals.

134 More on Lebesgue measure 68

Translation-invariance; non-measurable sets; inner and outer regularity; the Cantor set and function; *the Riemann integral.

135 The extended real line 79

The algebra of $\pm\infty$; Borel sets and convergent sequences in $[-\infty,\infty]$; measurable and integrable $[-\infty,\infty]$-valued functions.

*136 The Monotone Class Theorem 84

The σ-algebra generated by a family $\mathcal{I}$; algebras of sets.

Appendix to Volume 1

Introduction 89

1A1 Set theory 89

Notation; countable and uncountable sets.

1A2 Open and closed sets in $\mathbb{R}^r$ 92

Definitions; basic properties of open and closed sets; Cauchy's inequality; open balls.

1A3 Lim sups and lim infs 94

$\limsup_{n\to\infty} a_n$, $\liminf_{n\to\infty} a_n$ in $[-\infty,\infty]$.

Concordance 97

References for Volume 1 97

Index to Volume 1
Principal topics and results 98
General index 99

General Introduction

In this treatise I aim to give a comprehensive description of modern abstract measure theory, with some indication of its principal applications. The first two volumes are set at an introductory level; they are intended for students with a solid grounding in the concepts of real analysis, but possibly with rather limited detailed knowledge. The emphasis throughout is on the mathematical ideas involved, which in this subject are mostly to be found in the details of the proofs.

My intention is that the book should be usable both as a first introduction to the subject and as a reference work. For the sake of the first aim, I try to limit the ideas of the early volumes to those which are really essential to the development of the basic theorems. For the sake of the second aim, I try to express these ideas in their full natural generality, and in particular I take care to avoid suggesting any unnecessary restrictions in their applicability. Of course these principles are to to some extent contradictory. Nevertheless, I find that most of the time they are very nearly reconcilable, *provided* that I indulge in a certain degree of repetition. For instance, right at the beginning, the puzzle arises: should one develop Lebesgue measure first on the real line, and then in spaces of higher dimension, or should one go straight to the multidimensional case? I believe that there is no single correct answer to this question. Most students will find the one-dimensional case easier, and it therefore seems more appropriate for a first introduction, since even in that case the technical problems can be daunting. But certainly every student of measure theory must at a fairly early stage come to terms with Lebesgue area and volume as well as length; and with the correct formulations, the multidimensional case differs from the one-dimensional case only in a definition and a (substantial) lemma. So what I have done is to write them both out (in §§114-115), so that you can pass over the higher dimensions at first reading (by omitting §115) and at the same time have a complete and uncluttered argument for them (if you omit section §114). In the same spirit, I have been uninhibited, when setting out exercises, by the fact that many of the results I invite students to look for will appear in later chapters; I believe that throughout mathematics one has a better chance of understanding a theorem if one has previously attempted something similar alone.

The plan of the work is as follows:

Volume 1: The Irreducible Minimum
Volume 2: Broad Foundations
Volume 3: Measure Algebras
Volume 4: Topological Measure Spaces
Volume 5: Set-theoretic Measure Theory.

Volume 1 is intended for those with no prior knowledge of measure theory, but competent in the elementary techniques of real analysis. I hope that it will be found useful by undergraduates meeting Lebesgue measure for the first time. Volume 2 aims to lay out some of the fundamental results of pure measure theory (the Radon-Nikodým theorem, Fubini's theorem), but also gives short introductions to some of the most important applications of measure theory (probability theory, Fourier analysis). While I should like to believe that most of it is written at a level accessible to anyone who has mastered the contents of Volume 1, I should not myself have the courage to try to cover it in an undergraduate course, though I would certainly attempt to include some parts of it. Volumes 3 and 4 are set at a rather higher level, suitable to postgraduate courses; while Volume 5 will assume a wide-ranging competence over large parts of analysis and set theory.

There is a disclaimer which I ought to make in a place where you might see it in time to avoid paying for this book. I make no attempt to describe the history of the subject. This is not because I think the history uninteresting or unimportant; rather, it is because I have no confidence of saying anything which would not be seriously misleading. Indeed I have very little confidence in anything I have ever read concerning the history of ideas. So while I am happy to honour the names of Lebesgue and Kolmogorov and Maharam in more or less appropriate places, and I try to include in the bibliographies the works which I have myself consulted, I leave any consideration of the details to those bolder and better qualified than myself.

For the time being, at least, printing will be in short runs. I hope that readers will be energetic in commenting on errors and omissions, since it should be possible to correct these relatively promptly. An inevitable consequence of this is that paragraph references may go out of date rather quickly. I shall be most flattered if anyone chooses to rely on this book as a source for basic material; and I am willing to attempt to maintain a concordance to such references, indicating where migratory results have come to rest for the moment, if authors will supply me with copies of papers which use them. In the concordance to the present volume you will find notes on the items which have been referred to in other published volumes of this work.

I mention some minor points concerning the layout of the material. Most sections conclude with lists of 'basic exercises' and 'further exercises', which I hope will be generally instructive and occasionally entertaining. How many of these you should attempt must be for you and your teacher, if any, to decide, as no two students will have quite the same needs. I mark with a **>** those which seem to me to be particularly important. But while you may not need to write out solutions to all the 'basic exercises', if you are in any doubt as to your capacity to do so you should take this as a warning to slow down a bit. The 'further exercises' are unbounded in difficulty, and are unified only by a presumption that each has at least one solution based on ideas already introduced.

The impulse to write this treatise is in large part a desire to present a unified account of the subject. Cross-references are correspondingly abundant and wide-ranging. In order to be able to refer freely across the whole text, I have chosen a reference system which gives the same code name to a paragraph wherever it is being called from. Thus 132E is the fifth paragraph in the second section of Chapter 13, which is itself the third chapter of this volume, and is referred to by that name throughout. Let me emphasize that cross-references are supposed to help the reader, not distract him. Do not take the interpolation '(121A)' as an instruction, or even a recommendation, to turn back to §121. If you are happy with an argument as it stands, independently of the reference, then carry on. If, however, I seem to have made rather a large jump, or the notation has suddenly become opaque, local cross-references may help you to fill in the gaps.

Each volume will have an appendix of 'useful facts', in which I set out material which is called on somewhere in that volume, and which I do not feel I can take for granted. Typically the arrangement of material in these appendices is directed very narrowly at the particular applications I have in mind, and is unlikely to be a satisfactory substitute for conventional treatments of the topics touched on. Moreover, the ideas may well be needed only on rare and isolated occasions. So as a rule I recommend you to ignore the appendices until you have some direct reason to suppose that a fragment may be useful to you.

During the extended gestation of this project I have been helped by many people, and I hope that my friends and colleagues will be pleased when they recognise their ideas scattered through the pages below. But I am especially grateful to those who have taken the trouble to read through earlier drafts and comment on obscurities and errors. In particular, I should like to single out F.Nazarov and P.Wallace Thompson, whose thorough reading of the present volume corrected many faults.

Introduction to Volume 1

In this introductory volume I set out, at a level which I hope will be suitable for students with no prior knowledge of the Lebesgue (or even Riemann) integral and with only a basic (but thorough) preparation in the techniques of ϵ-δ analysis, the theory of measure and integration up to the convergence theorems (§123). I add a third chapter (Chapter 13) of miscellaneous additional results, mostly chosen as being relatively elementary material necessary for topics treated in Volume 2 which does not have a natural place there.

The title of this volume is a little more emphatic than I should care to try to justify *au pied de la lettre*. I would certainly characterise the construction of Lebesgue measure on $\mathbb{R}$ (§114), the definition of the integral on an abstract measure space (§122) and the convergence theorems (§123) as indispensable. But a teacher who wishes to press on to further topics will find that much of Chapter 13 can be set aside for a while. I say 'teacher' rather than 'student' here, because if you are studying on your own I think you should aim to go slower than the text requires rather than faster; in my view, these ideas are genuinely difficult, and I think you should take the time to get as much practice at relatively elementary levels as you can.

Perhaps this is a suitable moment at which to set down some general thoughts on the teaching of measure theory. I have been teaching analysis for over thirty years now, and one of the few constants over that period has been the feeling, almost universal among teachers of analysis, that we are not serving most of our students well. We have all encountered students who are not stupid – who are indeed quite good at mathematics – but who seem to have a disproportionate difficulty with rigorous analysis. They are so exhausted and demoralised by the technical problems that they cannot make sense or use even of the knowledge they achieve. The natural reaction to this is to try to make courses shorter and easier. But I think that this makes it even more likely that at the end of the semester your students will be stranded in thorn-bushes half way up the mountain. Specifically, with Lebesgue measure, you are in danger of spending twenty hours teaching them how to integrate the characteristic function of the rationals. This is not what the subject is for. Lebesgue's own presentations of the subject (Lebesgue 1904, Lebesgue 1918) emphasize the convergence theorems and the Fundamental Theorem of Calculus. I have put the former in Volume 1 and the latter in Volume 2, but it does seem to me that unless your students themselves want to know when one can expect to be able to interchange a limit and an integral, or which functions are indefinite integrals, or what the completions of $C([0,1])$ under the norms $\| \|_1$, $\| \|_2$ look like, then it is going to be very difficult for them to make anything of this material; and if you really cannot reach the point of explaining at least a couple of these matters in terms which they can appreciate, then it may not be worth starting. I would myself choose rather to omit a good many proofs than to come to the theorems for which the subject was created so late in the course that two thirds of my class have already given up before they are covered.

Of course I and others have followed that road too, with no better results (though usually with happier students) than we obtain by dotting every i and crossing every t in the proofs. Nearly every time I am consulted by a non-specialist who wants to be told a theorem which will solve his problem, I am reminded that pure mathematics in general, and analysis in particular, does not lie in the *theorems* but in the *proofs*. In so far as I have been successful in answering such questions, it has usually been by making a trifling adjustment to a standard argument to produce a non-standard theorem. The ideas are in the details. You have not understood Carathéodory's construction (§113) until you can, at the very least, reliably reproduce the argument which shows that it works. In the end, there is no alternative to going over every step of the ground, and while I have occasionally been ruthless in cutting out topics which seem to me to be marginal, I have tried to make sure – at the expense, frequently, of pedantry – that every necessary idea is signalled.

Faced, therefore, with any particular class, I believe that a teacher must compromise between scope and completeness. Exactly which compromises are most appropriate will depend on factors which it would be a waste of time for me to guess at. This volume is supposed to be a possible text on which to base a course; but I hope that no lecturer will set her class to read it at so many pages a week. My primary aim is to provide a concise and coherent basis on which to erect the structure of the later volumes. This involves me in pursuing, at more than one point, approaches which take slightly more difficult paths for the sake of developing a more refined technique. (Perhaps the most salient of these is my insistence that an integrable function need not be defined everywhere on the underlying measure space; see §§121-122.) It is the responsibility of the individual teacher to decide for herself whether such refinements are appropriate to the needs of her students, and, if not, to show them what translations are needed.

The above paragraphs are directed at teachers who are, supposedly, competent in the subject – certainly past the level treated in this volume – and who have access to some of the many excellent books already available, so that if they take the trouble to think out their aims, they should be able to choose which elements of my presentation are suitable. But I must also consider the position of a student who is setting out to learn this material on his own. I trust that you have understood from what I have already written that you should not be afraid to look ahead. You could, indeed, do worse than go to Volume 2, and take one of the wonderful theorems there – the Fundamental Theorem of Calculus (§222), for instance, or, if you are very ambitious, the strong law of large numbers (§273) – and use the index and the cross-references to try to extract a proof from first principles. If you are successful you will have every right to congratulate yourself. In the periods in which success seems elusive, however, you should be working systematically through the 'basic exercises' in the sections which seem to be relevant; and if all else fails, start again at the beginning. Mathematics is a difficult subject, that is why it is worth doing, and almost every section here contains some essential idea which you could not expect to find alone.

Note on second and third printings

For the second printing of this volume I made a few corrections, with a handful of new exercises. For the third printing I have done the same; in addition, I have given an elementary extra result and formal definitions of some almost standard terms. I have also allowed myself, in a couple of cases, to rearrange a set of exercises into what now seems to me a more natural order.

Note on second edition, 2011

For the new ('Lulu') edition of this volume, I have eliminated a number of further errors; no doubt many remain. There are some further exercises, and a little more material on upper and lower integrals (§133).

Chapter 11

Measure spaces

In this chapter I set out the fundamental concept of 'measure space', that is, a set in which some (not, as a rule, all) subsets may be assigned a 'measure', which you may wish to interpret as area, or mass, or volume, or thermal capacity, or indeed almost anything which you would expect to be additive – I mean, that the measure of the union of two disjoint sets should be the sum of their measures. The actual definition (in 112A) is not obvious, and depends essentially on certain technical features which make a preparatory section (§111) advisable. Furthermore, even with the definition well in hand, the original and most important examples of measures, Lebesgue measure on Euclidean space, remain elusive. I therefore devote a section (§113) to a method of constructing measures, before turning to the details of the arguments needed for Lebesgue measure in §§114-115. Thus the structure of the chapter is three sections of general theory followed by two (which are closely similar) on particular examples. I should say that the general theory is essentially easier; but it does rely on facility with certain manipulations of families of sets which may be new to you.

At some point I ought to comment on my arrangement of the material, and it may be helpful if I do so before you start work on this chapter. One of the many fundamental questions which any author on the subject must decide, is whether to begin with 'general' measure theory or with 'Lebesgue' measure and integration. The point is that Lebesgue measure is rather more than just the most important example of a measure space. It is so close to the heart of the subject that the great majority of the ideas of the elementary theory can be fully realised in theorems about Lebesgue measure. Looking ahead to Volume 2, I find that, with the exception of Chapter 21 – which is specifically devoted to extending your ideas of what measure spaces can be – only Chapter 27 and the second half of Chapter 25 really need the general theory to make sense, while Chapters 22, 26 and 28 are specifically about Lebesgue measure. Volume 3 is another matter, but even there more than half the mathematical content can be expressed in terms of Lebesgue measure. If you take the view, as I certainly do when it suits my argument, that the business of pure mathematics is to express and extend the logical capacity of the human mind, and that the actual theorems we work through are merely vehicles for the ideas, then you can correctly point out that all the really important things in the present volume can be done without going to the trouble of formulating a general theory of abstract measure spaces; and that by studying the relatively concrete example of Lebesgue measure on r-dimensional Euclidean space you can avoid a variety of irrelevant distractions.

If you are quite sure, as a teacher, that none of your pupils will wish to go beyond the elementary theory, there is something to be said for this view. I believe, however, that it becomes untenable if you wish to prepare any of your students for more advanced ideas. The difficulty is that, with the best will in the world, anyone who has worked through the full theory of Lebesgue measure, and then comes to the theory of abstract measure spaces, is likely to go through it too fast, and at the end find himself uncertain about just which ninety per cent of the facts he knows are generally applicable. I believe it is safer to keep the special properties of Lebesgue measure clearly labelled as such from the beginning.

It is of course the besetting sin of mathematics teachers at this level, to teach a class of twenty in a manner appropriate to perhaps two of them. But in the present case my own judgement is that very few students who are ready for the course at all will have any difficulty with the extra level of abstraction involved in 'Let (X, Σ, μ) be a measure space, ... '. I do assume knowledge of elementary linear algebra, and the grammar, at least, of arbitrary measure spaces is no worse than the grammar of arbitrary linear spaces. Moreover, the Lebesgue theory already involves statements of the form 'if E is a Lebesgue measurable set, ... ', and in my experience students who can cope with quantification over subsets of the reals are not deterred by quantification over sets of sets (which anyway is necessary for any elementary description of the σ-algebra of Borel sets). So I believe that here, at least, the extra generality of the 'professional' approach is not an obstacle to the amateur.

I have written all this here, rather than later in the chapter, because I do wish to give you the choice. And if your choice is to learn the Lebesgue theory first, and leave the general theory to later, this is how to do it. You should read

paragraphs 114A-114C
114D, with 113A-113B and 112Ba, 112Bc
114E, with 113C-113D, 111A, 112A, 112Bb
114F
114G, with 111G and 111C-111F,

and then continue with Chapter 12. At some point, of course, you should look at the exercises for §§112-113; but, as in Chapters 12-13, you will do so by translating 'Let (X, Σ, μ) be a measure space' into 'Let μ be Lebesgue measure on $\mathbb{R}$, and Σ the σ-algebra of Lebesgue measurable sets'. Similarly, when you look at 111X-111Y, you will take Σ to be *either* the σ-algebra of Lebesgue measurable sets *or* the σ-algebra of Borel subsets of $\mathbb{R}$.

111 σ-algebras

In the introduction to this chapter I remarked that a measure space is 'a set in which some (not, as a rule, all) subsets may be assigned a measure'. All ordinary concepts of 'length' or 'area' or 'volume' apply only to reasonably regular sets. Modern measure theory is remarkably powerful in that an extraordinary variety of sets are regular enough to be measured; but we must still expect some limitation, and when studying any measure a proper understanding of the class of sets which it measures will be central to our work. The basic definition here is that of 'σ-algebra of sets'; all measures in the standard theory are defined on such collections. I therefore begin with a statement of the definition, and a brief discussion of the properties, of these classes.

111A Definition Let X be a set. A **σ-algebra of subsets of** X (sometimes called a **σ-field**) is a family Σ of subsets of X such that

(i) $\emptyset \in \Sigma$;

(ii) for every $E \in \Sigma$, its complement $X \setminus E$ in X belongs to Σ;

(iii) for every sequence $\langle E_n \rangle_{n\in\mathbb{N}}$ in Σ, its union $\bigcup_{n\in\mathbb{N}} E_n$ belongs to Σ.

111B Remarks (a) Almost any new subject in pure mathematics is likely to begin with definitions. At this point there is no substitute for rote learning. These definitions encapsulate years, sometimes centuries, of thought by many people; you cannot expect that they will always correspond to familiar ideas.

(b) Nevertheless, you should always seek immediately to find ways of making new definitions more concrete by finding examples within your previous mathematical experience. In the case of 'σ-algebra', the really important examples, to be described below, are going to be essentially new – supposing, that is, that you need to read this chapter at all. However, two examples should be immediately accessible to you, and you should bear these in mind henceforth:

(i) for any X, $\Sigma = \{\emptyset, X\}$ is a σ-algebra of subsets of X;

(ii) for any X, $\mathcal{P}X$, the set of all subsets of X, is a σ-algebra of subsets of X.

These are of course the smallest and largest σ-algebras of subsets of X, and while we shall spend little time with them, both are in fact significant.

***(c)** The phrase **measurable space** is often used to mean a pair (X, Σ), where X is a set and Σ is a σ-algebra of subsets of X; but I myself prefer to avoid this terminology, unless greatly pressed for time, as in fact many of the most interesting examples of such objects have no useful measures associated with them.

111C Infinite unions and intersections If you have not seen infinite unions before, it is worth pausing over the formula $\bigcup_{n\in\mathbb{N}} E_n$. This is the set of points belonging to one or more of the sets E_n; we may write it as

$$\bigcup_{n\in\mathbb{N}} E_n = \{x : \exists\ n \in \mathbb{N}, x \in E_n\}$$
$$= E_0 \cup E_1 \cup E_2 \cup \dots .$$

(I write $\mathbb{N}$ for the set of natural numbers $\{0, 1, 2, 3, \dots\}$.) In the same way,

$$\bigcap_{n\in\mathbb{N}} E_n = \{x : x \in E_n\ \forall\ n \in \mathbb{N}\}$$
$$= E_0 \cap E_1 \cap E_2 \cap \dots .$$

It is characteristic of the elementary theory of measure spaces that it demands greater facility with the set-operations $\cup$, $\cap$, $\setminus$ ('set difference': $E \setminus F = \{x : x \in E, x \notin F\}$), $\triangle$ ('symmetric difference': $E\triangle F = (E \setminus F) \cup (F \setminus E) = (E \cup F) \setminus (E \cap F)$) than you have probably needed before, with the added complication of infinite unions and intersections. I strongly advise spending at least a little time with Exercise 111Xa at some point.

111D Elementary properties of σ-algebras If Σ is a σ-algebra of subsets of X, then it has the following properties.

(a) $E \cup F \in \Sigma$ for all E, $F \in \Sigma$. **P** For if E, $F \in \Sigma$, set $E_0 = E$, $E_n = F$ for $n \geq 1$; then $\langle E_n \rangle_{n\in\mathbb{N}}$ is a sequence in Σ and $E \cup F = \bigcup_{n\in\mathbb{N}} E_n \in \Sigma$. **Q**

(b) $E \cap F \in \Sigma$ for all E, $F \in \Sigma$. **P** By (ii) of the definition in 111A, $X \setminus E$ and $X \setminus F \in \Sigma$; by (a) of this paragraph, $(X \setminus E) \cup (X \setminus F) \in \Sigma$; by 111A(ii) again, $X \setminus ((X \setminus E) \cup (X \setminus F)) \in \Sigma$; but this is just $E \cap F$. **Q**

(c) $E \setminus F \in \Sigma$ for all E, $F \in \Sigma$. **P** $E \setminus F = E \cap (X \setminus F)$. **Q**

(d) Now suppose that $\langle E_n \rangle_{n\in\mathbb{N}}$ is a sequence in Σ, and consider

$$\begin{aligned}\bigcap_{n\in\mathbb{N}} E_n &= \{x : x \in E_n \ \forall\, n \in \mathbb{N}\}\\ &= E_0 \cap E_1 \cap E_2 \cap \dots\\ &= X \setminus \bigcup_{n\in\mathbb{N}} (X \setminus E_n);\end{aligned}$$

this also belongs to Σ.

111E More on infinite unions and intersections (a) So far I have considered infinite unions and intersections only in the context of sequences $\langle E_n \rangle_{n\in\mathbb{N}}$ indexed by the set $\mathbb{N}$ of natural numbers itself. Many others will arise more or less naturally in the pages ahead. Consider, for instance, sets of the form

$$\bigcup_{n\ge4} E_n = E_4 \cup E_5 \cup E_6 \cup \dots,$$

$$\bigcup_{n\in\mathbb{Z}} E_n = \{x : \exists\ n \in \mathbb{Z},\, x \in E_n\} = \dots \cup E_{-2} \cup E_{-1} \cup E_0 \cup E_1 \cup E_2 \cup \dots,$$

$$\bigcup_{q\in\mathbb{Q}} E_q = \{x : \exists\ q \in \mathbb{Q},\, x \in E_q\},$$

where I write $\mathbb{Z}$ for the set of all integers and $\mathbb{Q}$ for the set of rational numbers. If every E_n, E_q belongs to a σ-algebra Σ, so will these unions. On the other hand,

$$\bigcup_{t\in[0,1]} E_t = \{x : \exists\ t \in [0,1],\, x \in E_t\}$$

may fail to belong to a σ-algebra containing every E_t, and it is of the greatest importance to develop an intuition for those index sets, like $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$, which are 'safe' in this context, and those which are not.

(b) I rather hope that you have seen enough of Cantor's theory of infinite sets to make the following remarks a restatement of familiar material; but if not, I hope that they can stand as a first, and very partial, introduction to these ideas. The point about the first three examples is that we can re-index the families of sets involved as simple sequences of sets. For the first one, this is elementary; write $E'_n = E_{n+4}$ for $n \in \mathbb{N}$, and see that $\bigcup_{n\ge4} E_n = \bigcup_{n\in\mathbb{N}} E'_n \in \Sigma$. For the other two, we need to know something about the sets $\mathbb{Z}$ and $\mathbb{Q}$. We can find sequences $\langle k_n \rangle_{n\in\mathbb{N}}$ of integers, and $\langle q_n \rangle_{n\in\mathbb{N}}$ of rational numbers, such that every integer appears (at least once) as a k_n, and every rational number appears (at least once) as a q_n; that is, the functions $n \mapsto k_n : \mathbb{N} \to \mathbb{Z}$ and $n \mapsto q_n : \mathbb{N} \to \mathbb{Q}$ are surjective. **P** There are many ways of doing this; one is to set

$$\begin{aligned}k_n &= \frac{n}{2} \text{ for even } n,\\ &= -\tfrac{n+1}{2} \text{ for odd } n,\\ q_n &= \tfrac{n-m^3-m^2}{m+1} \text{ if } m \in \mathbb{N} \text{ and } m^3 \le n < (m+1)^3.\end{aligned}$$

(You should check carefully that these formulae do indeed do what I claim they do.) **Q** Now, to deal with $\bigcup_{n\in\mathbb{Z}} E_n$, we can set

$$E'_n = E_{k_n} \in \Sigma$$

for $n \in \mathbb{N}$, so that

$$\bigcup_{n\in\mathbb{Z}} E_n = \bigcup_{n\in\mathbb{N}} E_{k_n} = \bigcup_{n\in\mathbb{N}} E'_n \in \Sigma,$$

while for the other case we have

$$\bigcup_{q\in\mathbb{Q}} E_q = \bigcup_{n\in\mathbb{N}} E_{q_n} \in \Sigma.$$

Note that the first case $\bigcup_{n\ge4} E_n$ can be thought of as an application of the same principle; the map $n \mapsto n+4$ is a surjection from $\mathbb{N}$ onto $\{4, 5, 6, 7, \dots\}$.

111F Countable sets (a) The common feature of the sets $\{n : n \geq 4\}$, $\mathbb{Z}$ and $\mathbb{Q}$ which makes this procedure possible is that they are 'countable'. For our purposes here, the most natural definition of countability is the following: a set K is **countable** if either it is empty or there is a surjection from $\mathbb{N}$ onto K. In this case, if Σ is a σ-algebra of sets and $\langle E_k\rangle_{k\in K}$ is a family in Σ indexed by K, then $\bigcup_{k\in K} E_k \in \Sigma$. **P** For if $n \mapsto k_n : \mathbb{N} \to K$ is a surjection, then $E'_n = E_{k_n} \in \Sigma$ for every $n \in \mathbb{N}$, and $\bigcup_{k\in K} E_k = \bigcup_{n\in\mathbb{N}} E'_n \in \Sigma$. This leaves out the case $K = \emptyset$; but in this case the natural interpretation of $\bigcup_{k\in K} E_k$ is

$$\{x : \exists\ k \in \emptyset,\ x \in E_k\}$$

which is itself $\emptyset$, and therefore belongs to Σ by clause (i) of 111A. **Q** (In a sense this treatment of $\emptyset$ is a conventional matter; but there are various contexts in which we shall wish to discuss $\bigcup_{k\in K} E_k$ without checking whether K actually has any members, and we need to be clear about what we will do in such cases.)

(b) There is an extensive, and enormously important, theory concerning countable sets. The only fragments which I think we must have explicit at this point are the following. (In §1A1 I add a few words to link this presentation to conventional approaches.)

(i) If K is countable and $L \subseteq K$, then L is countable. **P** If $L = \emptyset$, this is immediate. Otherwise, take any $l^* \in L$, and a surjection $n \mapsto k_n : \mathbb{N} \to K$ (of course K also is not empty, as $l^* \in K$); set $l_n = k_n$ if $k_n \in L$, l^* otherwise; then $n \mapsto l_n : \mathbb{N} \to L$ is a surjection. **Q**

(ii) The Cartesian product $\mathbb{N}\times\mathbb{N} = \{(m,n) : m,\ n \in \mathbb{N}\}$ is countable. **P** For each $n \in \mathbb{N}$, let k_n, $l_n \in \mathbb{N}$ be such that $n+1 = 2^{k_n}(2l_n+1)$; that is, k_n is the power of 2 in the prime factorisation of $n+1$, and $2l_n+1$ is the (necessarily odd) number $(n+1)/2^{k_n}$. Now $n \mapsto (k_n, l_n)$ is a surjection from $\mathbb{N}$ to $\mathbb{N}\times\mathbb{N}$. **Q** It will be important to us later to know that $n \mapsto (k_n, l_n)$ is actually a bijection, as is readily checked.

(iii) It follows that if K and L are countable sets, so is $K\times L$. **P** If either K or L is empty, so is $K\times L$, so in this case $K\times L$ is certainly countable. Otherwise, let $\phi : \mathbb{N} \to K$ and $\psi : \mathbb{N} \to L$ be surjections; then we have a surjection $\theta : \mathbb{N}\times\mathbb{N} \to K\times L$ defined by setting $\theta(m,n) = (\phi(m), \psi(n))$ for all m, $n \in \mathbb{N}$. Now we know from (ii) just above that there is also a surjection $\chi : \mathbb{N} \to \mathbb{N}\times\mathbb{N}$, so that $\theta\chi : \mathbb{N} \to K\times L$ is a surjection, and $K\times L$ must be countable. **Q**

(iv) An induction on r now shows us that if K_1, $K_2, \dots, K_r$ are countable sets, so is $K_1\times\ldots\times K_r$. In particular, such sets as $\mathbb{Q}^r\times\mathbb{Q}^r$ will be countable, for any integer $r \geq 1$.

(c) Putting 111Dd above together with these ideas, we see that if Σ is a σ-algebra of sets, K is a non-empty countable set, and $\langle E_k\rangle_{k\in K}$ is a family in Σ, then

$$\bigcap_{k\in K} E_k = \{x : x \in E_k\ \forall\ k \in K\}$$

belongs to Σ. **P** Let $n \mapsto k_n : \mathbb{N} \to K$ be a surjection; then $\bigcap_{k\in K} E_k = \bigcap_{n\in\mathbb{N}} E_{k_n} \in \Sigma$, as in 111Dd. **Q**

Note that there is a difficulty with the notion of $\bigcap_{k\in K} E_k$ if $K = \emptyset$; the natural interpretation of this formula is to read it as the universal class. So ordinarily, when there is any possibility that K might be empty, one needs some such formulation as $X \cap \bigcap_{k\in K} E_k$.

(d) As an example of the way in which these ideas will be used, consider the following. Suppose that X is a set, Σ is a σ-algebra of subsets of X, and $\langle E_{qn}\rangle_{q\in\mathbb{Q},n\in\mathbb{N}}$ is a family in Σ. Then

$$E = \bigcap_{q\in\mathbb{Q},q<\sqrt{2}} \bigcup_{m\in\mathbb{N}} \bigcap_{n\geq m} E_{qn} = \bigcap_{q\in\mathbb{Q},q<\sqrt{2}} (\bigcup_{m\in\mathbb{N}} (\bigcap_{n\geq m} E_{qn})) \in \Sigma.$$

P Set $F_{qm} = \bigcap_{n\geq m} E_{qn} = \bigcap_{n\in\mathbb{N}} E_{q,m+n}$ for $q \in \mathbb{Q}$ and $m \in \mathbb{N}$; then every F_{qm} belongs to Σ, by 111Dd or (c) above. Set $G_q = \bigcup_{m\in\mathbb{N}} F_{qm}$ for $q \in \mathbb{Q}$; then every G_q belongs to Σ, by 111A(iii). Set $K = \{q : q \in \mathbb{Q},\ q < \sqrt{2}\}$; then K is countable, by 111E and (b-i) of this paragraph. So $\bigcap_{q\in K} G_q$ belongs to Σ, by (c). But $E = \bigcap_{q\in K} G_q$. **Q**

(e) And one final remark, which I give without proof here – though many proofs will be implicit in the work below, and I spell one out in 1A1Ha –

The set $\mathbb{R}$ of real numbers is not countable.

So you must resist any temptation to look for a list a_0, $a_1, \dots$ running over the whole set of real numbers.

111G Borel sets I can describe here one type of non-trivial σ-algebra; the formulation is rather abstract, but the technique is important and the terminology is part of the basic vocabulary of measure theory.

(a) Let X be a set, and let $\mathfrak{S}$ be any non-empty family of σ-algebras of subsets of X. (Thus a *member* of $\mathfrak{S}$ is itself a *family* of sets; $\mathfrak{S} \subseteq \mathcal{P}(\mathcal{P}X)$.) Then

$$\bigcap \mathfrak{S} = \{E : E \in \Sigma \text{ for every } \Sigma \in \mathfrak{S}\},$$

the intersection of all the σ-algebras belonging to $\mathfrak{S}$, is a σ-algebra of subsets of X. **P** (i) By hypothesis, $\mathfrak{S}$ is not empty; take $\Sigma_0 \in \mathfrak{S}$; then $\bigcap \mathfrak{S} \subseteq \Sigma_0 \subseteq \mathcal{P}X$, so every member of $\bigcap \mathfrak{S}$ is a subset of X. (ii) $\emptyset \in \Sigma$ for every $\Sigma \in \mathfrak{S}$, so $\emptyset \in \bigcap \mathfrak{S}$. (iii) If $E \in \bigcap \mathfrak{S}$ then $E \in \Sigma$ for every $\Sigma \in \mathfrak{S}$, so $X \setminus E \in \Sigma$ for every $\Sigma \in \mathfrak{S}$ and $X \setminus E \in \bigcap \mathfrak{S}$. (iv) Let $\langle E_n \rangle_{n \in \mathbb{N}}$ be any sequence in $\bigcap \mathfrak{S}$. Then for every $\Sigma \in \mathfrak{S}$, $\langle E_n \rangle_{n \in \mathbb{N}}$ is a sequence in Σ, so $\bigcup_{n \in \mathbb{N}} E_n \in \Sigma$; as Σ is arbitrary, $\bigcup_{n \in \mathbb{N}} E_n \in \bigcap \mathfrak{S}$. **Q**

(b) Now let $\mathcal{A}$ be any family of subsets of X. Consider

$$\mathfrak{S} = \{\Sigma : \Sigma \text{ is a } \sigma\text{-algebra of subsets of } X,\ \mathcal{A} \subseteq \Sigma\}.$$

By definition, $\mathfrak{S}$ is a family of σ-algebras of subsets of X; also, it is not empty, because $\mathcal{P}X \in \mathfrak{S}$. So $\Sigma_{\mathcal{A}} = \bigcap \mathfrak{S}$ is a σ-algebra of subsets of X. Because $\mathcal{A} \subseteq \Sigma$ for every $\Sigma \in \mathfrak{S}$, $\mathcal{A} \subseteq \Sigma_{\mathcal{A}}$; thus $\Sigma_{\mathcal{A}}$ itself belongs to $\mathfrak{S}$; it is the smallest σ-algebra of subsets of X including $\mathcal{A}$.

We say that $\Sigma_{\mathcal{A}}$ is the σ-algebra of subsets of X **generated by** $\mathcal{A}$.

Examples (i) For any X, the σ-algebra of subsets of X generated by $\emptyset$ is $\{\emptyset, X\}$.

(ii) The σ-algebra of subsets of $\mathbb{N}$ generated by $\{\{n\} : n \in \mathbb{N}\}$ is $\mathcal{P}\mathbb{N}$.

(c)(i) We say that a set $G \subseteq \mathbb{R}$ is **open** if for every $x \in G$ there is a $\delta > 0$ such that the open interval $]x - \delta, x + \delta[$ is included in G.

(ii) Similarly, for any $r \ge 1$, we say that a set $G \subseteq \mathbb{R}^r$ is **open** in $\mathbb{R}^r$ if for every $x \in G$ there is a $\delta > 0$ such that $\{y : \|y - x\| < \delta\} \subseteq G$, where for $z = (\zeta_1, \dots, \zeta_r) \in \mathbb{R}^r$ I write $\|z\| = \sqrt{\sum_{i=1}^r |\zeta_i|^2}$; thus $\|y - x\|$ is just the ordinary Euclidean distance from y to x.

(d) Now the **Borel sets** of $\mathbb{R}$, or of $\mathbb{R}^r$, are just the members of the σ-algebra of subsets of $\mathbb{R}$ or $\mathbb{R}^r$ generated by the family of open sets of $\mathbb{R}$ or $\mathbb{R}^r$; the σ-algebra itself is called the **Borel σ-algebra** in each case.

(e) Some readers will rightly feel that the development here gives very little idea of what a Borel set is 'really' like. (Open sets are much easier; see 111Ye.) In fact the importance of the concept derives largely from the fact that there are alternative, more explicit, and in a sense more concrete, ways of describing Borel sets. I shall return to this topic in Chapter 42 in Volume 4.

111X Basic exercises >(a) Practise the algebra of infinite unions and intersections until you can confidently interpret such formulae as

$$E \cap (\bigcup_{n \in \mathbb{N}} F_n), \qquad \bigcup_{n \in \mathbb{N}} (E_n \setminus F), \qquad E \cup (\bigcap_{n \in \mathbb{N}} F_n),$$

$$\bigcup_{n \in \mathbb{N}} (E \setminus F_n), \qquad E \setminus (\bigcup_{n \in \mathbb{N}} F_n), \qquad \bigcap_{n \in \mathbb{N}} (E_n \setminus F),$$

$$E \setminus (\bigcap_{n \in \mathbb{N}} F_n), \qquad \bigcap_{n \in \mathbb{N}} (E \cup F_n), \qquad (\bigcup_{n \in \mathbb{N}} E_n) \setminus F,$$

$$\bigcup_{n \in \mathbb{N}} (E \cap F_n), \qquad (\bigcap_{n \in \mathbb{N}} E_n) \setminus F, \qquad \bigcap_{n \in \mathbb{N}} (E \setminus F_n),$$

$$(\bigcup_{n \in \mathbb{N}} E_n) \cap (\bigcup_{n \in \mathbb{N}} F_n), \qquad \bigcap_{m,n \in \mathbb{N}} (E_m \setminus F_n), \qquad (\bigcap_{n \in \mathbb{N}} E_n) \cup (\bigcap_{n \in \mathbb{N}} F_n),$$

$$\bigcap_{m,n \in \mathbb{N}} (E_m \cup F_n), \qquad (\bigcap_{n \in \mathbb{N}} E_n) \setminus (\bigcup_{n \in \mathbb{N}} F_n), \qquad \bigcup_{m,n \in \mathbb{N}} (E_m \cap F_n),$$

and, in particular, can identify the nine pairs into which these formulae naturally fall.

>(b) In $\mathbb{R}$, show that all 'open intervals' $]a, b[$, $]-\infty, b[$, $]a, \infty[$ are open sets, and that all intervals (bounded or unbounded, open, closed or half-open) are Borel sets.

>(c) Let X and Y be sets and Σ a σ-algebra of subsets of X. Let $\phi : X \to Y$ be a function. Show that $\{F : F \subseteq Y,\ \phi^{-1}[F] \in \Sigma\}$ is a σ-algebra of subsets of Y. (See 1A1B for the notation here.)

>(d) Let X and Y be sets and T a σ-algebra of subsets of Y. Let $\phi : X \to Y$ be a function. Show that $\{\phi^{-1}[F] : F \in \mathrm{T}\}$ is a σ-algebra of subsets of X.

(e) Let X be a set, $\mathcal{A}$ a family of subsets of X, and Σ the σ-algebra of subsets of X generated by $\mathcal{A}$. Suppose that Y is another set and $\phi : Y \to X$ a function. Show that $\{\phi^{-1}[E] : E \in \Sigma\}$ is the σ-algebra of subsets of Y generated by $\{\phi^{-1}[A] : A \in \mathcal{A}\}$.

(f) Let X be a set, $\mathcal{A}$ a family of subsets of X, and Σ the σ-algebra of subsets of X generated by $\mathcal{A}$. Suppose that $Y \subseteq X$. Show that $\{E \cap Y : E \in \Sigma\}$ is the σ-algebra of subsets of Y generated by $\{A \cap Y : A \in \mathcal{A}\}$.

111Y Further exercises (a) In $\mathbb{R}^r$, where $r \geq 1$, show that $G + a = \{x + a : x \in G\}$ is open whenever $G \subseteq \mathbb{R}^r$ is open and $a \in \mathbb{R}^r$. Hence show that $E + a$ is a Borel set whenever $E \subseteq \mathbb{R}^r$ is a Borel set and $a \in \mathbb{R}^r$. (*Hint*: show that $\{E : E + a \text{ is a Borel set}\}$ is a σ-algebra containing all open sets.)

(b) Let X be a set, Σ a σ-algebra of subsets of X and A any subset of X. Show that $\{(E \cap A) \cup (F \setminus A) : E, F \in \Sigma\}$ is a σ-algebra of subsets of X, the σ-algebra generated by $\Sigma \cup \{A\}$.

(c) Let $G \subseteq \mathbb{R}^2$ be an open set. Show that all the horizontal and vertical sections

$$\{\xi : (\xi, \eta) \in G\}, \quad \{\xi : (\eta, \xi) \in G\}$$

of G are open subsets of $\mathbb{R}$.

(d) Let $E \subseteq \mathbb{R}^2$ be a Borel set. Show that all the horizontal and vertical sections

$$\{\xi : (\xi, \eta) \in E\}, \quad \{\xi : (\eta, \xi) \in E\}$$

of E are Borel subsets of $\mathbb{R}$. (*Hint*: show that the family of subsets of $\mathbb{R}^2$ whose sections are all Borel sets is a σ-algebra of subsets of $\mathbb{R}^2$ containing all the open sets.)

(e) Let $G \subseteq \mathbb{R}$ be an open set. Show that G is uniquely expressible as the union of a countable (possibly empty) family $\mathcal{I}$ of open intervals (the 'components' of G) no two of which have any point in common. (*Hint*: for x, $y \in G$ say that $x \sim y$ if every point between x and y belongs to G. Show that $\sim$ is an equivalence relation. Let $\mathcal{I}$ be the set of equivalence classes.)

111 Notes and comments I suppose that the most important concept in this section is the one introduced tangentially in 111E-111F, the idea of 'countable' set. While it is possible to avoid much of the formal theory of infinite sets for the time being, I do not think it is possible to make sense of this chapter without a firm notion of the difference between 'finite' and 'infinite', and some intuitions concerning 'countability'. In particular, you must remember that infinite sets are not, in general, countable, and that σ-algebras are not, in general, closed under arbitrary unions.

The next thing to be sure of is that you can cope with the set-theoretic manipulations here, so that such formulae as $\bigcap_{n\in\mathbb{N}} E_n = X \setminus \bigcup_{n\in\mathbb{N}}(X \setminus E_n)$ (111Dd) are, if not yet transparent, at least not alarming. A large proportion of the volume will be expressed in this language, and reasonable fluency is essential.

Finally, for those who are looking for an actual idea to work on straight away, I offer the concept of σ-algebra 'generated' by a collection $\mathcal{A}$ (111G). The point of the definition here is that it involves consideration of a family $\mathfrak{S} \in \mathcal{P}(\mathcal{P}(\mathcal{P}X))$, even though both $\mathcal{A}$ and $\Sigma_{\mathcal{A}}$ are subsets of $\mathcal{P}X$; we need to work a layer or two up in the hierarchy of power sets. You may have seen, for instance, the concept of 'linear subspace U generated by vectors $u_1, \dots, u_n$'. This can be defined as the intersection of all the linear subspaces containing the vectors $u_1, \dots, u_n$, which is the method corresponding to that of 111Ga-b; but it also has an 'internal' definition, as the set of vectors expressible as $\alpha_1 u_1 + \ldots + \alpha_n u_n$ for scalars α_i. For σ-algebras, however, there is no such simple 'internal' definition available (though there is a great deal to be said in this direction which I think we are not yet ready for; some ideas may be found in §136). This is primarily because of (iii) in the definition 111A; a σ-algebra must be closed under an infinitary operation, that is, the operation of union applied to infinite sequences of sets. By contrast, a linear subspace of a vector space need be closed only under the finitary operations of scalar multiplication and addition, each involving at most two vectors at a time.

112 Measure spaces

We are now, I hope, ready for the second major definition, the definition on which all the work of this treatise is based.

112A Definition A **measure space** is a triple (X, Σ, μ) where

(i) X is a set;

(ii) Σ is a σ-algebra of subsets of X;

(iii) $\mu : \Sigma \to [0, \infty]$ is a function such that

(α) $\mu\emptyset = 0$;

(β) if $\langle E_n \rangle_{n\in\mathbb{N}}$ is a disjoint sequence in Σ, then $\mu(\bigcup_{n\in\mathbb{N}} E_n) = \sum_{n=0}^{\infty} \mu E_n$.

In this context, members of Σ are called **measurable** sets, and μ is called a **measure on** X.

112B Remarks (a) The use of ∞ In (iii) of the definition above, I declare that μ is to be a function taking values in '$[0,\infty]$', that is, the set comprising the non-negative real numbers with '∞' adjoined. I expect that you have already encountered various uses of the symbol ∞ in analysis; I hope you have realised that it means rather different things in different contexts, and that it is necessary to establish clear conventions for its use each time. The '∞ of measure' corresponds to the notion of infinite length or area or volume. The basic operation we need to perform on it is addition: $\infty + a = a + \infty = \infty$ for every $a \in [0,\infty[$ (that is, every real number $a \ge 0$), and $\infty + \infty = \infty$. This renders $[0,\infty]$ a semigroup under addition. It will be reasonably safe to declare $\infty - a = \infty$ for every $a \in \mathbb{R}$; but we must absolutely decline to interpret the formula $\infty - \infty$. As for multiplication, it turns out that it is usually right to interpret $\infty \cdot \infty$, $a \cdot \infty$ and $\infty \cdot a$ as ∞ for $a > 0$, while $0 \cdot \infty = \infty \cdot 0$ can generally be taken as 0.

We also have a natural total ordering of $[0,\infty]$, writing $a < \infty$ for every $a \in [0,\infty[$. This gives an idea of supremum and infimum of an arbitrary (non-empty) subset of $[0,\infty]$; and it will often be right to interpret $\inf \emptyset$ as ∞, but I will try to signal this particular convention each time it is relevant. We also have a notion of limit; if $\langle u_n \rangle_{n\in\mathbb{N}}$ is a sequence in $[0,\infty]$, then it converges to $u \in [0,\infty]$ if

for every $v < u$ there is an $n_0 \in \mathbb{N}$ such that $v \le u_n$ for every $n \ge n_0$,

for every $v > u$ there is an $n_0 \in \mathbb{N}$ such that $v \ge u_n$ for every $n \ge n_0$.

Of course if $u = 0$ or $u = \infty$ then one of these clauses will be vacuously satisfied.

(See also §135.)

(b) I should say plainly what I mean by a 'disjoint' sequence: a sequence $\langle E_n \rangle_{n\in\mathbb{N}}$ is **disjoint** if no point belongs to more than one E_n, that is, if $E_m \cap E_n = \emptyset$ for all distinct m, $n \in \mathbb{N}$. Note that there is no bar here on one, or many, of the E_n being the empty set.

Similarly, if $\langle E_i \rangle_{i\in I}$ is a family of sets indexed by an arbitrary set I, it is **disjoint** if $E_i \cap E_j = \emptyset$ for all distinct i, $j \in I$.

(c) In interpreting clause (iii-β) of the definition above, we need to assign values to sums $\sum_{n=0}^{\infty} u_n$ for arbitrary sequences $\langle u_n \rangle_{n\in\mathbb{N}}$ in $[0,\infty]$. The natural way to do this is to say that $\sum_{n=0}^{\infty} u_n = \lim_{n\to\infty} \sum_{m=0}^{n} u_m$, using the definitions sketched in (a). If one of the u_m is itself infinite, say $u_k = \infty$, then $\sum_{m=0}^{n} u_m = \infty$ for every $n \ge k$, so of course $\sum_{n=0}^{\infty} u_n = \infty$. If all the u_m are finite, then, because they are all non-negative, the sequence $\langle \sum_{m=0}^{n} u_m \rangle_{n\in\mathbb{N}}$ of partial sums is monotonic non-decreasing, and either has a finite limit $\sum_{n=0}^{\infty} u_n \in \mathbb{R}$, or diverges to ∞; in which case we again interpret $\sum_{n=0}^{\infty} u_n$ as ∞.

(d) Once again, the important examples of measure spaces will have to wait until §§114 and 115 below. However, I can describe immediately one particular class of measure space, which should always be borne in mind, though it does not give a good picture of the most important and interesting parts of the subject. Let X be any set, and let $h : X \to [0,\infty]$ be any function. For every $E \subseteq X$ write $\mu E = \sum_{x\in E} h(x)$. To interpret this sum, note that there is no difficulty for finite sets E (taking $\sum_{x\in\emptyset} h(x) = 0$), while for infinite sets E we can take $\sum_{x\in E} h(x) = \sup\{\sum_{x\in I} h(x) : I \subseteq E$ is finite$\}$, because every $h(x)$ is non-negative. (You may well prefer to think about this at first with $X = \mathbb{N}$, so that $\sum_{n\in E} h(n) = \lim_{n\to\infty} \sum_{m\in E, m\le n} h(m)$; but I hope that a little thought will show you that the general case, in which X may even be uncountable, is not really more difficult.) Now $(X, \mathcal{P}X, \mu)$ is a measure space.

We are very far from being ready for the specialized vocabulary needed to describe different kinds of measure space, but when the time comes I will call measures of this kind **point-supported**.

Two particular cases recur often enough to be worth giving names to. If $h(x) = 1$ for every x, then μE is just the number of points of E if E is finite, and is ∞ if E is infinite. I will call this **counting measure** on X. If $x_0 \in X$, we can set $h(x_0) = 1$ and $h(x) = 0$ for $x \in X \setminus \{x_0\}$; then μE is 1 if $x_0 \in E$, and 0 for other E. I will call this the **Dirac measure on X concentrated at** x_0. Another simple example is with $X = \mathbb{N}$, $h(n) = 2^{-n-1}$ for every n; then $\mu X = \frac{1}{2} + \frac{1}{4} + \ldots = 1$.

(e) If (X, Σ, μ) is a measure space, then Σ is the domain of the function μ, and X is the largest member of Σ. We can therefore recover the whole triplet (X, Σ, μ) from its final component μ. This is not a game which is worth playing at this stage. However, it is convenient on occasion to introduce a measure without immediately giving a name to its domain, and when I do this I may say that 'μ **measures** E' or 'E is **measured by** μ' to mean that μE is defined, that is, that E belongs to the σ-algebra $\operatorname{dom}\mu$. **Warning!** Many authors use the phrase 'μ-measurable set' to mean something a little different from what I am discussing here.

112C Elementary properties of measure spaces Let (X, Σ, μ) be a measure space.

(a) If E, $F \in \Sigma$ and $E \cap F = \emptyset$ then $\mu(E \cup F) = \mu E + \mu F$.

(b) If E, $F \in \Sigma$ and $E \subseteq F$ then $\mu E \le \mu F$.

(c) $\mu(E \cup F) \le \mu E + \mu F$ for any E, $F \in \Sigma$.

(d) If $\langle E_n\rangle_{n\in\mathbb{N}}$ is any sequence in Σ, then $\mu(\bigcup_{n\in\mathbb{N}} E_n) \le \sum_{n=0}^{\infty} \mu E_n$.

(e) If $\langle E_n\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence in Σ (that is, $E_n \subseteq E_{n+1}$ for every $n\in\mathbb{N}$) then

$$\mu(\bigcup_{n\in\mathbb{N}} E_n) = \lim_{n\to\infty} \mu E_n = \sup_{n\in\mathbb{N}} \mu E_n.$$

(f) If $\langle E_n\rangle_{n\in\mathbb{N}}$ is a non-increasing sequence in Σ (that is, $E_{n+1} \subseteq E_n$ for every $n\in\mathbb{N}$), and if some μE_n is finite, then

$$\mu(\bigcap_{n\in\mathbb{N}} E_n) = \lim_{n\to\infty} \mu E_n = \inf_{n\in\mathbb{N}} \mu E_n.$$

proof (a) Set $E_0 = E$, $E_1 = F$, $E_n = \emptyset$ for $n \ge 2$; then $\langle E_n\rangle_{n\in\mathbb{N}}$ is a disjoint sequence in Σ and $\bigcup_{n\in\mathbb{N}} E_n = E\cup F$, so

$$\mu(E\cup F) = \sum_{n=0}^{\infty} \mu E_n = \mu E + \mu F$$

(because $\mu\emptyset = 0$).

(b) $F\setminus E \in \Sigma$ (111Dc) and $\mu(F\setminus E) \ge 0$ (because all values of μ are in $[0,\infty]$); so (using (a))

$$\mu F = \mu E + \mu(F\setminus E) \ge \mu E.$$

(c) $\mu(E\cup F) = \mu E + \mu(F\setminus E)$, by (a), and $\mu(F\setminus E) \le \mu F$, by (b).

(d) Set $F_0 = E_0$, $F_n = E_n \setminus \bigcup_{i<n} E_i$ for $n \ge 1$; then $\langle F_n\rangle_{n\in\mathbb{N}}$ is a disjoint sequence in Σ, $\bigcup_{n\in\mathbb{N}} F_n = \bigcup_{n\in\mathbb{N}} E_n$ and $F_n \subseteq E_n$ for every n. By (b) just above, $\mu F_n \le \mu E_n$ for each n; so

$$\mu(\bigcup_{n\in\mathbb{N}} E_n) = \mu(\bigcup_{n\in\mathbb{N}} F_n) = \sum_{n=0}^{\infty} \mu F_n \le \sum_{n=0}^{\infty} \mu E_n.$$

(e) Set $F_0 = E_0$, $F_n = E_n \setminus E_{n-1}$ for $n \ge 1$; then $\langle F_n\rangle_{n\in\mathbb{N}}$ is a disjoint sequence in Σ and $\bigcup_{n\in\mathbb{N}} F_n = \bigcup_{n\in\mathbb{N}} E_n$. Consequently $\mu(\bigcup_{n\in\mathbb{N}} E_n) = \sum_{n=0}^{\infty} \mu F_n$. But an easy induction on n, using (a) for the inductive step, shows that $\mu E_n = \sum_{m=0}^{n} \mu F_m$ for every n. So

$$\sum_{n=0}^{\infty} \mu F_n = \lim_{n\to\infty} \sum_{m=0}^{n} \mu F_m = \lim_{n\to\infty} \mu E_n.$$

Finally, $\lim_{n\to\infty} \mu E_n = \sup_{n\in\mathbb{N}} \mu E_n$ because (by (b)) $\langle \mu E_n\rangle_{n\in\mathbb{N}}$ is non-decreasing.

(f) Suppose that $\mu E_k < \infty$. Set $F_n = E_k \setminus E_{k+n}$ for $n\in\mathbb{N}$, $F = \bigcup_{n\in\mathbb{N}} F_n$; then $\langle F_n\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence in Σ, so $\mu F = \lim_{n\to\infty} \mu F_n$, by (e) just above. Also, $\mu F_n + \mu E_{k+n} = \mu E_k$; because $\mu E_k < \infty$, we may safely write $\mu F_n = \mu E_k - \mu E_{k+n}$, so that

$$\mu F = \lim_{n\to\infty} (\mu E_k - \mu E_{k+n}) = \mu E_k - \lim_{n\to\infty} \mu E_n.$$

Next, $F \subseteq E_k$, so $\mu F + \mu(E_k \setminus F) = \mu E_k$, and (again because μE_k is finite) $\mu F = \mu E_k - \mu(E_k\setminus F)$. Thus we must have $\mu(E_k\setminus F) = \lim_{n\to\infty} \mu E_n$. But $E_k \setminus F$ is just $\bigcap_{n\in\mathbb{N}} E_n$.

Finally, $\lim_{n\to\infty} \mu E_n = \inf_{n\in\mathbb{N}} \mu E_n$ because $\langle \mu E_n\rangle_{n\in\mathbb{N}}$ is non-increasing.

Remark Observe that in (f) above it is essential to have $\inf_{n\in\mathbb{N}} \mu E_n < \infty$. The construction in 112Bd is already enough to show this. Take $X = \mathbb{N}$ and let μ be counting measure on X. Set $E_n = \{i : i\in\mathbb{N},\ i \ge n\}$ for each n. Then $E_{n+1} \subseteq E_n$ for each n, but

$$\mu(\bigcap_{n\in\mathbb{N}} E_n) = \mu\emptyset = 0 < \infty = \lim_{n\to\infty} \mu E_n.$$

112D Negligible sets Let (X,Σ,μ) be any measure space.

(a) A set $A \subseteq X$ is **negligible** (or **null**) if there is a set $E \subseteq \Sigma$ such that $A \subseteq E$ and $\mu E = 0$. (If there seems to be a possibility of doubt about which measure is involved, I will write **μ-negligible**.)

(b) Let $\mathcal{N}$ be the family of negligible subsets of X. Then (i) $\emptyset \in \mathcal{N}$ (ii) if $A \subseteq B \in \mathcal{N}$ then $A \in \mathcal{N}$ (iii) if $\langle A_n\rangle_{n\in\mathbb{N}}$ is any sequence in $\mathcal{N}$, $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{N}$. **P** (i) $\mu(\emptyset) = 0$. (ii) There is an $E\in\Sigma$ such that $\mu E = 0$ and $B \subseteq E$; now $A \subseteq E$. (iii) For each $n\in\mathbb{N}$ choose an $E_n \in \Sigma$ such that $A_n \subseteq E_n$ and $\mu E_n = 0$. Now $E = \bigcup_{n\in\mathbb{N}} E_n \in \Sigma$ and $\bigcup_{n\in\mathbb{N}} A_n \subseteq \bigcup_{n\in\mathbb{N}} E_n$, and $\mu(\bigcup_{n\in\mathbb{N}} E_n) \le \sum_{n=0}^{\infty} \mu E_n$, by 112Cd, so $\mu(\bigcup_{n\in\mathbb{N}} E_n) = 0$. **Q**

I will call $\mathcal{N}$ the **null ideal** of the measure μ. (A family of sets satisfying the conditions (i)-(iii) here is called a **σ-ideal** of sets.)

(c) A set $A \subseteq X$ is **conegligible** if $X\setminus A$ is negligible; that is, there is a measurable set $E \subseteq A$ such that $\mu(X\setminus E) = 0$. Note that (i) X is conegligible (ii) if $A \subseteq B \subseteq X$ and A is conegligible then B is conegligible (iii) if $\langle A_n\rangle_{n\in\mathbb{N}}$ is a sequence of conegligible sets, then $\bigcap_{n\in\mathbb{N}} A_n$ is conegligible.

(d) It is convenient, and customary, to use some relatively informal language concerning negligible sets. If $P(x)$ is some assertion applicable to members x of the set X, we say that

'$P(x)$ for almost every $x \in X$'

or

'$P(x)$ a.e. (x)'

or

'P almost everywhere', 'P a.e.'

or, if it seems necessary to specify the measure involved,

'$P(x)$ for μ-almost every x', '$P(x)\,\mu$-a.e.(x)', '$P\,\mu$-a.e.',

to mean that

$$\{x : x \in X,\ P(x)\}$$

is conegligible in X, that is, that

$$\{x : x \in X,\ P(x) \text{ is false}\}$$

is negligible. Thus, for instance, if $f : X \to \mathbb{R}$ is a function, '$f > 0$ a.e.' means that $\{x : f(x) \le 0\}$ is negligible.

(e) The phrases '**almost surely**' (**a.s.**), '**presque partout**' (**p.p.**) are also used for 'almost everywhere'.

(f) I should call your attention to the fact that, on my definitions, a negligible set need not itself be measurable, though it must be included in some negligible measurable set. (Measure spaces in which all negligible sets are measurable are called **complete**. I will return to this question in §211.)

(g) When f and g are real-valued functions defined on conegligible subsets of a measure space, I will write $f =_{\text{a.e.}} g$, $f \le_{\text{a.e.}} g$ or $f \ge_{\text{a.e.}} g$ to mean, respectively,

$$f = g \text{ a.e., that is, } \{x : x \in \operatorname{dom}(f) \cap \operatorname{dom}(g),\ f(x) = g(x)\} \text{ is conegligible,}$$

$$f \le g \text{ a.e., that is, } \{x : x \in \operatorname{dom}(f) \cap \operatorname{dom}(g),\ f(x) \le g(x)\} \text{ is conegligible,}$$

$$f \ge g \text{ a.e., that is, } \{x : x \in \operatorname{dom}(f) \cap \operatorname{dom}(g),\ f(x) \ge g(x)\} \text{ is conegligible.}$$

112X Basic exercises >(a) Let (X, Σ, μ) be a measure space. Show that (i) $\mu(E \cup F) + \mu(E \cap F) = \mu E + \mu F$ (ii) $\mu(E \cup F \cup G) + \mu(E \cap F) + \mu(E \cap G) + \mu(F \cap G) = \mu E + \mu F + \mu G + \mu(E \cap F \cap G)$ for all E, F, $G \in \Sigma$. Generalize these results to longer sequences of sets. (You may prefer to begin with the case in which μE, μF and μG are all finite. But I hope you will be able to find arguments which deal with the general case.)

>(b) Let (X, Σ, μ) be a measure space and $\langle E_n \rangle_{n \in \mathbb{N}}$ a sequence in Σ. Show that

$$\mu(\bigcup_{n \in \mathbb{N}} \bigcap_{m \ge n} E_m) \le \liminf_{n \to \infty} \mu E_n.$$

(c) Let (X, Σ, μ) be a measure space, and E, $F \in \Sigma$; suppose that $\mu E < \infty$. Show that $\mu(F \triangle E) = \mu F - \mu E + 2\mu(E \setminus F)$.

(d) Let (X, Σ, μ) be a measure space and $\langle E_n \rangle_{n \in \mathbb{N}}$ a sequence of measurable sets such that $\mu(\bigcup_{n \in \mathbb{N}} E_n) < \infty$. (i) Show that $\limsup_{n \to \infty} \mu E_n \le \mu(\bigcap_{n \in \mathbb{N}} \bigcup_{m \ge n} E_m)$. (ii) Show that if $\bigcap_{n \in \mathbb{N}} \bigcup_{m \ge n} E_m = E = \bigcup_{n \in \mathbb{N}} \bigcap_{m \ge n} E_m$ then $\lim_{n \to \infty} \mu E_n$ exists and is equal to μE.

>(e) Let (X, Σ, μ) be a measure space, and $\mathfrak{F}$ the set of real-valued functions whose domains are conegligible subsets of X. (i) Show that $\{(f, g) : f, g \in \mathfrak{F}, f \le_{\text{a.e.}} g\}$ and $\{(f, g) : f, g \in \mathfrak{F}, f \ge_{\text{a.e.}} g\}$ are reflexive transitive relations on $\mathfrak{F}$, each the inverse of the other. (ii) Show that $\{(f, g) : f, g \in \mathfrak{F}, f =_{\text{a.e.}} g\}$ is their intersection, and is an equivalence relation on $\mathfrak{F}$.

(f) Let (X, Σ, μ) be a measure space, Y a set, and $\phi : X \to Y$ a function. Set $\mathrm{T} = \{F : F \subseteq Y,\ \phi^{-1}[F] \in \Sigma\}$ and $\nu F = \mu\phi^{-1}[F]$ for $F \in \mathrm{T}$. Show that ν is a measure on Y. (ν is called the **image measure** on Y, and I will generally denote it $\mu\phi^{-1}$.)

112Y Further exercises (a) Let X be a set and Σ a σ-algebra of subsets of X. Let μ_1 and μ_2 be two measures on X, both with domain Σ. Set

$$\mu E=\inf\{\mu_1(E\cap F)+\mu_2(E\setminus F):F\in\Sigma\}$$

for each $E\in\Sigma$. Show that μ is a measure on X, and that it is the greatest measure, with domain Σ, such that $\mu E\le\min(\mu_1 E,\mu_2 E)$ for every $E\in\Sigma$.

(b) Let X be a set and Σ a σ-algebra of subsets of X. Let μ_1 and μ_2 be two measures on X, both with domain Σ. Set

$$\mu E=\sup\{\mu_1(E\cap F)+\mu_2(E\setminus F):F\in\Sigma\}$$

for each $E\in\Sigma$. Show that μ is a measure on X, and that it is the least measure, with domain Σ, such that $\mu E\ge\max(\mu_1 E,\mu_2 E)$ for every $E\in\Sigma$.

(c) Let X be a set and Σ a σ-algebra of subsets of X.

(i) Suppose that $\nu_0,\dots,\nu_n$ are measures on X, all with domain Σ. Set

$$\mu E=\inf\{\textstyle\sum_{i=0}^n\nu_iF_i:F_0,\dots,F_n\in\Sigma,\ E\subseteq\bigcup_{i\le n}F_i\}$$

for $E\in\Sigma$. Show that μ is a measure on X.

(ii) Let N be a non-empty family of measures on X, all with domain Σ. Set

$$\mu E=\inf\{\sum_{n=0}^\infty\nu_nF_n:\langle\nu_n\rangle_{n\in\mathbb{N}}\text{ is a sequence in N},\ \langle F_n\rangle_{n\in\mathbb{N}}\text{ is a sequence in }\Sigma,\ E\subseteq\bigcup_{n\in\mathbb{N}}F_n\}$$

for $E\in\Sigma$. Show that μ is a measure on X.

(iii) Let N be a non-empty family of measures on X, all with domain Σ, and suppose that there is some $\nu'\in$ N such that $\nu'X<\infty$. Set

$$\mu E=\inf\{\textstyle\sum_{i=0}^n\nu_iF_i:n\in\mathbb{N},\ \nu_0,\dots,\nu_n\in\text{N},\ F_0,\dots,F_n\in\Sigma,E\subseteq\bigcup_{i\le n}F_i\}$$

for $E\in\Sigma$. Show that μ is a measure on X.

(iv) Suppose, in (iii), that N is downwards-directed, that is, for any ν_1, $\nu_2\in$ N there is a $\nu\in$ N such that $\nu E\le\min(\nu_1E,\nu_2E)$ for every $E\in\Sigma$. Show that $\mu E=\inf_{\nu\in\text{N}}\nu E$ for every $E\in\Sigma$.

(v) Show that in all the cases (i)-(iii) the measure constructed is the greatest measure μ with domain Σ such that $\mu E\le\inf_{\nu\in\text{N}}\nu E$ for every $E\in\Sigma$.

(d) Let X be a set and Σ a σ-algebra of subsets of X. Let N be a non-empty family of measures on X, all with domain Σ. Set

$$\mu E=\sup\{\sum_{i=0}^n\nu_iF_i:n\in\mathbb{N},\nu_0,\dots,\nu_n\in\text{N},\ F_0,\dots,F_n\text{ are disjoint subsets of }E\text{ belonging to }\Sigma\}$$

for $E\in\Sigma$. (i) Show that

$$\mu E=\sup\{\sum_{n=0}^\infty\nu_nF_n:\langle\nu_n\rangle_{n\in\mathbb{N}}\text{ is a sequence in N},\ \langle F_n\rangle_{n\in\mathbb{N}}\text{ is a disjoint sequence in }\Sigma,\ \bigcup_{n\in\mathbb{N}}F_n\subseteq E\}$$

for every $E\in\Sigma$. (ii) Show that μ is a measure on X, and that it is the least measure, with domain Σ, such that $\mu E\ge\sup_{\nu\in\text{N}}\nu E$ for every $E\in\Sigma$. (iii) Now suppose that N is upwards-directed, that is, for any ν_1, $\nu_2\in$ N there is a $\nu\in$ N such that $\nu E\ge\max(\nu_1E,\nu_2E)$ for every $E\in\Sigma$. Show that $\mu E=\sup_{\nu\in\text{N}}\nu E$ for every $E\in\Sigma$.

(e) Let (X,Σ,μ) be a measure space and $\langle E_n\rangle_{n\in\mathbb{N}}$ a sequence of measurable sets. For each $k\in\mathbb{N}$ set $H_k=\{x:x\in X,\ \#(\{n:x\in E_n\})\ge k\}$, the set of points belonging to E_n for k or more values of n. (i) Show that each H_k is measurable. (ii) Show that $\sum_{k=1}^\infty\mu H_k=\sum_{n=0}^\infty\mu E_n$. (*Hint*: start with the case in which $E_n=\emptyset$ for

$n \geq n_0$.) (iii) Show that if $\sum_{n=0}^{\infty} \mu E_n$ is finite, then almost every point of X belongs to only finitely many E_n, and $\sum_{n=0}^{\infty} \mu E_n = \sum_{k=0}^{\infty} k\mu G_k$, where

$$G_k = H_k \setminus H_{k+1} = \{x : \#(\{n : x \in E_n\}) = k\}.$$

(f) Let X be a set and μ, ν two measures on X, with domains Σ, T respectively. Set $\Lambda = \Sigma \cap \mathrm{T}$ and define $\lambda : \Lambda \to [0, \infty]$ by setting $\lambda E = \mu E + \nu E$ for every $E \in \Lambda$. Show that (X, Λ, λ) is a measure space.

112 Notes and comments The calculations in such results as 112Ca-112Cc, 112Xa and 112Xc, involving only finitely many sets, are common to any additive concept of measure; you may have encountered them in elementary probability theory, but of course I am now asking you to consider also the possibility that one or more of the sets has measure ∞. I hope you will find that these results are entirely natural and unsurprisin11g. I recommend Venn diagrams in this context; a result of this kind involving only finitely many measurable sets and only *addition*, with no *subtraction*, will be valid in general if and only if it is valid for the area of simple geometric shapes in the plane. The requirement '$\mu E < \infty$' in 112Xc is necessary only because we are subtracting μE; the corresponding additive result $\mu(F \triangle E) + \mu E = \mu F + 2\mu(E \setminus F)$ is true for all measurable E and F. Of course when sequences of sets enter the picture, we need to take a bit more care; the results 112Cd-112Cf are the basic ones to learn. I think however that the only trap is in the condition 'some μE_n is finite' in 112Cf. As noted in the remark at the end of 112C, this is essential, and for a decreasing sequence of measurable sets it is possible for the measure of the limit to be strictly less than the limit of the measures, though only when the latter is infinite.

113 Outer measures and Carathéodory's construction

I introduce the most important method of constructing measures.

113A Outer measures I come now to the third basic definition of this chapter.

Definition Let X be a set. An **outer measure** on X is a function $\theta : \mathcal{P}X \to [0, \infty]$ such that

(i) $\theta\emptyset = 0$,

(ii) if $A \subseteq B \subseteq X$ then $\theta A \leq \theta B$,

(iii) for every sequence $\langle A_n \rangle_{n \in \mathbb{N}}$ of subsets of X, $\theta(\bigcup_{n \in \mathbb{N}} A_n) \leq \sum_{n=0}^{\infty} \theta A_n$.

113B Remarks (a) For comments on the use of '∞', see 112B.

(b) Yet again, the most important outer measures must wait until §§114-115. The idea of the 'outer' measure of a set A is that it should be some kind of upper bound for the possible measure of A. If we are lucky, it may actually be the measure of A; but this is likely to be true only for sets with adequately smooth boundaries.

(c) Putting (i) and (iii) of the definition together, we see that if θ is an outer measure on X, and A, B are two subsets of X, then $\theta(A \cup B) \leq \theta A + \theta B$; compare 112Ca and 112Cc.

113C Carathéodory's Method: Theorem Let X be a set and θ an outer measure on X. Set

$$\Sigma = \{E : E \subseteq X,\ \theta A = \theta(A \cap E) + \theta(A \setminus E) \text{ for every } A \subseteq X\}.$$

Then Σ is a σ-algebra of subsets of X. Define $\mu : \Sigma \to [0, \infty]$ by writing $\mu E = \theta E$ for $E \in \Sigma$; then (X, Σ, μ) is a measure space.

proof (a) The first step is to note that for any E, $A \subseteq X$ we have $\theta(A \cap E) + \theta(A \setminus E) \geq \theta A$, by 113Bc; so that

$$\Sigma = \{E : E \subseteq X,\ \theta A \geq \theta(A \cap E) + \theta(A \setminus E) \text{ for every } A \subseteq X\}.$$

(b) Evidently $\emptyset \in \Sigma$, because

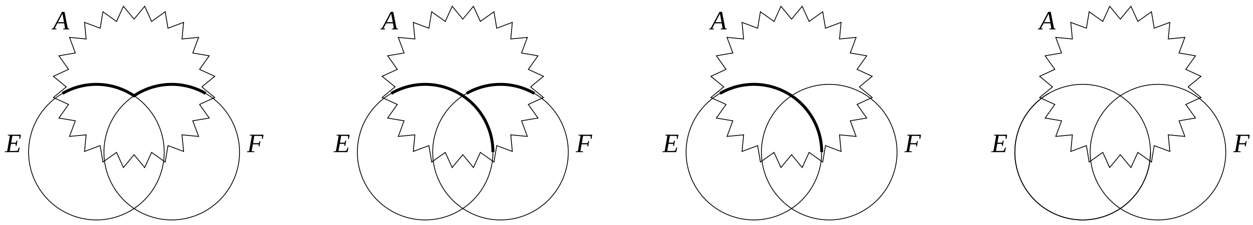
$$\theta(A \cap \emptyset) + \theta(A \setminus \emptyset) = \theta\emptyset + \theta A = \theta A$$

for every $A \subseteq X$. If $E \in \Sigma$, then $X \setminus E \in \Sigma$, because

$$\theta(A \cap (X \setminus E)) + \theta(A \setminus (X \setminus E)) = \theta(A \setminus E) + \theta(A \cap E) = \theta A$$

for every $A \subseteq X$.

(c) Now suppose that E, $F \in \Sigma$ and $A \subseteq X$. Then



(i) (ii) (iii) (iv)

$$\theta(A \cap (E \cup F)) + \theta(A \setminus (E \cup F)) \qquad \text{diagram (i)}$$

$$= \theta(A \cap (E \cup F) \cap E) + \theta(A \cap (E \cup F) \setminus E) + \theta(A \setminus (E \cup F)) \qquad \text{diag. (ii)}$$

(because $E \in \Sigma$ and $A \cap (E \cup F) \subseteq X$)

$$= \theta(A \cap E) + \theta((A \setminus E) \cap F) + \theta((A \setminus E) \setminus F)$$

$$= \theta(A \cap E) + \theta(A \setminus E) \qquad \text{diag. (iii)}$$

(because $F \in \Sigma$)

$$= \theta A \qquad \text{diag. (iv)}$$

(again because $E \in \Sigma$). Because A is arbitrary, $E \cup F \in \Sigma$.

(d) Thus Σ is closed under simple unions and complements, and contains $\emptyset$. Now suppose that $\langle E_n \rangle_{n\in\mathbb{N}}$ is a sequence in Σ, with $E = \bigcup_{n\in\mathbb{N}} E_n$. Set

$$G_n = \bigcup_{m \le n} E_m;$$

then $G_n \in \Sigma$ for each n, by induction on n. Set

$$F_0 = G_0 = E_0, \quad F_n = G_n \setminus G_{n-1} = E_n \setminus G_{n-1} \text{ for } n \ge 1;$$

then $E = \bigcup_{n\in\mathbb{N}} F_n = \bigcup_{n\in\mathbb{N}} G_n$.

Take any $n \ge 1$ and any $A \subseteq X$. Then

$$\begin{aligned}\theta(A \cap G_n) &= \theta(A \cap G_n \cap G_{n-1}) + \theta(A \cap G_n \setminus G_{n-1}) \\ &= \theta(A \cap G_{n-1}) + \theta(A \cap F_n).\end{aligned}$$

An induction on n shows that $\theta(A \cap G_n) = \sum_{m=0}^{n} \theta(A \cap F_m)$ for every $n \ge 0$.

Suppose that $A \subseteq X$. Then $A \cap E = \bigcup_{n\in\mathbb{N}} A \cap F_n$, so

$$\begin{aligned}\theta(A \cap E) &\le \sum_{n=0}^{\infty} \theta(A \cap F_n) \\ &= \lim_{n\to\infty} \sum_{m=0}^{n} \theta(A \cap F_m) = \lim_{n\to\infty} \theta(A \cap G_n).\end{aligned}$$

On the other hand,

$$\begin{aligned}\theta(A \setminus E) &= \theta(A \setminus \bigcup_{n\in\mathbb{N}} G_n) \\ &\le \inf_{n\in\mathbb{N}} \theta(A \setminus G_n) = \lim_{n\to\infty} \theta(A \setminus G_n),\end{aligned}$$

using 113A(ii) to see that $\langle \theta(A \setminus G_n) \rangle_{n\in\mathbb{N}}$ is non-increasing and that $\theta(A \setminus E) \le \theta(A \setminus G_n)$ for every n. Accordingly

$$\begin{aligned}\theta(A \cap E) + \theta(A \setminus E) &\le \lim_{n\to\infty} \theta(A \cap G_n) + \lim_{n\to\infty} \theta(A \setminus G_n) \\ &= \lim_{n\to\infty} (\theta(A \cap G_n) + \theta(A \setminus G_n)) = \theta A\end{aligned}$$

because every G_n belongs to Σ, so $\theta(A \cap G_n) + \theta(A \setminus G_n) = \theta A$ for every n. But A is arbitrary, so $E \in \Sigma$, by the remark in (a) above.

Because $\langle E_n\rangle_{n\in\mathbb{N}}$ is arbitrary, condition (iii) of 111A is satisfied, and Σ is a σ-algebra of subsets of X.

(e) Now let us turn to μ, the restriction of θ to Σ, and Definition 112A. Of course $\mu\emptyset=\theta\emptyset=0$. So let $\langle E_n\rangle_{n\in\mathbb{N}}$ be any disjoint sequence in Σ. Set $G_n=\bigcup_{m\le n}E_m$ for each n, as in (d), and

$$E=\bigcup_{n\in\mathbb{N}}E_n=\bigcup_{n\in\mathbb{N}}G_n.$$

As in (d),

$$\begin{aligned}\mu G_{n+1}=\theta G_{n+1}&=\theta(G_{n+1}\cap E_{n+1})+\theta(G_{n+1}\setminus E_{n+1})\\&=\theta E_{n+1}+\theta G_n=\mu E_{n+1}+\mu G_n\end{aligned}$$

for each n, so $\mu G_n=\sum_{m=0}^{n}\mu E_m$ for every n.

Now

$$\mu E=\theta E\le\sum_{n=0}^{\infty}\theta E_n=\sum_{n=0}^{\infty}\mu E_n.$$

But also

$$\mu E=\theta E\ge\theta G_n=\mu G_n=\sum_{m=0}^{n}\mu E_m$$

for each n, so $\mu E\ge\sum_{n=0}^{\infty}\mu E_n$.

Accordingly $\mu E=\sum_{n=0}^{\infty}\mu E_n$. As $\langle E_n\rangle_{n\in\mathbb{N}}$ is arbitrary, 112A(iii-β) is satisfied and (X,Σ,μ) is a measure space.

113D Remark Note from (a) in the proof above that in this construction

$$\Sigma=\{E:E\subseteq X,\ \theta(A\cap E)+\theta(A\setminus E)\le\theta A\text{ for every }A\subseteq X\}.$$

Since $\theta(A\cap E)+\theta(A\setminus E)$ is necessarily less than or equal to θA when $\theta A=\infty$,

$$\Sigma=\{E:E\subseteq X,\ \theta(A\cap E)+\theta(A\setminus E)\le\theta A\text{ whenever }A\subseteq X\text{ and }\theta A<\infty\}.$$

113X Basic exercises >(a) Let X be a set and θ an outer measure on X, and let μ be the measure on X defined from θ by Carathéodory's method. Show that if $\theta A=0$, then μ measures A, so that a set $A\subseteq X$ is μ-negligible iff $\theta A=0$, and μ is 'complete' in the sense of 112Df.

(b) Let X be a set. (i) Show that if θ_1, θ_2 are outer measures on X, so is $\theta_1+\theta_2$, setting $(\theta_1+\theta_2)(A)=\theta_1A+\theta_2A$ for every $A\subseteq X$. (ii) Show that if $\langle\theta_i\rangle_{i\in I}$ is any non-empty family of outer measures on X, so is $\theta=\sup_{i\in I}\theta_i$, setting $\theta A=\sup_{i\in I}\theta_iA$ for every $A\subseteq X$. (iii) Show that if θ_1, θ_2 are outer measures on X so is $\theta_1\wedge\theta_2$, setting

$$(\theta_1\wedge\theta_2)(A)=\inf\{\theta_1B+\theta_2(A\setminus B):B\subseteq A\}$$

for every $A\subseteq X$.

>(c) Let X and Y be sets, θ an outer measure on X, and $f:X\to Y$ a function. Show that the functional $B\mapsto\theta(f^{-1}[B]):\mathcal{P}Y\to[0,\infty]$ is an outer measure on Y.

>(d) Let X be a set and θ an outer measure on X; let Y be any subset of X. (i) Show that $\theta\restriction\mathcal{P}Y$, the restriction of θ to subsets of Y, is an outer measure on Y. (ii) Show that if $E\subseteq X$ is measured by the measure on X defined from θ by Carathéodory's method, then $E\cap Y$ is measured by the measure on Y defined from $\theta\restriction\mathcal{P}Y$.

>(e) Let X and Y be sets, θ an outer measure on Y, and $f:X\to Y$ a function. Show that the functional $A\mapsto\theta(f[A]):\mathcal{P}X\to[0,\infty]$ is an outer measure.

(f) Let X and Y be sets, θ an outer measure on X, and $R\subseteq X\times Y$ a relation. Show that the map $B\mapsto\theta(R^{-1}[B]):\mathcal{P}Y\to[0,\infty]$ is an outer measure on Y, where $R^{-1}[B]=\{x:\exists\,y\in B,\ (x,y)\in R\}$ (1A1Bc). Explain how this is a common generalization of (c), (d-i) and (e) above, and how it can be proved by putting them together.

(g) Let X be a set and θ an outer measure on X. Suppose that $E\subseteq X$ is measured by the measure on X defined from θ by Carathéodory's method. Show that $\theta(E\cap A)+\theta(E\cup A)=\theta E+\theta A$ for every $A\subseteq X$.

(h) Let X be a set and $\theta:\mathcal{P}X\to[0,\infty]$ a functional such that $\theta\emptyset=0$, $\theta A\le\theta B$ whenever $A\subseteq B\subseteq X$, and $\theta(A\cup B)\le\theta A+\theta B$ whenever A, $B\subseteq X$. Set

$$\Sigma=\{E:E\subseteq X,\ \theta A=\theta(A\cap E)+\theta(A\setminus E)\text{ for every }A\subseteq X\}.$$

Show that $\emptyset$, $X\setminus E$ and $E\cup F$ belong to Σ for all E, $F\in\Sigma$, so that $E\setminus F$, $E\cap F\in\Sigma$ for all E, $F\in\Sigma$. Show that $\theta(E\cup F)=\theta E+\theta F$ whenever E, $F\in\Sigma$ and $E\cap F=\emptyset$.

113Y Further exercises (a) Let (X,Σ,μ) be a measure space. For $A\subseteq X$ set $\mu^*A=\inf\{\mu E:E\in\Sigma,\,A\subseteq E\}$. Show that for every $A\subseteq X$ the infimum is attained, that is, there is an $E\in\Sigma$ such that $A\subseteq E$ and $\mu E=\mu^*A$. Show that μ^* is an outer measure on X.

(b) Let (X,Σ,μ) be a measure space and D any subset of X. Show that $\Sigma_D=\{E\cap D:E\in\Sigma\}$ is a σ-algebra of subsets of D. Set $\mu_D=\mu^*\restriction\Sigma_D$, the function with domain Σ_D such that $\mu_D B=\mu^*B$ for every $B\in\Sigma_D$, where μ^* is defined as in (a) above; show that (D,Σ_D,μ_D) is a measure space. (μ_D is the **subspace measure** on D.)

(c) Let (X,Σ,μ) be a measure space and let μ^* be the associated outer measure on X, as in 113Ya. Let $\check\mu$ be the measure on X constructed by Carathéodory's method from μ^*, and $\check\Sigma$ its domain. Show that $\Sigma\subseteq\check\Sigma$ and that $\check\mu$ extends μ.

(d) Let X be a set and $\tau:\mathcal{P}X\to[0,\infty]$ any function such that $\tau\emptyset=0$. For $A\subseteq X$ set

$$\theta A=\inf\{\sum_{j=0}^{\infty}\tau C_j:\langle C_j\rangle_{j\in\mathbb{N}}\text{ is a sequence of subsets of }X$$

$$\text{such that }A\subseteq\bigcup_{j\in\mathbb{N}}C_j\}.$$

Show that θ is an outer measure on X. (*Hint*: you will need 111F(b-ii) or something equivalent.)

(e) Let X be a set and θ_1, θ_2 two outer measures on X. Show that $\theta_1\wedge\theta_2$, as described in 113Xb(iii), is the outer measure derived by the process of 113Yd from the functional $\tau C=\min(\theta_1 C,\theta_2 C)$.

(f) Let X be a set and $\langle\theta_i\rangle_{i\in I}$ any non-empty family of outer measures on X. Set $\tau C=\inf_{i\in I}\theta_i C$ for each $C\subseteq X$. Show that the outer measure derived from τ by the process of 113Yd is the largest outer measure θ such that $\theta A\le\theta_i A$ whenever $A\subseteq X$ and $i\in I$.

(g) Let X be a set and $\phi:\mathcal{P}X\to[0,\infty]$ a functional such that

$\phi\emptyset=0$;

$\phi(A\cup B)\ge\phi A+\phi B$ for all disjoint A, $B\subseteq X$;

if $\langle A_n\rangle_{n\in\mathbb{N}}$ is a non-increasing sequence of subsets of X and $\phi A_0<\infty$ then $\phi(\bigcap_{n\in\mathbb{N}}A_n)=\lim_{n\to\infty}\phi A_n$;

if $\phi A=\infty$ and $a\in\mathbb{R}$ there is a $B\subseteq A$ such that $a\le\phi B<\infty$.

Set

$$\Sigma=\{E:E\subseteq X,\ \phi(A\cap E)+\phi(A\setminus E)=\phi A\text{ for every }A\subseteq X\}.$$

Show that $(X,\Sigma,\phi\restriction\Sigma)$ is a measure space.

(h) Let (X,Σ,μ) be a measure space and for $A\subseteq X$ set $\mu_*A=\sup\{\mu E:E\in\Sigma,\,E\subseteq A,\,\mu E<\infty\}$. Show that $\mu_*:\mathcal{P}X\to[0,\infty]$ satisfies the conditions of 113Yg, and that if $\mu X<\infty$ then the measure defined from μ_* by the method of 113Yg extends μ.

(i) Let X be a set and $\mathcal{A}$ an **algebra** of subsets of X, that is, a family of subsets of X such that

$\emptyset\in\mathcal{A}$,

$X\setminus E\in\mathcal{A}$ for every $E\in\mathcal{A}$,

$E\cup F\in\mathcal{A}$ whenever E, $F\in\mathcal{A}$.

Let $\phi:\mathcal{A}\to[0,\infty]$ be a function such that

$\phi\emptyset=0$,

$\phi(E\cup F)=\phi E+\phi F$ whenever E, $F\in\mathcal{A}$ and $E\cap F=\emptyset$,

$\phi E=\lim_{n\to\infty}\phi E_n$ whenever $\langle E_n\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence in $\mathcal{A}$ with union E.

Show that there is a measure μ on X extending ϕ. (*Hint*: set $\phi A=\infty$ for $A\in\mathcal{P}X\setminus\mathcal{A}$; define θ from ϕ as in 113Yd, and μ from θ.)

(j) (T.de Pauw) Let X be a set, T a σ-algebra of subsets of X, and θ an outer measure on X. Set $\Sigma=\{E:E\in\mathrm{T},\,\theta A=\theta(A\cap E)+\theta(A\setminus E)\text{ for every }A\in\mathrm{T}\}$. Show that Σ is a σ-algebra of subsets of X and that $\theta\restriction\Sigma$ is a measure.

(k) Let X, $\tau:\mathcal{P}X\to[0,\infty]$ and θ be as in 113Yd; let μ be the measure defined by Carathéodory's method from θ, and Σ the domain of μ. Suppose that $E\subseteq X$ is such that $\theta(C\cap E)+\theta(C\setminus E)\le\tau C$ whenever $C\subseteq X$ is such that $0<\tau C<\infty$. Show that $E\in\Sigma$.

113 Notes and comments We are proceeding by the easiest stages I can devise to the construction of a non-trivial measure space, that is, Lebesgue measure on $\mathbb{R}$. There are many constructions of Lebesgue measure, but in my view Carathéodory's method (113C) is the right one to begin with, because it is the most powerful and versatile single technique for constructing measures. It is, of course, abstract – it deals with arbitrary outer measures on arbitrary sets; but I really think that the Lebesgue theory, intertwined as it is with the rich structure of Euclidean space, is harder than the abstract theory of measure. We do at least have here a serious theorem for you to get your teeth into, mastery of which should be both satisfying and useful. I must say that I think it very remarkable that such a direct construction should be effective. Looking at the proof, it is perhaps worth while distinguishing between the 'algebraic' or 'finite' parts ((a)-(c)) and the parts involving sequences of sets ((d)-(e)); the former amount to a proof of 113Xh. Outer measures of various kinds appear throughout measure theory, and I sketch a few of the relevant constructions in 113X-113Y.

114 Lebesgue measure on $\mathbb{R}$

Following the very abstract ideas of §§111-113, we have an urgent need for a non-trivial example of a measure space. By far the most important example is the real line with Lebesgue measure, and I now proceed to a description of this measure (114A-114E), with a few of its basic properties.

The principal ideas of this section are repeated in §115, and if you have encountered Lebesgue measure before, or feel confident in your ability to deal with two- and three-dimensional spaces at the same time as doing some difficult analysis, you could go directly to that section, turning back to this one only when a specific reference is given.

114A Definitions (a) For the purposes of this section, a **half-open interval** in $\mathbb{R}$ is a set of the form $[a, b[= \{x : a \leq x < b\}$, where a, $b \in \mathbb{R}$.

Observe that I allow $b \leq a$ in this formula; in this case $[a, b[= \emptyset$ (see 1A1A).

(b) If $I \subseteq \mathbb{R}$ is a half-open interval, then either $I = \emptyset$ or $I = [\inf I, \sup I[$, so that its endpoints are well defined. We may therefore define the **length** λI of a half-open interval I by setting

$$\lambda\emptyset = 0, \quad \lambda\, [a, b[= b - a \text{ if } a < b.$$

114B Lemma If $I \subseteq \mathbb{R}$ is a half-open interval and $\langle I_j \rangle_{j\in\mathbb{N}}$ is a sequence of half-open intervals covering I, then $\lambda I \leq \sum_{j=0}^{\infty} \lambda I_j$.

proof (a) If $I = \emptyset$ then of course $\lambda I = 0 \leq \sum_{j=0}^{\infty} \lambda I_j$. Otherwise, take $I = [a, b[$, where $a < b$. For each $x \in \mathbb{R}$ let H_x be the half-line $]-\infty, x[$, and consider the set

$$A = \{x : a \leq x \leq b,\, x - a \leq \textstyle\sum_{j=0}^{\infty} \lambda(I_j \cap H_x)\}.$$

(Note that if $I_j = [c_j, d_j[$ then $I_j \cap H_x = [c_j, \min(d_j, x)[$, so $\lambda(I_j \cap H_x)$ is always defined.) We have $a \in A$ (because $a - a = 0 \leq \sum_{j=0}^{\infty} \lambda(I_j \cap H_a)$) and of course $A \subseteq [a, b]$, so $c = \sup A$ is defined, and belongs to $[a, b]$.

(b) We find now that $c \in A$.

$$\mathbf{P}\; c - a = \sup_{x\in A} x - a \leq \sup_{x\in A} \sum_{j=0}^{\infty} \lambda(I_j \cap H_x) \leq \sum_{j=0}^{\infty} \lambda(I_j \cap H_c).\ \mathbf{Q}$$

(c) ? Suppose, if possible, that $c < b$. Then $c \in [a, b[$, so there is some $k \in \mathbb{N}$ such that $c \in I_k$. Express I_k as $[c_k, d_k[$; then $x = \min(d_k, b) > c$. For each j, $\lambda(I_j \cap H_x) \geq \lambda(I_j \cap H_c)$, while

$$\lambda(I_k \cap H_x) = \lambda(I_k \cap H_c) + x - c.$$

So

$$\sum_{j=0}^{\infty} \lambda(I_j \cap H_x) \geq \sum_{j=0}^{\infty} \lambda(I_j \cap H_c) + x - c \geq c - a + x - c = x - a,$$

so $x \in A$; but $x > c$ and $c = \sup A$. **X**

(d) We conclude that $c = b$, so that $b \in A$ and

$$b - a \le \sum_{j=0}^{\infty} \lambda(I_j \cap H_b) \le \sum_{j=0}^{\infty} \lambda I_j,$$

as claimed.

114C Definition Now, and for the rest of this section, define $\theta : \mathcal{P}\mathbb{R} \to [0, \infty]$ by writing

$$\theta A = \inf\{\sum_{j=0}^{\infty} \lambda I_j : \langle I_j \rangle_{j\in\mathbb{N}} \text{ is a sequence of half-open intervals such that } A \subseteq \bigcup_{j\in\mathbb{N}} I_j\}.$$

Observe that every A can be covered by some sequence of half-open intervals – e.g., $A \subseteq \bigcup_{n\in\mathbb{N}} [-n, n[$; so that if we interpret the sums in $[0, \infty]$, as in 112Bc above, we always have a non-empty set to take the infimum of, and θA is always defined in $[0, \infty]$. This function θ is called **Lebesgue outer measure** on $\mathbb{R}$; the phrase is justified by (a) of the next proposition.

114D Proposition (a) θ is an outer measure on $\mathbb{R}$.

(b) $\theta I = \lambda I$ for every half-open interval $I \subseteq \mathbb{R}$.

proof (a)(i) θ takes values in $[0, \infty]$ because every θA is the infimum of a non-empty subset of $[0, \infty]$.

(ii) $\theta\emptyset = 0$ because (for instance) if we set $I_j = \emptyset$ for every j, then every I_j is a half-open interval (on the convention I am using) and $\emptyset \subseteq \bigcup_{j\in\mathbb{N}} I_j$, $\sum_{j=0}^{\infty} \lambda I_j = 0$.

(iii) If $A \subseteq B$ then whenever $B \subseteq \bigcup_{j\in\mathbb{N}} I_j$ we have $A \subseteq \bigcup_{j\in\mathbb{N}} I_j$, so θA is the infimum of a set at least as large as that involved in the definition of θB, and $\theta A \le \theta B$.

(iv) Now suppose that $\langle A_n \rangle_{n\in\mathbb{N}}$ is a sequence of subsets of $\mathbb{R}$, with union A. For any $\epsilon > 0$, we can choose, for each $n \in \mathbb{N}$, a sequence $\langle I_{nj} \rangle_{j\in\mathbb{N}}$ of half-open intervals such that $A_n \subseteq \bigcup_{j\in\mathbb{N}} I_{nj}$ and $\sum_{j=0}^{\infty} \lambda I_{nj} \le \theta A_n + 2^{-n}\epsilon$. (You should perhaps check that this formulation is valid whether θA_n is finite or infinite.) Now by 111F(b-ii) there is a bijection from $\mathbb{N}$ to $\mathbb{N} \times \mathbb{N}$; express this in the form $m \mapsto (k_m, l_m)$. Then $\langle I_{k_m,l_m} \rangle_{m\in\mathbb{N}}$ is a sequence of half-open intervals, and

$$A \subseteq \bigcup_{m\in\mathbb{N}} I_{k_m,l_m}.$$

P If $x \in A = \bigcup_{n\in\mathbb{N}} A_n$ there must be an $n \in \mathbb{N}$ such that $x \in A_n \subseteq \bigcup_{j\in\mathbb{N}} I_{nj}$, so there is a $j \in \mathbb{N}$ such that $x \in I_{nj}$. Now $m \mapsto (k_m, l_m)$ is surjective, so there is an $m \in \mathbb{N}$ such that $k_m = n$ and $l_m = j$, in which case $x \in I_{k_m,l_m}$. **Q**

Next,

$$\sum_{m=0}^{\infty} \lambda I_{k_m,l_m} \le \sum_{n=0}^{\infty} \sum_{j=0}^{\infty} \lambda I_{nj}.$$

P If $M \in \mathbb{N}$, then $N = \max(k_0, k_1, \dots, k_M, l_0, l_1, \dots, l_M)$ is finite; because every λI_{nj} is greater than or equal to 0, and any pair (n, j) can appear at most once as a (k_m, l_m),

$$\sum_{m=0}^{M} \lambda I_{k_m,l_m} \le \sum_{n=0}^{N} \sum_{j=0}^{N} \lambda I_{nj} \le \sum_{n=0}^{N} \sum_{j=0}^{\infty} \lambda I_{nj} \le \sum_{n=0}^{\infty} \sum_{j=0}^{\infty} \lambda I_{nj}.$$

So

$$\sum_{m=0}^{\infty} \lambda I_{k_m,l_m} = \lim_{M\to\infty} \sum_{m=0}^{M} \lambda I_{k_m,l_m} \le \sum_{n=0}^{\infty} \sum_{j=0}^{\infty} \lambda I_{nj}. \textbf{ Q}$$

Accordingly

$$\begin{aligned}\theta A &\le \sum_{m=0}^{\infty} \lambda I_{k_m,l_m} \\ &\le \sum_{n=0}^{\infty}\sum_{j=0}^{\infty} \lambda I_{nj} \\ &\le \sum_{n=0}^{\infty}(\theta A_n + 2^{-n}\epsilon) \\ &= \sum_{n=0}^{\infty}\theta A_n + \sum_{n=0}^{\infty} 2^{-n}\epsilon \\ &= \sum_{n=0}^{\infty}\theta A_n + 2\epsilon.\end{aligned}$$

Because ϵ is arbitrary, $\theta A \le \sum_{n=0}^{\infty}\theta A_n$ (again, you should check that this is valid whether or not $\sum_{n=0}^{\infty}\theta A_n$ is finite). As $\langle A_n\rangle_{n\in\mathbb{N}}$ is arbitrary, θ is an outer measure.

(b) Because we can always take $I_0 = I$, $I_j = \emptyset$ for $j \ge 1$, to obtain a sequence of half-open intervals covering I with $\sum_{j=0}^{\infty}\lambda I_j = \lambda I$, we surely have $\theta I \le \lambda I$. For the reverse inequality, use 114B: if $I \subseteq \bigcup_{j\in\mathbb{N}} I_j$, then $\lambda I \le \sum_{j=0}^{\infty}\lambda I_j$; as $\langle I_j\rangle_{j\in\mathbb{N}}$ is arbitrary, $\theta I \ge \lambda I$ and $\theta I = \lambda I$, as required.

Remark There is an ungainly shift in the argument of (a-iv) above, in the stage

$$‘\theta A \le \sum_{m=0}^{\infty}\lambda I_{k_m,l_m} \le \sum_{n=0}^{\infty}\sum_{j=0}^{\infty}\lambda I_{nj}’.$$

I dare say you would have believed me if I had suppressed the k_m, l_m altogether and simply written ‘because $A \subseteq \bigcup_{n,j\in\mathbb{N}} I_{nj}$, $\theta A \le \sum_{n=0}^{\infty}\sum_{j=0}^{\infty}\lambda I_{nj}$’. I hope that you will not find it too demoralizing if I suggest that such a jump is not quite safe. My reasons for interpolating a name for a bijection between $\mathbb{N}$ and $\mathbb{N}\times\mathbb{N}$, and taking a couple of lines to say explicitly that $\sum_{m=0}^{\infty}\lambda I_{k_m,l_m} \le \sum_{n=0}^{\infty}\sum_{j=0}^{\infty}\lambda I_{nj}$, are the following. To start with, there is the formal point that the definition 114C demands a simple sequence, not a double sequence. Is it really obvious that it doesn't matter here? If so, why? There can be no way to justify the shift which does not rely on the facts that $\mathbb{N}\times\mathbb{N}$ is countable and every λI_{nj} is non-negative. If either of those were untrue, the method would be in grave danger of failing.

At some point we shall certainly need to discuss sums over infinite index sets other than $\mathbb{N}$, including uncountable index sets. I have already touched on these in 112Bd, and I will return to them in 226A in Volume 2. For the moment, I feel that we have quite enough new ideas to cope with, and that what we need here is a reasonably honest expedient to deal with the question immediately before us.

You may have noticed, or guessed, that some of the inequalities ‘$\le$’ here must actually be equalities; if so, check your guess in 114Ya.

114E Definition Because Lebesgue outer measure (114C) is indeed an outer measure (114Da), we may use it to construct a measure μ, using Carathéodory's method (113C). This measure is **Lebesgue measure on** $\mathbb{R}$. The sets E measured by μ (that is, for which $\theta(A\cap E)+\theta(A\setminus E) = \theta A$ for every $A\subseteq\mathbb{R}$) are called **Lebesgue measurable**.

Sets which are negligible for μ are called **Lebesgue negligible**; note that these are just the sets A for which $\theta A = 0$, and are all Lebesgue measurable (113Xa).

114F Lemma Let $x\in\mathbb{R}$. Then $H_x = \,]-\infty, x[$ is Lebesgue measurable for every $x\in\mathbb{R}$.

proof (a) The point is that $\lambda I = \lambda(I\cap H_x) + \lambda(I\setminus H_x)$ for every half-open interval $I\subseteq\mathbb{R}$. **P** If either $I\subseteq H_x$ or $I\cap H_x = \emptyset$, this is trivial. Otherwise, I must be of the form $[a,b[$, where $a < x < b$. Now $I\cap H_x = [a,x[$ and $I\setminus H_x = [x,b[$ are both half-open intervals, and

$$\lambda(I\cap H_x) + \lambda(I\setminus H_x) = (x-a)+(b-x) = b-a = \lambda I. \quad \mathbf{Q}$$

(b) Now suppose that A is any subset of $\mathbb{R}$, and $\epsilon > 0$. Then we can find a sequence $\langle I_j\rangle_{j\in\mathbb{N}}$ of half-open intervals such that $A\subseteq\bigcup_{j\in\mathbb{N}} I_j$ and $\sum_{j=0}^{\infty}\lambda I_j \le \theta A+\epsilon$. Now $\langle I_j\cap H_x\rangle_{j\in\mathbb{N}}$ and $\langle I_j\setminus H_x\rangle_{j\in\mathbb{N}}$ are sequences of half-open intervals and $A\cap H_x\subseteq\bigcup_{j\in\mathbb{N}}(I_j\cap H_x)$, $A\setminus H_x\subseteq\bigcup_{j\in\mathbb{N}}(I_j\setminus H_x)$. So

$$\theta(A \cap H_x) + \theta(A \setminus H_x) \le \sum_{j=0}^{\infty} \lambda(I_j \cap H_x) + \sum_{j=0}^{\infty} \lambda(I_j \setminus H_x)$$
$$= \sum_{j=0}^{\infty} \lambda I_j \le \theta A + \epsilon.$$

Because ϵ is arbitrary, $\theta(A \cap H_x) + \theta(A \setminus H_x) \le \theta A$; because A is arbitrary, H_x is measurable, as remarked in 113D.

114G Proposition All Borel subsets of $\mathbb{R}$ are Lebesgue measurable; in particular, all open sets, and all sets of the following classes, together with countable unions of them:

(i) open intervals $]a, b[$, $]-\infty, b[$, $]a, \infty[$, $]-\infty, \infty[$, where $a < b \in \mathbb{R}$;

(ii) closed intervals $[a, b]$, where $a \le b \in \mathbb{R}$;

(iii) half-open intervals $[a, b[$, $]a, b]$, $]-\infty, b]$, $[a, \infty[$, where $a < b$ in $\mathbb{R}$.

We have moreover the following formula for the measures of such sets, writing μ for Lebesgue measure:

$$\mu\,]a, b[= \mu[a, b] = \mu\,[a, b[= \mu\,]a, b] = b - a$$

whenever $a \le b$ in $\mathbb{R}$, while all the unbounded intervals have infinite measure. It follows that every countable subset of $\mathbb{R}$ is measurable and of zero measure.

proof (a) I show first that all open subsets of $\mathbb{R}$ are measurable. **P** Let $G \subseteq \mathbb{R}$ be open. Let $K \subseteq \mathbb{Q} \times \mathbb{Q}$ be the set of pairs (q, q') of rational numbers such that $[q, q'[\subseteq G$. Now by the remarks in 111E-111F – specifically, 111Eb, showing that $\mathbb{Q}$ is countable, 111F(b-iii), showing that products of countable sets are countable, and 111F(b-i), showing that subsets of countable sets are countable – we see that K is countable. Also, every $[q, q'[$ is measurable, being $H_{q'} \setminus H_q$ in the language of 114F. So, by 111Fa, $G' = \bigcup_{(q,q') \in K} [q, q'[$ is measurable.

By the definition of K, $G' \subseteq G$. On the other hand, if $x \in G$, there is an $\epsilon > 0$ such that $]x - \epsilon, x + \epsilon[\subseteq G$. Now there are rational numbers $q \in]x - \epsilon, x]$ and $q' \in]x, x + \epsilon]$, so that $(q, q') \in K$ and $x \in [q, q'[\subseteq G'$. As x is arbitrary, $G = G'$ and G is measurable. **Q**

(b) Now the family Σ of Lebesgue measurable sets is a σ-algebra of subsets of $\mathbb{R}$ including the family of open sets, so must contain every Borel set, by the definition of Borel set (111G).

(c) Of the types of interval considered, all the open intervals are actually open sets, so are surely Borel. The complement of a closed interval is expressible as the union of at most two open intervals, so is Borel, and the closed interval, being the complement of a Borel set, is Borel. A bounded half-open interval is expressible as the intersection of an open interval with a closed interval, so is Borel; and finally an unbounded interval of the form $]-\infty, b]$ or $[a, \infty[$ is the complement of an open interval, so is also Borel.

(d) To compute the measures, we already know from 114Db that

$$\mu\,[a, b[= \theta\,[a, b[= b - a$$

if $a \le b$. For the other types of bounded interval, it is enough to note that $\mu\{a\} = 0$ for every $a \in \mathbb{R}$, as the different intervals differ only by one or two points; and this is so because $\{a\} \subseteq [a, a + \epsilon[$, so $\mu\{a\} \le \epsilon$, for every $\epsilon > 0$.

As for the unbounded intervals, they include arbitrarily long half-open intervals, so must have infinite measure.

(e) As just remarked, $\mu\{a\} = 0$ for every a. If $A \subseteq \mathbb{R}$ is countable, it is either empty or expressible as $\{a_n : n \in \mathbb{N}\}$. In the former case $\mu A = \mu\emptyset = 0$; in the latter, $A = \bigcup_{n \in \mathbb{N}} \{a_n\}$ is Borel and $\mu A \le \sum_{n=0}^{\infty} \mu\{a_n\} = 0$.

114X Basic exercises >(a) Let $g : \mathbb{R} \to \mathbb{R}$ be any non-decreasing function. For half-open intervals $I \subseteq \mathbb{R}$ define $\lambda_g I$ by setting

$$\lambda_g \emptyset = 0, \quad \lambda_g\,[a, b[= \lim_{x \uparrow b} g(x) - \lim_{x \uparrow a} g(x)$$

if $a < b$. For any set $A \subseteq \mathbb{R}$ set

$$\theta_g A = \inf\{\sum_{j=0}^{\infty} \lambda_g I_j : \langle I_j \rangle_{j \in \mathbb{N}} \text{ is a sequence of half-open intervals}$$
$$\text{such that } A \subseteq \bigcup_{j \in \mathbb{N}} I_j\}.$$

Show that θ_g is an outer measure on $\mathbb{R}$. Let μ_g be the measure defined from θ_g by Carathéodory's method; show that $\mu_g I$ is defined and equal to $\lambda_g I$ for every half-open interval $I \subseteq \mathbb{R}$, and that every Borel subset of $\mathbb{R}$ is in the domain of μ_g.

(μ_g is the **Lebesgue-Stieltjes** measure associated with g.)

(b) At which point would the argument of 114Xa break down if we wrote $\lambda_g\,[a,b[\,= g(b) - g(a)$ instead of using the formula given?

>(c) Write θ for Lebesgue outer measure and μ for Lebesgue measure on $\mathbb{R}$. Show that $\theta A = \inf\{\mu E : E$ is Lebesgue measurable, $A \subseteq E\}$ for every $A \subseteq \mathbb{R}$. (*Hint*: Consider sets E of the form $\bigcup_{j\in\mathbb{N}} I_j$, where $\langle I_j\rangle_{j\in\mathbb{N}}$ is a sequence of half-open intervals.)

(d) Let X be a set, $\mathcal{I}$ a family of subsets of X such that $\emptyset \in \mathcal{I}$, and $\lambda : \mathcal{I} \to [0,\infty[$ a function such that $\lambda\emptyset = 0$. Define $\theta : \mathcal{P}X \to [0,\infty]$ by writing

$$\theta A = \inf\{\textstyle\sum_{j=0}^{\infty} \lambda I_j : \langle I_j\rangle_{j\in\mathbb{N}} \text{ is a sequence in } \mathcal{I} \text{ such that } A \subseteq \bigcup_{j\in\mathbb{N}} I_j\},$$

interpreting $\inf\emptyset$ as ∞, so that $\theta A = \infty$ if A is not covered by any sequence in $\mathcal{I}$. Show that θ is an outer measure on X.

(e) Let $E \subseteq \mathbb{R}$ be a set of finite measure for Lebesgue measure μ. Show that for every $\epsilon > 0$ there is a disjoint family $I_0, \dots, I_n$ of half-open intervals such that $\mu(E \triangle \bigcup_{j\le n} I_j) \le \epsilon$. (*Hint*: let $\langle J_j\rangle_{j\in\mathbb{N}}$ be a sequence of half-open intervals such that $E \subseteq \bigcup_{j\in\mathbb{N}} J_j$ and $\sum_{j=0}^{\infty} \mu J_j \le \mu E + \frac{1}{2}\epsilon$. Now take a suitably large m and express $\bigcup_{j\le m} J_j$ as a disjoint union of half-open intervals.)

>(f) Write θ for Lebesgue outer measure and μ for Lebesgue measure on $\mathbb{R}$. Suppose that $c \in \mathbb{R}$. Show that $\theta(A+c) = \theta A$ for every $A \subseteq \mathbb{R}$, where $A + c = \{x + c : x \in A\}$. Show that if $E \subseteq \mathbb{R}$ is measurable so is $E + c$, and that in this case $\mu(E + c) = \mu E$.

(g) Write θ for Lebesgue outer measure and μ for Lebesgue measure on $\mathbb{R}$. Suppose that $c > 0$. Show that $\theta(cA) = c\theta(A)$ for every $A \subseteq \mathbb{R}$, where $cA = \{cx : x \in A\}$. Show that if $E \subseteq \mathbb{R}$ is measurable so is cE, and that in this case $\mu(cE) = c\mu E$.

114Y Further exercises (a) In (a-iv) of the proof of 114D, show that $\sum_{m=0}^{\infty} \lambda I_{k_m,l_m}$ is actually equal to $\sum_{n=0}^{\infty}\sum_{j=0}^{\infty} \lambda I_{nj}$.

(b) Let g, $h : \mathbb{R} \to \mathbb{R}$ be two non-decreasing functions, with sum $g + h$; let μ_g, μ_h, μ_{g+h} be the corresponding Lebesgue-Stieltjes measures (114Xa). Show that

$$\operatorname{dom}\mu_{g+h} = \operatorname{dom}\mu_g \cap \operatorname{dom}\mu_h, \quad \mu_{g+h}E = \mu_g E + \mu_h E \text{ for every } E \in \operatorname{dom}\mu_{g+h}.$$

(c) Let $\langle g_n\rangle_{n\in\mathbb{N}}$ be a sequence of non-decreasing functions from $\mathbb{R}$ to $\mathbb{R}$, and suppose that $g(x) = \sum_{n=0}^{\infty} g_n(x)$ is defined and finite for every $x \in \mathbb{R}$. Let μ_{g_n}, μ_g be the corresponding Lebesgue-Stieltjes measures. Show that

$$\operatorname{dom}\mu_g = \textstyle\bigcap_{n\in\mathbb{N}} \operatorname{dom}\mu_{g_n}, \quad \mu_g E = \sum_{n=0}^{\infty} \mu_{g_n} E \text{ for every } E \in \operatorname{dom}\mu_g.$$

(d)(i) Show that if $A \subseteq \mathbb{R}$ and $\epsilon > 0$, there is an open set $G \supseteq A$ such that $\theta G \le \theta A + \epsilon$, where θ is Lebesgue outer measure. (ii) Show that if $E \subseteq \mathbb{R}$ is Lebesgue measurable and $\epsilon > 0$, there is an open set $G \supseteq E$ such that $\mu(G \setminus E) \le \epsilon$, where μ is Lebesgue measure. (*Hint*: consider first the case of bounded E.) (iii) Show that if $E \subseteq \mathbb{R}$ is Lebesgue measurable, there are Borel sets H_1, H_2 such that $H_1 \subseteq E \subseteq H_2$ and $\mu(H_2 \setminus E) = \mu(E \setminus H_1) = 0$. (*Hint*: use (ii) to find H_2, and then consider the complement of E.)

(e) Write θ for Lebesgue outer measure on $\mathbb{R}$. Show that a set $E \subseteq \mathbb{R}$ is Lebesgue measurable iff $\theta([-n,n] \cap E) + \theta([-n,n] \setminus E) = 2n$ for every $n \in \mathbb{N}$. (*Hint*: Use 114Yd to show that for each n there are measurable sets F_n, H_n such that $F_n \subseteq [-n,n] \cap E \subseteq H_n$ and $H_n \setminus F_n$ is negligible.)

(f) Repeat 114Xc and 114Yd-114Ye for the Lebesgue-Stieltjes measures of 114Xa.

(g) Write $\mathcal{B}$ for the σ-algebra of Borel subsets of $\mathbb{R}$, and let $\nu : \mathcal{B} \to [0,\infty]$ be a measure. Let g, λ_g, θ_g and μ_g be as in 114Xa. Show that if $\nu I = \lambda_g I$ for every half-open interval I, then $\nu E = \mu_g E$ for every $E \in \mathcal{B}$. (*Hint*: first consider open sets E, and then use 114Yd(i) as extended in 114Yf.)

(h) Write $\mathcal{B}$ for the σ-algebra of Borel subsets of $\mathbb{R}$, and let $\nu : \mathcal{B} \to [0, \infty]$ be a measure such that $\nu[-n, n] < \infty$ for every $n \in \mathbb{N}$. Show that there is a function $g : \mathbb{R} \to \mathbb{R}$ which is non-decreasing, continuous on the left and such that $\nu E = \mu_g E$ for every $E \in \mathcal{B}$, where μ_g is defined as in 114Xa. Is g unique?

(i) Write $\mathcal{B}$ for the σ-algebra of Borel subsets of $\mathbb{R}$, and let ν_1, ν_2 be measures with domain $\mathcal{B}$ such that $\nu_1 I = \nu_2 I < \infty$ for every half-open interval $I \subseteq \mathbb{R}$. Show that $\nu_1 E = \nu_2 E$ for every $E \in \mathcal{B}$.

(j) Let $\mathcal{E}$ be any family of half-open intervals in $\mathbb{R}$. Show that (i) there is a countable $\mathcal{C} \subseteq \mathcal{E}$ such that $\bigcup \mathcal{E} = \bigcup \mathcal{C}$ (definition: 1A1F) (ii) that $\bigcup \mathcal{E}$ is a Borel set, so is Lebesgue measurable (iii) that there is a disjoint sequence $\langle I_n \rangle_{n \in \mathbb{N}}$ of half-open intervals in $\mathbb{R}$ such that $\bigcup \mathcal{E} = \bigcup_{n \in \mathbb{N}} I_n$.

(k) Show that for almost every $x \in \mathbb{R}$ (as measured by Lebesgue measure) the set

$$\{(m, n) : m \in \mathbb{Z},\ n \in \mathbb{N} \setminus \{0\},\ |x - \frac{m}{n}| \le \frac{1}{n^3}\}$$

is finite. (*Hint*: estimate the outer measure of $\bigcup_{n \ge n_0} \bigcup_{|m| \le kn} [\frac{m}{n} - \frac{1}{n^3}, \frac{m}{n} + \frac{1}{n^3}]$ for n_0, $k \ge 1$.) Repeat with $2 + \epsilon$ in the place of 3.

(l) Write μ for Lebesgue measure on $\mathbb{R}$. Show that there is a countable family $\mathcal{F}$ of Lebesgue measurable subsets of $\mathbb{R}$ such that whenever μE is defined and finite, and $\epsilon > 0$, there is an $F \in \mathcal{F}$ such that $\mu(E \triangle F) \le \epsilon$. (*Hint*: in 114Xe, show that we can take the I_j to have rational endpoints.)

114 Notes and comments My own interests are in 'abstract' measure theory, and from the point of view of the structure of this treatise, the chief object of this section is the description of a non-trivial measure space to provide a focus for the general theorems which follow. Let me enumerate the methods of constructing measure spaces already available to us. (a) We have the point-supported measures of 112Bd; in some ways, these are trivial; but they do occur in applications, and, just because they are generally easy to deal with, it is often right to test any new ideas on them. (b) We have Lebesgue measure on $\mathbb{R}$; a straightforward generalization of the construction yields the Lebesgue-Stieltjes measures (114Xa). (c) Next, we have ways of building new measures from old, starting with subspace measures (113Yb), image measures (112Xf) and sums of measures (112Yf). Perhaps the most important of these is 'Lebesgue measure on $[0, 1]$', I call it μ_1 for the moment, where the domain of μ_1 is $\{E : E \subseteq [0, 1]$ is Lebesgue measurable$\} = \{E \cap [0, 1] : E \subseteq \mathbb{R}$ is Lebesgue measurable$\}$, and $\mu_1 E$ is just the Lebesgue measure of E for each $E \in \operatorname{dom} \mu_1$. In fact the image measures of Lebesgue measure on $[0, 1]$ include a very large proportion of the probability measures (that is, measures giving measure 1 to the whole space) of importance in ordinary applications.

Of course Lebesgue measure is not only the dominant guiding example for general measure theory, but is itself the individual measure of greatest importance for applications. For this reason it would be possible – though in my view narrow-minded – to read chapters 12-13 of this volume, and a substantial proportion of Volume 2, as if they applied only to Lebesgue measure on $\mathbb{R}$. This is, indeed, the context in which most of these results were first developed. I believe, however, that it is often the case in mathematics, that one's understanding of a particular construction is deepened and strengthened by an acquaintance with related objects, and that one of the ways to an appreciation of the nature of Lebesgue measure is through a study of its properties in the more abstract context of general measure theory.

For any proper investigation of the applications of Lebesgue measure theory we must wait for Volume 2. But I include 114Yk as a hint of one of the ways in which this theory can be used.

115 Lebesgue measure on $\mathbb{R}^r$

Following the very abstract ideas of §§111-113, there is an urgent need for non-trivial examples of measure spaces. By far the most important examples are the Euclidean spaces $\mathbb{R}^r$ with Lebesgue measure, and I now proceed to a definition of these measures (115A-115E), with a few of their basic properties. Except at one point (in the proof of the fundamental lemma 115B) this section does not rely essentially on §114; but nevertheless most students encountering Lebesgue measure for the first time will find it easier to work through the one-dimensional case carefully before embarking on the multi-dimensional case.

115A Definitions (a) For practically the whole of this section (the exception is the proof of Lemma 115B) r will denote a fixed integer greater than or equal to 1. I will use Roman letters a, b, c, d, x, y to denote members of $\mathbb{R}^r$, and Greek letters for their coordinates, so that $a = (\alpha_1, \dots, \alpha_r)$, $b = (\beta_1, \dots, \beta_r)$, $x = (\xi_1, \dots, \xi_r)$.

(b) For the purposes of this section, a **half-open interval** in $\mathbb{R}^r$ is a set of the form $[a,b[\,=\{x:\alpha_i\le\xi_i<\beta_i\ \forall\, i\le r\}$, where a, $b\in\mathbb{R}^r$. Observe that I allow $\beta_i\le\alpha_i$ in this formula; if this happens for any i, then $[a,b[\,=\emptyset$.

(c) If $I=[a,b[\,\subseteq\mathbb{R}^r$ is a half-open interval, then either $I=\emptyset$ or

$$\alpha_i=\inf\{\xi_i:x\in I\},\quad \beta_i=\sup\{\xi_i:x\in I\}$$

for every $i\le r$; in the latter case, the expression of I as a half-open interval is unique. We may therefore define the r**-dimensional volume** λI of a half-open interval I by setting

$$\lambda\emptyset=0,\quad \lambda\,[a,b[\,=\prod_{i=1}^r\beta_i-\alpha_i \text{ if } \alpha_i<\beta_i \text{ for every } i.$$

115B Lemma If $I\subseteq\mathbb{R}^r$ is a half-open interval and $\langle I_j\rangle_{j\in\mathbb{N}}$ is a sequence of half-open intervals covering I, then $\lambda I\le\sum_{j=0}^\infty\lambda I_j$.

proof The proof is by induction on r. For this proof only, therefore, I write λ_r for the function defined on the half-open intervals of $\mathbb{R}^r$ by the formula of 115Ac.

(a) The argument for $r=1$, starting the induction, is similar to the inductive step; but rather than establish a suitable convention to set up a trivial case $r=0$, or ask you to work out the details yourself, I refer you to 114B, which is exactly the case $r=1$.

(b) For the inductive step to $r+1$, where $r\ge1$, take a half-open interval $I\subseteq\mathbb{R}^{r+1}$ and $\langle I_j\rangle_{j\in\mathbb{N}}$ a sequence of half-open intervals covering I. If $I=\emptyset$ then of course $\lambda_{r+1}I=0\le\sum_{j=0}^\infty\lambda_{r+1}I_j$. Otherwise, express I as $[a,b[$, where $\alpha_i<\beta_i$ for $i\le r+1$, and each I_j as $\left[a^{(j)},b^{(j)}\right[$. Write $\zeta=\prod_{i=1}^r\beta_i-\alpha_i$, so that $\lambda_{r+1}I=\zeta(\beta_{r+1}-\alpha_{r+1})$. Fix $\epsilon>0$. For each $\xi\in\mathbb{R}$ let H_ξ be the half-space $\{x:\xi_{r+1}<\xi\}$, and consider the set

$$A=\{\xi:\alpha_{r+1}\le\xi\le\beta_{r+1},\ \zeta(\xi-\alpha_{r+1})\le(1+\epsilon)\textstyle\sum_{j=0}^\infty\lambda_{r+1}(I_j\cap H_\xi)\}.$$

(Note that $I_j\cap H_\xi=\left[a^{(j)},\tilde b^{(j)}\right[$, where $\tilde\beta_i^{(j)}=\beta_i^{(j)}$ for $i\le r$ and $\tilde\beta_{r+1}^{(j)}=\min(\beta_{r+1}^{(j)},\xi)$, so $\lambda_{r+1}(I_j\cap H_\xi)$ is always defined.) We have $\alpha_{r+1}\in A$, because

$$\zeta(\alpha_{r+1}-\alpha_{r+1})=0\le(1+\epsilon)\textstyle\sum_{j=0}^\infty\lambda_{r+1}(I_j\cap H_{\alpha_{r+1}}),$$

and of course $A\subseteq[\alpha_{r+1},\beta_{r+1}]$, so $\gamma=\sup A$ is defined, and belongs to $[\alpha_{r+1},\beta_{r+1}]$.

(c) We find now that $\gamma\in A$.

$$\begin{aligned}\mathbf{P}\ \zeta(\gamma-\alpha_{r+1})&=\sup_{\xi\in A}\zeta(\xi-\alpha_{r+1})\\&\le(1+\epsilon)\sup_{\xi\in A}\sum_{j=0}^\infty\lambda_{r+1}(I_j\cap H_\xi)\le(1+\epsilon)\sum_{j=0}^\infty\lambda_{r+1}(I_j\cap H_\gamma).\ \mathbf{Q}\end{aligned}$$

(d) ? Suppose, if possible, that $\gamma<\beta_{r+1}$. Then $\gamma\in[\alpha_{r+1},\beta_{r+1}[$. Set

$$J=\{x:x\in\mathbb{R}^r,\ (x,\gamma)\in I\}=[a',b'[,$$

where $a'=(\alpha_1,\dots,\alpha_r)$, $b'=(\beta_1,\dots,\beta_r)$, and for each $j\in\mathbb{N}$ set

$$J_j=\{x:x\in\mathbb{R}^r,\ (x,\gamma)\in I_j\}.$$

Because $I\subseteq\bigcup_{j\in\mathbb{N}}I_j$, we must have $J\subseteq\bigcup_{j\in\mathbb{N}}J_j$. Of course both J and the J_j are half-open intervals in $\mathbb{R}^r$. (This is one of the places where it is helpful to count the empty set as a half-open interval.) By the inductive hypothesis, $\zeta=\lambda_rJ\le\sum_{j=0}^\infty\lambda_rJ_j$. As $\zeta>0$, there is an $m\in\mathbb{N}$ such that $\zeta\le(1+\epsilon)\sum_{j=0}^m\lambda_rJ_j$. Now for each $j\le m$, either $J_j=\emptyset$ or $\alpha_{r+1}^{(j)}\le\gamma<\beta_{r+1}^{(j)}$; set

$$\xi=\min(\{\beta_{r+1}\}\cup\{\beta_{r+1}^{(j)}:j\le m,\ J_j\ne\emptyset\})>\gamma.$$

Then

$$\lambda_{r+1}(I_j\cap H_\xi)\ge\lambda_{r+1}(I_j\cap H_\gamma)+(\xi-\gamma)\lambda_rJ_j$$

for every $j\le m$ such that J_j is non-empty, and therefore for every j. Consequently

$$\begin{aligned}\zeta(\xi-\alpha_{r+1}) &= \zeta(\gamma-\alpha_{r+1})+\zeta(\xi-\gamma)\\ &\le (1+\epsilon)\sum_{j=0}^{\infty}\lambda_{r+1}(I_j\cap H_\gamma)+(1+\epsilon)(\xi-\gamma)\sum_{j=0}^{m}\lambda_r J_j\\ &\le (1+\epsilon)\sum_{j=m+1}^{\infty}\lambda_{r+1}(I_j\cap H_\gamma)+(1+\epsilon)\sum_{j=0}^{m}\lambda_{r+1}(I_j\cap H_\xi)\\ &\le (1+\epsilon)\sum_{j=0}^{\infty}\lambda_{r+1}(I_j\cap H_\xi),\end{aligned}$$

and $\xi \in A$, which is impossible. **X**

(e) We conclude that $\gamma = \beta_{r+1}$, so that $\beta_{r+1} \in A$ and

$$\lambda_{r+1}I = \zeta(\beta_{r+1}-\alpha_{r+1}) \le (1+\epsilon)\textstyle\sum_{j=0}^{n}\lambda_{r+1}(I_j\cap H_{\beta_{r+1}}) \le (1+\epsilon)\sum_{j=0}^{\infty}\lambda_{r+1}I_j.$$

As ϵ is arbitrary,

$$\lambda_{r+1}I \le \textstyle\sum_{j=0}^{\infty}\lambda_{r+1}I_j,$$

as claimed.

Remark This proof is hard work, and not everybody makes such a mouthful of it. What is perhaps a more conventional approach is sketched in 115Ya, using the Heine-Borel theorem to reduce the problem to one of finite covers, and then (very often) saying that it is trivial. I do not use this method, partly because we do not need the Heine-Borel theorem elsewhere in this volume (though we shall certainly need it in Volume 2, and I write out a proof in 2A2F), and partly because I do not agree that the lemma is trivial when we have a finite sequence $I_0,\dots,I_m$ covering I. I invite you to consider this for yourself. It seems to me that any rigorous argument must involve an induction on the dimension, which is what I provide here. Of course dealing throughout with an infinite sequence makes it a little harder to keep track of what we are doing, and I note that in fact there is a crucial step which necessitates truncation of the sequence; I mean the formula

$$\xi = \min(\{\beta_{r+1}\}\cup\{\beta_{r+1}^{(j)} : j\le m,\ J_j\ne\emptyset\})$$

in part (d) of the proof. We certainly cannot take $\xi = \inf\{\beta_{r+1}^{(j)} : j\in\mathbb{N},\ J_j\ne\emptyset\}$, since this is very likely to be equal to γ. Accordingly I need some excuse for truncating, which is in the sentence

As $\zeta > 0$, there is an $m\in\mathbb{N}$ such that $\zeta \le (1+\epsilon)\sum_{j=0}^{m}\lambda_r J_j$.

And that step is the reason for introducing the slack ϵ into the definition of the set A at the beginning of the proof. Apart from this modification, the structure of the argument is supposed to reflect that of 114B; so I hope you can use the simpler formulae of 114B as a guide here.

115C Definition Now, and for the rest of this section, define $\theta : \mathcal{P}(\mathbb{R}^r)\to[0,\infty]$ by writing

$$\theta A = \inf\{\sum_{j=0}^{\infty}\lambda I_j : \langle I_j\rangle_{j\in\mathbb{N}} \text{ is a sequence of half-open intervals such that } A\subseteq\bigcup_{j\in\mathbb{N}} I_j\}.$$

Observe that every A can be covered by some sequence of half-open intervals – e.g., $A\subseteq\bigcup_{n\in\mathbb{N}}[-\mathbf{n},\mathbf{n}[$, writing $\mathbf{n}=(n,n,\dots,n)\in\mathbb{R}^r$; so that if we interpret the sums in $[0,\infty]$, as in 112Bc above, we always have a non-empty set to take the infimum of, and θA is always defined in $[0,\infty]$.

This function θ is called **Lebesgue outer measure** on $\mathbb{R}^r$; the phrase is justified by (a) of the next proposition.

115D Proposition (a) θ is an outer measure on $\mathbb{R}^r$.

(b) $\theta I = \lambda I$ for every half-open interval $I\subseteq\mathbb{R}^r$.

proof (a)(i) θ takes values in $[0,\infty]$ because every θA is the infimum of a non-empty subset of $[0,\infty]$.

(ii) $\theta\emptyset = 0$ because (for instance) if we set $I_j=\emptyset$ for every j, then every I_j is a half-open interval (on the convention I am using), $\emptyset\subseteq\bigcup_{j\in\mathbb{N}}I_j$ and $\sum_{j=0}^{\infty}\lambda I_j = 0$.

(iii) If $A \subseteq B$ then whenever $B \subseteq \bigcup_{j\in\mathbb{N}} I_j$ we have $A \subseteq \bigcup_{j\in\mathbb{N}} I_j$, so θA is the infimum of a set at least as large as that involved in the definition of θB, and $\theta A \le \theta B$.

(iv) Now suppose that $\langle A_n\rangle_{n\in\mathbb{N}}$ is a sequence of subsets of $\mathbb{R}^r$, with union A. For any $\epsilon > 0$, we can choose, for each $n \in \mathbb{N}$, a sequence $\langle I_{nj}\rangle_{j\in\mathbb{N}}$ of half-open intervals such that $A_n \subseteq \bigcup_{j\in\mathbb{N}} I_{nj}$ and $\sum_{j=0}^{\infty} \lambda I_{nj} \le \theta A_n + 2^{-n}\epsilon$. (You should perhaps check that this formulation is valid whether θA_n is finite or infinite.) Now by 111F(b-ii) there is a bijection from $\mathbb{N}$ to $\mathbb{N}\times\mathbb{N}$; express this in the form $m \mapsto (k_m, l_m)$. Then we find that

$$\sum_{m=0}^{\infty} \lambda I_{k_m,l_m} = \sum_{n=0}^{\infty}\sum_{j=0}^{\infty} \lambda I_{nj}.$$

(To see this, note that because every λI_{nj} is greater than or equal to 0, and $m \mapsto (k_m, l_m)$ is a bijection, both sums are equal to

$$\sup_{K\subseteq\mathbb{N}\times\mathbb{N} \text{ is finite}} \sum_{(n,j)\in K} \lambda I_{nj}.$$

Or look at the argument written out in 114D.) But now $\langle I_{k_m,l_m}\rangle_{m\in\mathbb{N}}$ is a sequence of half-open intervals and

$$A = \bigcup_{n\in\mathbb{N}} A_n \subseteq \bigcup_{n\in\mathbb{N}}\bigcup_{j\in\mathbb{N}} I_{nj} = \bigcup_{m\in\mathbb{N}} I_{k_m,l_m},$$

so

$$\begin{aligned}\theta A &\le \sum_{m=0}^{\infty} \lambda I_{k_m,l_m} = \sum_{n=0}^{\infty}\sum_{j=0}^{\infty} \lambda I_{nj}\\ &\le \sum_{n=0}^{\infty}(\theta A_n + 2^{-n}\epsilon) = \sum_{n=0}^{\infty}\theta A_n + \sum_{n=0}^{\infty} 2^{-n}\epsilon = \sum_{n=0}^{\infty}\theta A_n + 2\epsilon.\end{aligned}$$

Because ϵ is arbitrary, $\theta A \le \sum_{n=0}^{\infty}\theta A_n$ (again, you should check that this is valid whether or not $\sum_{n=0}^{\infty}\theta A_n$ is finite). As $\langle A_n\rangle_{n\in\mathbb{N}}$ is arbitrary, θ is an outer measure.

(b) Because we can always take $I_0 = I$, $I_j = \emptyset$ for $j \ge 1$, to obtain a sequence of half-open intervals covering I with $\sum_{j=0}^{\infty} \lambda I_j = \lambda I$, we surely have $\theta I \le \lambda I$. For the reverse inequality, use 115B; if $I \subseteq \bigcup_{j\in\mathbb{N}} I_j$, then $\lambda I \le \sum_{j=0}^{\infty} \lambda I_j$; as $\langle I_j\rangle_{j\in\mathbb{N}}$ is arbitrary, $\theta I \ge \lambda I$ and $\theta I = \lambda I$, as required.

115E Definition Because Lebesgue outer measure (115C) is indeed an outer measure (115Da), we may use it to construct a measure μ, using Carathéodory's method (113C). This measure is **Lebesgue measure on** $\mathbb{R}^r$. The sets E for which μE is defined (that is, for which $\theta(A\cap E) + \theta(A\setminus E) = \theta A$ for every $A \subseteq \mathbb{R}^r$) are called **Lebesgue measurable**.

Sets which are negligible for μ are called **Lebesgue negligible**; note that these are just the sets A for which $\theta A = 0$, and are all Lebesgue measurable (113Xa).

115F Lemma If $i \le r$ and $\xi \in \mathbb{R}$, then $H_{i\xi} = \{y : \eta_i < \xi\}$ is Lebesgue measurable.

proof Write H for $H_{i\xi}$.

(a) The point is that $\lambda I = \lambda(I\cap H) + \lambda(I\setminus H)$ for every half-open interval $I \subseteq \mathbb{R}^r$. **P** If either $I \subseteq H$ or $I \cap H = \emptyset$, this is trivial. Otherwise, I must be of the form $[a, b[$, where $\alpha_i < \xi < \beta_i$. Now $I \cap H = [a, x[$ and $I \setminus H = [y, b[$, where $\xi_j = \beta_j$ for $j \ne i$, $\xi_i = \xi$, $\eta_j = \alpha_j$ for $j \ne i$, $\eta_i = \xi$, so both are half-open intervals, and

$$\begin{aligned}\lambda(I\cap H) + \lambda(I\setminus H) &= (\xi - \alpha_i)\prod_{j\ne i}(\beta_j - \alpha_j) + (\beta_i - \xi)\prod_{j\ne i}(\beta_j - \alpha_j)\\ &= (\beta_i - \alpha_i)\prod_{j\ne i}(\beta_j - \alpha_j) = \lambda I. \textbf{ Q}\end{aligned}$$

(b) Now suppose that A is any subset of $\mathbb{R}^r$, and $\epsilon > 0$. Then we can find a sequence $\langle I_j\rangle_{j\in\mathbb{N}}$ of half-open intervals such that $A \subseteq \bigcup_{j\in\mathbb{N}} I_j$ and $\sum_{j=0}^{\infty} \lambda I_j \le \theta A + \epsilon$. In this case, $\langle I_j \cap H\rangle_{j\in\mathbb{N}}$ amd $\langle I_j \setminus H\rangle_{j\in\mathbb{N}}$ are sequences of half-open intervals, $A \cap H \subseteq \bigcup_{j\in\mathbb{N}}(I_j \cap H)$ and $A \setminus H \subseteq \bigcup_{j\in\mathbb{N}}(I_j \setminus H)$. So

$$\begin{aligned}\theta(A\cap H) + \theta(A\setminus H) &\le \sum_{j=0}^{\infty}\lambda(I_j\cap H) + \sum_{j=0}^{\infty}\lambda(I_j\setminus H)\\ &= \sum_{j=0}^{\infty}\lambda I_j \le \theta A + \epsilon.\end{aligned}$$

Because ϵ is arbitrary, $\theta(A \cap H) + \theta(A \setminus H) \le \theta A$; because A is arbitrary, H is measurable, as remarked in 113D.

115G Proposition All Borel subsets of $\mathbb{R}^r$ are Lebesgue measurable; in particular, all open sets, and all sets of the following classes, together with countable unions of them:

open intervals $]a, b[= \{x : x \in \mathbb{R}^r,\ \alpha_i < \xi_i < \beta_i\ \forall\ i \le r\}$, where $-\infty \le \alpha_i < \beta_i \le \infty$ for each $i \le r$;

closed intervals $[a, b] = \{x : x \in \mathbb{R}^r,\ \alpha_i \le \xi_i \le \beta_i\ \forall\ i \le r\}$, where $-\infty < \alpha_i < \beta_i < \infty$ for each $i \le r$.

We have moreover the following formula for the measures of such sets, writing μ for Lebesgue measure:

$$\mu\,]a, b[= \mu[a, b] = \textstyle\prod_{i=1}^r \beta_i - \alpha_i$$

whenever $a \le b$ in $\mathbb{R}^r$. Consequently every countable subset of $\mathbb{R}^r$ is measurable and of zero measure.

proof (a) I show first that all open subsets of $\mathbb{R}^r$ are measurable. **P** Let $G \subseteq \mathbb{R}^r$ be open. Let $K \subseteq \mathbb{Q}^r \times \mathbb{Q}^r$ be the set of pairs (c, d) of r-tuples of rational numbers such that $[c, d[\subseteq G$. Now by the remarks in 111E-111F – specifically, 111Eb, showing that $\mathbb{Q}$ is countable, 111F(b-iii), showing that the product of two countable sets is countable, and 111F(b-i), showing that subsets of countable sets are countable – we see, inducing on r, that $\mathbb{Q}^r$ is countable, and that K is countable. Also, every $[c, d[$ is measurable, being

$$\textstyle\bigcap_{i \le r} H_{i\delta_i} \setminus H_{i\gamma_i},$$

in the language of 115F, if $c = (\gamma_1, \dots, \gamma_r)$ and $d = (\delta_1, \dots, \delta_r)$. So, by 111Fa, $G' = \bigcup_{(r,s) \in K} [r, s[$ is measurable.

By the definition of K, $G' \subseteq G$. On the other hand, if $x \in G$, there is an $\epsilon > 0$ such that $y \in G$ whenever $\|y - x\| < \epsilon$. Now for each i there are rational numbers γ_i, δ_i such that $\gamma_i \le \xi_i < \delta_i$ and $\delta_i - \gamma_i \le \frac{\epsilon}{\sqrt{r}}$. If $y \in [c, d[$ then $|\eta_i - \xi_i| < \frac{\epsilon}{\sqrt{r}}$ for every i so $\|y - x\| < \epsilon$ and $y \in G$. Accordingly $(c, d) \in K$ and $x \in [c, d[\subseteq G'$. As x is arbitrary, $G = G'$ and G is measurable. **Q**

(b) Now the family Σ of Lebesgue measurable sets is a σ-algebra of subsets of $\mathbb{R}^r$ including the family of open sets, so must contain every Borel set, by the definition of Borel set (111G).

(c) Of the types of interval considered, all the open intervals are actually open sets, so are surely Borel. A closed interval $[a, b]$ is expressible as the intersection $\bigcap_{n \in \mathbb{N}}]a - 2^{-n}\mathbf{1}, b + 2^{-n}\mathbf{1}[$ of a sequence of open intervals, so is Borel.

(d) To compute the measures, we already know from 115Db that $\mu\,[a, b[= \prod_{i=1}^r \beta_i - \alpha_i$ if $a \le b$. For the other types of bounded interval, it is enough to note that if $-\infty < \alpha_i < \beta_i < \infty$ for every i, then

$$[a + \epsilon\mathbf{1}, b[\subseteq]a, b[\subseteq [a, b] \subseteq [a, b + \epsilon\mathbf{1}[$$

whenever $\epsilon > 0$ in $\mathbb{R}$. So

$$\mu\,]a, b[\le \mu[a, b] \le \inf_{\epsilon > 0} \mu\,[a, b + \epsilon\mathbf{1}[= \inf_{\epsilon > 0} \textstyle\prod_{i=1}^r (\beta_i - \alpha_1 + \epsilon) = \prod_{i=1}^r \beta_i - \alpha_i.$$

If $\beta_i = \alpha_i$ for any i, then we must have

$$\mu\,]a, b[= \mu[a, b] = 0 = \textstyle\prod_{i=1}^r \beta_i - \alpha_i.$$

If $\beta_i > \alpha_i$ for every i, then set $\epsilon_0 = \min_{i \le r} \beta_i - \alpha_i > 0$; then

$$\begin{aligned} \mu[a, b] \ge \mu\,]a, b[&\ge \sup_{0 < \epsilon \le \epsilon_0} \mu\,[a + \epsilon\mathbf{1}, b[\\ &= \sup_{0 < \epsilon \le \epsilon_0} \prod_{i=1}^r (\beta_i - \alpha_i - \epsilon) = \prod_{i=1}^r \beta_i - \alpha_i. \end{aligned}$$

So in this case

$$\textstyle\prod_{i=1}^r \beta_i - \alpha_i \le \mu\,]a, b[\le \mu[a, b] \le \prod_{i=1}^r \beta_i - \alpha_i$$

and

$$\mu\,]a, b[= \mu[a, b] = \textstyle\prod_{i=1}^r \beta_i - \alpha_i.$$

(e) By (d), $\mu\{a\} = \mu[a, a] = 0$ for every a. If $A \subseteq \mathbb{R}^r$ is countable, it is either empty or expressible as $\{a_n : n \in \mathbb{N}\}$. In the former case $\mu A = \mu\emptyset = 0$; in the latter, $A = \bigcup_{n \in \mathbb{N}} \{a_n\}$ is Borel and $\mu A \le \sum_{n=0}^\infty \mu\{a_n\} = 0$.

115X Basic exercises If you skipped §114, you should now return to 114X and assure yourself that you can do the exercises there as well as those below.

(a) Show that if I, J are half-open intervals in $\mathbb{R}^r$, then $I\setminus J$ is expressible as the union of at most $2r$ disjoint half-open intervals. Hence show that (i) any finite union of half-open intervals is expressible as a finite union of disjoint half-open intervals (ii) any countable union of half-open intervals is expressible as the union of a disjoint sequence of half-open intervals.

>(b) Write θ for Lebesgue outer measure, μ for Lebesgue measure on $\mathbb{R}^r$. Show that $\theta A=\inf\{\mu E: E$ is Lebesgue measurable, $A\subseteq E\}$ for every $A\subseteq\mathbb{R}^r$. (*Hint*: consider sets E of the form $\bigcup_{j\in\mathbb{N}}I_j$, where $\langle I_j\rangle_{j\in\mathbb{N}}$ is a sequence of half-open intervals.)

(c) Let $E\subseteq\mathbb{R}^r$ be a set of finite measure for Lebesgue measure μ. Show that for every $\epsilon>0$ there is a disjoint family $I_0,\dots,I_n$ of half-open intervals such that $\mu(E\triangle\bigcup_{j\le n}I_j)\le\epsilon$. (*Hint*: let $\langle J_j\rangle_{j\in\mathbb{N}}$ be a sequence of half-open intervals such that $E\subseteq\bigcup_{j\in\mathbb{N}}J_j$ and $\sum_{j=0}^{\infty}\mu J_j\le\mu E+\frac{1}{2}\epsilon$. Now take a suitably large m and express $\bigcup_{j\le m}J_j$ as a disjoint union of half-open intervals.)

>(d) Suppose that $c\in\mathbb{R}^r$. (i) Show that $\theta(A+c)=\theta A$ for every $A\subseteq\mathbb{R}^r$, where $A+c=\{x+c:x\in A\}$. (ii) Show that if $E\subseteq\mathbb{R}^r$ is measurable so is $E+c$, and that in this case $\mu(E+c)=\mu E$.

(e) Suppose that $\gamma>0$. (i) Show that $\theta(\gamma A)=\gamma^r\theta A$ for every $A\subseteq\mathbb{R}^r$, where $\gamma A=\{\gamma x:x\in A\}$. (ii) Show that if $E\subseteq\mathbb{R}^r$ is measurable so is γE, and that in this case $\mu(\gamma E)=\gamma^r\mu E$

115Y Further exercises (a) (i) Suppose that M is a strictly positive integer and k_i, l_i are integers for $1\le i\le r$. Set $\alpha_i=k_i/M$ and $\beta_i=l_i/M$ for each i, and $I=[a,b[$. Show that $\lambda I=\#(J)/M^r$, where J is $\{z:z\in\mathbb{Z}^r,\ \frac{1}{M}z\in I\}$. (ii) Show that if a half-open interval $I\subseteq\mathbb{R}^r$ is covered by a *finite* sequence $I_0,\dots,I_m$ of half-open intervals, and all the coordinates involved in specifying the intervals I, $I_0,\dots,I_m$ are rational, then $\lambda I\le\sum_{j=0}^{m}\lambda I_j$. (iii) Assuming the Heine-Borel theorem in the form

whenever $[a,b]$ is a closed interval in $\mathbb{R}^r$ which is covered by a sequence $\langle\,]a^{(j)},b^{(j)}[\,\rangle_{j\in\mathbb{N}}$ of open intervals, there is an $m\in\mathbb{N}$ such that $[a,b]\subseteq\bigcup_{j\le m}\,]a^{(j)},b^{(j)}[$,

prove 115B. (*Hint*: if $[a,b[\,\subseteq\bigcup_{j\in\mathbb{N}}\left[a^{(j)},b^{(j)}\right[$, replace $[a,b[$ by a smaller closed interval and each $\left[a^{(j)},b^{(j)}\right[$ by a larger open interval, changing the volumes by adequately small amounts.)

(b)(i) Show that if $A\subseteq\mathbb{R}^r$ and $\epsilon>0$, there is an open set $G\supseteq A$ such that $\theta G\le\theta A+\epsilon$, where θ is Lebesgue outer measure. (ii) Show that if $E\subseteq\mathbb{R}^r$ is Lebesgue measurable and $\epsilon>0$, there is an open set $G\supseteq E$ such that $\mu(G\setminus E)\le\epsilon$, where μ is Lebesgue measure. (*Hint*: consider first the case of bounded E.) (iii) Show that if $E\subseteq\mathbb{R}^r$ is Lebesgue measurable, there are Borel sets H_1, H_2 such that $H_1\subseteq E\subseteq H_2$ and $\mu(H_2\setminus E)=\mu(E\setminus H_1)=0$. (*Hint*: use (ii) to find H_2, and then consider the complement of E.)

(c) Write θ for Lebesgue outer measure on $\mathbb{R}^r$. Show that a set $E\subseteq\mathbb{R}^r$ is Lebesgue measurable iff $\theta([-\mathbf{n},\mathbf{n}]\cap E)+\theta([-\mathbf{n},\mathbf{n}]\setminus E)=(2n)^r$ for every $n\in\mathbb{N}$, writing $\mathbf{n}=(n,\dots,n)$. (*Hint*: use 115Yb to show that for each n there are measurable sets F_n, H_n such that $F_n\subseteq[-\mathbf{n},\mathbf{n}]\cap E\subseteq H_n$ and $H_n\setminus F_n$ is negligible.)

(d) Assuming that there is a set $A\subseteq\mathbb{R}$ which is not a Borel set, show that there is a family $\mathcal{E}$ of half-open intervals in $\mathbb{R}^2$ such that $\bigcup\mathcal{E}$ is not a Borel set. (*Hint*: consider $\mathcal{E}=\{[\xi,1+\xi[\,\times[-\xi,1-\xi[\,:\xi\in A\}$.)

(e) Let X be a set and $\mathcal{A}$ a **semiring** of subsets of X, that is, a family of subsets of X such that

$\emptyset\in\mathcal{A}$,

$E\cap F\in\mathcal{A}$ for all E, $F\in\mathcal{A}$,

whenever E, $F\in\mathcal{A}$ there are disjoint $E_0,\dots,E_n\in\mathcal{A}$ such that $E\setminus F=E_0\cup\ldots\cup E_n$.

Let $\lambda:\mathcal{A}\to[0,\infty]$ be a functional such that

$\lambda\emptyset=0$,

$\lambda E=\sum_{i=0}^{\infty}\lambda E_i$ whenever $E\in\mathcal{A}$ and $\langle E_i\rangle_{i\in\mathbb{N}}$ is a disjoint sequence in $\mathcal{A}$ with union E.

Show that there is a measure μ on X extending λ. (*Hint*: use the method of 113Yi.)

115 Notes and comments In the notes to §114 I ran over the methods so far available to us for the construction of measure spaces. To the list there we can now add Lebesgue measure on $\mathbb{R}^r$.

If you look back at §114, you will see that I have deliberately copied the exposition there. I hope that this duplication will help you to see the essential elements of the method, which are three: a primitive concept of volume (114A/115A); countable subadditivity (114B/115B); and measurability of building blocks (114F/115F).

Concerning the 'primitive concept of volume' there is not much to be said. The ideas of length of an interval, area of a rectangle and volume of a cuboid go back to the beginning of mathematics. I use 'half-open intervals', as defined in 114Aa/115Ab, for purely technical reasons, because they fit together neatly (see 115Xa and 115Ye); if we started with 'open' or 'closed' intervals the method would still work. One thing is perhaps worth mentioning: the blocks I use are all upright, with edges parallel to the coordinate axes. It is in fact a non-trivial exercise to prove that a block in any other orientation has the right Lebesgue measure, and I delay this until Chapter 26. For the moment we are looking for the shortest safe path to a precise definition, and the fact that rotating a set doesn't change its Lebesgue measure will have to wait.

The big step is 'countable subadditivity': the fact that if one block is covered by a sequence of other blocks, its volume is less than or equal to the sum of theirs. This is surely necessary if blocks are to be measurable with the right measures, by 112Cd. (What is remarkable is that it is so nearly sufficient.) Here we have some work to do, and in the r-dimensional case there is a substantial hill to climb. You can do the climb in two stages if you look up the Heine-Borel theorem (115Ya); but as I try to explain in the remarks following 115B, I do not think that this route avoids any of the real difficulties.

The third thing we must check is that blocks are measurable in the technical sense described by Carathéodory's theorem. This is because they are obtainable by the operations of intersection and union and complementation from half-spaces, and half-spaces are measurable for very straightforward reasons (114F/115F). Now we are well away, and I do very little more, only checking that open sets, and therefore Borel sets, are measurable, and that closed and open intervals have the right measures (114G/115G). Some more properties of Lebesgue measure can be found in §134. But every volume, if not quite every chapter, of this treatise will introduce further features of this extraordinary construction.

Chapter 12

Integration

If you look along the appropriate shelf of your college's library, you will see that the words 'measure' and 'integration' go together like Siamese twins. The linkage is both more complex and more intimate than any simple explanation can describe. But if we say that one of the concepts on which integration is based is that of 'area under a curve', then it is clear that any method of determining 'areas' ought to correspond to a method of integrating functions; and this has from the beginning been an essential part of the Lebesgue theory. For a literal description of the integral of a non-negative function in terms of the area of its ordinate set, I think it best to wait until Chapter 25 in Volume 2. In the present chapter I seek to give a concise description of the standard integral of a real-valued function on a general measure space, with the half-dozen most important theorems concerning this integral.

The construction bristles with technical difficulties at every step, and you will find it easy to understand why it was not done before 1901. What may be less clear is why it was ever done at all. So perhaps you should immediately read the statements of 123A-123D below. It is the case (some of the details will appear, rather late, in §436 in Volume 4) that any theory of integration powerful enough to have theorems of this kind must essentially encompass all the ideas of this chapter, and nearly all the ideas of the last.

121 Measurable functions

In this section, I take a step back to develop ideas relating to σ-algebras of sets, following §111; there will be no mention of 'measures' here, except in the exercises. The aim is to establish the concept of 'measurable function' (121C) and a variety of associated techniques. The best single example of a σ-algebra to bear in mind when reading this chapter is probably the σ-algebra of Borel subsets of $\mathbb{R}$ (111G); the σ-algebra of Lebesgue measurable subsets of $\mathbb{R}$ (114E) is a good second.

Throughout the exposition here (starting with 121A) I seek to deal with functions which are not defined on the whole of the space X under consideration. I believe that there are compelling reasons for facing up to such functions at an early stage (see 121G); but undeniably they add to the technical difficulties, and it would be fair to read through the chapter once with the mental reservation that all functions are taken to be defined everywhere, before returning to deal with the general case.

121A Lemma Let X be a set and Σ a σ-algebra of subsets of X. Let D be any subset of X and write

$$\Sigma_D = \{E \cap D : E \in \Sigma\}.$$

Then Σ_D is a σ-algebra of subsets of D.

proof (i) $\emptyset = \emptyset \cap D \in \Sigma_D$ because $\emptyset \in \Sigma$.

(ii) If $F \in \Sigma_D$, there is an $E \in \Sigma$ such that $F = E \cap D$; now $D \setminus F = (X \setminus E) \cap D \in \Sigma_D$ because $X \setminus E \in \Sigma$.

(iii) If $\langle F_n \rangle_{n\in\mathbb{N}}$ is any sequence in Σ_D, then for each $n \in \mathbb{N}$ we may choose an $E_n \in \Sigma$ such that $F_n = E_n \cap D$; now $\bigcup_{n\in\mathbb{N}} F_n = (\bigcup_{n\in\mathbb{N}} E_n) \cap D \in \Sigma_D$ because $\bigcup_{n\in\mathbb{N}} E_n \in \Sigma$.

Notation I will call Σ_D the **subspace σ-algebra** of subsets of D, and I will say that its members are **relatively measurable** in D. Σ_D is also sometimes called the **trace** of Σ on D.

121B Proposition Let X be a set, Σ a σ-algebra of subsets of X, and D a subset of X. Write Σ_D for the subspace σ-algebra of subsets of D. Then for any function $f : D \to \mathbb{R}$ the following assertions are equiveridical, that is, if one of them is true so are all the others:

(i) $\{x : f(x) < a\} \in \Sigma_D$ for every $a \in \mathbb{R}$;
(ii) $\{x : f(x) \le a\} \in \Sigma_D$ for every $a \in \mathbb{R}$;
(iii) $\{x : f(x) > a\} \in \Sigma_D$ for every $a \in \mathbb{R}$;
(iv) $\{x : f(x) \ge a\} \in \Sigma_D$ for every $a \in \mathbb{R}$.

proof (i)$\Rightarrow$(ii) Assume (i), and let $a \in \mathbb{R}$. Then

$$\{x : f(x) \le a\} = \bigcap_{n\in\mathbb{N}} \{x : f(x) < a + 2^{-n}\} \in \Sigma_D$$

because $\{x : f(x) < a + 2^{-n}\} \in \Sigma_D$ for every n and Σ_D is closed under countable intersections (111Dd). Because a is arbitrary, (ii) is true.

(ii)⇒(iii) Assume (ii), and let $a \in \mathbb{R}$. Then

$$\{x : f(x) > a\} = D \setminus \{x : f(x) \le a\} \in \Sigma_D$$

because $\{x : f(x) \le a\} \in \Sigma_D$ and Σ_D is closed under complementation. Because a is arbitrary, (iii) is true.

(iii)⇒(iv) Assume (iii), and let $a \in \mathbb{R}$. Then

$$\{x : f(x) \ge a\} = \bigcap_{n\in\mathbb{N}}\{x : f(x) > a - 2^{-n}\} \in \Sigma_D$$

because $\{x : f(x) > a - 2^{-n}\} \in \Sigma_D$ for every n and Σ_D is closed under countable intersections. Because a is arbitrary, (iv) is true.

(iv)⇒(i) Assume (iv), and let $a \in \mathbb{R}$. Then

$$\{x : f(x) < a\} = D \setminus \{x : f(x) \ge a\} \in \Sigma_D$$

because $\{x : f(x) \ge a\} \in \Sigma_D$ and Σ_D is closed under complementation. Because a is arbitrary, (i) is true.

121C Definition Let X be a set, Σ a σ-algebra of subsets of X, and D a subset of X. A function $f : D \to \mathbb{R}$ is called **measurable** (or **Σ-measurable**) if it satisfies any, or equivalently all, of the conditions (i)-(iv) of 121B.

If X is $\mathbb{R}$ or $\mathbb{R}^r$, and Σ is its Borel σ-algebra (111G), a Σ-measurable function is called **Borel measurable**. If X is $\mathbb{R}$ or $\mathbb{R}^r$, and Σ is the σ-algebra of Lebesgue measurable sets (114E, 115E), a Σ-measurable function is called **Lebesgue measurable**.

Remark Naturally the principal case here is when $D = X$. However, partially-defined functions are so common, and so important, in analysis (consider, for instance, the real function $\ln\sin$) that it seems worth while, from the beginning, to establish techniques for handling them efficiently.

Many authors develop a theory of 'extended real numbers' at this point, working with $[-\infty, \infty] = \mathbb{R} \cup \{-\infty, \infty\}$, and defining measurability for functions taking values in this set. I outline such a theory in §135 below.

121D Proposition Let X be $\mathbb{R}^r$ for some $r \ge 1$, D a subset of X, and $g : D \to \mathbb{R}$ a function.

(a) If g is Borel measurable it is Lebesgue measurable.
(b) If g is continuous it is Borel measurable.
(c) If $r = 1$ and g is monotonic it is Borel measurable.

proof (a) This is immediate from the definitions in 121C, if we recall that the Borel σ-algebra is included in the Lebesgue σ-algebra (114G, 115G).

(b) Take $a \in \mathbb{R}$. Set

$$\mathcal{G} = \{G : G \subseteq \mathbb{R}^r \text{ is open, } g(x) < a \ \forall\, x \in G \cap D\},$$

$$G_0 = \bigcup \mathcal{G} = \{x : \exists\ G \in \mathcal{G},\, x \in G\}.$$

Then G_0 is a union of open sets, therefore open (1A2Bd). Next,

$$\{x : g(x) < a\} = G_0 \cap D.$$

P (i) If $g(x) < a$, then (because g is continuous) there is a $\delta > 0$ such that $|g(y) - g(x)| < a - g(x)$ whenever $y \in D$ and $\|y - x\| < \delta$. But $\{y : \|y - x\| < \delta\}$ is open (1A2D), so belongs to $\mathcal{G}$ and is included in G_0, and $x \in G_0 \cap D$. **(ii)** If $x \in G_0 \cap D$, then there is a $G \in \mathcal{G}$ such that $x \in G$; now $g(y) < a$ for every $y \in G \cap D$, so, in particular, $g(x) < a$. **Q**

Finally, G_0, being open, is a Borel set. As a is arbitrary, g is Borel measurable.

(c) Suppose first that g is non-decreasing. Let $a \in \mathbb{R}$ and write $E = \{x : g(x) < a\}$. If $E = D$ or $E = \emptyset$ then of course it is the intersection of D with a Borel set. Otherwise, E is non-empty and bounded above in $\mathbb{R}$, so has a supremum $c \in \mathbb{R}$. Now E must be either $D \cap]-\infty, c[$ or $D \cap]-\infty, c]$, according to whether $c \in E$ or not, and in either case is the intersection of D with a Borel set (see 114G).

Similarly, if g is non-increasing, $\{x : g(x) > a\}$ will again be the intersection of D with either $\emptyset$ or $\mathbb{R}$ or $]-\infty, c]$ or $]-\infty, c[$ for some c. So in this case 121B(iii) will be satisfied.

Remark I see that in part (b) of the above proof I use some basic facts about open sets in $\mathbb{R}^r$. These are covered in detail in §1A2. If they are new to you it would probably be sensible to rehearse the arguments with $r = 1$, so that $D \subseteq \mathbb{R}$, before embracing the general case.

121E Theorem Let X be a set and Σ a σ-algebra of subsets of X. Let f and g be real-valued functions defined on domains $\operatorname{dom} f$, $\operatorname{dom} g \subseteq X$.

(a) If f is constant it is measurable.

(b) If f and g are measurable, so is $f+g$, where $(f+g)(x) = f(x)+g(x)$ for $x \in \operatorname{dom} f \cap \operatorname{dom} g$.

(c) If f is measurable and $c \in \mathbb{R}$, then cf is measurable, where $(cf)(x) = c \cdot f(x)$ for $x \in \operatorname{dom} f$.

(d) If f and g are measurable, so is $f \times g$, where $(f \times g)(x) = f(x) \times g(x)$ for $x \in \operatorname{dom} f \cap \operatorname{dom} g$.

(e) If f and g are measurable, so is f/g, where $(f/g)(x) = f(x)/g(x)$ when $x \in \operatorname{dom} f \cap \operatorname{dom} g$ and $g(x) \neq 0$.

(f) If f is measurable and $E \subseteq \mathbb{R}$ is a Borel set, then there is an $F \in \Sigma$ such that $f^{-1}[E] = \{x : f(x) \in E\}$ is equal to $F \cap \operatorname{dom} f$.

(g) If f is measurable and h is a Borel measurable function from a subset $\operatorname{dom} h$ of $\mathbb{R}$ to $\mathbb{R}$, then hf is measurable, where $(hf)(x) = h(f(x))$ for $x \in \operatorname{dom}(hf) = \{y : y \in \operatorname{dom} f,\ f(y) \in \operatorname{dom} h\}$.

(h) If f is measurable and A is any set, then $f \restriction A$ is measurable, where $\operatorname{dom}(f \restriction A) = A \cap \operatorname{dom} f$ and $(f \restriction A)(x) = f(x)$ for $x \in A \cap \operatorname{dom} f$.

proof For any $D \subseteq X$ write Σ_D for the subspace σ-algebra of subsets of D.

(a) If $f(x) = c$ for every $x \in \operatorname{dom} f$, then $\{x : f(x) < a\} = \operatorname{dom} f$ if $c < a$, $\emptyset$ otherwise, and in either case belongs to $\Sigma_{\operatorname{dom} f}$.

(b) Write $D = \operatorname{dom}(f+g) = \operatorname{dom} f \cap \operatorname{dom} g$. If $a \in \mathbb{R}$ then set $K = \{(q,q') : q,\ q' \in \mathbb{Q},\ q+q' \le a\}$. Then K is a subset of $\mathbb{Q} \times \mathbb{Q}$, so is countable (111Fb, 1A1E). For $q \in \mathbb{Q}$ choose sets F_q, $G_q \in \Sigma$ such that

$$\{x : f(x) < q\} = F_q \cap \operatorname{dom} f, \quad \{x : g(x) < q\} = G_q \cap \operatorname{dom} g.$$

For each $(q,q') \in K$, the set

$$E_{qq'} = \{x : f(x) < q,\ g(x) < q'\} = F_q \cap G_{q'} \cap D$$

belongs to Σ_D. Finally, if $(f+g)(x) < a$, then we can find $q \in \,]f(x), a-g(x)[$, $q' \in \,]g(x), a-q]$, so that $(q,q') \in K$ and $x \in E_{qq'}$; while if $(q,q') \in K$ and $x \in E_{qq'}$, then $(f+g)(x) < q+q' \le a$. Thus

$$\{x : (f+g)(x) < a\} = \bigcup_{(q,q') \in K} E_{qq'} \in \Sigma_D$$

by 111Fa. As a is arbitrary, $f+g$ is measurable.

(c) Write $D = \operatorname{dom} f$. Let $a \in \mathbb{R}$. If $c > 0$, then

$$\{x : cf(x) < a\} = \{x : f(x) < \tfrac{a}{c}\} \in \Sigma_D.$$

If $c < 0$, then

$$\{x : cf(x) < a\} = \{x : f(x) > \tfrac{a}{c}\} \in \Sigma_D.$$

While if $c = 0$, then $\{x : cf(x) < a\}$ is either D or $\emptyset$, as in (a) above, so belongs to Σ_D. As a is arbitrary, cf is measurable.

(d) Write $D = \operatorname{dom}(f \times g) = \operatorname{dom} f \cap \operatorname{dom} g$. Let $a \in \mathbb{R}$. Let K be

$$\{(q_1,q_2,q_3,q_4) : q_1, \dots, q_4 \in \mathbb{Q},\ uv < a \text{ whenever } u \in \,]q_1,q_2[\,,\ v \in \,]q_3,q_4[\}.$$

Then K is countable. For $q \in \mathbb{Q}$ choose sets F_q, F'_q, G_q, $G'_q \in \Sigma$ such that

$$\{x : f(x) < q\} = F_q \cap \operatorname{dom} f, \quad \{x : f(x) > q\} = F'_q \cap \operatorname{dom} f,$$

$$\{x : g(x) < q\} = G_q \cap \operatorname{dom} g, \quad \{x : g(x) > q\} = G'_q \cap \operatorname{dom} g.$$

For $(q_1,q_2,q_3,q_4) \in K$ set

$$\begin{aligned} E_{q_1q_2q_3q_4} &= \{x : f(x) \in \,]q_1,q_2[\,,\ g(x) \in \,]q_3,q_4[\} \\ &= D \cap F'_{q_1} \cap F_{q_2} \cap G'_{q_3} \cap G_{q_4} \in \Sigma_D; \end{aligned}$$

then $E = \bigcup_{(q_1,q_2,q_3,q_4) \in K} E_{q_1q_2q_3q_4} \in \Sigma_D$.

Now $E = \{x : (f \times g)(x) < a\}$. **P (i)** If $(f \times g)(x) < a$, set $u = f(x)$, $v = g(x)$. Set

$$\eta = \min(1, \tfrac{a-uv}{1+|u|+|v|}) > 0.$$

Take $q_1, \dots, q_4 \in \mathbb{Q}$ such that

$$u-\eta \le q_1 < u < q_2 \le u+\eta, \quad v-\eta \le q_3 < v < q_4 \le v+\eta.$$

If $u' \in \,]q_1,q_2[$, $v' \in \,]q_3,q_4[$, then $|u'-u| < \eta$ and $|v'-v| < \eta$, so

$$\begin{aligned} u'v'-uv &= (u'-u)(v'-v)+(u'-u)v+u(v'-v) \\ &< \eta^2+\eta|v|+|u|\eta \le \eta(1+|u|+|v|) \le a-uv, \end{aligned}$$

and $u'v' < a$. Accordingly $(q_1,q_2,q_3,q_4) \in K$. Also $x \in E_{q_1q_2q_3q_4}$, so $x \in E$. Thus $\{x : (f\times g)(x) < a\} \subseteq E$. **(ii)** On the other hand, if $x \in E$, there are $q_1,\dots,q_4$ such that $(q_1,q_2,q_3,q_4) \in K$ and $x \in E_{q_1q_2q_3q_4}$, so that $f(x) \in \,]q_1,q_2[$ and $g(x) \in \,]q_3,q_4[$ and $f(x)g(x) < a$. So $E \subseteq \{x : (f\times g)(x) < a\}$. **Q**

Thus $\{x : (f\times g)(x) < a\} \in \Sigma_D$. As a is arbitrary, $f\times g$ is measurable.

(e) In view of (d), it will be enough to show that $1/g$ is measurable. Now if $a > 0$, $\{x : 1/g(x) < a\} = \{x : g(x) > 1/a\} \cup \{x : g(x) < 0\}$; if $a < 0$, then $\{x : 1/g(x) < a\} = \{x : 1/a < g(x) < 0\}$; and if $a = 0$, then $\{x : 1/g(x) < a\} = \{x : g(x) < 0\}$. And all of these belong to $\Sigma_{\mathrm{dom}\,1/g}$.

(f) Write $D = \mathrm{dom}\, f$ and consider the set

$$\mathrm{T} = \{E : E \subseteq \mathbb{R},\ f^{-1}[E] \in \Sigma_D\}.$$

Then T is a σ-algebra of subsets of $\mathbb{R}$. **P** (i) $f^{-1}[\emptyset] = \emptyset \in \Sigma_D$, so $\emptyset \in \mathrm{T}$. (ii) If $E \in \mathrm{T}$, then $f^{-1}[\mathbb{R}\setminus E] = D\setminus f^{-1}[E] \in \Sigma_D$ so $\mathbb{R}\setminus E \in \mathrm{T}$. (iii) If $\langle E_n\rangle_{n\in\mathbb{N}}$ is a sequence in T, then $f^{-1}[\bigcup_{n\in\mathbb{N}} E_n] = \bigcup_{n\in\mathbb{N}} f^{-1}[E_n] \in \Sigma_D$ because Σ_D is a σ-algebra, so $\bigcup_{n\in\mathbb{N}} E_n \in \mathrm{T}$. **Q**

Next, T contains all sets of the form $H_a = \,]-\infty,a[$ for $a \in \mathbb{R}$, by the definition of measurability of f. The result follows by arguments already used in 114G above. First, all open subsets of $\mathbb{R}$ belong to T. **P** Let $G \subseteq \mathbb{R}$ be open. Let $K \subseteq \mathbb{Q}\times\mathbb{Q}$ be the set of pairs (q,q') of rational numbers such that $[q,q'[\, \subseteq G$. K is countable. Also, every $[q,q'[$ belongs to T, being $H_{q'}\setminus H_q$. So $G' = \bigcup_{(q,q')\in K} [q,q'[\, \in \mathrm{T}$.

By the definition of K, $G' \subseteq G$. On the other hand, if $x \in G$, there is a $\delta > 0$ such that $]x-\delta,x+\delta[\, \subseteq G$. Now there are rational numbers $q \in \,]x-\delta,x]$ and $q' \in \,]x,x+\delta]$, so that $(q,q') \in K$ and $x \in [q,q'[\, \subseteq G'$. As x is arbitrary, $G = G'$ and $G \in \mathrm{T}$. **Q**

Finally, T is a σ-algebra of subsets of $\mathbb{R}$ including the family of open sets, so must contain every Borel set, by the definition of Borel set (111G).

(g) If $a \in \mathbb{R}$, then $\{y : h(y) < a\}$ is of the form $E\cap \mathrm{dom}\, h$, where E is a Borel subset of $\mathbb{R}$. Next, $f^{-1}[E]$ is of the form $F\cap\mathrm{dom}\, f$, where $F \in \Sigma$, by (f) above. So

$$\{x : (hf)(x) < a\} = F\cap \mathrm{dom}\, hf \in \Sigma_{\mathrm{dom}\, hf}.$$

As a is arbitrary, hf is measurable.

(h) The point is that $\Sigma_{A\cap\mathrm{dom}\, f} = \{E\cap A : E \in \Sigma_{\mathrm{dom}\, f}\}$. So if $a \in \mathbb{R}$,

$$\{x : (f{\restriction}A)(x) < a\} = A\cap\{x : f(x) < a\} \in \Sigma_{\mathrm{dom}(f\restriction A)}.$$

Remarks Of course part (c) of this theorem is just a matter of putting (a) and (d) together, while (e) is a consequence of (d), (g) and the fact that continuous functions are Borel measurable (121Db).

I hope you will recognise the technique in the proof of part (d) as a version of arguments which may be used to prove that the limit of a product is the product of the limits, or that the product of continuous functions is continuous. In fact (b) and (d) here, as well as the theorems on sums and products of limits, are consequences of the fact that addition and multiplication are continuous functions. In 121K I give a general result which may be used to exploit such facts.

Really, part (f) here is the essence of the concept of 'measurable' real-valued function. The point of the definition in 121B-121C is that the Borel σ-algebra of $\mathbb{R}$ can be generated by any of the families $\{]-\infty,a[\, : a \in \mathbb{R}\}$, $\{]-\infty,a] : a \in \mathbb{R}\},\dots$. (See 121Yc(ii).) There are many routes covering this territory in rather fewer words than I have used, at the cost of slightly greater abstraction.

121F Theorem Let X be a set and Σ a σ-algebra of subsets of X. Let $\langle f_n\rangle_{n\in\mathbb{N}}$ be a sequence of Σ-measurable real-valued functions with domains included in X.

(a) Define a function $\lim_{n\to\infty} f_n$ by writing

$$(\lim_{n\to\infty} f_n)(x) = \lim_{n\to\infty} f_n(x)$$

for all those $x \in \bigcup_{n\in\mathbb{N}}\bigcap_{m\ge n} \mathrm{dom}\, f_m$ for which the limit exists in $\mathbb{R}$. Then $\lim_{n\to\infty} f_n$ is Σ-measurable.

(b) Define a function $\sup_{n\in\mathbb{N}} f_n$ by writing

$$(\sup_{n\in\mathbb{N}} f_n)(x) = \sup_{n\in\mathbb{N}} f_n(x)$$

for all those $x \in \bigcap_{n\in\mathbb{N}} \operatorname{dom} f_n$ for which the supremum exists in $\mathbb{R}$. Then $\sup_{n\in\mathbb{N}} f_n$ is Σ-measurable.

(c) Define a function $\inf_{n\in\mathbb{N}} f_n$ by writing

$$(\inf_{n\in\mathbb{N}} f_n)(x) = \inf_{n\in\mathbb{N}} f_n(x)$$

for all those $x \in \bigcap_{n\in\mathbb{N}} \operatorname{dom} f_n$ for which the infimum exists in $\mathbb{R}$. Then $\inf_{n\in\mathbb{N}} f_n$ is Σ-measurable.

(d) Define a function $\limsup_{n\to\infty} f_n$ by writing

$$(\limsup_{n\to\infty} f_n)(x) = \limsup_{n\to\infty} f_n(x)$$

for all those $x \in \bigcup_{n\in\mathbb{N}} \bigcap_{m\ge n} \operatorname{dom} f_m$ for which the lim sup exists in $\mathbb{R}$. Then $\limsup_{n\to\infty} f_n$ is Σ-measurable.

(e) Define a function $\liminf_{n\to\infty} f_n$ by writing

$$(\liminf_{n\to\infty} f_n)(x) = \liminf_{n\to\infty} f_n(x)$$

for all those $x \in \bigcup_{n\in\mathbb{N}} \bigcap_{m\ge n} \operatorname{dom} f_m$ for which the lim inf exists in $\mathbb{R}$. Then $\liminf_{n\in\mathbb{N}} f_n$ is Σ-measurable.

proof For $n \in \mathbb{N}$, $a \in \mathbb{R}$ choose $H_n(a) \in \Sigma$ such that $\{x : f_n(x) \le a\} = H_n(a) \cap \operatorname{dom} f_n$. The proofs are now a matter of observing the following facts:

(a) $\{x : (\lim_{n\to\infty} f_n)(x) \le a\} = \operatorname{dom}(\lim_{n\to\infty} f_n) \cap \bigcap_{k\in\mathbb{N}} \bigcup_{n\in\mathbb{N}} \bigcap_{m\ge n} H_m(a + 2^{-k})$;

(b) $\{x : (\sup_{n\in\mathbb{N}} f_n)(x) \le a\} = \operatorname{dom}(\sup_{n\in\mathbb{N}} f_n) \cap \bigcap_{n\in\mathbb{N}} H_n(a)$;

(c) $\inf_{n\in\mathbb{N}} f_n = -\sup_{n\in\mathbb{N}}(-f_n)$;

(d) $\limsup_{n\to\infty} f_n = \lim_{n\to\infty} \sup_{m\in\mathbb{N}} f_{m+n}$;

(e) $\liminf_{n\to\infty} f_n = -\limsup_{n\to\infty}(-f_n)$.

121G Remarks It is at this point that we first encounter clearly the problem of functions which are not defined everywhere. (The quotient f/g of 121Ee also need not be defined everywhere on the common domain of f and g, but it is less important and more easily dealt with.) The whole point of the theory of measure and integration, since Lebesgue, is that we can deal with limits of sequences of functions; and the set on which $\lim_{n\to\infty} f_n(x)$ exists can be decidedly irregular, even for apparently well-behaved functions f_n. (If you have encountered the theory of Fourier series, then an appropriate example to think of is the sequence of partial sums $f_n(x) = \frac{1}{2}a_0 + \sum_{k=1}^{n}(a_k \cos kx + b_k \sin kx)$ of a Fourier series in which $\sum_{k=1}^{\infty} |a_k| + |b_k| = \infty$, so that the series is not uniformly absolutely summable, but may be conditionally summable at certain points.)

I have tried to make it clear what domains I mean to attach to the functions $\sup_{n\in\mathbb{N}} f_n$, $\lim_{n\to\infty} f_n$, etc. The guiding principle is that they should be the set of all $x \in X$ for which the defining formulae $\sup_{n\in\mathbb{N}} f_n(x)$, $\lim_{n\to\infty} f_n(x)$ can be interpreted as real numbers. (As I noted in 121C, I am for the time being avoiding '∞' as a value of a function, though it gives little difficulty, and some formulae are more naturally interpreted by allowing it.) But in the case of lim, lim sup, lim inf it should be noted that I am using the restrictive definition, that $\lim_{n\to\infty} a_n$ can be regarded as existing only when there is some $n \in \mathbb{N}$ such that a_m is defined for every $m \ge n$. There are occasions when it would be more natural to admit the limit when we know only that a_m is defined for infinitely many m; but such a convention could make 121Fa false, unless care was taken.

As in 111E-111F, we can use the ideas of parts (b), (c) here to discuss functions of the form $\sup_{k\in K} f_k$, $\inf_{k\in K} f_k$ for any family $\langle f_k \rangle_{k\in K}$ of measurable functions indexed by a non-empty countable set K.

In this theorem and the last, the functions f, g, f_n have been permitted to have arbitrary domains, and consequently there is nothing that can be said about the domains of the constructed functions. However, it is of course the case that if the original functions have measurable domains, so do the functions constructed from them by the rules I propose. I spell out the details in the next proposition.

121H Proposition Let X be a set and Σ a σ-algebra of subsets of X; let f, g and f_n, for $n \in \mathbb{N}$, be Σ-measurable real-valued functions whose domains belong to Σ. Then all the functions

$$f+g, \quad f\times g, \quad f/g,$$

$$\sup_{n\in\mathbb{N}} f_n, \quad \inf_{n\in\mathbb{N}} f_n, \quad \lim_{n\to\infty} f_n, \quad \limsup_{n\to\infty} f_n, \quad \liminf_{n\to\infty} f_n$$

have domains belonging to Σ. Moreover, if h is a Borel measurable real-valued function defined on a Borel subset of $\mathbb{R}$, then $\operatorname{dom} hf \in \Sigma$.

proof For the first two, we have $\operatorname{dom}(f+g)=\operatorname{dom}(f\times g)=\operatorname{dom} f\cap\operatorname{dom} g$. Next, if E is a Borel subset of $\mathbb{R}$, there is an $H\in\Sigma$ such that $f^{-1}[E]=H\cap\operatorname{dom} f$; so $f^{-1}[E]\in\Sigma$. Thus

$$\operatorname{dom} hf=f^{-1}[\operatorname{dom} h]\in\Sigma.$$

Setting $h(a)=1/a$ for $a\in\mathbb{R}\setminus\{0\}$, we see that $\operatorname{dom}(1/f)\in\Sigma$. ($\operatorname{dom} h=\mathbb{R}\setminus\{0\}$ is a Borel set because it is open.) Similarly, $\operatorname{dom}(1/g)$ and $\operatorname{dom}(f/g)=\operatorname{dom} f\cap\operatorname{dom}(1/g)$ belong to Σ.

Now for the infinite combinations. Set $H_n(a)=\{x:x\in\operatorname{dom} f_n,\ f_n(x)<a\}$ for $n\in\mathbb{N}$, $a\in\mathbb{R}$; as just explained, every $H_n(a)$ belongs to Σ. Now

$$\operatorname{dom}(\sup_{n\in\mathbb{N}} f_n)=\bigcup_{m\in\mathbb{N}}\bigcap_{n\in\mathbb{N}} H_n(m)\in\Sigma.$$

Next, $|f_m-f_n|$ is measurable, with domain in Σ, for all m, $n\in\mathbb{N}$ (applying the results above to $-f_n=-1\cdot f_n$, $f_m-f_n=f_m+(-f_n)$ and $|f_m-f_n|=|\ |\circ(f_m-f_n)$), so

$$G_{mnk}=\{x:x\in\operatorname{dom} f_m\cap\operatorname{dom} f_n,\ |f_m(x)-f_n(x)|\le 2^{-k}\}\in\Sigma$$

for all m, n, $k\in\mathbb{N}$. Accordingly

$$\operatorname{dom}(\lim_{n\to\infty} f_n)=\{x:\exists\ n,\ \langle f_m(x)\rangle_{m\ge n}\text{ is Cauchy}\}=\bigcap_{k\in\mathbb{N}}\bigcup_{n\in\mathbb{N}}\bigcap_{m\ge n} G_{mnk}\in\Sigma.$$

Manipulating the above results as in (c), (d) and (e) of the proof of 121F, we easily complete the proof.

Remark Note the use of the General Principle of Convergence in the proof above. I am not sure whether this will strike you as 'natural', and there are alternative methods; but the formula

$$\{x:\lim_{n\to\infty} f_n(x)\text{ exists in }\mathbb{R}\}=\{x:\langle f_n(x)\rangle_{n\in\mathbb{N}}\text{ is Cauchy}\}$$

is one worth storing in your long-term memory.

***121I** I end this section with two results which can be safely passed by on first reading, but which you will need at some point to master if you wish to go farther into measure theory than the present chapter, as both are essential parts of the concept of 'measurable function'.

Proposition Let X be a set and Σ a σ-algebra of subsets of X. Let D be a subset of X and $f:D\to\mathbb{R}$ a function. Then f is measurable iff there is a measurable function $h:X\to\mathbb{R}$ extending f.

proof (a) If $h:X\to\mathbb{R}$ is measurable and $f=h\restriction D$, then f is measurable by 121Eh.

(b) Now suppose that f is measurable.

(i) For each $q\in\mathbb{Q}$, the set $D_q=\{x:x\in D,\ f(x)\le q\}$ belongs to the subspace σ-algebra Σ_D, that is, there is an $E_q\in\Sigma$ such that $D_q=E_q\cap D$. Set

$$F=X\setminus\bigcup_{q\in\mathbb{Q}} E_q,$$

$$G=\bigcap_{n\in\mathbb{N}}\bigcup_{q\in\mathbb{Q},q\le -n} E_q;$$

then both F and G belong to Σ, and are disjoint from D. **P** If $x\in D$, there is a $q\in\mathbb{Q}$ such that $f(x)\le q$, so that $x\in E_q$ and $x\notin F$. Also there is an $n\in\mathbb{N}$ such that $f(x)>-n$, so that $x\notin E_{q'}$ for $q'\le -n$ and $x\notin G$. **Q**

Set $H=X\setminus(F\cup G)\in\Sigma$. For $x\in H$,

$$\{q:q\in\mathbb{Q},\ x\in E_q\}$$

is non-empty and bounded below, so we may set

$$h(x)=\inf\{q:x\in E_q\};$$

for $x\in F\cup G$, set $h(x)=0$. This defines $h:X\to\mathbb{R}$.

(ii) $h(x)=f(x)$ for $x\in D$. **P** As remarked above, $x\in H$. If $q\in\mathbb{Q}$ and $x\in E_q$, then $f(x)\le q$; consequently $h(x)\ge f(x)$. On the other hand, given $\epsilon>0$, there is a $q\in\mathbb{Q}\cap[f(x),f(x)+\epsilon]$, and now $x\in E_q$, so $h(x)\le q\le f(x)+\epsilon$; as ϵ is arbitrary, $h(x)\le f(x)$. **Q**

(iii) h is measurable. **P** If $a>0$ then

$$\{x:h(x)<a\}=(H\cap\bigcup_{q<a} E_q)\cup(F\cup G)\in\Sigma,$$

while if $a\le 0$

$$\{x:h(x)<a\}=H\cap\bigcup_{q<a} E_q\in\Sigma.\ \mathbf{Q}$$

This completes the proof.

***121J** The next proposition may illuminate 121E, as well as being indispensable for the work of Volume 2. I start with a useful description of the Borel sets of $\mathbb{R}^r$.

Lemma Let $r \ge 1$ be an integer, and write $\mathcal{J}$ for the family of subsets of $\mathbb{R}^r$ of the form $\{x : \xi_i \le \alpha\}$ where $i \le r$, $\alpha \in \mathbb{R}$, writing $x = (\xi_1, \dots, \xi_r)$, as in §115. Then the σ-algebra of subsets of $\mathbb{R}^r$ generated by $\mathcal{J}$ is precisely the σ-algebra $\mathcal{B}$ of Borel subsets of $\mathbb{R}^r$.

proof (a) All the sets in $\mathcal{J}$ are closed, so must belong to $\mathcal{B}$; writing Σ for the σ-algebra generated by $\mathcal{J}$, we must have $\Sigma \subseteq \mathcal{B}$.

(b) The next step is to observe that all half-open intervals of the form

$$]a, b] = \{x : \alpha_i < \xi_i \le \beta_i \ \forall \ i \le r\}$$

belong to Σ; this is because

$$]a, b] = \bigcap_{i \le r}(\{x : \xi_i \le \beta_i\} \setminus \{x : \xi_i \le \alpha_i\}).$$

It follows that all open sets belong to Σ. **P** (Compare the proof of 121Ef.) Let $G \subseteq \mathbb{R}^r$ be an open set. Let $\mathcal{I}$ be the set of all intervals of the form $]q, q']$ which are included in G, where q, $q' \in \mathbb{Q}^r$. Then $\mathcal{I}$ is a countable subset of Σ, so (because Σ is a σ-algebra) $\bigcup \mathcal{I} \in \Sigma$. By the definition of $\mathcal{I}$, $\bigcup \mathcal{I} \subseteq G$. But also, if $x \in G$, there is a $\delta > 0$ such that the open ball $U(x, \delta)$ with centre x and radius δ is included in G (1A2A). Now, for each $i \le r$, we can find rational numbers α_i, β_i such that

$$\xi_i - \frac{\delta}{r} \le \alpha_i < \xi_i \le \beta_i < \xi_i + \frac{\delta}{r},$$

so that

$$x \in \,]a, b] \subseteq U(x, \delta) \subseteq G$$

and $x \in \,]a, b] \in \mathcal{I}$. Thus $x \in \bigcup \mathcal{I}$. As x is arbitrary, $G \subseteq \bigcup \mathcal{I}$ and $G = \bigcup \mathcal{I} \in \Sigma$. **Q**

(c) Thus Σ is a σ-algebra of sets containing every open set, and must include $\mathcal{B}$, the smallest such σ-algebra.

Remark Compare the proof of 115G.

***121K Proposition** Let X be a set and Σ a σ-algebra of subsets of X. Let $r \ge 1$ be an integer, and $f_1, \dots, f_r$ measurable functions defined on subsets of X. Set $D = \bigcap_{i \le r} \operatorname{dom} f_i$ and for $x \in D$ set $f(x) = (f_1(x), \dots, f_r(x)) \in \mathbb{R}^r$. Then

(a) for any Borel set $E \subseteq \mathbb{R}^r$, $f^{-1}[E]$ belongs to the subspace σ-algebra Σ_D;

(b) if h is a Borel measurable function from a subset $\operatorname{dom} h$ of $\mathbb{R}^r$ to $\mathbb{R}$, then the composition hf is measurable.

proof (a)(i) Consider the set

$$\mathrm{T} = \{E : E \subseteq \mathbb{R}^r, \ f^{-1}[E] \in \Sigma_D\}.$$

Then T is a σ-algebra of subsets of $\mathbb{R}^r$. **P** (Compare 121Ef.) **(α)** $f^{-1}[\emptyset] = \emptyset \in \Sigma_D$, so $\emptyset \in \mathrm{T}$. **(β)** If $E \in \mathrm{T}$, then $f^{-1}[\mathbb{R}^r \setminus E] = D \setminus f^{-1}[E] \in \Sigma_D$ so $\mathbb{R} \setminus E \in \mathrm{T}$. **($\gamma$)** If $\langle E_n \rangle_{n \in \mathbb{N}}$ is a sequence in T, then $f^{-1}[\bigcup_{n \in \mathbb{N}} E_n] = \bigcup_{n \in \mathbb{N}} f^{-1}[E_n] \in \Sigma_D$ because Σ_D is a σ-algebra, so $\bigcup_{n \in \mathbb{N}} E_n \in \mathrm{T}$. **Q**

(ii) Next, for any $i \le r$ and $\alpha \in \mathbb{R}$, $J = \{x : \xi_i \le \alpha\}$ belongs to T, because

$$f^{-1}[J] = \{x : x \in D, \ f_i(x) \le \alpha\} \in \Sigma_D.$$

So T includes the family $\mathcal{J}$ of 121J and therefore includes the σ-algebra $\mathcal{B}$ generated by $\mathcal{J}$, that is, contains every Borel subset of $\mathbb{R}^r$.

(b) Now the rest follows by the argument of 121Eg. If $a \in \mathbb{R}$, then $\{y : y \in \operatorname{dom} h, \ h(y) < a\}$ is of the form $E \cap \operatorname{dom} h$, where E is a Borel subset of $\mathbb{R}^r$, so $\{x : x \in \operatorname{dom}(hf), \ (hf)(x) < a\} = f^{-1}[E] \cap \operatorname{dom}(hf)$ belongs to $\Sigma_{\operatorname{dom} hf}$.

121X Basic exercises >(a) Let X be a set, Σ a σ-algebra of subsets of X, and $D \subseteq X$. Let $\langle D_n \rangle_{n \in \mathbb{N}}$ be a partition of D into relatively measurable sets and $\langle f_n \rangle_{n \in \mathbb{N}}$ a sequence of measurable real-valued functions such that $D_n \subseteq \operatorname{dom} f_n$ for each n. Define $f : D \to \mathbb{R}$ by setting $f(x) = f_n(x)$ whenever $n \in \mathbb{N}$, $x \in D_n$. Show that f is measurable.

(b) Let X be a set and Σ a σ-algebra of subsets of X. If f and g are measurable real-valued functions defined on subsets of X, show that f^+, f^-, $f \wedge g$ and $f \vee g$ are measurable, where

$$f^+(x) = \max(f(x), 0) \text{ for } x \in \operatorname{dom} f,$$

$$f^-(x) = \max(-f(x), 0) \text{ for } x \in \operatorname{dom} f,$$

$$(f \vee g)(x) = \max(f(x), g(x)) \text{ for } x \in \operatorname{dom} f \cap \operatorname{dom} g,$$

$$(f \wedge g)(x) = \min(f(x), g(x)) \text{ for } x \in \operatorname{dom} f \cap \operatorname{dom} g.$$

>(c) Let (X, Σ, μ) be a measure space. Write $\mathcal{L}^0$ for the set of real-valued functions f such that (α) $\operatorname{dom} f$ is a conegligible subset of X (β) there is a conegligible set $E \subseteq X$ such that $f \restriction E$ is measurable. (i) Show that the set E of clause (β) in the last sentence may be taken to belong to Σ and be included in $\operatorname{dom} f$. (ii) Show that if f, $g \in \mathcal{L}^0$ and $c \in \mathbb{R}$, then $f+g$, cf, $f \times g$, $|f|$, f^+, f^-, $f \wedge g$, $f \vee g$ all belong to $\mathcal{L}^0$. (iii) Show that if f, $g \in \mathcal{L}^0$ and $g \neq 0$ a.e. then $f/g \in \mathcal{L}^0$. (iv) Show that if $\langle f_n \rangle_{n \in \mathbb{N}}$ is a sequence in $\mathcal{L}^0$ then the functions

$$\lim_{n\to\infty} f_n, \quad \sup_{n\in\mathbb{N}} f_n, \quad \inf_{n\in\mathbb{N}} f_n, \quad \limsup_{n\to\infty} f_n, \quad \liminf_{n\to\infty} f_n$$

belong to $\mathcal{L}^0$ whenever they are defined almost everywhere as real-valued functions. (v) Show that if $f \in \mathcal{L}^0$ and $h : \mathbb{R} \to \mathbb{R}$ is Borel measurable then $hf \in \mathcal{L}^0$.

>(d) Consider the following four families of subsets of $\mathbb{R}$:

$$\mathcal{A}_1 = \{]-\infty, a[\, : a \in \mathbb{R}\}, \quad \mathcal{A}_2 = \{]-\infty, a] : a \in \mathbb{R}\},$$

$$\mathcal{A}_3 = \{]a, \infty[\, : a \in \mathbb{R}\}, \quad \mathcal{A}_4 = \{[a, \infty[\, : a \in \mathbb{R}\}.$$

Show that for each j the σ-algebra of subsets of $\mathbb{R}$ generated by $\mathcal{A}_j$ is the σ-algebra of Borel sets.

(e) Let D be any subset of $\mathbb{R}^r$, where $r \geq 1$. Write $\mathfrak{T}_D$ for the set $\{G \cap D : G \subseteq \mathbb{R}^r \text{ is open}\}$. (i) Show that $\mathfrak{T}_D$ satisfies the properties of open sets listed in 1A2B. (ii) Let $\mathcal{B}$ be the σ-algebra of Borel sets in $\mathbb{R}^r$, and $\mathcal{B}(D)$ the subspace σ-algebra on D. Show that $\mathcal{B}(D)$ is just the σ-algebra of subsets of D generated by $\mathfrak{T}_D$. (*Hint*: (α) observe that $\mathfrak{T}_D \subseteq \mathcal{B}(D)$ (β) consider $\{E : E \subseteq \mathbb{R}^r, \, E \cap D \text{ belongs to the } \sigma\text{-algebra generated by } \mathfrak{T}_D\}$.)

(f) Let (X, Σ, μ) be a measure space and define $\mathcal{L}^0$ as in 121Xc. Show that if $f_1, \dots, f_r$ belong to $\mathcal{L}^0$ and $h : \mathbb{R}^r \to \mathbb{R}$ is Borel measurable then $h(f_1, \dots, f_r)$ belongs to $\mathcal{L}^0$.

121Y Further exercises (a) Let X and Y be sets, Σ a σ-algebra of subsets of X, $\phi : X \to Y$ a function and g a real-valued function defined on a subset of Y. Set $\mathrm{T} = \{F : F \subseteq Y, \, \phi^{-1}[F] \in \Sigma\}$; then T is a σ-algebra of subsets of Y (see 111Xc). (i) Show that if g is T-measurable then $g\phi$ is Σ-measurable. (ii) Give an example in which $g\phi$ is Σ-measurable but g is not T-measurable. (iii) Show that if $g\phi$ is Σ-measurable and *either* ϕ is injective *or* $\operatorname{dom}(g\phi) \in \Sigma$ *or* $\phi[X] \subseteq \operatorname{dom} g$, then g is T-measurable.[1]

(b) Let X and Y be sets, T a σ-algebra of subsets of Y and $\phi : X \to Y$ a function. Set $\Sigma = \{\phi^{-1}[F] : F \in \mathrm{T}\}$, as in 111Xd. Show that a function $f : X \to \mathbb{R}$ is Σ-measurable iff there is a T-measurable function $g : Y \to \mathbb{R}$ such that $f = g\phi$.

(c) Let X and Y be sets and Σ, T σ-algebras of subsets of X, Y respectively. I say that a function $\phi : X \to Y$ is (Σ, T)**-measurable** if $\phi^{-1}[F] \in \Sigma$ for every $F \in \mathrm{T}$. (i) Show that if Σ, T, Υ are σ-algebras of subsets of X, Y, Z respectively, and $\phi : X \to Y$ is (Σ, T)-measurable, $\psi : Y \to Z$ is (T, Υ)-measurable, then $\psi\phi : X \to Z$ is (Σ, Υ)-measurable. (ii) Suppose that $\mathcal{A} \subseteq \mathrm{T}$ is such that T is the σ-algebra of subsets of Y generated by $\mathcal{A}$ (111Gb). Show that $\phi : X \to Y$ is (Σ, T)-measurable iff $\phi^{-1}[A] \in \Sigma$ for every $A \in \mathcal{A}$. (iii) For $r \geq 1$, write $\mathcal{B}_r$ for the σ-algebra of Borel subsets of $\mathbb{R}^r$. Show that if X is any set and Σ is a σ-algebra of subsets of X, then a function $f : X \to \mathbb{R}^r$ is $(\Sigma, \mathcal{B}_r)$-measurable iff $\pi_i f : X \to \mathbb{R}$ is $(\Sigma, \mathcal{B}_1)$-measurable for every $i \leq r$, writing $\pi_i(x) = \xi_i$ for $i \leq r$, $x = (\xi_1, \dots, \xi_r) \in \mathbb{R}^r$. (iv) Rewrite these ideas for partially-defined functions.

(d) Let X be a set and Σ a σ-algebra of subsets of X. For $r \geq 1$, $D \subseteq X$ say that a function $\phi : D \to \mathbb{R}^r$ is **measurable** if $\phi^{-1}[G]$ is relatively measurable in D for every open set $G \subseteq \mathbb{R}^r$. If $X = \mathbb{R}^s$ and Σ is the σ-algebra $\mathcal{B}_s$ of Borel subsets of $\mathbb{R}^s$, say that ϕ is **Borel measurable**. (i) Show that ϕ is measurable in this sense iff all its coordinate functions $\phi_i : D \to \mathbb{R}$ are measurable in the sense of 121C, taking $\phi(x) = (\phi_1(x), \dots, \phi_r(x))$ for $x \in D$. (In particular, this definition agrees with 121C when $r = 1$.) (ii) Show that $\phi : D \to \mathbb{R}^r$ is measurable iff it is $(\Sigma, \mathcal{B}_r)$-measurable in

[1] I am grateful to P.Wallace Thompson for pointing out the error in the original version of this exercise.

the sense of 121Yc. (iii) Show that if $\phi : D \to \mathbb{R}^r$ is measurable and $\psi : E \to \mathbb{R}^s$ is Borel measurable, where $E \subseteq \mathbb{R}^r$, then $\psi\phi : \phi^{-1}[E] \to \mathbb{R}^s$ is measurable. (iv) Show that any continuous function from a subset of $\mathbb{R}^s$ to $\mathbb{R}^r$ is Borel measurable.

(e) Let X be a set and θ an outer measure on X; let μ be the measure defined from θ by Carathéodory's method, and Σ its domain. Suppose that $f : X \to \mathbb{R}$ is a function such that

$$\theta\{x : x \in A,\ f(x) \le a\} + \theta\{x : x \in A,\ f(x) \ge b\} \le \theta A$$

whenever $A \subseteq X$ and $a < b$ in $\mathbb{R}$. Show that f is Σ-measurable. (*Hint*: suppose that $a \in \mathbb{R}$ and $\theta A < \infty$. Set

$$B_k = \{x : x \in A,\ a + \tfrac{1}{2k+2} \le f(x) \le a + \tfrac{1}{2k+1}\},$$

$$B'_k = \{x : x \in A,\ a + \tfrac{1}{2k+3} \le f(x) \le a + \tfrac{1}{2k+2}\}$$

for $k \in \mathbb{N}$. Show that $\sum_{k=0}^{\infty} \theta B_k \le \theta A$, and check a similar result for B'_k. Hence show that

$$\theta\{x : x \in A,\ f(x) > a\} = \lim_{k\to\infty} \theta\{x : x \in A,\ f(x) \ge a + \tfrac{1}{k}\}.)$$

121 Notes and comments I find myself offering no fewer than three definitions of 'measurable function', in 121C, 121Yc and 121Yd. It is in fact the last which is probably the most important and the best guide to further ideas. Nevertheless, the overwhelming majority of applications refer to real-valued functions, and the four equivalent conditions of 121B are the most natural and most convenient to use. The fact that they all coincide with the condition of 121Yd corresponds to the fact that they are all of the form

$$f^{-1}[E] \in \Sigma_D \text{ for every } E \in \mathcal{A}$$

where $\mathcal{A}$ is a family of subsets of $\mathbb{R}$ generating the Borel σ-algebra (121Xd, 121Yc(ii)).

The class of measurable functions may well be the widest you have yet seen, not counting the family of all real-valued functions; all easily describable functions between subsets of $\mathbb{R}$ are measurable. Not only is the space of measurable functions closed under addition and multiplication and composition with continuous functions (121E), but any natural operation acting on a sequence of measurable functions will produce a measurable function (121F, 121Xb, 121Xa). It is *not* however the case that the composition of two Lebesgue measurable functions from $\mathbb{R}$ to itself is always Lebesgue measurable; I offer a counter-example in 134Ib. The point here is that a function is called 'measurable' if it is $(\Sigma, \mathcal{B})$-measurable, in the language of 121Yc, where $\mathcal{B}$ is the σ-algebra of Borel sets. Such a function can well fail to be (Σ, Σ)-measurable, if Σ properly includes $\mathcal{B}$, so the natural argument for compositions (121Yc(i)) fails. Nevertheless, for reasons which I will hint at in §134, non-Lebesgue-measurable functions are hard to come by, and only in the most rarefied kinds of real analysis do they appear in any natural way. You may therefore approach the question of whether a particular function is Lebesgue measurable with reasonable confidence that it is, and that the proof is merely a challenge to your technique.

You will observe that the results of 121E are mostly covered by 121I-121K, which also include large parts of 114G and 115G; and that 121Kb is covered by 121Yd(iii). You can count yourself as having mastered this part of the subject when you find my exposition tediously repetitive. Of course, in order to deduce 121Ed from 121K, for instance, you have to know that multiplication, regarded as a function from $\mathbb{R}^2$ to $\mathbb{R}$, is continuous, therefore Borel measurable; the proof of this is embedded in the proof I give of 121Ed (look at the formula for η half way through).

122 Definition of the integral

I set out the definition of ordinary integration for real-valued functions defined on an arbitrary measure space, with its most basic properties.

122A Definitions Let (X, Σ, μ) be a measure space.

(a) For any set $A \subseteq X$, I write χA for the **characteristic function** of A, the function from X to $\{0, 1\}$ given by setting $\chi A(x) = 1$ if $x \in A$, 0 if $x \in X \setminus A$. (Of course this notation depends on it being understood which is the 'universal' set X under consideration; perhaps I should call it the 'characteristic function of A as a subset of X'.) Observe that χA is Σ-measurable, in the sense of 121C above, iff $A \in \Sigma$ (because $A = \{x : \chi A(x) > 0\}$).

(b) Now a **simple** function on X is a function of the form $\sum_{i=0}^{n} a_i \chi E_i$, where $E_0, \dots, E_n$ are measurable sets of finite measure and $a_0, \dots, a_n$ belong to $\mathbb{R}$. **Warning!** Some authors allow arbitrary sets E_i, so that a simple function on X is any function taking only finitely many values.

122B Lemma Let (X, Σ, μ) be a measure space.
(a) Every simple function on X is measurable.
(b) If f, $g : X \to \mathbb{R}$ are simple functions, so is $f + g$.
(c) If $f : X \to \mathbb{R}$ is a simple function and $c \in \mathbb{R}$, then $cf : X \to \mathbb{R}$ is a simple function.
(d) The constant zero function is simple.

proof (a) comes from the facts that χE is measurable for measurable E, and that sums and scalar multiples of measurable functions are measurable (121Eb-121Ec). (b)-(d) are trivial.

122C Lemma Let (X, Σ, μ) be a measure space.
(a) If $E_0, \dots, E_n$ are measurable sets of finite measure, there are disjoint measurable sets $G_0, \dots, G_m$ of finite measure such that each E_i is expressible as a union of some of the G_j.
(b) If $f : X \to \mathbb{R}$ is a simple function, it is expressible in the form $\sum_{j=0}^{m} b_j \chi G_j$ where $G_0, \dots, G_m$ are disjoint measurable sets of finite measure.
(c) If $E_0, \dots, E_n$ are measurable sets of finite measure, and $a_0, \dots, a_n \in \mathbb{R}$ are such that $\sum_{i=0}^{n} a_i \chi E_i(x) \geq 0$ for every $x \in X$, then $\sum_{i=0}^{n} a_i \mu E_i \geq 0$.

proof (a) Set $m = 2^{n+1} - 2$, and enumerate the non-empty subsets of $\{0, \dots, n\}$ as $I_0, \dots, I_m$. For each $j \leq m$, set

$$G_j = \bigcap_{i \in I_j} E_i \setminus \bigcup_{i \leq n, i \notin I_j} E_i.$$

Then every G_j is a measurable set, being obtained from finitely many measurable sets by the operations $\cup$, $\cap$ and $\setminus$, and has finite measure, because $I_j \neq \emptyset$ and $G_j \subseteq E_i$ if $i \in I_j$. Moreover, the G_j are disjoint, for if $i \in I_j \setminus I_k$ then $G_j \subseteq E_i$ and $G_k \cap E_i = \emptyset$. Finally, if $k \leq n$ and $x \in E_k$, there is a $j \leq m$ such that $I_j = \{i : i \leq n, x \in E_i\}$, and in this case $x \in G_j \subseteq E_k$; thus E_k is the union of those G_j which it includes.

(b) Express f as $\sum_{i=0}^{n} a_i \chi E_i$ where $E_0, \dots, E_n$ are measurable sets of finite measure and $a_0, \dots, a_n$ are real numbers. Let $G_0, \dots, G_m$ be disjoint measurable sets of finite measure such that every E_i is expressible as a union of appropriate G_j. Set $c_{ij} = 1$ if $G_j \subseteq E_i$, 0 otherwise, so that, because the G_j are disjoint, $\chi E_i = \sum_{j=0}^{m} c_{ij} \chi G_j$ for each i. Then

$$f = \sum_{i=0}^{n} a_i \chi E_i = \sum_{i=0}^{n} \sum_{j=0}^{m} a_i c_{ij} \chi G_j = \sum_{j=0}^{m} b_j \chi G_j,$$

setting $b_j = \sum_{i=0}^{n} a_i c_{ij}$ for each $j \leq m$.

(c) Set $f = \sum_{i=0}^{n} a_i \chi E_i$, and take G_j, c_{ij}, b_j as in (b). Then $b_j \mu G_j \geq 0$ for every j. **P** If $G_j = \emptyset$, this is trivial. Otherwise, let $x \in G_j$; then

$$0 \leq f(x) = \sum_{i=0}^{n} b_i \chi G_i(x) = b_j \chi G_j(x) = b_j,$$

so again $b_j \mu G_j \geq 0$. **Q** Next, because the G_j are disjoint,

$$\mu E_i = \sum_{j=0}^{m} c_{ij} \mu G_j$$

for each i, so

$$\sum_{i=0}^{n} a_i \mu E_i = \sum_{i=0}^{n} \sum_{j=0}^{m} a_i c_{ij} \mu G_j = \sum_{j=0}^{m} b_j \mu G_j \geq 0,$$

as required.

122D Corollary Let (X, Σ, μ) be a measure space. If

$$\sum_{i=0}^{m} a_i \chi E_i = \sum_{j=0}^{n} b_j \chi F_j,$$

where all the E_i and F_j are measurable sets of finite measure and the a_i, b_j are real numbers, then

$$\sum_{i=0}^{m} a_i \mu E_i = \sum_{j=0}^{n} b_j \mu F_j.$$

proof Apply 122Cc to $\sum_{i=0}^{m} a_i \chi E_i + \sum_{j=0}^{n} (-b_j) \chi F_j$ to see that $\sum_{i=0}^{m} a_i \mu E_i - \sum_{j=0}^{n} b_j \mu F_j \geq 0$; now reverse the roles of the two sums to get the opposite inequality.

122E Definition Let (X, Σ, μ) be a measure space. Then we may define the **integral** $\int f$ of f, for simple functions $f : X \to \mathbb{R}$, by saying that $\int f = \sum_{i=0}^{m} a_i \mu E_i$ whenever $f = \sum_{i=0}^{m} a_i \chi E_i$ and every E_i is a measurable set of finite measure; 122D promises us that it won't matter which representation of f we pick on for the calculation.

122F Proposition Let (X, Σ, μ) be a measure space.

(a) If f, $g : X \to \mathbb{R}$ are simple functions, then $f + g$ is a simple function and $\int f + g = \int f + \int g$.

(b) If f is a simple function and $c \in \mathbb{R}$, then cf is a simple function and $\int cf = c \int f$.

(c) If f, g are simple functions and $f(x) \le g(x)$ for every $x \in X$, then $\int f \le \int g$.

proof (a) and (b) are immediate from the formula given for $\int f$ in 122E. As for (c), observe that $g - f$ is a non-negative simple function, so that $\int g - f \ge 0$, by 122Cc; but this means that $\int g - \int f \ge 0$.

122G Lemma Let (X, Σ, μ) be a measure space. If $\langle f_n \rangle_{n\in\mathbb{N}}$ is a sequence of simple functions which is non-decreasing (in the sense that $f_n(x) \le f_{n+1}(x)$ for every $n \in \mathbb{N}$, $x \in X$) and f is a simple function such that $f(x) \le \sup_{n\in\mathbb{N}} f_n(x)$ for almost every $x \in X$ (allowing $\sup_{n\in\mathbb{N}} f_n(x) = \infty$ in this formula), then $\int f \le \sup_{n\in\mathbb{N}} \int f_n$.

proof Note that $f - f_0$ is a simple function, so $H = \{x : (f - f_0)(x) \ne 0\}$ is a finite union of sets of finite measure, and $\mu H < \infty$; also $f - f_0$ is bounded, so there is an $M \ge 0$ such that $(f - f_0)(x) \le M$ for every $x \in X$.

Let $\epsilon > 0$. For each $n \in \mathbb{N}$, set $H_n = \{x : (f - f_n)(x) \ge \epsilon\}$. Then each H_n is measurable (by 121E), and $\langle H_n \rangle_{n\in\mathbb{N}}$ is a non-increasing sequence of sets with intersection

$$\bigcap_{n\in\mathbb{N}} H_n = \{x : f(x) \ge \epsilon + \sup_{n\in\mathbb{N}} f_n(x)\} \subseteq \{x : f(x) > \sup_{n\in\mathbb{N}} f_n(x)\}.$$

Because $f(x) \le \sup_{n\in\mathbb{N}} f_n(x)$ for almost every x, $\{x : f(x) > \sup_{n\in\mathbb{N}} f_n(x)\}$ and $\bigcap_{n\in\mathbb{N}} H_n$ are negligible. Also $\mu H_0 < \infty$, because $H_0 \subseteq H$. Consequently

$$\lim_{n\to\infty} \mu H_n = \mu(\bigcap_{n\in\mathbb{N}} H_n) = 0$$

(112Cf). Let n be so large that $\mu H_n \le \epsilon$.

Consider the simple function $g = f_n + \epsilon\chi H + M\chi H_n$. Then $f \le g$, so

$$\int f \le \int g = \int f_n + \epsilon\mu H + M\mu H_n \le \int f_n + \epsilon(M + \mu H).$$

As ϵ is arbitrary, $\int f \le \sup_{n\in\mathbb{N}} \int f_n$.

122H Definition Let (X, Σ, μ) be a measure space. For the rest of this section, I will write U for the set of functions f such that

(i) the domain of f is a conegligible subset of X and $f(x) \in [0, \infty[$ for each $x \in \operatorname{dom} f$,

(ii) there is a non-decreasing sequence $\langle f_n \rangle_{n\in\mathbb{N}}$ of non-negative simple functions such that $\sup_{n\in\mathbb{N}} \int f_n < \infty$ and $\lim_{n\to\infty} f_n(x) = f(x)$ for almost every $x \in X$.

122I Lemma If f and $\langle f_n \rangle_{n\in\mathbb{N}}$ are as in 122H, then

$$\sup_{n\in\mathbb{N}} \int f_n = \sup\{\int g : g \text{ is a simple function and } g \le_{\text{a.e.}} f\}.$$

proof Of course

$$\sup_{n\in\mathbb{N}} \int f_n \le \sup\{\int g : g \text{ is a simple function and } g \le_{\text{a.e.}} f\}$$

because $f_n \le_{\text{a.e.}} f$ for each n. On the other hand, if g is a simple function and $g \le_{\text{a.e.}} f$, then $g(x) \le \sup_{n\in\mathbb{N}} f_n(x)$ for almost every x, so $\int g \le \sup_{n\in\mathbb{N}} \int f_n$ by 122G. Thus

$$\sup_{n\in\mathbb{N}} \int f_n \ge \sup\{\int g : g \text{ is a simple function and } g \le_{\text{a.e.}} f\},$$

as required.

122J Lemma Let (X, Σ, μ) be a measure space, and define U as in 122H.

(a) If f is a function defined on a conegligible subset of X and taking values in $[0, \infty[$, then $f \in U$ iff there is a conegligible measurable set $E \subseteq \operatorname{dom} f$ such that

(α) $f{\restriction}E$ is measurable,

(β) for every $\epsilon > 0$, $\mu\{x : x \in E, f(x) \ge \epsilon\} < \infty$,

(γ) $\sup\{\int g : g \text{ is a simple function}, g \le_{\text{a.e.}} f\} < \infty$.

(b) Suppose that $f \in U$ and that h is a function defined on a conegligible subset of X and taking values in $[0, \infty[$. Suppose that $h \le_{\text{a.e.}} f$ and there is a conegligible $F \subseteq X$ such that $h{\restriction}F$ is measurable. Then $h \in U$.

proof (a)(i) Suppose that $f \in U$. Then there is an non-decreasing sequence $\langle f_n \rangle_{n\in\mathbb{N}}$ of non-negative simple functions such that $f =_{\text{a.e.}} \lim_{n\to\infty} f_n$ and $\sup_{n\in\mathbb{N}} \int f_n = c < \infty$. The set $\{x : f(x) = \lim_{n\to\infty} f_n(x)\}$ is conegligible, so includes a measurable conegligible set E say. Now $f\restriction E = (\lim_{n\to\infty} f_n)\restriction E$ is measurable, by 121Fa and 121Eh; thus (α) is satisfied. Next, given $\epsilon > 0$, set $H_n = \{x : x \in E,\ f_n(x) \ge \frac{1}{2}\epsilon\}$; then $f_n \ge \frac{1}{2}\epsilon\chi H_n$, so

$$\frac{1}{2}\epsilon\mu H_n = \int \tfrac{1}{2}\epsilon\chi H_n \le \int f_n \le c,$$

for each n. Now $\langle H_n\rangle_{n\in\mathbb{N}}$ is non-decreasing, so

$$\mu(\bigcup_{n\in\mathbb{N}} H_n) = \sup_{n\in\mathbb{N}} \mu H_n \le 2c/\epsilon,$$

by 112Ce. Accordingly

$$\mu\{x : x \in E,\ f(x) \ge \epsilon\} \le \mu(\bigcup_{n\in\mathbb{N}} H_n) \le 2c/\epsilon < \infty.$$

As ϵ is arbitrary, (β) is satisfied. Finally, (γ) is satisfied by 122I.

(ii) Now suppose that the conditions (α)-(γ) are satisfied. Take an appropriate conegligible $E \in \Sigma$, and for each $n \in \mathbb{N}$ define $f_n : X \to \mathbb{R}$ by setting

$$\begin{aligned} f_n(x) &= 2^{-n}k \text{ if } x \in E,\ 0 \le k < 4^n,\ 2^{-n}k \le f(x) < 2^{-n}(k+1), \\ &= 0 \text{ if } x \in X \setminus E, \\ &= 2^n \text{ if } x \in E \text{ and } f(x) \ge 2^n. \end{aligned}$$

Then f_n is a non-negative simple function, being expressible as

$$f_n = \sum_{k=1}^{4^n} 2^{-n}\chi\{x : x \in E,\ f(x) \ge 2^{-n}k\};$$

all the sets $\{x : x \in E,\ f(x) \ge 2^{-n}k\}$ being measurable (because $f\restriction E$ is measurable) and of finite measure, by (β). Also it is easy to see that $\langle f_n\rangle_{n\in\mathbb{N}}$ is an non-decreasing sequence which converges to f at every point of E, so that $f =_{\text{a.e.}} \lim_{n\to\infty} f_n$. Finally,

$$\lim_{n\to\infty} \int f_n = \sup_{n\in\mathbb{N}} \int f_n \le \sup\{\int g : g \le f \text{ is simple}\} < \infty,$$

by (γ). So $f \in U$.

(b) Let E be a set as in (a). The sets E, F and $\{x : h(x) \le f(x)\}$ are all conegligible, so there is a conegligible measurable set E' included in their intersection. Now $E' \subseteq \operatorname{dom} h$, $h\restriction E'$ is measurable,

$$\mu\{x : x \in E',\ h(x) \ge \epsilon\} \le \mu\{x : x \in E,\ f(x) \ge \epsilon\} < \infty$$

for every $\epsilon > 0$, and

$$\sup\{\int g : g \text{ is simple, } g \le_{\text{a.e.}} h\} \le \sup\{\int g : g \text{ is simple, } g \le_{\text{a.e.}} f\} < \infty.$$

So $h \in U$.

122K Definition Let (X, Σ, μ) be a measure space, and define U as in 122H. For $f \in U$, set

$$\int f = \sup\{\int g : g \text{ is a simple function and } g \le_{\text{a.e.}} f\}.$$

By 122I, we see that $\int f = \lim_{n\to\infty} \int f_n$ whenever $\langle f_n\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence of simple functions converging to f almost everywhere; in particular, if $f \in U$ is itself a simple function, then $\int f$, as defined here, agrees with the original definition of $\int f$ in 122E, since we may take $f_n = f$ for every n.

122L Lemma Let (X, Σ, μ) be a measure space.

(a) If f, $g \in U$ then $f + g \in U$ and $\int f + g = \int f + \int g$.

(b) If $f \in U$ and $c \ge 0$ then $cf \in U$ and $\int cf = c\int f$.

(c) If f, $g \in U$ and $f \le_{\text{a.e.}} g$ then $\int f \le \int g$.

(d) If $f \in U$ and g is a function with domain a conegligible subset of X, taking values in $[0, \infty[$, and equal to f almost everywhere, then $g \in U$ and $\int g = \int f$.

(e) If f_1, g_1, f_2, $g_2 \in U$ and $f_1 - f_2 = g_1 - g_2$, then $\int f_1 - \int f_2 = \int g_1 - \int g_2$.

proof (a) We know that there are non-decreasing sequences $\langle f_n\rangle_{n\in\mathbb{N}}$, $\langle g_n\rangle_{n\in\mathbb{N}}$ of non-negative simple functions such that $f =_{\text{a.e.}} \lim_{n\to\infty} f_n$, $g =_{\text{a.e.}} \lim_{n\to\infty} g_n$, $\sup_{n\in\mathbb{N}} \int f_n < \infty$ and $\sup_{n\in\mathbb{N}} \int g_n < \infty$. Now $\langle f_n + g_n\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence of simple functions converging to $f + g$ a.e., and

$$\sup_{n\in\mathbb{N}} \int f_n + g_n = \lim_{n\to\infty} \int f_n + g_n = \lim_{n\to\infty} \int f_n + \lim_{n\to\infty} \int g_n = \int f + \int g.$$

Accordingly $f+g \in U$, and also, as remarked in 122K,

$$\int f+g = \lim_{n\to\infty} \int f_n + g_n = \int f + \int g.$$

(b) We know that there is a non-decreasing sequence $\langle f_n\rangle_{n\in\mathbb{N}}$ of non-negative simple functions such that $f =_{\text{a.e.}} \lim_{n\to\infty} f_n$ and $\sup_{n\in\mathbb{N}} \int f_n < \infty$. Now $\langle cf_n\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence of simple functions converging to cf a.e., and

$$\sup_{n\in\mathbb{N}} \int cf_n = \lim_{n\to\infty} \int cf_n = c\lim_{n\to\infty} \int f_n = c\int f.$$

Accordingly $cf \in U$, and also, as remarked in 122K,

$$\int cf = \lim_{n\to\infty} \int cf_n = c\int f.$$

(c) This is obvious from 122K.

(d) If $\langle f_n\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence of simple functions converging to f a.e., then it also converges to g a.e.; so $g \in U$ and

$$\int g = \lim_{n\to\infty} \int f_n = \int f.$$

(e) By (a), f_1+g_2 and f_2+g_1 both belong to U. Also, they are equal at any point at which all four functions are defined, which is almost everywhere. So

$$\int f_1 + \int g_2 = \int f_1 + g_2 = \int f_2 + g_1 = \int f_2 + \int g_1,$$

using (a) and (d). Shifting $\int g_2$ and $\int f_2$ across the equation, we have the result.

122M Definition Let (X, Σ, μ) be a measure space. Define U as in 122H. A real-valued function f is **integrable**, or **integrable over** X, or **μ-integrable over** X, if it is expressible as $f_1 - f_2$ with f_1, $f_2 \in U$, and in this case its **integral** is

$$\int f = \int f_1 - \int f_2.$$

122N Remarks (a) We see from 122Le that the integral $\int f$ is uniquely defined by the formula in 122M. Secondly, if $f \in U$, then $f = f - \mathbf{0}$ is integrable, and the integral here agrees with that defined in 122K. Finally, if f is a simple function, then it can be expressed as $f_1 - f_2$ where f_1, f_2 are non-negative simple functions (if $f = \sum_{i=0}^n a_i\chi E_i$, where each E_i is measurable and of finite measure, set

$$f_1 = \sum_{i=0}^n a_i^+\chi E_i, \quad f_2 = \sum_{i=0}^n a_i^-\chi E_i,$$

writing $a_i^+ = \max(a_i, 0)$, $a_i^- = \max(-a_i, 0)$); so that

$$\int f = \int f_1 - \int f_2 = \sum_{i=0}^n a_i\mu E_i,$$

and the definition of 122M is consistent with the definition of 122E.

(b) Alternative notations which I will use for $\int f$ are $\int_X f$, $\int f d\mu$, $\int f(x)\mu(dx)$, $\int f(x)dx$, $\int_X f(x)\mu(dx)$, etc., according to which aspects of the context seem due for emphasis.

When μ is Lebesgue measure on $\mathbb{R}$ or $\mathbb{R}^r$ we say that $\int f$ is the **Lebesgue integral** of f, and that f is **Lebesgue integrable** if this is defined.

(c) Note that when I say, in 122M, that 'f can be expressed as $f_1 - f_2$', I mean to interpret $f_1 - f_2$ according to the rules set out in 121E, so that $\operatorname{dom} f$ must be $\operatorname{dom}(f_1 - f_2) = \operatorname{dom} f_1 \cap \operatorname{dom} f_2$, and is surely conegligible.

122O Theorem Let (X, Σ, μ) be a measure space.

(a) If f and g are integrable over X then $f+g$ is integrable and $\int f+g = \int f + \int g$.

(b) If f is integrable over X and $c \in \mathbb{R}$ then cf is integrable and $\int cf = c\int f$.

(c) If f is integrable over X and $f \geq 0$ a.e. then $\int f \geq 0$.

(d) If f and g are integrable over X and $f \leq_{\text{a.e.}} g$ then $\int f \leq \int g$.

proof (a) Express f as $f_1 - f_2$ and g as $g_1 - g_2$ where f_1, f_2, g_1 and g_2 belong to U, as defined in 122H. Then $f + g = (f_1 + g_1) - (f_2 + g_2)$ is integrable because U is closed under addition (122La), and

$$\int f + g = \int f_1 + g_1 - \int f_2 + g_2 = \int f_1 + \int g_1 - \int f_2 - \int g_2 = \int f + \int g.$$

(b) Express f as $f_1 - f_2$ where f_1, f_2 belong to U. If $c \ge 0$ then $cf = cf_1 - cf_2$ is integrable because U is closed under multiplication by non-negative scalars (122Lb), and

$$\int cf = \int cf_1 - \int cf_2 = c\int f_1 - c\int f_2 = c\int f.$$

If $c = -1$ then $-f = f_2 - f_1$ is integrable and

$$\int(-f) = \int f_2 - \int f_1 = -\int f.$$

Putting these together we get the result for $c < 0$.

(c) Express f as $f_1 - f_2$ where f_1, $f_2 \in U$. Then $f_2 \le_{\text{a.e.}} f_1$, so $\int f_2 \le \int f_1$ (122Lc), and $\int f \ge 0$.

(d) Apply (c) to $g - f$.

122P Theorem Let (X, Σ, μ) be a measure space and f a real-valued function defined on a conegligible subset of X. Then the following are equiveridical:

(i) f is integrable;

(ii) $|f| \in U$, as defined in 122H, and there is a conegligible set $E \subseteq X$ such that $f\restriction E$ is measurable;

(iii) there are a $g \in U$ and a conegligible set $E \subseteq X$ such that $|f| \le_{\text{a.e.}} g$ and $f\restriction E$ is measurable.

proof (i)$\Rightarrow$(iii) Suppose that f is integrable. Let f_1, $f_2 \in U$ be such that $f = f_1 - f_2$. Then there are conegligible sets E_1, E_2 such that $f_1\restriction E_1$ and $f_2\restriction E_2$ are measurable; set $E = E_1 \cap E_2$, so that E also is a conegligible set. Now $f\restriction E = f_1\restriction E - f_2\restriction E$ is measurable. Next, $f_1 + f_2 \in U$ (122La) and $|f|(x) \le f_1(x) + f_2(x)$ for every $x \in \operatorname{dom} f$, so we may take $g = f_1 + f_2$.

(iii)$\Rightarrow$(ii) If $f\restriction E$ is measurable, so is $|f|\restriction E = |f\restriction E|$ (121Eg); so if $g \in U$ and $|f| \le_{\text{a.e.}} g$, then $|f| \in U$ by 122Jb.

(ii)$\Rightarrow$(i) Suppose that f satisfies the conditions of (ii). Set $f^+ = \frac{1}{2}(|f| + f)$ and $f^- = \frac{1}{2}(|f| - f)$. Of course $|f|\restriction E$, $f^+\restriction E$ and $f^-\restriction E$ are all measurable. Also $0 \le f^+(x) \le |f|(x)$ and $0 \le f^-(x) \le |f|(x)$ for every $x \in \operatorname{dom} f$, while $|f| \in U$ by hypothesis, so f^+ and f^- belong to U by 122Jb. Finally, $f = f^+ - f^-$, so f is integrable.

122Q Remark The condition 'there is a conegligible set E such that $f\restriction E$ is measurable' recurs so often that I think it worth having a phrase for it; I will call such functions **virtually measurable**, or **μ-virtually measurable** if it seems necessary to specify the measure.

122R Corollary Let (X, Σ, μ) be a measure space.

(a) A non-negative real-valued function, defined on a subset of X, is integrable iff it belongs to U, as defined in 122H.

(b) If f is integrable over X and h is a real-valued function, defined on a conegligible subset of X and equal to f almost everywhere, then h is integrable, with $\int h = \int f$.

(c) If f is integrable over X, $f \ge 0$ a.e. and $\int f \le 0$, then $f = 0$ a.e.

(d) If f and g are integrable over X, $f \le_{\text{a.e.}} g$ and $\int g \le \int f$, then $f =_{\text{a.e.}} g$.

(e) If f is integrable over X, so is $|f|$, and $|\int f| \le \int |f|$.

proof (a) If f is integrable then $f = |f| \in U$, by 122P(ii). If $f \in U$ then $f = f - \mathbf{0}$ is integrable, by 122M.

(b) Let E, F be conegligible sets such that $f\restriction E$ is measurable and $h\restriction F = f\restriction F$; then $E \cap F$ is conegligible and $h\restriction(E \cap F) = (f\restriction E)\restriction F$ is measurable. Next, there is a $g \in U$ such that $|f| \le_{\text{a.e.}} g$, and of course $|h| \le_{\text{a.e.}} g$. So h is integrable by 122P(iii). By 122Od, applied to f and h and then to h and f, $\int h = \int f$.

(c) ? Suppose, if possible, otherwise. Let $E \subseteq X$ be a conegligible set such that $f\restriction E$ is measurable (122P(ii)), and $E' \subseteq E \cap \operatorname{dom} f$ a conegligible measurable set. Then $F = \{x : x \in E', f(x) > 0\}$ must be non-negligible. Set $F_k = \{x : x \in E', f(x) \ge 2^{-k}\}$ for each $k \in \mathbb{N}$, so that $F = \bigcup_{k\in\mathbb{N}} F_k$ and there is a k such that $\mu F_k > 0$. But consider $g = 2^{-k}\chi F_k$. Because $f \ge 0$ a.e. and $f \ge 2^{-k}$ on F_k, $f \ge_{\text{a.e.}} g$, so that

$$0 < 2^{-k}\mu F_k = \int g \le \int f,$$

by 122Od; which is impossible. **X**

(d) Apply (c) to $g - f$.

(e) By (i)$\Rightarrow$(ii) of 122P, $|f|$ is integrable. Now $f^+ = \frac{1}{2}(|f| + f)$ and $f^- = \frac{1}{2}(|f| - f)$ are both integrable and non-negative, so have non-negative integrals, and

$$|\int f| = |\int f^+ - \int f^-| \le \int f^+ + \int f^- = \int |f|.$$

122X Basic exercises (a) Let (X, Σ, μ) be a measure space. (i) Show that if $f : X \to \mathbb{R}$ is simple so is $|f|$, setting $|f|(x) = |f(x)|$ for $x \in \operatorname{dom} f = X$. (ii) Show that if f, $g : X \to \mathbb{R}$ are simple functions so are $f \vee g$ and $f \wedge g$, as defined in 121Xb.

>(b) Let (X, Σ, μ) be a measure space and f a real-valued function which is integrable over X. Show that for every $\epsilon > 0$ there is a simple function $g : X \to \mathbb{R}$ such that $\int |f - g| \le \epsilon$. (*Hint*: consider non-negative f first.)

(c) Let (X, Σ, μ) be a measure space, and write $\mathcal{L}^1$ for the set of all real-valued functions which are integrable over X. Let $\Phi \subseteq \mathcal{L}^1$ be such that

(i) $\chi E \in \Phi$ whenever $E \in \Sigma$ and $\mu E < \infty$;

(ii) $f + g \in \Phi$ for all f, $g \in \Phi$;

(iii) $cf \in \Phi$ whenever $c \in \mathbb{R}$, $f \in \Phi$;

(iv) $f \in \Phi$ whenever $f \in \mathcal{L}^1$ is such that there is a non-decreasing sequence $\langle f_n \rangle_{n\in\mathbb{N}}$ in Φ for which $\lim_{n\to\infty} f_n = f$ almost everywhere.

Show that $\Phi = \mathcal{L}^1$.

>(d) Let μ be counting measure on $\mathbb{N}$ (112Bd). Show that a function $f : \mathbb{N} \to \mathbb{R}$ (that is, a sequence $\langle f(n) \rangle_{n\in\mathbb{N}}$) is μ-integrable iff it is absolutely summable, and in this case

$$\int f d\mu = \int_{\mathbb{N}} f(n)\mu(dn) = \sum_{n=0}^{\infty} f(n).$$

>(e) Let (X, Σ, μ) be a measure space and f, g two virtually measurable real-valued functions defined on subsets of X. (i) Show that $f + g$, $f \times g$ and f/g, defined as in 121E, are all virtually measurable. (ii) Show that if h is a Borel measurable real-valued function defined on any subset of $\mathbb{R}$, then the composition hf is virtually measurable.

>(f) Let (X, Σ, μ) be a measure space and $\langle f_n \rangle_{n\in\mathbb{N}}$ a sequence of virtually measurable real-valued functions defined on subsets of X. Show that $\lim_{n\to\infty} f_n$, $\sup_{n\in\mathbb{N}} f_n$, $\inf_{n\in\mathbb{N}} f_n$, $\limsup_{n\to\infty} f_n$ and $\liminf_{n\to\infty} f_n$, defined as in 121F, are virtually measurable.

>(g) Let (X, Σ, μ) be a measure space and f, g real-valued functions which are integrable over X. Show that $f \wedge g$ and $f \vee g$, as defined in 121Xb, are integrable.

>(h) Let (X, Σ, μ) be a measure space, f a real-valued function which is integrable over X, and g a bounded real-valued virtually measurable function defined on a conegligible subset of X. Show that $f \times g$, defined as in 121Ed, is integrable.

(i) Let X be a set, Σ a σ-algebra of subsets of X, and μ_1, μ_2 two measures with domain Σ. Set $\mu E = \mu_1 E + \mu_2 E$ for $E \in \Sigma$. Show that for any real-valued function f defined on a subset of X, $\int f d\mu = \int f d\mu_1 + \int f d\mu_2$ in the sense that if one side is defined as a real number so is the other, and they are then equal. (*Hint*: (α) Check that a subset of X is μ-conegligible iff it is μ_i-conegligible for both i. (β) Check the result for simple functions f. (γ) Now consider general non-negative f.)

122Y Further exercises (a) Let (X, Σ, μ) be a 'complete' measure space, that is, one in which all negligible sets are measurable (see, for instance, 113Xa). Show that if f is a virtually measurable real-valued function defined on a subset of X, then f is measurable. Use this fact to find appropriate simplifications of 122J and 122P for such spaces (X, Σ, μ).

(b) Write $\mathcal{L}^1$ for the set of all Lebesgue integrable real-valued functions on $\mathbb{R}$. Let $\Phi \subseteq \mathcal{L}^1$ be such that

(i) $\chi I \in \Phi$ whenever I is a bounded half-open interval in $\mathbb{R}$;

(ii) $f + g \in \Phi$ for all f, $g \in \Phi$;

(iii) $cf \in \Phi$ whenever $c \in \mathbb{R}$, $f \in \Phi$;

(iv) $f \in \Phi$ whenever $f \in \mathcal{L}^1$ is such that there is a non-decreasing sequence $\langle f_n \rangle_{n\in\mathbb{N}}$ in Φ for which $\lim_{n\to\infty} f_n = f$ almost everywhere.

Show that $\Phi = \mathcal{L}^1$. (*Hint*: show that (α) $\chi E \in \Phi$ whenever E is expressible as the union of finitely many half-open intervals (β) $\chi E \in \Phi$ whenever E has finite measure and is expressible as the union of a sequence of half-open intervals (γ) $\chi E \in \Phi$ whenever E is measurable and has finite measure.)

(c) Let X be any set, and let μ be counting measure on X. Let $f : X \to \mathbb{R}$ be a function; set $f^+(x) = \max(0, f(x))$, $f^-(x) = \max(0, -f(x))$ for $x \in X$. Show that the following are equiveridical: (i) $\int f d\mu$ exists in $\mathbb{R}$, and is equal to s; (ii) for every $\epsilon > 0$ there is a finite $K \subseteq X$ such that $|s - \sum_{i\in I} f(i)| \le \epsilon$ whenever $I \subseteq X$ is a finite set including K (iii) $\sum_{x\in X} f^+(x)$ and $\sum_{x\in X} f^-(x)$, defined as in 112Bd, are finite, and $s = \sum_{x\in X} f^+(x) - \sum_{x\in X} f^-(x)$.

(d) Let (X,Σ,μ) be a measure space. Let us say that a function $g:X\to\mathbb{R}$ is **quasi-simple** if it is expressible as $\sum_{i=0}^{\infty}a_i\chi G_i$, where $\langle G_i\rangle_{i\in\mathbb{N}}$ is a partition of X into measurable sets, $\langle a_i\rangle_{i\in\mathbb{N}}$ is a sequence in $\mathbb{R}$, and $\sum_{i=0}^{\infty}|a_i|\mu G_i<\infty$, counting $0\cdot\infty$ as 0, so that there can be G_i of infinite measure provided that the corresponding a_i are zero.

(i) Show that if g and h are quasi-simple functions so are $g+h$, $|g|$ and cg, for any $c\in\mathbb{R}$. (*Hint*: for $g+h$ you will need 111F(b-ii) or its equivalent.)

(ii) Show from first principles (I mean, without using anything later than 122F in this chapter) that if $g=\sum_{i=0}^{\infty}a_i\chi G_i$ and $h=\sum_{i=0}^{\infty}b_i\chi H_i$ are quasi-simple functions, and $g\le_{\text{a.e.}}h$, then $\sum_{i=0}^{\infty}a_i\mu G_i\le\sum_{i=0}^{\infty}b_i\mu H_i$.

(iii) Hence show that we have a functional I defined by saying that $I(g)=\sum_{i=0}^{\infty}a_i\mu G_i$ whenever g is a quasi-simple function represented as $\sum_{i=0}^{\infty}a_i\chi G_i$.

(iv) Show that if g and h are quasi-simple functions and $c\in\mathbb{R}$, then $I(g+h)=I(g)+I(h)$ and $I(cg)=cI(g)$, and that $I(g)\le I(h)$ if $g\le_{\text{a.e.}}h$.

(v) Show that if g is a quasi-simple function then g is integrable and $\int g=I(g)$. (I do now allow you to use 122G-122R.)

(vi) Show that a real-valued function f, defined on a conegligible subset of X, is integrable iff for every $\epsilon>0$ there are quasi-simple functions g, h such that $g\le_{\text{a.e.}}f\le_{\text{a.e.}}h$ and $I(h)-I(g)\le\epsilon$.

(e) Let μ be Lebesgue measure on $\mathbb{R}$. Let us say (for this exercise only) that a real-valued function g with $\operatorname{dom}g\subseteq\mathbb{R}$ is 'pseudo-simple' if it is expressible as $\sum_{i=0}^{\infty}a_i\chi J_i$, where $\langle J_i\rangle_{i\in\mathbb{N}}$ is a sequence of bounded half-open intervals (*not* necessarily disjoint) and $\sum_{i=0}^{\infty}|a_i|\mu J_i<\infty$. (Interpret the infinite sum $\sum_{i=0}^{\infty}a_i\chi J_i$ as in 121F, so that

$$\operatorname{dom}(\textstyle\sum_{i=0}^{\infty}a_i\chi J_i)=\{x:\lim_{n\to\infty}\sum_{i=0}^{n}a_i(\chi J_i)(x)\text{ exists in }\mathbb{R}\}.)$$

(i) Show that if g, h are pseudo-simple functions so are $g+h$ and cg, for any $c\in\mathbb{R}$.

(ii) Show that if g is a pseudo-simple function then g is integrable.

(iii) Show that a real-valued function f, defined on a conegligible subset of $\mathbb{R}$, is integrable iff for every $\epsilon>0$ there are pseudo-simple functions g, h such that $g\le_{\text{a.e.}}f\le_{\text{a.e.}}h$ and $\int h-g\,d\mu\le\epsilon$. (*Hint*: Take Φ to be the set of integrable functions with this property, and show that it satisfies the conditions of 122Yb.)

(f) Repeat 122Yb and 122Ye for Lebesgue measure on $\mathbb{R}^r$, where $r>1$.

(g) Let (X,Σ,μ) be a measure space, and assume that there is at least one partition of X into infinitely many non-empty measurable sets. Let f be a real-valued function defined on a conegligible subset of X, and $a\in\mathbb{R}$. Show that the following are equiveridical:

(i) f is integrable, with $\int f=a$;

(ii) for every $\epsilon>0$ there is a partition $\langle E_n\rangle_{n\in\mathbb{N}}$ of X into non-empty measurable sets such that

$$\textstyle\sum_{n=0}^{\infty}|f(t_n)|\mu E_n<\infty,\qquad|a-\sum_{n=0}^{\infty}f(t_n)\mu E_n|\le\epsilon$$

whenever $\langle t_n\rangle_{n\in\mathbb{N}}$ is a sequence such that $t_n\in E_n\cap\operatorname{dom}f$ for every n. (As usual, take $0\cdot\infty=0$ in these formulae.) (*Hint*: use 122Yd.)

(h) Find a re-formulation of (g) which covers the case of measure spaces which can *not* be partitioned into sequences of non-empty measurable sets.

(i) Let X be a set, Σ a σ-algebra of subsets of X, and $\langle\mu_n\rangle_{n\in\mathbb{N}}$ a sequence of measures with domain Σ. Set $\mu E=\sum_{n=0}^{\infty}\mu_nE$ for $E\in\Sigma$. (i) Show that μ is a measure. (ii) Show that for any real-valued function f defined on a subset of X, f is μ-integrable iff it is μ_n-integrable for every n and $\sum_{n=0}^{\infty}\int|f|d\mu_n$ is finite, and that then $\int fd\mu=\sum_{n=0}^{\infty}\int fd\mu_n$.

(j) Let X be a set, Σ a σ-algebra of subsets of X, and $\langle\mu_i\rangle_{i\in I}$ a family of measures with domain Σ. Set $\mu E=\sum_{i\in I}\mu_iE$ for $E\in\Sigma$. (i) Show that μ is a measure. (ii) Show that for any Σ-measurable function $f:X\to\mathbb{R}$, f is μ-integrable iff it is μ_i-integrable for every i and $\sum_{i\in I}\int|f|d\mu_i$ is finite.

122 Notes and comments Just as in §121, some extra technical problems are caused by my insistence on trying to integrate (i) functions which are not defined on the whole of the measure space under consideration (ii) functions which are not, strictly speaking, measurable, but are only measurable on some conegligible set. There is nothing in the present section to justify either of these elaborations. In the next section, however, we shall be looking at the limits of sequences of functions, and these limits need not be defined at every point; and the examples in which the limits are not everywhere defined are not in any sense pathological, but are central to the most important applications of the theory.

The question of integrating not-quite-measurable functions is more disputable. I will discuss this point further after formally introducing 'complete' measure spaces in Chapter 21. For the moment, I will say only that I think it is worth taking the trouble to have a formalisation which integrates as many functions as is reasonably possible; the original point of the Lebesgue integral being, in part, that it enables us to integrate more functions than its predecessors.

The definition of 'integration' here proceeds in three distinguishable stages: (i) integration of simple functions (122A-122G); (ii) integration of non-negative functions (122H-122L); (iii) integration of general real-valued functions (122M-122R). I have taken each stage slowly, passing to non-negative integrable functions only when I have a full set of the requisite lemmas on simple functions, for instance. There are, of course, innumerable alternative routes; see, for instance, 122Yd, which offers a definition using two steps rather than three. I prefer the longer, gentler climb partly because (to my eye) it gives a clearer view of the ideas and partly because it corresponds to an almost canonical method of proving properties of integrable functions: we prove them first for simple functions, then for non-negative integrable functions, then for general integrable functions. (The hint I give for 122Yb conforms to this philosophy. See also 122Xc; but I do not give this as a formally expressed theorem, because the exact order of proof varies from case to case, and I think it is best remembered as a method of attack rather than as a specific result to apply.)

You have a right to feel that this section has been singularly abstract, and gives very little idea of which of your favourite functions are likely to be integrable, let alone what the integrals are. I hope that Chapter 13 will provide some help in this direction, though I have to say that for really useful methods for calculating integrals we must wait for Chapters 22, 25 and 26 in the next volume. If you want to know the true centre of the arguments of this section, I would myself locate it in 122G, 122H and 122K. The point is that the ideas of 122A-122F apply to a much wider class of structures (X,Σ,μ), because they involve only operations on finitely many members of Σ at a time; there is no mention of sequences of sets. The key that makes all the rest possible is 122G, which is founded on 112Cf. And after 122G-122K, the rest of the section, although by no means elementary, really is no more than a careful series of checks to ensure that the functional defined in 122K behaves as we expect it to.

Many of the results of this section (including the key one, 122G) will be superseded by stronger results in the following section. But I should remark on Lemma 122Ja, which will periodically recur as a most useful criterion for integrability of non-negative functions (see 122Ra).

There is another point about the standard integral as defined here. It is an 'absolute' integral, meaning that if f is integrable so is $|f|$ (122P). This means that although the Lebesgue integral extends the 'proper' Riemann integral (see 134K below), there are functions with finite 'improper' Riemann integrals which are not Lebesgue integrable; a typical example is $f(x)=\frac{\sin x}{x}$, where $\lim_{a\to\infty}\int_0^a f$ exists in $\mathbb{R}$, while $\lim_{a\to\infty}\int_0^a |f|=\infty$, so that f is not integrable, in the sense defined here, over the whole interval $]0,\infty[$. (For full proofs of these assertions, see 283D and 282Xm in Volume 2.) If you have encountered the theory of 'absolutely' and 'conditionally' summable series, you will be aware that the latter can exhibit very confusing behaviour, and will appreciate that restricting the notion of 'integrable' to mean 'absolutely integrable' is a great convenience.

Indeed, it is more than just a convenience; it is necessary to make the definition work at the level of abstraction used in this chapter. Consider the example of counting measure μ on $\mathbb{N}$ (112Bd, 122Xd). The structure $(\mathbb{N},\mathcal{P}\mathbb{N},\mu)$ is invariant under permutations; that is, $\mu(\pi[A])=\mu A$ for every $A\subseteq\mathbb{N}$ and every permutation $\pi:\mathbb{N}\to\mathbb{N}$. Consequently, *any* definition of integration which depends only on the structure $(\mathbb{N},\mathcal{P}\mathbb{N},\mu)$ must also be invariant under permutations, that is,

$$\int f(\pi(n))\mu(dn)=\int f(n)\mu(dn)$$

for any integrable function f and any permutation π. But of course (as I hope you have been told) a series $\langle f(n)\rangle_{n\in\mathbb{N}}$ such that $\sum_{n=0}^{\infty}f(\pi(n))=\sum_{n=0}^{\infty}f(n)\in\mathbb{R}$ for any permutation π must be absolutely summable. Thus if we are to define an integral on an abstract measure space (X,Σ,μ) in terms depending only on Σ and μ, we are nearly inevitably forced to define an absolute integral.

Naturally there are important contexts in which this restriction is an embarrassment, and in which some kind of 'improper' integral seems appropriate. A typical one is the theory of Fourier transforms, in which we find ourselves looking at $\lim_{a\to\infty}\int_{-a}^{a}f$ in place of $\int_{-\infty}^{\infty}f$ (see §283). A vast number of more or less abstract forms of improper integral have been proposed; many are interesting and some are important; but none rivals the 'standard' integral as described in this chapter. (For an attempt at a systematic examination of a particular class of such improper integrals, see Chapter 48 in Volume 4.)

Much less work has been done on the integration of non-measurable functions – to speak more exactly, of functions which are not equal almost everywhere to a measurable integrable function. I am sure that this is simply because there are too few important problems to show us which way to turn. In 134C below I mention the question of whether there is *any* non-measurable real-valued function on $\mathbb{R}$. The standard answer is 'yes', but no such function can possibly arise as a result of any ordinary construction. Consequently the majority of questions concerning non-measurable functions appear in very special contexts, and so far I have seen none which gives any useful hint of what a generally appropriate extension of the notion of 'integrability' might be.

123 The convergence theorems

The great labour we have gone through so far has not yet been justified by any theorems powerful enough to make it worth while. We come now to the heart of the modern theory of integration, the 'convergence theorems', describing conditions under which we can integrate the limit of a sequence of integrable functions.

123A B.Levi's theorem Let (X,Σ,μ) be a measure space and $\langle f_n\rangle_{n\in\mathbb{N}}$ a sequence of real-valued functions, all integrable over X, such that (i) $f_n \le_{\text{a.e.}} f_{n+1}$ for every $n\in\mathbb{N}$ (ii) $\sup_{n\in\mathbb{N}}\int f_n < \infty$. Then $f = \lim_{n\to\infty} f_n$ is integrable, and $\int f = \lim_{n\to\infty}\int f_n$.

Remarks I ought to repeat at once the conventions I am following here. Each of the functions f_n is taken to be defined on a conegligible set $\operatorname{dom} f_n \subseteq X$, as in 122Nc, and the limit function f is taken to have domain

$$\{x : x\in\bigcup_{n\in\mathbb{N}}\bigcap_{m\ge n}\operatorname{dom} f_m,\ \lim_{n\to\infty} f_n(x) \text{ is defined in } \mathbb{R}\},$$

as in 121Fa. You would miss no important idea if you supposed that every f_n was defined everywhere on X; but the statement 'f is integrable' includes the assertion 'f is defined, as a real number, almost everywhere', and this is an essential part of the theorem.

proof (a) Let us first deal with the case in which $f_0 = 0$ a.e. Write $c = \sup_{n\in\mathbb{N}}\int f_n = \lim_{n\to\infty}\int f_n$ (noting that, by 122Od, $\langle\int f_n\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence).

(i) All the sets $\operatorname{dom} f_n$, $\{x : f_0(x) = 0\}$, $\{x : f_n(x) \le f_{n+1}(x)\}$ are conegligible, so their intersection F also is. For each $n\in\mathbb{N}$ there is a conegligible set E_n such that $f_n\restriction E_n$ is measurable (122P); let E^* be a measurable conegligible set included in the conegligible set $F\cap\bigcap_{n\in\mathbb{N}} E_n$.

(ii) For $a > 0$ and $n\in\mathbb{N}$ set $H_n(a) = \{x : x\in E^*,\ f_n(x) \ge a\}$; then $H_n(a)$ is measurable because $f_n\restriction E_n$ is measurable and E^* is a measurable subset of E_n. Also $a\chi H_n(a) \le f_n$ everywhere on E^*, so

$$a\mu H_n(a) = \int a\chi H_n(a) \le \int f_n \le c.$$

Because $f_n(x) \le f_{n+1}(x)$ for every $x\in E^*$, $H_n(a)\subseteq H_{n+1}(a)$ for every $n\in\mathbb{N}$, and writing $H(a) = \bigcup_{n\in\mathbb{N}} H_n(a)$, we have

$$\mu H(a) = \lim_{n\to\infty}\mu H_n(a) \le \frac{c}{a}$$

(112Ce). In particular, $\mu H(a) < \infty$ for every a. Furthermore,

$$\mu(\bigcap_{k\ge 1} H(k)) \le \inf_{k\ge 1}\mu H(k) \le \inf_{k\ge 1}\frac{c}{k} = 0.$$

Set $E = E^*\setminus\bigcap_{k\ge 1} H(k)$; then E is conegligible.

(iii) If $x\in E$, there is some k such that $x\notin H(k)$, that is, $x\notin\bigcup_{n\in\mathbb{N}} H_n(k)$, that is, $f_n(x) < k$ for every n; moreover, $\langle f_n(x)\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence, so $f(x) = \lim_{n\to\infty} f_n(x) = \sup_{n\in\mathbb{N}} f_n(x)$ is defined in $\mathbb{R}$. Thus the limit function f is defined almost everywhere. Because every $f_n\restriction E$ is measurable (121Eh), so is $f\restriction E = \lim_{n\to\infty} f_n\restriction E$ (121Fa). If $\epsilon > 0$ then $\{x : x\in E,\ f(x)\ge\epsilon\}$ is included in $H(\frac{1}{2}\epsilon)$, so has finite measure.

(iv) Now suppose that g is a simple function and that $g \le_{\text{a.e.}} f$. As in the proof of 122G, $H = \{x : g(x)\ne 0\}$ has finite measure, and g is bounded above by M say.

Let $\epsilon > 0$. For each $n\in\mathbb{N}$ set $G_n = \{x : x\in E,\ (g - f_n)(x)\ge\epsilon\}$. Then each G_n is measurable, and $\langle G_n\rangle_{n\in\mathbb{N}}$ is a non-increasing sequence with intersection

$$\{x : x\in E,\ g(x)\ge\epsilon + \sup_{n\in\mathbb{N}} f_n(x)\}\subseteq\{x : g(x) > f(x)\},$$

which is negligible. Also $\mu G_0 < \infty$ because $G_0\subseteq H$. Consequently $\lim_{n\to\infty}\mu G_n = 0$ (112Cf). Let n be such that $\mu G_n \le \epsilon$. Then, for any $x\in E$,

$$g(x) \le f_n(x) + \epsilon\chi H(x) + M\chi G_n(x),$$

so

$$g \le_{\text{a.e.}} f_n + M\chi G_n + \epsilon\chi H$$

and

$$\int g \le \int f_n + M\mu G_n + \epsilon\mu H \le c + \epsilon(M + \mu H).$$[2]

As ϵ is arbitrary, $\int g \le c$.

(v) Accordingly, $f{\restriction}E$ (which is non-negative) satisfies the conditions of Lemma 122Ja, and is integrable. Moreover, its integral is at most c, by Definition 122K. Because $f =_{\text{a.e.}} f{\restriction}E$, f also is integrable, with the same integral (122Rb). On the other hand, $f \ge_{\text{a.e.}} f_n$ for each n, so $\int f \ge \sup_{n\in\mathbb{N}} \int f_n = c$, by 122Od.

This completes the proof when $f_0 = 0$ a.e.

(b) For the general case, consider the sequence $\langle f'_n\rangle_{n\in\mathbb{N}} = \langle f_n - f_0\rangle_{n\in\mathbb{N}}$. By (a), $f' = \lim_{n\to\infty} f'_n$ is integrable, and $\int f' = \lim_{n\to\infty} \int f'_n$; now $\lim_{n\to\infty} f_n =_{\text{a.e.}} f' + f_0$, so is integrable, with integral $\int f' + \int f_0 = \lim_{n\to\infty} \int f_n$.

Remark You may have observed, without surprise, that the argument of (a-iv) in the proof here repeats that used for the special case 122G.

123B Fatou's Lemma Let (X, Σ, μ) be a measure space, and $\langle f_n\rangle_{n\in\mathbb{N}}$ a sequence of real-valued functions, all integrable over X. Suppose that every f_n is non-negative a.e., and that $\liminf_{n\to\infty} \int f_n < \infty$. Then $\liminf_{n\to\infty} f_n$ is integrable, and $\int \liminf_{n\to\infty} f_n \le \liminf_{n\to\infty} \int f_n$.

Remark Once again, this theorem includes the assertion that $\liminf_{n\to\infty} f_n(x)$ is defined in $\mathbb{R}$ for almost every $x \in X$.

proof Set $c = \liminf_{n\to\infty} \int f_n$ and $f = \liminf_{n\to\infty} f_n$. For each $n \in \mathbb{N}$, let E_n be a conegligible set such that $f'_n = f_n{\restriction}E_n$ is measurable and non-negative. Set $g_n = \inf_{m\ge n} f'_m$; then each g_n is measurable (121Fc), non-negative and defined on the conegligible set $\bigcap_{m\ge n} E_m$, and $g_n \le_{\text{a.e.}} f_n$, so g_n is integrable (122P) with $\int g_n \le \inf_{m\ge n} \int f_m \le c$. Next, $g_n(x) \le g_{n+1}(x)$ for every $x \in \operatorname{dom} g_n$, so $\langle g_n\rangle_{n\in\mathbb{N}}$ satisfies the conditions of B.Levi's theorem (123A), and $g = \lim_{n\to\infty} g_n$ is integrable, with $\int g = \lim_{n\to\infty} \int g_n \le c$. Finally, because every f'_n is equal to f_n almost everywhere, $g = \liminf_{n\to\infty} f'_n =_{\text{a.e.}} f$, and $\int f$ exists, equal to $\int g \le c$.

123C Lebesgue's Dominated Convergence Theorem Let (X, Σ, μ) be a measure space and $\langle f_n\rangle_{n\in\mathbb{N}}$ a sequence of real-valued functions, all integrable over X, such that $f(x) = \lim_{n\to\infty} f_n(x)$ exists in $\mathbb{R}$ for almost every $x \in X$. Suppose moreover that there is an integrable function g such that $|f_n| \le_{\text{a.e.}} g$ for every n. Then f is integrable, and $\lim_{n\to\infty} \int f_n$ exists and is equal to $\int f$.

proof Consider $\tilde f_n = f_n + g$ for each $n \in \mathbb{N}$. Then $0 \le \tilde f_n \le 2g$ a.e. for each n, so $\tilde c = \liminf_{n\to\infty} \int \tilde f_n$ exists in $\mathbb{R}$, and $\tilde f = \liminf_{n\to\infty} \tilde f_n$ is integrable, with $\int \tilde f \le \tilde c$, by Fatou's Lemma (123B). But observe that $f =_{\text{a.e.}} \tilde f - g$, since $f(x) = \tilde f(x) - g(x)$ at least whenever $f(x)$ and $g(x)$ are both defined, so f is integrable, with

$$\int f = \int \tilde f - \int g \le \liminf_{n\to\infty} \int \tilde f_n - \int g = \liminf_{n\to\infty} \int f_n.$$

Similarly, considering $\langle -f_n\rangle_{n\in\mathbb{N}}$,

$$\int(-f) \le \liminf_{n\to\infty} \int(-f_n),$$

that is,

$$\int f \ge \limsup_{n\to\infty} \int f_n.$$

So $\lim_{n\to\infty} \int f_n$ exists and is equal to $\int f$.

Remark We have at last reached the point where the technical problems associated with partially-defined functions are reducing, or rather, are being covered efficiently by the conventions I am using concerning the interpretation of such formulae as '$\limsup$'.

123D To try to show the power of these theorems, I give a result here which is one of the standard applications of the theory.

Corollary Let (X, Σ, μ) be a measure space and $]a, b[$ a non-empty open interval in $\mathbb{R}$. Let $f : X \times {]a, b[} \to \mathbb{R}$ be a function such that

(i) the integral $F(t) = \int f(x, t)dx$ is defined for every $t \in {]a, b[}$;

(ii) the partial derivative $\frac{\partial f}{\partial t}$ of f with respect to the second variable is defined everywhere in $X \times {]a, b[}$;

(iii) there is an integrable function $g : X \to [0, \infty[$ such that $|\frac{\partial f}{\partial t}(x, t)| \le g(x)$ for every $x \in X$ and $t \in {]a, b[}$.

[2] I am grateful to P.Wallace Thompson for noticing a fault at this stage in previous editions.

Then the derivative $F'(t)$ and the integral $\int \frac{\partial f}{\partial t}(x,t)dx$ exist for every $t \in \,]a,b[$, and are equal.

proof (a) Let t be any point of $]a,b[$, and $\langle t_n\rangle_{n\in\mathbb{N}}$ any sequence in $]a,b[\setminus\{t\}$ converging to t. Consider

$$\frac{F(t_n)-F(t)}{t_n-t} = \int \frac{f(x,t_n)-f(x,t)}{t_n-t}\mu(dx)$$

for each n. (This step uses 122O.) If we set

$$f_n(x) = \frac{f(x,t_n)-f(x,t)}{t_n-t},$$

for $x \in X$, then we see from the Mean Value Theorem that there is a τ (depending, of course, on both n and x), lying between t_n and t, such that $f_n(x) = \frac{\partial f}{\partial t}(x,\tau)$, so that $|f_n(x)| \le g(x)$. At the same time, $\lim_{n\to\infty} f_n(x) = \frac{\partial f}{\partial t}(x,t)$ for every x. So Lebesgue's Dominated Convergence Theorem tells us that $\int \frac{\partial f}{\partial t}(x,t)dx$ exists and is equal to

$$\lim_{n\to\infty}\int f_n(x)dx = \lim_{n\to\infty}\frac{F(t_n)-F(t)}{t_n-t}.$$

(b) Because $\langle t_n\rangle_{n\in\mathbb{N}}$ is arbitrary,

$$\lim_{s\to t}\frac{F(s)-F(t)}{s-t} = \int \frac{\partial f}{\partial t}(x,t)dx,$$

as claimed.

Remark In the next volume I offer a variation on this theorem, with both hypotheses and conclusion weakened (252Ye).

123X Basic exercises >(a) Let (X,Σ,μ) be a measure space, and $\langle f_n\rangle_{n\in\mathbb{N}}$ a sequence of real-valued functions, all integrable over X, such that $\sum_{n=0}^{\infty}\int|f_n|$ is finite. Show that $f(x) = \sum_{n=0}^{\infty} f_n(x)$ is defined in $\mathbb{R}$ for almost every $x \in X$, and that $\int f = \sum_{n=0}^{\infty}\int f_n$. (*Hint*: consider first the case in which every f_n is non-negative.)

(b) Let (X,Σ,μ) be a measure space. Suppose that T is any subset of $\mathbb{R}$, and $\langle f_t\rangle_{t\in T}$ a family of functions, all integrable over X, such that, for any $t \in T$,

$$f_t(x) = \lim_{s\in T, s\to t} f_s(x)$$

for almost every $x \in X$. Suppose moreover that there is an integrable function g such that $|f_t| \le_{\text{a.e.}} g$ for every $t \in T$. Show that $t \mapsto \int f_t : T \to \mathbb{R}$ is continuous.

>(c) Let f be a real-valued function defined everywhere on $[0,\infty[$, endowed with Lebesgue measure. Its (real) **Laplace transform** is the function F defined by

$$F(s) = \int_0^{\infty} e^{-sx} f(x)dx$$

for all those real numbers s for which the integral is defined.

(i) Show that if $s \in \operatorname{dom} F$ and $s' \ge s$ then $s' \in \operatorname{dom} F$ (because $e^{-s'x}e^{sx} \le 1$ for all x). (How do you know that $x \mapsto e^{-s'x}e^{sx}$ is measurable?)

(ii) Show that F is differentiable on the interior of its domain. (*Hint*: note that if $a_0 \in \operatorname{dom} F$ and $a_0 < a < b$ then there is some M such that $xe^{-sx}|f(x)| \le Me^{-a_0x}|f(x)|$ whenever $x \in [0,\infty[$, $s \in [a,b]$.)

(iii) Show that if F is defined anywhere then $\lim_{s\to\infty} F(s) = 0$. (*Hint*: use Lebesgue's Dominated Convergence Theorem to show that $\lim_{n\to\infty} F(s_n) = 0$ whenever $\lim_{n\to\infty} s_n = \infty$.)

(iv) Show that if f, g have Laplace transforms F, G then the Laplace transform of $f+g$ is $F+G$, at least on $\operatorname{dom} F \cap \operatorname{dom} G$.

(d) Let (X,Σ,μ) be a measure space and $\langle f_n\rangle_{n\in\mathbb{N}}$ a sequence of real-valued functions, all integrable over X, such that there is an integrable function g such that $|f_n| \le_{\text{a.e.}} g$ for every n. Show that $\limsup_{n\to\infty} f_n$ is integrable and that $\int \limsup_{n\to\infty} f_n \ge \limsup_{n\to\infty} f_n$.

123Y Further exercises (a) Let (X,Σ,μ) be a measure space, Y any set and $\phi : X \to Y$ any function; let $\mu\phi^{-1}$ be the image measure on Y (112Xf). Show that if $h : Y \to \mathbb{R}$ is any function, then h is $\mu\phi^{-1}$-integrable iff $h\phi$ is μ-integrable, and the integrals are then equal.

(b) Explain how to adapt 123Xc to the case in which f is undefined on a negligible subset of $\mathbb{R}$.

(c) Let (X,Σ,μ) be a measure space and $a<b$ in $\mathbb{R}$. Let $f:X\times\,]a,b[\,\to[0,\infty[$ be a function such that $\int f(x,t)dx$ is defined for every $t\in\,]a,b[$ and $t\mapsto f(x,t)$ is continuous for every $x\in X$. Suppose that $c\in\,]a,b[$ is such that $\liminf_{t\to c}\int f(x,t)dx<\infty$. Show that $\int\liminf_{t\to c}f(x,t)dx$ is defined and less than or equal to $\liminf_{t\to c}\int f(x,t)dx$.

(d) Show that there is a function $f:\mathbb{R}^2\to\{0,1\}$ such that (i) the Lebesgue integral $\int f(x,t)dx$ is defined and equal to 1 for every $t\neq 0$ (ii) the function $x\mapsto\liminf_{t\to 0}f(x,t)$ is not Lebesgue measurable. (*Remark*: you will of course have to start your construction from a non-measurable subset of $\mathbb{R}$; see 134B for such a set.)

(e) Let (Y,T,ν) be a measure space. Let X be a set, Σ a σ-algebra of subsets of X, and $\langle\mu_y\rangle_{y\in Y}$ a family of measures on X such that μ_yX is finite for every y and $\mu E=\int\mu_yE\,\nu(dy)$ is defined for every $E\in\Sigma$. (i) Show that $\mu:\Sigma\to[0,\infty[$ is a measure. (ii) Show that if $f:X\to[0,\infty[$ is a Σ-measurable function, then f is μ-integrable iff it is μ_y-integrable for almost every $y\in Y$ and $\int\bigl(\int fd\mu_y\bigr)\nu(dy)$ is defined, and that this is then $\int fd\mu$.

(f) Let (X,Σ,μ) be a measure space, and $\langle f_n\rangle_{n\in\mathbb{N}}$ a sequence of virtually measurable real-valued functions all defined almost everywhere in X. Suppose that $\sum_{n=0}^{\infty}\int|f_n(x)-1|\mu(dx)<\infty$. Show that $\prod_{n=0}^{\infty}f_n(x)$ is defined in $\mathbb{R}$ for almost every $x\in X$.

123 Notes and comments I hope that 123D and its special case 123Xc will help you to believe that the theory here has useful applications.

All the theorems of this section can be thought of as 'exchange of limit' theorems, setting out conditions under which

$$\lim_{n\to\infty}\int f_n=\int\lim_{n\to\infty}f_n,$$

or

$$\frac{\partial}{\partial t}\int f\,dx=\int\frac{\partial f}{\partial t}dx.$$

Even for functions which are accessible to much more primitive methods of integration (e.g., the Riemann integral), theorems of this type can involve laborious validation of inequalities. The power of Lebesgue's integral is that it gives general theorems which cover a reasonable proportion of the important cases which arise in practice. (I have to admit, however, that nothing is more typical of applied analysis than its need for special results which are related to, but not derivable from, the standard general theorems.) For instance, in 123Xc, the fact that the range of integration is the unbounded interval $[0,\infty[$ adds no difficulty. Of course this is connected with the fact that we consider only integrals of functions with integrable absolute values.

The limits used in 123A-123C are all limits of sequences; it is of course part of the essence of measure theory that we expect to be able to handle countable families of sets or functions, but that anything larger is alarming. Nevertheless, there are many contexts in which we can take other types of limit. I describe some in 123D, 123Xb and 123Xc(iii). The point is that in such limits as $\lim_{t\to u}\phi(t)$, where $u\in[-\infty,\infty]$, we shall have $\lim_{t\to u}\phi(t)=a$ iff $\lim_{n\to\infty}\phi(t_n)=a$ whenever $\langle t_n\rangle_{n\in\mathbb{N}}$ converges to u; so that when seeking a limit $\lim_{t\to u}\int f_t$, for some family $\langle f_t\rangle_{t\in T}$ of functions, it will be sufficient if we can find $\lim_{n\to\infty}\int f_{t_n}$ for enough sequences $\langle t_n\rangle_{n\in\mathbb{N}}$. This type of argument will be effective for any of the standard limits $\lim_{t\uparrow a}$, $\lim_{t\downarrow a}$, $\lim_{t\to a}$, $\lim_{t\to\infty}$, $\lim_{t\to-\infty}$ of basic calculus, and can be used in conjunction either with B.Levi's theorem or with Lebesgue's theorem. I should perhaps remark that a difficulty arises with a similar extension of Fatou's lemma (123Yc-123Yd).

Chapter 13

Complements

In this chapter I collect a number of results which do not lie in the direct line of the argument from 111A (the definition of 'σ-algebra') to 123C (Lebesgue's Dominated Convergence Theorem), but which nevertheless demand inclusion in this volume, being both relatively elementary, essential for further developments and necessary to a proper comprehension of what has already been done. The longest section is §134, dealing with a few of the elementary special properties of Lebesgue measure; in particular, its translation-invariance, the existence of non-measurable sets and functions, and the Cantor set. The other sections are comparatively lightweight. §131 discusses (measurable) subspaces and the interpretation of the formula $\int_E f$, generalizing the idea of an integral $\int_a^b f$ of a function over an interval. §132 introduces the outer measure associated with a measure, a kind of inverse of Carathéodory's construction of a measure from an outer measure. §§133 and 135 lay out suitable conventions for dealing with 'infinity' and complex numbers (separately! they don't mix well) as values either of integrands or of integrals; at the same time I mention 'upper' and 'lower' integrals. Finally, in §136, I give some theorems on σ-algebras of sets, describing how they can (in some of the most important cases) be generated by relatively restricted operations.

131 Measurable subspaces

Very commonly we wish to integrate a function over a subset of a measure space; for instance, to form an integral $\int_a^b f(x)dx$, where $a < b$ in $\mathbb{R}$. As often as not, we wish to do this when the function is partly or wholly undefined outside the subset, as in such expressions as $\int_0^1 \ln x\,dx$. The natural framework in which to perform such operations is that of 'subspace measures'. If (X, Σ, μ) is a measure space and $H \in \Sigma$, there is a natural subspace measure μ_H on H, which I describe in this section. I begin with the definition of this subspace measure (131A-131C), with a description of integration with respect to it (131E-131H); this gives a solid foundation for the concept of 'integration over a (measurable) subset' (131D).

131A Proposition Let (X, Σ, μ) be a measure space, and $H \in \Sigma$. Set $\Sigma_H = \{E : E \in \Sigma,\ E \subseteq H\}$ and let μ_H be the restriction of μ to Σ_H. Then (H, Σ_H, μ_H) is a measure space.

proof Of course Σ_H is just $\{E \cap H : E \in \Sigma\}$, and I have noted already (in 121A) that this is a σ-algebra of subsets of H. It is now obvious that μ_H satisfies (iii) of 112A, so that (H, Σ_H, μ_H) is a measure space.

131B Definition If (X, Σ, μ) is any measure space and $H \in \Sigma$, then μ_H, as defined in 131A, is the **subspace measure** on H.

When $X = \mathbb{R}^r$, where $r \geq 1$, and μ is Lebesgue measure on $\mathbb{R}^r$, I will call a subspace measure μ_H **Lebesgue measure on** H.

It is worth noting the following elementary facts.

131C Lemma Let (X, Σ, μ) be a measure space, $H \in \Sigma$, and μ_H the subspace measure on H, with domain Σ_H. Then

(a) for any $A \subseteq H$, A is μ_H-negligible iff it is μ-negligible;

(b) if $G \in \Sigma_H$ then $(\mu_H)_G$, the subspace measure on G when G is regarded as a measurable subset of H, is identical to μ_G, the subspace measure on G when G is regarded as a measurable subset of X.

131D Integration over subsets: Definition Let (X, Σ, μ) be a measure space, $H \in \Sigma$ and f a real-valued function defined on a subset of X. By $\int_H f$ (or $\int_H f(x)\mu(dx)$, etc.) I shall mean $\int (f{\restriction}H) d\mu_H$, if this exists, following the definitions of 131A-131B and 122M, and taking $\operatorname{dom}(f{\restriction}H) = H \cap \operatorname{dom} f$, $(f{\restriction}H)(x) = f(x)$ for $x \in H \cap \operatorname{dom} f$.

131E Proposition Let (X, Σ, μ) be a measure space, $H \in \Sigma$, and f a real-valued function defined on a subset $\operatorname{dom} f$ of H. Set $\tilde{f}(x) = f(x)$ if $x \in \operatorname{dom} f$, 0 if $x \in X \setminus H$. Then $\int f d\mu_H = \int \tilde{f} d\mu$ if either is defined in $\mathbb{R}$.

proof (a) If f is μ_H-simple, it is expressible as $\sum_{i=0}^n a_i \chi E_i$, where $E_0, \dots, E_n \in \Sigma_H$, $a_0, \dots, a_n \in \mathbb{R}$ and $\mu_H E_i < \infty$ for each i. Now $\tilde{f}$ also is equal to $\sum_{i=0}^n a_i \chi E_i$ if this is now interpreted as a function from X to $\mathbb{R}$. So

$$\sum_{i=0}^n a_i \mu_H E_i = \sum_{i=0}^n a_i \mu E_i = \int \tilde{f} d\mu.$$

(b) If f is a non-negative μ_H-integrable function, there is a non-decreasing sequence $\langle f_n\rangle_{n\in\mathbb{N}}$ of non-negative μ_H-simple functions converging to f μ_H-almost everywhere; now $\langle \tilde{f}_n\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence of μ-simple functions converging to $\tilde{f}$ μ-a.e. (131Ca), and

$$\sup_{n\in\mathbb{N}}\int \tilde{f}_n d\mu = \sup_{n\in\mathbb{N}}\int f_n d\mu_H = \int f d\mu_H < \infty,$$

so $\int \tilde{f} d\mu$ exists and is equal to $\int f d\mu_H$.

(c) If f is μ_H-integrable, it is expressible as $f_1 - f_2$ where f_1 and f_2 are non-negative μ_H-integrable functions, so that $\tilde{f} = \tilde{f}_1 - \tilde{f}_2$ and

$$\int \tilde{f} d\mu = \int \tilde{f}_1 d\mu - \int \tilde{f}_2 d\mu = \int f_1 d\mu_H - \int f_2 d\mu_H = \int f d\mu_H.$$

(d) Now suppose that $\tilde{f}$ is μ-integrable. In this case there is a μ-conegligible $E \in \Sigma$ such that $E \subseteq \operatorname{dom}\tilde{f}$ and $\tilde{f}\restriction E$ is Σ-measurable (122P). Of course $\mu(H\setminus E) = 0$ so $E\cap H$ is μ_H-conegligible; also, for any $a\in\mathbb{R}$,

$$\{x : x\in E\cap H,\, f(x)\ge a\} = H\cap\{x : x\in E,\, \tilde{f}(x)\ge a\}\in\Sigma_H,$$

so $f\restriction E\cap H$ is Σ_H-measurable, and f is μ_H-virtually measurable and defined μ_H-a.e. Next, for $\epsilon>0$,

$$\mu_H\{x : x\in E\cap H,\, |f(x)|\ge\epsilon\} = \mu\{x : x\in E,\, |\tilde{f}(x)|\ge\epsilon\} < \infty,$$

while if g is a μ_H-simple function and $g\le|f|$ μ_H-a.e. then $\tilde{g}\le|\tilde{f}|$ μ-a.e. and

$$\int g\, d\mu_H = \int \tilde{g}\, d\mu \le \int |\tilde{f}| d\mu < \infty.$$

By the criteria of 122J and 122P, f is μ_H-integrable, so that again we have $\int f d\mu_H = \int \tilde{f} d\mu$.

131F Corollary Let (X,Σ,μ) be a measure space and f a real-valued function defined on a subset $\operatorname{dom} f$ of X.

(a) If $H\in\Sigma$ and f is defined almost everywhere in X, then $f\restriction H$ is μ_H-integrable iff $f\times\chi H$ is μ-integrable, and in this case $\int_H f = \int f\times\chi H$.

(b) If f is μ-integrable, then $f\ge 0$ a.e. iff $\int_H f\ge 0$ for every $H\in\Sigma$.

(c) If f is μ-integrable, then $f = 0$ a.e. iff $\int_H f = 0$ for every $H\in\Sigma$.

proof (a) Because $\operatorname{dom} f$ is μ-conegligible, $(f\restriction H)^\sim$, as defined in 131E, is equal to $f\times\chi H$ μ-a.e., so that, by 131E,

$$\int_H f d\mu = \int (f\restriction H)^\sim d\mu = \int (f\times\chi H) d\mu$$

if any one of the integrals exists.

(b)(i) If $f\ge 0$ μ-a.e., then for any $H\in\Sigma$ we must have $f\restriction H\ge 0$ μ_H-a.e., so $\int_H f = \int (f\restriction H) d\mu_H \ge 0$.

(ii) If $\int_H f\ge 0$ for every $H\in\Sigma$, let $E\in\Sigma$ be a conegligible subset of $\operatorname{dom} f$ such that $f\restriction E$ is measurable. Set $F = \{x : x\in E,\, f(x)<0\}$. Then $\int_F f\ge 0$; by 122Rc, it follows that $f\restriction F = 0$ μ_F-a.e., which is possible only if $\mu F = 0$, in which case $f\ge 0$ μ-a.e.

(c) Apply (b) to f and to $-f$ to see that $f\le 0\le f$ a.e.

131G Corollary Let (X,Σ,μ) be a measure space and $H\in\Sigma$ a conegligible set. If f is any real-valued function defined on a subset of X, $\int_H f = \int f$ if either is defined.

proof In the language of 131E, $f = (f\restriction H)^\sim$ μ-almost everywhere, so that

$$\int f = \int (f\restriction H)^\sim = \int_H f$$

if any of the integrals is defined.

131H Corollary Let (X,Σ,μ) be a measure space and f, g two μ-integrable real-valued functions.

(a) If $\int_H f\ge\int_H g$ for every $H\in\Sigma$ then $f\ge g$ a.e.

(b) If $\int_H f = \int_H g$ for every $H\in\Sigma$ then $f = g$ a.e.

proof Apply 131Fb-131Fc to $f-g$.

131X Basic exercises >(a) Let (X,Σ,μ) be a measure space, and f a real-valued function which is integrable over X. For $E\in\Sigma$ set $\nu E = \int_E f$. (i) Show that if E, F are disjoint members of Σ then $\nu(E\cup F) = \nu E + \nu F$. (*Hint*: 131E.) (ii) Show that if $\langle E_n\rangle_{n\in\mathbb{N}}$ is a disjoint sequence in Σ then $\nu(\bigcup_{n\in\mathbb{N}} E_n) = \sum_{n=0}^\infty \nu E_n$. (*Hint*: 123C.) (iii) Show that if f is non-negative then (X,Σ,ν) is a measure space.

>(b) Let μ be Lebesgue measure on $\mathbb{R}$. (i) Show that whenever $a \le b$ in $\mathbb{R}$ and f is a real-valued function with $\operatorname{dom} f \subseteq \mathbb{R}$, then

$$f d\mu = \int_{[a,b[} f d\mu = \int_{]a,b]} f d\mu = \int_{[a,b]} f d\mu$$

if any of these is defined. (*Hint*: apply 131E to four different versions of $\tilde{f}$.) Write $\int_a^b f d\mu$ for the common value. (ii) Show that if $a \le b \le c$ in $\mathbb{R}$ then, for any real-valued function f, $\int_a^c f d\mu = \int_a^b f d\mu + \int_b^c f d\mu$ if either side is defined. (iii) Show that if f is integrable over $\mathbb{R}$, then $(a,b) \mapsto \int_a^b f d\mu$ is continuous. (*Hint*: *Either* consider simple functions f first *or* consider $\lim_{n\to\infty} \int_{a_n}^b f d\mu$ for monotonic sequences $\langle a_n\rangle_{n\in\mathbb{N}}$.)

(c) Let $g : \mathbb{R} \to \mathbb{R}$ be a non-decreasing function and μ_g the associated Lebesgue-Stieltjes measure (114Xa). (i) Show that if $a \le b \le c$ in $\mathbb{R}$ then, for any real-valued function f, $\int_{[a,c[} f d\mu_g = \int_{[a,b[} f d\mu_g + \int_{[b,c[} f d\mu_g$ if either side is defined. (ii) Show that if f is μ_g-integrable over $\mathbb{R}$, then $(a,b) \mapsto \int_{[a,b[} f d\mu_g$ is continuous on $\{(a,b) : a \le b,\ g \text{ is continuous at both } a \text{ and } b\}$.

131Y Further exercises (a) Let (X, Σ, μ) be a measure space and $E \in \Sigma$ a measurable set of finite measure. Let $\langle f_n\rangle_{n\in\mathbb{N}}$ be a sequence of measurable real-valued functions, with measurable domains[1], such that $f = \lim_{n\to\infty} f_n$ is defined almost everywhere in E (following the conventions of 121Fa). Show that for every $\epsilon > 0$ there is a measurable set $F \subseteq E$ such that $\mu(E \setminus F) \le \epsilon$ and $\langle f_n\rangle_{n\in\mathbb{N}}$ converges uniformly to f on F. (This is **Egorov's theorem**.)

131 Notes and comments If you want a quick definition of $\int_H f$ for measurable H, the simplest seems to be that of 131E, which enables you to avoid the concept of 'subspace measure' entirely. I think however that we really do need to be able to speak of 'Lebesgue measure on $[0,1]$', for instance, meaning the subspace measure $\mu_{[0,1]}$ where μ is Lebesgue measure on $\mathbb{R}$.

This section has a certain amount of detailed technical analysis. The point is that from 131A on we generally have at least two measures in play, and the ordinary language of measure theory – words like 'measurable' and 'integrable' – becomes untrustworthy in such contexts, since it omits the crucial declarations of which σ-algebras or measures are under consideration. Consequently I have to use less elegant and more explicit terminology. I hope however that once you have worked carefully through such results as 131F you will feel that the pattern formed is reasonably coherent. The general rule is that for *measurable* subspaces there are no serious surprises (131Cb, 131Fa).

I ought to remark that there is also a standard definition of subspace measure on *non-measurable* subsets of a measure space. I have given the definition already in 113Yb; for the theory of integration, extending the results above, I will wait until §214. There are significant extra difficulties and the extra generality is not often needed in elementary applications.

Let me call your attention to 131Fb-131Fc and 131Xa-131Xc; these are first steps to understanding 'indefinite integrals', the functionals $E \mapsto \int_E f : \Sigma \to \mathbb{R}$ where f is an integrable function. I will return to these in Chapters 22 and 23.

132 Outer measures from measures

The next topic I wish to mention is a simple construction with applications everywhere in measure theory. With any measure there is associated, in a straightforward way, a standard outer measure (132A-132B). If we start with Lebesgue measure we just return to Lebesgue outer measure (132C). I take the opportunity to introduce the idea of 'measurable envelope' (132D-132E).

132A Proposition Let (X, Σ, μ) be a measure space. Define $\mu^* : \mathcal{P}X \to [0,\infty]$ by writing

$$\mu^* A = \inf\{\mu E : E \in \Sigma,\ A \subseteq E\}$$

for every $A \subseteq X$. Then

(a) for every $A \subseteq X$ there is an $E \in \Sigma$ such that $A \subseteq E$ and $\mu^* A = \mu E$;

(b) μ^* is an outer measure on X;

(c) $\mu^* E = \mu E$ for every $E \in \Sigma$;

(d) a set $A \subseteq X$ is μ-negligible iff $\mu^* A = 0$;

(e) $\mu^*(\bigcup_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty} \mu^* A_n$ for every non-decreasing sequence $\langle A_n\rangle_{n\in\mathbb{N}}$ of subsets of X;

[1] I am grateful to P.Wallace Thompson for pointing out that this clause, or something with similar effect, is necessary.

(f) $\mu^*A = \mu^*(A \cap F) + \mu^*(A \setminus F)$ whenever $A \subseteq X$ and $F \in \Sigma$.

proof (a) For each $n \in \mathbb{N}$ we can choose an $E_n \in \Sigma$ such that $A \subseteq E_n$ and $\mu E_n \le \mu^*A + 2^{-n}$; now $E = \bigcap_{n\in\mathbb{N}} E_n \in \Sigma$, $A \subseteq E$ and

$$\mu^*A \le \mu E \le \inf_{n\in\mathbb{N}} \mu E_n \le \mu^*A.$$

(b)(i) $\mu^*\emptyset = \mu\emptyset = 0$. (ii) If $A \subseteq B \subseteq X$ then $\{E : A \subseteq E \in \Sigma\} \supseteq \{E : B \subseteq E \in \Sigma\}$ so $\mu^*A \le \mu^*B$. (iii) If $\langle A_n\rangle_{n\in\mathbb{N}}$ is any sequence in $\mathcal{P}X$, then for each $n \in \mathbb{N}$ there is an $E_n \in \Sigma$ such that $A_n \subseteq E_n$ and $\mu E_n = \mu^*A_n$; now $\bigcup_{n\in\mathbb{N}} A_n \subseteq \bigcup_{n\in\mathbb{N}} E_n \in \Sigma$ so

$$\mu^*(\bigcup_{n\in\mathbb{N}} A_n) \le \mu(\bigcup_{n\in\mathbb{N}} E_n) \le \sum_{n=0}^{\infty} \mu E_n = \sum_{n=0}^{\infty} \mu^*A_n.$$

(c) This is just because $\mu E \le \mu F$ whenever E, $F \in \Sigma$ and $E \subseteq F$.

(d) By (a), $\mu^*A = 0$ iff there is an $E \in \Sigma$ such that $A \subseteq E$ and $\mu E = 0$; but this is the definition of 'negligible set'.

(e) Of course $\langle \mu^*A_n\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence with limit at most μ^*A, writing $A = \bigcup_{n\in\mathbb{N}} A_n$, just because $\mu^*B \le \mu^*C$ whenever $B \subseteq C \subseteq X$. For each $n \in \mathbb{N}$, let $E_n \in \Sigma$ be such that $A_n \subseteq E_n$ and $\mu E_n = \mu^*A_n$. Set $F_n = \bigcap_{m\ge n} E_m$ for each n; then $\langle F_n\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence in Σ, and $A_n \subseteq F_n \subseteq E_n$, so $\mu^*A_n = \mu F_n$ for each $n \in \mathbb{N}$. Set $F = \bigcup_{n\in\mathbb{N}} F_n$; then $A \subseteq F$ so

$$\mu^*A \le \mu F = \lim_{n\to\infty} \mu F_n = \lim_{n\to\infty} \mu^*A_n.$$

Thus $\mu^*A = \lim_{n\to\infty} \mu^*A_n$, as claimed.

(f) Of course $\mu^*A \le \mu^*(A \cap F) + \mu^*(A \setminus F)$, by (b). On the other hand, there is an $E \in \Sigma$ such that $A \subseteq E$ and $\mu E = \mu^*A$, by (a), and now $A \cap F \subseteq E \cap F \in \Sigma$, $A \setminus F \subseteq E \setminus F \in \Sigma$ so

$$\mu^*(A \cap F) + \mu^*(A \setminus F) \le \mu(E \cap F) + \mu(E \setminus F) = \mu E = \mu^*A.$$

132B Definition If (X, Σ, μ) is a measure space, I will call μ^*, as defined in 132A, **the outer measure defined from** μ.

Remark If we start with an outer measure θ on a set X, construct a measure μ from θ by Carathéodory's method, and then construct the outer measure μ^* from μ, it is not necessarily the case that $\mu^* = \theta$. **P** Take any set X with at least three members, and set $\theta A = 0$ if $A = \emptyset$, 1 if $A = X$, $\frac{1}{2}$ otherwise. Then $\operatorname{dom}\mu = \{\emptyset, X\}$ and $\mu^*A = 1$ for every non-empty $A \subseteq X$. **Q**

However, this problem does not arise with Lebesgue outer measure. I state the next proposition in terms of Lebesgue measure on $\mathbb{R}^r$, but if you skipped §115 I hope that you will still be able to make sense of this, and later results, in terms of Lebesgue measure on $\mathbb{R}$, by setting $r = 1$.

132C Proposition If θ is Lebesgue outer measure on $\mathbb{R}^r$ and μ is Lebesgue measure, then μ^*, as defined in 132A, is equal to θ.

proof Let $A \subseteq \mathbb{R}^r$.

(a) If E is measurable and $A \subseteq E$, then $\theta A \le \theta E = \mu E$; so $\theta A \le \mu^*A$.

(b) If $\epsilon > 0$, there is a sequence $\langle I_n\rangle_{n\in\mathbb{N}}$ of half-open intervals, covering A, with $\sum_{n=0}^{\infty} \mu I_n \le \theta A + \epsilon$ (using 114G/115G to identify μI_n with the volume λI_n used in the definition of θ), so

$$\mu^*A \le \mu(\bigcup_{n\in\mathbb{N}} I_n) \le \sum_{n=0}^{\infty} \mu I_n \le \theta A + \epsilon.$$

As ϵ is arbitrary, $\mu^*A \le \theta A$.

Remark Accordingly it will henceforth be unnecessary to distinguish θ from μ^* when speaking of 'Lebesgue outer measure'. (In the language of 132Xa below, Lebesgue outer measure is 'regular'.) In particular (using 132Aa), if $A \subseteq \mathbb{R}^r$ there is a measurable set $E \supseteq A$ such that $\mu E = \theta A$ (compare 134Fc).

132D Measurable envelopes The following is a useful concept in this context. If (X, Σ, μ) is a measure space and $A \subseteq X$, a **measurable envelope** (or **measurable cover**) of A is a set $E \in \Sigma$ such that $A \subseteq E$ and $\mu(F \cap E) = \mu^*(F \cap A)$ for every $F \in \Sigma$. In general, not every set in a measure space has a measurable envelope (I suggest examples in 216Yc in Volume 2). But we do have the following.

132E Lemma Let (X, Σ, μ) be a measure space.

(a) If $A \subseteq E \in \Sigma$, then E is a measurable envelope of A iff $\mu F = 0$ whenever $F \in \Sigma$ and $F \subseteq E \setminus A$.

(b) If $A \subseteq E \in \Sigma$ and $\mu E < \infty$ then E is a measurable envelope of A iff $\mu E = \mu^* A$.

(c) If E is a measurable envelope of A and $H \in \Sigma$, then $E \cap H$ is a measurable envelope of $A \cap H$.

(d) Let $\langle A_n \rangle_{n \in \mathbb{N}}$ be a sequence of subsets of X. Suppose that each A_n has a measurable envelope E_n. Then $\bigcup_{n \in \mathbb{N}} E_n$ is a measurable envelope of $\bigcup_{n \in \mathbb{N}} A_n$.

(e) If $A \subseteq X$ can be covered by a sequence of sets of finite measure, then A has a measurable envelope.

(f) In particular, if μ is Lebesgue measure on $\mathbb{R}^r$, then every subset of $\mathbb{R}^r$ has a measurable envelope for μ.

proof (a) If E is a measurable envelope of A, $F \in \Sigma$ and $F \subseteq E \setminus A$, then

$$\mu F = \mu(F \cap E) = \mu^*(F \cap A) = 0.$$

If E is not a measurable envelope of A, there is an $H \in \Sigma$ such that $\mu^*(A \cap H) < \mu(E \cap H)$. Let $G \in \Sigma$ be such that $A \cap H \subseteq G$ and $\mu G = \mu^*(A \cap H)$, and set $F = E \cap H \setminus G$. Since $\mu G < \mu(E \cap H)$, $\mu F > 0$; but also $F \subseteq E$ and $F \cap A \subseteq H \cap A \setminus G$ is empty.

(b) If E is a measurable envelope of A then we must have

$$\mu^* A = \mu^*(A \cap E) = \mu(E \cap E) = \mu E.$$

If $\mu E = \mu^* A$, and $F \in \Sigma$ is a subset of $E \setminus A$, then $A \subseteq E \setminus F$, so $\mu(E \setminus F) = \mu E$; because μE is finite, it follows that $\mu F = 0$, so the condition of (a) is satisfied and E is a measurable envelope of A.

(c) If $F \in \Sigma$ and $F \subseteq E \cap H \setminus A$, then $F \subseteq E \setminus A$, so $\mu F = 0$, by (a); as F is arbitrary, $E \cap H$ is a measurable envelope of $A \cap H$, by (a) again.

(d) Write A for $\bigcup_{n \in \mathbb{N}} A_n$ and E for $\bigcup_{n \in \mathbb{N}} E_n$. Then $A \subseteq E$. If $F \in \Sigma$ and $F \subseteq E \setminus A$, then, for every $n \in \mathbb{N}$, $F \cap E_n \subseteq E_n \setminus A_n$, so $\mu(F \cap E_n) = 0$, by (a). Consequently $F = \bigcup_{n \in \mathbb{N}} F \cap E_n$ is negligible; as F is arbitrary, E is a measurable envelope of A.

(e) Let $\langle F_n \rangle_{n \in \mathbb{N}}$ be a sequence of sets of finite measure covering A. For each $n \in \mathbb{N}$, let $E_n \in \Sigma$ be such that $A \cap F_n \subseteq E_n$ and $\mu E_n = \mu^*(A \cap F_n)$ (using 132Aa above); by (b), E_n is a measurable envelope of $A \cap F_n$. By (d), $\bigcup_{n \in \mathbb{N}} E_n$ is a measurable envelope of $\bigcup_{n \in \mathbb{N}} A \cap F_n = A$.

(f) In the case of Lebesgue measure on $\mathbb{R}^r$, of course, the same sequence $\langle B_n \rangle_{n \in \mathbb{N}}$ will work for every A, if we take B_n to be the half-open interval $[-\mathbf{n}, \mathbf{n}[$ for each $n \in \mathbb{N}$, writing $\mathbf{n} = (n, \dots, n)$ as in §115.

132F Full outer measure This is a convenient moment at which to introduce the following term. If (X, Σ, μ) is a measure space, a set $A \subseteq X$ is **of full outer measure** or **thick** if X is a measurable envelope of A; that is, if $\mu^*(F \cap A) = \mu F$ for every $F \in \Sigma$; equivalently, if $\mu F = 0$ whenever $F \in \Sigma$ and $F \subseteq X \setminus A$. If $\mu X < \infty$, $A \subseteq X$ has full outer measure iff $\mu^* A = \mu X$.

132X Basic exercises >(a) Let X be a set and θ an outer measure on X; let μ be the measure on X defined by Carathéodory's method from θ, and μ^* the outer measure defined from μ by the construction of 132A. (i) Show that $\mu^* A \ge \theta A$ for every $A \subseteq X$. (ii) θ is said to be a **regular outer measure** if $\theta = \mu^*$. Show that if there is any measure ν on X such that $\theta = \nu^*$ then θ is regular. (iii) Show that if θ is regular and $\langle A_n \rangle_{n \in \mathbb{N}}$ is a non-decreasing sequence of subsets of X, then $\theta(\bigcup_{n \in \mathbb{N}} A_n) = \lim_{n \to \infty} \theta A_n$.

(b) Let (X, Σ, μ) be a measure space and H any member of Σ. Let μ_H be the subspace measure on H (131B) and μ^*, μ_H^* the outer measures defined from μ, μ_H respectively. Show that $\mu_H^* = \mu^* \restriction \mathcal{P} H$.

(c) Give an example of a measure space (X, Σ, μ) such that the measure $\check{\mu}$ defined by Carathéodory's method from the outer measure μ^* is a proper extension of μ. (*Hint*: take $\mu X = 0$.)

>(d) Let (X, Σ, μ) be a measure space and A a subset of X. Suppose that $\langle E_n \rangle_{n \in \mathbb{N}}$ is a sequence in Σ such that $\langle A \cap E_n \rangle_{n \in \mathbb{N}}$ is disjoint. Show that $\mu^*(A \cap \bigcup_{n \in \mathbb{N}} E_n) = \sum_{n=0}^{\infty} \mu^*(A \cap E_n)$. (*Hint*: replace E_n by $E'_n = E_n \setminus \bigcup_{i<n} E_i$, and use 132Ae-132Af.)

(e) Let (X, Σ, μ) be a measure space and $\langle A_n \rangle_{n \in \mathbb{N}}$ any sequence of subsets of X. Show that the outer measure of $\bigcup_{n \in \mathbb{N}} \bigcap_{i \ge n} A_i$ is at most $\liminf_{n \to \infty} \mu^* A_n$.

(f) Let (X, Σ, μ) be a measure space and suppose that $A \subseteq B \subseteq X$ are such that $\mu^*A = \mu^*B < \infty$. Show that $\mu^*(A \cap E) = \mu^*(B \cap E)$ for every $E \in \Sigma$. (*Hint*: a measurable envelope of B is a measurable envelope of A.)

>(g) Let ν_g be a Lebesgue-Stieltjes measure on $\mathbb{R}$, constructed as in 114Xa from a non-decreasing function $g : \mathbb{R} \to \mathbb{R}$. Show that (i) the outer measure ν_g^* derived from ν_g (132A) coincides with the outer measure θ_g of 114Xa; (ii) if $A \subseteq \mathbb{R}$ is any set, then A has a measurable envelope for the measure ν_g.

>(h) Let $A \subseteq \mathbb{R}^r$ be a set which is not measured by Lebesgue measure μ. Show that there is a bounded measurable set E such that $\mu^*(E \cap A) = \mu^*(E \setminus A) = \mu E > 0$. (*Hint*: take $E = E' \cap E'' \cap B$, where E' is a measurable envelope for A, E'' is a measurable envelope for $\mathbb{R}^r \setminus A$, and B is a suitable bounded set.)

(i) Let μ be Lebesgue measure on $\mathbb{R}^r$ and Σ its domain, and f a real-valued function, defined on a subset of $\mathbb{R}^r$, which is not Σ-measurable. Show that there are $q < q'$ in $\mathbb{Q}$ and a bounded measurable set E such that

$$\mu^*\{x : x \in E \cap \operatorname{dom} f,\ f(x) \le q\} = \mu^*\{x : x \in E \cap \operatorname{dom} f,\ f(x) \ge q'\} = \mu E > 0.$$

(*Hint*: take E_q, E'_q to be measurable envelopes for $\{x : f(x) \le q\}$, $\{x : f(x) > q\}$ for each q. Find q such that $\mu(E_q \cap E'_q) > 0$ and q' such that $\mu(E_q \cap E'_{q'}) > 0$.)

(j) Check that you can do exercise 113Yc.

(k) Let (X, Σ, μ) be a measure space and μ^* the outer measure defined from μ. Show that $\mu^*(A \cup B) + \mu^*(A \cap B) \le \mu^*A + \mu^*B$ for all A, $B \subseteq X$.

132Y Further exercises (a) Let (X, Σ, μ) be a measure space and $\langle f_n \rangle_{n \in \mathbb{N}}$ a sequence of real-valued functions defined almost everywhere in X. Suppose that $\langle \epsilon_n \rangle_{n \in \mathbb{N}}$ is a sequence of non-negative real numbers such that

$$\sum_{n=0}^{\infty} \epsilon_n < \infty, \quad \sum_{n=0}^{\infty} \mu^*\{x : |f_{n+1}(x) - f_n(x)| \ge \epsilon_n\} < \infty.$$

Show that $\lim_{n \to \infty} f_n$ is defined (as a real-valued function) almost everywhere.

(b) Let (X, Σ, μ) be a measure space, Y a set and $f : X \to Y$ a function. Let ν be the image measure μf^{-1} (112Xf). Show that $\nu^* f[A] \ge \mu^* A$ for every $A \subseteq X$.

(c) Let (X, Σ, μ) be a measure space with $\mu X < \infty$. Let $\langle A_n \rangle_{n \in \mathbb{N}}$ be a sequence of subsets of X such that $\bigcup_{n \in \mathbb{N}} A_n$ has full outer measure in X. Show that there is a partition $\langle E_n \rangle_{n \in \mathbb{N}}$ of X into measurable sets such that $\mu E_n = \mu^*(A_n \cap E_n)$ for every $n \in \mathbb{N}$.

(d) Let (X, Σ, μ) be a measure space and $\mathcal{A}$ a family of subsets of X such that $\bigcap_{n \in \mathbb{N}} A_n$ has full outer measure for every sequence $\langle A_n \rangle_{n \in \mathbb{N}}$ in $\mathcal{A}$. Show that there is a measure ν on X, extending μ, such that every member of $\mathcal{A}$ is ν-conegligible.

(e) Check that you can do exercises 113Yg-113Yh.

(f) Let (X, Σ, μ) be a measure space. Show that $\mu^* : \mathcal{P}X \to [0, \infty]$ is **alternating of all orders**, that is,

$$\sum_{J \subseteq I, \#(J) \text{ is even}} \mu^*(A \cup \bigcup_{i \in J} A_i) \le \sum_{J \subseteq I, \#(J) \text{ is odd}} \mu^*(A \cup \bigcup_{i \in J} A_i)$$

whenever I is a non-empty finite set, $\langle A_i \rangle_{i \in I}$ is a family of subsets of X and A is another subset of X.

132 Notes and comments Almost the most fundamental fact in measure theory is that in all important measure spaces there are non-measurable sets. (For Lebesgue measure see 134B below.) One can respond to this fact in a variety of ways. An approach which works quite well is just to ignore it. The point is that, for very deep reasons, the sets and functions which arise in ordinary applications nearly always are measurable, or can be made so by elementary manipulations; the only exceptions I know of in applied mathematics appear in generalized control theory. As a pure mathematician I am uncomfortable with such an approach, and as a measure theorist I think it closes the door on some of the most subtle ideas of the theory. In this treatise, therefore, non-measurable sets will always be present, if only subliminally. In this section I have described two of the basic methods of dealing with them: the move from a measure to an outer measure, which at least assigns some sort of size to an arbitrary set, and the idea of 'measurable envelope', which (when defined) describes the region in which the non-measurable set has to be taken into account. In both cases we seek to describe the non-measurable set from the outside, so to speak. There are no real difficulties, and the only points to take note of are that (i) outside the boundary marked by 132Ee measurable envelopes need not exist (ii) Carathéodory's construction of a measure from an outer measure, and the construction here of an outer measure from a measure, are closely related (132C, 132Xg, 113Yc, 132Xa(i)), but are not quite inverses of each other in general (132B, 132Xc).

133 Wider concepts of integration

There are various contexts in which it is useful to be able to assign a value to the integral of a function which is not quite covered by the basic definition in 122M. In this section I offer suggestions concerning the assignment of the values $\pm\infty$ to integrals of real-valued functions (133A), the integration of complex-valued functions (133C-133H) and upper and lower integrals (133I-133L). In §135 below I will discuss a further elaboration of the ideas of Chapter 12.

133A Infinite integrals It is normal to restrict the phrase 'f is integrable' to functions f to which a finite integral $\int f$ can be assigned (just as a series is called 'summable' only when a finite sum can be assigned to it). But for non-negative functions it is sometimes convenient to write '$\int f=\infty$' if, in some sense, the only way in which f fails to be integrable is that the integral is too large; that is, f is defined almost everywhere, is μ-virtually measurable, and either

$$\{x:x\in\operatorname{dom}f,\,f(x)\ge\epsilon\}$$

includes a set of infinite measure for some $\epsilon>0$, or

$$\sup\{\int h:h\text{ is simple},\,h\le_{\text{a.e.}}f\}=\infty.$$

(Compare 122J.) Under this rule, we shall still have

$$\int f_1+f_2=\int f_1+\int f_2,\quad\int cf=c\int f$$

whenever $c\in[0,\infty[$ and f_1, f_2, f are non-negative functions for which $\int f_1$, $\int f_2$, $\int f$ are defined in $[0,\infty]$.

We can therefore repeat the definition 122M and say that

$$\int f_1-f_2=\int f_1-\int f_2$$

whenever f_1, f_2 are real-valued functions such that $\int f_1$, $\int f_2$ are defined in $[0,\infty]$ and are not both infinite; the last condition being imposed to avoid the possibility of being asked to calculate $\infty-\infty$.

We still have the rules that

$$\int f+g=\int f+\int g,\quad\int(cf)=c\int f,\quad\int|f|\ge|\int f|$$

at least when the right-hand-sides can be interpreted, allowing $0\cdot\infty=0$, but not allowing any interpretation of $\infty-\infty$; and $\int f\le\int g$ whenever both integrals are defined and $f\le_{\text{a.e.}}g$. (But of course it is now possible to have $f\le g$ and $\int f=\int g=\pm\infty$ without f and g being equal almost everywhere.)

Setting $f^+(x)=\max(f(x),0)$, $f^-(x)=\max(-f(x),0)$ for $x\in\operatorname{dom}f$, then

$$\int f=\infty\iff\int f^+=\infty\text{ and }f^-\text{ is integrable},$$

$$\int f=-\infty\iff f^+\text{ is integrable and }\int f^-=\infty.$$

For further ideas in this direction, see §135 below.

133B Functions with exceptional values It is also convenient to allow as 'integrable' functions f which take occasional values which are not real – typically, where a formula for $f(x)$ allows the value '∞' on some convention. For such a function I will write $\int f=\int\tilde f$ if $\int\tilde f$ is defined, where

$$\operatorname{dom}\tilde f=\{x:x\in\operatorname{dom}f,\,f(x)\in\mathbb{R}\},\quad\tilde f(x)=f(x)\text{ for }x\in\operatorname{dom}\tilde f.$$

Since in this convention I still require $\tilde f$ to be defined almost everywhere in X, the set $\{x:x\in\operatorname{dom}f,\,f(x)\notin\mathbb{R}\}$ will have to be negligible.

133C Complex-valued functions All the theory of measurable and integrable functions so far developed has been devoted to real-valued functions. There are no substantial new ideas required to deal with complex-valued functions, but perhaps I should spell out some of the details, since there are many applications in which complex-valued functions are the most natural context in which to work.

133D Definitions (a) Let X be a set and Σ a σ-algebra of subsets of X. If $D\subseteq X$ and $f:D\to\mathbb{C}$ is a function, then we say that f is **measurable** if its real and imaginary parts $\mathcal{R}\mathrm{e}\,f$, $\mathcal{I}\mathrm{m}\,f$ are measurable in the sense of 121B-121C.

(b) Let (X,Σ,μ) be a measure space. If f is a complex-valued function defined on a conegligible subset of X, we say that f is **integrable** if its real and imaginary parts are integrable, and then

$$\int f=\int\mathcal{R}\mathrm{e}\,f+i\int\mathcal{I}\mathrm{m}\,f.$$

(c) Let (X, Σ, μ) be a measure space, $H \in \Sigma$ and f a complex-valued function defined on a subset of X. Then $\int_H f$ is $\int (f \restriction H) d\mu_H$ if this is defined in the sense of (b), taking the subspace measure μ_H to be that of 131A-131B.

133E Lemma (a) If X is a set, Σ is a σ-algebra of subsets of X, and f and g are measurable complex-valued functions with domains $\operatorname{dom} f$, $\operatorname{dom} g \subseteq X$, then

(i) $f + g : \operatorname{dom} f \cap \operatorname{dom} g \to \mathbb{C}$ is measurable;
(ii) $cf : \operatorname{dom} f \to \mathbb{C}$ is measurable, for every $c \in \mathbb{C}$;
(iii) $f \times g : \operatorname{dom} f \cap \operatorname{dom} g \to \mathbb{C}$ is measurable;
(iv) $f/g : \{x : x \in \operatorname{dom} f \cap \operatorname{dom} g,\ g(x) \neq 0\} \to \mathbb{C}$ is measurable;
(v) $|f| : \operatorname{dom} f \to \mathbb{R}$ is measurable.

(b) If $\langle f_n \rangle_{n \in \mathbb{N}}$ is a sequence of measurable complex-valued functions defined on subsets of X, then $f = \lim_{n\to\infty} f_n$ is measurable, if we take $\operatorname{dom} f$ to be

$$\begin{aligned}\{x : x \in \bigcup_{n\in\mathbb{N}} \bigcap_{m\ge n} \operatorname{dom} f_m,\ \lim_{n\to\infty} f_n(x) \text{ exists in } \mathbb{C}\}\\ = \operatorname{dom}(\lim_{n\to\infty} \mathcal{R}\mathrm{e}\, f_n) \cap \operatorname{dom}(\lim_{n\to\infty} \mathcal{I}\mathrm{m}\, f_n).\end{aligned}$$

proof (a) All are immediate from 121E, if you write down the formulae for the real and imaginary parts of $f+g, \dots , |f|$ in terms of the real and imaginary parts of f and g.

(b) Use 121Fa.

133F Proposition Let (X, Σ, μ) be a measure space.

(a) If f and g are integrable complex-valued functions defined on conegligible subsets of X, then $f + g$ and cf are integrable, $\int f + g = \int f + \int g$ and $\int cf = c \int f$, for every $c \in \mathbb{C}$.

(b) If f is a complex-valued function defined on a conegligible subset of X, then f is integrable iff $|f|$ is integrable and f is μ-virtually measurable, that is, $\mathcal{R}\mathrm{e}\, f$ and $\mathcal{I}\mathrm{m}\, f$ are μ-virtually measurable.

proof (a) Use 122Oa-122Ob.

(b) The point is that $|\mathcal{R}\mathrm{e}\, f|$, $|\mathcal{I}\mathrm{m}\, f| \le |f| \le |\mathcal{R}\mathrm{e}\, f| + |\mathcal{I}\mathrm{m}\, f|$; now we need only apply 122P an adequate number of times.

133G Lebesgue's Dominated Convergence Theorem Let (X, Σ, μ) be a measure space and $\langle f_n \rangle_{n\in\mathbb{N}}$ a sequence of integrable complex-valued functions on X such that $f(x) = \lim_{n\to\infty} f_n(x)$ exists in $\mathbb{C}$ for almost every $x \in X$. Suppose moreover that there is a real-valued integrable function g on X such that $|f_n| \le_{\text{a.e.}} g$ for each n. Then f is integrable and $\lim_{n\to\infty} \int f_n$ exists and is equal to $\int f$.

proof Apply 123C to the sequences $\langle \mathcal{R}\mathrm{e}\, f_n \rangle_{n\in\mathbb{N}}$ and $\langle \mathcal{I}\mathrm{m}\, f_n \rangle_{n\in\mathbb{N}}$.

133H Corollary Let (X, Σ, μ) be a measure space and $]a, b[$ a non-empty open interval in $\mathbb{R}$. Let $f : X \times]a, b[\to \mathbb{C}$ be a function such that

(i) the integral $F(t) = \int f(x, t)dx$ is defined for every $t \in]a, b[$;
(ii) the partial derivative $\frac{\partial f}{\partial t}$ of f with respect to the second variable is defined everywhere in $X \times]a, b[$;
(iii) there is an integrable function $g : X \to [0, \infty[$ such that $|\frac{\partial f}{\partial t}(x, t)| \le g(x)$ for every $x \in X$, $t \in]a, b[$.

Then the derivative $F'(t)$ and the integral $\int \frac{\partial f}{\partial t}(x, t)dx$ exist for every $t \in]a, b[$, and are equal.

proof Apply 123D to $\mathcal{R}\mathrm{e}\, f$ and $\mathcal{I}\mathrm{m}\, f$.

133I Upper and lower integrals I return now to real-valued functions. Let (X, Σ, μ) be a measure space and f a real-valued function defined almost everywhere in X. Its **upper integral** is

$$\overline{\int} f = \inf\{\int g : \int g \text{ is defined in the sense of 133A and } f \le_{\text{a.e.}} g\},$$

allowing ∞ for $\inf\{\infty\}$ or $\inf \emptyset$ and $-\infty$ for $\inf \mathbb{R}$. Similarly, the **lower integral** of f is

$$\underline{\int} f = \sup\{\int g : \int g \text{ is defined, } f \ge_{\text{a.e.}} g\},$$

allowing $-\infty$ for $\sup\{-\infty\}$ or $\sup \emptyset$ and ∞ for $\sup \mathbb{R}$.

133J Proposition Let (X,Σ,μ) be a measure space.

(a) Let f be a real-valued function defined almost everywhere in X.

(i) If $\overline{\int} f$ is finite, then there is an integrable g such that $f \leq_{\text{a.e.}} g$ and $\int g = \overline{\int} f$. In this case,

$$\{x : x \in \operatorname{dom} f \cap \operatorname{dom} g,\ g(x) \leq f(x) + g_0(x)\}$$

has full outer measure for every measurable function $g_0 : X \to]0,\infty[$.

(ii) If $\underline{\int} f$ is finite, then there is an integrable h such that $h \leq_{\text{a.e.}} f$ and $\int h = \underline{\int} f$. In this case,

$$\{x : x \in \operatorname{dom} f \cap \operatorname{dom} h,\ f(x) \leq h(x) + h_0(x)\}$$

has full outer measure for every measurable function $h_0 : X \to]0,\infty[$.

(b) For any real-valued functions f, g defined on conegligible subsets of X and any $c \geq 0$,

(i) $\underline{\int} f \leq \overline{\int} f$,

(ii) $\overline{\int} f + g \leq \overline{\int} f + \overline{\int} g$,

(iii) $\overline{\int} cf = c\overline{\int} f$,

(iv) $\underline{\int}(-f) = -\overline{\int} f$,

(v) $\underline{\int} f + g \geq \underline{\int} f + \underline{\int} g$,

(vi) $\underline{\int} cf = c\underline{\int} f$

whenever the right-hand-sides do not involve adding ∞ to $-\infty$.

(c) If $f \leq_{\text{a.e.}} g$ then $\overline{\int} f \leq \overline{\int} g$ and $\underline{\int} f \leq \underline{\int} g$.

(d) A real-valued function f defined almost everywhere in X is integrable iff

$$\overline{\int} f = \underline{\int} f = a \in \mathbb{R},$$

and in this case $\int f = a$.

(e) $\mu^* A = \overline{\int} \chi A$ for every $A \subseteq X$.

proof (a)(i) For each $n \in \mathbb{N}$, choose a function g_n such that $f \leq_{\text{a.e.}} g_n$ and $\int g_n$ is defined and at most $2^{-n} + \overline{\int} f$; as $\overline{\int} f \leq \int g_n$, $\int g_n$ is finite, so g_n is integrable. Set $h_n = \inf_{i \leq n} g_i$ for each n; then h_n is integrable (because $|h_n - g_0| \leq \sum_{i=0}^{n} |g_i - g_0|$ on $\bigcap_{i \leq n} \operatorname{dom} g_i$), and $f \leq_{\text{a.e.}} h_n$, so

$$\overline{\int} f \leq \int h_n \leq \int g_n \leq 2^{-n} + \overline{\int} f.$$

By B.Levi's theorem (123A), applied to $\langle -h_n \rangle_{n\in\mathbb{N}}$, $g(x) = \inf_{n\in\mathbb{N}} h_n(x) \in \mathbb{R}$ for almost every x, and $\int g = \inf_{n\in\mathbb{N}} \int h_n = \overline{\int} f$; also, of course, $f \leq_{\text{a.e.}} g$.

Now take a measurable function $g_0 : X \to]0,\infty[$, and consider the set

$$A = \{x : x \in \operatorname{dom} f \cap \operatorname{dom} g,\ g(x) \leq f(x) + g_0(x)\}.$$

? If A does not have full outer measure, there is a non-negligible measurable set $F \subseteq X \setminus A$. Since g_0 is strictly positive, $F = \bigcup_{n\in\mathbb{N}} F_n$ where $F_n = \{x : x \in F,\ g_0(x) \geq 2^{-n}\}$, and there is an $n \in \mathbb{N}$ such that $\mu F_n > 0$. Consider the function $g_1 = g - 2^{-n}\chi F$. Then $f \leq_{\text{a.e.}} g_1$. Also $\int g_1 = \int g - 2^{-n}\mu F_n$ is strictly less than $\int g$, so $\overline{\int} f < \int g$. **X**

(ii) Argue similarly, or use (b-iv).

(b)(i) If either $\underline{\int} f = -\infty$ or $\overline{\int} f = \infty$ this is trivial. Otherwise it follows at once from the fact that if $g \leq_{\text{a.e.}} f \leq_{\text{a.e.}} h$ then $\int g \leq \int h$ if the integrals are defined (in the wide sense).

(ii) If $a > \overline{\int} f + \overline{\int} g$, neither $\overline{\int} f$ nor $\overline{\int} g$ can be ∞, so there must be functions f_1, g_1 such that $f \leq_{\text{a.e.}} f_1$, $g \leq_{\text{a.e.}} g_1$ and $\int f_1 + \int g_1 \leq a$. Now $f + g \leq_{\text{a.e.}} f_1 + g_1$, so

$$\overline{\int} f + g \leq \int f_1 + g_1 \leq a.$$

As a is arbitrary, we have the result.

(iii)(α) If $c = 0$ this is trivial. **(β)** If $c > 0$ and $a > c\overline{\int} f$, there must be an f_1 such that $f \leq_{\text{a.e.}} f_1$ and $c\int f_1 \leq a$. Now $cf \leq_{\text{a.e.}} cf_1$ and $\int cf_1 \leq a$, so $\overline{\int} cf \leq a$. As a is arbitrary, $\overline{\int} cf \leq c\overline{\int} f$. **(γ)** Still supposing that $c > 0$, we also have

$$c\overline{\int} f = c\overline{\int} c^{-1}cf \leq cc^{-1}\overline{\int} cf = \overline{\int} cf,$$

so we get equality.

(iv) This is just because $\overline{\int}(-f_1) = -\underline{\int} f_1$ for any function f_1 for which either integral is defined.

(v)-(vi) Use (iv) to turn $\underline{\int}$ into $\overline{\int}$, and apply (ii) or (iii).

(c) These are immediate from the definitions, because (for instance) if $g \leq_{\text{a.e.}} h$ then $f \leq_{\text{a.e.}} h$.

(d) If f is integrable, then

$$\overline{\int} f = \int f = \underline{\int} f$$

by 122Od. If $\overline{\int} f = \underline{\int} f = a \in \mathbb{R}$, then, by (a), there are integrable g, h such that $g \leq_{\text{a.e.}} f \leq_{\text{a.e.}} h$ and $\int g = \int h = a$, so that $g =_{\text{a.e.}} h$, by 122Rc, $g =_{\text{a.e.}} f =_{\text{a.e.}} h$ and f is integrable, by 122Rb.

(e) If $E \supseteq A$ is measurable, then

$$\mu E = \int \chi E \geq \overline{\int} \chi A;$$

as E is arbitrary, $\mu^* A \geq \overline{\int} \chi A$. If $\int g$ is defined and $\chi A \leq_{\text{a.e.}} g$, let $E \subseteq \operatorname{dom} g$ be a conegligible measurable set such that $g \restriction E$ is measurable, and set $F = \{x : x \in E,\ g(x) \geq 1\}$. Then $A \setminus F$ is negligible, so $\mu^* A \leq \mu F \leq \int g$; as g is arbitrary, $\mu^* A \leq \overline{\int} \chi A$.

Remark I hope that the formulae here remind you of $\limsup$, $\liminf$.

133K Convergence theorems for upper integrals We have the following versions of B.Levi's theorem and Fatou's Lemma.

Proposition Let (X, Σ, μ) be a measure space, and $\langle f_n \rangle_{n \in \mathbb{N}}$ a sequence of real-valued functions defined almost everywhere in X.

(a) If, for each n, $f_n \leq_{\text{a.e.}} f_{n+1}$, and $-\infty < \sup_{n \in \mathbb{N}} \overline{\int} f_n < \infty$, then $f(x) = \sup_{n \in \mathbb{N}} f_n(x)$ is defined in $\mathbb{R}$ for almost every $x \in X$, and $\overline{\int} f = \sup_{n \in \mathbb{N}} \overline{\int} f_n$.

(b) If, for each n, $f_n \geq 0$ a.e., and $\liminf_{n \to \infty} \overline{\int} f_n < \infty$, then $f(x) = \liminf_{n \to \infty} f_n(x)$ is defined in $\mathbb{R}$ for almost every $x \in X$, and $\overline{\int} f \leq \liminf_{n \to \infty} \overline{\int} f_n$.

proof (a) Set $c = \sup_{n \in \mathbb{N}} \overline{\int} f_n$. For each n, there is an integrable function g_n such that $f_n \leq_{\text{a.e.}} g_n$ and $\int g_n = \overline{\int} f_n$ (133J(a-i)). Set $g_n' = \min(g_n, g_{n+1})$; then g_n' is integrable and $f_n \leq_{\text{a.e.}} g_n' \leq_{\text{a.e.}} g_n$, so

$$\overline{\int} f_n \leq \int g_n' \leq \int g_n = \overline{\int} f_n$$

and g_n' must be equal to g_n a.e. Consequently $g_n \leq_{\text{a.e.}} g_{n+1}$, for each n, while $\sup_{n \in \mathbb{N}} \int g_n = c < \infty$. By B.Levi's theorem, $g = \sup_{n \in \mathbb{N}} g_n$ is defined, as a real-valued function, almost everywhere in X, and $\int g = c$. Now of course $f(x)$ is defined, and not greater than $g(x)$, for any $x \in \operatorname{dom} g \cap \bigcap_{n \in \mathbb{N}} \operatorname{dom} f_n$ such that $f_n(x) \leq g_n(x)$ for every n, that is, for almost every x; so $\overline{\int} f \leq \int g = c$. On the other hand, $f_n \leq_{\text{a.e.}} f$, so $\overline{\int} f_n \leq \overline{\int} f$, for every $n \in \mathbb{N}$; it follows that $\overline{\int} f$ must be at least c, and is therefore equal to c, as required.

(b) The argument follows that of 123B. Set $c = \liminf_{n \to \infty} \overline{\int} f_n$. For each n, set $g_n = \inf_{m \geq n} f_n$; then $\overline{\int} g_n \leq \inf_{m \geq n} \overline{\int} f_m \leq c$. We have $g_n(x) \leq g_{n+1}(x)$ for every $x \in \operatorname{dom} g_n$, that is, almost everywhere, for each n; so, by (a),

$$\overline{\int} g = \sup_{n \in \mathbb{N}} \overline{\int} g_n \leq c,$$

where

$$g = \sup_{n \in \mathbb{N}} g_n =_{\text{a.e.}} \liminf_{n \to \infty} f_n,$$

and $\overline{\int} \liminf_{n \to \infty} f_n \leq c$, as claimed.

***133L** The following is at a less fundamental level than the results in 133J, but is still important.

Proposition Let (X, Σ, μ) be a measure space and f a real-valued function defined almost everywhere in X. Suppose that h_1, h_2 are non-negative virtually measurable functions defined almost everywhere in X. Then

$$\overline{\int} f \times (h_1 + h_2) = \overline{\int} f \times h_1 + \overline{\int} f \times h_2,$$

where here, for once, we can interpret $\infty + (-\infty)$ or $(-\infty) + \infty$ as ∞ if called for on the right-hand side.

proof (a) If either $\overline{\int} f \times h_1 = \infty$ or $\overline{\int} f \times h_2 = \infty$ then $\overline{\int} f \times (h_1 + h_2) = \infty$. **P?** Otherwise, there is a g such that $f \times (h_1 + h_2) \leq_{\text{a.e.}} g$ and $\int g < \infty$. In this case,

$$f \times h_1 \leq_{\text{a.e.}} f^+ \times h_1 \leq_{\text{a.e.}} f^+ \times (h_1 + h_2) = (f \times (h_1 + h_2))^+ \leq_{\text{a.e.}} g^+$$

so $\overline{\int} f \times h_1 \leq \int g^+ < \infty$. Similarly, $\overline{\int} f \times h_2 < \infty$; contradicting our hypothesis. **XQ** So in this case, under the local rule $\infty + (-\infty) = (-\infty) + \infty = \infty$, we have the result.

(b) Now suppose that the upper integrals $\overline{\int} f \times h_1$ and $\overline{\int} f \times h_2$ are both less than ∞, so that their sum can be interpreted by the usual rules. By 133J(b-ii), $\overline{\int} f \times (h_1 + h_2) \leq \overline{\int} f \times h_1 + \overline{\int} f \times h_2 < \infty$. In the other direction, suppose that $g \geq_{\text{a.e.}} f \times (h_1 + h_2)$ and $\int g < \infty$. For $i = 1, 2$ set

$$\begin{aligned} g_i(x) &= \tfrac{g(x)h_i(x)}{h_1(x)+h_2(x)} \text{ if } x \in \operatorname{dom} g \cap \operatorname{dom} h_1 \cap \operatorname{dom} h_2 \text{ and } h_1(x) + h_2(x) > 0, \\ &= 0 \text{ for other } x \in X. \end{aligned}$$

Then, for both i, g_i is virtually measurable, $g_i^+ \leq_{\text{a.e.}} g^+$ and $g_i \geq_{\text{a.e.}} f \times h_i$; while $g \geq_{\text{a.e.}} g_1 + g_2$. **P** The set

$$H = \{x : x \in \operatorname{dom} f \cap \operatorname{dom} g \cap \operatorname{dom} h_1 \cap \operatorname{dom} h_2,\ g(x) \geq f(x)(h_1(x) + h_2(x))\}$$

is conegligible, and for $x \in H$

$$\begin{aligned} g(x) &= g_1(x) + g_2(x) \text{ if } h_1(x) + h_2(x) > 0, \\ &\geq 0 = g_1(x) + g_2(x) \text{ if } h_1(x) + h_2(x) = 0. \quad \mathbf{Q} \end{aligned}$$

So

$$\overline{\int} f \times h_1 + \overline{\int} f \times h_2 \leq \int g_1 + \int g_2 = \int g_1 + g_2 \leq \int g$$

(because $\int g_1$ and $\int g_2$ are both at most $\int g^+ < \infty$, so we can add them on the usual rules). As g is arbitrary, $\overline{\int} f \times h_1 + \overline{\int} f \times h_2 \leq \overline{\int} f \times (h_1 + h_2)$ and we must have equality.

133X Basic exercises >(a) Let (X, Σ, μ) be a measure space, and $f : X \to [0, \infty[$ a measurable function. Show that

$$\begin{aligned} \int f d\mu &= \sup_{n \in \mathbb{N}} 2^{-n} \sum_{k=1}^{4^n} \mu\{x : f(x) \geq 2^{-n}k\} \\ &= \lim_{n \to \infty} 2^{-n} \sum_{k=1}^{4^n} \mu\{x : f(x) \geq 2^{-n}k\} \end{aligned}$$

in $[0, \infty]$.

(b) Let (X, Σ, μ) be a measure space and f a complex-valued function defined on a subset of X. (i) Show that if $E \in \Sigma$, then $f{\restriction}E$ is μ_E-integrable iff $\tilde{f}$ is μ-integrable, writing μ_E for the subspace measure on E and $\tilde{f}(x) = f(x)$ if $x \in E \cap \operatorname{dom} f$, 0 if $x \in X \setminus E$; and in this case $\int_E f d\mu_E = \int \tilde{f} d\mu$. (ii) Show that if $E \in \Sigma$ and f is defined μ-almost everywhere, then $f{\restriction}E$ is μ_E-integrable iff $f \times \chi E$ is μ-integrable, and in this case $\int_E f = \int f \times \chi E$. (iii) Show that if $\int_E f = 0$ for every $E \in \Sigma$, then $f = 0$ a.e.

(c) Suppose that (X, Σ, μ) is a measure space and that G is an open subset of $\mathbb{C}$, that is, a set such that for every $w \in G$ there is a $\delta > 0$ such that $\{z : |z - w| < \delta\} \subseteq G$. Let $f : X \times G \to \mathbb{C}$ be a function, and suppose that the derivative $\frac{\partial f}{\partial z}$ of f with respect to the second variable exists for all $x \in X$, $z \in G$. Suppose moreover that (i) $F(z) = \int f(x, z)dx$ exists for every $z \in G$ (ii) there is an integrable function g such that $|\frac{\partial f}{\partial z}(x, z)| \leq g(x)$ for every $x \in X$, $z \in G$. Show that the derivative F' of F exists everywhere in G, and $F'(z) = \int \frac{\partial f}{\partial z}(x, z)dx$ for every $z \in G$. (*Hint*: you will need to check that $|f(x, z) - f(x, w)| \leq |z - w|g(x)$ whenever $x \in X$, $z \in G$ and w is close to z.)

>(d) Let f be a complex-valued function defined almost everywhere on $[0, \infty[$, endowed as usual with Lebesgue measure. Its **Laplace transform** is the function F defined by writing

$$F(s) = \int_0^\infty e^{-sx} f(x)dx$$

for all those complex numbers s for which the integral is defined in $\mathbb{C}$.

(i) Show that if $s \in \operatorname{dom} F$ and $\mathcal{R}\mathrm{e}\, s' \ge \mathcal{R}\mathrm{e}\, s$ then $s' \in \operatorname{dom} F$ (because $|e^{-s'x}e^{sx}| \le 1$ for all x).

(ii) Show that F is analytic (that is, differentiable as a function of a complex variable) on the interior of its domain. (*Hint*: 133Xc.)

(iii) Show that if F is defined anywhere then $\lim_{\mathcal{R}\mathrm{e}\, s\to\infty} F(s) = 0$.

(iv) Show that if f, g have Laplace transforms F, G then the Laplace transform of $f + g$ is $F + G$, at least on $\operatorname{dom} F \cap \operatorname{dom} G$.

>(e) Let f be an integrable complex-valued function defined almost everywhere in $\mathbb{R}$, endowed as usual with Lebesgue measure. Its **Fourier transform** is the function $\hat{f}$ defined by

$$\hat{f}(s) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-isx} f(x)dx$$

for all real s.

(i) Show that $\hat{f}$ is continuous. (*Hint*: use Lebesgue's Dominated Convergence Theorem on sequences of the form $f_n(x) = e^{-is_n x} f(x)$.)

(ii) Show that if f, g have Fourier transforms $\hat{f}$, $\hat{g}$ then the Fourier transform of $f + g$ is $\hat{f} + \hat{g}$.

(iii) Show that if $\int xf(x)dx$ exists then $\hat{f}$ is differentiable, with $\hat{f}'(s) = -\frac{i}{\sqrt{2\pi}}\int xe^{-isx}f(x)dx$ for every s.

(f) Let (X, Σ, μ) be a measure space and $\langle f_n\rangle_{n\in\mathbb{N}}$ a sequence of real-valued functions each defined almost everywhere in X. Suppose that there is an integrable real-valued function g such that $|f_n| \le_{\text{a.e.}} g$ for each n. Show that

$$\overline{\int} \liminf_{n\to\infty} f_n \le \liminf_{n\to\infty} \overline{\int} f_n, \quad \underline{\int} \limsup_{n\to\infty} f_n \ge \limsup_{n\to\infty} \underline{\int} f_n.$$

133Y Further exercises (a) Use the ideas of 133C-133H to develop a theory of measurable and integrable functions taking values in $\mathbb{R}^r$, where $r \ge 2$.

(b) Let X be a set and Σ a σ-algebra of subsets of X. Let Y be a subset of X and $f : Y \to \mathbb{C}$ a Σ_Y-measurable function, where $\Sigma_Y = \{E \cap Y : E \in \Sigma\}$. Show that there is a Σ-measurable function $\tilde{f} : X \to \mathbb{C}$ extending f. (*Hint*: 121I.)

(c) Let f be an integrable complex-valued function defined almost everywhere in $\mathbb{R}^r$, endowed as usual with Lebesgue measure, where $r \ge 1$. Its **Fourier transform** is the function $\hat{f}$ defined by

$$\hat{f}(s) = \frac{1}{(\sqrt{2\pi})^r}\int e^{-is\centerdot x} f(x)dx$$

for all $s \in \mathbb{R}^r$, writing $s\centerdot x$ for $\sigma_1\xi_1 + \ldots + \sigma_r\xi_r$ if $s = (\sigma_1, \ldots , \sigma_r)$, $x = (\xi_1, \ldots , \xi_r) \in \mathbb{R}^r$.

(i) Show that $\hat{f}$ is continuous.

(ii) Show that if f, g have Fourier transforms $\hat{f}$, $\hat{g}$ then the Fourier transform of $f + g$ is $\hat{f} + \hat{g}$.

(iii) Show that if $\int \|x\||f(x)|dx$ is finite (taking $\|x\| = \sqrt{\xi_1^2 + \ldots + \xi_r^2}$ if $x = (\xi_1, \ldots, \xi_r)$), then $\hat{f}$ is differentiable, with

$$\frac{\partial\hat{f}}{\partial\sigma_k}(s) = -\frac{i}{(\sqrt{2\pi})^r}\int \xi_k e^{-is\centerdot x} f(x)dx$$

for every $s \in \mathbb{R}^r$, $k \le r$.

(d) Recall the definition of 'quasi-simple' function from 122Yd. Show that for any measure space (X, Σ, μ) and any real-valued function f defined almost everywhere in X,

$$\overline{\int} f = \inf\{\int g : g \text{ is quasi-simple, } f \le_{\text{a.e.}} g\},$$

$$\underline{\int} f = \sup\{\int g : g \text{ is quasi-simple, } f \ge_{\text{a.e.}} g\},$$

allowing ∞ for $\inf\emptyset$ and $\sup\mathbb{R}$ and $-\infty$ for $\inf\mathbb{R}$ and $\sup\emptyset$.

(e) State and prove a similar result concerning the 'pseudo-simple' functions of 122Ye.

133 Notes and comments I have spelt this section out in detail, even though there is nothing that can really be called a new idea in it, because it gives us an opportunity to review the previous work, and because the manipulations which are by now, I hope, becoming 'obvious' to you are in fact justifiable only through difficult theorems, and I believe that it is at least some of the time right to look back to the exact points at which justifications were written out.

You may have noticed similarities between results involving 'upper integrals', as described here, and those of §132 concerning 'outer measure' (132Ae and 133Ka, for instance, or 132Xe and 133Kb). These are not a coincidence; an explanation of sorts can be found in 252Ym in Volume 2.

134 More on Lebesgue measure

The special properties of Lebesgue measure will take up a substantial proportion of this treatise. In this section I present a miscellany of relatively easy basic results. In 134A-134F, r will be a fixed integer greater than or equal to 1, μ will be Lebesgue measure on $\mathbb{R}^r$ and μ^* will be Lebesgue outer measure (see 132C); when I say that a set or a function is 'measurable', then it is to be understood that (unless otherwise stated) this means 'measurable with respect to the σ-algebra of Lebesgue measurable sets', while 'negligible' means 'negligible for Lebesgue measure'. Most of the results will be expressed in terms adapted to the multi-dimensional case; but if you are primarily interested in the real line, you will miss none of the ideas if you read the whole section as if $r = 1$.

134A Proposition Both Lebesgue outer measure and Lebesgue measure are translation-invariant; that is, setting $A + x = \{a + x : a \in A\}$ for $A \subseteq \mathbb{R}^r$, $x \in \mathbb{R}^r$, we have

(a) $\mu^*(A + x) = \mu^* A$ for every $A \subseteq \mathbb{R}^r$, $x \in \mathbb{R}^r$;

(b) whenever $E \subseteq \mathbb{R}^r$ is measurable and $x \in \mathbb{R}^r$, then $E + x$ is measurable, with $\mu(E + x) = \mu E$.

proof The point is that if $I \subseteq \mathbb{R}^r$ is a half-open interval, as defined in 114Aa/115Ab, then so is $I+x$, and $\lambda(I+x) = \lambda I$ for every $x \in \mathbb{R}^r$, where λ is defined as in 114Ab/115Ac; this is immediate from the definition, since $[a, b[+ x = [a + x, b + x[$.

(a) If $A \subseteq \mathbb{R}^r$ and $x \in \mathbb{R}^r$ and $\epsilon > 0$, we can find a sequence $\langle I_j \rangle_{j\in\mathbb{N}}$ of half-open intervals such that $A \subseteq \bigcup_{j\in\mathbb{N}} I_j$ and $\sum_{j=0}^{\infty} \lambda I_j \le \mu^* A + \epsilon$. Now $A + x \subseteq \bigcup_{j\in\mathbb{N}}(I_j + x)$ so

$$\mu^*(A + x) \le \sum_{j=0}^{\infty} \lambda(I_j + x) = \sum_{j=0}^{\infty} \lambda I_j \le \mu^* A + \epsilon.$$

As ϵ is arbitrary, $\mu^*(A + x) \le \mu^* A$. Similarly

$$\mu^* A = \mu^*((A + x) + (-x)) \le \mu^*(A + x),$$

so $\mu^*(A + x) = \mu^* A$, as claimed.

(b) Now suppose that $E \subseteq \mathbb{R}^r$ is measurable and $x \in \mathbb{R}^r$, and that $A \subseteq \mathbb{R}^r$. Then, using (a) repeatedly,

$$\begin{aligned}\mu^*(A \cap (E + x)) + \mu^*(A \setminus (E + x)) &= \mu^*(((A - x) \cap E) + x) + \mu^*(((A - x) \setminus E) + x)\\ &= \mu^*((A - x) \cap E) + \mu^*((A - x) \setminus E)\\ &= \mu^*(A - x) = \mu^* A,\end{aligned}$$

writing $A - x$ for $A + (-x) = \{a - x : a \in A\}$. As A is arbitrary, $E + x$ is measurable. Now

$$\mu(E + x) = \mu^*(E + x) = \mu^* E = \mu E.$$

134B Theorem Not every subset of $\mathbb{R}^r$ is Lebesgue measurable.

proof Set $\mathbf{0} = (0, \dots , 0)$, $\mathbf{1} = (1, \dots , 1) \in \mathbb{R}^r$. On

$$[\mathbf{0}, \mathbf{1}[= \{(\xi_1, \dots , \xi_r) : \xi_i \in [0, 1[\text{ for every } i \le r\},$$

consider the relation $\sim$, defined by saying that $x \sim y$ iff $y - x \in \mathbb{Q}^r$. It is easy to see that this is an equivalence relation, so divides $[\mathbf{0}, \mathbf{1}[$ into equivalence classes. Choose one point from each of these equivalence classes, and let A be the set of points obtained in this way. Then $\mu^* A \le \mu^* [\mathbf{0}, \mathbf{1}[= 1$.

Consider $A + \mathbb{Q}^r = \{a + q : a \in A,\ q \in \mathbb{Q}^r\} = \bigcup_{q\in\mathbb{Q}^r} A + q$. This is equal to $\mathbb{R}^r$. **P** If $x \in \mathbb{R}^r$, there is an $e \in \mathbb{Z}^r$ such that $x - e \in [\mathbf{0}, \mathbf{1}[$; there is an $a \in A$ such that $a \sim x - e$, that is, $x - e - a \in \mathbb{Q}^r$; now $x = a + (e + x - e - a) \in A + \mathbb{Q}^r$. **Q** Next, $\mathbb{Q}^r$ is countable (111F(b-iv)), so we have

$$\infty = \mu\mathbb{R}^r \le \sum_{q\in\mathbb{Q}^r} \mu^*(A + q),$$

and there must be some $q \in \mathbb{Q}^r$ such that $\mu^*(A+q) > 0$; but as μ^* is translation-invariant (134A), $\mu^*A > 0$.

Take $n \in \mathbb{N}$ such that $n > 2^r/\mu^*A$, and distinct $q_1, \ldots, q_n \in [\mathbf{0}, \mathbf{1}[\cap \mathbb{Q}^r$. If a, $b \in A$ and $1 \le i < j \le n$, then $a + q_i \ne b + q_j$; for if $a = b$ then $q_i \ne q_j$, while if $a \ne b$ then $a \not\sim b$ so $b - a \ne q_i - q_j$. Thus $A + q_1, \ldots, A + q_n$ are disjoint. On the other hand, all are subsets of $[\mathbf{0}, \mathbf{2}[$. So we have

$$\sum_{i=1}^n \mu^*(A+q_i) = n\mu^*A > 2^r = \mu\,[\mathbf{0},\mathbf{2}[\ge \mu^*(\bigcup_{1\le i\le n}(A+q_i)).$$

It follows that not all the $A + q_i$ can be measurable; as Lebesgue measure is translation-invariant, we see that A itself is not measurable. In any case we have found a non-measurable set.

***134C Remark** 134B is known as 'Vitali's construction'.

Observe that at the beginning of the proof I asked you to *choose* one member of each of the equivalence classes for $\sim$. This is of course an appeal to the Axiom of Choice. So far I have made rather few appeals to the axiom of choice. One was in (a-iv) of the proof of 114D/115D; an earlier one was in 112Db; yet another in 121A. See also 1A1F. In all of these, only 'countable choice' was involved; that is, I needed to choose simultaneously one member of each of a named sequence of sets. Because there are surely uncountably many equivalence classes for $\sim$, the form of choice needed for the example above is essentially stronger than that needed for the positive results so far. It is in fact the case that very large parts of measure theory can be developed without appealing to the full strength of the axiom of choice.

The significance of this is that it suggests the possibility that there might be a consistent mathematical system in which enough of the axiom of choice is valid to make measure theory possible, without having enough to construct a non-Lebesgue-measurable set. Such a system has indeed been worked out by R.M.Solovay (SOLOVAY 70). (In a formal sense there is room for a residual doubt concerning its consistency. In my view this is of no importance.) In Volume 5 I will return to the question of what Lebesgue measure looks like with a weak axiom of choice, or none at all. For the moment, I have to say that nearly all measure theory continues to proceed in directions at least consistent with the full axiom of choice, so that non-measurable sets are constantly present, at least potentially; and that will be my normal position in this treatise. But I mention the point at this early stage because I believe that it could happen at any time that the focus of interest might switch to systems in which the axiom of choice is false; and in this case measure theory without non-measurable sets might become important to many pure mathematicians, and even to applied mathematicians, who have no reason, other than the convenience of being able to quote results from books like this one, for loyalty to the axiom of choice.

I ought to remark that while we need a fairly strong form of the axiom of choice to construct a non-Lebesgue-measurable set, a non-Borel set can be constructed in much weaker set theories. One possible construction is outlined in §423 in Volume 4.

Of course there is a non-Lebesgue-measurable subset of $\mathbb{R}$ iff there is a non-Lebesgue-measurable function from $\mathbb{R}$ to $\mathbb{R}$; for if every set is measurable, then the definition 121C makes it plain that every real-valued function on any subset of $\mathbb{R}$ is measurable; while if $A \subseteq \mathbb{R}$ is not measurable, then $\chi A : \mathbb{R} \to \mathbb{R}$ is not measurable.

***134D** In fact there are much stronger results than 134B concerning the existence of non-measurable sets (provided, of course, that we allow ourselves to use the axiom of choice). Here I give one which can be reached by a slight refinement of the methods of 134B.

Proposition There is a set $C \subseteq \mathbb{R}^r$ such that $F \cap C$ is not measurable for any measurable set F of non-zero measure; so that both C and its complement have full outer measure in $\mathbb{R}^r$.

proof (a) Start from a set $A \subseteq [\mathbf{0}, \mathbf{1}[\subseteq \mathbb{R}^r$ such that $\langle A + q\rangle_{q\in\mathbb{Q}^r}$ is a partition of $\mathbb{R}^r$, as constructed in the proof of 134B. As in 134B, the outer measure μ^*A of A must be greater than 0. The argument there shows in fact that $\mu F = 0$ for every measurable set $F \subseteq A$. **P** For every n we can find distinct $q_1, \ldots, q_n \in [\mathbf{0}, \mathbf{1}[\cap \mathbb{Q}^r$, and now

$$n\mu F = \mu(\bigcup_{1\le i\le n} F + q_i) \le \mu\,[\mathbf{0}, \mathbf{2}[= 2^r,$$

so that $\mu F \le 2^r/n$; as n is arbitrary, $\mu F = 0$. **Q**

(b) Now let $E \subseteq [\mathbf{0}, \mathbf{1}[$ be a measurable envelope of A (132Ef). Then $E + q$ is a measurable envelope of $A + q$ for any q. **P** I hope that this will very soon be 'an obvious consequence of the translation-invariance of Lebesgue measure'. In detail: $A + q \subseteq E + q$, $E + q$ is measurable and, for any measurable F,

$$\begin{aligned}\mu(F \cap (E+q)) &= \mu(((F-q)\cap E)+q) = \mu((F-q)\cap E)\\ &= \mu^*((F-q)\cap A) = \mu^*(((F-q)\cap A)+q) = \mu^*(F\cap(A+q)),\end{aligned}$$

using 134A repeatedly. **Q** Also E is a measurable envelope of $A' = E \setminus A$. **P** Of course E is a measurable set including A'. If $F \subseteq E \setminus A'$ is measurable then $F \subseteq A$, so $\mu F = 0$, by (a); now 132Ea tells us that E is a measurable envelope of A'. **Q** It follows that $E + q$ is a measurable envelope of $A' + q$ for every q.

(c) Let $\langle q_n\rangle_{n\in\mathbb{N}}$ be a sequence running over $\mathbb{Q}^r$. Then

$$\bigcup_{n\in\mathbb{N}} E+q_n \supseteq \bigcup_{n\in\mathbb{N}} A+q_n = \mathbb{R}^r.$$

Write E_n for $E+q_n\setminus\bigcup_{i<n} E+q_i$ for $n\in\mathbb{N}$, so that $\langle E_n\rangle_{n\in\mathbb{N}}$ is disjoint and $\bigcup_{n\in\mathbb{N}} E_n=\mathbb{R}^r$.

Now set

$$C=\bigcup_{n\in\mathbb{N}} E_n\cap(A+q_n).$$

This is a set with the required properties.

P (i) Let $F\subseteq\mathbb{R}^r$ be any non-negligible measurable set. Then there must be some $n\in\mathbb{N}$ such that $\mu(F\cap E_n)>0$. But this means that

$$\mu^*(F\cap E_n\cap C)\ge\mu^*(F\cap E_n\cap(A+q_n))=\mu(F\cap E_n\cap(E+q_n))=\mu(F\cap E_n),$$

$$\begin{aligned}\mu^*(F\cap E_n\setminus C)&\ge\mu^*(F\cap E_n\cap((E+q_n)\setminus(A+q_n)))\\&=\mu(F\cap E_n\cap(E+q_n))=\mu(F\cap E_n).\end{aligned}$$

Since

$$\mu(F\cap E_n)\le\mu(E+q_n)=\mu E\le 1,$$

$\mu^*(F\cap E_n\cap C)+\mu^*(F\cap E_n\setminus C)>\mu(F\cap E_n)$, and $F\cap C$ cannot be measurable.

(ii) In particular, no measurable subset of $\mathbb{R}^r\setminus C$ can have non-zero measure, and C has full outer measure; similarly, C has no measurable subset of non-zero measure, and $\mathbb{R}^r\setminus C$ has full outer measure. **Q**

Remark In fact it is the case that for any sequence $\langle D_n\rangle_{n\in\mathbb{N}}$ of subsets of $\mathbb{R}^r$ there is a set $C\subseteq\mathbb{R}^r$ such that

$$\mu^*(E\cap D_n\cap C)=\mu^*(E\cap D_n\setminus C)=\mu^*(E\cap D_n)$$

for every measurable set $E\subseteq\mathbb{R}^r$ and every $n\in\mathbb{N}$. But for the proof of this result we must wait for Volume 5.

134E Borel sets and Lebesgue measure on $\mathbb{R}^r$ Recall from 111G that the family $\mathcal{B}$ of Borel sets in $\mathbb{R}^r$ is the σ-algebra generated by the family of open sets. In 114G/115G I showed that every Borel set in $\mathbb{R}^r$ is Lebesgue measurable. It is time we returned to the topic and looked more closely at the very intimate connexion between Borel and measurable sets.

Recall that a set $A\subseteq\mathbb{R}^r$ is **bounded** if there is an M such that $A\subseteq B(\mathbf{0},M)=\{x:\|x\|\le M\}$; equivalently, if $\sup_{x\in A}|\xi_j|<\infty$ for every $j\le r$ (writing $x=(\xi_1,\dots,\xi_r)$, as in §115).

134F Proposition (a) If $A\subseteq\mathbb{R}^r$ is any set, then

$$\mu^*A=\inf\{\mu G: G\text{ is open}, G\supseteq A\}=\min\{\mu H: H\text{ is Borel}, H\supseteq A\}.$$

(b) If $E\subseteq\mathbb{R}^r$ is measurable, then

$$\mu E=\sup\{\mu F: F\text{ is closed and bounded}, F\subseteq E\},$$

and there are Borel sets H_1, H_2 such that $H_1\subseteq E\subseteq H_2$ and

$$\mu(H_2\setminus H_1)=\mu(H_2\setminus E)=\mu(E\setminus H_1)=0.$$

(c) If $A\subseteq\mathbb{R}^r$ is any set, then A has a measurable envelope which is a Borel set.

(d) If f is a Lebesgue measurable real-valued function defined on a subset of $\mathbb{R}^r$, then there is a conegligible Borel set $H\subseteq\mathbb{R}^r$ such that $f\restriction H$ is Borel measurable.

proof (a)(i) First note that if $I\subseteq\mathbb{R}^r$ is a half-open interval, and $\epsilon>0$, then either $I=\emptyset$ is already open, or I is expressible as $[a,b[$ where $a=(\alpha_1,\dots,\alpha_r)$, $b=(\beta_1,\dots,\beta_r)$ and $\alpha_i<\beta_i$ for every i. In the latter case, $G=\,]a-\epsilon(b-a),b[$ is an open set including I, and

$$\mu G=\prod_{i=1}^r(1+\epsilon)(\beta_i-\alpha_i)=(1+\epsilon)^r\mu I,$$

by the formula in 114G/115G.

(ii) Now, given $\epsilon>0$, there is a sequence $\langle I_n\rangle_{n\in\mathbb{N}}$ of half-open intervals, covering A, such that $\sum_{n=0}^\infty\mu I_n\le\mu^*A+\epsilon$. For each n, let $G_n\supseteq I_n$ be an open set of measure at most $(1+\epsilon)^r\mu I_n$. Then $G=\bigcup_{n\in\mathbb{N}} G_n$ is open (1A2Bd), and $A\subseteq G$; also

$$\mu G\le\sum_{n=0}^\infty\mu G_n\le(1+\epsilon)^r\sum_{n=0}^\infty\mu I_n\le(1+\epsilon)^r(\mu^*A+\epsilon).$$

As ϵ is arbitrary, $\mu^*A \ge \inf\{\mu G : G \text{ is open}, G \supseteq A\}$.

(iii) Next, using (ii), we can choose for each $n \in \mathbb{N}$ an open set $G_n \supseteq A$ such that $\mu G_n \le \mu^*A + 2^{-n}$. Set $H_0 = \bigcap_{n\in\mathbb{N}} G_n$; then H_0 is a Borel set, $A \subseteq H_0$, and

$$\mu H_0 \le \inf_{n\in\mathbb{N}} \mu G_n \le \mu^* A.$$

(iv) On the other hand, we surely have $\mu^*A \le \mu^*H = \mu H$ for every Borel set $H \supseteq A$. So we must have

$$\mu^*A \le \inf\{\mu G : G \text{ is open}, G \supseteq A\},$$

and

$$\mu^*A = \mu H_0 = \min\{\mu H : H \text{ is Borel}, H \supseteq A\}.$$

(b)(i) For each $n \in \mathbb{N}$, set $E_n = E \cap B(\mathbf{0}, n)$. Let $G_n \supseteq E_n$ be an open set of measure at most $\mu E_n + 2^{-n}$; then (because $\mu B(\mathbf{0}, n) < \infty$) $\mu(G_n \setminus E_n) \le 2^{-n}$. Now, for each n, set $G'_n = \bigcup_{m\ge n} G_m$; then G'_n is open, $E = \bigcup_{m\ge n} E_m \subseteq G'_n$, and

$$\mu(G'_n \setminus E) \le \sum_{m=n}^{\infty} \mu(G_m \setminus E) \le \sum_{m=n}^{\infty} \mu(G_m \setminus E_m) \le \sum_{m=n}^{\infty} 2^{-m} = 2^{-n+1}.$$

Setting $H_2 = \bigcap_{n\in\mathbb{N}} G_n$, we see that H_2 is a Borel set including E and that $\mu(H_2 \setminus E) = 0$.

(ii) Repeating the argument of (i) with $\mathbb{R}^r \setminus E$ in place of E, we obtain a Borel set $\tilde{H}_2 \supseteq \mathbb{R}^r \setminus E$ such that $\mu(\tilde{H}_2 \setminus (\mathbb{R}^r \setminus E)) = 0$; now $H_1 = \mathbb{R}^r \setminus \tilde{H}_2$ is a Borel set included in E and

$$\mu(E \setminus H_1) = \mu(\tilde{H}_2 \setminus (\mathbb{R}^r \setminus E)) = 0.$$

Of course we now also have

$$\mu(H_2 \setminus H_1) = \mu(H_2 \setminus E) + \mu(E \setminus H_1) = 0.$$

(iii) Again using the idea of (i), there is for each $n \in \mathbb{N}$ an open set $\tilde{G}_n \supseteq B(\mathbf{0}, n) \setminus E$ such that

$$\mu(\tilde{G}_n \cap E_n) \le \mu(\tilde{G}_n \setminus (B(\mathbf{0}, n) \setminus E)) \le 2^{-n}.$$

Set

$$F_n = B(\mathbf{0}, n) \setminus \tilde{G}_n = B(\mathbf{0}, n) \cap (\mathbb{R}^r \setminus \tilde{G}_n);$$

then F_n is closed (1A2Fd) and bounded and $F_n \subseteq E_n \subseteq E$. Also

$$\mu E_n = \mu F_n + \mu(E_n \setminus F_n) = \mu F_n + \mu(\tilde{G}_n \cap E_n) \le \mu F_n + 2^{-n}.$$

So

$$\mu E = \lim_{n\to\infty} \mu E_n \le \sup_{n\in\mathbb{N}} \mu F_n \le \sup\{\mu F : F \text{ is closed and bounded}, F \subseteq E\},$$

and

$$\mu E = \sup\{\mu F : F \text{ is closed and bounded}, F \subseteq E\}.$$

(c) Let E be any measurable envelope of A (132Ef), and $H \supseteq E$ a Borel set such that $\mu(H \setminus E) = 0$; then $\mu^*(F \cap A) = \mu(F \cap E) = \mu(F \cap H)$ for every measurable set F, so H is a measurable envelope of A.

(d) Set $D = \operatorname{dom} f$ and write $\mathcal{B}$ for the σ-algebra of Borel sets. For each rational number q, let E_q be a measurable set such that $\{x : f(x) \le q\} = E_q \cap D$. Let H_q, $H'_q \in \mathcal{B}$ be such that $H_q \subseteq E_q \subseteq H'_q$ and $\mu(H'_q \setminus H_q) = 0$. Let H be the conegligible Borel set $\mathbb{R}^r \setminus \bigcup(H'_q \setminus H_q)$. Then

$$\{x : (f\restriction H)(x) \le q\} = H \cap E_q \cap D = H_q \cap D \cap H$$

belongs to the subspace σ-algebra $\mathcal{B}(D)$ for every $q \in \mathbb{Q}$. For irrational $a \in \mathbb{R}$, set $H_a = \bigcap_{q\in\mathbb{Q}, q\ge a} H_q$; then $H_a \in \mathcal{B}$, and

$$\{x : (f\restriction H)(x) \le a\} = H_a \cap \operatorname{dom}(f\restriction H).$$

Thus $f\restriction H$ is Borel measurable.

Remark The emphasis on closed *bounded* sets in part (b) of this proposition is on account of their important topological properties, in particular, the fact that they are 'compact'. This is one of the most important facts about Lebesgue measure, as will appear in Volume 4. I will discuss 'compactness' briefly in §2A2 of Volume 2.

134G The Cantor set One of the purposes of the theory of Lebesgue measure and integration is to study rather more irregular sets and functions than can be dealt with by more primitive methods. In the next few paragraphs I discuss *measurable* sets and functions which from the point of view of the present theory are amenable without being trivial. From now on, μ will be Lebesgue measure on $\mathbb{R}$.

(a) The 'Cantor set' $C \subseteq [0,1]$ is defined as the intersection of a sequence $\langle C_n\rangle_{n\in\mathbb{N}}$ of sets, constructed as follows. $C_0 = [0,1]$. Given that C_n consists of 2^n disjoint closed intervals each of length 3^{-n}, take each of these intervals and delete the middle third to produce two closed intervals each of length 3^{-n-1}; take C_{n+1} to be the union of the 2^{n+1} closed intervals so formed, and continue. Observe that $\mu C_n = (\frac{2}{3})^n$ for each n.

Approaching the Cantor set

The **Cantor set** is $C = \bigcap_{n\in\mathbb{N}} C_n$. Its measure is

$$\mu C = \lim_{n\to\infty} \mu C_n = \lim_{n\to\infty}(\frac{2}{3})^n = 0.$$

(b) Each C_n can also be described as the set of real numbers expressible as $\sum_{j=1}^{\infty} 3^{-j}\epsilon_j$ where every ϵ_j is either 0, 1 or 2, and $\epsilon_j \ne 1$ for $j \le n$. Consequently C itself is the set of numbers expressible as $\sum_{j=1}^{\infty} 3^{-j}\epsilon_j$ where every ϵ_j is either 0 or 2; that is, the set of numbers between 0 and 1 expressible in ternary form without 1's. The expression in each case will be unique, so we have a bijection $\phi : \{0,1\}^{\mathbb{N}} \to C$ defined by writing

$$\phi(z) = \frac{2}{3}\sum_{j=0}^{\infty} 3^{-j} z(j)$$

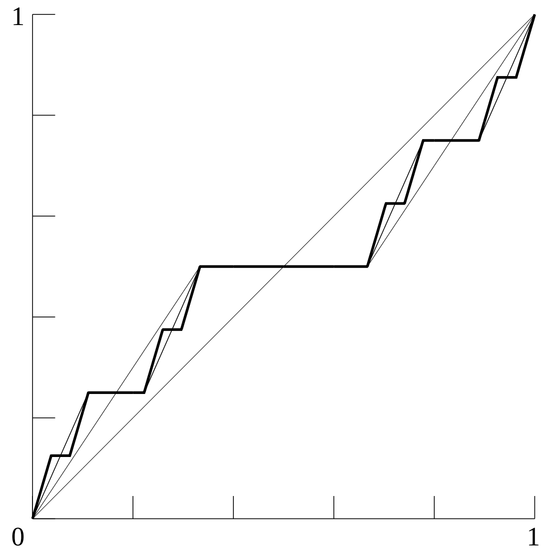
for every $z \in \{0,1\}^{\mathbb{N}}$.

134H The Cantor function Continuing from 134G, we have the following construction.

(a) For each $n \in \mathbb{N}$ we define a function $f_n : [0,1] \to [0,1]$ by setting

$$f_n(x) = (\frac{3}{2})^n \mu(C_n \cap [0,x])$$

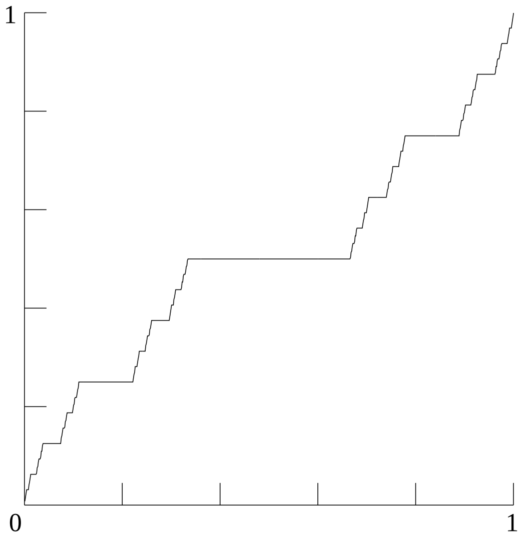
for each $x \in [0,1]$. Because C_n is just a finite union of intervals, f_n is a polygonal function, with $f_n(0) = 0$, $f_n(1) = 1$; f_n is constant on each of the $2^n - 1$ open intervals composing $[0,1] \setminus C_n$, and rises with slope $(\frac{3}{2})^n$ on each of the 2^n closed intervals composing C_n.



Approaching the Cantor function: the functions f_0, f_1, f_2, $\boldsymbol{f_3}$

If the jth interval of C_n, counting from the left, is $[a_{nj}, b_{nj}]$, then $f_n(a_{nj}) = 2^{-n}(j-1)$ and $f_n(b_{nj}) = 2^{-n}j$. Also, $a_{nj} = a_{n+1,2j-1}$ and $b_{nj} = b_{n+1,2j}$; hence, or otherwise, $f_{n+1}(a_{nj}) = f_n(a_{nj})$ and $f_{n+1}(b_{nj}) = f_n(b_{nj})$, and f_{n+1} agrees with f_n on all the endpoints of the intervals of C_n, and therefore on $[0,1] \setminus C_n$.

Within any particular interval $[a_{nj}, b_{nj}]$ of C_n, the greatest difference between $f_n(x)$ and $f_{n+1}(x)$ is at the new endpoints within that interval, viz., $b_{n+1,2j-1}$ and $a_{n+1,2j}$; and the magnitude of the difference is $\frac{1}{6}2^{-n}$ (because, for instance, $f_n(b_{n+1,2j-1}) = \frac{2}{3}f_n(a_{nj}) + \frac{1}{3}f_n(b_{nj})$, while $f_{n+1}(b_{n+1,2j-1}) = \frac{1}{2}f_n(a_{nj}) + \frac{1}{2}f_n(b_{nj})$). Thus we have $|f_{n+1}(x) - f_n(x)| \le \frac{1}{6}2^{-n}$ for every $n \in \mathbb{N}$, $x \in [0,1]$. Because $\sum_{n=0}^{\infty} \frac{1}{6}2^{-n} < \infty$, $\langle f_n \rangle_{n\in\mathbb{N}}$ is uniformly convergent to a function $f : [0,1] \to [0,1]$, and f will be continuous. f is the **Cantor function** or **Devil's Staircase**.



The Cantor function

(b) Because every f_n is non-decreasing, so is f. If $x \in [0,1] \setminus C$, there is an n such that $x \in [0,1] \setminus C_n$; let I be the open interval of $[0,1] \setminus C_n$ containing x; then f_{m+1} agrees on I with f_m for every $m \ge n$, so f agrees on I with f_n, and f is constant on I. Thus, in particular, the derivative $f'(x)$ exists and is 0 for every $x \in [0,1] \setminus C$; so f' is zero almost everywhere in $[0,1]$. Also, of course, $f(0) = 0$ and $f(1) = 1$, because $f_n(0) = 0$, $f_n(1) = 1$ for every n. It follows that $f : [0,1] \to [0,1]$ is surjective (by the Intermediate Value Theorem).

(c) Let $\phi : \{0,1\}^{\mathbb{N}} \to C$ be the function described in 134Gb. Then $f(\phi(z)) = \frac{1}{2}\sum_{j=0}^{\infty} 2^{-j}z(j)$ for every $z \in \{0,1\}^{\mathbb{N}}$. **P** Fix $z = (\zeta_0, \zeta_1, \zeta_2, \dots)$ in $\{0,1\}^{\mathbb{N}}$, and for each n take I_n to be the component interval of C_n containing $\phi(z)$. Then I_{n+1} will be the left-hand third of I_n if $\zeta_n = 0$ and the right-hand third if $\zeta_n = 1$. Taking a_n to be the left-hand endpoint of I_n, we see that

$$a_{n+1} = a_n + \frac{2}{3}3^{-n}\zeta_n, \quad f_{n+1}(a_{n+1}) = f_n(a_n) + \frac{1}{2}2^{-n}\zeta_n$$

for each n. Now

$$\phi(z) = \lim_{n\to\infty} a_n, \quad f(\phi(z)) = \lim_{n\to\infty} f(a_n) = \lim_{n\to\infty} f_n(a_n) = \frac{1}{2}\sum_{j=0}^{\infty} 2^{-j}\zeta_j,$$

as claimed. **Q**

In particular, $f[C] = [0,1]$. **P** Any $x \in [0,1]$ is expressible as $\sum_{j=0}^{\infty} 2^{-j-1}z(j) = f(\phi(z))$ for some $z \in \{0,1\}^{\mathbb{N}}$. **Q**

134I The Cantor function modified I continue the argument of 134G-134H.

(a) Consider the formula

$$g(x) = \frac{1}{2}(x + f(x)),$$

where f is the Cantor function, as defined in 134H; this defines a continuous function $g : [0,1] \to [0,1]$ which is strictly increasing (because f is non-decreasing) and has $g(0) = 0$, $g(1) = 1$; consequently, by the Intermediate Value Theorem, g is bijective, and its inverse $g^{-1} : [0,1] \to [0,1]$ is continuous.

Now $g[C]$ is a closed set and $\mu g[C] = \frac{1}{2}$. **P** Because g is a permutation of the points of $[0,1]$, $[0,1]\setminus g[C] = g[\,[0,1]\setminus C]$. For each of the open intervals $I_{nj} = \,]b_{nj}, a_{n,j+1}[$ making up $[0,1]\setminus C_n$, we see that $g[I_{nj}] = \,]g(b_{nj}), g(a_{n,j+1})[$ has length just half the length of I_{nj}. Consequently $g[\,[0,1]\setminus C] = \bigcup_{n\ge1, 1\le j<2^n} g[I_{nj}]$ is open, and

$$\begin{aligned}\mu(g[\,[0,1]\setminus C_n]) &= \sum_{j=1}^{2^n-1} g(a_{n,j+1}) - g(b_{nj}) = \frac{1}{2}\sum_{j=1}^{2^n-1} a_{n,j+1} - b_{nj} \\ &= \frac{1}{2}\mu([0,1]\setminus C_n) = \frac{1}{2}(1 - (\frac{2}{3})^n)\end{aligned}$$

(134Ga). Because $\langle [0,1]\setminus C_n\rangle_{n\in\mathbb{N}}$ is an increasing sequence of sets with union $[0,1]\setminus C$,

$$\mu g([\,[0,1]\setminus C]) = \lim_{n\to\infty} \mu g([\,[0,1]\setminus C_n]) = \frac{1}{2}.$$

So $g[C] = [0,1]\setminus g[\,[0,1]\setminus C]$ is closed and $\mu g[C] = \frac{1}{2}$. **Q**

(b) By 134D there is a set $D \subseteq \mathbb{R}$ such that

$$\mu^*(g[C]\cap D) = \mu^*(g[C]\setminus D) = \mu g[C] = \frac{1}{2};$$

set $A = g[C]\cap D$. Of course A cannot be measurable, since $\mu^*A + \mu^*(g[C]\setminus A) > \mu g[C]$. However, $g^{-1}[A] \subseteq C$ must be measurable, because $\mu^*C = 0$. This means that if we set $h = \chi(g^{-1}[A]) : [0,1] \to \mathbb{R}$, then h is measurable; but $hg^{-1} = \chi A : [0,1] \to \mathbb{R}$ is not.

Thus **the composition of a measurable function with a continuous function need not be measurable**. Contrast this with 121Eg.

134J More examples I think it is worth taking the space to spell out two more of the basic examples of Lebesgue measurable set in detail.

(a) As already observed in 114G, every countable subset of $\mathbb{R}$ is negligible. In particular, $\mathbb{Q}$ is negligible (111Eb). We can say more. Let $\langle q_n\rangle_{n\in\mathbb{N}}$ be a sequence running over $\mathbb{Q}$, and for each $n \in \mathbb{N}$ set

$$I_n = \,]q_n - 2^{-n}, q_n + 2^{-n}[,$$

$$G_n = \bigcup_{k\ge n} I_k.$$

Then G_n is an open set of measure at most $\sum_{k=n}^{\infty} 2\cdot 2^{-k} = 4\cdot 2^{-n}$, and it contains all but finitely many points of $\mathbb{Q}$, so is dense (that is, meets every non-trivial interval). Set $F_n = \mathbb{R}\setminus G_n$; then F_n is closed, $\mu(\mathbb{R}\setminus F_n) \le 4/2^n$, but F_n does not contain q_k for any $k \ge n$, so F_n cannot include any non-trivial interval. Observe that $\langle G_n\rangle_{n\in\mathbb{N}}$ is non-increasing so $\langle F_n\rangle_{n\in\mathbb{N}}$ is non-decreasing.

(b) We can elaborate the above construction, as follows. There is a measurable set $E \subseteq \mathbb{R}$ such that $\mu(I\cap E) > 0$ and $\mu(I\setminus E) > 0$ for every non-trivial interval $I \subseteq \mathbb{R}$. **P** First note that if k, $n \in \mathbb{N}$, there is a $j \ge n$ such that $q_j \in I_k$, so that $I_k \cap I_j \ne \emptyset$ and $\mu(I_k \setminus F_n) > 0$. Now there must be an $l > n$ such that $\mu G_l < \mu(I_k\setminus F_n)$, so that

$$\mu(I_k \cap F_l \setminus F_n) = \mu((I_k \setminus F_n) \setminus G_l) > 0.$$

Choose $n_0 < n_1 < n_2 < \ldots$ as follows. Start with $n_0 = 0$. Given n_{2k}, where $k \in \mathbb{N}$, choose n_{2k+1}, n_{2k+2} such that

$$\mu(I_k \cap F_{n_{2k+1}} \setminus F_{n_{2k}}) > 0, \quad \mu(I_k \cap F_{n_{2k+2}} \setminus F_{n_{2k+1}}) > 0.$$

Continue.

On completing the induction, set

$$E = \bigcup_{k\in\mathbb{N}} F_{n_{2k+1}} \setminus F_{n_{2k}}, \quad H = \bigcup_{k\in\mathbb{N}} F_{n_{2k+2}} \setminus F_{n_{2k+1}}.$$

Because $\langle F_k\rangle_{k\in\mathbb{N}}$ is non-decreasing, $E \cap H = \emptyset$. If $k \in \mathbb{N}$, $E \cap I_k$ and $H \cap I_k$ both have positive measure.

Now suppose that $I \subseteq \mathbb{R}$ is an interval with more than one point; suppose that a, $b \in I$ and $a < b$. Then there is an $m \in \mathbb{N}$ such that $4 \cdot 2^{-m} \le b - a$; now there is a $k \ge m$ such that $q_k \in [a + 2^{-m}, b - 2^{-m}]$, so that $I_k \subseteq I$ and

$$\mu(I \cap E) \ge \mu(E \cap I_k) > 0, \quad \mu(I \setminus E) \ge \mu(H \cap I_k) > 0. \ \mathbf{Q}$$

(c) This shows that E and its complement are measurable sets which are not merely both dense (like $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$), but 'essentially' dense in that they meet every non-empty open interval in a set of positive measure, so that (for instance) $E \setminus A$ is dense for every negligible set A.

***134K Riemann integration** I have tried, in writing this book, to assume as little prior knowledge as possible. In particular, it is not necessary to have studied Riemann integration. Nevertheless, if you have worked through the basic theory of the Riemann integral – which is, indeed, not only a splendid training in the techniques of ϵ-δ analysis, but also a continuing source of ideas for the subject – you will, I hope, wish to connect it with the material we are looking at here; both because you will not want to feel that your labour has been wasted, and because you have probably developed a number of intuitions which will continue to be valuable, if suitably adapted to the new context. I therefore give a brief account of the relationship between the Riemann and Lebesgue methods of integration on the real line.

(a) There are many ways of describing the Riemann integral; I choose one of the popular ones. If $[a, b]$ is a non-trivial closed interval in $\mathbb{R}$, then I say that a **dissection** of $[a, b]$ is a finite list $D = (a_0, a_1, \ldots , a_n)$, where $n \ge 1$, such that $a = a_0 < a_1 < \ldots < a_n = b$. If now f is a real-valued function defined (at least) on $[a, b]$ and bounded on $[a, b]$, the **upper sum** and **lower sum** of f on $[a, b]$ derived from D are

$$S_D(f) = \sum_{i=1}^n (a_i - a_{i-1}) \sup_{x\in]a_{i-1},a_i[} f(x),$$

$$s_D(f) = \sum_{i=1}^n (a_i - a_{i-1}) \inf_{x\in]a_{i-1},a_i[} f(x).$$

You have to prove that if D and D' are two dissections of $[a, b]$, then $s_D(f) \le S_{D'}(f)$. Now define the **upper Riemann integral** and **lower Riemann integral** of f to be

$$U_{[a,b]}(f) = \inf\{S_D(f) : D \text{ is a dissection of } [a, b]\},$$

$$L_{[a,b]}(f) = \sup\{s_D(f) : D \text{ is a dissection of } [a, b]\}.$$

Check that $L_{[a,b]}(f)$ is necessarily less than or equal to $U_{[a,b]}(f)$. Finally, declare f to be **Riemann integrable over** $[a, b]$ if $U_{[a,b]}(f) = L_{[a,b]}(f)$, and in this case take the common value to be the **Riemann integral** $\mathrm{R}\!\!\int_a^b f$ of f over $[a, b]$.

(b) If $f : [a, b] \to \mathbb{R}$ is Riemann integrable, it is Lebesgue integrable, with the same integral. $\mathbf{P}$ For any dissection $D = (a_0, \ldots , a_n)$ of $[a, b]$, define g_D, $h_D : [a, b] \to \mathbb{R}$ by saying

$$g_D(x) = \inf\{f(y) : y \in]a_{i-1}, a_i[\} \text{ if } a_{i-1} < x < a_i, \quad g_D(a_i) = f(a_i) \text{ for each } i,$$

$$h_D(x) = \sup\{f(y) : y \in]a_{i-1}, a_i[\} \text{ if } a_{i-1} < x < a_i, \quad h_D(a_i) = f(a_i) \text{ for each } i.$$

Then g_D and h_D are constant on each interval $]a_{i-1}, a_i[$, so all sets $\{x : g_D(x) \le c\}$, $\{x : h_D(x) \le c\}$ are finite unions of intervals, and g_D and h_D are measurable; moreover,

$$\int g_D d\mu = s_D(f), \quad \int h_D d\mu = S_D(f).$$

Consequently

$$\begin{aligned} \mathrm{R}\!\!\int_a^b f = L_{[a,b]}(f) = \sup_D \int g_D d\mu \le \underline{\int} f d\mu \\ \le \overline{\int} f d\mu \le \inf_D \int h_D d\mu = U_{[a,b]}(f) = \mathrm{R}\!\!\int_a^b f, \end{aligned}$$

and $\overline{\int} f d\mu = \underline{\int} f d\nu = \mathcal{R}\!\!\int_a^b f$, so that $\int f d\mu$ exists and is equal to $\mathcal{R}\!\!\int_a^b f$ (133Jd). **Q**

(c) The discussion above is of the 'proper' Riemann integral, of bounded functions on bounded intervals. For unbounded functions and unbounded intervals, one uses various forms of 'improper' integral; for instance, the improper Riemann integral $\int_0^\infty \frac{\sin x}{x} dx$ is taken to be $\lim_{a\to\infty} \int_0^a \frac{\sin x}{x} dx$, while $\int_0^1 \ln x\,dx$ is taken to be $\lim_{a\downarrow 0} \int_a^1 \ln x\,dx$. Of these, the second exists as a Lebesgue integral, but the first does not, because $\int_0^\infty |\frac{\sin x}{x}| dx = \infty$. The power of the Lebesgue integral to deal directly with 'absolutely integrable' unbounded functions on unbounded domains means that what one might call 'conditionally integrable' functions are pushed into the background of the theory. In Chapter 48 of Volume 4 I will discuss the general theory of such functions, but for the time being I will deal with them individually, on the rare occasions when they arise.

***134L** There is in fact a beautiful characterisation of the Riemann integrable functions, as follows.

Proposition If $a < b$ in $\mathbb{R}$, a bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable iff it is continuous almost everywhere in $[a, b]$.

proof (a) Suppose that f is Riemann integrable. For each $x \in [a, b]$, set

$$g(x) = \sup_{\delta>0} \inf_{y\in[a,b],|y-x|\le\delta} f(y),$$

$$h(x) = \inf_{\delta>0} \sup_{y\in[a,b],|y-x|\le\delta} f(y),$$

so that f is continuous at x iff $g(x) = h(x)$. We have $g \le f \le h$, so if D is any dissection of $[a, b]$ then $S_D(g) \le S_D(f) \le S_D(h)$ and $s_D(g) \le s_D(f) \le s_D(h)$. But in fact $S_D(f) = S_D(h)$ and $s_D(g) = s_D(f)$, because on any open interval $]c, d[\subseteq [a, b]$ we must have

$$\inf_{x\in]c,d[} g(x) = \inf_{x\in]c,d[} f(x), \quad \sup_{x\in]c,d[} f(x) = \sup_{x\in]c,d[} h(x).$$

It follows that

$$L_{[a,b]}(f) = L_{[a,b]}(g) \le U_{[a,b]}(g) \le U_{[a,b]}(f),$$

$$L_{[a,b]}(f) \le L_{[a,b]}(h) \le U_{[a,b]}(h) = U_{[a,b]}(f).$$

Because f is Riemann integrable, both g and h must be Riemann integrable, with integrals equal to $\mathcal{R}\!\!\int_a^b f$. By 134Kb, they are both Lebesgue integrable, with the same integral. But $g \le h$, so $g =_{\text{a.e.}} h$, by 122Rd. Now f is continuous at any point where g and h agree, so f is continuous a.e.

(b) Now suppose that f is continuous a.e. For each $n \in \mathbb{N}$, let D_n be the dissection of $[a, b]$ into 2^n equal portions. Set

$$h_n(x) = \sup_{y\in]c,d[} f(y), \quad g_n(x) = \inf_{y\in]c,d[} f(y)$$

if $]c, d[$ is an open interval of D_n containing x; for definiteness, say $h_n(x) = g_n(x) = f(x)$ if x is one of the points of the list D_n. Then $\langle g_n\rangle_{n\in\mathbb{N}}$, $\langle h_n\rangle_{n\in\mathbb{N}}$ are, respectively, increasing and decreasing sequences of functions, each function constant on each of a finite family of intervals covering $[a, b]$; and $s_{D_n}(f) = \int g_n d\mu$, $S_{D_n}(f) = \int h_n d\mu$. Next,

$$\lim_{n\to\infty} g_n(x) = \lim_{n\to\infty} h_n(x) = f(x)$$

at any point x at which f is continuous; so $f =_{\text{a.e.}} \lim_{n\to\infty} g_n =_{\text{a.e.}} \lim_{n\to\infty} h_n$. By Lebesgue's Dominated Convergence Theorem (123C),

$$\lim_{n\to\infty} \int g_n d\mu = \int f d\mu = \lim_{n\to\infty} \int h_n d\mu;$$

but this means that

$$L_{[a,b]}(f) \ge \int f d\mu \ge U_{[a,b]}(f),$$

so these are all equal and f is Riemann integrable.

134X Basic exercises >(a) Show that if f is an integrable real-valued function on $\mathbb{R}^r$, then $\int f(x + a)dx$ exists and is equal to $\int f$ for every $a \in \mathbb{R}^r$. (*Hint*: start with simple functions f.)

(b) More generally, show that if $E \subseteq \mathbb{R}^r$ is measurable and f is a real-valued function which is integrable over E in the sense of 131D, then $\int_{E-a} f(x + a)dx$ exists and is equal to $\int_E f$ for every $a \in \mathbb{R}^r$.

(c) Show that if $C \subseteq \mathbb{R}$ is any non-negligible set, it has a non-measurable subset. (*Hint*: use the method of 134B, taking the relation $\sim$ on a suitable bounded subset of C in place of $[\mathbf{0}, \mathbf{1}[$.)

>(d) Let ν_g be a Lebesgue-Stieltjes measure on $\mathbb{R}$, constructed as in 114Xa from a non-decreasing function $g : \mathbb{R} \to \mathbb{R}$, and Σ_g its domain. (See also 132Xg.) Show that

(i) if $A \subseteq \mathbb{R}$ is any set, then

$$\begin{aligned}\nu_g^* A &= \inf\{\nu_g G : G \text{ is open, } G \supseteq A\}\\ &= \min\{\nu_g H : H \text{ is Borel, } H \supseteq A\};\end{aligned}$$

(ii) if $E \in \Sigma_g$, then

$$\nu_g E = \sup\{\nu_g F : F \text{ is closed and bounded, } F \subseteq E\},$$

and there are Borel sets H_1, H_2 such that $H_1 \subseteq E \subseteq H_2$ and $\nu_g(H_2 \setminus H_1) = \nu_g(H_2 \setminus E) = \nu_g(E \setminus H_1) = 0$;

(iii) if $A \subseteq \mathbb{R}$ is any set, then A has a measurable envelope which is a Borel set;

(iv) if f is a Σ_g-measurable real-valued function defined on a subset of $\mathbb{R}$, then there is a ν_g-conegligible Borel set $H \subseteq \mathbb{R}$ such that $f{\restriction}H$ is Borel measurable.

(e) Let $E \subseteq \mathbb{R}^r$ be a measurable set, and $\epsilon > 0$. (i) Show that there is an open set $G \supseteq E$ such that $\mu(G \setminus E) \le \epsilon$. (*Hint*: apply 134Fa to each set $E \cap B(\mathbf{0}, n)$.) (ii) Show that there is a closed set $F \subseteq E$ such that $\mu(E \setminus F) \le \epsilon$.

(f) Let $C \subseteq [0,1]$ be the Cantor set. Show that $\{x + y : x,\, y \in C\} = [0,2]$ and $\{x - y : x,\, y \in C\} = [-1,1]$.

(g) Let f, g be functions from $\mathbb{R}$ to itself. Show that (i) if f and g are both Borel measurable, so is their composition fg (ii) if f is Borel measurable and g is Lebesgue measurable, then fg is Lebesgue measurable (iii) if f is Lebesgue measurable and g is Borel measurable, then fg need not be Lebesgue measurable.

(h) Show that for any integer $r \ge 1$ there is a measurable set $E \subseteq \mathbb{R}^r$ such that E and $\mathbb{R}^r \setminus E$ both meet every non-empty open interval in a set of strictly positive measure.

(i) Give $[0,1]$ its subspace measure. (i) Show that there is a disjoint sequence $\langle A_n \rangle_{n\in\mathbb{N}}$ of subsets of $[0,1]$ all of outer measure 1. (ii) Show that there is a function $f : [0,1] \to \,]0,1[$ such that $\underline{\int} f = 0$ and $\overline{\int} f = 1$.

(j) Let f be a measurable real function and g a real function such that $\operatorname{dom} g \setminus \operatorname{dom} f$ and $\{x : x \in \operatorname{dom} g \cap \operatorname{dom} f,\ g(x) \ne f(x)\}$ are both negligible. Show that g is measurable.

134Y Further exercises (a) Fix $c > 0$. For $A \subseteq \mathbb{R}^r$ set $cA = \{cx : x \in A\}$. (i) Show that $\mu^*(cA) = c^r \mu^* A$ for every $A \subseteq \mathbb{R}^r$. (ii) Show that cE is measurable for every measurable $E \subseteq \mathbb{R}^r$.

(b) Let $\langle f_{mn} \rangle_{m,n\in\mathbb{N}}$, $\langle f_m \rangle_{m\in\mathbb{N}}$, f be real-valued measurable functions defined almost everywhere in $\mathbb{R}^r$ and such that $f_m =_{\text{a.e.}} \lim_{n\to\infty} f_{mn}$ for each m and $f =_{\text{a.e.}} \lim_{m\to\infty} f_m$. Show that there is a sequence $\langle n_k \rangle_{k\in\mathbb{N}}$ such that $f =_{\text{a.e.}} \lim_{k\to\infty} f_{k,n_k}$. (*Hint*: take n_k such that the measure of $\{x : \|x\| \le k,\ |f_k(x) - f_{k,n_k}(x)| \ge 2^{-k}\}$ is at most 2^{-k} for each k.)

(c) Let f be a measurable real-valued function defined almost everywhere in $\mathbb{R}^r$. Show that there is a sequence $\langle f_n \rangle_{n\in\mathbb{N}}$ of continuous functions converging to f almost everywhere. (*Hint*: Deal successively with the cases (i) $f = \chi I$ where I is a half-open interval (ii) $f = \chi(\bigcup_{j\le n} I_j)$ where $I_0, \dots, I_n$ are disjoint half-open intervals (iii) $f = \chi E$ where E is a measurable set of finite measure (iv) f is a simple function (v) general f, using 134Yb at steps (iii) and (v).)

(d) Let f be a real-valued function defined on a subset of $\mathbb{R}^r$. Show that the following are equiveridical: (i) f is measurable (ii) whenever $E \subseteq \mathbb{R}^r$ is measurable and $\mu E > 0$, there is a measurable set $F \subseteq E$ such that $\mu F > 0$ and $f{\restriction}F$ is continuous (iii) whenever $E \subseteq \mathbb{R}^r$ is measurable and $\gamma < \mu E$, there is a measurable $F \subseteq E$ such that $\mu F \ge \gamma$ and $f{\restriction}F$ is continuous. (*Hint*: for (i)$\Rightarrow$(iii), use 134Yc and 131Ya; for (ii)$\Rightarrow$(i) use 121D. This is a version of **Lusin's theorem**.)

(e) Let ν be a measure on $\mathbb{R}$ which is translation-invariant in the sense of 134Ab, and such that $\nu[0,1]$ is defined and equal to 1. Show that ν agrees with Lebesgue measure on the Borel sets of $\mathbb{R}$. (*Hint*: Show first that $[a,1]$ belongs to the domain of ν for every $a \in [0,1]$, and hence that every half-open interval of length at most 1 belongs to the domain of ν; show that $\nu\,[a, a+2^{-n}[\, = 2^{-n}$ for every $a \in \mathbb{R}$, $n \in \mathbb{N}$, and hence that $\nu\,[a,b[\, = b-a$ whenever $a < b$.)

(f) Let ν be a measure on $\mathbb{R}^r$ which is translation-invariant in the sense of 134Ab, where $r > 1$, and such that $\nu[\mathbf{0}, \mathbf{1}]$ is defined and equal to 1. Show that ν agrees with Lebesgue measure on the Borel sets of $\mathbb{R}^r$.

(g) Show that if f is any real-valued integrable function on $\mathbb{R}$, and $\epsilon > 0$, there is a continuous function $g : \mathbb{R} \to \mathbb{R}$ such that $\{x : g(x) \neq 0\}$ is bounded and $\int |f - g| \le \epsilon$. (*Hint*: show that the set Φ of functions f with this property satisfies the conditions of 122Yb.)

(h) Repeat 134Yg for real-valued integrable functions on $\mathbb{R}^r$, where $r > 1$.

(i) Repeat 134Fd, 134Xa, 134Xb, 134Yb, 134Yc, 134Yd, 134Yg and 134Yh for complex-valued functions.

(j) Show that if $G \subseteq \mathbb{R}^r$ is open and not empty, it is expressible as a disjoint union of a sequence of half-open intervals each of the form $\{x : 2^{-m} n_i \le \xi_i < 2^{-m}(n_i + 1)$ for every $i \le r\}$ where $m \in \mathbb{N}$, $n_1, \dots , n_r \in \mathbb{Z}$.

(k) Show that a set $E \subseteq \mathbb{R}^r$ is Lebesgue negligible iff there is a sequence $\langle C_n \rangle_{n\in\mathbb{N}}$ of hypercubes in $\mathbb{R}^r$ such that $E \subseteq \bigcap_{n\in\mathbb{N}} \bigcup_{k\ge n} C_k$ and $\sum_{k=0}^{\infty} (\operatorname{diam} C_k)^r < \infty$, writing $\operatorname{diam} C_k$ for the diameter of C_k.

(l) Show that there is a continuous function $f : [0, 1] \to [0, 1]^2$ such that $\mu_1 f^{-1}[E] = \mu_2 E$ for every measurable $E \subseteq [0, 1]^2$, writing μ_1, μ_2 for Lebesgue measure on $\mathbb{R}$, $\mathbb{R}^2$ respectively. (*Hint*: for each $n \in \mathbb{N}$, express $[0, 1]^2$ as the union of 4^n closed squares of side 2^{-n}; call the set of these squares $\mathcal{D}_n$. Construct continuous $f_n : [0, 1] \to [0, 1]^2$, families $\langle I_D \rangle_{D\in\mathcal{D}_n}$ inductively in such a way that each I_D is a closed interval of length 4^{-n} and $f_m[I_D] \subseteq D$ whenever $D \in \mathcal{D}_n$ and $m \ge n$. The induction will proceed more smoothly if you suppose that the path f_n enters each square in $\mathcal{D}_n$ at a corner and leaves at an adjacent corner. Take $f = \lim_{n\to\infty} f_n$. This is a special kind of **Peano** or **space-filling** curve.)

(m) Show that if $r \le s$ there is a continuous function $f : [0, 1]^r \to [0, 1]^s$ such that $\mu_r f^{-1}[E] = \mu_s E$ for every measurable $E \subseteq [0, 1]^s$, writing μ_r, μ_s for Lebesgue measure on $\mathbb{R}^r$, $\mathbb{R}^s$ respectively.

(n) Show that there is a continuous function $f : \mathbb{R} \to \mathbb{R}^2$ such that $\mu_1 f^{-1}[E] = \mu_2 E$ for every measurable $E \subseteq \mathbb{R}^2$, writing μ_1, μ_2 for Lebesgue measure on $\mathbb{R}$, $\mathbb{R}^2$ respectively.

(o) Show that the function $f : [0, 1] \to [0, 1]^2$ of 134Yl may be chosen in such a way that $\mu_2 f[E] = \mu_1 E$ for every Lebesgue measurable set $E \subseteq [0, 1]$. (*Hint*: using the construction suggested in 134Yl, and setting $H = f^{-1}[([0, 1] \setminus \mathbb{Q})^2]$, $f \restriction H$ will be an isomorphism between $(H, \mu_{1,H})$ and $(f[H], \mu_{2,f[H]})$, writing $\mu_{1,H}$ and $\mu_{2,f[H]}$ for the subspace measures.)

(p) Show that $\mathbb{R}$ can be expressed as the union of a disjoint sequence $\langle E_n \rangle_{n\in\mathbb{N}}$ of sets of finite measure such that $\mu(I \cap E_n) > 0$ for every non-empty open interval $I \subseteq \mathbb{R}$ and every $n \in \mathbb{N}$.

(q) Show that for any $r \ge 1$, $\mathbb{R}^r$ can be expressed as the union of a disjoint sequence $\langle E_n \rangle_{n\in\mathbb{N}}$ of sets of finite measure such that $\mu(G \cap E_n) > 0$ for every non-empty open set $G \subseteq \mathbb{R}^r$ and every $n \in \mathbb{N}$.

(r) Show that there is a disjoint sequence $\langle A_n \rangle_{n\in\mathbb{N}}$ of subsets of $\mathbb{R}$ such that $\mu^*(A_n \cap E) = \mu E$ for every measurable set E and every $n \in \mathbb{N}$. (*Remark*: in fact there is a disjoint family $\langle A_t \rangle_{t\in\mathbb{R}}$ with this property, but I think a new idea is needed for this extension. See 419I in Volume 4.)

(s) Repeat 134Yr for $\mathbb{R}^r$, where $r > 1$.

(t) Describe a Borel measurable function $f : [0, 1] \to [0, 1]$ such that $f \restriction A$ is discontinuous at every point of A whenever $A \subseteq [0, 1]$ is a set of full outer measure.

(u) Let $\langle E_n \rangle_{n\in\mathbb{N}}$ be a sequence of non-negligible measurable subsets of $\mathbb{R}^r$. Show that there is a measurable set $E \subseteq \mathbb{R}^r$ such that all the sets $E_n \cap E$, $E_n \setminus E$ are non-negligible.

134 Notes and comments Lebesgue measure enjoys an enormous variety of special properties, corresponding to the richness of the real line, with its algebraic and topological and order structures. Here I have only been able to hint at what is possible.

There are many methods of constructing non-measurable sets, all significant; the one I give in 134B is perhaps the most accessible, and shows that translation-invariance is (subject to the axiom of choice) an insuperable barrier to measuring every subset of $\mathbb{R}$.

In 134F I list some of the basic relationships between the measure and the topology of Euclidean space. Others are in 134Yc, 134Yd and 134Yg; see also 134Xd. A systematic analysis of these will take up a large part of Volume 4.

The Cantor set and function (134G-134I) form one of the basic examples in the theory. Here I present them just as an interesting design and as a counter-example to a natural conjecture. But they will reappear in three different chapters of Volume 2 as illustrations of three quite different phenomena.

The relationship between the Lebesgue and Riemann integrals goes a good deal deeper than I wish to explore just at present; the fact that the Lebesgue integral extends the Riemann integral (134Kb) is only a small part of the story, and I should be sorry if you were left with the impression that the Lebesgue integral therefore renders the Riemann integral obsolete. Without going into the details here, I hope that 134F and 134Yg make it plain that the Lebesgue integral is in some sense the canonical extension of the Riemann integral. (This, at least, I shall return to in Chapter 43.) Another way of looking at this is 134Yf; the Lebesgue integral is the basic translation-invariant integral on $\mathbb{R}^r$.

135 The extended real line

It is often convenient to allow '∞' into our formulae, and in the context of measure theory the appropriate manipulations are sufficiently consistent for it to be possible to develop a theory of the **extended real line**, the set $[-\infty,\infty]=\mathbb{R}\cup\{-\infty,\infty\}$, sometimes written $\overline{\mathbb{R}}$. I give a brief account without full proofs, as I hope that by the time this material becomes necessary to the arguments I use it will all appear thoroughly elementary.

135A The algebraic structure of $[-\infty,\infty]$ (a) If we write

$$a+\infty=\infty+a=\infty,\quad a+(-\infty)=(-\infty)+a=-\infty$$

for every $a\in\mathbb{R}$, and

$$\infty+\infty=\infty,\quad (-\infty)+(-\infty)=-\infty,$$

but refuse to define $\infty+(-\infty)$ or $(-\infty)+\infty$, we obtain a partially-defined binary operation on $[-\infty,\infty]$, extending ordinary addition on $\mathbb{R}$. This is *associative* in the sense that

if u, v, $w\in[-\infty,\infty]$ and one of $u+(v+w)$, $(u+v)+w$ is defined, so is the other, and they are then equal,

and *commutative* in the sense that

if u, $v\in[-\infty,\infty]$ and one of $u+v$, $v+u$ is defined, so is the other, and they are then equal.

It has an *identity* 0 such that $u+0=0+u=u$ for every $u\in[-\infty,\infty]$; but ∞ and $-\infty$ lack inverses.

(b) If we define

$$a\cdot\infty=\infty\cdot a=\infty,\quad a\cdot(-\infty)=(-\infty)\cdot a=-\infty$$

for real $a>0$,

$$a\cdot\infty=\infty\cdot a=-\infty,\quad a\cdot(-\infty)=(-\infty)\cdot a=\infty$$

for real $a<0$,

$$\infty\cdot\infty=(-\infty)\cdot(-\infty)=\infty,\quad (-\infty)\cdot\infty=\infty\cdot(-\infty)=-\infty,$$

$$0\cdot\infty=\infty\cdot 0=0\cdot(-\infty)=(-\infty)\cdot 0=0$$

then we obtain a binary operation on $[-\infty,\infty]$ extending ordinary multiplication on $\mathbb{R}$, which is associative and commutative and has an identity 1; 0, ∞ and $-\infty$ lack inverses.

(c) We have a *distributive law*, a little weaker than the associative and commutative laws of addition:

if u, v, $w\in[-\infty,\infty]$ and both $u(v+w)$ and $uv+uw$ are defined, then they are equal.

(But note the problems which arise with such combinations as $\infty(1+(-2))$, $0\cdot\infty+0\cdot(-\infty)$.)

(d) While ∞ and $-\infty$ do not have inverses in the semigroup $([-\infty,\infty],\cdot)$, there seems no harm in writing $a/\infty=a/(-\infty)=0$ for every $a\in\mathbb{R}$. But of course such an extension of the notion of division must be watched carefully in such formulae as $u\cdot\frac{v}{u}$.

135B The order structure of $[-\infty,\infty]$ (a) If we write

$$-\infty\le u\le\infty \text{ for every } u\in[-\infty,\infty],$$

we obtain a relation on $[-\infty,\infty]$, extending the usual ordering of $\mathbb{R}$, which is a *total* ordering, that is,

for any u, v, $w \in [-\infty,\infty]$, if $u \le v$ and $v \le w$ then $u \le w$,

$u \le u$ for every $u \in [-\infty,\infty]$,

for any u, $v \in [-\infty,\infty]$, if $u \le v$ and $v \le u$ then $u = v$,

for any u, $v \in [-\infty,\infty]$, either $u \le v$ or $v \le u$.

Moreover, every subset of $[-\infty,\infty]$ has a supremum and an infimum, if we write $\sup\emptyset = -\infty$, $\inf\emptyset = \infty$.

(b) The ordering is 'translation-invariant' in the weak sense that

if u, v, $w \in [-\infty,\infty]$ and $v \le w$ and $u+v$, $u+w$ are both defined, then $u+v \le u+w$.

It is preserved by non-negative multiplications in the sense that

if u, v, $w \in [-\infty,\infty]$ and $0 \le u$ and $v \le w$, then $uv \le uw$,

while it is reversed by non-positive multiplications in the sense that

if u, v, $w \in [-\infty,\infty]$ and $u \le 0$ and $v \le w$, then $uw \le uv$.

135C The Borel structure of $[-\infty,\infty]$ We say that a set $E \subseteq [-\infty,\infty]$ is a **Borel set** in $[-\infty,\infty]$ if $E \cap \mathbb{R}$ is a Borel subset of $\mathbb{R}$. It is easy to check that the family of such sets is a σ-algebra of subsets of $[-\infty,\infty]$. See also 135Xb below.

135D Convergent sequences in $[-\infty,\infty]$ We can say that a sequence $\langle u_n\rangle_{n\in\mathbb{N}}$ in $[-\infty,\infty]$ **converges** to $u \in [-\infty,\infty]$ if

whenever $v < u$ there is an $n_0 \in \mathbb{N}$ such that $v \le u_n$ for every $n \ge n_0$, and whenever $u < v$ there is an $n_0 \in \mathbb{N}$ such that $u_n \le v$ for every $n \ge n_0$;

alternatively,

either $u \in \mathbb{R}$ and for every $\delta > 0$ there is an $n_0 \in \mathbb{N}$ such that $u_n \in [u-\delta, u+\delta]$ for every $n \ge n_0$

or $u = -\infty$ and for every $a \in \mathbb{R}$ there is an $n_0 \in \mathbb{N}$ such that $u_n \le a$ for every $n \ge n_0$

or $u = \infty$ and for every $a \in \mathbb{R}$ there is an $n_0 \in \mathbb{N}$ such that $u_n \ge a$ for every $n \ge n_0$.

(Compare the notion of convergence in 112Ba.)

135E Measurable functions Let X be any set and Σ a σ-algebra of subsets of X.

(a) Let D be a subset of X and Σ_D the subspace σ-algebra (121A). For any function $f : D \to [-\infty,\infty]$, the following are equiveridical:

(i) $\{x : f(x) < u\} \in \Sigma_D$ for every $u \in [-\infty,\infty]$;
(ii) $\{x : f(x) \le u\} \in \Sigma_D$ for every $u \in [-\infty,\infty]$;
(iii) $\{x : f(x) > u\} \in \Sigma_D$ for every $u \in [-\infty,\infty]$;
(iv) $\{x : f(x) \ge u\} \in \Sigma_D$ for every $u \in [-\infty,\infty]$;
(v) $\{x : f(x) \le q\} \in \Sigma_D$ for every $q \in \mathbb{Q}$.

P The proof is almost identical to that of 121B. The only modifications are:

– in (i)⇒(ii), $\{x : f(x) \le \infty\}$ and $\{x : f(x) \le -\infty\}$ are not necessarily equal to $\bigcap_{n\in\mathbb{N}}\{x : f(x) < \infty + 2^{-n}\}$, $\bigcap_{n\in\mathbb{N}}\{x : f(x) < -\infty + 2^{-n}\}$; but the former is D, so surely belongs to Σ_D, and the latter is $\bigcap_{n\in\mathbb{N}}\{x : f(x) < -n\}$, so belongs to Σ_D.

– In (iii)⇒(iv), similarly, we have to use the facts that

$$\{x : f(x) \ge -\infty\} = D \in \Sigma_D, \quad \{x : f(x) \ge \infty\} = \bigcap_{n\in\mathbb{N}}\{x : f(x) > n\} \in \Sigma_D.$$

– Concerning the extra condition (v), of course we have (ii)⇒(v), but also we have (v)⇒(i), because

$$\{x : f(x) < u\} = \bigcup_{q\in\mathbb{Q},q<u}\{x : f(x) \le q\}$$

for every $u \in [-\infty,\infty]$. **Q**

(b) We may therefore say, as in 121C, that a function taking values in $[-\infty,\infty]$ is **measurable** if it satisfies these equivalent conditions.

(c) Note that if $f : D \to [-\infty,\infty]$ is Σ-measurable, then

$$E_\infty(f) = f^{-1}[\{\infty\}] = \{x : f(x) \ge \infty\}, \quad E_{-\infty}(f) = f^{-1}[\{-\infty\}] = \{x : f(x) \le -\infty\}$$

must belong to Σ_D, while $f_{\mathbb{R}} = f\restriction D \setminus (E_\infty(f) \cup E_{-\infty}(f))$, the 'real-valued part of f', is measurable in the sense of 121C.

(d) Conversely, if E_∞ and $E_{-\infty}$ belong to Σ_D, and $f_{\mathbb{R}} : D \setminus (E_\infty \cup E_{-\infty}) \to \mathbb{R}$ is measurable, then $f : D \to [-\infty,\infty]$ will be measurable, where $f(x) = \infty$ if $x \in E_\infty$, $f(x) = -\infty$ if $x \in E_{-\infty}$ and $f(x) = f_{\mathbb{R}}(x)$ for other $x \in D$.

(e) It follows that if f, g are measurable functions from subsets of X to $[-\infty,\infty]$, then $f+g$, $f \times g$ and f/g are measurable. **P** This can be proved either by adapting the arguments of 121Eb, 121Ed and 121Ee, or by applying those results to $f_{\mathbb{R}}$ and $g_{\mathbb{R}}$ and considering separately the sets on which one or both are infinite. **Q**

(f) We can say that a function h from a subset D of $[-\infty,\infty]$ to $[-\infty,\infty]$ is **Borel measurable** if it is measurable (in the sense of (b) above) with respect to the Borel σ-algebra of $[-\infty,\infty]$ (as defined in 135C). Now if X is a set, Σ is a σ-algebra of subsets of X, f is a measurable function from a subset of X to $[-\infty,\infty]$ and h is a Borel measurable function from a subset of $[-\infty,\infty]$ to $[-\infty,\infty]$, then hf is measurable. **P** Apply 121Eg to $h^* f_{\mathbb{R}}$, where $h^* = h{\restriction}(\mathbb{R} \cap h^{-1}[\mathbb{R}])$, and then look separately at the sets $\{x : f(x) = \pm\infty\}$, $\{x : hf(x) = \pm\infty\}$. **Q**

(g) Let X be a set and Σ a σ-algebra of subsets of X. Let $\langle f_n \rangle_{n\in\mathbb{N}}$ be a sequence of measurable functions from subsets of X to $[-\infty,\infty]$. Then $\lim_{n\to\infty} f_n$, $\sup_{n\in\mathbb{N}} f_n$ and $\inf_{n\in\mathbb{N}} f_n$ are measurable, if, following the principles set out in 121F, we take their domains to be

$$\{x : x \in \bigcup_{n\in\mathbb{N}} \bigcap_{m\ge n} \operatorname{dom} f_m,\ \lim_{n\to\infty} f_n(x) \text{ exists in } [-\infty,\infty]\},$$

$$\bigcap_{n\in\mathbb{N}} \operatorname{dom} f_n.$$

P Follow the method of 121Fa-121Fc. **Q**

135F $[-\infty,\infty]$-valued integrable functions (a) We are surely not going to admit a function as 'integrable' unless it is finite almost everywhere, and for such functions the remarks in 133B are already adequate.

(b) However, it is possible to make a consistent extension of the idea of an infinite integral, elaborating slightly the ideas of 133A. If (X,Σ,μ) is a measure space and f is a function, defined almost everywhere in X, taking values in $[0,\infty]$, and virtually measurable (that is, such that $f{\restriction}E$ is measurable in the sense of 135E for some conegligible set E), then we can safely write '$\int f = \infty$' whenever f is not integrable. We shall find that for such functions we have $\int f + g = \int f + \int g$ and $\int cf = c\int f$ for every $c \in [0,\infty]$, using the definitions given above for addition and multiplication on $[0,\infty]$. Consequently, as in 122M-122O, we can say that for a general virtually measurable function f, defined almost everywhere in X, taking values in $[-\infty,\infty]$, $\int f = \int f_1 - \int f_2$ whenever f is expressible as a difference $f_1 - f_2$ of non-negative functions such that $\int f_1$ and $\int f_2$ are both defined and not both infinite. Now we have, as always, the basic formulae

$$\int f + g = \int f + \int g, \quad \int cf = c\int f, \quad \int |f| \ge |\int f|$$

whenever the right-hand-sides are defined, and $\int f \le \int g$ whenever $f \le_{\text{a.e.}} g$ and both integrals are defined. It is important to note that $\int f$ can be finite, on this definition, only when f is finite almost everywhere.

135G We now have versions of B.Levi's theorem and Fatou's Lemma (compare 133K).

Proposition Let (X,Σ,μ) be a measure space, and $\langle f_n \rangle_{n\in\mathbb{N}}$ a sequence of $[-\infty,\infty]$-valued functions defined almost everywhere in X which have integrals defined in $[-\infty,\infty]$.

(a) If $f_n \le_{\text{a.e.}} f_{n+1}$ for every n and $-\infty < \sup_{n\in\mathbb{N}} \int f_n$, then $\int \sup_{n\in\mathbb{N}} f_n = \sup_{n\in\mathbb{N}} \int f_n$.

(b) If, for each n, $f_n \ge 0$ a.e., then $\int \liminf_{n\to\infty} f_n \le \liminf_{n\to\infty} \int f_n$.

proof (a) Note that $f = \sup_{n\in\mathbb{N}} f_n$ is defined everywhere on $\bigcap_{n\in\mathbb{N}} \operatorname{dom} f_n$, which is almost everywhere; and that there is a conegligible set E such that $f_n{\restriction}E$ is measurable for every n, so that $f{\restriction}E$ is measurable. Now if $u = \sup_{n\in\mathbb{N}} \int f_n$ is finite, then all but finitely many of the f_n must be finite almost everywhere, and the result is a consequence of B.Levi's theorem for real-valued functions; while if $u = \infty$ then surely $\int \sup_{n\in\mathbb{N}} f_n$ is infinite.

(b) As in 123B or 133Kb, this now follows, applying (a) to $g_n = \inf_{m\ge n} f_m$.

135H Upper and lower integrals again (a) To handle functions taking values in $[-\infty,\infty]$ we need to adapt the definitions in 133I. Let (X,Σ,μ) be a measure space and f a $[-\infty,\infty]$-valued function defined almost everywhere in X. Its **upper integral** is

$$\overline{\int} f = \inf\{\int g : \int g \text{ is defined in the sense of 135F and } f \le_{\text{a.e.}} g\},$$

allowing ∞ for $\inf\{\infty\}$ and $-\infty$ for $\inf\,]-\infty,\infty]$ or $\inf[-\infty,\infty]$. Similarly, the **lower integral** of f is

$$\underline{\int} f = \sup\{\int g : \int g \text{ is defined, } f \geq_{\text{a.e.}} g\}.$$

With this modification, all the results of 133J are valid for functions taking values in $[-\infty,\infty]$ rather than in $\mathbb{R}$.

(b) Corresponding to 133Ka, we have the following. Let (X,Σ,μ) be a measure space, and $\langle f_n\rangle_{n\in\mathbb{N}}$ a sequence of $[-\infty,\infty]$-valued functions defined almost everywhere in X.

(i) If $f_n \leq_{\text{a.e.}} f_{n+1}$ for every n and $\sup_{n\in\mathbb{N}} \overline{\int} f_n > -\infty$, then $\overline{\int} \sup_{n\in\mathbb{N}} f_n = \sup_{n\in\mathbb{N}} \overline{\int} f_n$.

(ii) If, for each n, $f_n \geq 0$ a.e., then $\overline{\int} \liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty} \overline{\int} f_n$.

135I Subspace measures We need to re-examine the ideas of §131 in the new context.

Proposition Let (X,Σ,μ) be a measure space, and $H \in \Sigma$; write Σ_H for the subspace σ-algebra on H and μ_H for the subspace measure. For any $[-\infty,\infty]$-valued function f defined on a subset of H, write $\tilde f$ for the extension of f defined by saying that $\tilde f(x) = f(x)$ if $x \in \operatorname{dom} f$, 0 if $x \in X\setminus H$.

(a) Suppose that f is a $[-\infty,\infty]$-valued function defined on a subset of H.

(i) $\operatorname{dom} f$ is μ_H-conegligible iff $\operatorname{dom}\tilde f$ is μ-conegligible.

(ii) f is μ_H-virtually measurable iff $\tilde f$ is μ-virtually measurable.

(iii) $\int_H f d\mu_H = \int_X \tilde f d\mu$ if either is defined in $[-\infty,\infty]$.

(b) Suppose that h is a $[-\infty,\infty]$-valued function defined almost everywhere in X. Then $\int_H (h\restriction H) d\mu_H = \int h\times\chi H\,d\mu$ if either is defined in $[-\infty,\infty]$.

(c) If h is a $[-\infty,\infty]$-valued function and $\int_X h\,d\mu$ is defined in $[-\infty,\infty]$, then $\int_H (h\restriction H)d\mu_H$ is defined in $[-\infty,\infty]$.

(d) Suppose that h is a $[-\infty,\infty]$-valued function defined almost everywhere in X. Then

$$\overline{\int}_H (h\restriction H)d\mu_H = \overline{\int}_X h\times\chi H d\mu.$$

proof (a)(i) This is immediate from 131Ca, since $H\setminus\operatorname{dom} f = X\setminus\operatorname{dom}\tilde f$.

(ii)(α) If f is μ_H-virtually measurable, there is a μ_H-conegligible $E\in\Sigma_H$ such that $f\restriction E$ is Σ_H-measurable. There is an $F\in\Sigma$ such that $E = F\cap H$; now $G = F\cup(X\setminus H)$ belongs to Σ and $E = G\cap H$ and G is μ-conegligible. Also, for $q\in\mathbb{Q}$,

$$\begin{aligned}\{x : x\in G,\ \tilde f(x)\leq q\} &= \{x : x\in E,\ f(x)\leq q\}\in\Sigma_H\subseteq\Sigma \text{ if } q<0,\\ &= \{x : x\in E,\ f(x)\leq q\}\cup(X\setminus H)\in\Sigma \text{ if } q\geq 0,\end{aligned}$$

so $\tilde f\restriction G$ is Σ-measurable and $\tilde f$ is μ-virtually measurable.

(β) If $\tilde f$ is μ-virtually measurable, there is a μ-conegligible $G\in\Sigma$ such that $\tilde f\restriction G$ is Σ-measurable. Now $E = G\cap H$ belongs to Σ_H and is μ_H-conegligible, and for $q\in\mathbb{Q}$

$$\{x : x\in E,\ f(x)\leq q\} = H\cap\{x : x\in G,\ f(x)\leq q\}\in\Sigma_H.$$

So $f\restriction E$ is Σ_H-measurable and f is μ_H-virtually measurable.

(iii) Assume that at least one of the integrals is defined. Then (ii) tells us that there is a μ-conegligible $E\in\Sigma$ such that $\tilde f\restriction E$ is Σ-measurable, in which case $f\restriction H\cap E$ is Σ_H-measurable.

(α) Suppose that f is non-negative everywhere on its domain. Then $\int_H f d\mu_H$ and $\int_X \tilde f d\mu$ are both defined in $[0,\infty]$. If both are infinite, we can stop. Otherwise,

$$G = \{x : x\in E\cap H,\ f(x)<\infty\} = \{x : x\in E,\ \tilde f(x)<\infty\}$$

must be conegligible. Set $g = f\restriction G\cap H$; then $\tilde g = \tilde f\restriction G$, so $g = f$ μ_H-a.e. and $\tilde g = \tilde f$ μ-a.e. Accordingly $\int_H f d\mu_H = \int_H g\,d\mu_H$ and $\int_X \tilde f d\mu = \int_X \tilde g\,d\mu$. Now we are supposing that at least one of these is finite. But in this case we can apply 131E to see that $\int_H g\,d\mu = \int_X \tilde g\,d\mu$, so $\int_H f\,d\mu = \int_X \tilde f\,d\mu$.

(β) In general, express f as $f^+ - f^-$, where

$$f^+(x) = \max(0,f(x)),\qquad f^-(x) = \max(0,-f(x))$$

for $x\in\operatorname{dom} f$. Then $(f^+)^\sim = \tilde f^+$ and $(f^-)^\sim = \tilde f^-$. So

$$\int_H f d\mu_H = \int_H f^+ d\mu_H - \int_H f^- d\mu_H = \int_X \tilde f^+ d\mu - \int_X \tilde f^- d\mu = \int_X \tilde f d\mu$$

if any of the four expressions is defined in $[-\infty,\infty]$.

(b) Set $f=h\restriction H$; then $(h\times\chi H)(x)=\tilde f(x)$ for every $x\in\operatorname{dom}h$, so (a-iii) tells us that

$$\int_X h\times\chi H\,d\mu=\int_X\tilde f d\mu=\int_H(h\restriction H)d\mu_H$$

if any of the three is defined in $[-\infty,\infty]$.

(c) Setting $h^+(x)=\max(0,h(x))$ and $h^-(x)=\max(0,-h(x))$ for $x\in\operatorname{dom}h$, both $\int_X h^+d\mu$ and $\int_X h^-d\mu$ are defined in $[0,\infty]$, and at most one of them is infinite. In particular, both are μ-virtually measurable and defined μ-almost everywhere, so the same is true of $h^+\times\chi H$ and $h^-\times\chi H$. As $\int_X h^+\times\chi H d\mu\le\int_X h^+d\mu$ and $\int_X h^-\times\chi H d\mu\le\int_X h^-d\mu$, at most one of $\int_X h^+\times\chi H d\mu$, $\int_X h^-\times\chi H d\mu$ is infinite, and

$$\int_X h\times\chi H d\mu=\int_X h^+\times\chi H d\mu-\int_X h^-\times\chi H d\mu$$

is defined in $[-\infty,\infty]$. By (b) above, $\int_H(h\restriction H)d\mu_H$ is defined in $[-\infty,\infty]$.

(d)(i) Suppose that $\int_X g\,d\mu$ is defined in $[-\infty,\infty]$ and that $h\times\chi H\le g$ μ-a.e. Then

$$\int_H(g\restriction H)d\mu_H=\int_X g\times\chi H d\mu$$

is defined, by (c); and as $g(x)\ge 0$ for μ-almost every $x\in X\setminus H$, $g\times\chi H\le_{\text{a.e.}} g$. So

$$\overline{\int}_H(h\restriction H)d\mu_H\le\int_H(g\restriction H)d\mu_H=\int_X g\times\chi H d\mu\le\int_X g\,d\mu.$$

As g is arbitrary, $\overline{\int}_H(h\restriction H)d\mu_H\le\overline{\int}_X h\times\chi H\,d\mu$.

(ii) Suppose that $\int_H f d\mu_H$ is defined in $[-\infty,\infty]$ and that $h\restriction H\le f$ μ_H-a.e. Then $\int_X\tilde f d\mu$ is defined in $[-\infty,\infty]$ and $h\times\chi H\le\tilde f$ μ-a.e., so

$$\overline{\int}_X h\times\chi H\,d\mu\le\int_X\tilde f d\mu=\int_H f\,d\mu_H.$$

As f is arbitrary, $\overline{\int}_X h\times\chi H\,d\mu\le\overline{\int}_H(h\restriction H)d\mu_H$.

135X Basic exercises (a) We say that a set $G\subseteq[-\infty,\infty]$ is **open** if (i) $G\cap\mathbb{R}$ is open in the usual sense as a subset of $\mathbb{R}$ (ii) if $\infty\in G$, then there is some $a\in\mathbb{R}$ such that $]a,\infty]\subseteq G$ (iii) if $-\infty\in G$ then there is some $a\in\mathbb{R}$ such that $[-\infty,a[\,\subseteq G$. Show that the family $\mathfrak{T}$ of open subsets of $[-\infty,\infty]$ has the properties corresponding to (a)-(d) of 1A2B.

(b) Show that the Borel sets of $[-\infty,\infty]$ as defined in 135C are precisely the members of the σ-algebra of subsets of $[-\infty,\infty]$ generated by the open sets as defined in 135Xa.

>(c) Define $\phi:[-\infty,\infty]\to[-1,1]$ by setting

$$\phi(-\infty)=-1,\quad \phi(x)=\tanh x=\tfrac{e^{2x}-1}{e^{2x}+1}\text{ if }-\infty<x<\infty,\quad \phi(\infty)=1.$$

Show that (i) ϕ is an order-isomorphism between $[-\infty,\infty]$ and $[-1,1]$ (ii) for any sequence $\langle u_n\rangle_{n\in\mathbb{N}}$ in $[-\infty,\infty]$, $\langle u_n\rangle_{n\in\mathbb{N}}\to u$ iff $\langle\phi(u_n)\rangle_{n\in\mathbb{N}}\to\phi(u)$ (iii) for any set $E\subseteq[-\infty,\infty]$, E is Borel in $[-\infty,\infty]$ iff $\phi[E]$ is a Borel subset of $\mathbb{R}$ (iv) a real-valued function h defined on a subset of $[-\infty,\infty]$ is Borel measurable iff $h\phi^{-1}$ is Borel measurable.

>(d) Let X be a set, Σ a σ-algebra of subsets of X and f a function from a subset of X to $[-\infty,\infty]$. Show that f is measurable iff the composition ϕf is measurable, where ϕ is the function of 135Xc. Use this to reduce 135Ef and 135Eg to the corresponding results in §121.

(e) Let $\phi:[-\infty,\infty]\to[-1,1]$ be the function described in 135Xc. Show that the functions

$$(t,u)\mapsto\phi(\phi^{-1}(t)+\phi^{-1}(u)):[-1,1]^2\setminus\{(-1,1),(1,-1)\}\to[-1,1],$$

$$(t,u)\mapsto\phi(\phi^{-1}(t)\phi^{-1}(u)):[-1,1]^2\to[-1,1],$$

$$(t,u)\mapsto\phi(\phi^{-1}(t)/\phi^{-1}(u)):\big([-1,1]\times([-1,1]\setminus\{0\})\big)\setminus\{(\pm1,\pm1)\}\to[-1,1]$$

are Borel measurable. Use this with 121K to prove 135Ee.

(f) Following the conventions of 135Ab and 135Ad, give full descriptions of the cases in which $uu'/vv'=(u/v)(u'/v')$ and in which $uw/vw=u/v$.

(g) Let (X, Σ, μ) be a measure space and suppose that $E \in \Sigma$ has non-zero finite measure. Let f be a virtually measurable $[-\infty, \infty]$-valued function defined on a subset of X and suppose that $f(x)$ is defined and greater than α for almost every $x \in E$. Show that $\int_E f > \alpha\mu E$.

135Y Further exercises (a) Let X be a set and Σ a σ-algebra of subsets of X. Show that if $f : X \to [0, \infty]$ is Σ-measurable, there is a sequence $\langle E_n \rangle_{n\in\mathbb{N}}$ in Σ such that $f = \sum_{n=0}^{\infty} \frac{1}{n+1}\chi E_n$.

(b) Let (X, Σ, μ) be a measure space, and f, g two $[-\infty, \infty]$-valued functions, defined on subsets of X, such that $\int f$ and $\int g$ are both defined in $[-\infty, \infty]$. (i) Show that $\int f \vee g$ and $\int f \wedge g$ are defined in $[-\infty, \infty]$, where $(f \vee g)(x) = \max(f(x), g(x))$, $(f \wedge g)(x) = \min(f(x), g(x))$ for $x \in \operatorname{dom} f \cap \operatorname{dom} g$. (ii) Show that $\int f \vee g + \int f \wedge g = \int f + \int g$ in the sense that if one of the sums is defined in $[-\infty, \infty]$ so is the other, and they are then equal.

(c) Let (X, Σ, μ) be a measure space, $f : X \to [-\infty, \infty]$ a function and $g : X \to [0, \infty]$, $h : X \to [0, \infty]$ measurable functions. Show that $\overline{\int} f \times (g + h) = \overline{\int} f \times g + \overline{\int} f \times h$, where here we interpret $\infty + (-\infty)$ as ∞, as in 133L.

135 Notes and comments I have taken this exposition into a separate section partly because of its length, and partly because I wish to emphasize that these techniques are incidental to the principal ideas of this volume. Really all I am trying to do here is give a coherent account of the language commonly used to deal with a variety of peripheral cases. As a general rule, '∞' enters these arguments only as a shorthand for certain types of triviality. When we find ourselves wishing to assign the values $\pm\infty$ to a function, either this happens on a negligible set – in which case it is often right, if slightly less comforting, to think of the function as undefined on that set – or things have got completely out of hand, and the theory has little useful to tell us.

Of course it is not difficult to incorporate the theory of the extended real line directly into the arguments of Chapter 12, so that the results of this section become the basic ones. I have avoided this route partly in an attempt to reduce the number of new ideas needed in the technically very demanding material of Chapter 12 – believing, as I do, that independently of our treatment of $\pm\infty$ it is absolutely necessary to be able to deal with partially-defined functions – and partly because I do not think that the real line should really be regarded as a substructure of the extended real line. I think that they are different structures with different properties, and that the original real line is overwhelmingly more important. But it is fair to say that in terms of the ideas treated in this volume they are so similar that when you are properly familiar with this work you will be able to move freely from one to the other, so freely indeed that you can safely leave the distinction to formal occasions, such as when you are presenting the statement of a theorem.

*136 The Monotone Class Theorem

For the final section of this volume, I present two theorems on σ-algebras, with some simple corollaries. They are here because I find no natural home for them in Volume 2. While they (especially 136B) are part of the basic technique of measure theory, and have many and widespread applications, they are not central to the particular approach I have chosen, and can if you wish be left on one side until they come to be needed.

136A Lemma Let X be a set, and $\mathcal{A}$ a family of subsets of X. Then the following are equiveridical:

(i) $X \in \mathcal{A}$, $B \setminus A \in \mathcal{A}$ whenever A, $B \in \mathcal{A}$ and $A \subseteq B$, and $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$ whenever $\langle A_n \rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence in $\mathcal{A}$;

(ii) $\emptyset \in \mathcal{A}$, $X \setminus A \in \mathcal{A}$ for every $A \in \mathcal{A}$, and $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$ whenever $\langle A_n \rangle_{n\in\mathbb{N}}$ is a disjoint sequence in $\mathcal{A}$.

proof (i)⇒(ii) Suppose that (i) is true. Then of course $\emptyset = X \setminus X$ belongs to $\mathcal{A}$ and $X \setminus A \in \mathcal{A}$ for every $A \in \mathcal{A}$. If A, $B \in \mathcal{A}$ are disjoint, then $A \subseteq X \setminus B \in \mathcal{A}$, so $(X \setminus B) \setminus A$ and its complement $A \cup B$ belong to $\mathcal{A}$. So if $\langle A_n \rangle_{n\in\mathbb{N}}$ is a disjoint sequence in $\mathcal{A}$, $\bigcup_{i\le n} A_i \in \mathcal{A}$ for every n, and $\bigcup_{n\in\mathbb{N}} A_n$ is the union of a non-decreasing sequence in $\mathcal{A}$, so belongs to $\mathcal{A}$. Thus (ii) is true.

(ii)⇒(i) If (ii) is true, then of course $X = X \setminus \emptyset$ belongs to $\mathcal{A}$. If A and B are members of $\mathcal{A}$ such that $A \subseteq B$, then $X \setminus B$ belongs to $\mathcal{A}$ and is disjoint from A, so $A \cup (X \setminus B)$ and its complement $B \setminus A$ belong to $\mathcal{A}$. Thus the second clause of (i) is satisfied. As for the third, if $\langle A_n \rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence in $\mathcal{A}$, then $A_0, A_1 \setminus A_0, A_2 \setminus A_1, \ldots$ is a disjoint sequence in $\mathcal{A}$, so its union $\bigcup_{n\in\mathbb{N}} A_n$ belongs to $\mathcal{A}$.

Definition If $\mathcal{A} \subseteq \mathcal{P}X$ satisfies the conditions of (i) and/or (ii) above, it is called a **Dynkin class** of subsets of X.

136B Monotone Class Theorem Let X be a set and $\mathcal{A}$ a Dynkin class of subsets of X. Suppose that $\mathcal{I} \subseteq \mathcal{A}$ is such that $I \cap J \in \mathcal{I}$ for all I, $J \in \mathcal{I}$. Then $\mathcal{A}$ includes the σ-algebra of subsets of X generated by $\mathcal{I}$.

proof (a) Let $\mathfrak{S}$ be the family of Dynkin classes of subsets of X including $\mathcal{I}$. Then it is easy to check, using either (i) or (ii) of 136A, that the intersection $\Sigma = \bigcap \mathfrak{S}$ also is a Dynkin class (compare 111Ga). Because $\mathcal{A} \in \mathfrak{S}$, $\Sigma \subseteq \mathcal{A}$.

(b) If $H \in \Sigma$, then

$$\Sigma_H = \{E : E \in \Sigma,\, E \cap H \in \Sigma\}$$

is a Dynkin class. **P** (α) $X \cap H = H \in \Sigma$ so $X \in \Sigma_H$. (β) If A, $B \in \Sigma_H$ and $A \subseteq B$ then $A \cap H$, $B \cap H$ belong to Σ and $A \cap H \subseteq B \cap H$; consequently

$$(B \setminus A) \cap H = (B \cap H) \setminus (A \cap H) \in \Sigma$$

and $B \setminus A \in \Sigma_H$. (γ) If $\langle A_n \rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence in Σ_H, then $\langle A_n \cap H\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence in Σ, so

$$(\bigcup_{n\in\mathbb{N}} A_n) \cap H = \bigcup_{n\in\mathbb{N}}(A_n \cap H) \in \Sigma$$

and $\bigcup_{n\in\mathbb{N}} A_n \in \Sigma_H$. **Q**

It follows that if $I \cap H \in \Sigma$ for every $I \in \mathcal{I}$, so that $\Sigma_H \supseteq \mathcal{I}$, then $\Sigma_H \in \mathfrak{S}$ and must be equal to Σ.

(c) We find next that $G \cap H \in \Sigma$ for all G, $H \in \Sigma$. **P** Take I, $J \in \mathcal{I}$. We know that $I \cap J \in \mathcal{I}$. As I is arbitrary, $\Sigma_J = \Sigma$ and $H \in \Sigma_J$, that is, $H \cap J \in \Sigma$. As J is arbitrary, $\Sigma_H = \Sigma$ and $G \in \Sigma_H$, that is, $G \cap H \in \Sigma$. **Q**

(d) Since Σ is a Dynkin class, $\emptyset = X \setminus X \in \Sigma$. Also

$$G \cup H = X \setminus ((X \setminus G) \cap (X \setminus H)) \in \Sigma$$

for any G, $H \in \Sigma$ (using (c)). So if $\langle G_n \rangle_{n\in\mathbb{N}}$ is any sequence in Σ, $G'_n = \bigcup_{i\le n} G_i \in \Sigma$ for each n (inducing on n). But $\langle G'_n\rangle_{n\in\mathbb{N}}$ is now a non-decreasing sequence in Σ, so

$$\bigcup_{n\in\mathbb{N}} G_n = \bigcup_{n\in\mathbb{N}} G'_n \in \Sigma.$$

This means that Σ satisfies all the conditions of 111A and is a σ-algebra of subsets of X. Since $\mathcal{I} \subseteq \Sigma$, Σ must include the σ-algebra Σ' of subsets of X generated by $\mathcal{I}$. So $\Sigma' \subseteq \Sigma \subseteq \mathcal{A}$, as required.

(Actually, of course, $\Sigma = \Sigma'$, because $\Sigma' \in \mathfrak{S}$.)

Remark I have seen this result called the **Sierpiński Class Theorem** and the **π-λ Theorem**.

136C Corollary Let X be a set, and μ, ν two measures defined on X with domains Σ, T respectively. Suppose that $\mu X = \nu X < \infty$, and that $\mathcal{I} \subseteq \Sigma \cap \mathrm{T}$ is a family of sets such that $\mu I = \nu I$ for every $I \in \mathcal{I}$ and $I \cap J \in \mathcal{I}$ for all I, $J \in \mathcal{I}$. Then $\mu E = \nu E$ for every E in the σ-algebra of subsets of X generated by $\mathcal{I}$.

proof The point is that

$$\mathcal{A} = \{H : H \in \Sigma \cap \mathrm{T},\, \mu H = \nu H\}$$

is a Dynkin class of subsets of X. **P** I work from (ii) of 136A. Of course $\emptyset \in \mathcal{A}$. If $A \in \mathcal{A}$ then

$$\mu(X \setminus A) = \mu X - \mu A = \nu X - \nu A = \nu(X \setminus A)$$

(because $\mu X = \nu X < \infty$, so the subtraction is safe), and $X \setminus A \in \mathcal{A}$. If $\langle A_n \rangle_{n\in\mathbb{N}}$ is a disjoint sequence in $\mathcal{A}$, then

$$\mu A = \sum_{n=0}^{\infty} \mu A_n = \sum_{n=0}^{\infty} \nu A_n = \nu A,$$

and $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$. **Q**

Since $\mathcal{I} \subseteq \mathcal{A}$, 136B tells us that the σ-algebra Σ' generated by $\mathcal{I}$ is included in $\mathcal{A}$, that is, μ and ν agree on Σ'.

136D Corollary Let μ, ν be two measures on $\mathbb{R}^r$, where $r \ge 1$, both defined, and agreeing, on all intervals of the form

$$]-\infty, a] = \{x : x \le a\} = \{(\xi_1, \dots, \xi_r) : \xi_i \le \alpha_i \text{ for every } i \le r\}$$

for $a = (\alpha_1, \dots, \alpha_r) \in \mathbb{R}^r$. Suppose further that $\mu\mathbb{R}^r < \infty$. Then μ and ν agree on all the Borel subsets of $\mathbb{R}^r$.

proof In 136C, take $X = \mathbb{R}^r$ and $\mathcal{I}$ the set of intervals $]-\infty, a]$. Then $I \cap J \in \mathcal{I}$ for all I, $J \in \mathcal{I}$, since $]-\infty, a] \cap]-\infty, b] =]-\infty, a \wedge b]$, writing $a \wedge b = (\min(\alpha_1, \beta_1), \dots, \min(\alpha_r, \beta_r))$ if $a = (\alpha_1, \dots, \alpha_r)$, $b = (\beta_1, \dots, \beta_r) \in \mathbb{R}^r$. Also, setting $\mathbf{n} = (n, \dots, n)$ for $n \in \mathbb{N}$,

$$\nu\mathbb{R}^r = \lim_{n\to\infty} \nu\,]-\infty,\mathbf{n}] = \lim_{n\to\infty} \mu\,]-\infty,\mathbf{n}] = \mu\mathbb{R}^r.$$

So all the conditions of 136C are satisfied and μ, ν agree on the σ-algebra Σ generated by $\mathcal{I}$. But this is just the σ-algebra of Borel sets, by 121J.

136E Algebras of sets: Definition Let X be a set. A family $\mathcal{E} \subseteq \mathcal{P}X$ is an **algebra** or **field** of subsets of X if

(i) $\emptyset \in \mathcal{E}$;

(ii) for every $E \in \mathcal{E}$, its complement $X \setminus E$ belongs to $\mathcal{E}$;

(iii) for every E, $F \in \mathcal{E}$, $E \cup F \in \mathcal{E}$.

136F Remarks (a) I could very well have introduced this notion in Chapter 11, along with 'σ-algebras'. I omitted it, apart from some exercises, because there seemed to be quite enough new definitions in §111 already, and because I had nothing substantial to say about algebras of sets.

(b) If $\mathcal{E}$ is an algebra of subsets of X, then

$$E \cap F = X \setminus ((X \setminus E) \cup (X \setminus F)), \quad E \setminus F = E \cap (X \setminus F),$$

$$E_0 \cup E_1 \cup \ldots \cup E_n, \quad E_0 \cap E_1 \cap \ldots \cap E_n$$

belong to $\mathcal{E}$ for all E, F, $E_0, \ldots, E_n \in \mathcal{E}$. (Induce on n for the last.)

(c) A σ-algebra of subsets of X is (of course) an algebra of subsets of X.

136G Theorem Let X be a set and $\mathcal{E}$ an algebra of subsets of X. Suppose that $\mathcal{A} \subseteq \mathcal{P}X$ is a family of sets such that

(α) $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$ for every non-decreasing sequence $\langle A_n\rangle_{n\in\mathbb{N}}$ in $\mathcal{A}$,

(β) $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{A}$ for every non-increasing sequence $\langle A_n\rangle_{n\in\mathbb{N}}$ in $\mathcal{A}$,

(γ) $\mathcal{E} \subseteq \mathcal{A}$.

Then $\mathcal{A}$ includes the σ-algebra of subsets of X generated by $\mathcal{E}$.

proof I use the same ideas as in 136B.

(a) Let $\mathfrak{S}$ be the family of all sets $\mathcal{S} \subseteq \mathcal{P}X$ satisfying (α)-(γ). Then its intersection $\Sigma = \bigcap\mathfrak{S}$ also satisfies the conditions. Because $\mathcal{A} \in \mathfrak{S}$, $\Sigma \subseteq \mathcal{A}$.

(b) If $H \in \Sigma$, then

$$\Sigma_H = \{E : E \in \Sigma,\, E \cap H \in \Sigma\}$$

satisfies conditions (α)-(β). **P** (α) If $\langle A_n\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence in Σ_H, then $\langle A_n\cap H\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence in Σ, so

$$(\bigcup_{n\in\mathbb{N}} A_n) \cap H = \bigcup_{n\in\mathbb{N}}(A_n \cap H) \in \Sigma$$

and $\bigcup_{n\in\mathbb{N}} A_n \in \Sigma_H$. ($\beta$) Similarly, if $\langle A_n\rangle_{n\in\mathbb{N}}$ is a non-increasing sequence in Σ_H, then $\bigcap_{n\in\mathbb{N}} A_n \cap H \in \Sigma$ so $\bigcap_{n\in\mathbb{N}} A_n \in \Sigma_H$. **Q**

It follows that if $E \cap H \in \Sigma$ for every $E \in \mathcal{E}$, so that Σ_H also satisfies (γ), then $\Sigma_H \in \mathfrak{S}$ and must be equal to Σ.

(c) Consequently $G \cap H \in \Sigma$ for all G, $H \in \Sigma$. **P** Take E, $F \in \mathcal{E}$. We know that $E \cap F \in \mathcal{E}$. As E is arbitrary, $\Sigma_F = \Sigma$ and $H \in \Sigma_F$, that is, $H \cap F \in \Sigma$. As F is arbitrary, $\Sigma_H = \Sigma$ and $G \in \Sigma_H$, that is, $G \cap H \in \Sigma$. **Q**

(d) Next, $\Sigma^* = \{X \setminus H : H \in \Sigma\} \in \mathfrak{S}$. **P** ($\alpha$) If $\langle A_n\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence in Σ^*, then $\langle X \setminus A_n\rangle_{n\in\mathbb{N}}$ is a non-increasing sequence in Σ, so

$$\bigcup_{n\in\mathbb{N}} A_n = X \setminus \bigcap_{n\in\mathbb{N}}(X \setminus A_n) \in \Sigma^*.$$

(β) Similarly, if $\langle A_n\rangle_{n\in\mathbb{N}}$ is a non-increasing sequence in Σ^*, then

$$\bigcap_{n\in\mathbb{N}} A_n = X \setminus \bigcup_{n\in\mathbb{N}}(X \setminus A_n) \in \Sigma^*.$$

(γ) If $E \in \mathcal{E}$ then $X \setminus E \in \mathcal{E}$ so $X \setminus E \in \Sigma$ and $E \in \Sigma^*$. **Q** It follows that $\Sigma \subseteq \Sigma^*$, that is, that $X \setminus H \in \Sigma$ for every $H \in \Sigma$.

(e) Putting (c) and (d) together with the fact that $X \in \Sigma$ (because $X \in \mathcal{E}$) and the union of a non-decreasing sequence in Σ belongs to Σ (by condition (α)), we see that the same argument as in part (d) of the proof of 136B shows that Σ is a σ-algebra of subsets of X. So, just as in 136B, we conclude that the σ-algebra generated by $\mathcal{E}$ is included in Σ and therefore in $\mathcal{A}$.

***136H Proposition** Let (X,Σ,μ) be a measure space such that $\mu X<\infty$, and $\mathcal{E}$ a subalgebra of Σ; let Σ' be the σ-algebra of subsets of X generated by $\mathcal{E}$. If $F\in\Sigma'$ and $\epsilon>0$, there is an $E\in\mathcal{E}$ such that $\mu(E\cap F)\le\epsilon$.

proof Let $\mathcal{A}$ be the family of sets $F\in\Sigma$ such that

$$\text{for every } \epsilon>0 \text{ there is an } E\in\mathcal{E} \text{ such that } \mu(F\triangle E)\le\epsilon.$$

Then $\mathcal{A}$ is a Dynkin class. **P** I check the three conditions of 136A(i). (α) $X\in\mathcal{A}$ because $X\in\mathcal{E}$. (β) If F_1, $F_2\in\mathcal{A}$ and $\epsilon>0$, there are E_1, $E_2\in\mathcal{E}$ such that $\mu(F_i\triangle E_i)\le\frac{1}{2}\epsilon$ for both i; now $E_1\setminus E_2\in\mathcal{E}$ and

$$(F_1\setminus F_2)\triangle(E_1\setminus E_2)\subseteq(F_1\triangle E_1)\cup(F_2\triangle E_2),$$

so

$$\mu((F_1\setminus F_2)\triangle(E_1\setminus E_2))\le\mu(F_1\triangle E_1)+\mu(F_2\triangle E_2)\le\epsilon.$$

As ϵ is arbitrary, $F_1\setminus F_2\in\mathcal{A}$. (γ) If $\langle F_n\rangle_{n\in\mathbb{N}}$ is a non-decreasing sequence in $\mathcal{A}$, with union F, and $\epsilon>0$, then

$$\lim_{n\to\infty}\mu F_n=\mu F\le\mu X<\infty,$$

so there is an $n\in\mathbb{N}$ such that $\mu(F\setminus F_n)\le\frac{1}{2}\epsilon$. Now there is an $E\in\mathcal{E}$ such that $\mu(F_n\triangle E)\le\frac{1}{2}\epsilon$; as $F\triangle E\subseteq(F\setminus F_n)\cup(F_n\triangle E)$, $\mu(F\triangle E)\le\epsilon$. As ϵ is arbitrary, $F\in\mathcal{A}$. **Q**

Since $\mathcal{E}\subseteq\mathcal{A}$ and $\mathcal{E}$ is closed under $\cap$, $\mathcal{A}$ includes the σ-algebra Σ' generated by $\mathcal{E}$, as claimed.

136X Basic exercises >(a) Let X be a set and $\mathcal{A}$ a family of subsets of X. Show that the following are equiveridical:

(i) $X\in\mathcal{A}$ and $B\setminus A\in\mathcal{A}$ whenever A, $B\in\mathcal{A}$ and $A\subseteq B$;

(ii) $\emptyset\in\mathcal{A}$, $X\setminus A\in\mathcal{A}$ for every $A\in\mathcal{A}$ and $A\cup B\in\mathcal{A}$ whenever A, $B\in\mathcal{A}$ are disjoint.

(b) Suppose that X is a set and $\mathcal{A}\subseteq\mathcal{P}X$. Show that $\mathcal{A}$ is a σ-algebra of subsets of X iff it is a Dynkin class and $A\cap B\in\mathcal{A}$ whenever A, $B\in\mathcal{A}$.

(c) Let X be a set, and $\mathcal{I}$ a family of subsets of X such that $I\cap J\in\mathcal{I}$ for all I, $J\in\mathcal{I}$; let Σ be the σ-algebra of subsets of X generated by $\mathcal{I}$. Show that $\mu E=\nu E$ whenever $E\in\Sigma$ is covered by a sequence in $\mathcal{I}$. (*Hint*: For $J\in\mathcal{I}$, set $\mu_J E=\mu(E\cap J)$, $\nu_J E=\nu(E\cap J)$ for $E\in\Sigma$. Use 136C to show that $\mu_J=\nu_J$ for each J.)

>(d) Set $X=\{0,1,2,3\}$, $\mathcal{I}=\{X,\{0,1\},\{0,2\}\}$. Find two distinct measures μ, ν on X, both defined on the σ-algebra $\mathcal{P}X$ and with $\mu I=\nu I<\infty$ for every $I\in\mathcal{I}$.

(e) Let Σ be the family of subsets of $[0,1[$ expressible as finite unions of half-open intervals $[a,b[$. Show that Σ is an algebra of subsets of $[0,1[$.

(f) Let X be a set, and $\mathcal{I}$ a family of subsets of X such that $I\cap J\in\mathcal{I}$ whenever I, $J\in\mathcal{I}$. Let Σ be the smallest family of sets such that $X\in\Sigma$, $F\setminus E\in\Sigma$ whenever E, $F\in\Sigma$ and $E\subseteq F$, and $\mathcal{I}\subseteq\Sigma$. Show that Σ is an algebra of subsets of X.

(g) Let X be a set, and $\mathcal{E}$ an algebra of subsets of X. A functional $\nu:\mathcal{E}\to\mathbb{R}$ is called (**finitely**) **additive** if $\nu(E\cup F)=\nu E+\nu F$ whenever E, $F\in\mathcal{E}$ and $E\cap F=\emptyset$. (i) Show that in this case $\nu(E\cup F)+\nu(E\cap F)=\nu E+\nu F$ for all E, $F\in\mathcal{E}$. (ii) Show that if $\nu E\ge 0$ for every $E\in\mathcal{E}$ then $\nu(\bigcup_{i\le n}E_i)\le\sum_{i=0}^{n}\nu E_i$ for all $E_0,\dots,E_n\in\mathcal{E}$.

>(h) Let X be a set, and $\mathcal{A}$ a family of subsets of X such that (α) $\emptyset$, X belong to $\mathcal{A}$ (β) $A\cap B\in\mathcal{A}$ for all A, $B\in\mathcal{A}$ (γ) $A\cup B\in\mathcal{A}$ whenever A, $B\in\mathcal{A}$ and $A\cap B=\emptyset$. Show that $\{A:A\in\mathcal{A},\ X\setminus A\in\mathcal{A}\}$ is an algebra of subsets of X.

>(i) Let X be a set, and $\mathcal{A}$ a family of subsets of X such that (α) $\emptyset$, X belong to $\mathcal{A}$ (β) $\bigcap_{n\in\mathbb{N}}A_n\in\mathcal{A}$ for every sequence $\langle A_n\rangle_{n\in\mathbb{N}}$ in $\mathcal{A}$ (γ) $\bigcup_{n\in\mathbb{N}}A_n\in\mathcal{A}$ for every disjoint sequence $\langle A_n\rangle_{n\in\mathbb{N}}$ in $\mathcal{A}$. Show that $\{A:A\in\mathcal{A},\ X\setminus A\in\mathcal{A}\}$ is a σ-algebra of subsets of X.

>(j) Let $\mathcal{A}$ be a family of subsets of $\mathbb{R}$ such that (i) $\bigcap_{n\in\mathbb{N}}A_n\in\mathcal{A}$ for every sequence $\langle A_n\rangle_{n\in\mathbb{N}}$ in $\mathcal{A}$ (ii) $\bigcup_{n\in\mathbb{N}}A_n\in\mathcal{A}$ for every disjoint sequence $\langle A_n\rangle_{n\in\mathbb{N}}$ in $\mathcal{A}$ (iii) every open interval $]a,b[$ belongs to $\mathcal{A}$. Show that every Borel subset of $\mathbb{R}$ belongs to $\mathcal{A}$. (*Hint*: show that every half-open interval $[a,b[$, $]a,b]$ belongs to $\mathcal{A}$, and therefore all intervals $]-\infty,a]$, $[a,\infty[$; now use 136Xi.)

>(k) Let X be a set, $\mathcal{E}$ an algebra of subsets of X, and $\mathcal{A}$ a family of subsets of X such that (α) $\bigcap_{n\in\mathbb{N}}A_n\in\mathcal{A}$ for every non-increasing sequence $\langle A_n\rangle_{n\in\mathbb{N}}$ in $\mathcal{A}$ (β) $\bigcup_{n\in\mathbb{N}}A_n\in\mathcal{A}$ for every disjoint sequence in $\mathcal{A}$ (γ) $\mathcal{E}\subseteq\mathcal{A}$. Show that the σ-algebra of sets generated by $\mathcal{E}$ is included in $\mathcal{A}$. (*Hint*: use the method of 136B to reduce to the case in which $A\cap B\in\mathcal{A}$ for every A, $B\in\mathcal{A}$; now use 136Xi.)

136Y Further exercises (a) Let X be a set and $\mathcal{E}$ an algebra of subsets of X. Let $\nu : \mathcal{E} \to [0,\infty[$ be a non-negative functional which is additive in the sense of 136Xg. Define $\theta : \mathcal{P}X \to [0,\infty[$ by setting

$$\theta A = \inf\{\sum_{n=0}^{\infty} \nu E_n : \langle E_n\rangle_{n\in\mathbb{N}} \text{ is a sequence in } \mathcal{E} \text{ covering } A\}$$

for every $A \subseteq X$. (i) Show that θ is an outer measure on X and that $\theta E \le \nu E$ for every $E \in \mathcal{E}$. (ii) Let μ be the measure on X defined from θ by Carathéodory's method, and Σ its domain. Show that $\mathcal{E} \subseteq \Sigma$ and that $\mu E \le \nu E$ for every $E \in \mathcal{E}$. (iii) Show that the following are equiveridical: (α) $\mu E = \nu E$ for every $E \in \mathcal{E}$ (β) $\theta X = \nu X$ (γ) whenever $\langle E_n\rangle_{n\in\mathbb{N}}$ is a non-increasing sequence in $\mathcal{E}$ with empty intersection, $\lim_{n\to\infty} \nu E_n = 0$.

(b) Let X be a set, $\mathcal{E}$ an algebra of subsets of X, and ν a non-negative additive functional on $\mathcal{E}$. Let Σ be the σ-algebra of subsets of X generated by $\mathcal{E}$. Show that there is at most one measure μ on X with domain Σ extending ν, and that there is such a measure iff $\lim_{n\to\infty} \nu E_n = 0$ for every non-increasing sequence $\langle E_n\rangle_{n\in\mathbb{N}}$ in $\mathcal{E}$ with empty intersection.

(c) Let X be a set. Let $\mathcal{G}$ be a family of subsets of X such that (i) $G \cap H \in \mathcal{G}$ for all G, $H \in \mathcal{G}$ (ii) for every $G \in \mathcal{G}$ there is a sequence $\langle G_n\rangle_{n\in\mathbb{N}}$ in $\mathcal{G}$ such that $X \setminus G = \bigcap_{n\in\mathbb{N}} G_n$. Let $\mathcal{A}$ be a family of subsets of X such that (α) $\emptyset$, $X \in \mathcal{A}$ (β) $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{A}$ for every non-increasing sequence $\langle A_n\rangle_{n\in\mathbb{N}}$ in $\mathcal{A}$ (γ) $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$ for every disjoint sequence in $\mathcal{A}$ (δ) $\mathcal{G} \subseteq \mathcal{A}$. Show that the σ-algebra of sets generated by $\mathcal{G}$ is included in $\mathcal{A}$.

136 Notes and comments The most useful result here is 136B; it will be needed in Chapter 27, and helpful at various other points in Volume 2, often through its corollaries 136C and 136Xc. Of course 136C, like its corollary 136D and its special case 136Yb, can be used directly only on measures which do not take the value ∞, since we have to know that $\mu(F \setminus E) = \mu F - \mu E$ for measurable sets $E \subseteq F$; that is why it comes into prominence only when we specialize to probability measures (for which the whole space has measure 1). So I include 136Xc to indicate a technique that can take us a step farther. I do not feel that we are really ready for general measures on the Borel sets of $\mathbb{R}^r$, but I mention 136D to show what kind of class $\mathcal{I}$ can appear in 136B.

The two theorems here (136B, 136G) both address the question: given a family of sets $\mathcal{I}$, what operations must we perform in order to build the σ-algebra Σ generated by $\mathcal{I}$? For arbitrary $\mathcal{I}$, of course, we expect to need complements and unions of sequences. The point of the theorems here is that if $\mathcal{I}$ has a certain amount of structure then we can reach Σ with more limited operations; thus if $\mathcal{I}$ is an algebra of sets, then *monotonic* unions and intersections are enough (136G). Of course there are innumerable variations on this theme. I offer 136Xh-136Xj as a typical result which will actually be used in Volume 4, and 136Xk and 136Yc as examples of possible modifications. There is an abstract version of 136B in 313G in Volume 3.

Having once started to consider the extension of an algebra of sets to a σ-algebra, it is natural to ask for conditions under which a functional on an algebra of sets can be extended to a measure. The condition of additivity (136Xg) is obviously necessary, and almost equally obviously not sufficient. I include 136Ya-136Yb as the most important of many necessary and sufficient conditions for an additive functional to be extendable to a measure. We shall have to return to this in Volume 4.

Appendix to Volume 1

Useful Facts

Each volume of this treatise will have an appendix, containing very brief accounts of material which many readers will have met before but some may not, and which is relevant to some topic dealt with in the volume. For this first volume the appendix is short, partly because the volume itself is short, but mostly because the required basic knowledge of analysis is so fundamental that it must be done properly from a regular textbook or in a regular course. However I do set out a few details that might be omitted from some first courses in analysis, describing some not-quite-standard notation and the elementary theory of countable sets (§1A1), open and closed sets in Euclidean space (§1A2) and upper and lower limits of sequences and functions (§1A3).

1A1 Set theory

In 111E-111F I briefly discussed 'countable' sets. The approach there was along what seemed to be the shortest path to the facts immediately needed, and it is perhaps right that I should here indicate a more conventional route. I take the opportunity to list some notation which I find convenient but is not universally employed.

1A1A Square bracket notations I use square brackets [and] in a variety of ways; the context will I hope always make it clear what interpretation is expected.

(a) For a, $b \in \mathbb{R}$, I write

$$[a,b] = \{x : a \le x \le b\}, \quad]a,b[= \{x : a < x < b\},$$

$$[a,b[= \{x : a \le x < b\}, \quad]a,b] = \{x : a < x \le b\}.$$

It is natural, when these formulae appear, to jump to the conclusion that $a < b$; but just occasionally it is useful to interpret them when $b \le a$, in which case I follow the formulae above literally, so that

$$[a,a] = \{a\}, \quad]a,a[= [a,a[=]a,a] = \emptyset,$$

$$[a,b] =]a,b[= [a,b[=]a,b] = \emptyset \text{ if } b < a.$$

(b) We can interpret the formulae with infinite a or b; for example,

$$]-\infty,b[= \{x : x < b\}, \quad]a,\infty[= \{x : a < x\}, \quad]-\infty,\infty[= \mathbb{R},$$

$$[a,\infty[= \{x : x \ge a\}, \quad]-\infty,b] = \{x : x \le b\},$$

and even

$$[0,\infty] = \{x : x \in \mathbb{R},\, x \ge 0\} \cup \{\infty\}, \quad [-\infty,\infty] = \mathbb{R} \cup \{-\infty,\infty\}.$$

(c) With some circumspection – since further choices have to be made, which it is safer to set out explicitly when the occasion arises – we can use similar formulae for 'intervals' in multidimensional space $\mathbb{R}^r$; see, for instance, 115A or 136D; and even in general partially ordered sets, though these will not be important to us before Volume 3.

(d) Perhaps I owe you an explanation for my choice of $]a,b[$, $[a,b[$ in favour of (a,b), $[a,b)$, which are both commoner and more pleasing to the eye. In the first instance it is simply because the formula

$$(1,2) \in]0,2[\times]1,3[$$

makes better sense than its translation. Generally, it leads to a slightly better balance in the number of appearances of (and [, even allowing for the further uses of $[\dots]$ which I am about to specify.

1A1B Direct and inverse images I now describe an entirely different use of square brackets, belonging to abstract set theory rather than to the theory of the real number system.

(a) If f is a function and A is a set, I write

$$f[A] = \{f(x) : x \in A \cap \operatorname{dom} f\}$$

for the **direct image** of A under f. Note that while A will often be a subset of the domain of f, this is not assumed.

(b) If f is a function and B is a set, I write

$$f^{-1}[B] = \{x : x \in \operatorname{dom} f,\ f(x) \in B\}$$

for the **inverse image** of B under f. This time, it is important to note that there is no presumption that f is injective, or that f^{-1} is a function; the formula $f^{-1}[\ ]$ is being given a meaning independent of any meaning of the expression f^{-1}. But it is easy to see that when f is injective, so that we have a true inverse function f^{-1} (defined on the set of values of f, $f[\operatorname{dom} f]$), then $f^{-1}[B]$, as defined here, agrees with its interpretation under (a).

(c) Now suppose that R is a relation, that is, a set of ordered pairs, and A, B are sets. Then I write

$$R[A] = \{y : \exists\, x \in A \text{ such that } (x, y) \in R\},$$

$$R^{-1}[B] = \{x : \exists\, y \in B \text{ such that } (x, y) \in R\}.$$

If we write

$$R^{-1} = \{(y, x) : (x, y) \in R\},$$

then we have an alternative interpretation of $R^{-1}[B]$ which agrees with the one just given. Moreover, if R is the graph of a function f, that is, if for every x there is at most one y such that $(x, y) \in R$, then the formulae here agree with those of (a)-(b) above.

(d) (The following is addressed exclusively to readers who have been taught to distinguish between the words 'set' and 'class'.) I have used the word 'set' more than once above. But that was purely for euphony. The same formulae can be used with arbitrary classes, though in some set theories the expressions involved may not be recognised as 'terms' in the technical sense.

1A1C Countable sets In 111Fa I defined 'countable set' as follows: a set K is countable if either it is empty or there is a surjective function from $\mathbb{N}$ to K. A commoner formulation is to say that a set K is countable iff either it is finite or there is a bijection between $\mathbb{N}$ and K. So I should check at once that these two formulations agree.

1A1D Proposition Let K be a set. Then the following are equiveridical:
(i) either K is empty or there is a surjection from $\mathbb{N}$ onto K;
(ii) either K is finite or there is a bijection between $\mathbb{N}$ and K;
(iii) there is an injection from K to $\mathbb{N}$.

proof (a)(i)$\Rightarrow$(iii) Assume (i). If K is empty, then the empty function is an injection from K to $\mathbb{N}$. Otherwise, there is a surjection $\phi : \mathbb{N} \to K$. Now, for each $k \in K$, set

$$\psi(k) = \min\{n : n \in \mathbb{N},\ \phi(n) = k\};$$

this is always well-defined because ϕ is surjective, so that $\{n : \phi(n) = k\}$ is never empty, and must have a least member. Because $\phi\psi(k) = k$ for every k, ψ must be injective, so is the required injection from K to $\mathbb{N}$.

(b)(iii)$\Rightarrow$(ii) Assume (iii); let $\psi : K \to \mathbb{N}$ be an injection, and set $A = \psi[K] \subseteq \mathbb{N}$. Then ψ is a bijection between K and A. If K is finite, then of course (ii) is satisfied. Otherwise, A must also be infinite. Define $\phi : A \to \mathbb{N}$ by setting

$$\phi(m) = \#(\{i : i \in A,\ i < m\}),$$

the number of elements of A less than m, for each $m \in A$; thus $\phi(m)$ is the position of m if the elements of A are listed from the bottom, starting at 0 for the least element of A. Then $\phi : A \to \mathbb{N}$ is a bijection, because A is infinite, and $\phi\psi : K \to \mathbb{N}$ is a bijection.

(c)(ii)$\Rightarrow$(i) If K is empty, surely it satisfies (i). If K is finite and not empty, list its members as $k_0, \ldots, k_n$; now set $\phi(i) = k_i$ for $i \le n$, k_0 for $i > n$ to get a surjection $\phi : \mathbb{N} \to K$. If K is infinite, there is a bijection from $\mathbb{N}$ to K, which is of course also a surjection from $\mathbb{N}$ to K. So (i) is true in all cases.

Remark I referred to the 'empty function' in the proof above. This is the function with domain $\emptyset$; having said this, any, or no, rule for calculating the function will have the same effect, since it will never be applied. By examining your feelings about this construction you can learn something about your basic attitude to mathematics. You may feel that it is an artificial irrelevance, or you may feel that it is as necessary as the number 0. Both are entirely legitimate feelings, and the fully rounded mathematician alternates between them; but I have to say that I myself tend to the latter more often than the former, and that when I say 'function' in this treatise the empty function will generally be in the back of my mind as a possibility.

1A1E Properties of countable sets Let me recapitulate the basic properties of countable sets:

(a) If K is countable and $\phi : K \to L$ is a surjection, then L is countable. **P** If K is empty then so is L. Otherwise there is a surjection $\psi : \mathbb{N} \to K$, so $\phi\psi$ is a surjection from $\mathbb{N}$ onto L, and L is countable. **Q**

(b) If K is countable and $\phi : L \to K$ is an injection, then L is countable. **P** By 1A1D(iii), there is an injection $\psi : K \to \mathbb{N}$; now $\psi\phi : L \to \mathbb{N}$ is injective, so L is countable. **Q**

(c) In particular, any subset of a countable set is countable (as in 111F(b-i)).

(d) The Cartesian product of finitely many countable sets is countable (111Fb(iii)-(iv)).

(e) $\mathbb{Z}$ is countable. **P** The map $(m,n) \mapsto m-n : \mathbb{N}\times\mathbb{N} \to \mathbb{Z}$ is surjective. **Q**

(f) $\mathbb{Q}$ is countable. **P** The map $(m,n) \mapsto \frac{m}{n+1} : \mathbb{Z}\times\mathbb{N} \to \mathbb{Q}$ is surjective. **Q**

1A1F Another fundamental property is worth distinguishing from these, as it relies on a slightly deeper argument.

Theorem If $\mathcal{K}$ is a countable collection of countable sets, then

$$\bigcup\mathcal{K} = \{x : \exists\, K \in \mathcal{K},\ x \in K\}$$

is countable.

proof Set

$$\mathcal{K}' = \mathcal{K} \setminus \{\emptyset\} = \{K : K \in \mathcal{K},\ K \neq \emptyset\};$$

then $\mathcal{K}' \subseteq \mathcal{K}$, so is countable, and $\bigcup\mathcal{K}' = \bigcup\mathcal{K}$. If $\mathcal{K}' = \emptyset$, then

$$\bigcup\mathcal{K} = \bigcup\mathcal{K}' = \emptyset$$

is surely countable. Otherwise, let $m \mapsto K_m : \mathbb{N} \to \mathcal{K}'$ be a surjection. For each $m \in \mathbb{N}$, K_m is a non-empty countable set, so there is a surjection $n \mapsto k_{mn} : \mathbb{N} \to K_m$. Now $(m,n) \mapsto k_{mn} : \mathbb{N}\times\mathbb{N} \to \bigcup\mathcal{K}$ is a surjection (if $k \in \bigcup\mathcal{K}$, there is a $K \in \mathcal{K}'$ such that $k \in K$; there is an $m \in \mathbb{N}$ such that $K = K_m$; there is an $n \in \mathbb{N}$ such that $k = k_{mn}$). So $\bigcup\mathcal{K}$ is countable, as required.

***1A1G Remark** I divide this result from the 'elementary' facts in 1A1E partly because it uses a different principle of argument from any necessary for the earlier work. In the middle of the proof I wrote 'so there is a surjection $n \mapsto k_{mn} : \mathbb{N} \to K_m$'. That there is a surjection from $\mathbb{N}$ onto K_m does indeed follow from the immediately preceding statement 'K_m is a non-empty countable set'. The sleight of hand lies in immediately naming such a surjection as '$n \mapsto k_{mn}$'. There may of course be many surjections from $\mathbb{N}$ to K_m – as a rule, indeed, there will be uncountably many – and what I am in effect doing here is picking arbitrarily on one of them. The choice has to be arbitrary, because I am working in a totally abstract context, and while in any particular application of this theorem there may be some natural surjection to use, I have no way of forecasting what approach, if any, might offer a criterion for distinguishing a particular function here. Now it has been a basic method of mathematical argument, from Euclid's time at least, that we are willing to give a name to an object, a 'general point' or an 'arbitrary number', without specifying exactly which object we are naming. But here I am picking out simultaneously infinitely many objects, all arbitrary members of certain sets. This is a use of the **Axiom of Choice**.

I do not recall ever having had a student criticise an argument in the form of that in 1A1F on the grounds that it uses a new, and possibly illegitimate, principle; I am sure that it never occurred to me that anything exceptionable was being done in these cases, until someone pointed it out. If you find that discussions of this kind are irrelevant to your own mathematical interests, you can certainly pass them by, at least until you reach Volume 5. Mathematical systems have been studied in which the axiom of choice is false; they are of great interest but so far remain peripheral to the subject. Systems in which the axiom of choice is so false that the union of countably many countable sets is sometimes uncountable have a character all of their own, and in particular the theory of Lebesgue measure is transformed; I will come to this possibility in Chapter 56 of Volume 5.

For a brief comment on other ways of using the axiom of choice, see 134C.

1A1H Some uncountable sets Of course not all sets are countable. In 114G/115G I remark that all countable subsets of Euclidean space are negligible for Lebesgue measure; consequently, any set which is not negligible – for instance, any non-trivial interval – must be uncountable. But perhaps it will be helpful if I offer here elementary arguments to show that $\mathbb{R}$ and $\mathcal{P}\mathbb{N}$ are not countable.

(a) There is no surjection from $\mathbb{N}$ onto $\mathbb{R}$. **P** Let $n \mapsto a_n : \mathbb{N} \to \mathbb{R}$ be any function. For each $n \in \mathbb{N}$, express a_n in decimal form as

$$a_n = k_n + 0 \cdot \epsilon_{n1}\epsilon_{n2} \ldots = k_n + \sum_{i=1}^{\infty} 10^{-i}\epsilon_{ni},$$

where $k_n \in \mathbb{Z}$ is the greatest integer not greater than a_n, and each ϵ_{ni} is an integer between 0 and 9; for definiteness, if a_n happens to be an exact decimal, use the terminating expansion, so that the ϵ_{ni} are eventually 0 rather than eventually 9.

Now define ϵ_i, for $i \ge 1$, by saying that

$$\begin{aligned}\epsilon_i &= 6 \text{ if } \epsilon_{ii} < 6,\\ &= 5 \text{ if } \epsilon_{ii} \ge 6.\end{aligned}$$

Consider $a = k_0 + 1 + \sum_{i=1}^{\infty} 10^{-i}\epsilon_i$, so that $a = k_0 + 1 + 0 \cdot \epsilon_1\epsilon_2 \ldots$ in decimal form. I claim that $a \ne a_n$ for every n. Of course $a \ne a_0$ because $a_0 < k_0 + 1 \le a$. If $n \ge 1$, then $\epsilon_n \ne \epsilon_{nn}$; because no ϵ_i is either 0 or 9, there is no alternative decimal expansion of a, so the expansion $a_n = k_n + 0 \cdot \epsilon_{n1}\epsilon_{n2} \ldots$ cannot represent a, and $a \ne a_n$.

Thus I have constructed a real number which is not in the list $a_0, a_1, \ldots$. As $\langle a_n \rangle_{n\in\mathbb{N}}$ is arbitrary, there is no surjection from $\mathbb{N}$ onto $\mathbb{R}$. **Q**

Thus $\mathbb{R}$ is uncountable.

(b) There is no surjection from $\mathbb{N}$ onto its power set $\mathcal{P}\mathbb{N}$. **P** Let $n \mapsto A_n : \mathbb{N} \to \mathcal{P}\mathbb{N}$ be any function. Set

$$A = \{n : n \in \mathbb{N},\ n \notin A_n\}.$$

If $n \in \mathbb{N}$, then

either $n \in A_n$, in which case $n \notin A$,

or $n \notin A_n$, in which case $n \in A$.

Thus in both cases we have $n \in A \triangle A_n$, so that $A \ne A_n$. As n is arbitrary, $A \notin \{A_n : n \in \mathbb{N}\}$ and $n \mapsto A_n$ is not a surjection. As $\langle A_n \rangle_{n\in\mathbb{N}}$ is arbitrary, there is no surjection from $\mathbb{N}$ onto $\mathcal{P}\mathbb{N}$. **Q**

Thus $\mathcal{P}\mathbb{N}$ is also uncountable.

1A1I Remark In fact it is the case that there is a bijection between $\mathbb{R}$ and $\mathcal{P}\mathbb{N}$ (2A1Ha); so that the uncountability of both can be established by just one of the arguments above.

1A1J Notation For definiteness, I remark here that I will say that a family $\mathcal{A}$ of sets is a **partition** of a set X whenever $\mathcal{A}$ is a disjoint cover of X, that is, $X = \bigcup \mathcal{A}$ and $A \cap A' = \emptyset$ for all distinct A, $A' \in \mathcal{A}$; in particular, the empty set may or may not belong to $\mathcal{A}$. Similarly, an indexed family $\langle A_i \rangle_{i\in I}$ is a **partition** partition of X if $\bigcup_{i\in I} A_i = X$ and $A_i \cap A_j = \emptyset$ for all distinct i, $j \in I$; again, one or more of the A_i may be empty.

1A1 Notes and comments The ideas of 1A1C-1A1I are essentially due to G.F.Cantor. These concepts are fundamental to modern set theory, and indeed to very large parts of modern pure mathematics. The notes above hardly begin to suggest the extraordinary fertility of these ideas, which need a book of their own for their proper expression; my only aim here has been to try to make sense of those tiny parts of the subject which are needed in the present volume. In later volumes I will present results which call on substantially more advanced ideas, which I will discuss in appendices to those volumes.

1A2 Open and closed sets in $\mathbb{R}^r$

In 111G I gave the definition of an open set in $\mathbb{R}$ or $\mathbb{R}^r$, and in 121D I used, in passing, some of the basic properties of these sets; perhaps it will be helpful if I set out a tiny part of the elementary theory.

1A2A Open sets Recall that a set $G \subseteq \mathbb{R}$ is **open** if for every $x \in G$ there is a $\delta > 0$ such that $]x - \delta, x + \delta[\subseteq G$; similarly, a set $G \subseteq \mathbb{R}^r$ is **open** if for every $x \in G$ there is a $\delta > 0$ such that $U(x, \delta) \subseteq G$, where $U(x, \delta) = \{y : \|y - x\| < \delta\}$, writing $\|z\|$ for $\sqrt{\zeta_1^2 + \ldots + \zeta_r^2}$ if $z = (\zeta_1, \ldots, \zeta_r)$. Henceforth I give the arguments for general r; if you are at present interested only in the one-dimensional case, you should have no difficulty in reading them as if $r = 1$ throughout.

1A2B The family of all open sets Let $\mathfrak{T}$ be the family of open sets of $\mathbb{R}^r$. Then $\mathfrak{T}$ has the following properties.

(a) $\emptyset \in \mathfrak{T}$, that is, the empty set is open. **P** Because the definition of '$\emptyset$ is open' begins with 'for every $x \in \emptyset$, ...', it must be vacuously satisfied by the empty set. **Q**

(b) $\mathbb{R}^r \in \mathfrak{T}$, that is, the whole space under consideration is an open set. **P** $U(x,1) \subseteq \mathbb{R}^r$ for every $x \in \mathbb{R}^r$. **Q**

(c) If G, $H \in \mathfrak{T}$ then $G \cap H \in \mathfrak{T}$; that is, the intersection of two open sets is always an open set. **P** Let $x \in G \cap H$. Then there are δ_1, $\delta_2 > 0$ such that $U(x,\delta_1) \subseteq G$ and $U(x,\delta_2) \subseteq H$. Set $\delta = \min(\delta_1,\delta_2) > 0$; then

$$U(x,\delta) = \{y : \|y-x\| < \min(\delta_1,\delta_2)\} = U(x,\delta_1) \cap U(x,\delta_2) \subseteq G \cap H.$$

As x is arbitrary, $G \cap H$ is open. **Q**

(d) If $\mathcal{G} \subseteq \mathfrak{T}$, then

$$\bigcup \mathcal{G} = \{x : \exists\, G \in \mathcal{G},\, x \in G\} \in \mathfrak{T};$$

that is, the union of any family of open sets is open. **P** Let $x \in \bigcup \mathcal{G}$. Then there is a $G \in \mathcal{G}$ such that $x \in G$. Because $G \in \mathfrak{T}$, there is a $\delta > 0$ such that

$$U(x,\delta) \subseteq G \subseteq \bigcup \mathcal{G}.$$

As x is arbitrary, $\bigcup \mathcal{G} \in \mathfrak{T}$. **Q**

1A2C Cauchy's inequality: Proposition For all x, $y \in \mathbb{R}^r$, $\|x+y\| \le \|x\| + \|y\|$.

proof Express x as $(\xi_1,\dots,\xi_r)$, y as $(\eta_1,\dots,\eta_r)$; set $\alpha = \|x\|$, $\beta = \|y\|$. Then both α and β are non-negative. If $\alpha = 0$ then $\sum_{j=1}^r \xi_j^2 = 0$ so every $\xi_j = 0$ and $x = \mathbf{0}$, so $\|x+y\| = \|y\| = \|x\| + \|y\|$; if $\beta = 0$, then $y = \mathbf{0}$ and $\|x+y\| = \|x\| = \|x\| + \|y\|$. Otherwise, consider

$$\begin{aligned}
\alpha\beta\|x+y\|^2 &\le \alpha\beta\|x+y\|^2 + \|\alpha y - \beta x\|^2 \\
&= \alpha\beta\sum_{j=1}^r (\xi_j + \eta_j)^2 + \sum_{j=1}^r (\alpha\eta_j - \beta\xi_j)^2 \\
&= \sum_{j=1}^r \alpha\beta\xi_j^2 + \alpha\beta\eta_j^2 + \alpha^2\eta_j^2 + \beta^2\xi_j^2 \\
&= \alpha^3\beta + \alpha\beta^3 + \alpha^2\beta^2 + \beta^2\alpha^2 \\
&= \alpha\beta(\alpha+\beta)^2 = \alpha\beta(\|x\| + \|y\|)^2.
\end{aligned}$$

Dividing both sides by $\alpha\beta$ and taking square roots we have the result.

1A2D Corollary $U(x,\delta)$, as defined in 1A2A, is open, for every $x \in \mathbb{R}^r$ and $\delta > 0$.

proof If $y \in U(x,\delta)$, then $\eta = \delta - \|y-x\| > 0$. Now if $z \in U(y,\eta)$,

$$\|z-x\| = \|(z-y)+(y-x)\| \le \|z-y\| + \|y-x\| < \eta + \|y-x\| = \delta,$$

and $z \in U(x,\delta)$; thus $U(y,\eta) \subseteq U(x,\delta)$. As y is arbitrary, $U(x,\delta)$ is open.

1A2E Closed sets: Definition A set $F \subseteq \mathbb{R}^r$ is **closed** if $\mathbb{R}^r \setminus F$ is open. (*Warning!* 'Most' subsets of $\mathbb{R}^r$ are neither open nor closed; two subsets of $\mathbb{R}^r$, viz., $\emptyset$ and $\mathbb{R}^r$, are both open and closed.) Corresponding to the list in 1A2B, we have the following properties of the family $\mathcal{F}$ of closed subsets of $\mathbb{R}^r$.

1A2F Proposition Let $\mathcal{F}$ be the family of closed subsets of $\mathbb{R}^r$.

(a) $\emptyset \in \mathcal{F}$ (because $\mathbb{R}^r \in \mathfrak{T}$).

(b) $\mathbb{R}^r \in \mathcal{F}$ (because $\emptyset \in \mathfrak{T}$).

(c) If E, $F \in \mathcal{F}$ then $E \cup F \in \mathcal{F}$, because

$$\mathbb{R}^r \setminus (E \cup F) = (\mathbb{R}^r \setminus E) \cap (\mathbb{R}^r \setminus F) \in \mathfrak{T}.$$

(d) If $\mathcal{E} \subseteq \mathcal{F}$ is a *non-empty* family of closed sets, then

$$\bigcap \mathcal{E} = \{x : x \in F\ \forall\, F \in \mathcal{E}\} = \mathbb{R}^r \setminus \textstyle\bigcup_{F\in\mathcal{E}}(\mathbb{R}^r \setminus F) \in \mathcal{F}.$$

Remark In (d), we need to assume that $\mathcal{E} \ne \emptyset$ to ensure that $\bigcap \mathcal{E} \subseteq \mathbb{R}^r$.

1A2G Corresponding to 1A2D, we have the following fact:

Lemma If $x \in \mathbb{R}^r$ and $\delta \geq 0$ then $B(x,\delta) = \{y : \|y-x\| \leq \delta\}$ is closed.

proof Set $G = \mathbb{R}^r \setminus B(x,\delta)$. If $y \in G$, then $\eta = \|y-x\| - \delta > 0$; if $z \in U(y,\eta)$, then

$$\delta + \eta = \|y-x\| \leq \|y-z\| + \|z-x\| < \eta + \|z-x\|,$$

so $\|z-x\| > \delta$ and $z \in G$. So $U(y,\eta) \subseteq G$. As y is arbitrary, G is open and $B(x,\delta)$ is closed.

1A3 Lim sups and lim infs

It occurs to me that not every foundation course in real analysis has time to deal with the concepts lim sup and lim inf.

1A3A Definition (a) For a real sequence $\langle a_n \rangle_{n\in\mathbb{N}}$, write

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \sup_{m\geq n} a_m = \inf_{n\in\mathbb{N}} \sup_{m\geq n} a_m,$$

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} \inf_{m\geq n} a_m = \sup_{n\in\mathbb{N}} \inf_{m\geq n} a_m;$$

if we allow the values $\pm\infty$, both for suprema and infima and for limits (see 112Ba), these will always be defined, because the sequences

$$\langle \sup_{m\geq n} a_m \rangle_{n\in\mathbb{N}}, \quad \langle \inf_{m\geq n} a_m \rangle_{n\in\mathbb{N}}$$

are monotonic.

(b) Explicitly:

$$\limsup_{n\to\infty} a_n = \infty \iff \{a_n : n \in \mathbb{N}\} \text{ is unbounded above},$$

$$\limsup_{n\to\infty} a_n = -\infty \iff \lim_{n\to\infty} a_n = -\infty,$$

that is, if and only if for every $a \in \mathbb{R}$ there is an $n_0 \in \mathbb{N}$ such that $a_n \leq a$ for every $n \geq n_0$;

$$\liminf_{n\to\infty} a_n = -\infty \iff \{a_n : n \in \mathbb{N}\} \text{ is unbounded below},$$

$$\liminf_{n\to\infty} a_n = \infty \iff \lim_{n\to\infty} a_n = \infty,$$

that is, if and only if for every $a \in \mathbb{R}$ there is an $n_0 \in \mathbb{N}$ such that $a_n \geq a$ for every $n \geq n_0$.

(c) For finite $a \in \mathbb{R}$, we have

$\limsup_{n\to\infty} a_n = a$ iff (i) for every $\epsilon > 0$ there is an $n_0 \in \mathbb{N}$ such that $a_n \leq a + \epsilon$ for every $n \geq n_0$ (ii) for every $\epsilon > 0$, $n_0 \in \mathbb{N}$ there is an $n \geq n_0$ such that $a_n \geq a - \epsilon$,

while

$\liminf_{n\to\infty} a_n = a$ iff (i) for every $\epsilon > 0$ there is an $n_0 \in \mathbb{N}$ such that $a_n \geq a - \epsilon$ for every $n \geq n_0$ (ii) for every $\epsilon > 0$, $n_0 \in \mathbb{N}$ there is an $n \geq n_0$ such that $a_n \leq a + \epsilon$.

Generally, for $u \in [-\infty,\infty]$, we can say that

$\limsup_{n\to\infty} a_n = u$ iff (i) for every $v > u$ (if any) there is an $n_0 \in \mathbb{N}$ such that $a_n \leq v$ for every $n \geq n_0$ (ii) for every $v < u$, $n_0 \in \mathbb{N}$ there is an $n \geq n_0$ such that $a_n \geq v$,

$\liminf_{n\to\infty} a_n = u$ iff (i) for every $v < u$ there is an $n_0 \in \mathbb{N}$ such that $a_n \geq v$ for every $n \geq n_0$ (ii) for every $v > u$, $n_0 \in \mathbb{N}$ there is an $n \geq n_0$ such that $a_n \leq v$.

1A3B We have the following basic results.

Proposition For any sequences $\langle a_n \rangle_{n\in\mathbb{N}}$, $\langle b_n \rangle_{n\in\mathbb{N}}$ in $\mathbb{R}$,

(a) $\liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n$,

(b) $\lim_{n\to\infty} a_n = u \in [-\infty,\infty]$ iff $\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = u$,

(c) $\liminf_{n\to\infty} a_n = -\limsup_{n\to\infty}(-a_n)$,

(d) $\limsup_{n\to\infty}(a_n + b_n) \leq \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n$,

(e) $\liminf_{n\to\infty}(a_n + b_n) \geq \liminf_{n\to\infty} a_n + \liminf_{n\to\infty} b_n$,

(f) $\limsup_{n\to\infty} ca_n = c\limsup_{n\to\infty} a_n$ if $c \geq 0$,

(g) $\liminf_{n\to\infty} ca_n = c\liminf_{n\to\infty} a_n$ if $c \geq 0$, with the proviso in (d) and (e) that we must be able to interpret the right-hand-side of the inequality according to the rules in 135A, while in (f) and (g) we take $0 \cdot \infty = 0 \cdot (-\infty) = 0$.

proof (a) $\sup_{m \ge n} a_m \ge \inf_{m \ge n} a_m$ for every n, so

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \sup_{m\ge n} a_m \ge \lim_{n\to\infty} \inf_{m\ge n} a_m = \limsup_{n\to\infty} a_n.$$

(b) Using the last description of $\limsup_{n\to\infty}$ and $\liminf_{n\to\infty}$ in 1A3Ac, and a corresponding description of $\lim_{n\to\infty}$, we have

$$\begin{aligned}
&\lim_{n\to\infty} a_n = u \\
&\iff \text{ for every } v > u \text{ there is an } n_1 \in \mathbb{N} \text{ such that } a_n \le v \text{ for every } n \ge n_1 \\
&\qquad \text{and for every } v < u \text{ there is an } n_2 \in \mathbb{N} \text{ such that } a_n \ge v \text{ for every } n \ge n_2 \\
&\iff \text{ for every } v > u \text{ there is an } n_1 \in \mathbb{N} \text{ such that } a_n \le v \text{ for every } n \ge n_1 \\
&\qquad \text{and for every } v < u,\ n_0 \in \mathbb{N} \text{ there is an } n \ge n_0 \text{ such that } a_n \ge v \\
&\qquad \text{and for every } v < u \text{ there is an } n_2 \in \mathbb{N} \text{ such that } a_n \ge v \text{ for every } n \ge n_2 \\
&\qquad \text{and for every } v > u,\ n_0 \in \mathbb{N} \text{ there is an } n \ge n_0 \text{ such that } a_n \le v \\
&\iff \limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = u.
\end{aligned}$$

(c) This is just a matter of turning the formulae upside down:

$$\begin{aligned}
\liminf_{n\to\infty} a_n &= \sup_{n\in\mathbb{N}} \inf_{m\ge n} a_m = \sup_{n\in\mathbb{N}}(-\sup_{m\ge n}(-a_m)) \\
&= -\inf_{n\in\mathbb{N}} \sup_{m\ge n}(-a_m) = -\limsup_{n\to\infty}(-a_n).
\end{aligned}$$

(d) If $v > \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n$, there are v_1, v_2 such that $v_1 > \limsup_{n\to\infty} a_n$, $v_2 > \limsup_{n\to\infty} b_n$ and $v_1 + v_2 = v$. Now there are n_1, $n_2 \in \mathbb{N}$ such that $\sup_{m\ge n_1} a_n \le v_1$ and $\sup_{m \ge n_2} b_n \le v_2$; so that

$$\begin{aligned}
\sup_{m\ge\max(n_1,n_2)} a_m + b_m &\le \sup_{m\ge\max(n_1,n_2)} a_m + \sup_{m\ge\max(n_1,n_2)} b_m \\
&\le \sup_{m\ge n_1} a_m + \sup_{m\ge n_2} b_m \le v_1 + v_2 = v.
\end{aligned}$$

As v is arbitrary,

$$\limsup_{n\to\infty} a_n + b_n = \inf_{n\in\mathbb{N}} \sup_{m\ge n} a_m + b_m \le \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n.$$

(e) Putting (c) and (d) together,

$$\begin{aligned}
\liminf_{n\to\infty} a_n + b_n &= -\limsup_{n\to\infty}(-a_n) + (-b_n) \\
&\ge -\limsup_{n\to\infty}(-a_n) - \limsup_{n\to\infty}(-b_n) = \liminf_{n\to\infty} a_n + \liminf_{n\to\infty} b_n.
\end{aligned}$$

(f) Because $c \ge 0$,

$$\begin{aligned}
\limsup_{n\to\infty} ca_n &= \inf_{n\in\mathbb{N}} \sup_{m\ge n} ca_m = \inf_{n\in\mathbb{N}} c \sup_{m\ge n} a_m \\
&= c \inf_{n\in\mathbb{N}} \sup_{m\ge n} a_m = c \limsup_{n\to\infty} a_n.
\end{aligned}$$

(g) Finally,

$$\liminf_{n\to\infty} ca_n = -\limsup_{n\to\infty} c(-a_n) = -c \limsup_{n\to\infty}(-a_n) = c \liminf_{n\to\infty} a_n.$$

1A3C Remark Of course the familiar results that $\lim_{n\to\infty} a_n + b_n = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n$, $\lim_{n\to\infty} ca_n = c \lim_{n\to\infty} a_n$ are immediate corollaries of 1A3B.

***1A3D Other expressions of the same idea** The concepts of $\limsup$ and $\liminf$ may be applied in any context in which we can consider the limit of a real-valued function. For instance, if f is a real-valued function defined (at least) on a punctured interval of the form $\{x : 0 < |c - x| \le \epsilon\}$ where $c \in \mathbb{R}$ and $\epsilon > 0$, then

$$\limsup_{t\to c} f(t) = \lim_{\delta\downarrow 0} \sup_{0<|t-c|\le\delta} f(t) = \inf_{0<\delta\le\epsilon} \sup_{0<|t-c|\le\delta} f(t),$$

$$\liminf_{t\to c} f(t) = \lim_{\delta\downarrow 0} \inf_{0<|t-c|\le\delta} f(t) = \sup_{0<\delta\le\epsilon} \inf_{0<|t-c|\le\delta} f(t),$$

allowing ∞ and $-\infty$ whenever they seem called for. Or if f is defined on the half-open interval $]c, c + \epsilon]$, we can say

$$\limsup_{t\downarrow c} f(t) = \lim_{\delta\downarrow 0} \sup_{c<t\le c+\delta} f(t) = \inf_{0<\delta\le\epsilon} \sup_{c<t\le c+\delta} f(t),$$

$$\liminf_{t\downarrow c} f(t) = \lim_{\delta\downarrow 0} \inf_{c<t\le c+\delta} f(t) = \sup_{0<\delta\le\epsilon} \inf_{c<t\le c+\delta} f(t).$$

Similarly, if f is defined on $[M, \infty[$ for some $M \in \mathbb{R}$, we have

$$\limsup_{t\to\infty} f(t) = \lim_{a\to\infty} \sup_{t\ge a} f(t) = \inf_{a\ge M} \sup_{t\ge a} f(t),$$

$$\liminf_{t\to\infty} f(t) = \lim_{a\to\infty} \inf_{t\ge a} f(t) = \sup_{a\ge M} \inf_{t\ge a} f(t).$$

A further extension of the idea is examined briefly in 2A3S in Volume 2.

Concordance

I list here the section and paragraph numbers which have (to my knowledge) appeared in print in references to this volume, and which have since been changed.

112E-112F Image measures These paragraphs, referred to in the 2001 and 2003 editions of Volume 2, and the 2003 and 2006 editions of Volume 4, have been moved to 234C-234D in Volume 2.

112Ya Sums of measures This material, referred to in the 2001 and 2003 editions of Volume 2, has been moved to 234G in Volume 2.

121Yb (Σ, T)-measurable functions Exercise 121Yb in the 2000 and 2001 editions, referred to in the 2001 and 2003 editions of Volume 2, has been moved to 121Yc.

132E Measurable envelopes Parts (d) and (e) of 132E in the 2000 and 2001 editions, referred to in the 2001 edition of Volume 2 and the 2002 edition of Volume 3, are now parts (e) and (f).

132G Pull-back measures Proposition 132G, referred to in the 2006 edition of Volume 4, has been moved to 234F.

References for Volume 1

In addition to those (very few) works which I have mentioned in the course of this volume, I list some of the books from which I myself learnt measure theory, as a mark of grateful respect, and to give you an opportunity to sample alternative approaches.

Bartle R.G. [66] *The Elements of Integration.* Wiley, 1966.

Cohn D.L. [80] *Measure Theory.* Birkhäuser, 1980.

Facenda J.A. & Freniche F.J. [02] *Integración de funciones de varias variables.* Ediciones Pirámide, 2002.

Halmos P.R. [50] *Measure Theory.* Van Nostrand, 1950.

Lebesgue H. [1901] 'Sur une généralisation de l'intégrale définie', C.R.Acad.Sci. (Paris) 132 (1901) 1025-1027. Reprinted in Lebesgue 72, vol. 1.

Lebesgue H. [1904] *Leçons sur l'intégration et la recherche des fonctions primitives.* Gauthier-Villars, 1904. Reprinted in Lebesgue 72, vol. 2. [Vol. 1 *intro.*]

Lebesgue H. [1918] 'Remarques sur les théories de la mesure et de l'intégration', Ann. Sci. Ecole Norm. Sup. 35 (1918) 191-250. Reprinted in Lebesgue 72, vol. 2. [Vol. 1 *intro.*]

Lebesgue H. [72] *Oeuvres Scientifiques.* L'Enseignement Mathématique, Institut de Mathématiques, Univ. de Genève, 1972.

Munroe M.E. [53] *Introduction to Measure and Integration.* Addison-Wesley, 1953.

Royden H.L. [63] *Real Analysis.* Macmillan, 1963.

Solovay R.M. [70] 'A model of set theory in which every set of reals is Lebesgue measurable', Annals of Math. 92 (1970) 1-56. [134C.]

Widom H. [69] *Lectures on Measure and Integration.* Van Nostrand Reinhold, 1969.

Williamson J.H. [62] *Lebesgue Integration.* Holt, Rinehart & Winston, 1962.

Index to volume 1

Principal topics and results

The general index below is intended to be comprehensive. Inevitably the entries are voluminous to the point that they are often unhelpful. I have therefore prepared a shorter, better-annotated, index which will, I hope, help readers to focus on particular areas. It does not mention definitions, as the bold-type entries in the main index are supposed to lead efficiently to these; and if you draw blank here you should always, of course, try again in the main index. Entries in the form of mathematical assertions frequently omit essential hypotheses and should be checked against the formal statements in the body of the work.

Borel sets in $\mathbb{R}^r$ 111G
— and Lebesgue measure 114G, 115G, 134F

Cantor set and function 134G, 134H
Carathéodory's construction of measures from outer measures 113C
construction of measures
— from outer measures (Carathéodory's method) 113C
— subspace measures 131A
convergence theorems (B.Levi, Fatou, Lebesgue) §123
countable sets 111F, 1A1C *et seq.*
counting measure 112Bd

extended real line §135

Fatou's Lemma ($\int \liminf \le \liminf \int$ for sequences of non-negative functions) 123B

inner regularity of measures
— (with respect to compact sets) Lebesgue measure 134F
integration of real-valued functions, construction §122
— as a positive linear functional 122O
— characterization of integrable functions 122P, 122R
— functions and integrals with values in $[-\infty, \infty]$ 133A, 135F

Lebesgue measure, construction of §114, §115
— further properties §134
Lebesgue's Dominated Convergence Theorem ($\int \lim = \lim \int$ for dominated sequences of functions) 123C
B.Levi's theorem ($\int \lim = \lim \int$ for monotonic sequences of functions) 123A

measurable envelopes
— elementary properties 132E
measurable functions
— (real-valued) §121
— — sums, products and other operations on finitely many functions 121E
— — limits, infima, suprema 121F
measure space §112
Monotone Class Theorem 136B

non-measurable set (for Lebesgue measure) 134B

outer measures constructed from measures 132A *et seq.*
outer regularity of Lebesgue measure 134F
indexvheaderRothberger

subspace measures
— for measurable subspaces §131

σ-algebras of sets §111
— generated by given families 136B, 136G

General index

References in **bold** type refer to definitions; references in *italics* are passing references. Definitions marked with > are those in which my usage is dangerously at variance with that of some other author or authors.

additive functional on an algebra of sets *see* finitely additive (**136Xg**)
algebra of sets 113Yi, **136E**, 136F, 136G, 136Xg, 136Xh, 136Xk, 136Ya, 136Yb; *see also* σ-algebra (**111A**)
almost every, almost everywhere **112Dd**
almost surely **112De**
alternating functional **132Yf**
analytic (complex) function 133Xc
axiom *see* countable choice

Borel algebra *see* Borel σ-algebra (**111Gd**, **135C**)
Borel measurable function **121C**, 121D, 121Eg, 121H, 121K, **121Yd**, 134Fd, 134Xg, 134Yt, **135Ef**, 135Xc, 135Xe
Borel sets in $\mathbb{R}$, $\mathbb{R}^r$ **111G**, *111Yd*, 114G, 114Yd, 115G, 115Yb, *115Yd*, *121Ef*, *121K*, 134F, 134Xd, *135C*, 136D, 136Xj
Borel σ-algebra (of subsets of $\mathbb{R}^r$) **111Gd**, *114Yg-114Yi*, 121J, 121Xd, 121Xe, 121Yc, 121Yd;
— (of other spaces) **135C**, 135Xb
bounded set (in $\mathbb{R}^r$) **134E**

Cantor function **134H**, 134I
Cantor set **134G**, 134H, 134I, 134Xf
Carathéodory's method (of constructing measures) 113C, 113D, 113Xa, 113Xd, 113Xg, 113Yc, 113Yk, 114E, 114Xa, 115E, *121Ye*, 132Xc, 136Ya
characteristic function (of a set) **122Aa**
choice, axiom of 134C, 1A1G; *see also* countable choice
closed interval (in $\mathbb{R}$ or $\mathbb{R}^r$) 114G, 115G, 1A1A
closed set (in a topological space) 134Fb, 134Xd, **1A2E**, 1A2F, 1A2G
complete measure (space) **112Df**, 113Xa, 122Ya
complex-valued function §133
component (in a topological space) *111Ye*
conegligible set **112Dc**
continuous function 121D, 121Yd
convergent sequence **135D**
countable (set) **111F**, 114G, 115G, §1A1
countable choice, axiom of *134C*
counting measure **112Bd**, *122Xd*, 122 *notes*
cover *see* measurable envelope (**132D**)

Dedekind complete ordered set *135Ba*
derivative of a function*see* partial derivative
Devil's Staircase *see* Cantor function (**134H**)
differentiable function (of one variable) 123D
differentiating through an integral 123D
Dirac measure **112Bd**
direct image (of a set under a function or relation) 1A1B
disintegration of a measure 123Ye
disjoint family (of sets) **112Bb**
Dominated Convergence Theorem *see* Lebesgue's Dominated Convergence Theorem (**123C**)
Dynkin class **136A**, 136B, 136Xb

Egorov's theorem 131Ya
envelope *see* measurable envelope (**132D**)
equiveridical **121B**
Euclidean topology §1A2
extended real line *121C*, §135
extension of finitely additive functionals 113Yi
extension of measures 113Yc, 113Yh, 132Yd

Fatou's Lemma **123B**, 133Kb, 135Gb, 135Hb
field (of sets) *see* algebra (**136E**)
finitely additive function(al) on an algebra of sets 136Xg, 136Ya, 136Yb
Fourier series *121G*
Fourier transform **133Xd**, **133Yc**
full outer measure **132F**, 132Yd, 134D, 134Yt
function 1A1B

generated (σ-)algebra of sets **111Gb**, 111Xe, 111Xf, *121J*, 121Xd, 121Yc, 136B, 136C, 136G, 136H, 136Xc, 136Xk, 136Yb, 136Yc

half-open interval (in $\mathbb{R}$ or $\mathbb{R}^r$) **114Aa**, 114G, *114Yj*, **115Ab**, 115Xa, *115Xc*, *115Yd*

ideal in an algebra of sets *see* σ-ideal (**112Db**)
image measure **112Xf**, 123Ya, 132Yb
indefinite integral 131Xa
infinity 112B, 133A, §135
inner measure 113Yg
—— —— defined by a measure 113Yh
integrable function §122 (**122M**), *123Ya*, **133B**, **133Db**, 133F, 133Jd, 133Xa, *135Fa*
integral §122 (**122E**, **122K**, **122M**), **133B**, **133D**; *see also* integrable function, Lebesgue integral (**122Nb**), lower integral (**133I**), Riemann integral (**134K**), upper integral (**133I**)
interval *see* half-open interval (**114Aa**, **115Ab**), open interval (**111Xb**
inverse image (of a set under a function or relation) **1A1B**
inverse-measure-preserving function 134Yl-134Yn

Laplace transform **123Xc**, *123Yb*, **133Xd**
Lebesgue, H. Vol. 1 *intro.*
Lebesgue's Dominated Convergence Theorem 123C, 133G
Lebesgue integrable function **122Nb**, 122Yb, 122Ye, *122Yf*
Lebesgue integral **122Nb**
Lebesgue measurable function **121C**, 121D, 134Xg, 134Xj
Lebesgue measurable set **114E**, 114F, 114G, *114Xe*, 114Ye, **115E**, 115F, 115G
Lebesgue measure (on $\mathbb{R}$) §114 (**114E**), 131Xb, 133Xd, 133Xe, 134G-134L
—— —— (on $\mathbb{R}^r$) §115 (**115E**), 132C, *132Ef*, 133Yc, §134
—— —— (on other subsets of $\mathbb{R}^r$) **131B**
Lebesgue negligible set **114E**, **115E**, 134Yk
Lebesgue outer measure **114C**, 114D, 114Xc, *114Yd*, **115C**, 115D, 115E, 115Xb, 115Xd, 115Yb, 132C, 134A, 134D, 134Fa
Lebesgue-Stieltjes measure **114Xa**, *114Xb*, 114Yb, 114Yc, 114Yf, 131Xc, 132Xg, 134Xd
length of an interval **114Ab**
B.Levi's theorem **123A**, 123Xa, 133Ka, 135Ga, 135Hb
lower integral **133I**, 133J, 133Xf, 133Yd, 135Ha
lower Riemann integral **134Ka**
Lusin's theorem 134Yd

measurable cover *see* measurable envelope (**132D**)
measurable envelope **132D**, 132E, 132F, 132Xg, 134Fc, 134Xd; *see also* full outer measure (**132F**)
measurable function (taking values in $\mathbb{R}$) §121 (**121C**), 122Ya
—— —— (taking values in $\mathbb{R}^r$) **121Yd**
—— —— (taking values in other spaces) **133Da**, 133E, 133Yb, **135E**, *135Xd*, 135Yf
—— —— ((Σ, T)-measurable function) **121Yc**
—— —— *see also* Borel measurable, Lebesgue measurable
measurable set **112A**; *see also* relatively measurable (**121A**)
measurable space **111Bc**
measurable transformation *see* inverse-measure-preserving function
measure **112A**
—— (in 'μ measures E', 'E is measured by μ') **>112Be**

measure space §112 (**112A**)
Monotone Class Theorem 136B-136D, 136Xc, 136Xf
Monotone Convergence Theorem *see* B.Levi's theorem (123A)
monotonic function 121D

negligible set **112D**, 131Ca; *see also* Lebesgue negligible (**114E**, **115E**), null ideal (**112Db**)
non-decreasing sequence of sets **112Ce**
non-increasing sequence of sets **112Cf**
non-measurable set 134B, 134D, 134Xc
null ideal of a measure **112Db**
null set *see* negligible (**112Da**)

open interval *111Xb*, 114G, 115G, 1A1A
open set (in $\mathbb{R}^r$) **111Gc**, *111Yc*, *114Yd*, *115G*, *115Yb*, **133Xc**, 134Fa, 134Xe, **135Xa**, **1A2A**, 1A2B, 1A2D; (in $\mathbb{R}$) 111Gc, 111Ye, *114G*, 134Xd
outer measure §113 (**113A**), 114Xd, **132B**, *132Xg*, 136Ya; *see also* Lebesgue outer measure (**114C**, **115C**), regular outer measure (**132Xa**)
—— —— defined from a measure 113Ya, 132A-132E (**132B**), 132Xa-132Xi, 132Xk, 132Ya-132Yc, 133Je

partial derivative *123D*
partition **1A1J**
Peano curve 134Yl-134Yo
point-supported measure **112Bd**;*see also* Dirac measure (**112Bd**)
power set $\mathcal{P}\mathbb{N}$ 1A1Hb
presque partout **112De**
pseudo-simple function **122Ye**, 133Ye

quasi-simple function **122Yd**, 133Yd

regular outer measure 132C, **132Xa**
relation 1A1B
relatively measurable set **121A**
Riemann integrable function **134K**, 134L
Riemann integral **134K**

semi-ring of sets **115Ye**
Sierpiński Class Theorem *see* Monotone Class Theorem (136B)
simple function §122 (**>122A**)
space-filling curve 134Yl
Stieltjes measure *see* Lebesgue-Stieltjes measure (**114Xa**)
subspace measure 113Yb
—— —— on a measurable subset 131A, **131B**, 131C, 132Xb, 135I
—— —— (integration with respect to a subspace measure) **131D**, 131E-131H, 131Xa-131Xc, *133Dc*, 133Xb, 135I
subspace σ-algebra **121A**
sum over arbitrary index set 112Bd
sum of measures 112Yf, 122Xi
supported *see* point-supported (**112Bd**)

thick set *see* full outer measure (**132F**)
totally ordered set *135Ba*
trace (of a σ-algebra) *see* subspace σ-algebra (**121A**)
translation-invariant measure 114Xf, 115Xd, 134A, 134Ye, 134Yf

upper integral **133I**, 133J-133L, 133Xf, 133Yd, **135H**, 135I, 135Yc
upper Riemann integral **134Ka**

virtually measurable function **122Q**, 122Xe, 122Xf, 135Ia
Vitali's construction of a non-measurable set 134B
volume **115Ac**

a.e. ('almost everywhere') **112Dd**
a.s. ('almost surely') **112De**

$\mathbb{C}$ = the set of complex numbers

diam (in $\operatorname{diam} A$) = diameter
dom (in $\operatorname{dom} f$): the domain of a function f

$\mathcal{L}^0$ (in $\mathcal{L}^0(\mu)$) 121Xb
$\mathcal{L}^1$ (in $\mathcal{L}^1(\mu)$) *122Xc*
lim inf (in $\liminf_{n\to\infty}$) §1A3 (**1A3Aa**); (in $\liminf_{t\downarrow 0}$) **1A3D**
lim sup (in $\limsup_{n\to\infty}$) §1A3 (**1A3Aa**); (in $\limsup_{t\downarrow 0}$) **1A3D**

$\mathbb{N}\times\mathbb{N}$ 111Fb

$\mathcal{P}$ *see* power set
p.p. ('presque partout') **112De**

$\mathbb{Q}$ (the set of rational numbers) 111Eb, 1A1Ef

$\mathbb{R}$ (the set of real numbers) 111Fe, 1A1Ha
$\overline{\mathbb{R}}$ *see* extended real line (§135)

U (in $U(x,\delta)$) **1A2A**

$\mathbb{Z}$ (the set of integers) 111Eb, 1A1Ee
ZFC *see* Zermelo-Fraenkel set theory

π-λ Theorem *see* Monotone Class Theorem (136B)

σ-algebra of sets **111A**, 111B, 111D-111G, 111Xc-111Xf, 111Yb, 136Xb, 136Xi; *see also* Borel σ-algebra (**111G**))
σ-field *see* σ-algebra (**111A**)
σ-ideal (of sets) **112Db**

$\sum_{i\in I} a_i$ **112Bd**

χ (in χA, where A is a set) **122Aa**

$\setminus$ (in $E\setminus F$, 'set difference') **111C**
$\triangle$ (in $E\triangle F$, 'symmetric difference') **111C**
$\bigcup$ (in $\bigcup_{n\in\mathbb{N}} E_n$) **111C**; (in $\bigcup\mathcal{A}$) **1A1F**
$\bigcap$ (in $\bigcap_{n\in\mathbb{N}} E_n$) **111C**; (in $\bigcap\mathcal{E}$) **1A2F**
$\vee$, $\wedge$ (in a lattice) 121Xb
$\restriction$ (in $f\restriction A$, the restriction of a function to a set) **121Eh**
$=_{\text{a.e.}}$ **112Dg**, 112Xe
$\leq_{\text{a.e.}}$ **112Dg**, 112Xe
$\geq_{\text{a.e.}}$ **112Dg**, 112Xe
* (in μ^*) *see* outer measure defined by a measure (**132B**)
$_*$ (in μ_*) *see* inner measure defined by a measure (113Yh)
$\int$ (in $\int f$, $\int f d\mu$, $\int f(x)\mu(dx)$) **122E**, **122K**, **122M**, 122Nb; *see also* upper integral, lower integral (**133I**)
— (in $\int_A f$) **131D**; *see also* subspace measure
$\overline{\int}$ *see* upper integral (**133I**)
$\underline{\int}$ *see* lower integral (**133I**)
$\mathrm{R}\!\!\int$ *see* Riemann integral (**134K**)
$^+$ (in f^+, where f is a function) **121Xa**
$^-$ (in f^-, where f is a function) **121Xa**
∞ *see* infinity
[ ] (in $[a,b]$) *see* closed interval (**115G**, **1A1A**); (in $f[A]$, $f^{-1}[B]$, $R[A]$, $R^{-1}[B]$) **1A1B**
[[(in $[a,b[$) *see* half-open interval (**115Ab**, **1A1A**)
]] (in $]a,b]$) *see* half-open interval (**1A1A**)
] [(in $]a,b[$) *see* open interval (**115G**, **1A1A**)