OXFORD

Introduction to Modern Analysis

Second Edition

Shmuel Kantorovitz and Ami Viselter

OXFORD GRADUATE TEXTS IN MATHEMATICS | 29

Introduction to Modern Analysis

Oxford Graduate Texts in Mathematics

OXFORD GRADUATE TEXTS IN MATHEMATICS

1. Keith Hannabuss: *An Introduction to Quantum Theory*
2. Reinhold Meise and Dietmar Vogt: *Introduction to Functional Analysis*
3. James G. Oxley: *Matroid Theory*
4. N. J. Hitchin, G. B. Segal, and R. S. Ward: *Integrable Systems: Twistors, Loop Groups, and Riemann Surfaces*
5. Wulf Rossmann: *Lie Groups: An Introduction Through Linear Groups*
6. Qing. Liu: *Algebraic Geometry and Arithmetic Curves*
7. Martin R. Bridson and Simon M, Salamon (eds): *Invitations to Geometry and Topology*
8. Shmuel Kantorovitz: *Introduction to Modern Analysis*
9. Terry Lawson: *Topology: A Geometric Approach*
10. Meinolf Geck: *An Introduction to Algebraic Geometry and Algebraic Groups*
11. Alastair Fletcher and Vladimir Markovic: *Quasiconformal Maps and Teichmüller Theory*
12. Dominic Joyce: *Riemannian Holonomy Groups and Calibrated Geometry*
13. Fernando Villegas: *Experimental Number Theory*
14. Péter Medvegyev: *Stochastic Integration Theory*
15. Martin A. Guest: *From Quantum Cohomology to Integrable Systems*
16. Alan D. Rendall: *Partial Differential Equations in General Relativity*
17. Yves Félix, John Oprea, and Daniel Tanré: *Algebraic Models in Geometry*
18. Jie Xiong: *Introduction to Stochastic Filtering Theory*
19. Maciej Dunajski: *Solitons, Instantons, and Twistors*
20. Graham R. Allan: *Introduction to Banach Spaces and Algebras*
21. James Oxley: *Matroid Theory, Second Edition*
22. Simon Donaldson: *Riemann Surfaces*
23. Clifford Henry Taubes: *Differential Geometry: Bundles, Connections, Metrics and Curvature*
24. Gopinath Kallianpur and P. Sundar: *Stochastic Analysis and Diffusion Processes*
25. Selman Akbulut: *4-Manifolds*
26. Fon-Che Liu: *Real Analysis*
27. Dusa McDuff and Dietmar Salamon: *Introduction to Symplectic Topology, Third Edition*
28. Chris Heunen, Jamie Vicary: *Categories for Quantum Theory: An Introduction*
29. Shmuel Kantorovitz, Ami Viselter: *Introduction to Modern Analysis, Second Edition*

Introduction to Modern Analysis

Second Edition

Shmuel Kantorovitz

Bar Ilan University, Israel

Ami Viselter

University of Haifa, Israel

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

Impression: 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2022933188

ISBN 978–0–19–284954–0 (hbk)
ISBN 978–0–19–284955–7 (pbk)

DOI: 10.1093/oso/9780192849540.001.0001

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

To Ita, Bracha, Pnina, Pinchas, Ruth, and Lilach

Contents

Preface to the First Edition

This book grew out of lectures given since 1964 at Yale University, the University of Illinois at Chicago, and Bar Ilan University. The material covers the usual topics of Measure Theory and Functional Analysis, with applications to Probability Theory and to the theory of linear partial differential equations. Some relatively advanced topics are included in each chapter (excluding the first two): the Riesz–Markov representation theorem and differentiability in Euclidean spaces (Chapter 3); Haar measure (Chapter 4); Marcinkiewicz's interpolation theorem (Chapter 5); the Gelfand–Naimark–Segal representation theorem (Chapter 7); the von Neumann double commutant theorem (Chapter 8); the spectral representation theorem for normal operators (Chapter 9); the extension theory for unbounded symmetric operators (Chapter 10); the Lyapounov Central Limit theorem and the Kolmogoroff "Three Series theorem" (Application I); the Hormander–Malgrange theorem, fundamental solutions of linear partial differential equations with variable coefficients, and Hormander's theory of convolution operators, with an application to integration of pure imaginary order (Application I). Some important complementary material is included in the 'Exercises' sections, with step-by-step detailed hints leading to the wanted results. Solutions to the end of chapter exercises may be found on the companion website for this text: http://www.oup.co.uk/academic/companion/mathematics/kantorovitz.

Ramat Gan S. K.
July 2002

Preface to the Second Edition

The purpose of the second edition is to make our *Introduction to Modern Analysis* more modern. We did this mostly by broadening and deepening the presentation of operator algebras, which form a central area in functional analysis. There are three new chapters: Chapter 11 on C^*-algebras, Chapter 12 on von Neumann algebras, and Chapter 13 on constructions of C^*-algebras. They contain much more material on these subjects than the first edition. These chapters are also more advanced than the previous parts of the book and require more from the reader, occasionally in the form of guided exercises. Nevertheless, what we give here is merely a taste of operator algebras.

In addition, we made numerous corrections and added quite a lot of exercises. There are also new subjects of independent interest: fixed-point theorems (Chapter 5); the bounded *weak**-topology (Chapter 5); the Arens products (Chapter 7); tensor products of vector spaces and of Hilbert spaces (Chapter 8); and quadratic forms (Chapter 10).

Ramat Gan and Haifa S. K. and A. V.
December 2021

1

Measures

This chapter begins the study of measure theory, which spans Chapters 1–3 and most of Chapter 4. Let us explain first what necessitated this theory.

Consider the set of all Riemann integrable functions on an interval $[a, b]$. It becomes a semi-normed space with respect to the semi-norm $\|f\| := \int_a^b |f(x)|dx$. The problem is that this space is *not complete*: it admits non-convergent Cauchy sequences. As discussed later in the book, in modern analysis it is especially important for (semi-) normed spaces to be complete. Measure theory, invented by H. L. Lebesgue, introduces the concept of a *measure space*, which is a triple $(X, \mathcal{A}, \mu)$, where X is a set, $\mathcal{A}$ is the σ-algebra of all *measurable subsets* of X, and $\mu : \mathcal{A} \to [0, \infty]$ is a *measure*. To each such triple there is an associated *Lebesgue integral*. In the particular case when $X = [a, b]$ and μ is the *Lebesgue measure* (very roughly, $\mu([c, d]) = d - c$, which justifies the word "measure"), every Riemann integrable function is also Lebesgue integrable (but not conversely!) and the integrals coincide.

One big virtue of Lebesgue integration is that the space of integrable functions that comes out of it is complete. In fact, to every measure space we associate not one complete normed space, but a continuum of them—the so-called L^p-spaces ($p \in [1, \infty]$).

The chapter is structured as follows. We first introduce positive measure spaces and Lebesgue integration on them and prove several *convergence theorems* that are fundamental in the theory. We then define the L^p-spaces and prove that they are *Banach spaces*, that is, complete normed spaces. Next, we prove a few basic facts on Hilbert spaces culminating in the *"little" Riesz representation theorem* (we return to Hilbert spaces in Chapter 8). Hilbert spaces are needed in the proof of the *Lebesgue–Radon–Nikodym theorem*, a deep result about the relationship between two arbitrary measures. *Complex* measures are then introduced and studied. A few notions of *convergence* of sequences of measurable functions are defined and the relations between them are explained, including a surprising theorem of *Egoroff* saying that on *finite* measure spaces, pointwise convergence is "almost uniform". A short treatment of the *distribution function* of

Introduction to Modern Analysis. Second Edition. Shmuel Kantorovitz and Ami Viselter, Oxford University Press.
 DOI: 10.1093/oso/9780192849540.003.0001

a measurable function follows. The chapter ends with the notion of a *truncation* of a function.

1.1 Measurable sets and functions

The setting of abstract measure theory is a family $\mathcal{A}$ of so-called *measurable* subsets of a given set X, and a function

$$\mu : \mathcal{A} \to [0, \infty],$$

so that the *measure* $\mu(E)$ of the set $E \in \mathcal{A}$ has some "intuitively desirable" property, such as "countable additivity":

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i),$$

for mutually disjoint sets $E_i \in \mathcal{A}$. In order to make sense, this setting has to deal with a family $\mathcal{A}$ that is closed under countable unions. We then arrive to the concept of a *measurable space*.

Definition 1.1. Let X be a (non-empty) set. A σ-algebra of subsets of X (briefly, a σ-algebra *on* X) is a subfamily $\mathcal{A}$ of the family $\mathbb{P}(X)$ of all subsets of X, with the following properties:

(1) $X \in \mathcal{A}$;

(2) if $E \in \mathcal{A}$, then the complement E^{c} of E belongs to $\mathcal{A}$;

(3) if $\{E_i\}$ is a *sequence* of sets in $\mathcal{A}$, then its *union* belongs to $\mathcal{A}$.

The ordered pair $(X, \mathcal{A})$, with $\mathcal{A}$ a σ-algebra on X, is called a *measurable space*. The sets of the family $\mathcal{A}$ are called *measurable sets* (or $\mathcal{A}$-measurable sets) in X.

Observe that by (1) and (2), the empty set $\emptyset$ belongs to the σ-algebra $\mathcal{A}$. Taking then $E_i = 0$ for all $i > n$ in (3), we see that $\mathcal{A}$ is closed under *finite unions*; if this weaker condition replaces (3), $\mathcal{A}$ is called an *algebra* of subsets of X (briefly, an algebra *on* X).

By (2) and (3), and DeMorgan's Law, $\mathcal{A}$ is closed under countable intersections (finite intersections, in the case of an algebra). In particular, any algebra on X is closed under differences $E - F := E \cap F^{\mathrm{c}}$.

The intersection of an arbitrary family of σ-algebras on X is a σ-algebra on X. If all the σ-algebras in the family contain some fixed collection $\mathcal{E} \subset \mathbb{P}(X)$, the said intersection is the smallest σ-algebra on X (with respect to set inclusion) that contains $\mathcal{E}$; it is called *the σ-algebra generated by $\mathcal{E}$*, and is denoted by $[\mathcal{E}]$.

An important case comes up naturally when X is a topological space (for some topology τ). The σ-algebra $[\tau]$ generated by the topology is called the *Borel* (σ)-algebra [denoted $\mathcal{B}(X)$], and the sets in $\mathcal{B}(X)$ are the *Borel sets* in X. For example, the countable intersection of τ-open sets (a so-called G_δ-set) and the countable union of τ-closed sets (a so-called F_σ-set) are Borel sets.

Definition 1.2. Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces. A map $f : X \to Y$ is *measurable* if for each $B \in \mathcal{B}$, the set

$$f^{-1}(B) := \{x \in X; f(x) \in B\} := [f \in B]$$

belongs to $\mathcal{A}$.

A constant map $f(x) = p \in Y$ is trivially measurable, since $[f \in B]$ is either $\emptyset$ or X (when $p \in B^{\mathrm{c}}$ and $p \in B$, respectively), and so belongs to $\mathcal{A}$.

When Y is a topological space, we shall usually take $\mathcal{B} = \mathcal{B}(Y)$, the Borel algebra on Y. In particular, for $Y = \mathbb{R}$ (the real line), $Y = [-\infty, \infty]$ (the "extended real line"), or $Y = \mathbb{C}$ (the complex plane), with their usual topologies, we shall call the measurable map a *measurable function* (more precisely, an $\mathcal{A}$-measurable function). If X is a topological space, a $\mathcal{B}(X)$-measurable map (function) is called a Borel map (function).

Given a measurable space $(X, \mathcal{A})$ and a map $f : X \to Y$, for an arbitrary set Y, the family

$$\mathcal{B}_f := \{F \in \mathbb{P}(Y); f^{-1}(F) \in \mathcal{A}\}$$

is a σ-algebra on Y (because the inverse image operation preserves the set theoretical operations: $f^{-1}(\bigcup_\alpha F_\alpha) = \bigcup_\alpha f^{-1}(F_\alpha)$, etc.), and it is the largest σ-algebra on Y for which f is measurable.

If Y is a topological space, and $f^{-1}(V) \in \mathcal{A}$ for every *open* V, then $\mathcal{B}_f$ contains the topology τ, and so contains $\mathcal{B}(Y)$; that is, f is measurable. Since $\tau \subset \mathcal{B}(Y)$, the converse is trivially true.

Lemma 1.3. *A map f from a measurable space $(X, \mathcal{A})$ to a topological space Y is measurable if and only if $f^{-1}(V) \in \mathcal{A}$ for every open $V \subset Y$.*

In particular, if X is also a topological space, and $\mathcal{A} = \mathcal{B}(X)$, it follows that every continuous map $f : X \to Y$ is a Borel map.

Lemma 1.4. *A map f from a measurable space $(X, \mathcal{A})$ to $[-\infty, \infty]$ is measurable if and only if*

$$[f > c] \in \mathcal{A}$$

for all real c.

The non-trivial direction in the lemma follows from the fact that $(c, \infty] \in \mathcal{B}_f$ by hypothesis for all real c; therefore, the σ-algebra $\mathcal{B}_f$ contains the sets

$$\bigcup_{n=1}^{\infty} (b - 1/n, \infty]^{\mathrm{c}} = \bigcup_{n=1}^{\infty} [-\infty, b - 1/n] = [-\infty, b)$$

and $(a, b) = [-\infty, b) \cap (a, \infty]$ for every real $a < b$, and so contains all countable unions of "segments" of the above type, that is, all open subsets of $[-\infty, \infty]$.

The sets $[f > c]$ in the condition of Lemma 1.4 can be replaced by any of the sets $[f \geq c], [f < c]$, or $[f \leq c]$ (for all real c), respectively. The proofs are analogous.

For $f : X \to [-\infty, \infty]$ measurable and α real, the function αf (defined pointwise, with the usual arithmetics $\alpha \cdot \infty = \infty$ for $\alpha > 0, = 0$ for $\alpha = 0$, and $= -\infty$ for $\alpha < 0$, and similarly for $-\infty$) is measurable, because for all real $c, [\alpha f > c] = [f > c/\alpha]$ for $\alpha > 0, = [f < c/\alpha]$ for $\alpha < 0$, and αf is constant for $\alpha = 0$.

If $\{a_n\} \subset [-\infty, \infty]$, one denotes the superior (inferior) limit, that is, the "largest" ("smallest") limit point, of the sequence by $\limsup a_n$ ($\liminf a_n$, respectively).

Let $b_n := \sup_{k \geq n} a_k$. Then $\{b_n\}$ is a decreasing sequence, and therefore

$$\exists \lim_n b_n = \inf_n b_n.$$

Let $\alpha := \limsup a_n$ and $\beta = \lim b_n$. For any given $n \in \mathbb{N}, a_k \leq b_n$ for all $k \geq n$, and therefore $\alpha \leq b_n$. Hence $\alpha \leq \beta$.

On the other hand, for any $t > \alpha, a_k > t$ for at most *finitely many* indices k. Therefore, there exists n_0 such that $a_k \leq t$ for all $k \geq n_0$, hence $b_{n_0} \leq t$. But then $b_n \leq t$ for all $n \geq n_0$ (because $\{b_n\}$ is decreasing), and so $\beta \leq t$. Since $t > \alpha$ was arbitrary, it follows that $\beta \leq \alpha$, and the conclusion $\alpha = \beta$ follows. We showed

$$\limsup a_n = \lim_n \left(\sup_{k \geq n} a_k \right) = \inf_{n \in \mathbb{N}} \left(\sup_{k \geq n} a_k \right). \tag{1}$$

Similarly

$$\liminf a_n = \lim_n \left(\inf_{k \geq n} a_k \right) = \sup_{n \in \mathbb{N}} \left(\inf_{k \geq n} a_k \right). \tag{2}$$

Lemma 1.5. *Let $\{f_n\}$ be a sequence of measurable $[-\infty, \infty]$-valued functions on the measurable space $(X, \mathcal{A})$. Then the functions $\sup f_n, \inf f_n, \limsup f_n, \lim$ $\inf f_n$, and $\lim f_n$ (when it exists), all defined pointwise, are measurable.*

Proof. Let $h = \sup f_n$. Then for all real c,

$$[h > c] = \bigcup_n [f_n > c] \in \mathcal{A},$$

so that h is measurable by Lemma 1.4.

As remarked, $-f_n = (-1)f_n$ are measurable, and therefore $\inf f_n = -\sup$ $(-f_n)$ is measurable.

The proof is completed by the relations (1), (2), and

$$\lim f_n = \limsup f_n = \liminf f_n,$$

when the second equality holds (i.e. if and only if $\lim f_n$ exists). □

In particular, taking a sequence with $f_k = f_n$ for all $k > n$, we see that $\max\{f_1, \ldots, f_n\}$ and $\min\{f_1, \ldots, f_n\}$ are measurable, when $f_1, \ldots, f_n$ are measurable functions into $[-\infty, \infty]$. For example, the *positive* (*negative*) parts $f^+ := \max\{f, 0\}$ ($f^- := -\min\{f, 0\}$) of a measurable function $f : X \to [-\infty, \infty]$ are (non-negative) measurable functions. Note the decompositions

$$f = f^+ - f^-; \quad |f| = f^+ + f^-.$$

Lemma 1.6. *Let $(X, \mathcal{A})$, $(Y, \mathcal{B})$ and $(Z, \mathcal{C})$ be measurable spaces. If $f : X \to Y$ and $g : Y \to Z$ are measurable, then so is the composite function $h := g \circ f : X \to Z$.*

Indeed, for every $C \in \mathcal{C}$ we have $g^{-1}(C) \in \mathcal{B}$ by measurability of g, thus $h^{-1}(C) = f^{-1}(g^{-1}(C)) \in \mathcal{A}$ by measurability of f.

In particular, if Y, Z are topological spaces and $g : Y \to Z$ is continuous, then $g \circ f$ is measurable.

If

$$Y = \prod_{k=1}^{n} Y_k$$

is the product space of topological spaces Y_k, the projections $p_k : Y \to Y_k$ are continuous. Therefore, if $f : X \to Y$ is measurable, so are the "component functions" $f_k(x) := p_k(f(x)) : X \to Y_k$ $(k = 1, \dots, n)$, by Lemma 1.6. Conversely, if the topologies on Y_k have countable bases (for all k), a countable base for the topology of Y consists of sets of the form $V = \prod_{k=1}^{n} V_k$ with V_k varying in a countable base for the topology of Y_k (for each k). Now,

$$[f \in V] = \bigcap_{k=1}^{n} [f_k \in V_k] \in \mathcal{A}$$

if all f_k are measurable. Since every open $W \subset Y$ is a countable union of sets of the above type, $[f \in W] \in \mathcal{A}$, and f is measurable. We proved:

Lemma 1.7. *Let Y be the Cartesian product of topological spaces $Y_1, \dots, Y_n$ with countable bases to their topologies. Let $(X, \mathcal{A})$ be a measurable space. Then $f : X \to Y$ is measurable iff the components f_k are measurable for all k.*

For example, if $f_k : X \to \mathbb{C}$ are measurable for $k = 1, \dots, n$, then $f := (f_1, \dots, f_n) : X \to \mathbb{C}^n$ is measurable, and since $g(z_1, \dots, z_n) := \Sigma \alpha_k z_k (\alpha_k \in \mathbb{C})$ and $h(z_1, \dots, z_n) = z_1 \dots z_n$ are continuous from $\mathbb{C}^n$ to $\mathbb{C}$, it follows from Lemma 1.6 that (finite) linear combinations and products of complex measurable functions are measurable. Thus, the complex measurable functions form an algebra over the complex field (similarly, the real measurable functions form an algebra over the real field), for the usual pointwise operations.

If f has values in $\mathbb{R}, [-\infty, \infty]$, or $\mathbb{C}$, its measurability implies that of $|f|$, by Lemma 1.6.

By Lemma 1.7, a complex function is measurable iff its real part $\Re f$ and imaginary part $\Im f$ are both measurable.

If f, g are measurable with values in $[0, \infty]$, the functions $f + g$ and fg are well-defined pointwise (with values in $[0, \infty]$) and measurable, since the functions $(s, t) \to s + t$ and $(s, t) \to st$ from $[0, \infty]^2$ to $[0, \infty]$ are Borel (cf. Lemmas 1.6 and 1.7).

The function $f : X \to \mathbb{C}$ is *simple* if its range is a *finite* set $\{c_1, \dots, c_n\} \subset \mathbb{C}$. Let $E_k := [f = c_k]$, $\quad k = 1, \dots, n$. Then X is the disjoint union of the sets E_k, and

$$f = \sum_{k=1}^{n} c_k I_{E_k},$$

where I_E denotes the *indicator* of E (also called the *characteristic function* of E by non-probabilists, while probabilists reserve the later name to a different concept):

$$I_E(x) = 1 \quad \text{for } x \in E \quad \text{and} \quad = 0 \quad \text{for } x \in E^{\mathrm{c}}.$$

Since a singleton $\{c\} \subset \mathbb{C}$ is closed, it is a Borel set. Suppose now that the simple (complex) function f is defined on a measurable space $(X, \mathcal{A})$. If f is measurable, then $E_k := [f = c_k]$ is measurable for all $k = 1, \ldots, n$. Conversely, if all E_k are measurable, then for each open $V \subset \mathbb{C}$,

$$[f \in V] = \bigcup_{\{k; c_k \in V\}} E_k \in \mathcal{A},$$

so that f is measurable. In particular, an indicator I_E is measurable iff $E \in \mathcal{A}$.

Let $B(X, \mathcal{A})$ denote the complex algebra of all *bounded* complex $\mathcal{A}$-measurable functions on X (for the pointwise operations), and denote

$$\|f\| = \sup_X |f| \quad (f \in B(X, \mathcal{A})).$$

The map $f \to \|f\|$ of $B(X, \mathcal{A})$ into $[0, \infty)$ has the following properties:

(1) $\|f\| = 0$ iff $f = 0$ (the zero function);

(2) $\|\alpha f\| = |\alpha| \, \|f\|$ for all $\alpha \in \mathbb{C}$ and $f \in B(X, \mathcal{A})$;

(3) $\|f + g\| \leq \|f\| + \|g\|$ for all $f, g \in B(X, \mathcal{A})$;

(4) $\|fg\| \leq \|f\| \, \|g\|$ for all $f, g \in B(X, \mathcal{A})$.

For example, (3) is verified by observing that for all $x \in X$,

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \sup_X |f| + \sup_X |g|.$$

A map $\|\cdot\|$ from any (complex) vector space Z to $[0, \infty)$ with Properties (1)–(3) is called a *norm* on Z. The previous example is the *supremum norm* or *uniform norm* on the vector space $Z = B(X, \mathcal{A})$. Property (1) is the *definiteness* of the norm; Property (2) is its *homogeneity*; Property (3) is the *triangle inequality*. A vector space with a specified norm is a *normed space*. If Z is an algebra, and the specified norm satisfies Property (4) also, Z is called a *normed algebra*. Thus, $B(X, \mathcal{A})$ is a normed algebra with respect to the supremum norm. Any normed space Z is a metric space for the metric *induced by the norm*

$$d(u, v) := \|u - v\| \quad u, v \in Z.$$

Convergence in Z is convergence with respect to this metric (unless stated otherwise). Thus, convergence in the normed space $B(X, \mathcal{A})$ is precisely *uniform convergence* on X (this explains the name “uniform norm”).

If $x, y \in Z$, the triangle inequality implies $\|x\| = \|(x-y)+y\| \leq \|x-y\|+\|y\|$, so that $\|x\| - \|y\| \leq \|x - y\|$. Since we may interchange x and y, we have

$$\big|\|x\| - \|y\|\,\big| \leq \|x - y\|.$$

In particular, *the norm function is continuous on Z.*

The simple functions in $B(X, \mathcal{A})$ form a subalgebra $B_0(X, \mathcal{A})$; it is *dense* in $B(X, \mathcal{A})$:

Theorem 1.8 (Approximation theorem). *Let $(X, \mathcal{A})$ be a measurable space. Then:*

(1) $B_0(X, \mathcal{A})$ is dense in $B(X, \mathcal{A})$ (i.e., every bounded complex measurable function is the uniform limit of a sequence of simple measurable complex functions).

(2) If $f : X \to [0, \infty]$ is measurable, then there exists a sequence of measurable simple functions

$$0 \leq \phi_1 \leq \phi_2 \leq \cdots \leq f,$$

such that $f = \lim \phi_n$.

Proof. (1) Since any $f \in B(X, \mathcal{A})$ can be written as

$$f = u^+ - u^- + \mathrm{i}v^+ - \mathrm{i}v^-$$

with $u = \Re f$ and $v = \Im f$, it suffices to prove (1) for f with range in $[0, \infty)$. Let N be the first integer such that $N > \sup f$. For $n = 1, 2, \ldots$, set

$$\phi_n := \sum_{k=1}^{N2^n} \frac{k-1}{2^n} I_{E_{n,k}},$$

where

$$E_{n,k} := f^{-1}\left(\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right)\right).$$

The simple functions ϕ_n are measurable,

$$0 \leq \phi_1 \leq \phi_2 \leq \cdots \leq f,$$

and

$$0 \leq f - \phi_n < \frac{1}{2^n},$$

so that indeed $\|f - \phi_n\| \leq (1/2^n)$, as wanted.

If f has range in $[0, \infty]$, set

$$\phi_n := \sum_{k=1}^{n2^n} \frac{k-1}{2^n} I_{E_{n,k}} + nI_{F_n},$$

where $F_n := [f \geq n]$. Again $\{\phi_n\}$ is a non-decreasing sequence of non-negative measurable simple functions $\leq f$. If $f(x) = \infty$ for some $x \in X$, then $x \in F_n$

for all n, and therefore $\phi_n(x) = n$ for all n; hence $\lim_n \phi_n(x) = \infty = f(x)$. If $f(x) < \infty$ for some x, let $n > f(x)$. Then there exists a unique $k, 1 \leq k \leq n2^n$, such that $x \in E_{n,k}$. Then $\phi_n(x) = ((k-1)/2^n)$ while $((k-1)/2^n) \leq f(x) < (k/2^n)$, so that

$$0 \leq f(x) - \phi_n(x) < 1/2^n \quad (n > f(x)).$$

Hence $f(x) = \lim_n \phi_n(x)$ for all $x \in X$. □

1.2 Positive measures

Definition 1.9. Let $(X, \mathcal{A})$ be a measurable space. A (*positive*) *measure* on $\mathcal{A}$ is a function

$$\mu : \mathcal{A} \to [0, \infty]$$

such that $\mu(\emptyset) = 0$ and

$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mu(E_k) \tag{1}$$

for any sequence of mutually disjoint sets $E_k \in \mathcal{A}$. Property (1) is called *σ-additivity* of the function μ. The ordered triple $(X, \mathcal{A}, \mu)$ will be called a (*positive*) *measure space.*

Taking in particular $E_k = \emptyset$ for all $k > n$, it follows that

$$\mu\left(\bigcup_{k=1}^{n} E_k\right) = \sum_{k=1}^{n} \mu(E_k) \tag{2}$$

for any *finite* collection of mutually disjoint sets $E_k \in \mathcal{A}, k = 1, \ldots, n$. We refer to Property (2) by saying that μ is (*finitely*) *additive.*

Any finitely additive function $\mu \geq 0$ on an algebra $\mathcal{A}$ is necessarily *monotonic*, that is, $\mu(E) \leq \mu(F)$ when $E \subset F (E, F \in \mathcal{A})$; indeed

$$\mu(F) = \mu(E \cup (F - E)) = \mu(E) + \mu(F - E) \geq \mu(E).$$

If $\mu(E) < \infty$, we get

$$\mu(F - E) = \mu(F) - \mu(E).$$

Lemma 1.10. *Let $(X, \mathcal{A}, \mu)$ be a positive measure space, and let*

$$E_1 \subset E_2 \subset E_3 \subset \cdots$$

be measurable sets with union E. Then

$$\mu(E) = \lim_n \mu(E_n).$$

Proof. The sets E_n and E can be written as *disjoint* unions

$$E_n = E_1 \cup (E_2 - E_1) \cup (E_3 - E_2) \cup \cdots \cup (E_n - E_{n-1}),$$
$$E = E_1 \cup (E_2 - E_1) \cup (E_3 - E_2) \cup \cdots ,$$

where all differences belong to $\mathcal{A}$. Set $E_0 = \emptyset$. By σ-additivity,

$$\begin{aligned}\mu(E) &= \sum_{k=1}^{\infty} \mu(E_k - E_{k-1}) \\ &= \lim_n \sum_{k=1}^{n} \mu(E_k - E_{k-1}) = \lim_n \mu(E_n).\end{aligned}$$

□

In general, if E_j belong to an *algebra* $\mathcal{A}$ of subsets of X, set $A_0 = \emptyset$ and $A_n = \bigcup_{j=1}^n E_j$, $n = 1, 2, \ldots$. The sets $A_j - A_{j-1}, 1 \le j \le n$, are disjoint $\mathcal{A}$-measurable subsets of E_j with union A_n. If μ is a non-negative *additive* set function on $\mathcal{A}$, then

$$\mu\Big(\bigcup_{j=1}^{n} E_j\Big) = \mu(A_n) = \sum_{j=1}^{n} \mu(A_j - A_{j-1}) \le \sum_{j=1}^{n} \mu(E_j). \tag{*}$$

This is the *subadditivity property* of non-negative additive set functions (on algebras).

If $\mathcal{A}$ is a σ-algebra and μ is a positive measure on $\mathcal{A}$, then since $A_1 \subset A_2 \subset \cdots$ and $\bigcup_{n=1}^\infty A_n = \bigcup_{j=1}^\infty E_j$, letting $n \to \infty$ in (*), it follows from Lemma 1.10 that

$$\mu\left(\bigcup_{j=1}^{\infty} E_j\right) \le \sum_{j=1}^{\infty} \mu(E_j).$$

This property of positive measures is called σ-*subadditivity.*

For *decreasing* sequences of measurable sets, the "dual" of Lemma 1.10 is false in general, unless we assume that the sets have *finite* measure:

Lemma 1.11. *Let $\{E_k\} \subset \mathcal{A}$ be a decreasing sequence (with respect to set-inclusion) such that $\mu(E_1) < \infty$. Let $E = \bigcap_k E_k$. Then*

$$\mu(E) = \lim_n \mu(E_n).$$

Proof. The sequence $\{E_1 - E_k\}$ is increasing, with union $E_1 - E$. By Lemma 1.10 and the *finiteness* of the measures of E and E_k (subsets of E_1!),

$$\begin{aligned}\mu(E_1) - \mu(E) &= \mu\left(\bigcup_k (E_1 - E_k)\right) \\ &= \lim \mu(E_1 - E_n) = \mu(E_1) - \lim \mu(E_n),\end{aligned}$$

and the result follows by cancelling the *finite* number $\mu(E_1)$. □

If $\{E_k\}$ is an *arbitrary* sequence of subsets of X, set $F_n = \bigcap_{k\geq n} E_k$ and $G_n = \bigcup_{k\geq n} E_k$. Then $\{F_n\}$ ($\{G_n\}$) is increasing (decreasing, respectively), and $F_n \subset E_n \subset G_n$ for all n.

One defines

$$\liminf_n E_n := \bigcup_n F_n; \quad \limsup_n E_n := \bigcap_n G_n.$$

These sets belong to $\mathcal{A}$ if $E_k \in \mathcal{A}$ for all k. The set $\liminf E_n$ consists of all x that *belong to E_n for all but finitely many n*; the set $\limsup E_n$ consists of all x that *belong to E_n for infinitely many n*. By Lemma 1.10,

$$\mu(\liminf E_n) = \lim_n \mu(F_n) \leq \liminf \mu(E_n). \tag{3}$$

If the measure of G_1 is finite, we also have by Lemma 1.11

$$\mu(\limsup E_n) = \lim_n \mu(G_n) \geq \limsup \mu(E_n). \tag{4}$$

1.3 Integration of non-negative measurable functions

Definition 1.12. Let $(X, \mathcal{A}, \mu)$ be a positive measure space, and $\phi : X \to [0, \infty)$ a measurable simple function. The *integral over X of ϕ with respect to μ*, denoted

$$\int_X \phi \, d\mu$$

or briefly

$$\int \phi \, d\mu,$$

is the finite sum

$$\sum_k c_k \mu(E_k) \in [0, \infty],$$

where

$$\phi = \sum_k c_k I_{E_k}, \quad E_k = [\phi = c_k],$$

and c_k are the distinct values of ϕ.

Note that

$$\int I_E \, d\mu = \mu(E) \quad E \in \mathcal{A}$$

and

$$0 \leq \int \phi \, d\mu \leq \|\phi\| \, \mu([\phi \neq 0]). \tag{1}$$

For an *arbitrary* measurable function $f : X \to [0, \infty]$, consider the (non-empty) set S_f of measurable simple functions ϕ such that $0 \le \phi \le f$, and define

$$\int f \, d\mu := \sup_{\phi \in S_f} \int \phi \, d\mu. \tag{2}$$

For any $E \in \mathcal{A}$, the integral *over* E of f is defined by

$$\int_E f \, d\mu := \int f I_E \, d\mu. \tag{3}$$

Let ϕ, ψ be measurable simple functions; let c_k, d_j be the distinct values of ϕ and ψ, taken on the (mutually disjoint) sets E_k and F_j, respectively. Denote $Q := \{(k, j) \in \mathbb{N}^2; E_k \cap F_j \neq \emptyset\}$.

If $\phi \le \psi$, then $c_k \le d_j$ for $(k, j) \in Q$. Hence

$$\begin{aligned} \int \phi \, d\mu &= \sum_k c_k \mu(E_k) = \sum_{(k,j) \in Q} c_k \mu(E_k \cap F_j) \\ &\le \sum_{(k,j) \in Q} d_j \mu(E_k \cap F_j) = \sum_j d_j \mu(F_j) = \int \psi \, d\mu. \end{aligned}$$

Thus, the integral is *monotonic on simple functions.*

If f is simple, then $\int \phi \, d\mu \le \int f \, d\mu$ for all $\phi \in S_f$ (by monotonicity of the integral on simple functions), and therefore the supremum in (2) is less than or equal to the integral of f as a simple function; since $f \in S_f$, the reverse inequality is trivial, so that the two definitions of the integral of f coincide for f simple.

Since $S_{cf} = cS_f := \{c\phi; \phi \in S_f\}$ for $0 \le c < \infty$, we have (for f as above)

$$\int cf \, d\mu = c \int f \, d\mu \quad (0 \le c < \infty). \tag{4}$$

If $f \le g$ (f, g as above), $S_f \subset S_g$, and therefore $\int f \, d\mu \le \int g \, d\mu$ (*monotonicity* of the integral with respect to the "integrand").

In particular, if $E \subset F$ (both measurable), then $f I_E \le f I_F$, and therefore $\int_E f \, d\mu \le \int_F f \, d\mu$ (monotonicity of the integral with respect to the set of integration).

If $\mu(E) = 0$, then any $\phi \in S_{f I_E}$ assumes its non-zero values c_k on the sets $E_k \cap E$, that have measure 0 (as measurable subsets of E), and therefore $\int \phi \, d\mu = 0$ for all such ϕ, hence $\int_E f \, d\mu = 0$.

If $f = 0$ on E (for some $E \in \mathcal{A}$), then $f I_E$ is the zero function, hence has zero integral (by definition of the integral of simple functions!); this means that $\int_E f \, d\mu = 0$ when $f = 0$ on E.

Consider now the set function

$$\nu(E) := \int_E \phi \, d\mu \quad E \in \mathcal{A}, \tag{5}$$

for a *fixed* simple measurable function $\phi \ge 0$. As a special case of the preceding remark, $\nu(\emptyset) = 0$. Write $\phi = \sum c_k I_{E_k}$, and let $A_j \in \mathcal{A}$ be mutually disjoint ($j = 1, 2, \ldots$) with union A. Then

$$\phi I_A = \sum c_k I_{E_k \cap A},$$

so that, by the σ-additivity of μ and the possibility of interchanging summation order when the summands are non-negative,

$$\begin{aligned} \nu(A) : &= \sum_k c_k \mu(E_k \cap A) = \sum_k c_k \sum_j \mu(E_k \cap A_j) \\ &= \sum_j \sum_k c_k \mu(E_k \cap A_j) = \sum_j \nu(A_j). \end{aligned}$$

Thus ν is a positive measure. This is actually true for *any* measurable $\phi \ge 0$ (not necessarily simple), but this will be proved later.

If ψ, χ are *simple* functions as above (the distinct values of ψ and χ being $a_1, \ldots, a_p$ and $b_1, \ldots, b_q$, assumed on the measurable sets $F_1, \ldots, F_p$ and $G_1, \ldots, G_q$, respectively), then the simple measurable function $\phi := \psi + \chi$ assumes the constant value $a_i + b_j$ on the set $F_i \cap G_j$, and therefore, defining the measure ν as shown, we have

$$\nu(F_i \cap G_j) = (a_i + b_j)\mu(F_i \cap G_j). \tag{6}$$

But a_i and b_j are the constant values of ψ and χ on the set $F_i \cap G_j$ (respectively), so that the right-hand side of (6) equals $\nu'(F_i \cap G_j) + \nu''(F_i \cap G_j)$, where ν' and ν'' are the measures defined as ν, with the integrands ψ and χ instead of ϕ. Summing over all i, j, since X is the disjoint union of the sets $F_i \cap G_j$, the additivity of the measures ν, ν', and ν'' implies that $\nu(X) = \nu'(X) + \nu''(X)$, that is,

$$\int (\psi + \chi)\, d\mu = \int \psi\, d\mu + \int \chi\, d\mu. \tag{7}$$

Property (7) is the additivity of the integral over non-negative measurable simple functions. This property too is extended later to *arbitrary* non-negative measurable functions.

Theorem 1.13. *Let $(X, \mathcal{A}, \mu)$ be a positive measure space. Let*

$$f_1 \le f_2 \le f_3 \le \cdots : X \to [0, \infty]$$

be measurable, and denote $f = \lim f_n$ (defined pointwise). Then

$$\int f\, d\mu = \lim \int f_n\, d\mu. \tag{8}$$

This is the Monotone Convergence theorem *of Lebesgue.*

Proof. By Lemma 1.5, f is measurable (with range in $[0,\infty]$). The monotonicity of the integral (and the fact that $f_n \le f_{n+1} \le f$) implies that

$$\int f_n \, d\mu \le \int f_{n+1} \, d\mu \le \int f \, d\mu,$$

and therefore the limit in (8) exists ($:= c \in [0,\infty]$) and the *inequality* $\ge$ holds in (8). It remains to show the inequality $\le$ in (8). Let $0 < t < 1$. Given $\phi \in S_f$, denote

$$A_n = [t\phi \le f_n] = [f_n - t\phi \ge 0] \quad (n = 1, 2, \ldots).$$

Then $A_n \in \mathcal{A}$ and $A_1 \subset A_2 \subset \cdots$ (because $f_1 \le f_2 \le \cdots$). If $x \in X$ is such that $\phi(x) = 0$, then $x \in A_n$ (for all n). If $x \in X$ is such that $\phi(x) > 0$, then $f(x) \ge \phi(x) > t\phi(x)$, and there exists therefore n, for which $f_n(x) \ge t\phi(x)$, that is, $x \in A_n$ (for that n). This shows that $\bigcup_n A_n = X$. Consider the measure ν defined by (5) (for the simple function $t\phi$). By Lemma 1.10,

$$t \int \phi \, d\mu = \nu(X) = \lim_n \nu(A_n) = \lim_n \int_{A_n} t\phi \, d\mu.$$

However, $t\phi \le f_n$ on A_n, so the integrals on the right are $\le \int_{A_n} f_n \, d\mu \le \int_X f_n \, d\mu$ (by the monotonicity property of integrals with respect to the set of integration). Therefore $t \int \phi \, d\mu \le c$, and so $\int \phi \, d\mu \le c$ by the arbitrariness of $t \in (0,1)$. Taking the supremum over all $\phi \in S_f$, we conclude that $\int f \, d\mu \le c$ as wanted. $\square$

For *arbitrary* sequences of non-negative measurable functions we have the following *inequality*:

Theorem 1.14 (Fatou's lemma). *Let $f_n : X \to [0,\infty]$, $n = 1, 2, \ldots,$ be measurable. Then*

$$\int \liminf_n f_n \, d\mu \le \liminf_n \int f_n \, d\mu.$$

Proof. We have

$$\liminf_n f_n := \lim_n \big(\inf_{k \ge n} f_k \big).$$

Denote the infimum on the right by g_n. Then $g_n, n = 1, 2, \ldots,$ are measurable, $g_n \le f_n$,

$$0 \le g_1 \le g_2 \le \cdots,$$

and $\lim_n g_n = \liminf_n f_n$. By Theorem 1.13,

$$\int \liminf_n f_n \, d\mu = \int \lim g_n \, d\mu = \lim \int g_n \, d\mu.$$

But the integrals on the right are $\le \int f_n \, d\mu$, therefore their limit is $\le \liminf \int f_n \, d\mu$. $\square$

Another consequence of Theorem 1.13 is the additivity of the integral of non-negative measurable functions.

Theorem 1.15. *Let $f, g : X \to [0, \infty]$ be measurable. Then*

$$\int (f+g)\, d\mu = \int f\, d\mu + \int g\, d\mu.$$

Proof. By the Approximation theorem (Theorem 1.8), there exist simple measurable functions ϕ_n, ψ_n such that

$$\begin{aligned} 0 \le \phi_1 \le \phi_2 \le \dots, &\quad \lim \phi_n = f, \\ 0 \le \psi_1 \le \psi_2 \le \dots, &\quad \lim \psi_n = g. \end{aligned}$$

Then the measurable simple functions $\chi_n = \phi_n + \psi_n$ satisfy

$$0 \le \chi_1 \le \chi_2 \le \dots, \quad \lim \chi_n = f + g.$$

By Theorem 1.13 and the additivity of the integral of (non-negative measurable) *simple* functions (cf. (7)), we have

$$\begin{aligned} \int (f+g)\, d\mu &= \lim \int \chi_n\, d\mu = \lim \int (\phi_n + \psi_n)\, d\mu \\ &= \lim \int \phi_n\, d\mu + \lim \int \psi_n\, d\mu = \int f\, d\mu + \int g\, d\mu. \end{aligned}$$

□

The additivity property of the integral is also true for *infinite* sums of non-negative measurable functions:

Theorem 1.16 (Beppo Levi). *Let $f_n : X \to [0, \infty], n = 1, 2, \dots,$ be measurable. Then*

$$\int \sum_{n=1}^{\infty} f_n\, d\mu = \sum_{n=1}^{\infty} \int f_n\, d\mu.$$

Proof. Let

$$g_k = \sum_{n=1}^{k} f_n; \quad g = \sum_{n=1}^{\infty} f_n.$$

The measurable functions g_k satisfy

$$0 \le g_1 \le g_2 \le \dots, \quad \lim g_k = g,$$

and by Theorem 1.15 (and induction)

$$\int g_k\, d\mu = \sum_{n=1}^{k} \int f_n\, d\mu.$$

Therefore, by Theorem 1.13

$$\int g\,d\mu = \lim_k \int g_k\,d\mu = \lim_k \sum_{n=1}^{k} \int f_n\,d\mu = \sum_{n=1}^{\infty} \int f_n\,d\mu.$$

□

We may extend now the measure property of ν, defined earlier with a *simple* integrand, to the general case of a non-negative measurable integrand.

Theorem 1.17. *Let $f : X \to [0,\infty]$ be measurable, and set*

$$\nu(E) := \int_E f\,d\mu, \quad E \in \mathcal{A}.$$

Then ν is a (positive) measure on $\mathcal{A}$, and for any measurable $g : X \to [0,\infty]$,

$$\int g\,d\nu = \int gf\,d\mu. \tag{*}$$

Proof. Let $E_j \in \mathcal{A}, j = 1,2,\ldots$ be mutually disjoint, with union E. Then

$$fI_E = \sum_{j=1}^{\infty} fI_{E_j},$$

and therefore, by Theorem 1.16,

$$\nu(E) := \int fI_E\,d\mu = \sum_j \int fI_{E_j}\,d\mu = \sum_j \nu(E_j).$$

Thus, ν is a measure.

If $g = I_E$ for some $E \in \mathcal{A}$, then

$$\int g\,d\nu = \nu(E) = \int I_E f\,d\mu = \int gf\,d\mu.$$

By (4) and Theorem 1.15 (for the measures μ and ν), (*) is valid for g *simple.* Finally, for general g, the Approximation theorem (Theorem 1.8) provides a sequence of simple measurable functions

$$0 \le \phi_1 \le \phi_2 \le \cdots; \quad \lim \phi_n = g.$$

Then the measurable functions $\phi_n f$ satisfy

$$0 \le \phi_1 f \le \phi_2 f \le \cdots; \quad \lim \phi_n f = gf,$$

and Theorem 1.13 implies that

$$\int g\,d\nu = \lim_n \int \phi_n\,d\nu = \lim_n \int \phi_n f\,d\mu = \int gf\,d\mu. \qquad \square$$

Relation (*) is conveniently abbreviated as

$$d\nu = f\,d\mu.$$

Observe that if f_1 and f_2 coincide *almost everywhere* (briefly, "a.e." or μ-a.e., if the measure needs to be specified), that is, if they coincide except on a *null set* $A \in \mathcal{A}$ (more precisely, a μ-*null set*, that is, a measurable set A such that $\mu(A) = 0$), then the corresponding measures ν_i are equal, and in particular $\int f_1\,d\mu = \int f_2\,d\mu$. Indeed, for all $E \in \mathcal{A}$, $\mu(E \cap A) = 0$, and therefore

$$\nu_i(E \cap A) = \int_{E\cap A} f_i\,d\mu = 0, \quad i = 1, 2$$

by one of the observations following Definition 1.12. Hence

$$\begin{aligned}\nu_1(E) &= \nu_1(E \cap A) + \nu_1(E \cap A^c) = \nu_1(E \cap A^c)\\ &= \nu_2(E \cap A^c) = \nu_2(E).\end{aligned}$$

1.4 Integrable functions

Let $(X, \mathcal{A}, \mu)$ be a positive measure space, and let f be a measurable function with range in $[-\infty, \infty]$ or $\overline{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$ (the Riemann sphere). Then $|f| : X \to [0, \infty]$ is measurable, and has therefore an integral ($\in [0, \infty]$). In case this integral is *finite*, we shall say that f is *integrable*. In that case, the measurable set $[|f| = \infty]$ has measure zero. Indeed, it is contained in $[|f| > n]$ for all $n = 1, 2, \ldots$, and

$$n\mu([|f| > n]) = \int_{[|f|>n]} n\,d\mu \le \int_{[|f|>n]} |f|\,d\mu \le \int |f|\,d\mu.$$

Hence for all n

$$0 \le \mu([|f| = \infty]) \le \frac{1}{n} \int |f|\,d\mu,$$

and since the integral on the right is finite, we must have $\mu([|f| = \infty]) = 0$.

In other words, an integrable function is *finite a.e.*

We observed previously that non-negative measurable functions that coincide a.e. have equal integrals. This property is desirable in the general case now considered. If f is measurable, and if we redefine it as the finite arbitrary constant c on a set $A \in \mathcal{A}$ of measure zero, then the new function g is also measurable. Indeed, for any open set V in the range space,

$$[g \in V] = \{[g \in V] \cap A^c\} \cup \{[g \in V] \cap A\}.$$

The second set on the right is empty if $c \in V^c$, and is A if $c \in V$, thus belongs to $\mathcal{A}$ in any case. The first set on the right is equal to $[f \in V] \cap A^c \in \mathcal{A}$, by the measurability of f. Thus $[g \in V] \in \mathcal{A}$.

If f is integrable, we can redefine it as an arbitrary *finite* constant on the set $[|f| = \infty]$ (that has measure zero) and obtain a new *finite-valued* measurable function, whose integral should be the same as the integral of f (by the "desirable" property mentioned before). This discussion shows that we may restrict ourselves to *complex* (or, as a special case, to *real*) valued measurable functions.

Definition 1.18. Let $(X, \mathcal{A}, \mu)$ be a positive measure space. The function $f : X \to \mathbb{C}$ is *integrable* if it is measurable and

$$\|f\|_1 := \int |f|\, d\mu < \infty.$$

The set of all (complex) *integrable* functions will be denoted by

$$L^1(X, \mathcal{A}, \mu),$$

or briefly by $L^1(\mu)$ or $L^1(X)$ or L^1, when the unmentioned "objects" of the measure space are understood.

Defining the operations pointwise, L^1 is a complex vector space, since the inequality

$$|\alpha f + \beta g| \le |\alpha|\,|f| + |\beta|\,|g|$$

implies, by monotonicity, additivity, and homogeneity of the integral of non-negative measurable functions:

$$\|\alpha f + \beta g\|_1 \le |\alpha|\,\|f\|_1 + |\beta|\,\|g\|_1 < \infty,$$

for all $f, g \in L^1$ and $\alpha, \beta \in \mathbb{C}$.

In particular, $\|\cdot\|_1$ satisfies the triangle inequality (take $\alpha = \beta = 1$), and is trivially homogeneous.

Suppose $\|f\|_1 = 0$. For any $n = 1, 2, \dots,$

$$\begin{aligned} 0 \le \mu([|f| > 1/n]) &= \int_{[|f|>1/n]} d\mu = n \int_{[|f|>1/n]} (1/n)\, d\mu \\ &\le n \int_{|f|>1/n]} |f|\, d\mu \le n\,\|f\|_1 = 0, \end{aligned}$$

so $\mu([|f| > 1/n]) = 0$. Now the set where f is *not* zero is

$$[|f| > 0] = \bigcup_{n=1}^{\infty} [|f| > 1/n],$$

and by the σ-subadditivity property of positive measures, it follows that this set has measure zero. Thus, the vanishing of $\|f\|_1$ implies that $f = 0$ a.e.

(the converse is trivially true). One verifies easily that the relation "$f = g$ a.e." is an equivalence relation for complex measurable functions (transitivity follows from the fact that the union of two sets of measure zero has measure zero, by subadditivity of positive measures). All the functions f in the same equivalence class have the same value of $\|f\|_1$ (cf. discussion following Theorem 1.17).

We use the same notation L^1 for the space of all *equivalence classes* of integrable functions, with operations performed as usual on representatives of the classes, and with the $\|\cdot\|_1$-norm of a class equal to the norm of any of its representatives; L^1 is a normed space (for the norm $\|\cdot\|_1$). It is customary, however, to think of the elements of L^1 as functions (rather than equivalence classes of functions!).

If $f \in L^1$, then $f = u + \mathrm{i}v$ with $u := \Re f$ and $v := \Im f$ real measurable functions (cf. discussion following Lemma 1.7), and since $|u|, |v| \leq |f|$, we have $\|u\|_1, \|v\|_1 \leq \|f\|_1 < \infty$, that is, u, v are *real* elements of L^1 (conversely, if u, v are real elements of L^1, then $f = u + \mathrm{i}v \in L^1$, since L^1 is a complex vector space).

Writing $u = u^+ - u^-$ (and similarly for v), we obtain four non-negative (finite) measurable functions (cf. remarks following Lemma 1.5), and since $u^+ \leq |u| \leq |f|$ (and similarly for u^-, etc.), they have *finite integrals.* It makes sense therefore to *define*

$$\int u\,d\mu := \int u^+\,d\mu - \int u^-\,d\mu$$

(on the right, one has the difference of two *finite* non-negative real numbers!).

Doing the same with v, we then let

$$\int f\,d\mu := \int u\,d\mu + \mathrm{i}\int v\,d\mu.$$

Note that according to this definition,

$$\Re\int f\,d\mu = \int \Re f\,d\mu,$$

and similarly for the imaginary part.

Theorem 1.19. *The map $f \to \int f\,d\mu \in \mathbb{C}$ is a continuous linear functional on the normed space $L^1(\mu)$.*

Proof. Consider first real-valued functions $f, g \in L^1$. Let $h = f + g$. Then

$$h^+ - h^- = (f^+ - f^-) + (g^+ - g^-),$$

and since all functions here have *finite* values,

$$h^+ + f^- + g^- = h^- + f^+ + g^+.$$

By Theorem 1.15,

$$\int h^+\,d\mu + \int f^-\,d\mu + \int g^-\,d\mu = \int h^-\,d\mu + \int f^+\,d\mu + \int g^+\,d\mu.$$

All integrals above are finite, so we may subtract $\int h^- + \int f^- + \int g^-$ from both sides of the equation. This yields:

$$\int h\,d\mu = \int f\,d\mu + \int g\,d\mu.$$

The additivity of the integral extends trivially to *complex* functions in L^1.

If $f \in L^1$ is real and $c \in [0,\infty)$, $(cf)^+ = cf^+$ and similarly for f^-. Therefore, by (4) (following Definition 1.12),

$$\int cf\,d\mu = \int cf^+\,d\mu - \int cf^-\,d\mu = c\int f^+\,d\mu - c\int f^-\,d\mu = c\int f\,d\mu.$$

If $c \in (-\infty,0)$, $(cf)^+ = -cf^-$ and $(cf)^- = -cf^+$, and a similar calculation shows again that $\int(cf) = c\int f$. For $f \in L^1$ complex and c real, write $f = u + \mathrm{i}v$. Then

$$\int cf = \int (cu + \mathrm{i}cv) := \int (cu) + \mathrm{i}\int (cv) = c\Big(\int u + \mathrm{i}\int v\Big) := c\int f.$$

Note next that

$$\int (\mathrm{i}f) = \int (-v + \mathrm{i}u) = -\int v + \mathrm{i}\int u = \mathrm{i}\int f.$$

Finally, if $c = a + \mathrm{i}b$ (a, b real), then by additivity of the integral and the previous remarks,

$$\int (cf) = \int (af + \mathrm{i}bf) = \int (af) + \int \mathrm{i}bf = a\int f + \mathrm{i}b\int f = c\int f.$$

Thus

$$\int (\alpha f + \beta g)\,d\mu = \alpha\int f\,d\mu + \beta\int g\,d\mu$$

for all $f, g \in L^1$ and $\alpha, \beta \in \mathbb{C}$.

For $f \in L^1$, let $\lambda := \int f\,d\mu (\in \mathbb{C})$. Then, since the left-hand side of the following equation is *real*,

$$\begin{aligned}|\lambda| = \mathrm{e}^{\mathrm{i}\theta}\lambda = \mathrm{e}^{\mathrm{i}\theta}\int f\,d\mu = \int (\mathrm{e}^{\mathrm{i}\theta} f)\,d\mu = \Re\int (\mathrm{e}^{\mathrm{i}\theta} f)\,d\mu = \int \Re(\mathrm{e}^{\mathrm{i}\theta} f)\,d\mu \\ \le \int |\mathrm{e}^{\mathrm{i}\theta} f|\,d\mu = \int |f|\,d\mu.\end{aligned}$$

We thus obtained the important inequality

$$\Big|\int f\,d\mu\Big| \le \int |f|\,d\mu. \tag{1}$$

If $f, g \in L^1$, it follows from the linearity of the integral and (1) that

$$\Big|\int f\,d\mu - \int g\,d\mu\Big| = \Big|\int (f - g)\,d\mu\Big| \le \|f - g\|_1. \tag{2}$$

In particular, if f and g represent the same equivalence class, then $\|f-g\|_1 = 0$, and therefore $\int f\,d\mu = \int g\,d\mu$. This means that the functional $f \to \int f\,d\mu$ is *well-defined* as a functional on the *normed space* $L^1(\mu)$ (of equivalence classes!), and its continuity follows trivially from (2). □

In term of sequences, continuity of the integral on the normed space L^1 means that if $\{f_n\} \subset L^1$ converges to f in the L^1-metric, then

$$\int f_n\,d\mu \to \int f\,d\mu. \tag{3}$$

A useful sufficient condition for convergence in the L^1-metric, and therefore, for the validity of (3), is contained in the *Dominated Convergence theorem* of Lebesgue:

Theorem 1.20. *Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $\{f_n\}$ be a sequence of complex measurable functions on X that converge pointwise to the function f. Suppose there exists $g \in L^1(\mu)$ (with values in $[0, \infty)$) such that*

$$|f_n| \le g \quad (n = 1, 2, \ldots). \tag{4}$$

Then $f, f_n \in L^1(\mu)$ for all n, and $f_n \to f$ in the $L^1(\mu)$-metric.

In particular, (3) is valid.

Proof. By Lemma 1.5, f is measurable. By (4) and monotonicity

$$\|f\|_1, \|f_n\|_1 \le \|g\|_1 < \infty,$$

so that $f, f_n \in L^1$.

Since $|f_n - f| \le 2g$, the measurable functions $2g - |f_n - f|$ are non-negative. By Fatou's Lemma (Theorem 1.14),

$$\int \liminf_n (2g - |f_n - f|)\,d\mu \le \liminf_n \int (2g - |f_n - f|)\,d\mu. \tag{5}$$

The left-hand side of (5) is $\int 2g\,d\mu$. The integral on the right-hand side is $\int 2g\,d\mu + (-\|f_n - f\|_1)$, and its lim inf is

$$= \int 2g\,d\mu + \liminf_n (-\|f_n - f\|_1) = \int 2g\,d\mu - \limsup_n \|f_n - f\|_1.$$

Subtracting the finite number $\int 2g\,d\mu$ from both sides of the inequality, we obtain

$$\limsup_n \|f_n - f\|_1 \le 0.$$

However, if a non-negative sequence $\{a_n\}$ satisfies $\limsup a_n \le 0$, then it converges to 0 (because $0 \le \liminf a_n \le \limsup a_n \le 0$ implies $\liminf a_n = \limsup a_n = 0$). Thus $\|f_n - f\|_1 \to 0$. □

Rather than assuming pointwise convergence of the sequence $\{f_n\}$ at every point of X, we may assume that the sequence converges *almost everywhere*, that is, $f_n \to f$ on a set $E \in \mathcal{A}$ and $\mu(E^c) = 0$. The functions f_n could be defined only a.e., and we could include the countable union of all the sets where these functions are not defined (which is a set of measure zero, by the σ-subadditivity of measures) in the "exceptional" set E^c. The limit function f is defined a.e., in any case. For such a function, measurability means that $[f \in V] \cap E \in \mathcal{A}$ for each open set V.

If f_n (defined a.e.) converge pointwise a.e. to f, then with E as mentioned, the restrictions $f_n|_E$ are $\mathcal{A}_E$-measurable, where $\mathcal{A}_E$ is the σ-algebra $\mathcal{A} \cap E$, because

$$[f_n|_E \in V] = [f_n \in V] \cap E \in \mathcal{A}_E.$$

By Lemma 1.5, $f|_E := \lim f_n|_E$ is $\mathcal{A}_E$-measurable, and therefore the a.e.-defined function f is "measurable" in the above sense. We may define f as an arbitrary constant $c \in \mathbb{C}$ on E^c; the function thus extended to X is $\mathcal{A}$-measurable, as seen by the argument preceding Definition 1.18.

Now $f_n I_E$ are $\mathcal{A}$-measurable, converge pointwise everywhere to $f I_E$, and if $|f_n| \le g \in L^1$ for all n (wherever the functions are defined), then $|f_n I_E| \le g \in L^1$ (everywhere!). By Theorem 1.20,

$$\|f_n - f\|_1 = \|f_n I_E - f I_E\|_1 \to 0.$$

We then have the following *a.e. version* of the Lebesgue Dominated Convergence theorem:

Theorem 1.21. *Let $\{f_n\}$ be a sequence of a.e.-defined measurable complex functions on X, converging a.e. to the function f. Let $g \in L^1$ be such that $|f_n| \le g$ for all n (at all points where f_n is defined). Then f and f_n are in L^1, and $f_n \to f$ in the L^1-metric (in particular, $\int f_n\, d\mu \to \int f\, d\mu$).*

A useful "almost everywhere" proposition is the following:

Proposition 1.22. *If $f \in L^1(\mu)$ satisfies $\int_E f\, d\mu = 0$ for every $E \in \mathcal{A}$, then $f = 0$ a.e.*

Proof. Let $E = [u := \Re f \ge 0]$. Then $E \in \mathcal{A}$, so

$$\|u^+\|_1 = \int_E u\, d\mu = \Re \int_E f\, d\mu = 0,$$

and therefore $u^+ = 0$ a.e. Similarly $u^- = v^+ = v^- = 0$ a.e. (where $v := \Im f$), so that $f = 0$ a.e. $\square$

We should remark that, in general, a measurable a.e.-defined function f can be extended as a measurable function on X only by defining it as *constant* on the exceptional null set E^c. Indeed, the null set E^c could have a *non-measurable* subset A. Suppose $f : E \to \mathbb{C}$ is not onto, and let $a \in f(E)^c$. If we assign on A the (constant complex) value a, and any value $b \in f(E)$ on $E^c - A$, then the extended function is not measurable, because $[f = a] = A \notin \mathcal{A}$.

In order to be able to extend f in an arbitrary fashion and always get a measurable function, it is sufficient that subsets of null sets should be measurable (recall that a "null set" is measurable by definition!). A measure space with this property is called a *complete* measure space. Indeed, let f' be an arbitrary extension to X of an a.e.-defined measurable function f, defined on $E \in \mathcal{A}$, with E^c null. Then for any open $V \subset \mathbb{C}$,

$$[f' \in V] = ([f' \in V] \cap E) \cup ([f' \in V] \cap E^c).$$

The first set in the union is in $\mathcal{A}$, by measurability of the a.e.-defined function f; the second set is in $\mathcal{A}$ as a subset of the null set E^c (by completeness of the measure space). Hence $[f' \in V] \in \mathcal{A}$, and f' is measurable.

We say that the measure space $(X, \mathcal{M}, \nu)$ is an *extension* of the measure space $(X, \mathcal{A}, \mu)$ (both on X!) if $\mathcal{A} \subset \mathcal{M}$ and $\nu = \mu$ on $\mathcal{A}$. It is important to know that any measure space $(X, \mathcal{A}, \mu)$ has a (unique) "minimal" complete extension $(X, \mathcal{M}, \nu)$, where minimality means that if $(X, \mathcal{N}, \sigma)$ is any complete extension of $(X, \mathcal{A}, \mu)$, then it is an extension of $(X, \mathcal{M}, \nu)$. Uniqueness is of course trivial. The existence is proved below by a "canonical" construction.

Theorem 1.23. *Any measure space $(X, \mathcal{A}, \mu)$ has a unique minimal complete extension $(X, \mathcal{M}, \nu)$ (called the completion of the given measure space).*

Proof. We let $\mathcal{M}$ be the collection of all subsets E of X for which there exist $A, B \in \mathcal{A}$ such that

$$A \subset E \subset B, \quad \mu(B - A) = 0. \tag{6}$$

If $E \in \mathcal{A}$, we may take $A = B = E$ in (6), so $\mathcal{A} \subset \mathcal{M}$. In particular $X \in \mathcal{M}$.

If $E \in \mathcal{M}$ and A, B are as in (6), then $A^c, B^c \in \mathcal{A}$,

$$B^c \subset E^c \subset A^c$$

and $\mu(A^c - B^c) = \mu(B - A) = 0$, so that $E^c \in \mathcal{M}$.

If $E_j \in \mathcal{M}, j = 1, 2, \ldots$ and A_j, B_j are as in (6) (for E_j), then if E, A, B are the respective unions of E_j, A_j, B_j, we have $A, B \in \mathcal{A}$, $A \subset E \subset B$, and

$$B - A = \bigcup_j (B_j - A) \subset \bigcup_j (B_j - A_j).$$

The union on the right is a null set (as a countable union of null sets, by σ-subadditivity of measures), and therefore $B - A$ is a null set (by monotonicity of measures). This shows that $E \in \mathcal{M}$, and we conclude that $\mathcal{M}$ is a σ-algebra.

For $E \in \mathcal{M}$ and A, B as in (6), we let $\nu(E) = \mu(A)$. The function ν is *well defined* on $\mathcal{M}$, that is, the definition does not depend on the choice of A, B as in (6). Indeed, if A', B' satisfy (6) with E, then

$$A - A' \subset E - A' \subset B' - A',$$

so that $A - A'$ is a null set. Hence by additivity of μ, $\mu(A) = \mu(A \cap A') + \mu(A - A') = \mu(A \cap A')$. Interchanging the roles of A and A', we also have $\mu(A') = \mu(A \cap A')$, and therefore $\mu(A) = \mu(A')$, as wanted.

If $E \in \mathcal{A}$, we could choose $A = B = E$, and so $\nu(E) = \mu(E)$. In particular, $\nu(\emptyset) = 0$. If $\{E_j\}$ is a sequence of mutually disjoint sets in $\mathcal{M}$ with union E, and A_j, B_j are as in (6) (for E_j), we observed above that we could choose A for E (for (6)) as the union of the sets A_j. Since $A_j \subset E_j, j = 1, 2, \ldots$ and E_j are mutually disjoint, so are the sets A_j. Hence

$$\nu(E) := \mu(A) = \sum_j \mu(A_j) := \sum_j \nu(E_j),$$

and we conclude that $(X, \mathcal{M}, \nu)$ is a measure space extending $(X, \mathcal{A}, \mu)$. It is complete, because if $E \in \mathcal{M}$ is ν-null and A, B are as in (6), then for any $F \subset E$, we have

$$\emptyset \subset F \subset B,$$

and since $\mu(B - A) = 0$,

$$\mu(B - \emptyset) = \mu(B) = \mu(A) := \nu(E) = 0,$$

so that $F \in \mathcal{M}$.

Finally, suppose $(X, \mathcal{N}, \sigma)$ is any complete extension of $(X, \mathcal{A}, \mu)$, let $E \in \mathcal{M}$, and let A, B be as in (6). Write $E = A \cup (E - A)$. The set $B - A \in \mathcal{A} \subset \mathcal{N}$ is σ-null ($\sigma(B - A) = \mu(B - A) = 0$). By completeness of $(X, \mathcal{N}, \sigma)$, the subset $E - A$ of $B - A$ belongs to $\mathcal{N}$ (and is of course σ-null). Since $A \in \mathcal{A} \subset \mathcal{N}$, we conclude that $E \in \mathcal{N}$ and $\mathcal{M} \subset \mathcal{N}$. Also since $\sigma = \mu$ on $\mathcal{A}$, $\sigma(E) = \sigma(A) + \sigma(E - A) = \mu(A) := \nu(E)$, so that $\sigma = \nu$ on $\mathcal{M}$. $\square$

1.5 L^p-spaces

Let $(X, \mathcal{A}, \mu)$ be a (positive) measure space, and let $p \in [1, \infty)$. If $f : X \to [0, \infty]$ is measurable, so is f^p by Lemma 1.6, and therefore $\int f^p \, d\mu \in [0, \infty]$ is well defined. We denote

$$\|f\|_p := \left(\int f^p \, d\mu \right)^{1/p}.$$

Theorem 1.24 (Holder's inequality). *Let $p, q \in (1, \infty)$ be conjugate exponents, that is,*

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{1}$$

Then for all measurable functions $f, g : X \to [0, \infty]$,

$$\int fg \, d\mu \leq \|f\|_p \, \|g\|_q. \tag{2}$$

Proof. If $\|f\|_p = 0$, then $\|f^p\|_1 = 0$, and therefore $f = 0$ a.e.; hence $fg = 0$ a.e., and the left-hand side of (2) vanishes (as well as the right-hand side). By symmetry, the same holds true if $\|g\|_q = 0$. So we may consider only the case where $\|f\|_p$ and $\|g\|_q$ are both *positive.* Now if one of these quantities is infinite, the right-hand side of (2) is infinite, and (2) is trivially true. So we may assume that both quantities belong to $(0,\infty)$ (*positive and finite*). Denote

$$u = f/\|f\|_p, \quad v = g/\|g\|_q. \tag{3}$$

Then

$$\|u\|_p = \|v\|_q = 1. \tag{4}$$

It suffices to prove that

$$\int uv\,d\mu \le 1, \tag{5}$$

because (2) would follow by substituting (3) in (5).

The logarithmic function is concave ($(\log t)'' = -(1/t^2) < 0$). Therefore, by (1)

$$\frac{1}{p}\log s + \frac{1}{q}\log t \le \log\left(\frac{s}{p} + \frac{t}{q}\right)$$

for all $s,t \in (0,\infty)$. Equivalently,

$$s^{1/p}t^{1/q} \le \frac{s}{p} + \frac{t}{q}, \quad s,t \in (0,\infty). \tag{6}$$

When $x \in X$ is such that $u(x), v(x) \in (0,\infty)$, we substitute $s = u(x)^p$ and $t = v(x)^q$ in (6) and obtain

$$u(x)v(x) \le \frac{u(x)^p}{p} + \frac{v(x)^q}{q}, \tag{7}$$

and this inequality is trivially true when $u(x), v(x) \in \{0,\infty\}$. Thus (7) is valid on X, and integrating the inequality over X, we obtain by (4) and (1)

$$\int uv\,d\mu \le \frac{\|u\|_p^p}{p} + \frac{\|v\|_q^q}{q} = \frac{1}{p} + \frac{1}{q} = 1.$$

$\square$

Theorem 1.25 (Minkowski's inequality). *For any measurable functions $f, g : X \to [0,\infty]$,*

$$\|f+g\|_p \le \|f\|_p + \|g\|_p \quad (1 \le p < \infty). \tag{8}$$

Proof. Since (8) is trivial for $p = 1$ (by the additivity of the integral of non-negative measurable functions, we get even an equality), we consider $p \in (1,\infty)$. The case $\|f+g\|_p = 0$ is trivial. By convexity of the function t^p (for $p > 1$), $((s+t)/2)^p \le (s^p + t^p)/2$ for $s,t \in (0,\infty)$. Therefore, if $x \in X$ is such that $f(x), g(x) \in (0,\infty)$,

$$(f(x)+g(x))^p \le 2^{p-1}[f(x)^p + g(x)^p], \tag{9}$$

and (9) is trivially true if $f(x), g(x) \in \{0, \infty\}$, and holds therefore on X. Integrating, we obtain

$$\|f+g\|_p^p \le 2^{p-1}[\|f\|_p^p + \|g\|_p^p]. \tag{10}$$

If $\|f+g\|_p = \infty$, it follows from (10) that at least one of the quantities $\|f\|_p, \|g\|_p$ is infinite, and (8) is then valid (as the trivial equality $\infty = \infty$). This discussion shows that we may restrict our attention to the case

$$0 < \|f+g\|_p < \infty. \tag{11}$$

We write

$$(f+g)^p = f(f+g)^{p-1} + g(f+g)^{p-1}. \tag{12}$$

By Holder's inequality,

$$\int f(f+g)^{p-1}\, d\mu \le \|f\|_p \|(f+g)^{p-1}\|_q = \|f\|_p \|f+g\|_p^{p/q},$$

since $(p-1)q = p$ for conjugate exponents p, q. A similar estimate holds for the integral of the second summand on the right-hand side of (12). Adding these estimates, we obtain

$$\|f+g\|_p^p \le (\|f\|_p + \|g\|_p)\|f+g\|_p^{p/q}.$$

By (11), we may divide this inequality by $\|f+g\|_p^{p/q}$, and (8) follows since $p - p/q = 1$. □

In a manner analogous to that used for L^1, if $p \in [1, \infty)$, we consider the set

$$L^p(X, \mathcal{A}, \mu)$$

(or briefly, $L^p(\mu)$, or $L^p(X)$, or L^p, when the unmentioned parameters are understood) of all (*equivalence classes*) of measurable complex functions f on X, with

$$\|f\|_p := \|\,|f|\,\|_p < \infty.$$

Since $\|\cdot\|_p$ is trivially homogeneous, it follows from (8) that L^p is a normed space (over $\mathbb{C}$) for the pointwise operations and the norm $\|\cdot\|_p$. We can restate Holder's inequality in the form:

Theorem 1.26. *Let $p, q \in (1, \infty)$ be conjugate exponents. If $f \in L^p$ and $g \in L^q$, then $fg \in L^1$, and*

$$\|fg\|_1 \le \|f\|_p\, \|g\|_q.$$

A sufficient condition for convergence in the L^p-metric follows at once from Theorem 1.21:

Proposition. *Let $\{f_n\}$ be a sequence of a.e.-defined measurable complex functions on X, converging a.e. to the function f. For some $p \in [1, \infty)$, suppose* *there exists $g \in L^p$ such that $|f_n| \le g$ for all n (with the usual equivalence class ambiguity). Then $f, f_n \in L^p$, and $f_n \to f$ in the L^p-metric.*

Proof. The first statement follows from the inequalities $|f|^p, |f_n|^p \le g^p \in L^1$. Since $|f - f_n|^p \to 0$ a.e. and $|f - f_n|^p \le (2g)^p \in L^1$, the second statement follows from Theorem 1.21. □

The positive measure space $(X, \mathcal{A}, \mu)$ is said to be *finite* if $\mu(X) < \infty$. When this is the case, the Holder inequality implies that $L^p(\mu) \subset L^r(\mu)$ topologically (i.e., the inclusion map is continuous) when $1 \le r < p < \infty$. Indeed, if $f \in L^p(\mu)$, then by Holder's inequality with the conjugate exponents p/r and $s := p/(p-r)$,

$$\begin{aligned} \|f\|_r^r &= \int |f|^r.1\, d\mu \\ &\le \left[\int (|f|^r)^{p/r}\, d\mu\right]^{r/p} \left[\int 1^s\, d\mu\right]^{1/s} = \mu(X)^{1/s}\|f\|_p^r. \end{aligned}$$

Since $1/rs = (1/r) - (1/p)$, we obtain

$$\|f\|_r \le \mu(X)^{1/r-1/p}\|f\|_p. \tag{13}$$

Hence, $f \in L^r(\mu)$, and (13) (with $f - g$ replacing f) shows the continuity of the inclusion map of $L^p(\mu)$ into $L^r(\mu)$.

Taking in particular $r = 1$, we get that $L^p(\mu) \subset L^1(\mu)$ (topologically) for all $p \ge 1$, and

$$\|f\|_1 \le \mu(X)^{1/q}\|f\|_p, \tag{14}$$

where q is the conjugate exponent of p.

We formalize this discussion for future reference.

Proposition. *Let $(X, \mathcal{A}, \mu)$ be a finite positive measure space. Then $L^p(\mu) \subset L^r(\mu)$ (topologically) for $1 \le r < p < \infty$, and the norms inequality (13) holds.*

Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces, and let $h : X \to Y$ be a measurable map (cf. Definition 1.2). If μ is a measure on $\mathcal{A}$, the function $\nu : \mathcal{B} \to [0, \infty]$ given by

$$\nu(E) = \mu(h^{-1}(E)), \quad E \in \mathcal{B} \tag{15}$$

is well defined, and is clearly a measure on $\mathcal{B}$. Since $I_{h^{-1}(E)} = I_E \circ h$, we can write (15) in the form

$$\int_Y I_E\, d\nu = \int_X I_E \circ h\, d\mu.$$

By linearity of the integral, it follows that

$$\int_Y f\, d\nu = \int_X f \circ h\, d\mu \tag{16}$$

for every $\mathcal{B}$-measurable simple function f on Y. If $f : Y \to [0, \infty]$ is $\mathcal{B}$-measurable, use the Approximation Theorem 1.8 to obtain a non-decreasing sequence $\{f_n\}$ of $\mathcal{B}$-measurable non-negative simple functions converging pointwise to f; then $\{f_n \circ h\}$ is a similar sequence converging to $f \circ h$, and the Monotone Convergence Theorem shows that (16) is true for all such f.

If $f : Y \to \mathbb{C}$ is $\mathcal{B}$-measurable, then $f \circ h$ is a (complex) $\mathcal{A}$-measurable function on X, and for any $1 \leq p < \infty$,

$$\int_Y |f|^p \, d\nu = \int_X |f|^p \circ h \, d\mu = \int_X |f \circ h|^p \, d\mu.$$

Thus, $f \in L^p(\nu)$ for some $p \in [1, \infty)$ if and only if $f \circ h \in L^p(\mu)$, and

$$\|f\|_{L^p(\nu)} = \|f \circ h\|_{L^p(\mu)}.$$

In particular (case $p = 1$), f is ν-integrable on Y if and only if $f \circ h$ is μ-integrable on X. When this is the case, writing f as a linear combination of four non-negative ν-integrable functions, we see that (16) is valid for all such f.

Proposition. *Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces, and let $h : X \to Y$ be a measurable map. For any (positive) measure μ on $\mathcal{A}$, define $\nu(E) := \mu(h^{-1}(E))$ for $E \in \mathcal{B}$. Then:*

(1) ν is a (positive) measure on $\mathcal{B}$;

(2) if $f : Y \to [0, \infty]$ is $\mathcal{B}$-measurable, then $f \circ h$ is $\mathcal{A}$-measurable and (16) is valid;

(3) if $f : Y \to \mathbb{C}$ is $\mathcal{B}$-measurable, then $f \circ h$ is $\mathcal{A}$-measurable; $f \in L^p(\nu)$ for some $p \in [1, \infty)$ if and only if $f \circ h \in L^p(\mu)$, and in that case, the map $f \to f \circ h$ is norm-preserving*; in the special case $p = 1$, the map is* integral preserving *(i.e. (16) is valid).*

If ϕ is a simple complex measurable function with distinct *non-zero* values c_j assumed on E_j, then

$$\|\phi\|_p^p = \sum_j |c_j|^p \mu(E_j)$$

is finite if and only if $\mu(E_j) < \infty$ for all j, that is, equivalently, iff

$$\mu([|\phi| > 0]) < \infty.$$

Thus, the simple functions in L^p (for *any* $p \in [1, \infty)$) are the (measurable) simple functions vanishing outside a measurable set of *finite* measure (depending on the function). These functions are *dense* in L^p. Indeed, if $0 \leq f \in L^p$ (without loss of generality, we assume that f is everywhere defined!), the Approximation Theorem provides a sequence of simple measurable functions

$$0 \leq \phi_1 \leq \phi_2 \leq \cdots \leq f$$

such that $\phi_n \to f$ pointwise. By the proposition following Theorem 1.26, $\phi_n \to f$ in the L^p-metric.

For $f \in L^p$ complex, we may write $f = \sum_{k=0}^{3} \mathrm{i}^k g_k$ with $0 \leq g_k \in L^p$ ($g_0 := u^+$, etc., where $u = \Re f$). We then obtain four sequences $\{\phi_{n,k}\}$ of simple functions in L^p converging, respectively, to $g_k, k = 0, \ldots, 3$, in the L^p-metric; if $\phi_n := \sum_{k=0}^{3} \mathrm{i}^k \phi_{n,k}$, then ϕ_n are simple L^p-functions, and $\phi_n \to f$ in the L^p-metric. We proved

Theorem 1.27. *For any $p \in [1, \infty)$, the simple functions in L^p are dense in L^p.*

Actually, L^p is the *completion* of the normed space of all measurable simple functions vanishing outside a set of finite measure, with respect to the L^p-metric (induced by the L^p-norm). The meaning of this statement is made clear by the following definition.

Definition 1.28. Let Z be a metric space, with metric d. A *Cauchy sequence* in Z is a sequence $\{z_n\} \subset Z$ such that $d(z_n, z_m) \to 0$ when $n, m \to \infty$. The space Z is *complete* if every Cauchy sequence in Z converges in Z. If $Y \subset Z$ is dense in Z, and Z is complete, we also say that Z is the *completion* of Y (for the metric d). The completion of Y (for the metric d) is *unique* in a suitable sense.

A *complete normed space* is called a *Banach space*.

In order to get the conclusion preceding Definition 1.28, we still have to prove that L^p is complete:

Theorem 1.29. *L^p is a Banach space for each $p \in [1, \infty)$.*

We first prove the following.

Lemma 1.30. *Let $\{f_n\}$ be a Cauchy sequence in $L^p(\mu)$. Then it has a subsequence converging pointwise μ-a.e.*

Proof of Lemma. Since $\{f_n\}$ is Cauchy, there exists $m_k \in \mathbb{N}$ such that $\|f_n - f_m\|_p < 1/2^k$ for all $n > m > m_k$. Set

$$n_k = k + \max(m_1, \ldots, m_k).$$

Then $n_{k+1} > n_k > m_k$, and therefore $\{f_{n_k}\}$ is a subsequence of $\{f_n\}$ satisfying

$$\|f_{n_{k+1}} - f_{n_k}\|_p < 1/2^k \quad k = 1, 2, \ldots \tag{17}$$

Consider the series

$$g = \sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}|, \tag{18}$$

and its partial sums g_m. By Theorem 1.25 and (17),

$$\|g_m\|_p \le \sum_{k=1}^{m} \|f_{n_{k+1}} - f_{n_k}\|_p < \sum_{k=1}^{\infty} 1/2^k = 1$$

for all m. By Fatou's lemma,

$$\int g^p \, d\mu \le \liminf_m \int g_m^p \, d\mu = \liminf_m \|g_m\|_p^p \le 1.$$

Therefore, $g < \infty$ a.e., that is, the series (18) converges a.e., that is, the series

$$f_{n_1} + \sum_{k=1}^{\infty} (f_{n_{k+1}} - f_{n_k}) \tag{19}$$

converges absolutely pointwise a.e. to its sum f (extended as 0 on the null set where the series does not converge). Since the partial sums of (19) are precisely f_{n_m}, the lemma is proved. □

Proof of Theorem 1.29. Let $\{f_n\} \subset L^p$ be Cauchy. Thus for any $\epsilon > 0$, there exists $n_\epsilon \in \mathbb{N}$ such that

$$\|f_n - f_m\|_p < \epsilon \tag{20}$$

for all $n, m > n_\epsilon$. By the lemma, let then $\{f_{n_k}\}$ be a subsequence converging pointwise a.e. to the (measurable) complex function f. Applying Fatou's lemma to the non-negative measurable functions $|f_{n_k} - f_m|$, we obtain

$$\|f - f_m\|_p^p = \int \lim_k |f_{n_k} - f_m|^p \, d\mu \le \liminf_k \|f_{n_k} - f_m\|_p^p \le \epsilon^p \tag{21}$$

for all $m > n_\epsilon$. In particular, $f - f_m \in L^p$, and therefore $f = (f - f_m) + f_m \in L^p$, and (21) means that $f_m \to f$ in the L^p-metric. □

Definition 1.31. Let $(X, \mathcal{A}, \mu)$ be a positive measure space, and let $f : X \to \mathbb{C}$ be a measurable function. We say that $M \in [0, \infty]$ is an a.e. upper bound for $|f|$ if $|f| \le M$ a.e. The infimum of all the a.e. upper bounds for $|f|$ is called the *essential supremum* of $|f|$, and is denoted $\|f\|_\infty$. The set of all (*equivalence classes of*) measurable complex functions f on X with $\|f\|_\infty < \infty$ will be denoted by $L^\infty(\mu)$ (or $L^\infty(X)$, or $L^\infty(X, \mathcal{A}, \mu)$, or L^∞, depending on which "parameter" we wish to stress, if at all).

By definition of the essential supremum, we have

$$|f| \le \|f\|_\infty \quad \text{a.e.} \tag{22}$$

In particular, $\|f\|_\infty = 0$ implies that $f = 0$ a.e. (that is, f is the zero class).

If $f, g \in L^\infty$, then by (22), $|f + g| \le |f| + |g| \le \|f\|_\infty + \|g\|_\infty$ a.e., and so $\|f + g\|_\infty \le \|f\|_\infty + \|g\|_\infty$.

The homogeneity $\|\alpha f\|_\infty = |\alpha| \, \|f\|_\infty$ is trivial if either $\alpha = 0$ or $\|f\|_\infty = 0$. Assume then $|\alpha|, \|f\|_\infty > 0$. For any $t \in (0, 1)$, $t\|f\|_\infty < \|f\|_\infty$, hence it is not an a.e. upper bound for $|f|$, so that $\mu([|f| > t\|f\|_\infty]) > 0$, that is, $\mu([|\alpha f| > t|\alpha| \, \|f\|_\infty]) > 0$. Therefore, $\|\alpha f\|_\infty \ge t|\alpha| \, \|f\|_\infty$ for all $t \in (0, 1)$, hence $\|\alpha f\|_\infty \ge |\alpha| \, \|f\|_\infty$. The reversed inequality follows trivially from (22), and the homogeneity of $\| \cdot \|_\infty$ follows. We conclude that L^∞ is a normed space (over $\mathbb{C}$) for the pointwise operations and the L^∞-*norm* $\| \cdot \|_\infty$.

We verify its completeness as follows. Let $\{f_n\}$ be a Cauchy sequence in L^∞. In particular, it is a bounded set in L^∞. Let then $K = \sup_n \|f_n\|_\infty$. By (22), the sets $F_k := [|f_k| > K] \qquad (k \in \mathbb{N})$ and

$$E_{n,m} := [|f_n - f_m| > \|f_n - f_m\|_\infty] \quad (n, m \in \mathbb{N})$$

are μ-null, so their (countable) union E is null. For all $x \in E^c$,

$$|f_n(x) - f_m(x)| \le \|f_n - f_m\|_\infty \to 0$$

as $n, m \to \infty$ and $|f_n(x)| \le K$. By completeness of $\mathbb{C}$, the limit $f(x) := \lim_n f_n(x)$ exists for all $x \in E^c$ and $|f(x)| \le K$. Defining $f(x) = 0$ for all $x \in E$, we obtain a measurable function on X such that $|f| \le K$, that is, $f \in L^\infty$. Given $\epsilon > 0$, let $n_\epsilon \in \mathbb{N}$ be such that

$$\|f_n - f_m\|_\infty < \epsilon \quad (n, m > n_\epsilon).$$

Since $|f_n(x) - f_m(x)| < \epsilon$ for all $x \in E^c$ and $n, m > n_\epsilon$, letting $m \to \infty$, we obtain $|f_n(x) - f(x)| \le \epsilon$ for all $x \in E^c$ and $n > n_\epsilon$, and since $\mu(E) = 0$,

$$\|f_n - f\|_\infty \le \epsilon \quad (n > n_\epsilon),$$

that is, $f_n \to f$ in the L^∞-metric. We proved

Theorem 1.32. *L^∞ is a Banach space.*

Defining the conjugate exponent of $p = 1$ to be $q = \infty$ (so that $(1/p) + (1/q) = 1$ is formally valid in the usual sense), Holder's inequality remains true for this pair of conjugate exponents. Indeed, if $f \in L^1$ and $g \in L^\infty$, then $|fg| \le \|g\|_\infty |f|$ a.e., and therefore $fg \in L^1$ and

$$\|fg\|_1 \le \|g\|_\infty \|f\|_1.$$

Formally

Theorem 1.33. *Holder's inequality (Theorem 1.26) is valid for conjugate exponents $p, q \in [1, \infty]$.*

1.6 Inner product

For the conjugate pair $(p, q) = (2, 2)$, Theorem 1.26 asserts that if $f, g \in L^2$, then the product $f\bar{g}$ is integrable, so we may define

$$(f, g) := \int f\bar{g}\, d\mu \tag{1}$$

($\bar{g}$ denotes here the complex conjugate of g). The function (or *form*) $(\cdot\,,\cdot)$ has obviously the following properties on $L^2 \times L^2$:

(i) $(f, f) \ge 0$, and $(f, f) = 0$ if and only if $f = 0$ (the zero element);

(ii) $(\cdot\,, g)$ is linear for each given $g \in L^2$;

(iii) $(g, f) = \overline{(f, g)}$.

Property (i) is called *positive definiteness* of the form $(\cdot\,,\cdot)$; Properties (ii) and (iii) (together) are referred to as *sesquilinearity* or *Hermitianity* of the form. We may also consider the weaker condition

$$(\mathrm{i}')(f,f) \geq 0 \text{ for all } f,$$

called (positive) *semi-definiteness* of the form.

Definition 1.34. Let X be a complex vector space (with elements $x, y, \ldots$). A *(semi)-inner product on* X is a (semi)-definite sesquilinear form $(\cdot\,,\cdot)$ on X. The space X with a given (semi)-inner product is called a (*semi*)-*inner product space.*

If X is a semi-inner product space, the non-negative square root of (x,x) is denoted $\|x\|$.

Thus L^2 is an inner product space for the inner product (1) and $\|f\| := (f,f)^{1/2} = \|f\|_2$. By Theorem 1.26 with $p = q = 2$,

$$|(f,g)| \leq \|f\|_2 \, \|g\|_2 \tag{2}$$

for all $f, g \in L^2$. This special case of the Holder inequality is called the *Cauchy–Schwarz inequality.* We demonstrate below that it is valid in *any* semi-inner product space.

Observe that any sesquilinear form $(\cdot\,,\cdot)$ is *conjugate linear* with respect to its second variable, that is, for each given $x \in X$,

$$(x, \alpha u + \beta v) = \bar{\alpha}(x,u) + \bar{\beta}(x,v) \tag{3}$$

for all $\alpha, \beta \in \mathbb{C}$ and $u, v \in X$.

In particular

$$(x,0) = (0,y) = 0 \tag{4}$$

for all $x, y \in X$.

By (ii) and (3), for all $\lambda \in \mathbb{C}$ and $x, y \in X$,

$$(x + \lambda y, x + \lambda y) = (x,x) + \bar{\lambda}(x,y) + \lambda(y,x) + |\lambda|^2(y,y).$$

Since $\lambda(y,x)$ is the conjugate of $\bar{\lambda}(x,y)$ by (iii), we obtain the identity (for all $\lambda \in \mathbb{C}$ and $x, y \in X$)

$$\|x + \lambda y\|^2 = \|x\|^2 + 2\Re[\bar{\lambda}(x,y)] + |\lambda|^2\|y\|^2. \tag{5}$$

In particular, for $\lambda = 1$ and $\lambda = -1$, we have the identities

$$\|x + y\|^2 = \|x\|^2 + 2\Re(x,y) + \|y\|^2 \tag{6}$$

and

$$\|x - y\|^2 = \|x\|^2 - 2\Re(x,y) + \|y\|^2. \tag{7}$$

Adding, we obtain the so-called *parallelogram identity* for *any* s.i.p. (semi-inner product):

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2. \tag{8}$$

Subtracting (7) from (6), we obtain

$$4\Re(x,y) = \|x+y\|^2 - \|x-y\|^2. \tag{9}$$

If we replace y by iy in (9), we obtain

$$4\Im(x,y) = 4\Re[-\mathrm{i}(x,y)] = 4\Re(x,\mathrm{i}y) = \|x+\mathrm{i}y\|^2 - \|x-\mathrm{i}y\|^2. \tag{10}$$

By (9) and (10),

$$(x,y) = \frac{1}{4}\sum_{k=0}^{3} \mathrm{i}^k \|x+\mathrm{i}^k y\|^2, \tag{11}$$

where $\mathrm{i} = \sqrt{-1}$. This is the so-called *polarization identity* (which expresses the s.i.p. in terms of "induced norms").

By (5),

$$0 \le \|x\|^2 + 2\Re[\bar\lambda(x,y)] + |\lambda|^2\|y\|^2 \tag{12}$$

for all $\lambda \in \mathbb{C}$ and $x,y \in X$. If $\|y\| > 0$, take $\lambda = -(x,y)/\|y\|^2$; then $|(x,y)|^2/\|y\|^2 \le \|x\|^2$, and therefore

$$|(x,y)| \le \|x\|\,\|y\|. \tag{13}$$

If $\|y\| = 0$ but $\|x\| > 0$, interchange the roles of x and y and use (iii) to reach the same conclusion. If both $\|x\|$ and $\|y\|$ vanish, take $\lambda = -(x,y)$ in (12): we get $0 \le -2|(x,y)|^2$, hence $|(x,y)| = 0 = \|x\|\,\|y\|$, and we conclude that (13) is valid for *all* $x,y \in X$. This is the general Cauchy–Schwarz inequality for semi-inner products.

By (6) and (13),

$$\|x+y\|^2 \le \|x\|^2 + 2|(x,y)| + \|y\|^2 \le \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 = (\|x\|+\|y\|)^2,$$

hence

$$\|x+y\| \le \|x\| + \|y\|$$

for all $x,y \in X$. Taking $x = 0$ in (5), we get $\|\lambda y\| = |\lambda|\,\|y\|$ for all $\lambda \in \mathbb{C}$ and $y \in X$. We conclude that $\|\cdot\|$ is a semi-norm on X; it is a norm iff the s.i.p. is an *inner product*, that is, iff it is *definite*. Thus, an inner product space X is a normed space for the norm $\|x\| := (x,x)^{1/2}$ *induced* by its inner product (unless stated otherwise, this will be the standard norm for such spaces). In case X is *complete*, it is called a *Hilbert space*. Thus Hilbert spaces are special cases of Banach spaces.

The norm induced by the inner product (1) on L^2 is the usual L^2-norm $\|\cdot\|_2$, so that, by Theorem 1.29, L^2 is a Hilbert space.

1.7 Hilbert space: a first look

We consider some "geometric" properties of Hilbert spaces.

Theorem 1.35 (Distance theorem). *Let X be a Hilbert space, and let $K \subset X$ be non-empty, closed, and convex (i.e., $(x+y)/2 \in K$ whenever $x, y \in K$). Then for each $x \in X$, there exists a unique $k \in K$ such that*

$$d(x,k) = d(x,K). \tag{1}$$

The notation $d(x,y)$ is used for the metric induced by the norm, $d(x,y) := \|x-y\|$. As in any metric space, $d(x,K)$ denotes the *distance from x to K*, that is,

$$d(x,K) := \inf_{y \in K} d(x,y). \tag{2}$$

Proof. Let $d = d(x,K)$. Since $d^2 = \inf_{y \in K} \|x-y\|^2$, there exist $y_n \in K$ such that

$$(d^2 \leq) \|x - y_n\|^2 < d^2 + 1/n, \quad n = 1, 2, \ldots. \tag{3}$$

By the parallelogram identity,

$$\begin{aligned} \|y_n - y_m\|^2 &= \|(x-y_m) - (x-y_n)\|^2 \\ &= 2\|x-y_m\|^2 + 2\|x-y_n\|^2 - \|(x-y_m) + (x-y_n)\|^2. \end{aligned}$$

Rewrite the last term on the right-hand side in the form

$$4\|x - (y_m + y_n)/2\|^2 \geq 4d^2,$$

since $(y_m + y_n)/2 \in K$, by hypothesis. Hence by (3)

$$\|y_n - y_m\|^2 \leq 2/m + 2/n \to 0$$

as $m, n \to \infty$. Thus, the sequence $\{y_n\}$ is Cauchy. Since X is *complete*, the sequence converges in X, and its limit k in necessarily in K because $y_n \in K$ for all n and K is closed. By continuity of the norm on X, letting $n \to \infty$ in (3), we obtain $\|x-k\| = d$, as wanted.

To prove uniqueness, suppose $k, k' \in K$ satisfy

$$\|x-k\| = \|x-k'\| = d.$$

Again by the parallelogram identity,

$$\begin{aligned} \|k - k'\|^2 &= \|(x-k') - (x-k)\|^2 \\ &= 2\|x-k'\|^2 + 2\|x-k\|^2 - \|(x-k') + (x-k)\|^2. \end{aligned}$$

As before, write the last term as $4\|x-(k+k')/2\|^2 \geq 4d^2$ (since $(k+k')/2 \in K$ by hypothesis). Hence

$$\|k-k'\|^2 \leq 2d^2 + 2d^2 - 4d^2 = 0,$$

and therefore $k = k'$. □

We say that the vector $y \in X$ is *orthogonal* to the vector x if $(x,y) = 0$. In that case also $(y,x) = \overline{(x,y)} = 0$, so that the orthogonality relation is symmetric. For x given, let $x^\perp$ denote the set of all vectors orthogonal to x. This is the kernel of the linear functional $\phi = (\cdot\,, x)$, that is, the set $\phi^{-1}(\{0\})$. As such a kernel, it is a subspace. Since $|\phi(y) - \phi(z)| = |(y-z,x)| \leq \|y-z\|\,\|x\|$ by Schwarz's inequality, ϕ is continuous, and therefore $x^\perp = \phi^{-1}(\{0\})$ is closed. Thus, $x^\perp$ is a closed subspace. More generally, for any non-empty subset A of X, define

$$A^\perp := \bigcap_{x \in A} x^\perp = \{y \in Y; (y,x) = 0 \text{ for all } x \in A\}.$$

As the intersection of closed subspaces, $A^\perp$ is a closed subspace of X.

Theorem 1.36 (Orthogonal decomposition theorem). *Let Y be a closed subspace of the Hilbert space X. Then X is the direct sum of Y and $Y^\perp$, that is, each $x \in X$ has the unique orthogonal decomposition $x = y + z$ with $y \in Y$ and $z \in Y^\perp$.*

Note that the so-called *components* y and z of x (in Y and $Y^\perp$, respectively) are orthogonal.

Proof. As a closed subspace of X, Y is a non-empty, closed, convex subset of X. By the distance theorem, there exists a unique $y \in Y$ such that

$$\|x-y\| = d := d(x,Y).$$

Letting $z := x - y$, the existence part of the theorem will follow if we show that $(z,u) = 0$ for all $u \in Y$. Since Y is a subspace, and $Y \neq \{0\}$ without loss of generality, every $u \in Y$ is a scalar multiple of a unit vector in Y, so it suffices to prove that $(z,u) = 0$ for *unit vectors* $u \in Y$. For all $\lambda \in \mathbb{C}$, by the identity (5) (following Definition 1.34),

$$\|z - \lambda u\|^2 = \|z\|^2 - 2\Re[\bar{\lambda}(z,u)] + |\lambda|^2.$$

The left-hand side is

$$\|x - (y + \lambda u)\|^2 \geq d^2,$$

since $y + \lambda u \in Y$. Since $\|z\| = d$, we obtain

$$0 \leq -2\Re[\bar{\lambda}(z,u)] + |\lambda|^2.$$

Choose $\lambda = (z,u)$. Then $0 \leq -|(z,u)|^2$, so that $(z,u) = 0$ as claimed.

If $x = y + z = y' + z'$ are two decompositions with $y, y' \in Y$ and $z, z' \in Y^{\perp}$, then $y - y' = z' - z \in Y \cap Y^{\perp}$, so that in particular $y - y'$ is orthogonal to itself (i.e., $(y - y', y - y') = 0$), which implies that $y - y' = 0$, whence $y = y'$ and $z = z'$. □

We observed in passing that for each given $y \in X$, the function $\phi := (\cdot\,, y)$ is a continuous linear functional on the inner product space X. For *Hilbert* spaces, this is the *general* form of continuous linear functionals:

Theorem 1.37 ("Little" Riesz representation theorem). *Let $\phi : X \to \mathbb{C}$ be a continuous linear functional on the Hilbert space X. Then there exists a unique $y \in X$ such that $\phi = (\cdot\,, y)$.*

Proof. If $\phi = 0$ (the zero functional), take $y = 0$. Assume then that $\phi \neq 0$, so that its kernel Y is a closed subspace $\neq X$. Therefore $Y^{\perp} \neq \{0\}$, by Theorem 1.36. Let then $z \in Y^{\perp}$ be a unit vector. Since $Y \cap Y^{\perp} = \{0\}$, $z \notin Y$, so that $\phi(z) \neq 0$. For any given $x \in X$, we may then define

$$u := x - \frac{\phi(x)}{\phi(z)} z.$$

By linearity,

$$\phi(u) = \phi(x) - \frac{\phi(x)}{\phi(z)} \phi(z) = 0,$$

that is, $u \in Y$, and

$$x = u + \frac{\phi(x)}{\phi(z)} z \tag{4}$$

is the (unique) orthogonal decomposition of x (corresponding to the particular subspace Y, the kernel of ϕ). Define now $y = \overline{\phi(z)} z (\in Y^{\perp})$. By (4),

$$(x, y) = (u, y) + \frac{\phi(x)}{\phi(z)} \phi(z)(z, z) = \phi(x)$$

since $(u, y) = 0$ and $\|z\| = 1$. This proves the existence part of the theorem. Suppose now that $y, y' \in X$ are such that $\phi(x) = (x, y) = (x, y')$ for all $x \in X$. Then $(x, y - y') = 0$ for all x, hence in particular $(y - y', y - y') = 0$, which implies that $y = y'$. □

1.8 The Lebesgue–Radon–Nikodym theorem

We apply the Riesz representation theorem to prove the Lebesgue decomposition theorem and the Radon–Nikodym theorem for (positive) measures.

We start with a measure-theoretic lemma.

The positive measure space $(X, \mathcal{A}, \mu)$ is *σ-finite* if there exists a *sequence* of mutually disjoint measurable sets X_j with union X, such that $\mu(X_j) < \infty$ for all j.

Lemma 1.38 (The averages lemma). *Let $(X, \mathcal{A}, \sigma)$ be a σ-finite positive measure space. Let $g \in L^1(\sigma)$ be such that, for all $E \in \mathcal{A}$ with $0 < \sigma(E) < \infty$, the "averages"*

$$A_E(g) := \frac{1}{\sigma(E)} \int_E g \, d\sigma$$

are contained in some given closed *set $F \subset \mathbb{C}$. Then $g(x) \in F$ σ-a.e.*

Proof. We need to prove that $g^{-1}(F^c)$ is σ-null. Write the *open* set F^c as the countable union of the closed discs

$$\Delta_n := \{z \in \mathbb{C}; |z - a_n| \leq r_n\}, \quad n = 1, 2, \ldots.$$

Then

$$g^{-1}(F^c) = \bigcup_{n=1}^{\infty} g^{-1}(\Delta_n),$$

and it suffices to prove that $E_\Delta := g^{-1}(\Delta)$ is σ-null whenever Δ is a closed disc (with center a and radius r) contained in F^c.

Write X as the countable union of mutually disjoint measurable sets X_k with $\sigma(X_k) < \infty$. Set $E_{\Delta,k} := E_\Delta \cap X_k$, and suppose $\sigma(E_{\Delta,k}) > 0$ for some Δ as above and some k. Since $|g(x) - a| \leq r$ on $E := E_{\Delta,k}$, and $0 < \sigma(E) < \infty$, we have

$$|A_E(g) - a| = |A_E(g - a)| \leq \frac{1}{\sigma(E)} \int_E |g - a| \, d\sigma \leq r,$$

so that $A_E(g) \in \Delta \subset F^c$, contradicting the hypothesis. Hence $\sigma(E_{\Delta,k}) = 0$ for all k and therefore $\sigma(E_\Delta) = 0$ for all Δ as above. □

Lemma 1.39. *Let $0 \leq \lambda \leq \sigma$ be finite measures on the measurable space $(X, \mathcal{A})$. Then there exists a measurable function $g : X \to [0, 1]$ such that*

$$\int f \, d\lambda = \int f g \, d\sigma \tag{1}$$

for all $f \in L^2(\sigma)$.

Proof. By Definition 1.12, the relation $\lambda \leq \sigma$ between positive measures implies that $\int f \, d\lambda \leq \int f \, d\sigma$ for all non-negative measurable functions f. Hence $L^2(\sigma) \subset L^2(\lambda) (\subset L^1(\lambda)$, by the second proposition following Theorem 1.26.)

For all $f \in L^2(\sigma)$, we have then by Schwarz's inequality:

$$\left| \int f \, d\lambda \right| \leq \int |f| \, d\lambda \leq \int |f| \, d\sigma \leq \sigma(X)^{1/2} \|f\|_{L^2(\sigma)}.$$

Replacing f by $f - h$ (with $f, h \in L^2(\sigma)$), we get

$$\left| \int f \, d\lambda - \int h \, d\lambda \right| = \left| \int (f - h) d\lambda \right| \leq \sigma(X)^{1/2} \|f - h\|_{L^2(\sigma)},$$

so that the functional $f \to \int f\,d\lambda$ is a continuous linear functional on $L^2(\sigma)$. By the Riesz representation theorem for the Hilbert space $L^2(\sigma)$, there exists an element $g_1 \in L^2(\sigma)$ such that this functional is $(\cdot\,, g_1)$. Letting $g = \overline{g_1}$ $(\in L^2(\sigma))$, we get the wanted relation (1).

Since $I_E \in L^2(\sigma)$ (because σ is a finite measure), we have in particular

$$\lambda(E) = \int I_E\,d\lambda = \int_E g\,d\sigma$$

for all $E \in \mathcal{A}$. If $\sigma(E) > 0$,

$$\frac{1}{\sigma(E)} \int_E g\,d\sigma = \frac{\lambda(E)}{\sigma(E)} \in [0,1].$$

By the Averages Lemma 1.38, $g(x) \in [0,1]$ σ-a.e., and we may then choose a representative of the equivalence class g with range in $[0,1]$. $\square$

Terminology. Let $(X, \mathcal{A}, \lambda)$ be a positive measure space. We say that the set $A \in \mathcal{A}$ *carries* the measure λ (or that λ is supported by A) if $\lambda(E) = \lambda(E \cap A)$ for all $E \in \mathcal{A}$.

This is, of course, equivalent to $\lambda(E) = 0$ for all measurable subsets E of A^c.

Two (positive) measures λ_1, λ_2 on $(X, \mathcal{A})$ are *mutually singular* (notation $\lambda_1 \perp \lambda_2$) if they are carried by *disjoint* measurable sets A_1, A_2. Equivalently, each measure is carried by a null set relative to the other measure.

On the other hand, if $\lambda_2(E) = 0$ whenever $\lambda_1(E) = 0$ (for $E \in \mathcal{A}$), we say that λ_2 is *absolutely continuous* with respect to λ_1 (notation: $\lambda_2 \ll \lambda_1$).

Equivalently, $\lambda_2 \ll \lambda_1$ if and only if any (measurable) set that carries λ_1 also carries λ_2.

Theorem 1.40 (Lebesgue–Radon–Nikodym). *Let $(X, \mathcal{A}, \mu)$ be a σ-finite positive measure space, and let λ be a* finite *positive measure on $(X, \mathcal{A})$. Then*

(a) λ has the unique (so-called) Lebesgue decomposition

$$\lambda = \lambda_a + \lambda_s$$

with $\lambda_a \ll \mu$ and $\lambda_s \perp \mu$;

(b) there exists a unique $h \in L^1(\mu)$ such that

$$\lambda_a(E) = \int_E h\,d\mu$$

for all $E \in \mathcal{A}$.

(part (a) is the Lebesgue decomposition theorem; part (b) is the Radon–Nikodym theorem.)

Proof. Case $\mu(X) < \infty$.

Let $\sigma := \lambda + \mu$. Then the finite positive measures λ, σ satisfy $\lambda \le \sigma$, so that by Lemma 1.39, there exists a measurable function $g : X \to [0,1]$ such that (1) holds, that is, after rearrangement,

$$\int f(1-g)\, d\lambda = \int fg\, d\mu \tag{2}$$

for all $f \in L^2(\sigma)$. Define

$$A := g^{-1}([0,1)); \quad B := g^{-1}(\{1\}).$$

Then A, B are disjoint measurable sets with union X.

Taking $f = I_B$ ($\in L^2(\sigma)$, since σ is a finite measure) in (2), we obtain $\mu(B) = 0$ (since $g = 1$ on B). Therefore, the measure λ_s defined on $\mathcal{A}$ by

$$\lambda_s(E) := \lambda(E \cap B)$$

satisfies $\lambda_s \perp \mu$.

Define similarly $\lambda_a(E) := \lambda(E \cap A)$; this is a positive measure on $\mathcal{A}$, mutually singular with λ_s (since it is carried by $A = B^{\mathrm{c}}$), and by additivity of measures,

$$\lambda(E) = \lambda(E \cap A) + \lambda(E \cap B) = \lambda_a(E) + \lambda_s(E),$$

so that the Lebesgue decomposition will follow if we show that $\lambda_a \ll \mu$. This follows trivially from the integral representation (b), which we proceed to prove.

For each $n \in \mathbb{N}$ and $E \in \mathcal{A}$, take in (2)

$$f = f_n := (1 + g + \cdots + g^n) I_E.$$

(Since $0 \le g \le 1$, f is a bounded measurable function, hence $f \in L^2(\sigma)$.) We obtain

$$\int_E (1 - g^{n+1})\, d\lambda = \int_E (g + g^2 + \cdots + g^{n+1})\, d\mu. \tag{3}$$

Since $g = 1$ on B, the left-hand side equals $\int_{E \cap A} (1 - g^{n+1})\, d\lambda$. However, $0 \le g < 1$ on A, so that the integrands form a non-decreasing sequence of non-negative measurable functions converging pointwise to 1. By the monotone convergence theorem, the left-hand side of (3) converges therefore to $\lambda(E \cap A) = \lambda_a(E)$. The integrands on the right-hand side of (3) form a non-decreasing sequence of non-negative measurable functions converging pointwise to the (measurable) function

$$h := \sum_{n=1}^{\infty} g^n.$$

Again, by monotone convergence, the right-hand side of (3) converges to $\int_E h\, d\mu$, and the representation (b) follows. Taking in particular $E = X$, we get

$$\|h\|_{L^1(\mu)} = \int_X h\, d\mu = \lambda_a(X) = \lambda(A) < \infty,$$

so that $h \in L^1(\mu)$, and the existence part of the theorem is proved in case $\mu(X) < \infty$.

General case. Let $X_j \in \mathcal{A}$ be mutually disjoint, with union X, such that $0 < \mu(X_j) < \infty$. Define

$$w = \sum_j \frac{1}{2^j \mu(X_j)} I_{X_j}.$$

This is a strictly positive μ-integrable function, with $\|w\|_1 = 1$. Consider the positive measure

$$\nu(E) = \int_E w \, d\mu.$$

Then $\nu(X) = \|w\|_1 = 1$, and $\nu \ll \mu$. On the other hand, if $\nu(E) = 0$, then $\sum_j (1/2^j \mu(X_j))\mu(E \cap X_j) = 0$, hence $\mu(E \cap X_j) = 0$ for all j, and therefore $\mu(E) = 0$. This shows that $\mu \ll \nu$ as well (one says that the measures μ and ν are *mutually absolutely continuous*, or *equivalent*).

Since ν is a finite measure, the first part of the proof gives the decomposition $\lambda = \lambda_a + \lambda_s$ with $\lambda_a \ll \nu$ (hence $\lambda_a \ll \mu$ by the trivial transitivity of the relation $\ll$), and $\lambda_s \perp \nu$ (hence $\lambda_s \perp \mu$, because λ_s is supported by a ν-null set, which is also μ-null, since $\mu \ll \nu$). The first part of the proof gives also the representation (cf. Theorem 1.17)

$$\lambda_a(E) = \int_E h \, d\nu = \int_E hw \, d\mu = \int_E \tilde{h} \, d\mu,$$

where $\tilde{h} := hw$ is non-negative, measurable, and

$$\|\tilde{h}\|_1 = \int_X \tilde{h} \, d\mu = \lambda_a(X) \le \lambda(X) < \infty.$$

This completes the proof of the "existence part" of the theorem in the general case.

To prove the *uniqueness* of the Lebesgue decomposition, suppose

$$\lambda = \lambda_a + \lambda_s = \lambda'_a + \lambda'_s,$$

with

$$\lambda_a, \lambda'_a \ll \mu \quad \text{and} \quad \lambda_s, \lambda'_s \perp \mu.$$

Let B be a μ-null set that carries both λ_s and λ'_s. Then

$$\lambda_a(B) = \lambda'_a(B) = 0 \quad \text{and} \quad \lambda_s(B^{\mathrm{c}}) = \lambda'_s(B^{\mathrm{c}}) = 0,$$

so that for all $E \in \mathcal{A}$,

$$\begin{aligned} \lambda_a(E) &= \lambda_a(E \cap B^{\mathrm{c}}) = \lambda(E \cap B^{\mathrm{c}}) \\ &= \lambda'_a(E \cap B^{\mathrm{c}}) = \lambda'_a(E), \end{aligned}$$

hence also $\lambda_s(E) = \lambda'_s(E)$.

In order to prove the uniqueness of h in (b), suppose $h, h' \in L^1(\mu)$ satisfy

$$\lambda_a(E) = \int_E h\,d\mu = \int_E h'\,d\mu.$$

Then $h-h' \in L^1(\mu)$ satisfies $\int_E (h-h')\,d\mu = 0$ for all $E \in \mathcal{A}$, and it follows from Proposition 1.22 that $h-h' = 0$ μ-a.e., that is, $h = h'$ *as elements of* $L^1(\mu)$. $\square$

If the measure λ is absolutely continuous with respect to μ, it has the trivial Lebesgue decomposition $\lambda = \lambda + 0$, with the zero measure as singular part. By uniqueness, it follows that $\lambda_a = \lambda$, and therefore Part 2 of the theorem gives the representation $\lambda(E) = \int_E h\,d\mu$ for all $E \in \mathcal{A}$. Conversely, such an integral representation of λ implies trivially that $\lambda \ll \mu$ (if $\mu(E) = 0$, the function $hI_E = 0$ μ-a.e., and therefore $\lambda(E) = \int f I_E\,d\mu = 0$). Thus

Theorem 1.41 (Radon–Nikodym). *Let $(X, \mathcal{A}, \mu)$ be a σ-finite positive measure space. A finite positive measure λ on $\mathcal{A}$ is absolutely continuous with respect to μ if and only if there exists $h \in L^1(\mu)$ such that*

$$\lambda(E) = \int_E h\,d\mu \quad (E \in \mathcal{A}). \tag{*}$$

By Theorem 1.17, Relation (*) implies that

$$\int g\,d\lambda = \int gh\,d\mu \tag{**}$$

for all non-negative measurable functions g on X. Since we may take $g = I_E$ in (**), this last relation implies (*). As mentioned after Theorem 1.17, these equivalent relations are *symbolically* written in the form $d\lambda = h\,d\mu$. It follows easily from Theorem 1.17 that, in that case, if $g \in L^1(\lambda)$, then $gh \in L^1(\mu)$ and (**) is valid for such (complex) functions g. The function h is called the *Radon–Nikodym derivative* of λ with respect to μ, and is denoted $d\lambda/d\mu$.

1.9 Complex measures

Definition 1.42. Let $(X, \mathcal{A})$ be an arbitrary measurable space. A *complex measure* on $\mathcal{A}$ is a σ-additive function $\mu : \mathcal{A} \to \mathbb{C}$, that is,

$$\mu\Big(\bigcup_n E_n\Big) = \sum_n \mu(E_n) \tag{1}$$

for any sequence of mutually disjoint sets $E_n \in \mathcal{A}$.

Since the left-hand side of (1) is independent of the order of the sets E_n and is a complex number, the right-hand side converges in $\mathbb{C}$ unconditionally, hence *absolutely*. Taking $E_n = \emptyset$ for all n, the convergence of (1) shows that $\mu(\emptyset) = 0$. It follows that μ is (finitely) additive, and since its values are complex numbers, it is "subtractive" as well (i.e., $\mu(E - F) = \mu(E) - \mu(F)$ whenever $E, F \in \mathcal{A}$, $F \subset E$).

A *partition* of $E \in \mathcal{A}$ is a sequence of mutually disjoint sets $A_k \in \mathcal{A}$ with union equal to E. We set

$$|\mu|(E) := \sup \sum_k |\mu(A_k)|, \tag{2}$$

where the supremum is taken over all partitions of E.

Theorem 1.43. *Let μ be a complex measure on $\mathcal{A}$, and define $|\mu|$ by (2). Then $|\mu|$ is a finite positive measure on $\mathcal{A}$ that dominates μ (i.e., $|\mu(E)| \le |\mu|(E)$ for all $E \in \mathcal{A}$).*

Proof. Let $E = \bigcup E_n$ with $E_n \in \mathcal{A}$ mutually disjoint ($n \in \mathbb{N}$). For any partition $\{A_k\}$ of E, $\{A_k \cap E_n\}_k$ is a partition of E_n ($n = 1, 2, \ldots$), so that

$$\sum_k |\mu(A_k \cap E_n)| \le |\mu|(E_n), \quad n = 1, 2, \ldots.$$

We sum these inequalities over all n, interchange the order of summation in the double sum (of non-negative terms!), and use the triangle inequality to obtain

$$\sum_n |\mu|(E_n) \ge \sum_k |\sum_n \mu(A_k \cap E_n)| = \sum_k |\mu(A_k)|,$$

since $\{A_k \cap E_n\}_n$ is a partition of A_k, for each $k \in \mathbb{N}$. Taking now the supremum over all partitions $\{A_k\}$ of E, it follows that

$$\sum_n |\mu|(E_n) \ge |\mu|(E). \tag{3}$$

On the other hand, given $\epsilon > 0$, there exists a partition $\{A_{n,k}\}_k$ of E_n such that

$$\sum_k |\mu(A_{n,k})| > |\mu|(E_n) - \epsilon/2^n, \quad n = 1, 2, \ldots.$$

Since $\{A_{n,k}\}_{n,k}$ is a partition of E, we obtain

$$|\mu|(E) \ge \sum_{n,k} |\mu(A_{n,k})| > \sum_n |\mu|(E_n) - \epsilon.$$

Letting $\epsilon \to 0+$ and using (3), we conclude that $|\mu|$ is σ-additive. Since $|\mu|(\emptyset) = 0$ is trivial, $|\mu|$ is indeed a positive measure on $\mathcal{A}$. $\square$

In order to show that the measure $|\mu|$ is finite, we need the following.