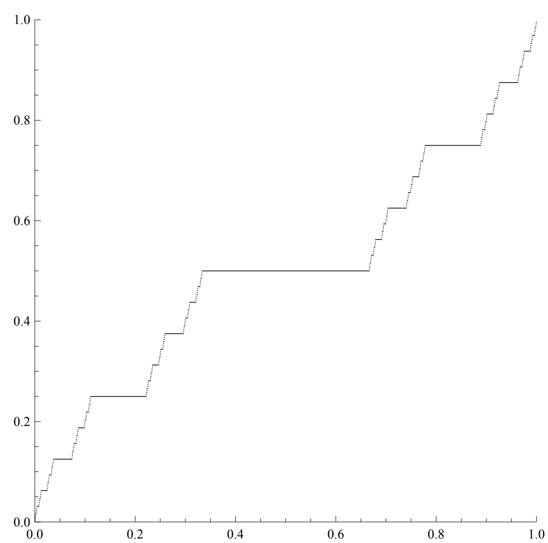
REAL ANALYSIS

Lecture notes for MA 645/646

Rudi Weikard



Version of August 3, 2022

Contents

Preface

For millennia humankind has concerned itself with the concepts of patterns, shape, and quantity. Out of these grew mathematics which, accordingly, comprises three branches: algebra, geometry, and analysis. We are concerned here with the latter. As quantity is commonly expressed using real numbers, analysis begins with a careful study of those. Next are the concepts of continuity, derivative, and integral. While at least the ideas, if not the formal definition, of the former two have been rather stable since the advent of the calculus the same is not true for the concept of integral. Initially integrals were thought of as anti-derivatives until, in the 19th century, Augustin-Louis Cauchy (1789 – 1857) and Bernhard Riemann (1826 – 1866) defined the integral of a function over an interval $[a, b]$ by partitioning the interval into shorter and shorter subintervals. However, the Riemann integral proved to have severe shortcomings leading many mathematicians at the end of the 19th century to search for alternatives. In 1901 Henri Lebesgue (1875 – 1941) presented a new idea using more general subsets than subintervals in the sums approximating the integral. In order to make this work he assumed having a concept of the size, i.e., a measure, of such sets compatible with the length of intervals. Clearly, one should at least require that the measure of the union of two disjoint sets equals the sum of their respective measures but it turned out to be much more fruitful to require such a property of countable collections of pairwise disjoint sets, a property which is called countable additivity. One wishes, of course, to assign a measure to any subset; alas this is not always possible and one may have to be satisfied with a domain for the measure smaller than the power set. For this the framework of a σ-algebra gained widespread acceptance. Lebesgue's approach to the integral had a tremendous impact on analysis.

It is an important consequence of Lebesgue's ideas that the mere presence of a countably additive measure defined on a σ-algebra in a set X allows to develop a theory of integration and thus a rather abstract approach to the subject. In this course we will take this abstract approach in the beginning. Later we will define Lebesgue measure and investigate more concrete problems for functions defined on $\mathbb{R}$ or $\mathbb{R}^n$.

These notes were informed by the following texts: Bennewitz [1], Folland [2], Gordon [3], Henze [4], Hewitt and Stromberg [5], Kolmogorov and Fomin [6], Riesz and Sz.-Nagy [7], and Rudin [8] and [9]. Thanks are also owed to several generations of students who had to work through previous versions of these notes finding errors and suggesting improvements. Steven Redolfi deserves a particular mention in this context.

CHAPTER 1

Abstract Integration

1.1. Integration of non-negative functions

1.1.1 Dealing with infinity. The set $[0,\infty] = [0,\infty) \cup \{\infty\}$ becomes a totally ordered set after declaring $x \leq \infty$ for any $x \in [0,\infty]$ (and maintaining the usual order in $[0,\infty)$). Recall that a totally ordered space is a topological space, since intervals of the types $\{x : x < b\}$, $\{x : a < x < b\}$, and $\{x : a < x\}$ (with $a, b \in [0,\infty]$) form a base of a topology. In fact, it is sufficient to choose a and b in $\mathbb{Q}$ and this implies that every open set is a union of at most countably many of such basic sets.

We also extend the usual arithmetic from $[0,\infty)$ to $[0,\infty]$ by declaring $x+\infty = \infty+x = \infty$ for all x and $x\cdot\infty = \infty\cdot x = \infty$ unless $x = 0$ in which case we set instead $0\cdot\infty = \infty\cdot 0 = 0$. Addition and multiplication are then associative and commutative and multiplication is distributive over addition. We also define $\infty^p = \infty$ when $0 < p < \infty$.

Suppose $n \mapsto a_n$ is a sequence in $[0,\infty]$. We define $\sum_{n=1}^\infty a_n = \infty$, if $a_k = \infty$ for some $k \in \mathbb{N}$ or if the series is numerical but not convergent. Otherwise, if the series is convergent, the symbol $\sum_{n=1}^\infty a_n$ represents its limit.

1.1.2 σ-algebras. A collection $\mathcal{M}$ of subsets of a set X is called a *σ-algebra* in X if $\mathcal{M}$ has the following three properties: (i) $X \in \mathcal{M}$; (ii) $A \in \mathcal{M}$ implies that A^c, the complement of A in X, is in $\mathcal{M}$, too; and (iii) $A_n \in \mathcal{M}$ for all $n \in \mathbb{N}$ implies that $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{M}$.

For example, the set $\{\emptyset, X\}$ and the power set of X (denoted by $\mathcal{P}(X)$) are σ-algebras in X.

Any σ-algebra contains the empty set as well as finite unions and finite and countable intersections (by de Morgan's laws) of its elements.

The elements of $\mathcal{M}$ are called *measurable sets*. The pair $(X,\mathcal{M})$ (or X itself, if no confusion can arise) is called a *measurable space*.

1.1.3 Measures. Let $(X,\mathcal{M})$ be a measurable space. A function $\mu : \mathcal{M} \to [0,\infty]$ or $\mu : \mathcal{M} \to \mathbb{C}$ is called a *measure* on $\mathcal{M}$, if $\mu(\emptyset) = 0$ and if

$$\mu(\bigcup_{n=1}^\infty A_n) = \sum_{n=1}^\infty \mu(A_n)$$

provided that the A_n are pairwise disjoint measurable sets. The latter property is called *countable additivity* or *σ-additivity*. We call μ a *positive measure* if its values are in $[0,\infty]$ and a *complex measure* if they are in $\mathbb{C}$. We may speak of a measure on X, if no confusion about the underlying σ-algebra can arise. A *measure space* $(X,\mathcal{M},\mu)$ is a measurable space with a measure defined on its σ-algebra.

If μ is a measure, it has the following continuity property: if $n \mapsto A_n$ is a non-decreasing sequence (i.e., $A_n \subset A_{n+1}$ for all n) of measurable sets, then $\mu(\bigcup_{n=1}^\infty A_n) = \lim_{n\to\infty} \mu(A_n)$.

Moreover, a positive measure μ is *monotone*, i.e., if $A \subset B$, then $\mu(A) \leq \mu(B)$.

1.1.4 First examples. The following are simple examples of measures which may be defined for any set X.

1. The *Dirac measure* (or *unit mass measure*) is defined on the power set $\mathcal{P}(X)$ of X. Fix $x_0 \in X$. Then define, for any subset E of X,

$$\mu_{x_0}(E) = \begin{cases} 1 & \text{if } x_0 \in E, \\ 0 & \text{if } x_0 \notin E. \end{cases}$$

2. The *counting measure* is also defined on the power set of X. If E has finitely many elements then $\mu(E)$ is their number and otherwise infinity. The case where $X = \mathbb{N}$ and $\mathcal{M} = \mathcal{P}(\mathbb{N})$ is particularly interesting.

1.1.5 Simple functions. Suppose $(X, \mathcal{M})$ is a measurable space. A *simple function* is a function on X which assumes only finitely many values in $[0, \infty)$, each one on a measurable set. A simple function may always be represented as $\sum_{k=1}^{n} \alpha_k \chi_{A_k}$ with pairwise distinct values α_k and pairwise disjoint non-empty measurable sets A_k whose union is X. This representation, being uniquely determined, is called the *canonical representation* of the function.

Sums and products of simple functions are again simple functions. Moreover, if s_1 and s_2 are simple functions, then so are $\max\{s_1, s_2\}$ and $\min\{s_1, s_2\}$.

1.1.6 Integrals of simple functions. Let $(X, \mathcal{M}, \mu)$ be a measure space with a positive measure μ. If s is a simple function with canonical representation $s = \sum_{k=1}^{n} \alpha_k \chi_{A_k}$ we define its *integral* with respect to the measure μ to be

$$\int_\mu s = \sum_{k=1}^{n} \alpha_k \mu(A_k).$$

If the measure under consideration is known from the context, we will generally drop the corresponding subscript on the integral sign.

For all $c \in [0, \infty)$ we have $\int cs = c \int s$ and the function $\phi : \mathcal{M} \to [0, \infty]$ defined by $\phi(E) = \int_\mu s\chi_E = \sum_{k=1}^{n} \alpha_k \mu(A_k \cap E)$ is again a positive measure. It follows that the integral is also *additive* (i.e., $\int (s_1 + s_2) = \int s_1 + \int s_2$). Moreover, if s_1 and s_2 satisfy $s_1 \leq s_2$, then $\int s_1 \leq \int s_2$. One may now show that $\int_\mu s = \sum_{k=1}^{n} \alpha_k \mu(A_k)$, if $s = \sum_{k=1}^{n} \alpha_k \chi_{A_k}$, even if this is not the canonical representation of s.

1.1.7 Non-negative measurable functions. Suppose $(X, \mathcal{M})$ is a measurable space. A function $f : X \to [0, \infty]$ is called *measurable*, if it is the pointwise limit of a non-decreasing sequence of simple functions. Note that $f + g$ and fg are measurable, if f and g are.

THEOREM. $f : X \to [0, \infty]$ is measurable, if and only if the preimage of any open set in $[0, \infty]$ is measurable.

SKETCH OF PROOF. Assume first that f is measurable and that $n \mapsto s_n$ is a non-decreasing sequence of simple functions converging to it. Then note that $\{x : \alpha < f(x)\} = \bigcup_{n=1}^{\infty} \{x : \alpha < s_n(x)\}$ and $\{x : f(x) < \beta\} = \bigcup_{k=1}^{\infty} \{x : \beta - 1/k < f(x)\}^c$. Conversely, assume that $\{x : q < f(x)\}$ is measurable for all non-negative $q \in \mathbb{Q}$ and let $k \mapsto q_k$ be an enumeration of these numbers. Then $s_n = \max\{q_k \chi_{\{t : q_k < f(t)\}} : k \leq n\}$ defines a non-decreasing sequence of simple functions converging to f. □

COROLLARY. Suppose $n \mapsto f_n$ is a sequence of non-negative measurable functions. Then $\sup\{f_n : n \in \mathbb{N}\}$, $\inf\{f_n : n \in \mathbb{N}\}$, $\limsup_{n\to\infty} f_n$, and $\liminf_{n\to\infty} f_n$ are measurable, too.

1.1.8 Integrals of non-negative functions. Assume that $(X, \mathcal{M}, \mu)$ is a measure space with a positive measure μ. The integral of a measurable function $f : X \to [0, \infty]$ with respect to μ is defined by

$$\int_\mu f = \sup \left\{ \int_\mu s : s \text{ simple and } 0 \le s \le f \right\}.$$

If $\int_\mu f < \infty$, then f is called *integrable.*

1.1.9 Basic properties of integrals. Suppose that $(X, \mathcal{M}, \mu)$ is a measure space with a positive measure μ and that $f, g : X \mapsto [0, \infty]$ are measurable. Then the following statements hold:

(1) If $f \le g$ then $\int f \le \int g$.
(2) If $f = 0$ then $\int f = 0$ even if $\mu(X) = \infty$.
(3) $\int f = 0$ if and only if $\mu(\{x : f(x) > 0\}) = 0$.
(4) If $c \in [0, \infty]$ then $\int cf = c \int f$.

1.1.10 The monotone convergence theorem. We now turn to the monotone convergence theorem, the cornerstone of integration theory.

THEOREM. Let $(X, \mathcal{M}, \mu)$ be a measure space with a positive measure μ and $n \mapsto f_n$ a sequence of non-negative measurable functions on X such that $0 \le f_1 \le f_2 \le ... \le \infty$ and $\lim f_n = f$ pointwise. Then f is measurable and

$$\lim_{n \to \infty} \int f_n = \int f.$$

SKETCH OF PROOF. The measurability of f was shown in Corollary 1.1.7.

The sequence $n \mapsto \int f_n$ is non-decreasing and converges to some $a \in [0, \int f]$. Thus it is sufficient to show that $\int s \le a$ for any simple function s such that $0 \le s \le f$. To do this introduce the measure $\phi(E) = \int s\chi_E$, fix a $c \in (0, 1)$, and define the sets $E_n = \{x : f_n(x) \ge cs(x)\}$. The E_n are measurable, $E_n \subset E_{n+1}$, and $\bigcup_{n=1}^\infty E_n = X$. This implies $\int s = \lim_{n\to\infty} \phi(E_n) \le a/c$ and, since c may be arbitrarily close to 1, $\int s \le a$. □

One important consequence of this theorem is that we may find integrals of non-negative measurable functions by taking limits of integrals of non-decreasing sequences of simple functions. In particular, we obtain the additivity of the integral, i.e., $\int(f + g) = \int f + \int g$ whenever f and g are non-negative measurable functions.

1.1.11 Fatou's lemma. If $(X, \mathcal{M}, \mu)$ is a measure space with a positive measure μ and $f_n : X \to [0, \infty]$ are measurable for all $n \in \mathbb{N}$, then

$$\int \liminf_{n\to\infty} f_n \le \liminf_{n\to\infty} \int f_n.$$

SKETCH OF PROOF. Consider the sequence $g_k = \inf\{f_k, f_{k+1}, ...\}$. □

1.1.12 Measures induced by positive functions. Let $(X, \mathcal{M}, \mu)$ be a measure space with a positive measure μ and $f : X \to [0, \infty]$ a measurable function. Define $\phi : \mathcal{M} \to [0, \infty]$ by $\phi(E) = \int_\mu f\chi_E$. Then ϕ is a positive measure. Moreover, if $g : X \to [0, \infty]$ is measurable then gf is measurable and

$$\int_\phi g = \int_\mu gf.$$

1.1.13 Double series of non-negative numbers. Let f be a non-negative function on $\mathbb{N}^2$ and ν the counting measure on $\mathbb{N}^2$. Then $E \mapsto \mu(E) = \int_\nu f\chi_E$ is a positive measure on $\mathbb{N}^2$. Moreover, if ρ is a bijection from $\mathbb{N}$ to $\mathbb{N}^2$, then

$$\mu(\mathbb{N}^2) = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} f(j,k) = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} f(j,k) = \sum_{k=1}^{\infty} f(\rho(k)).$$

1.1.14 The notion of *almost everywhere*. Suppose $(X, \mathcal{M}, \mu)$ is a measure space with a positive measure μ. Any statement depending on a parameter $x \in X$ is said to be true for *almost all* $x \in X$ or *almost everywhere*, if it is true for all $x \in X \setminus N$ where N is a set of measure zero. For instance, we say that a function f vanishes almost everywhere, if $f(x) = 0$ for all x outside a set of measure 0. Of course, this notion depends on the measure under consideration, and the measure has to be mentioned explicitly in case of doubt.

1.2. Integration of complex-valued functions

Throughout this section $(X, \mathcal{M}, \mu)$ denotes a measure space with a positive measure μ.

1.2.1 Measurable functions. Suppose Y is a topological space and S is a measurable subset of X. Then we say that $f : S \to Y$ is *measurable* if the preimage of every open set in Y is measurable. Note that, in view of Theorem 1.1.7, this definition is compatible with the earlier one when $Y = [0, \infty]$ and $S = X$. We are particularly interested in the case where $Y = \mathbb{C}$ and $\mu(S^c) = 0$ (e.g., $S = X$).

The following statements hold.

(1) Let Y be $[0, \infty]$, $\mathbb{R}$ or $\mathbb{C}$. Then χ_E is measurable if and only if E is.
(2) Continuous functions of measurable functions are measurable.
(3) $f : S \to \mathbb{R}^2$ is measurable if and only if its components are.
(4) $h : S \to \mathbb{C}$ is measurable if and only if its real and imaginary parts are.
(5) If $f, g : S \to \mathbb{C}$ are measurable so are $f + g$ and fg.
(6) The complex-valued measurable functions form a vector space.
(7) If $h : S \to \mathbb{C}$ is measurable then so is $|h|$ and there is a measurable function α such that $|\alpha| = 1$ and $h = \alpha|h|$.

1.2.2 Splitting a complex-valued function into four parts. When f is a real-valued function on X we define $f^\pm = \max\{\pm f, 0\}$, the *positive* and *negative parts* of f. Note that $f_+, f_- \geq 0$, $f_+ f_- = 0$, $f = f_+ - f_-$, and $|f| = f_+ + f_-$. Thus, when f is a complex-valued function on X we may split it into four parts, viz., the positive and negative parts of both the real and imaginary parts of f. Thus we may write

$$f = (\operatorname{Re} f)_+ - (\operatorname{Re} f)_- + i(\operatorname{Im} f)_+ - i(\operatorname{Im} f)_-.$$

A complex-valued function on X is measurable if and only if the four parts $(\operatorname{Re} f)_\pm$ and $(\operatorname{Im} f)_\pm$ are all measurable.

1.2.3 Integrable functions and definition of the integral. A measurable function $f : X \to \mathbb{C}$ is called *integrable* if $\int |f| < \infty$. It is integrable if and only if each of its four parts $(\operatorname{Re} f)_\pm$ and $(\operatorname{Im} f)_\pm$ are.

If $f : X \to \mathbb{C}$ is integrable we define

$$\int f = \int (\operatorname{Re} f)_+ - \int (\operatorname{Re} f)_- + i \int (\operatorname{Im} f)_+ - i \int (\operatorname{Im} f)_-$$

and note that $\int f$ is a complex number.

1.2.4 Linearity of the integral. The set of integrable functions is a vector space and the integral is a linear functional on it, i.e.,

$$\int(\alpha f+\beta g)=\alpha\int f+\beta\int g$$

whenever f, g are integrable and $\alpha, \beta \in \mathbb{C}$.

Sketch of proof. Any constant multiple of an integrable function is integrable in view of part (4) of 1.1.9. The additivity of the integral for non-negative functions was shown in 1.1.10. Thus, taking part (1) of 1.1.9 into account, it follows that the sum of two integrable functions is integrable and hence that the set of integrable functions is a vector space.

Next one proves linearity of the integral for real functions and real constants using the fact that then $(f+g)_+ + f_- + g_- = (f+g)_- + f_+ + g_+$. This shows that $\int f = \int \operatorname{Re} f + i \int \operatorname{Im} f$ from which the general claim follows easily. □

1.2.5 Integrals of almost everywhere defined functions. If f is defined only on $S \subset X$ and if $\mu(S^c)=0$, let $\tilde{f}$ be any measurable extension of f to all of X. The value of $\int \tilde{f}$ is then independent of the choice of that extension and we will simply write $\int f$ for $\int \tilde{f}$.

1.2.6 Generalized triangle inequality. If f is integrable then

$$\left|\int f\right| \le \int |f|. \tag{1}$$

Remark: If $X=\{1,2\}$ and μ is the counting measure then (1) becomes $|f(1)+f(2)| \le |f(1)|+|f(2)|$.

1.2.7 Sequences of measurable functions. Suppose $n \mapsto f_n : X \to \mathbb{C}$ is a sequence of measurable functions. If $n \mapsto f_n(x)$ converges for every $x \in S \in \mathcal{M}$, define $f : S \to \mathbb{C}$ by $f(x)=\lim_{n\to\infty} f_n(x)$. Since the convergence of f_n implies the convergence of each of $(\operatorname{Re} f_n)_\pm$ and $(\operatorname{Im} f_n)_\pm$ it follows that f is measurable.

1.2.8 Lebesgue's dominated convergence theorem. We turn now to the most powerful tool of integration theory.

Theorem. Let $(X, \mathcal{M}, \mu)$ be a measure space with a positive measure μ and $n \mapsto f_n : X \to \mathbb{C}$ a sequence of measurable functions. Let S be the set of those x where $f(x)=\lim_{n\to\infty} f_n(x)$ exists and assume that $\mu(S^c)=0$. If there exists an integrable function $g : X \to [0,\infty]$ such that $|f_n| \le g$ for all n, then f is integrable,

$$\lim_{n\to\infty}\int |f_n - f| = 0, \tag{2}$$

and

$$\lim_{n\to\infty}\int f_n = \int f. \tag{3}$$

Sketch of proof. Since $f_n \to f$ one gets that $|f_n| \to |f|$ and hence that $|f| \le g$. Equation (2) follows after applying Fatou's lemma to $2g - |f_n - f| \ge 0$. Then linearity and $|\int (f_n - f)| \le \int |f_n - f|$ imply equation (3). □

1.2.9 Averages. Suppose μ is a positive measure on X and that $X=\bigcup_{n=1}^\infty X_n$ with $\mu(X_n)<\infty$. Also assume that $f : X \to \mathbb{C}$ is integrable and that S is a closed subset of $\mathbb{C}$.

If the *averages*

$$A_f(E) = \frac{1}{\mu(E)} \int_\mu f\chi_E$$

are in S whenever $0 < \mu(E) < \infty$, then $f(x) \in S$ for almost all $x \in X$.

Sketch of proof. S^c is a countable union of closed balls. If B is one of these balls and $E = f^{-1}(B) \cap X_n$ has positive measure, then $A_f(E) \in B$, a contradiction. □

1.3. Convex functions and Jensen's inequality

1.3.1 Convex functions. A function $\varphi : (a, b) \to \mathbb{R}$ is called *convex* on (a, b) if

$$\varphi((1-t)x + ty) \leq (1-t)\varphi(x) + t\varphi(y)$$

for all $t \in [0, 1]$ and all $x, y \in (a, b)$. (We allow for $a = -\infty$ and $b = \infty$.)

If φ is convex and $a < u < v < w < b$, then

$$\frac{\varphi(v) - \varphi(u)}{v - u} \leq \frac{\varphi(w) - \varphi(u)}{w - u} \leq \frac{\varphi(w) - \varphi(v)}{w - v}.$$

Conversely, if

$$\frac{\varphi(v) - \varphi(u)}{v - u} \leq \frac{\varphi(w) - \varphi(v)}{w - v}$$

whenever $a < u < v < w < b$, then φ is convex.

A convex function is continuous. A differentiable function is convex if and only if its derivative is non-decreasing.

1.3.2 Jensen's inequality. Let μ be a positive measure on a set X such that $\mu(X) = 1$. Suppose $f : X \to (a, b)$ is integrable and φ is convex on (a, b). Then

$$\varphi\left(\int f\right) \leq \int \varphi \circ f.$$

Sketch of proof. Let $\beta(v) = \sup\{(\varphi(v) - \varphi(u))/(v - u) : u \in (a, v)\}$. Then

$$\varphi(y) \geq \varphi(v) + \beta(v)(y - v)$$

for all $v, y \in (a, b)$, in particular for $y = f(x)$ and $v = \int f$. Integration establishes the inequality. □

1.4. L^p-spaces

Throughout this section $(X, \mathcal{M}, \mu)$ denotes a measure space with a positive measure μ.

1.4.1 Conjugate exponents. If $p, q > 1$ and $p + q = pq$ or, equivalently, $1/p + 1/q = 1$, then p and q are called *conjugate exponents*. We also regard 1 and ∞ as conjugate exponents.

1.4.2 Hölder's inequality. Suppose $p, q > 1$ are conjugate exponents. Let f, g be measurable, nonnegative functions on X. Then

$$\int fg \leq \left(\int f^p\right)^{1/p} \left(\int g^q\right)^{1/q}.$$

Assuming the right-hand side is finite, equality holds here if and only if there are $\alpha, \beta \in [0, \infty)$, not both zero, such that $\alpha f^p = \beta g^q$ almost everywhere.

Sketch of proof. Let A and B be the factors on the right side of the inequality. The inequality is trivial if either of A and B is zero or infinity. Hence consider $0 < A, B < \infty$. Defining $F = f/A$ and $G = g/B$ we get $\int F^p = \int G^q = 1$.

Whenever $F(x)$ and $G(x)$ are positive there are numbers $s(x), t(x)$ such that $F(x) = \mathrm{e}^{s(x)/p}$ and $G(x) = \mathrm{e}^{t(x)/q}$ which implies that $F(x)G(x) \le F(x)^p/p + G(x)^q/q$ since the exponential function is convex. Now integrate. □

1.4.3 Minkowski's inequality. Suppose $1 < p < \infty$ and let f, g be measurable, nonnegative functions on X. Then

$$\left(\int (f+g)^p\right)^{1/p} \le \left(\int f^p\right)^{1/p} + \left(\int g^p\right)^{1/p}.$$

Assuming the right-hand side is finite, equality holds here if and only if there are $\alpha, \beta \in [0, \infty)$, not both zero, such that $\alpha f = \beta g$ almost everywhere.

Sketch of proof. If the left-hand side is in $(0, \infty)$ use that $(f+g)^p = f(f+g)^{p-1} + g(f+g)^{p-1}$. If it is infinity use the convexity of $t \mapsto t^p$, i.e., $((f+g)/2)^p \le (f^p + g^p)/2$. □

1.4.4 p-semi-norms. For $0 < p < \infty$ and a complex-valued, measurable function f define

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

If this is finite, f is called p-integrable.

If $1 \le p < \infty$ Minkowski's inequality 1.4.3 allows to show that the set of all p-integrable functions forms a vector space. The function $f \mapsto \|f\|_p$ is a semi-norm on it.

1.4.5 Essentially bounded functions. If f is a complex-valued, measurable function defined on X let $S = \{\alpha \in [0, \infty) : |f(x)| \le \alpha \text{ almost everywhere}\}$. Unless S is empty we call f *essentially bounded* and set

$$\|f\|_\infty = \inf S.$$

If $S = \emptyset$ we define instead $\|f\|_\infty = \infty$. It follows that $|f(x)| \le \|f\|_\infty$ almost everywhere. The number $\|f\|_\infty$ is called the *essential supremum* of $|f|$.

The function $f \mapsto \|f\|_\infty$ is a semi-norm on the vector space of all essentially bounded functions.

1.4.6 L^p-spaces. Fix $p \in [1, \infty]$. If $1 \le p < \infty$ the p-integrable functions form a semi-normed space. So do the essentially bounded functions when $p = \infty$. Note that $\|f - g\|_p = 0$ if and only if f and g are equal almost everywhere. Equality almost everywhere is an equivalence relation which partitions the given space into equivalence classes. The class of all functions related to f is denoted by $[f]$. The collection of these classes is again a vector space (addition and scalar multiplication interact well the classes) which we denote by $L^p(\mu)$. The function $[f] \mapsto \|f\|_p$ is well defined and turns $L^p(\mu)$ into a normed vector space. Products of classes are also well defined, i.e., $[f][g] = [fg]$. Henceforth, abusing notation, we will generally write f even when we mean its class $[f]$.

1.4.7 Hölder's and Minkowski's inequalities. If p and q are conjugate exponents in $[1, \infty]$, if $f \in L^p(\mu)$ and $g \in L^q(\mu)$, then $fg \in L^1(\mu)$ and

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

This follows from Hölder's inequality 1.4.2 except when the exponents are one and infinity when it follows from basic properties of integrals.

Also, if $p \in [1, \infty]$ and $f, g \in L^p(\mu)$, then, as we saw in 1.4.6,

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

1.4.8 Cauchy sequences of p-integrable and essentially bounded functions. If $n \mapsto f_n$ is a Cauchy sequence in the set of essentially bounded functions, it converges almost everywhere to a measurable function. If $1 \leq p < \infty$ and $n \mapsto f_n$ is a Cauchy sequence in the set of p-integrable functions, then there exists a subsequence $k \mapsto f_{n_k}$ converging pointwise almost everywhere to some measurable function f.

Sketch of proof. If $p = \infty$ let $B_{n,m} = \{x : |f_n(x) - f_m(x)| > \|f_n - f_m\|_\infty\}$. The union of all these sets, B, has measure zero. If $x \in B^c$ then $f_n(x)$ is a Cauchy sequence in $\mathbb{C}$ and hence converges to some number $f(x)$.

Now let $p < \infty$. There is a subsequence $k \mapsto f_{n_k}$ such that $\|f_{n_{k+1}} - f_{n_k}\|_p < 2^{-k}$. Define $g_k = \sum_{j=1}^k |f_{n_{j+1}} - f_{n_j}|$ and $g = \lim g_k$. Use Fatou's lemma and Minkowski's inequality to show that $\|g\|_p \leq 1$. Since g is finite almost everywhere $f_{n_{k+1}} = f_{n_1} + \sum_{j=1}^k (f_{n_{j+1}} - f_{n_j})$ converges almost everywhere to some function f as $k \to \infty$. □

1.4.9 $L^p(\mu)$ is a Banach space. If $p \in [1, \infty]$ then $L^p(\mu)$ is a Banach space.

Sketch of proof. Let $n \mapsto f_n$ be a Cauchy sequence in $L^p(\mu)$, and f, f_{n_j}, and B as in Theorem 1.4.8. If $1 \leq p < \infty$ then, by Fatou, $\|f_n - f\|_p^p \leq \varepsilon^p$ for sufficiently large n. Therefore $f = f_n + (f - f_n)$ gives rise to an element in $L^p(\mu)$ and $\|f_n - f\|_p \to 0$.

A similar argument works for $p = \infty$. □

1.4.10 Approximation by simple functions. We extend our definition of simple functions to include all complex linear combinations of characteristic functions of measurable sets. Denote the set of all simple functions by S and the set of simple functions which are different from zero only on a set of finite measure by S_0. Then S is dense in $L^\infty(\mu)$ and S_0 is dense in $L^p(\mu)$ for every $p \in [1, \infty)$.

1.5. Exercises

1.1. Find all σ-algebras for $X = \{1, 2, 3\}$.

1.2. Investigate the cancelation "laws" in $[0, \infty]$ which identify conditions under which either of the statements $a + b = c + b$ or $a \cdot b = c \cdot b$ implies $a = c$.

1.3. Suppose μ is a measure. Prove that the sequence $n \mapsto \mu(A_n)$ converges to $\mu(A)$ provided that $A_1 \supset A_2 \supset ... \supset A = \bigcap_{n=1}^\infty A_n$, and $\mu(A_1)$ is finite.

1.4. Give an example of a measure space and a sequence A_n of measurable sets such that $A_n \supset A_{n+1}$ but $\lim_{n\to\infty} \mu(A_n) \neq \mu\left(\bigcap_{n=1}^\infty A_n\right)$.

1.5. Show that any positive measure is countably subadditive, i.e., that $\mu(\bigcup_{j=1}^\infty A_j) \leq \sum_{j=1}^\infty \mu(A_j)$ if μ is a positive measure and the A_j are measurable sets.

1.6 (Numerical series). Suppose X is a subset of the integers, $\mathcal{M}$ the corresponding power set, and μ the counting measure. Show that integrals turn into sums or series. In particular, show that $\int f = \sum_{n=1}^\infty f(n)$, if $X = \mathbb{N}$. Which property of series corresponds to the concept of integrability?

1.7. Suppose $(X, \mathcal{M}, \mu)$ is a measure space with a positive measure μ and f is a complex-valued integrable function on X. If $\int f\chi_E = 0$ for every $E \in \mathcal{M}$, prove that $f = 0$ almost everywhere on X.

1.8. Suppose $(X, \mathcal{M}, \mu)$ is a measure space with a positive measure μ and f is integrable with respect to μ. If $|\int f| = \int |f|$, prove the existence of an $\alpha \in \mathbb{C}$ such that $|f| = \alpha f$ almost everywhere.

1.9 (Convergence in measure). Suppose $(X, \mathcal{M}, \mu)$ is a measure space with a positive measure μ. The sequence $f_n : X \to \mathbb{C}$ of measurable functions is said to converge in measure to $f : X \to \mathbb{C}$ if $\lim_{n\to\infty} \mu(\{x \in X : |f_n(x) - f(x)| \geq \delta\}) = 0$ for every $\delta > 0$. Show that $f_n \to f$ in measure, if $f_n \to f$ pointwise almost everywhere and $\mu(X) < \infty$. Give a counterexample for the case when $\mu(X) = \infty$.

1.10 (Egorov's theorem). Suppose $\mu(X) < \infty$ and f_n is a sequence of complex-valued measurable functions on X converging pointwise to some function f. Then, for every $\varepsilon > 0$, there exists a measurable set $E \subset X$ such that $\mu(E) < \varepsilon$ and $f_n \to f$ uniformly on E^c.

1.11. Prove that the geometric mean of n positive numbers is not larger than their arithmetic mean: $\sqrt[n]{y_1 ... y_n} \leq (y_1 + ... + y_n)/n$, in particular $2ab \leq a^2 + b^2$ for real numbers a and b.

1.12. If a_1, ..., a_n are non-negative real numbers and $p \in [1, \infty)$, prove that

$$\big(\sum_{k=1}^{n} a_k\big)^p \leq n^{p-1} \sum_{k=1}^{n} a_k^p.$$

1.13. Suppose $(X, \mathcal{M}, \mu)$ is a measure space with a positive measure μ. Construct a Cauchy sequence $n \mapsto f_n$ in $L^p(\mu)$ which does not converge for any $x \in X$. You may assume that μ is a measure which assigns to intervals their lengths.

CHAPTER 2

Measures

2.1. Types of measures

2.1.1 Generating a σ-algebra. Let X be a set. The intersection of all elements of a collection of σ-algebras in X is again a σ-algebra in X. Therefore, if $\mathcal{A}$ is any collection of subsets of X, there exists a smallest σ-algebra $\mathcal{M}(\mathcal{A})$ in X containing $\mathcal{A}$. $\mathcal{M}(\mathcal{A})$ is called the σ-algebra generated by $\mathcal{A}$.

2.1.2 The Borel σ-algebra. If X is a topological space, the σ-algebra generated by the topology is called the *Borel σ-algebra* and is denoted by $\mathcal{B}(X)$. Its elements are called *Borel sets*. A measure defined on a σ-algebra containing all the Borel sets is called a *Borel measure*. All closed sets, all F_σ sets (countable unions of closed sets) and all G_δ sets (countable intersections of open sets) are Borel sets.

If X and Y are topological spaces, we call the function $f : X \to Y$ Borel measurable (or simply Borel), if it is measurable with respect to $\mathcal{B}(X)$, i.e., if the preimage of any open set in Y is in $\mathcal{B}(X)$. Note that every continuous function $f : X \to Y$ is Borel.

If $(X, \mathcal{M})$ is a measurable space, Y a topological space, and $f : X \to Y$ is measurable, then $\mathcal{B}(Y) \subset \{E \subset Y : f^{-1}(E) \in \mathcal{M}\}$.

2.1.3 Restriction of a measure. Let $(X, \mathcal{M}, \mu)$ be a measure space and W a measurable subset of X. Then $\mathcal{M}_W = \{E \cap W : E \in \mathcal{M}\}$ is a σ-algebra in W and $\mu_W = \mu|_{\mathcal{M}_W}$ is a measure on $\mathcal{M}_W$.

2.1.4 Finite and σ-finite measures. A positive measure on X is called *finite*, if $\mu(X)$ is finite. It is called *σ-finite* if there is countable collection of sets of finite measure whose union is X.

2.1.5 Probability measures. A positive measure μ on X such that $\mu(X) = 1$ is called a *probability measure*. The measurable sets are then called *events*. For instance the measure space $(X, \mathcal{M}, \mu)$ where $X = \{1, ..., 6\}$, $\mathcal{M} = \mathcal{P}(X)$, $\mu(\{k\}) = 1/6$ represents a die. $\mu(A)$ gives the probability that rolling the die yields one of the elements contained in A.

2.1.6 Complete measures. A positive measure μ is called *complete* if every subset of a set of measure zero is measurable.

2.1.7 Completion of positive measures. If $(X, \mathcal{M}, \mu)$ is a measure space with positive measure μ then,

$$\overline{\mathcal{M}} = \{E \subset X : \exists A, B \in \mathcal{M} : A \subset E \subset B, \mu(B \setminus A) = 0\}$$

is a σ-algebra and $\overline{\mu} : \overline{\mathcal{M}} \to [0, \infty] : E \mapsto \overline{\mu}(E) = \mu(A)$ if $A \subset E \subset B$ and $\mu(B \setminus A) = 0$ is a complete positive measure which extends μ.

$\overline{\mu}$ is called the *completion* of μ and $\overline{\mathcal{M}}$ is called the *completion* of $\mathcal{M}$ with respect to μ.

2.1.8 Measurability of functions and completion of measures. Suppose $\overline{\mathcal{M}}$ is the completion of a σ-algebra $\mathcal{M}$ with respect to the positive measure μ and f is a complex-valued $\overline{\mathcal{M}}$-measurable function. Then there is an $\mathcal{M}$-measurable function g and a set $N \in \mathcal{M}$ such that $\mu(N) = 0$ and $f = g$ outside N.

Sketch of proof. One proves this, in turn, for characteristic functions, simple functions, non-negative functions, and, finally, for complex-valued functions. □

2.1.9 Regular positive measures. If μ is a positive Borel measure a measurable set E is called *outer regular* with respect to μ if

$$\mu(E) = \inf\{\mu(V) : E \subset V,\, V \text{ open}\}.$$

It is called *inner regular* with respect to μ if

$$\mu(E) = \sup\{\mu(K) : K \subset E,\, K \text{ compact}\}.$$

If E is both outer and inner regular it is called simply *regular*. The measure μ is called outer or inner regular or just regular if every measurable set has the respective property.

2.2. Construction of measures

2.2.1 Outer measure. Let X be a set. A function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ is called *monotone* if $\mu^*(A) \le \mu^*(B)$ whenever $A \subset B$. μ^* is called *countably subadditive* (or σ*-subadditive*) if always $\mu^*(\bigcup_{j=1}^\infty A_j) \le \sum_{j=1}^\infty \mu^*(A_j)$.

A monotone, countably subadditive function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ for which $\mu^*(\emptyset) = 0$ is called an *outer measure* on $\mathcal{P}(X)$.

2.2.2 Constructing outer measures. Let X be a set. Suppose $\mathcal{E} \subset \mathcal{P}(X)$ and $|\cdot| : \mathcal{E} \to [0, \infty]$ are such that $\emptyset \in \mathcal{E}$, X is the countable union of elements of $\mathcal{E}$, and $|\emptyset| = 0$. Then $\mu^* : \mathcal{P}(X) \to [0, \infty]$ defined by

$$\mu^*(A) = \inf\{\sum_{j=1}^\infty |E_j| : E_j \in \mathcal{E}, A \subset \bigcup_{j=1}^\infty E_j\}$$

is an outer measure on $\mathcal{P}(X)$. If $A \subset \bigcup_{j=1}^\infty E_j$ with $E_j \in \mathcal{E}$ we call $\{E_j : j \in \mathbb{N}\}$ a *countable cover* of A by elements of $\mathcal{E}$.

Sketch of proof. Obviously, μ^* is well-defined, $\mu^*(\emptyset) = 0$, and μ^* is monotone. To show countable subadditivity note that one may assume that $\mu^*(A_j) < \infty$ for all $j \in \mathbb{N}$ and that $E_{j,k}$ can be chosen so that $\sum_{k=1}^\infty |E_{j,k}| - \mu^*(A_j)$ is sufficiently small. □

2.2.3 Carathéodory's construction of a measure. Let X be a set and μ^* an outer measure on $\mathcal{P}(X)$. Define

$$\mathcal{C} = \{A \subset X : \forall B \subset X : \mu^*(B) = \mu^*(A \cap B) + \mu^*(A^c \cap B)\}$$

and $\mu = \mu^*|_{\mathcal{C}}$. Then $\mathcal{C}$ is a σ-algebra and μ is a complete positive measure.

Sketch of proof. Obviously $\mathcal{C}$ contains X and the complement of any of its elements. Thus we need to show closedness of $\mathcal{C}$ under countable unions but we begin with finite unions. If $E_1, E_2 \in \mathcal{C}$ one shows that $\mu^*((E_1 \cup E_2) \cap B) + \mu^*((E_1 \cup E_2)^c \cap B) = \mu^*(B)$ by employing Carathéodory's criterion and subadditivity (twice). It follows that $\mathcal{C}$ contains unions of two and, after induction, also unions of finitely many of its elements. Of course, $\mathcal{C}$ is also closed under finite intersections. Moreover, if $E_1 \cap E_2 = \emptyset$, Carathéodory's criterion gives

$\mu^*((E_1 \cup E_2) \cap R) = \mu^*(E_1 \cap R) + \mu^*(E_2 \cap R)$ by choosing $B = (E_1 \cup E_2) \cap R$ and $A = E_1$. Induction shows now that

$$\mu^*((E_1 \cup ... \cup E_k) \cap R) = \sum_{j=1}^{k} \mu^*(E_j \cap R) \tag{4}$$

when E_1, ..., $E_k \in \mathcal{C}$ are pairwise disjoint, and, in particular, the finite additivity of $\mu^*|_{\mathcal{C}}$.

Now suppose $A_k \in \mathcal{C}$ for each $k \in \mathbb{N}$ and $A = \bigcup_{j=1}^{\infty} A_j$. Define $R_k = \bigcup_{j=1}^{k} A_j$ and $E_k = A_k \setminus R_{k-1}$. These are in $\mathcal{C}$. The latter are pairwise disjoint, $\bigcup_{j=1}^{k} E_j = R_k$, and $\bigcup_{j=1}^{\infty} E_j = A$. Therefore we find, using (4) and monotonicity of μ^*, that

$$\mu^*(B) = \sum_{j=1}^{k} \mu^*(E_j \cap B) + \mu^*(R_k^c \cap B) \geq \sum_{j=1}^{k} \mu^*(E_j \cap B) + \mu^*(A^c \cap B)$$

for every $k \in \mathbb{N}$. Taking the limit and using the countable subadditivity of μ^* we get $\mu^*(B) \geq \mu^*(A \cap B) + \mu^*(A^c \cap B)$. Since the opposite inequality always holds, $\mathcal{C}$ is closed under countable unions and hence a σ-algebra.

Equation (4) for $R = X$, monotonicity, and subadditivity give

$$\sum_{j=1}^{k} \mu^*(E_j) \leq \mu^*(\bigcup_{j=1}^{\infty} E_j) \leq \sum_{j=1}^{\infty} \mu^*(E_j)$$

for all $k \in \mathbb{N}$ when the E_j are pairwise disjoint. This proves countable additivity of μ.

Completeness of μ follows when we show that every set of outer measure 0 is in $\mathcal{C}$. Thus assume $\mu^*(A) = 0$. Then $\mu^*(A^c \cap B) = \mu^*(A \cap B) + \mu^*(A^c \cap B) \geq \mu^*(B) \geq \mu^*(A^c \cap B)$. ☐

2.3. Lebesgue measure on $\mathbb{R}$

2.3.1 Lengths of open intervals. Let $\mathcal{E} = \{(a, b) : -\infty < a \leq b < \infty\}$ be the set of finite open intervals and note that $\emptyset = (a, a) \in \mathcal{E}$. We define the *length* of such an interval (a, b) to be $b - a$ and denote it by $|(a, b)|$. By 2.2.2 the set $\mathcal{E}$ and the length function $|\cdot|$ give rise to an outer measure m^* on $\mathcal{P}(\mathbb{R})$.

2.3.2 m^* is an extension of length. We need to show that $\mathrm{m}^*((a, b)) = b - a$. It is obvious that $\mathrm{m}^*((a, b)) \leq b - a$ and we assume, by way of contradiction, that $\mathrm{m}^*((a, b)) = b - a - 3\delta$ for some positive δ. Thus there are intervals (a_k, b_k), $k \in \mathbb{N}$, such that $(a, b) \subset \bigcup_{k=1}^{\infty}(a_k, b_k)$ and $\sum_{k=1}^{\infty}(b_k - a_k) < b - a - 2\delta$. Since $[a + \delta, b - \delta]$ is compact there is a $K \in \mathbb{N}$ such that

$$[a + \delta, b - \delta] \subset \bigcup_{k=1}^{K} (a_k, b_k).$$

Now $a + \delta$ is in one of these K intervals, i.e., there is a $k_1 \in \{1, ..., K\}$ such that $a_{k_1} < a + \delta < b_{k_1}$. Unless $b_{k_1} > b - \delta$ there is a k_2 such that $a_{k_2} < b_{k_1} < b_{k_2}$. Induction shows now the existence of an $L \leq K$ such that $a_{k_\ell} < b_{k_{\ell-1}} < b_{k_\ell}$ for $\ell = 2, ..., L$ and $b_{k_L} > b - \delta$. This proves

$$\sum_{k=1}^{\infty}(b_k - a_k) > b - a - 2\delta$$

which is in contradiction with the previous estimate.

2.3.3 The Lebesgue σ-algebra and Lebesgue measure. The set

$$\mathcal{L}(\mathbb{R}) = \{A \subset \mathbb{R} : \forall B \subset \mathbb{R} : \mathrm{m}^*(B) = \mathrm{m}^*(A \cap B) + \mathrm{m}^*(B \cap A^c)\}$$

is a σ-algebra called the *Lebesgue σ-algebra*. Its elements are called *Lebesgue measurable* sets.

The restriction of m^* to $\mathcal{L}(\mathbb{R})$ is a complete positive measure. It is called *Lebesgue measure* and is usually denoted by m., i.e.,

$$\mathrm{m} = \mathrm{m}^*|_{\mathcal{L}(\mathbb{R})}.$$

2.3.4 Important properties of Lebesgue measure. The following statements hold:

(1) A countable set is Lebesgue measurable and has measure zero.
(2) The outer measure of each of the intervals $[a, b]$, $[a, b)$ and $(a, b]$ is $b - a$.
(3) Open intervals are Lebesgue measurable.
(4) The Borel σ-algebra $\mathcal{B}(\mathbb{R})$ is contained in the Lebesgue σ-algebra $\mathcal{L}(\mathbb{R})$.
(5) Lebesgue measure is translation-invariant, i.e., $\mathrm{m}(A) = \mathrm{m}(x + A)$ where $x + A = \{x + a : a \in A\}$.

2.3.5 Regularity of Lebesgue measure. Recall that Lebesgue measure is a Borel measure since the Borel σ-algebra is contained in the Lebesgue σ-algebra.

THEOREM. Lebesgue measure on $\mathbb{R}$ is regular. Moreover the following statements hold:

(1) If E is a Lebesgue measurable set and ε a positive number then there is a closed set C and an open set V such that $C \subset E \subset V$ and $\mathrm{m}(V \setminus C) < \varepsilon$.
(2) If E is a Lebesgue measurable set then there is an F_σ-set F and a G_δ-set G such that $F \subset E \subset G$ and $\mathrm{m}(G \setminus F) = 0$.

SKETCH OF PROOF. Let E be a measurable set and $\varepsilon > 0$ be given. For each k we have an open set V_k such that $[-k, k] \cap E \subset V_k$ and $\mathrm{m}([-k, k] \cap E) + \varepsilon/2^{k+1} > \mathrm{m}(V_k)$. Now notice that $V = \bigcup_{k=1}^\infty V_k$ is open, that $E \subset V$, and that $\mathrm{m}(V \setminus E) < \varepsilon/2$. This proves outer regularity of Lebesgue measure.

This proves also the existence of an open set U which contains E^c and satisfies $\mathrm{m}(U \setminus E^c) < \varepsilon/2$. Let $C = U^c$ and note that $U \setminus E^c = E \setminus U^c$ to obtain a closed set C such that $C \subset E$ and $\mathrm{m}(E \setminus C) < \varepsilon/2$. Since $C = \bigcup_{k=1}^\infty ([-k, k] \cap C)$ is a countable union of compact sets inner regularity of Lebesgue measure follows.

Statements (1) and (2) are now also immediate. □

2.3.6 Completion of the Borel σ-algebra. The Lebesgue σ-algebra $\mathcal{L}(\mathbb{R})$ is the completion of the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ with respect to Lebesgue measure.

SKETCH OF PROOF. If $E \in \mathcal{L}(\mathbb{R})$, then $E \in \overline{\mathcal{B}(\mathbb{R})}$, the completion of $\mathcal{B}(\mathbb{R})$, by regularity. For the converse one checks Carathéodory's criterion. □

2.4. Comparison of the Riemann and the Lebesgue integral

2.4.1 Partitions. A *partition* P of $[a, b]$ is a finite subset of $[a, b]$ which contains both a and b. If the number of elements in P is $n + 1$ we will label them so that

$$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b.$$

A partition P' is called a *refinement* of P if $P \subset P'$. $P \cup Q$ is called the *common refinement* of the partitions P and Q.

2.4.2 Upper and lower sums. If P is a partition of $[a, b]$ with $n+1$ elements and $f : [a, b] \to \mathbb{R}$ is a bounded function we define the *lower Riemann sum* $L(f, P)$ and the *upper Riemann sum* $U(f, P)$ by

$$L(f, P) = \sum_{j=1}^{n} m_j (x_j - x_{j-1})$$

and

$$U(f, P) = \sum_{j=1}^{n} M_j (x_j - x_{j-1})$$

where $m_j = \inf\{f(x) : x_{j-1} \le x \le x_j\}$ and $M_j = \sup\{f(x) : x_{j-1} \le x \le x_j\}$.

If P' is a refinement of the partition P then

$$L(f, P) \le L(f, P') \le U(f, P') \le U(f, P).$$

2.4.3 Definition of the Riemann integral. Let $[a, b]$ be a bounded interval in $\mathbb{R}$ and $f : [a, b] \to \mathbb{R}$ a bounded function. The numbers $\sup\{L(f, P) : P \text{ is a partition of } [a, b]\}$ and $\inf\{U(f, P) : P \text{ is a partition of } [a, b]\}$ are called lower and upper Riemann integrals of f over $[a, b]$, respectively. Note that the lower Riemann integral is never larger than the upper Riemann integral.

The function f is called *Riemann integrable* over $[a, b]$ if its lower and upper Riemann integrals coincide. This common value is then called the *Riemann integral* of f over $[a, b]$ and is denoted by $\int_a^b f$. We emphasize that, by definition, Riemann integrable functions are bounded and are defined on finite intervals.

2.4.4 Comparison of the Riemann and the Lebesgue integral. Let $[a, b]$ be a bounded interval in $\mathbb{R}$. If $f : [a, b] \to \mathbb{R}$ is Riemann integrable over $[a, b]$ then it is Lebesgue integrable and $\int_a^b f = \int_{\mathrm{m}} f$ (here m is the restriction of Lebesgue measure to $[a, b]$).

Sketch of proof. Let $k \mapsto P_k$ be a sequence of successive refinements of partitions such that $\lim_{k\to\infty} L(f, P_k)$ equals the lower Riemann integral, and that $\lim_{k\to\infty} U(f, P_k)$ equals the upper Riemann integral. This uses 2.4.2.

The lower and upper Riemann sums $L(f, P_k)$ and $U(f, P_k)$ can be represented as Lebesgue integrals of simple functions ℓ_k and u_k. The sequences $k \mapsto \ell_k$ and $k \mapsto u_k$ have pointwise limits defining measurable functions ℓ and u such that $\ell(x) \le f(x) \le u(x)$.

The dominated convergence theorem implies that $\int_{\mathrm{m}} \ell = \lim_{k\to\infty} L(f, P_k)$ equals the lower Riemann integral of f. Likewise we have that $\int_{\mathrm{m}} u = \lim_{k\to\infty} U(f, P_k)$ equals the upper Riemann integral of f.

Now suppose that f is Riemann integrable. Then $\int_{\mathrm{m}} \ell = \int_{\mathrm{m}} u$ and this shows that, almost everywhere, $\ell = f = u$. Hence f is measurable and its Lebesgue integral is equal to the Riemann integral. □

2.4.5 The set of Riemann integrable functions. The (bounded) function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if and only if it is almost everywhere continuous.

Sketch of proof. We use the notation from 2.4.4. We may choose the partitions in such a way that the distance of adjacent points of P_k is at most $1/k$. The statement follows from the observation that, if x belongs to none of the points of the partitions P_k, then f is continuous at x if and only if $\ell(x) = u(x)$. □

2.5. Complex measures and their total variation

2.5.1 Complex measures. Recall that a *complex measure* is a complex-valued countably additive function on a given σ-algebra. This includes, of course, the case when all values of the measure are real in which case we will speak of a *real measure*. Note that when λ is a complex measure and E_1, E_2, ... are pairwise disjoint and measurable sets, then $\sum_{j=1}^{\infty} \lambda(E_j)$ converges absolutely.

Given two complex measures μ and λ defined on the same σ-algebra and a complex number c we may define the complex measures $\mu + \lambda$ and $c\mu$ by $(\mu + \lambda)(E) = \mu(E) + \lambda(E)$ and $(c\mu)(E) = c\mu(E)$, respectively. Thus the set of all complex measures on a given σ-algebra is a complex vector space.

If λ is a complex measure define the *real* and *imaginary part* of λ by $(\operatorname{Re}\lambda)(E) = \operatorname{Re}(\lambda(E))$ and $(\operatorname{Im}\lambda)(E) = \operatorname{Im}(\lambda(E))$, respectively. Both $\operatorname{Re}\lambda$ and $\operatorname{Im}\lambda$ are real measures.

2.5.2 Total variation of a measure. If the sets E_n, $n \in \mathbb{N}$, are pairwise disjoint and their union is E we call $\{E_n : n \in \mathbb{N}\}$ a *partition* of E.

Let μ be a complex measure on a σ-algebra $\mathcal{M}$. Define a function $|\mu| : \mathcal{M} \to [0, \infty]$ by

$$|\mu|(E) = \sup\{\sum_{n=1}^{\infty} |\mu(E_n)| : \{E_n : n \in \mathbb{N}\} \text{ is a partition of } E\}.$$

The function $|\mu|$ is called the *total variation* of μ. The number $|\mu|(A)$ is called the *total variation* of A (when the meaning of μ is clear from the context).

Theorem. The total variation of a complex measure is a positive measure which satisfies $|\mu|(E) \geq |\mu(E)|$ for all $E \in \mathcal{M}$. If λ is a positive measure satisfying $\lambda(E) \geq |\mu(E)|$ for all $E \in \mathcal{M}$ then $\lambda \geq |\mu|$.

Sketch of proof. Obviously, $|\mu|(\emptyset) = 0$ and $|\mu|(E) \geq |\mu(E)|$. To prove countable additivity choose a partition of E approximating $|\mu|(E)$ to establish one inequality and partitions approximating the $|\mu|(E_n)$ to establish the other. Finally note that $\lambda(E) \geq |\mu|(E)$ follows from $\lambda(E) = \sum_{n=1}^{\infty} \lambda(E_n) \geq \sum_{n=1}^{\infty} |\mu(E_n)|$. □

2.5.3 The total variation of a complex measure is a finite measure. Let $(X, \mathcal{M})$ be a measurable space and μ a complex measure on $\mathcal{M}$. Then $|\mu|(X) < \infty$.

Sketch of proof. Since $\sum_{k=1}^{\infty} \mu(A_k)$ is convergent when the A_k are pairwise disjoint, there can be no sequence $k \mapsto A_k$ of pairwise disjoint sets satisfying $|\mu(A_k)| > 1$. However, using induction and the lemma below, the assumption $|\mu|(X) = \infty$ allows to show the existence of a sequence $k \mapsto A_k$ of pairwise disjoint sets satisfying $|\mu(A_k)| > 1$. This contradiction proves the claim. □

Lemma. If $|\mu|(E) = \infty$ then there is a measurable set $A \subset E$ such that $|\mu|(A) = \infty$ $|\mu(A)| > 1$, and $|\mu(E \setminus A)| > 1$.

Sketch of proof. Since $|\mu|(E) = \infty$ there are measurable, pairwise disjoint subsets E_1, ..., E_N of E such that $\sum_{k=1}^{N} |\mu(E_k)| > \pi(1 + |\mu(E)|)$. Set $\mu(E_k) = r_k \mathrm{e}^{i\alpha_k}$ and define the function $f : [-\pi, \pi] \to [0, \infty)$ by

$$f(t) = \sum_{k=1}^{N} r_k \cos_+(\alpha_k - t).$$

Since f is continuous it has a maximum attained at a point t_0. Let $S = \{k \in \{1, ..., N\} : \cos(\alpha_k - t_0) > 0\}$ and $B = \bigcup_{k \in S} E_k$. At least one of B and $E \setminus B$ has infinite variation. Since

$$|\mu(B)| = \left|\sum_{k \in S} \mu(E_k)\right| \geq f(t_0) \geq \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)dt = \frac{1}{\pi} \sum_{k=1}^{N} |\mu(E_k)| > 1 + |\mu(B) + \mu(E \setminus B)|,$$

we have both $|\mu(B)| > 1$ and $|\mu(E \setminus B)| > 1$. $\square$

2.5.4 A norm for complex measures. Given a complex measure μ define $\|\mu\| = |\mu|(X)$. This function $\|\cdot\|$ is a norm turning the set of complex measures into a normed vector space.

2.5.5 Positive and negative variations. Suppose now that μ is a real measure. Define

$$\mu^+ = \frac{1}{2}(|\mu| + \mu), \quad \mu^- = \frac{1}{2}(|\mu| - \mu).$$

Then both μ^+ and μ^- are finite positive measures called the *positive* and *negative variation* of μ, respectively. The pair (μ^+, μ^-) is called the *Jordan decomposition* of μ.

2.6. Absolute continuity and mutually singular measures

2.6.1 Absolute continuity. A measure λ (complex or positive) is called *absolutely continuous* with respect to a positive measure μ if all sets of μ-measure zero also have λ-measure zero. We denote this relationship by $\lambda \ll \mu$.

2.6.2 A complex measure is absolutely continuous with respect to its total variation. Let $(X, \mathcal{M}, \lambda)$ be a measure space with a complex measure λ. Since $|\lambda(E)| \leq |\lambda|(E)$ it follows that $\lambda \ll |\lambda|$.

2.6.3 A criterion for absolute continuity. Suppose λ is a complex measure and μ is a positive measure. Then $\lambda \ll \mu$ if and only if for every $\varepsilon > 0$ exists a $\delta > 0$ such that for every measurable set E the condition $\mu(E) < \delta$ implies $|\lambda(E)| < \varepsilon$.

Sketch of proof. The "if" direction is simple. For the "only if" direction we prove the contrapositive. Thus assume there is an $\varepsilon > 0$ and there are sets E_n such that $\mu(E_n) < 2^{-n}$ but $|\lambda(E_n)| \geq \varepsilon$. If $F = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n$ we have $\mu(F) = 0$ but $|\lambda|(F) \geq \varepsilon$. Finally note that $\lambda \ll \mu$ if and only if $|\lambda| \ll \mu$. $\square$

2.6.4 Mutually singular measures. Suppose μ is a measure. If there is a measurable set A such that $\mu(E) = \mu(E \cap A)$ for all $E \in \mathcal{M}$ we say that the measure μ is *concentrated* on A. It follows that μ is concentrated on A if and only if $\mu(E) = 0$ for all measurable $E \subset A^c$. Lebesgue measure, for instance, is concentrated on the set of irrational numbers.

If μ is concentrated on A and $A \subset B$, then μ is concentrated on B. If μ is concentrated on A_1 and also on A_2, then μ is concentrated on $A_1 \cap A_2$.

Two measures μ and λ are called *mutually singular* if they are concentrated on disjoint sets. This is indicated by $\lambda \perp \mu$.

2.6.5 Basic properties. Suppose that μ, λ, λ_1, and λ_2 are measures on $\mathcal{M}$ and that μ is positive. Then the following statements hold:

(1) λ is concentrated on A if and only if $|\lambda|$ is.
(2) $\lambda_1 \perp \lambda_2$ if and only if $|\lambda_1| \perp |\lambda_2|$.
(3) $\lambda \ll \mu$ if and only if $|\lambda| \ll \mu$.
(4) If $\lambda_1 \ll \mu$ and $\lambda_2 \perp \mu$ then $\lambda_1 \perp \lambda_2$.
(5) If $\lambda \ll \mu$ and $\lambda \perp \mu$ then $\lambda = 0$.

We get two more statements when it makes sense to define $\alpha\lambda_1 + \beta\lambda_2$, i.e., when λ_1 and λ_2 are complex measures and $\alpha, \beta \in \mathbb{C}$ or when they are both positive measure and $\alpha, \beta \geq 0$.

(1) If $\lambda_1 \perp \lambda$ and $\lambda_2 \perp \lambda$, then $\alpha\lambda_1 + \beta\lambda_2 \perp \lambda$.
(2) If $\lambda_1 \ll \mu$ and $\lambda_2 \ll \mu$, then $\alpha\lambda_1 + \beta\lambda_2 \ll \mu$.

2.6.6 Discrete and continuous measures. Let $(X, \mathcal{M})$ be a measurable space such that $\mathcal{M}$ contains all countable subsets of X. A measure μ on $\mathcal{M}$ is called *discrete*, if it is concentrated on a countable set. It is called *continuous*, if $\mu(\{x\}) = 0$ for all $x \in X$. Discrete and continuous measures are mutually singular.

A complex or sigma-finite positive measure μ can be expressed uniquely as the sum of a discrete and a continuous measure.

2.7. Exercises

2.1. Let $\mathcal{F} = \{\{n\} : n \in \mathbb{Z}\}$ and $\mathcal{G} = \{\{r\} : r \in \mathbb{R}\}$. Find the smallest σ-algebra in $\mathbb{Z}$ containing $\mathcal{F}$ and the smallest σ-algebra in $\mathbb{R}$ containing $\mathcal{G}$.

2.2. Construct the measure modeling the rolling of two dice. That is identify X, $\mathcal{M}$, and μ.

2.3. Suppose $(X, \mathcal{M}, \mu)$ is a complete measure space, g is a measurable function, and $f = g$ almost everywhere. Show that f is also measurable.

2.4. Suppose $(X, \mathcal{M}, \mu)$ is a measure space with a positive measure μ. Define $\lambda^*(E) = \inf\{\mu(A) : A \in \mathcal{M}, E \subset A\}$ for every $E \subset X$ and $\mathcal{C} = \{E \subset X : \forall T \subset X : \lambda^*(T) = \lambda^*(E \cap T) + \lambda^*(E^c \cap T)\}$.

Prove the following statements:

(1) λ^* is an outer measure on $\mathcal{P}(X)$.
(2) $\mathcal{M} \subset \mathcal{C}$.
(3) $\lambda^*(E) = \mu(E)$ whenever $E \in \mathcal{M}$.
(4) If μ is σ-finite, $\mathcal{C}$ is the completion of $\mathcal{M}$ with respect to μ.

2.5. Determine the regularity properties of the counting measure and the Dirac measure on $\mathbb{R}^n$.

2.6. Find a function which is Lebesgue integrable but not Riemann integrable.

2.7. Find a function f on $[0, \infty)$ such that the improper Riemann integral

$$\int_0^\infty f = \lim_{R\to\infty} \int_0^R f$$

exists and is finite, but f is not Lebesgue integrable.

2.8. A step function is a simple function $s : \mathbb{R} \to \mathbb{C}$ such that $\{x : s(x) = \alpha\}$ is a finite union of bounded intervals for all α (allowing, of course, for the empty set). Show that step functions are dense in $L^p(\mathrm{m})$.

2.9. Show that the 2.6.3 may fail, if λ is a positive measure which is allowed to assume the value ∞.

CHAPTER 3

Integration on Product Spaces

3.1. Product measure spaces

Throughout this section $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ denote measure spaces with positive measures μ and ν.

3.1.1 Measurable rectangles. For every $A \in \mathcal{M}$ and $B \in \mathcal{N}$ the set $A \times B$ is called a *measurable rectangle* in $X \times Y$. The σ-algebra generated by the measurable rectangles is denoted by $\mathcal{M} \otimes \mathcal{N}$. It is called a *product σ-algebra*.

3.1.2 Carathéodory's construction. Both $\emptyset = \emptyset \times \emptyset$ and $X \times Y$ are measurable rectangles so that by 2.2.2 the function

$$\lambda^*(E) = \inf \{ \sum_{j=1}^{\infty} \mu(A_j)\nu(B_j) : A_j \in \mathcal{M}, B_j \in \mathcal{N}, E \subset \bigcup_{j=1}^{\infty} A_j \times B_j \}$$

is an outer measure on $\mathcal{P}(X \times Y)$. By 2.2.3 the set

$$\mathcal{C} = \{E \subset X \times Y : \forall T \subset X \times Y : \lambda^*(T) = \lambda^*(E \cap T) + \lambda^*(E^c \cap T)\}$$

is a σ-algebra and $\lambda = \lambda^*|_{\mathcal{C}}$ is a complete positive measure.

The measurable rectangles are contained in $\mathcal{C}$. Moreover, aided by the monotone convergence theorem, one shows that λ is an extension of the map $E \times F \mapsto \mu(E)\nu(F)$ defined on the measurable rectangles.

3.1.3 Product σ-algebra and product measure. Let $\mathcal{C}$ and λ be as in 3.1.2 and assume that μ and ν are σ-finite. Then the completion $\overline{\mathcal{M} \otimes \mathcal{N}}$ of $\mathcal{M} \otimes \mathcal{N}$ with respect to λ is $\mathcal{C}$.

Sketch of proof. Since the measurable rectangles are in both $\mathcal{M} \otimes \mathcal{N}$ and $\mathcal{C}$ we have $\mathcal{M} \otimes \mathcal{N} \subset \mathcal{C}$ but $\lambda|_{\mathcal{M} \otimes \mathcal{N}}$ is, in general, not a complete measure even if μ and ν are. Suppose $A, B \in \mathcal{M} \otimes \mathcal{N}$, $A \subset E \subset B$, and $\lambda(B \setminus A) = 0$. Then E satisfies Carathéodory's criterion so that $\overline{\mathcal{M} \otimes \mathcal{N}} \subset \mathcal{C}$. To show the converse assume $E \in \mathcal{C}$ and, at first, that $\lambda(E) < \infty$. For every $k \in \mathbb{N}$ there is a sequence of measurable rectangles $A_{k,n} \times B_{k,n}$ so that $E \subset D_k = \bigcup_{n=1}^{\infty} A_{k,n} \times B_{k,n}$ and $\lambda(D_k \setminus E) \leq 1/k$. Let $D = \bigcap_{k=1}^{\infty} D_k$. Then $D \in \mathcal{M} \otimes \mathcal{N}$, $E \subset D$, and $\lambda(D \setminus E) = 0$. Since μ and ν are σ-finite the same result holds even if $\lambda(E) = \infty$. Similarly there is a $C \in \mathcal{M} \otimes \mathcal{N}$ such that $C \subset E$ and $\lambda(E \setminus C) = 0$. Thus we have $E \in \overline{\mathcal{M} \otimes \mathcal{N}}$. $\square$

The measure λ defined in 3.1.2 (defined on either $\mathcal{M} \otimes \mathcal{N}$ or $\overline{\mathcal{M} \otimes \mathcal{N}}$) is called a *product measure* and will henceforth be denoted by $\mu \otimes \nu$.

3.1.4 Sections of sets. Let E be a subset of $X \times Y$ and $x \in X$ and $y \in Y$. Then $E_x = \{y : (x, y) \in E\} \subset Y$ and $E^y = \{x : (x, y) \in E\} \subset X$ are called the *x-section* and the *y-section* of E, respectively.

Since $\{E \in \mathcal{M} \otimes \mathcal{N} : \forall x \in X : E_x \in \mathcal{N}\}$ is a σ-algebra, it equals $\mathcal{M} \otimes \mathcal{N}$. Thus one obtains the following statement: If $E \in \mathcal{M} \otimes \mathcal{N}$, $x \in X$, and $y \in Y$, then $E_x \in \mathcal{N}$ and $E^y \in \mathcal{M}$.

3.2. Fubini's theorem

Throughout this section $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ denote measure spaces with complete, positive, σ-finite measures μ and ν. The proof of Fubini's theorem is rather involved. The material in 3.2.1 – 3.2.3 serves only preparatory purposes.

3.2.1 Monotone classes. Let X be a set. A *monotone class* $\mathcal{A}$ is a collection of subsets of X with the property that it contains the union of any non-decreasing sequence of elements of $\mathcal{A}$ as well as the intersection of any non-increasing sequence of elements of $\mathcal{A}$.

Every σ-algebra is a monotone class and the intersection of monotone classes is again a monotone class. The monotone class generated by a collection $\mathcal{E}$ of subsets of X is the intersection of all monotone classes containing $\mathcal{E}$.

Theorem. $\mathcal{M} \otimes \mathcal{N}$ is the smallest monotone class containing all finite unions of measurable rectangles.

Sketch of proof. Let $\mathcal{A}$ be the smallest monotone class containing $\mathcal{E}$, the set of all finite unions of measurable rectangles. Then $\mathcal{A} \subset \mathcal{M} \otimes \mathcal{N}$. The claim follows if we can show that $\mathcal{A}$ contains complements and finite unions of its elements since this turns $\mathcal{A}$ into a σ-algebra. Thus, for $P \subset X \times Y$ define $\Omega(P) = \{Q \subset X \times Y : P \setminus Q, Q \setminus P, P \cup Q \in \mathcal{A}\}$ and note that (i) $P \in \Omega(Q)$ if and only if $Q \in \Omega(P)$ and (ii) $\Omega(P)$ is a monotone class. Now one shows, if $P \in \mathcal{E}$, then $\mathcal{E} \subset \Omega(P)$ and hence $\mathcal{A} \subset \Omega(P)$. Also, if $Q \in \mathcal{A}$, then $\mathcal{A} \subset \Omega(Q)$. □

3.2.2 Fubini's theorem – preliminary version. If $E \in \mathcal{M} \otimes \mathcal{N}$, then the function $x \mapsto \nu(E_x)$ is $\mathcal{M}$-measurable and the function $y \mapsto \mu(E^y)$ is $\mathcal{N}$-measurable. Moreover $(\mu \otimes \nu)(E) = \int_\mu \nu(E_x) = \int_\nu \mu(E^y)$ (here we write $\nu(E_x)$ and $\mu(E^y)$ as abbreviations for $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$, respectively).

Sketch of proof. Let Ω be the set of all those elements $E \in \mathcal{M} \otimes \mathcal{N}$ for which the conclusions of the theorem hold. The proof takes the following steps: (i) Any finite union of pairwise disjoint measurable rectangles is in Ω. (ii) Any finite union of measurable rectangles is a finite union of pairwise disjoint measurable rectangles. (iii) Ω contains the union of a non-decreasing sequence of its elements. (iv) Ω contains the intersection of a non-increasing sequence of its elements, provided they are contained in a rectangle $A \times B$ with $\mu(A), \nu(B) < \infty$. (v) Since μ and ν are σ-finite we have non-decreasing sequences $X_k \in \mathcal{M}$ and $Y_k \in \mathcal{N}$ such that $X \times Y = \bigcup_{k=1}^\infty X_k \times Y_k$ and $\mu(X_k), \nu(Y_k) < \infty$. Define $\Omega'_k = \{E \in \mathcal{M} \otimes \mathcal{N} : E \cap (X_k \times Y_k) \in \Omega\}$. Since Ω'_k is a monotone class containing all finite unions of measurable rectangles, it equals $\mathcal{M} \otimes \mathcal{N}$. (vi) Now suppose that E is the intersection of a non-increasing sequence $n \mapsto E_n \in \Omega$. Then $E = \bigcup_{k=1}^\infty (E \cap (X_k \times Y_k)) = \bigcup_{k=1}^\infty \bigcap_{n=1}^\infty (E_n \cap (X_k \times Y_k))$. Since $E_n \in \mathcal{M} \otimes \mathcal{N} = \Omega'_k$, the sequence $n \mapsto E_n \cap (X_k \times Y_k)$ is a non-increasing sequence in Ω so that, by (iv), $E \cap (X_k \times Y_k)$ is also in Ω. Finally, by (iii), E itself is in Ω. Thus Ω is a monotone class and hence equal to $\mathcal{M} \otimes \mathcal{N}$. □

3.2.3 Sections of functions. If f is a complex-valued function defined on $X \times Y$ and $x \in X$ we denote the function $y \mapsto f(x, y)$ by f_x and call it the *x-section* of f. Similarly f^y denotes the function $x \mapsto f(x, y)$ when y is a fixed element of Y.

If f is a $\overline{\mathcal{M} \otimes \mathcal{N}}$-measurable function and $f = g + h$ where g is $\mathcal{M} \otimes \mathcal{N}$-measurable and $h = 0$ almost everywhere with respect to $\mu \otimes \nu$, then the following statements are true:

(1) If $x \in X$, then g_x is $\mathcal{N}$-measurable.
(2) If $y \in Y$, then g^y is $\mathcal{M}$-measurable.
(3) For almost all $x \in X$ the function h_x is $\mathcal{N}$-measurable and zero almost everywhere.
(4) For almost all $y \in Y$ the function h^y is $\mathcal{M}$-measurable and zero almost everywhere.

Sketch of proof. Statements and (1) and (2) hold by 3.1.4. To prove (3) and (4) let P be the set where h does not vanish. Then $P \subset Q$ for some $Q \in \mathcal{M} \otimes \mathcal{N}$ and $(\mu \otimes \nu)(Q) = 0$. From 3.2.2 we get $0 = \int_\mu \nu(Q_x)$ so that $0 = \nu(Q_x) = \nu(P_x)$ for all x outside a set S of measure zero. For $x \notin S$ the function h_x is zero outside P_x, i.e., almost everywhere. Hence h_x is then $\mathcal{N}$-measurable. □

Recall that, by 2.1.8, there is always a $\mathcal{M} \otimes \mathcal{N}$-measurable function g which coincides with f almost everywhere.

3.2.4 Fubini's theorem. Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be measure spaces with complete, positive, σ-finite measures μ and ν and f a $\overline{\mathcal{M} \otimes \mathcal{N}}$-measurable function on $X \times Y$. Define, when it makes sense to do so, $\varphi(x) = \int_\nu f_x$, $\psi(y) = \int_\mu f^y$, $\varphi^*(x) = \int_\nu |f_x|$, and $\psi^*(y) = \int_\mu |f^y|$.

Then the following statements hold:

(1) If $0 \le f \le \infty$, then φ and ψ are $\mathcal{M}$-measurable and $\mathcal{N}$-measurable, respectively. Moreover,
$$\int_\mu \varphi = \int_{\mu \otimes \nu} f = \int_\nu \psi.$$
(2) If f is complex-valued and $\int_\mu \varphi^* < \infty$ or $\int_\nu \psi^* < \infty$, then $f \in L^1(\mu \otimes \nu)$.
(3) If $f \in L^1(\mu \otimes \nu)$, then $f_x \in L^1(\nu)$ for almost every $x \in X$, $f^y \in L^1(\mu)$ for almost every $y \in Y$, $\varphi \in L^1(\mu)$, and $\psi \in L^1(\nu)$. Moreover,
$$\int_\mu \varphi = \int_{\mu \otimes \nu} f = \int_\nu \psi. \tag{5}$$

Equation (5) is often written in terms of iterated integrals as
$$\int_{\mu(x)} \Big(\int_{\nu(y)} f(x, y) \Big) = \int_{\mu \otimes \nu} f = \int_{\nu(y)} \Big(\int_{\mu(x)} f(x, y) \Big).$$

Sketch of proof. Show (1), in turn, when f is the characteristic function of a measurable set, for simple functions, and then in general. Statement (2) follows from applying (1) to $|f|$. For (3) split f in real and imaginary parts and those in positive and negative parts. □

3.2.5 Counterexamples. Iterated integrals do not coincide in the following situations:

(1) Let $X = Y = \mathbb{N}$ and $\mu = \nu$ the counting measure (then $\mu \otimes \nu$ is the counting measure on $X \times Y$). Suppose $f(j, k) = 1$ for $j = k$, $f(j, k) = -1$ if $j = k + 1$, and zero otherwise.
(2) Let $X = Y = [0, 1]$, μ Lebesgue measure, ν the counting measure, and f the characteristic function of the main diagonal of $[0, 1]^2$.

3.3. Lebesgue measure on $\mathbb{R}^n$

3.3.1 Lebesgue measure on $\mathbb{R}^n$. We define Lebesgue measure m_n on $\mathbb{R}^n$ recursively. Specifically, given that m_n is already defined we set $\mathrm{m}_{n+1} = \mathrm{m}_n \otimes \mathrm{m}_1$ on $\mathcal{L}(\mathbb{R}^{n+1}) = \overline{\mathcal{L}(\mathbb{R}^n) \otimes \mathcal{L}(\mathbb{R})}$.

3.3.2 Basics. The following statements are true:

(1) Any open rectangular box $B = (a_1, b_1) \times ... \times (a_n, b_n)$ is an element of $\mathcal{L}(\mathbb{R}^n)$.
(2) If $B = (a_1, b_1) \times ... \times (a_n, b_n)$, then $\mathrm{m}_n(B)$ equals $|B| = \prod_{k=1}^n (b_k - a_k)$, the *volume* of the box (or area of the rectangle if $n = 2$) in geometry.
(3) m_n is a Borel measure on $\mathbb{R}^n$.
(4) Lebesgue measure on $\mathbb{R}^n$ is translation invariant, i.e., $\mathrm{m}_n(A) = \mathrm{m}_n(x+A)$ whenever $A \in \mathcal{L}(\mathbb{R}^n)$ and $x + A = \{x + a : a \in A\}$.

3.3.3 Further properties of Lebesgue measure. Let $\mathcal{E}_n$ be the set of all open rectangular boxes in $\mathbb{R}^n$ and $\phi^*(E) = \inf\{\sum_{j=1}^\infty |A_j| : A_j \in \mathcal{E}_n, E \subset \bigcup_{j=1}^\infty A_j\}$. Then ϕ^* is an outer measure on $\mathcal{P}(\mathbb{R}^n)$. Clearly $\mathrm{m}_n^*(E) \le \phi^*(E)$ for all $E \subset \mathbb{R}^n$. The opposite inequality also holds; to see this choose appropriate measurable rectangles $A_k \times B_k$ which cover E, then appropriate boxes and intervals $E_{k,j}$ and $F_{k,\ell}$ which cover A_k and B_k, respectively. In conclusion, $\mathrm{m}_n^* = \phi^*$, i.e.,

$$\mathrm{m}_n(E) = \inf\{\sum_{j=1}^\infty |A_j| : A_j \in \mathcal{E}_n, E \subset \bigcup_{j=1}^\infty A_j\}.$$

This result has the following two consequences:

(1) m_n is regular.
(2) $\mathcal{L}(\mathbb{R}^n)$ is the completion of $\mathcal{B}(\mathbb{R}^n)$ with respect to m_n.
(3) If $p + q = n$, then $\mathrm{m}_n = \mathrm{m}_p \otimes \mathrm{m}_q$.

3.3.4 The volume of balls. Let $B_n(x, R) \subset \mathbb{R}^n$ be the open ball of radius R centered at x. For each natural number n there exists a positive number σ_n such that $\mathrm{m}_n(B_n(x, R)) = \sigma_n R^n$ whenever $x \in \mathbb{R}^n$ and $R > 0$.

Sketch of proof. It is enough to consider $x = 0$. When $n = 1$, the claim holds with $\sigma_1 = 2$. Let $r(x) = \sqrt{R^2 - x^2}$. Then we have

$$\chi_{B_{n+1}(0,R)}(x_1, ..., x_{n+1}) = \chi_{(-R,R)}(x_{n+1})\chi_{B_n(0,r(x_{n+1}))}(x_1, ..., x_n)$$

and the claim follows by induction. The numbers σ_n may be computed explicitly in terms of the integrals $\int_0^1 (1 - x^2)^{(n-1)/2}\, dx = \int_0^{\pi/2} (\cos t)^n\, dt$. □

3.4. Exercises

3.1. Let $A = \{(x, y) \in \mathbb{R}^2 : x \in \mathbb{Q} \text{ or } y \in \mathbb{Q}\}$. Is A Lebesgue measurable? If so, what is its measure?

3.2. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. Show that the set $\{(x, f(x)) : x \in \mathbb{R}\}$ is in $\mathcal{L}(\mathbb{R}^2)$ and that its Lebesgue measure is 0.

3.3. Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are measure spaces with complete, positive, σ-finite measures μ and ν. Construct a function h with the following properties: (i) h is $\overline{\mathcal{M} \otimes \mathcal{N}}$-measurable and 0 almost everywhere, (ii) there is a point $x \in X$ for which h_x is not measurable, and (iii) there is a point $x \in X$ for which h_x is measurable but not 0 almost everywhere.

CHAPTER 4

The Lebesgue-Radon-Nikodym theorem

4.1. The Lebesgue-Radon-Nikodym theorem

4.1.1 The Lebesgue-Radon-Nikodym theorem. Let μ be a positive σ-finite measure on a σ-algebra $\mathcal{M}$ in X and λ a complex measure on $\mathcal{M}$. Then

(1) there exists a unique pair of complex measures λ_a and λ_s on $\mathcal{M}$ such that

$$\lambda = \lambda_a + \lambda_s, \quad \lambda_a \ll \mu, \quad \lambda_s \perp \mu.$$

If λ is positive and finite then so are λ_a and λ_s.

(2) There exists a unique function $h \in L^1(\mu)$ such that $\lambda_a(E) = \int_\mu h\chi_E$ for every $E \in \mathcal{M}$.

The pair λ_a, λ_s associated with λ is called the *Lebesgue decomposition* of λ.

Sketch of proof. This is a deep theorem. The proof below is due to von Neumann. We shall break it in several parts.

(1) Showing uniqueness of λ_a, λ_s and h is straightforward.
(2) It is sufficient to prove the theorem for a positive (but finite) measure λ.
(3) There exists a function $w \in L^1(\mu)$ such that $0 < w < 1$.
(4) Suppose there was a measurable function g such that $0 \le g \le 1$ and

$$\int_\lambda f(1-g) = \int_\mu fgw \tag{6}$$

for bounded, nonnegative, measurable functions f. Define $A = \{x : 0 \le g(x) < 1\}$ and $B = A^c = \{x : g(x) = 1\}$ as well as

$$\lambda_a(E) = \lambda(E \cap A) \quad \text{and} \quad \lambda_s(E) = \lambda(E \cap B).$$

Then λ_a and λ_s are finite positive measures satisfying $\lambda_a + \lambda_s = \lambda$ and $\lambda_s \perp \mu$.

(5) Now let $f = (1 + g + ... + g^n)\chi_E$ in (6). Then

$$\int_\lambda (1-g^{n+1})\chi_{E\cap A} = \int_\lambda (1-g^{n+1})\chi_E = \int_\mu (g + ... + g^{n+1})w\chi_E.$$

The monotone convergence theorem shows then that $\lambda_a(E) = \lambda(E \cap A) = \int_\mu h\chi_E$ where $h = wg/(1-g)$ on A. Taking $E = X$ shows $h \in L^1(\mu)$. It also follows that $\lambda_a \ll \mu$.

(6) It remains to show the existence of g. To this end define the finite measure φ by $\varphi(E) = \lambda(E) + \int_\mu w\chi_E$. Then

$$\int_\varphi f = \int_\lambda f + \int_\mu fw$$

for any nonnegative measurable function f.

(7) The inequality
$$\left|\int_\lambda f\right| \le \int_\lambda |f| \le \int_\varphi |f| \le \|f\|_{L^2(\varphi)}(\varphi(X))^{1/2}$$
holds for all $f \in L^2(\varphi)$. Therefore $T : L^2(\varphi) \to \mathbb{C} : f \mapsto \int_\lambda f$ is a bounded linear functional. The representation theorem A.2.4 guarantees now the existence of a $g \in L^2(\varphi)$ such that
$$\int_\lambda f = Tf = \langle f, \overline{g}\rangle_{L^2(\varphi)} = \int_\varphi fg.$$
Putting here $f = \chi_E$ for any E with $\varphi(E) > 0$ shows, using 1.2.9, that $0 \le g \le 1$ almost everywhere with respect to φ and hence with respect to both λ and μ. One may now choose a representative such that $0 \le g \le 1$.

□

4.1.2 Allowing for σ-finite measures. If λ is positive and σ-finite rather than finite the Lebesgue-Radon-Nikodym theorem takes the following form: There are unique positive σ-finite measures λ_a and λ_s on $\mathcal{M}$ such that $\lambda = \lambda_a + \lambda_s$, $\lambda_a \ll \mu$, and $\lambda_s \perp \mu$. Moreover, there is a unique non-negative measurable function h (not necessarily integrable) such that $\lambda_a(E) = \int_\mu h\chi_E$ for every $E \in \mathcal{M}$.

To prove this let $X = \bigcup_{n=1}^\infty X_n$ with pairwise disjoint sets X_n of finite λ-measure, define the functions g_n on X_n in analogy to 4.1.1, and extend them by 0 to all of X. Then let $g = \sum_{n=1}^\infty g_n$ and, as before, $A = \{x : g(x) < 1\}$, $\lambda_a(E) = \lambda(E \cap A)$, $\lambda_s(E) = \lambda(E \cap A^c)$, and $h = wg/(1-g)$ on A. Then λ_a, λ_s, and h have the required properties.

4.1.3 The Radon-Nikodym derivative. Suppose μ is a σ-finite positive measure on a σ-algebra $\mathcal{M}$ and λ is either a complex measure or else a σ-finite positive measure also defined on $\mathcal{M}$. If λ is absolutely continuous with respect to μ, then either 4.1.1 or else 4.1.2 defines a function h through the equation $\lambda(E) = \int_\mu h\chi_E$. This function is called the *Radon-Nikodym derivative* of λ with respect to μ. We denote it by (λ/μ).

4.1.4 The Radon-Nikodym derivative of a measure with respect to its total variation. Suppose λ is a complex measure defined on the σ-algebra $\mathcal{M}$. Since it is absolutely continuous with respect to its total variation $|\lambda|$, there is a Radon-Nikodym derivative $h = (\lambda/|\lambda|)$ in $L^1(|\lambda|)$. In fact, $|h| = 1$ almost everywhere with respect to $|\lambda|$. Of course, we may choose h so that $|h| = 1$ everywhere.

4.1.5 The chain rule for Radon-Nikodym derivatives. Suppose κ, λ and μ are σ-finite positive measures on a σ-algebra $\mathcal{M}$ and that $\kappa \ll \lambda \ll \mu$. Then $\kappa \ll \mu$ and
$$(\kappa/\mu) = (\kappa/\lambda)(\lambda/\mu).$$
This statement is also true when κ is a complex measure.

4.1.6 The Lebesgue decomposition of a total variation measure. Let μ be a positive σ-finite measure on a σ-algebra $\mathcal{M}$ in X and λ a complex measure on $\mathcal{M}$ so that $\lambda(E) = \lambda_a(E) + \lambda_s(E)$ with $\lambda_a \ll \mu$ and $\lambda_s \perp \mu$. Then $|\lambda|_a = |\lambda_a|$ and $|\lambda|_s = |\lambda_s|$. Moreover, if $g = (\lambda_a/\mu)$, then $|g| = (|\lambda_a|/\mu)$.

Sketch of proof. Since $\lambda_a \perp \lambda_s$ it is easy to see that $|\lambda_a| \perp |\lambda_s|$ and $|\lambda| = |\lambda_a| + |\lambda_s|$. Now note that $\lambda_a \ll |\lambda_a| \ll \mu$. Let $h = (\lambda_a/|\lambda_a|)$ and $k = (|\lambda_a|/\mu)$. Thus $g = hk$ by 4.1.5. Using 1.2.9 and 4.1.4 we get $k \ge 0$ and $|h| = 1$, respectively. Hence $k = |g|$. □

4.2. Integration with respect to a complex measure

4.2.1 Integration with respect to complex measures. Suppose $(X, \mathcal{M}, \lambda)$ is a measure space with a complex measure λ. Let h be the Radon-Nikodym derivative of λ with respect to $|\lambda|$ and recall that $|h| = 1$. Now we define

$$\int_\lambda g = \int_{|\lambda|} gh$$

whenever g is integrable with respect to $|\lambda|$.

4.2.2 Integration and the Radon-Nikodym derivative. We have now the following extension of 1.1.12. Let $(X, \mathcal{M}, \mu)$ be a measure space with a σ-finite positive measure μ. Suppose ϕ is either a complex measure or else a σ-finite positive measure defined on $\mathcal{M}$ which is absolutely continuous with respect to μ and let k be the corresponding Radon-Nikodym derivative. Then $g \in L^1(|\phi|)$ if and only if $kg \in L^1(\mu)$. In this case

$$\int_\phi g = \int_\mu gk.$$

SKETCH OF PROOF. When ϕ is a positive measure, show this, in turn, for g being a characteristic function, a simple function, a positive function, and an integrable function. Otherwise take also 4.2.1, 4.1.5, and 4.1.6 into account. □

CHAPTER 5

Radon Functionals on Locally Compact Hausdorff Spaces

Throughout this chapter X denotes a locally compact Hausdorff space.

5.1. Preliminaries

5.1.1 Locally compact Hausdorff spaces. Recall that a topological space is called *locally compact* if every point has an open neighborhood with compact closure. A topological space is called a *Hausdorff space* if any two distinct points have disjoint neighborhoods.

Note that $\mathbb{R}^n$ as well as its open and closed subsets are locally compact Hausdorff spaces for any $n \in \mathbb{N}$.

5.1.2 Compactly supported continuous functions. Let $f : X \to \mathbb{C}$ be a function. The set $\operatorname{supp} f = \overline{\{x : f(x) \neq 0\}}$ is called the *support* of a function f. The set of compactly supported continuous functions defined on a topological space X is denoted by $C_c^0(X)$. $C_c^0(X)$ is a normed vector space under the norm $f \mapsto \|f\|_\infty = \sup\{|f(x)| : x \in X\}$[1].

5.1.3 Urysohn's lemma. The notation $K \prec f$ indicates that K is compact in X, that $f \in C_c^0(X)$, $0 \leq f \leq 1$, and that $f(x) = 1$ for all $x \in K$. The notation $f \prec V$ indicates that V is open in X, that $f \in C_c^0(X)$, $0 \leq f \leq 1$, and that the support of f is in V.

The following theorem is well-known from topology.

Theorem. Suppose X is a locally compact Hausdorff space, K compact, V open, and $K \subset V \subset X$. Then there exists $f \in C_c^0(X)$ such that $K \prec f \prec V$.

5.1.4 Partitions of unity. The following theorem, whose proof depends on Urysohn's lemma, is well-known from topology. If U_1, ..., U_n are open subsets of X and if the compact set K is contained in $\bigcup_{k=1}^n U_k$, then there are functions h_k, $k = 1, ..., n$, such that $h_k \prec U_k$ and $K \prec \sum_{k=1}^n h_k$. The collection $\{h_1, ..., h_n\}$ is called a *partition of unity* on K with respect to the cover $\{U_1, ..., U_n\}$.

5.2. Approximation by continuous functions

5.2.1 Compactly supported continuous functions in $L^p(\mu)$. Suppose X is a locally compact Hausdorff space and μ a regular, positive Borel measure on X. If $1 \leq p < \infty$ then $C_c^0(X)$ is dense in $L^p(\mu)$.

Sketch of proof. Suppose $f \in L^p(\mu)$ and $\varepsilon > 0$ is given. According to 1.4.10 there is a simple function $s = \sum_{j=1}^n \alpha_j \chi_{A_j}$ with $\mu(A_j) < \infty$ such that $\|f - s\|_p < \varepsilon/2$. Set $M = \max\{|\alpha_1|, ..., |\alpha_n|\}$. For each A_j there is then a continuous function g_j such that $\mu(\{x : \chi_{A_j}(x) \neq g_j(x)\}) < \frac{1}{n}(\frac{\varepsilon}{2nM})^p$. For $g = \sum_{j=1}^n \alpha_j g_j$ we get $\|s - g\|_p < \varepsilon/2$. □

[1]This notation must not be confused with the one in 1.4.5 for functions which are essentially bounded with respect to a positive measure.

5.2.2 Functions vanishing at infinity. A complex-valued function f on a locally compact Hausdorff space is said to *vanish at infinity* if for every $\varepsilon > 0$ there exists a compact set K such that $|f(x)| < \varepsilon$ if $x \notin K$. The class of continuous functions which vanish at infinity is denoted by $C_0^0(X)$.

Of course, $C_c^0(X) \subset C_0^0(X)$. The converse is true if X is compact.

5.2.3 $C_0^0(X)$ is a Banach space. If X is a locally compact Hausdorff space then $C_0^0(X)$ is the completion of $C_c^0(X)$ with respect to the norm $\|\cdot\|_\infty$.

Sketch of proof. To prove denseness consider the function $h = fg$ where $K \prec g$ and K being such that $|f| < \varepsilon$ outside K. To prove completeness show that a Cauchy sequence has pointwise limits giving rise to a continuous function vanishing at infinity. □

5.3. Riesz's representation theorem

5.3.1 Radon functionals. A *Radon functional* on $C_c^0(X)$ is a function $\phi : C_c^0(X) \to \mathbb{C}$ which is linear and has the property that for every compact set $K \subset X$ there is a number C_K such that

$$|\phi(f)| \leq C_K \|f\|_\infty$$

whenever $f \in C_c^0(X)$ and $\operatorname{supp} f \subset K$.

Given two Radon functionals ϕ and ψ and a complex number c we may define $\phi + \psi$ and $c\phi$ by $(\phi + \psi)(f) = \phi(f) + \psi(f)$ and $(c\phi)(f) = c\phi(f)$, respectively. Thus the set of all Radon functionals on $C_c^0(X)$ is a complex vector space.

Given a Radon functional ϕ we define its *conjugate* $\overline{\phi}$ by $\overline{\phi}(f) = \overline{\phi(\overline{f})}$. Note that $\overline{\overline{\phi}} = \phi$. A Radon functional ϕ is called real if $\overline{\phi} = \phi$. This is equivalent with the requirement that $\phi(f)$ is real whenever f assumes only real values. Note, however, that ϕ being real does not mean it is always real-valued.

Define $\operatorname{Re}\phi$ and $\operatorname{Im}\phi$ by $(\operatorname{Re}\phi)(f) = (\phi(f)+\overline{\phi}(f))/2$ and $(\operatorname{Im}\phi)(f) = (\phi(f)-\overline{\phi}(f))/(2i)$, respectively. Then $\phi = \operatorname{Re}\phi + i\operatorname{Im}\phi$.

$\overline{\phi}$, $\operatorname{Re}\phi$, and $\operatorname{Im}\phi$ are Radon functionals when ϕ is one. Also, $\operatorname{Re}\phi$ and $\operatorname{Im}\phi$ are real.

5.3.2 Positive linear functionals. If $\phi : C_c^0(X) \to \mathbb{C}$ is linear and if $\phi(f) \geq 0$ whenever $f \geq 0$, then ϕ is called a *positive linear functional* on $C_c^0(X)$.

Positive linear functionals are *monotone* in the sense that $\phi(f) \leq \phi(g)$ if $f \leq g$. Conversely, if $\phi : C_c^0(X) \to \mathbb{C}$ is linear and monotone, then it is a positive linear functional.

Theorem. Any positive linear functional is a Radon functional.

Sketch of proof. Choose $g \in C_c^0(X)$ such that $K \prec g$. Write $f = \sum_{k=0}^3 i^k f_k$ where $f_0 = (\operatorname{Re} f)_+$ etc. If $\operatorname{supp} f \subset K$, then $f_k \leq \|f\|_\infty g$. Choose $C_K = 4\phi(g)$. □

5.3.3 The representation theorem for positive linear functionals. This major theorem was first proved by F. Riesz in 1909 in the case where $X = [0, 1]$.

Theorem. Suppose ϕ is a positive linear functional on $C_c^0(X)$. Then there exist a unique positive measure μ on $\mathcal{B}(X)$ with the following properties: (1) μ is outer regular, (2) every open set is inner regular, and (3) $\phi(f) = \int_\mu f$ for all $f \in C_c^0(X)$.

Moreover, the measure μ satisfies $\mu(K) = \inf\{\phi(f) : K \prec f\} < \infty$ whenever K is compact and $\mu(U) = \sup\{\phi(f) : f \prec U\}$ whenever U is open.

Sketch of proof. If $K \prec f \prec U$ then, by the monotonicity of μ, we have $\mu(K) \leq \phi(f) \leq \mu(U)$. Since U is inner regular ϕ determines its measure uniquely. Outer regularity establishes now the uniqueness of μ.

We show existence by constructing the measure explicitly. Let $\mathcal{E}$ be the set of open sets in X and define $|\cdot|:\mathcal{E}\to[0,\infty]$ by $|U|=\sup\{\phi(f):f\prec U\}$. By 2.2.2 $|\cdot|$ defines an outer measure μ^* on $\mathcal{P}(X)$. By Carathéodory's construction 2.2.3

$$\mathcal{M}=\{A\subset X:\forall B\subset X:\mu^*(B)=\mu^*(A\cap B)+\mu^*(A^c\cap B)\}$$

is a σ-algebra and $\mu=\mu^*|_{\mathcal{M}}$ is a complete positive measure on it. We show that (i) μ^* is an extension of $|\cdot|$, (ii) that $\mathcal{B}(X)\subset\mathcal{M}$ and that μ is outer regular, (iii) that $\mu(K)=\inf\{\phi(f):K\prec f\}<\infty$ when K is compact and that open sets are inner regular, and (iv) that $\phi(f)=\int_\mu f$ for all $f\in C_c^0(X)$.

For (i) it is sufficient to show that $|U|\le\sum_{k=1}^\infty|U_k|$ if U and the U_k are open and if $U\subset\bigcup_{k=1}^\infty U_k$. This proof uses a partition of unity on $\operatorname{supp} f$ with respect to the cover $\{U_1,...,U_n\}$ whenever $f\prec U$.

To prove (ii) note first that, by (i), there is, for any $B\subset X$ and any $\varepsilon>0$, an open set $V\supset B$ such that $\mu^*(B)+\varepsilon\ge\mu^*(V)$. If U is now any open set, there is an $f\prec U\cap V$ and a $g\prec V\setminus\operatorname{supp} f$ such that $\mu^*(U\cap V)\le\phi(f)+\varepsilon$, $\mu^*(U^c\cap V)\le\phi(g)+\varepsilon$, and $\phi(f+g)\le\mu^*(V)$.

We now turn to (iii) which relies on the monotonicity of ϕ. If $K\prec f$ we have that $g\le cf$ whenever $c>1$ and $g\prec V=\{x:1<cf(x)\}$. Since $K\subset V$, we get $\mu(K)\le c\phi(f)$, and since $c>1$ is arbitrary, we get this inequality also with $c=1$. Choosing f so that $K\prec f\prec U$ with $\mu(U)$ close to $\mu(K)$ proves that $\mu(K)=\inf\{\phi(f):K\prec f\}$. To show that an open set U is inner regular, choose, for any $g\prec U$ and $\varepsilon>0$, an f satisfying $\operatorname{supp} g\prec f$ and $\phi(f)\le\mu(\operatorname{supp} g)+\varepsilon$.

To prove (iv) note first that we need only consider real-valued f and that, due to the linearity of ϕ, it is sufficient to prove $\phi(f)\le\int_\mu f$. Now let $n\in\mathbb{N}$ be given, define $K=\operatorname{supp} f$, and assume $f(K)\subset(-a,a)$ for a suitable $a>0$. Let $\Delta=2a/n$ and $E_j=\{x\in K:-a+(j-1)\Delta<f(x)\le -a+j\Delta\}$ for $j=1,...,n$. Note that the E_j are pairwise disjoint and that their union is K. For each j there is an open set U_j such that $E_j\subset U_j\subset\{x:f(x)<-a+(j+1)\Delta\}$ and $\mu(E_j)\ge\mu(U_j)-1/n^2$. Let $\{h_1,...,h_n\}$ be a partition of unity on K with respect to $\{U_1,...,U_n\}$ so that $\phi(h_j)\le\mu(U_j)$. Using these inequalities and the monotonicity of ϕ this implies

$$\phi(f)\le\sum_{j=1}^n(-a+(j+1)\Delta)\phi(h_j)\le -a\sum_{j=1}^n\phi(h_j)+\sum_{j=1}^n\int_\mu(f+a+2\Delta)\chi_{E_j}+\frac{(n+1)\Delta}{n}.$$

Since $\mu(K)\le\sum_{j=1}^n\phi(h_j)$ one obtains $\phi(f)\le\int_\mu f+2\Delta(\mu(K)+1)$. □

5.3.4 Total variation of a Radon functional. Suppose $\phi:C_c^0(X)\to\mathbb{C}$ is a Radon functional. Then there is a unique positive linear functional $|\phi|$ such that (i) $|\phi(f)|\le|\phi|(|f|)$ for all $f\in C_c^0(X)$ and (ii) $|\phi|(|f|)\le\lambda(|f|)$, if $\lambda:C_c^0(X)\to\mathbb{C}$ is any positive linear functional such that $|\phi(f)|\le\lambda(|f|)$ for all $f\in C_c^0(X)$. $|\phi|$ is called the *total variation functional* of ϕ.

Sketch of proof. For non-negative $f\in C_c^0(X)$ define

$$|\phi|(f)=\sup\{|\phi(g)|:g\in C_c^0(X),|g|\le f\}.$$

Then $0\le|\phi|<\infty$ and it satisfies $|\phi|(cf)=c|\phi|(f)$ whenever $c\in[0,\infty)$. If $|g_1|\le f_1$, $|g_2|\le f_2$, and $|\alpha_1|=|\alpha_2|=1$, then $|\alpha_1g_1+\alpha_2g_2|\le f_1+f_2$, so that we get $|\phi|(f_1)+|\phi|(f_2)\le|\phi|(f_1+f_2)$. To prove the opposite inequality assume that $|g|\le f_1+f_2$ and set $g_j=(f_jg)/(f_1+f_2)$, $j=1,2$ if the denominator does not vanish and otherwise 0. We have now proved that $|\phi|$ is additive on the non-negative functions in $C_c^0(X)$. For complex valued f define

$$|\phi|(f) = |\phi|((\operatorname{Re} f)_+) - |\phi|((\operatorname{Re} f)_-) + i|\phi|((\operatorname{Im} f)_+) - i|\phi|((\operatorname{Im} f)_-).$$

Then $|\phi|$ is a positive linear functional such that $|\phi(f)| \leq |\phi|(|f|)$. The minimality property follows since λ, as a positive functional, is monotone, i.e., $\lambda(|g|) \leq \lambda(|f|)$ if $|g| \leq |f|$. □

5.3.5 Regularity for spaces with σ-finite measures. Suppose μ is a positive, σ-finite Borel measure on X, which is outer regular, inner regular on open sets, and finite on compact sets. Then μ is also inner regular.

Sketch of proof. Suppose $X = \bigcup_{n=1}^\infty X_n$ where $\mu(X_n) < \infty$ and that E is any measurable set. Then there are open sets $U_n \supset E \cap X_n$ such that $\mu(U_n \setminus (E \cap X_n)) < \varepsilon 2^{-n}$. Hence $E \subset U = \bigcup_{n=1}^\infty U_n$ and $\mu(U \setminus E) < \varepsilon$. The set $U \setminus E$, in turn, is contained in an open set V whose measure is still less than ϵ.

If $\mu(E) = \infty$ and M is a positive number, there is a compact set $F \subset U$ such that $\mu(F) > M$. Let $K = F \setminus V$ and note that K is a compact subset of E. We also have, that $\mu(K) \geq M - \varepsilon$ and hence that E is inner regular.

If $\mu(E) < \infty$ the proof is similar. □

5.3.6 Regularity for σ-compact spaces. In addition to our standard assumption that X is a locally compact Hausdorff space suppose now also that X is *σ-compact*, i.e., that X is a countable union of compact sets. If μ is a positive, outer regular Borel measure which is finite on compact sets, then μ is σ-finite and inner regular.

Sketch of proof. It is clear that μ is σ-finite. Inner regularity may be proved by imitating 2.3.5. □

5.3.7 The representation theorem for general Radon functionals.

Theorem. Suppose X is a σ-compact and locally compact Hausdorff space. If ϕ is a Radon functional on $C_c^0(X)$, then there exists a unique positive regular measure μ on $\mathcal{B}(X)$ and a measurable function h of absolute value 1 such that

$$\phi(f) = \int_\mu fh$$

whenever $f \in C_c^0(X)$. Additionally, μ is finite on compact sets.

Sketch of proof. Set $\psi_0 = (|\operatorname{Re}\phi| + \operatorname{Re}\phi)/2$, $\psi_2 = (|\operatorname{Re}\phi| - \operatorname{Re}\phi)/2$, $\psi_1 = (|\operatorname{Im}\phi| + \operatorname{Im}\phi)/2$, and $\psi_3 = (|\operatorname{Im}\phi| - \operatorname{Im}\phi)/2$. Then the ψ_k are positive Radon functionals and $\phi = \sum_{k=0}^3 i^k \psi_k$. By 5.3.3 and 5.3.6 each of the ψ_k is associated with a unique positive regular and σ-finite measure μ_k. These measures are absolutely continuous with respect to their sum $\tilde\mu$ which is also regular. Denote the Radon-Nikodym derivatives of μ_k with respect to $\tilde\mu$ by g_k. By 1.2.9 we have $0 \leq g_k \leq 1$. With $g = \sum_{k=0}^3 i^k g_k$ we obtain

$$\phi(f) = \sum_{k=0}^3 i^k \psi_k(f) = \sum_{k=0}^3 i^k \int_{\mu_k} f = \sum_{k=0}^3 i^k \int_{\tilde\mu} f g_k = \int_{\tilde\mu} fg.$$

Now define $h(x) = g(x)/|g(x)|$ when $g(x) \neq 0$ and $h(x) = 1$ otherwise. Also define μ by $\mu(E) = \int_{\tilde\mu} |g|\chi_E$. Then μ is again σ-finite and regular and

$$\phi(f) = \int_{\tilde\mu} fh|g| = \int_\mu fh.$$

□

5.3.8 The dual of $C_0^0(X)$. If ϕ is bounded linear functional on $C_0^0(X)$ then there is a unique regular complex measure μ on $\mathcal{B}(X)$ such that $\phi(f) = \int_\mu f$. Thus, the dual of $C_0^0(X)$ is the space of all regular complex measures on $\mathcal{B}(X)$.

Sketch of proof. First suppose that ϕ is positive, i.e., $\phi(f) \geq 0$ whenever $f \geq 0$. The restriction of ϕ to $C_c^0(X)$ is then a positive Radon functional on $C_c^0(X)$. Theorem 5.3.3 gives the existence of a positive Borel measure μ such that $\phi(f) = \int_\mu f$ for all $f \in C_c^0(X)$. Note that $\mu(X) = \sup\{\phi(f) : f \prec X\} \leq C\|f\|_\infty \leq C$, i.e., μ is a finite measure. The boundedness of ϕ and the dominated convergence theorem show then that $\phi(f) = \int_\mu f$ for all $f \in C_0^0(X)$.

Inner regularity follows from 5.3.5.

In general, if ϕ is bounded, then so are the positive functionals $(|\operatorname{Re}\phi| \pm \operatorname{Re}\phi)/2$ and $(|\operatorname{Im}\phi| \pm \operatorname{Im}\phi)/2$. □

5.4. Exercises

5.1 (Lusin's theorem). Suppose X is a locally compact Hausdorff space and μ is complete, regular, positive Borel measure on X which is finite on compact sets. Assume that $f : X \to \mathbb{C}$ is measurable, that $A \subset X$ is of finite measure, that $f(x) = 0$ if $x \notin A$, and that $\varepsilon > 0$. Then there exists a continuous function $g : X \to \mathbb{C}$ of compact support such that

$$\mu(\{x : f(x) \neq g(x)\}) < \varepsilon$$

and, if f is bounded,

$$\sup\{|g(x)| : x \in X\} \leq \sup\{|f(x)| : x \in X\}.$$

CHAPTER 6

Differentiation

In this chapter we study functions defined on $\mathbb{R}^d$, $\mathbb{R}$, or on compact intervals $[a,b] \subset \mathbb{R}$. Throughout m denotes Lebesgue measure on $\mathbb{R}$ or $\mathbb{R}^d$. The open balls of radius r centered at x are denoted by $B(x,r)$. Unless stated otherwise "almost everywhere" means "almost everywhere with respect to Lebesgue measure".

6.1. Derivatives of measures

6.1.1 A covering lemma. If $\mathcal{C}$ is a non-empty collection of open balls in $\mathbb{R}^d$ and $c < \mathrm{m}(\bigcup_{B\in\mathcal{C}} B)$, then there are pairwise disjoint balls $A_1, ..., A_k \in \mathcal{C}$ such that $3^d \sum_{j=1}^k \mathrm{m}(A_j) > c$.

Sketch of proof. Inner regularity of Lebesgue measure gives a compact set $K \subset \bigcup_{B\in\mathcal{C}} B$ with $\mathrm{m}(K) > c$. K will be covered by finitely many of the balls in $\mathcal{C}$. One of those with maximal radius is A_1. Among the balls disjoint from A_1 there is again one with maximal radius, A_2. After the k-th step of this process no balls disjoint from the chosen ones are left and it comes to an end. Enlarging the radii of the chosen balls by a factor of 3 gives balls which still cover K. □

6.1.2 Hardy-Littlewood's maximal function. If μ is a complex Borel measure on $\mathbb{R}^d$ define the *maximal function*

$$M_\mu(x) = \sup\{|\mu|(B(x,r))/\mathrm{m}(B(x,r)) : r > 0\}.$$

Then M_μ is Borel measurable and

$$\mathrm{m}(\{x : M_\mu(x) > \alpha\}) \le \frac{3^d}{\alpha}|\mu|(\mathbb{R}^d)$$

whenever $\alpha > 0$.

Sketch of proof. We want to show that the set $V = \{x : M_\mu(x) > a\}$ is open. Hence suppose $x_0 \in V$ and note that then $|\mu|(B(x_0,r))/\mathrm{m}(B(x_0,r)) = b > a$ for some $r > 0$. There is a $\delta > 0$ such that $(r+\delta)^d < br^d/a$. Since $B(x_0,r) \subset B(x,r+\delta)$ if $|x - x_0| < \delta$ it follows that $B(x_0,\delta) \subset V$.

For the second statement choose r_x such that $|\mu|(B(x,r_x))/\mathrm{m}(B(x,r_x)) > \alpha$ for all $x \in \{x : M_\mu(x) > \alpha\}$. Now use the covering lemma 6.1.1. □

Recall from 1.1.12 that, when $0 \le f \in L^1(\mathbb{R}^d)$, then $A \mapsto \mu(A) = \int_A f d\mathrm{m}$ is a finite positive measure on $\mathcal{L}(\mathbb{R}^d)$. We denote the associated maximal function by M_f.

6.1.3 Lebesgue points. If $f : \mathbb{R}^d \to \mathbb{C}$ is integrable, then $x \in \mathbb{R}^d$ is called a Lebesgue point of f if

$$\lim_{r\to 0} \frac{1}{\mathrm{m}(B(x,r))} \int_{\mathrm{m}} |f - f(x)|\chi_{B(x,r)} = 0.$$

If f is continuous at x, then x is a Lebesgue point of f.

If x is a Lebesgue point of f, then

$$f(x) = \lim_{r\to 0} \frac{1}{\mathrm{m}(B(x,r))} \int_{\mathrm{m}} f\chi_{B(x,r)}.$$

Note, however, that the converse is not true.

THEOREM. If $f : \mathbb{R}^d \to \mathbb{C}$ is integrable, then almost every point of $\mathbb{R}^d$ (with respect to m) is a Lebesgue point of f.

SKETCH OF PROOF. Let $\varepsilon > 0$ be given. By 5.2.1 there is, for every $n \in \mathbb{N}$, a $g_n \in C_c^0(\mathbb{R}^d)$ such that $\|f - g_n\|_1 \leq 1/n$. Let $A_n = \{x : |f(x) - g_n(x)| > \varepsilon/3\}$, $B_n = \{x : M_{|f-g_n|}(x) > \varepsilon/3\}$, and $E = \bigcap_{n=1}^{\infty}(A_n \cup B_n)$. Then $\mathrm{m}(E) = 0$ and we may prove the claim when $x \in E^c$. □

6.1.4 Nicely shrinking sets. Let $x \in \mathbb{R}^d$. We say that Borel sets $E_j \subset \mathbb{R}^d$ *shrink nicely* to x if there exists an $\alpha > 0$ and a sequence of balls $B(x, r_j)$ with $r_j \to 0$ such that, for all j, $E_j \subset B(x, r_j)$ and $\mathrm{m}(E_j) \geq \alpha \mathrm{m}(B(x, r_j))$.

6.1.5 Derivatives of measures. Let μ be a complex Borel measure on $\mathbb{R}^d$ with Lebesgue decomposition $\mu(E) = \int_{\mathrm{m}} f\chi_E + \mu_s(E)$ and assume that $|\mu_s|$ is regular (here f is the Radon-Nikodym derivative of μ with respect to m). Then there is a set $M \subset \mathbb{R}^d$ whose complement has measure zero with the following property. If $x \in M$ and $j \mapsto E_j(x)$ a sequence of sets shrinking nicely to x, then

$$\lim_{j\to\infty} \frac{\mu(E_j(x))}{\mathrm{m}(E_j(x))} = f(x).$$

In particular, if $\mu \perp \mathrm{m}$ and $x \in M$ then $\lim_{j\to\infty} \mu(E_j(x))/\mathrm{m}(E_j(x)) = 0$.

SKETCH OF PROOF. Suppose μ is positive and $\mu \perp \mathrm{m}$. Let A be such that $\mu(A) = \mathrm{m}(A^c) = 0$ and

$$F_k = \bigcap_{r>0} \bigcup_{0<s<r} \{x \in A : \frac{\mu(B(x,s))}{\mathrm{m}(B(x,s))} > \frac{1}{k}\}.$$

Since μ is regular there is an open set U such that $A \subset U$ and $\mu(U) < \varepsilon$ for any $\varepsilon > 0$. Use the covering lemma 6.1.1 to show that $\mathrm{m}(F_k) = 0$. Now prove the claim for $x \in A\backslash \bigcup_{k=1}^{\infty} F_k$ assuming the E_j are balls.

We are done when μ is positive, $\mu \perp \mathrm{m}$, and $E_j(x)$ is a sequence of balls. The general case follows from this, the definition of nicely shrinking sets, the fact that $|\mu_s(E)| \leq |\mu_s|(E)$, and 6.1.3. □

6.2. Exercises

6.1. Give an example of a sequence of sets shrinking nicely and one which does not.

CHAPTER 7

Functions of Bounded Variation and Lebesgue-Stieltjes Measures

7.1. Functions of bounded variation

Throughout this section $I \subset \mathbb{R}$ denotes an interval of positive or infinite length.

7.1.1 Variation. Let f be a complex-valued function on I. Then we define the *variation* of f over I as

$$\operatorname{Var}_f(I) = \sup \left\{ \sum_{j=1}^{n} |f(x_j) - f(x_{j-1})| : \text{all } x_j \in I \text{ and } x_0 < x_1 < ... < x_n \right\}.$$

If $\operatorname{Var}_f(I) < \infty$, we say that f is of *bounded variation* on I. If $\operatorname{Var}_f(K) < \infty$ whenever K is a compact subinterval of I, we say that f is of *locally bounded variation* on I.

7.1.2 Basic properties of functions of bounded variation. Every non-decreasing function f on I is of locally bounded variation since, in this case, $\operatorname{Var}_f([x,y]) = f(y) - f(x)$. If f is of bounded variation on I, then it is bounded there.

Variation is additive in the sense that $\operatorname{Var}_f([x,y]) + \operatorname{Var}_f([y,z]) = \operatorname{Var}_f([x,z])$ whenever $x < y < z$. Analogous formulas hold when x and/or z are removed from the intervals under consideration.

If f is of (locally) bounded variation, then so are $\operatorname{Re}(f)$ and $\operatorname{Im}(f)$.

7.1.3 The vector space of functions of bounded variation. The complex-valued functions defined on an interval I which are of locally bounded variation form a complex vector space which is denoted by $\mathrm{BV}_{\mathrm{loc}}(I)$. In fact, since the product of two functions of locally bounded variation is again of locally bounded variation, $\mathrm{BV}_{\mathrm{loc}}(I)$ is an algebra over $\mathbb{C}$. The set of functions of bounded variation on I is denoted by $\mathrm{BV}(I)$. It is a subalgebra of $\mathrm{BV}_{\mathrm{loc}}(I)$.

7.1.4 Variation functions. Let I be an interval with endpoints a and b where $a < b$. Suppose $f \in \mathrm{BV}_{\mathrm{loc}}(I)$ and c is a fixed point in I. For $x \in I$ let

$$V_f(x) = \begin{cases} \operatorname{Var}_f([c,x]) & \text{if } x > c, \\ 0 & \text{if } x = c, \\ -\operatorname{Var}_f([x,c]) & \text{if } x < c. \end{cases}$$

The function V_f is called a *variation function* for f. If $a \in I$, it is customary to choose $c = a$ so that $V_f(a) = 0$ and $V_f(x) = \operatorname{Var}_f([a,x])$ when $x > a$.

If $a \notin I$, we may still define $V_f(x) = \operatorname{Var}_f((a,x])$ provided f is of bounded variation near a (i.e., f is of bounded variation on (a,d) for some $d \in I$).

Any two variation functions for f differ only by a constant.

7.1.5 Bounded variation and monotonicity. If $f : I \to \mathbb{R}$ is of locally bounded variation, then V_f as well as $V_f \pm f$ are non-decreasing. Therefore any real-valued function of locally bounded variation may be expressed as the difference of two non-decreasing functions.

Complex-valued functions which are of locally bounded variation may be written as a combination of four non-decreasing functions.

Since non-decreasing functions are Borel measurable, it follows that all functions of locally bounded variation are Borel measurable.

7.1.6 Discontinuities. Let f be a complex-valued function on the interval I. If f is not continuous at the point $c \in I$ but $\lim_{x\downarrow c} f(x)$ and $\lim_{x\uparrow c} f(x)$ exist[1], then c is called a *jump discontinuity* or a *discontinuity of the first kind*. Otherwise it is called a *discontinuity of the second kind*.

If f is a complex-valued function on the interval I which has no discontinuities of the second kind, we define $f^\pm : I \to \mathbb{C}$ by setting $f^+(x) = \lim_{t\downarrow x} f(t)$ and $f^-(x) = \lim_{t\uparrow x} f(t)$. If f is non-decreasing we have $f^+(x) = \inf\{f(t) : t > x\}$ and $f^-(x) = \sup\{f(t) : t < x\}$.

7.1.7 Bounded variation and continuity. A function in $\mathrm{BV}_{\mathrm{loc}}(I)$ has at most countably many discontinuities. Each of these is a discontinuity of the first kind. Moreover, if f is of bounded variation on (a, b), then $\lim_{x\downarrow a} f(x)$ and $\lim_{x\uparrow b} f(x)$ exist.

7.1.8 Left-continuous and right-continuous functions. Suppose f is a function of locally bounded variation on the interval I. Then f^+ is right-continuous while f^- is left-continuous.

If $f \in \mathrm{BV}_{\mathrm{loc}}((a, b))$ is left-continuous, then so is its variation function V_f.

We also find that $\lim_{x\downarrow a} \mathrm{Var}_f((a, x]) = 0$ and $\lim_{x\uparrow b} \mathrm{Var}_f((a, x]) = \mathrm{Var}_f((a, b))$, if $f \in \mathrm{BV}((a, b))$.

7.2. Lebesgue-Stieltjes measures

In this section (a, b) is a non-empty open interval in $\mathbb{R}$. We allow $a = -\infty$ and $b = \infty$.

7.2.1 Positive Lebesgue-Stieltjes measures. Suppose $F : (a, b) \to \mathbb{R}$ is a non-decreasing function. Let $\mathcal{E}$ be the set of all open intervals (c, d) such that $a < c \le d < b$ and define $|(c, d)| = F^-(d) - F^+(c)$ when $c < d$ and $|\emptyset| = |(c, c)| = 0$. Then

$$\mu_F^*(A) = \inf\{\sum_{j=1}^\infty |(c_j, d_j)| : (c_j, d_j) \in \mathcal{E}, A \subset \bigcup_{j=1}^\infty (c_j, d_j)\}$$

defines an outer measure which yields, employing Carathéodory's construction 2.2.3, a complete positive measure μ_F on a σ-algebra $\mathcal{M}_F$ in (a, b). μ_F is an extension of $|\cdot| : \mathcal{E} \to [0, \infty]$ and $\mathcal{M}_F$ contains all Borel sets. Moreover, μ_F assumes finite values on compact sets. μ_F is called a *positive Lebesgue-Stieltjes measure* on (a, b). We shall say that μ_F is *generated* by F.

Sketch of proof. To prove that the outer measure μ_F^* is an extension of $(c, d) \mapsto F^-(d) - F^+(c)$ note that, given $\delta > 0$, there exist α and β such that $c < \alpha < \beta < d$, $F(\alpha) - \delta < F^+(c) \le F(\alpha)$, and $F(\beta) \le F^-(d) < F(\beta) + \delta$. Since $\mu_F^*(\{c\}) = F^+(c) - F^-(c) \ge 0$ we get $\mu_F^*([c, d)) \le F^-(d) - F^-(c)$ and $\mu_F^*((c, d]) \le F^+(d) - F^+(c)$. Using this one shows that open rays and hence open intervals are measurable. □

[1] If c is an endpoint of I we consider, of course, only one of these limits.

Note that, if C is a real number, replacing F by $F^- + C$ (or by $F^+ + C$) yields the same measure.

If F and G are non-decreasing functions on (a, b) and $\alpha, \beta \in [0, \infty)$, then $\mu_{\alpha F + \beta G} = \alpha\mu_F + \beta\mu_G$.

7.2.2 Examples. The following basic examples are instructive.

(1) If F is the identity the associate Lebesgue-Stieltjes measure is Lebesgue measure.
(2) If $F = \chi_{(x_0,\infty)}$ the associate Lebesgue-Stieltjes measure is the Dirac measure concentrated on $\{x_0\}$.

7.2.3 Regularity of positive Lebesgue-Stieltjes measures. An argument very similar to that in 2.3.5 shows that positive Lebesgue-Stieltjes measures are regular.

7.2.4 Complex Lebesgue-Stieltjes measures. If F is of bounded variation, then there is a complex Borel measure μ_F on (a, b) such that $\mu_F([c, d)) = F^-(d) - F^-(c)$ for all intervals $[c, d)$ such that $[c, d] \subset (a, b)$. The measure μ_F is called a *complex Lebesgue-Stieltjes measure*. Again we say that μ_F is *generated* by F.

If $F, G \in \mathrm{BV}((a, b))$ and $\alpha, \beta \in \mathbb{C}$, then $\mu_{\alpha F+\beta G} = \alpha\mu_F + \beta\mu_G$.

7.2.5 Cumulative distribution functions. Suppose μ is a Borel measure, which is finite on compact sets. If c is a fixed point in (a, b), we call the function defined by

$$F_\mu(x) = \begin{cases} \mu([c, x)) & \text{if } x > c, \\ 0 & \text{if } x = c, \\ -\mu([x, c)) & \text{if } x < c \end{cases}$$

a *cumulative distribution function* (cdf) or *distribution function* for short. This function is of locally bounded variation.

If μ is a finite Borel measure (in particular when it is a probability measure) it is customary to define

$$F_\mu(x) = \mu((a, x)).$$

In this case F_μ is of bounded variation.

In either case F_μ is left-continuous. Moreover, if μ is a positive measure, then F_μ is non-decreasing.

7.2.6 Cumulative distribution functions and Lebesgue-Stieltjes measures. Suppose the function $F : (a, b) \to \mathbb{C}$ is non-decreasing or of bounded variation and μ_F is the associated Lebesgue-Stieltjes measure. If c is a fixed point in (a, b), the cumulative distribution function F_{μ_F} of μ_F satisfies $F_{\mu_F}(x) = F^-(x) - F^-(c)$ for all $x \in (a, b)$.

On the other hand, suppose μ is a Borel measure on (a, b), which is finite on compact sets, and F is (one of) its cumulative distribution function(s). Then the Lebesgue-Stieltjes measure μ_F generated by F satisfies $\mu_F(E) = \mu(E)$ for all $E \in \mathcal{B}((a, b))$.

Sketch of proof. Assume μ is positive. Clearly $\mu_F(I) = \mu(I)$ for all intervals I with compact closure in (a, b) and hence for all open sets. If $E \in \mathcal{B}((a, b))$ we have $\mu(E) \le \mu(U) = \mu_F(U)$ for any open set U which contains E. Hence $\mu(E) \le \inf\{\mu(U) : U \text{ open and } E \subset U\} = \mu_F(E)$. Also

$$\mu_F(E) \le \mu_F(U) = \mu(U) = \mu(E) + \mu(U \setminus E) \le \mu(E) + \mu_F(U \setminus E) \le \mu(E) + \varepsilon$$

for a suitable open set U which contains E on account of the outer regularity of μ_F. The general case follows. $\square$

In summary we see that there is a one-to-one correspondence between positive (or complex) Lebesgue-Stieltjes measures and non-decreasing left-continuous functions (or left-continuous functions of bounded variation) which vanish at a given point $c \in (a, b)$.

7.2.7 Total variation of Lebesgue-Stieltjes measures. Suppose F is a left-continuous function of bounded variation. Then the total variation of the complex Lebesgue-Stieltjes measure μ_F is given by $|\mu_F| = \mu_{V_F}$.

Sketch of proof. Note that V_F is a left-continuous, non-decreasing, and bounded function such that $\lim_{x\downarrow a} V_F(x) = 0$. Now let $G(x) = |\mu_F|((a, x))$. Then $V_F \le G$. Next prove that $|\mu_F(E)| \le \mu_{V_F}(E)$ in turn for intervals of the type $[c, d)$, for open intervals, for open sets and finally for measurable sets. This implies that $\mu_G = |\mu_F| \le \mu_{V_F}$ from which we obtain $G \le V_F$. □

7.2.8 Notation for Lebesgue-Stieltjes integrals. If μ_F is a Lebesgue-Stieltjes measure on (a, b) generated by the function $F \in \mathrm{BV}_{\mathrm{loc}}((a, b))$ and $g \in L^1(|\mu_F|)$ it is customary to write

$$\int_{\mu_F} g\chi_E = \int_E g dF.$$

In particular, thinking of x as the identity function which generates Lebesgue measure, an integral with respect to Lebesgue measure may be written as $\int g dx$.

7.2.9 Integration by parts. Suppose F and G are in $\mathrm{BV}_{\mathrm{loc}}((a, b))$. Then the following integration by parts formulas hold whenever $[c, d] \subset (a, b)$.

$$\int \chi_{[c,d]} F^+ dG + \int \chi_{[c,d]} G^- dF = (FG)^+(d) - (FG)^-(c)$$

and

$$\int \chi_{(c,d)} F^+ dG + \int \chi_{(c,d)} G^- dF = (FG)^-(d) - (FG)^+(c).$$

Sketch of proof. First suppose that F and G are non-decreasing and note that $\mu_G \otimes \mu_F$ is a product measure on $[c, d] \times [c, d]$. Let $Q = \{(t, u) \in [c, d] \times [c, d] : t \ge u\}$ so that $Q^u = [u, d]$ and $Q_t = [c, t]$. Then the baby version of Fubini's theorem 3.2.2 gives

$$\int (G^+(d) - G^-(u)) dF(u) = \int \mu_G([u, d]) dF(u)$$
$$= \int \mu_F([c, t]) dG(t) = \int (F^+(t) - F^-(c)) dG(t),$$

i.e., the first formula. The complex case follows by an obvious computation. The second formula (or other similar ones) are proved analogously. □

7.2.10 Consequences of the Lebesgue-Radon-Nikodym theorem for bounded variation functions. Suppose F is a left-continuous function of bounded variation. Then the following statements are true.

(1) F is almost everywhere differentiable and $F' \in L^1((a, b))$ is the Radon-Nikodym derivative of the absolutely continuous part of μ_F.
(2) $\mu_F \perp \mathrm{m}$ if and only if $F' = 0$ almost everywhere.
(3) $\mu_F \ll \mathrm{m}$ if and only if $F(d) - F(c) = \int_{\mathrm{m}} F' \chi_{[c,d)}$ whenever $c < d$.

Here "almost everywhere" is with respect to Lebesgue measure.

Sketch of proof. Let $\mu_F(E) = \int_{\mathrm{m}} f\chi_E + \nu(E)$ be the Lebesgue decomposition of μ_F, in particular $\nu \perp \mathrm{m}$. As a Lebesgue-Stieltjes measure $|\nu|$ is regular. By 4.1.1 $f \in L^1((a,b))$ and by 6.1.5, using $E_j(x) = [x, x+h_j)$ or $[x-h_j, x)$, $F' = f$ almost everywhere (regardless of the choice of the sequence h_j), proving the first and second claim. For the third recall that $\mu_F([c,d)) = F(d)-F(c)$ which implies immediately the "only if" part of the statement. To prove the "if" part define $\lambda(E) = \int_{\mathrm{m}} |F'|\chi_E$ and note that $\lambda \ll \mathrm{m}$. Thus 2.6.3 shows that there is, for every $\varepsilon > 0$, a $\delta > 0$ such that $\mathrm{m}(U) < \delta$ implies $\lambda(U) < \varepsilon$. Note that $\mu_F(E) = \int_{\mathrm{m}} F'\chi_E$ holds for intervals of the type $[c,d)$ and hence for open intervals and open sets. Now assume $\mathrm{m}(E) = 0$ and choose an open set U such that $E \subset U$, $\mathrm{m}(U) < \delta$, and $|\mu_F|(U \setminus E) < \varepsilon$. Then $|\mu_F(E)| \le |\mu_F(U)| + |\mu_F|(U \setminus E) < 2\varepsilon$. □

7.2.11 Lebesgue decomposition of Lebesgue-Stieltjes measures. Suppose μ is a Lebesgue-Stieltjes measure. Then there are unique Lebesgue-Stieltjes measures μ_{ac}, μ_{sc}, and μ_d such that

$$\mu = \mu_{ac} + \mu_{sc} + \mu_d$$

where μ_{ac} is continuous and $\mu_{ac} \ll \mathrm{m}$, μ_{sc} is continuous and $\mu_{sc} \perp \mathrm{m}$, while μ_d is discrete and $\mu_d \perp \mathrm{m}$.

7.3. Absolutely continuous functions

$I \subset \mathbb{R}$ denotes an interval of positive length in this section.

7.3.1 Absolutely continuous functions. A complex-valued function defined on an interval I is called *absolutely continuous* on I if, for every positive ε, there is a positive δ such that $\sum_{j=1}^n |f(y_j) - f(x_j)| < \varepsilon$ whenever $\{(x_j, y_j) : 1 \le j \le n\}$ is a collection of pairwise disjoint intervals in I such that $\sum_{j=1}^n (y_j - x_j) < \delta$. If f is absolutely continuous on every compact subinterval of I it is called *locally absolutely continuous* on I.

The set of locally absolutely continuous functions on I is a complex vector space denoted by $\mathrm{AC}_{\mathrm{loc}}(I)$. $\mathrm{AC}(I)$, the space of absolutely continuous functions on I, is a subspace of $\mathrm{AC}_{\mathrm{loc}}(I)$.

7.3.2 Basic properties of absolutely continuous functions. If $f : I \to \mathbb{C}$ is absolutely continuous, then it is uniformly continuous and of locally bounded variation on I. In particular, absolutely continuous functions are differentiable almost everywhere.

The product of two locally absolutely continuous functions is again locally absolutely continuous. Thus $\mathrm{AC}_{\mathrm{loc}}(I)$ is an algebra over $\mathbb{C}$.

If f is absolutely continuous, then so are $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$.

7.3.3 Compositions of absolutely continuous functions. A function $h : (A,B) \to \mathbb{R}$ is called *Lipschitz continuous*, if there is a positive number C such that $|h(x)-h(y)| \le C|x-y|$ whenever $x, y \in (A,B)$. Any Lipschitz continuous function is absolutely continuous.

Suppose $g : (\alpha,\beta) \to (A,B)$ is absolutely continuous, $f : (a,b) \to (\alpha,\beta)$ is absolutely continuous and strictly increasing, and $h : (A,B) \to \mathbb{R}$ is Lipschitz continuous. Then $g \circ f : (a,b) \to (A,B)$ and $h \circ g : (\alpha,\beta) \to \mathbb{R}$ are also absolutely continuous.

7.3.4 The variation function of an absolutely continuous function. If $f : I \to \mathbb{C}$ is (locally) absolutely continuous, then so is any of its variation functions.

7.3.5 Absolutely continuous functions and absolutely continuous measures. Suppose F is a left-continuous function of bounded variation. Then F is absolutely continuous if and only if $\mu_F \ll \mathrm{m}$. In particular, $\int g\, dF = \int gF'\, dx$ whenever $g \in L^1(|\mu_F|)$.

Sketch of proof. Recall that $\mu_F \ll \mathrm{m}$ if and only if $|\mu_F| \ll \mathrm{m}$. For the "if" direction of the claim use 2.6.3. For the "only if" direction use that V_F is locally absolutely continuous and the outer regularity of m. □

7.3.6 The fundamental theorem of calculus. Let $F : [a, b] \to \mathbb{C}$ be a measurable function. Then F is absolutely continuous on $[a, b]$ if and only if F is differentiable almost everywhere, F' is integrable, and

$$F(x') - F(x) = \int_{\mathrm{m}} F' \chi_{(x,x')}$$

whenever $a \le x \le x' \le b$.

7.3.7 More on the variation function. Let $I \subset \mathbb{R}$ be an interval and $[x_0, x] \subset I$. Suppose $F : I \to \mathbb{R}$ is left-continuous. Then the following statements hold:

(1) If F is non-decreasing, then $\int_{\mathrm{m}} F' \chi_{(x_0,x)} \le F(x) - F(x_0)$.
(2) If F is of locally bounded variation, then $\int_{\mathrm{m}} |F'| \chi_{(x_0,x)} \le V_F(x) - V_F(x_0)$.
(3) If F is locally absolutely continuous, then $\int_{\mathrm{m}} |F'| \chi_{(x_0,x)} = V_F(x) - V_F(x_0)$. In particular, $V_F' = |F'|$ almost everywhere with respect to Lebesgue measure.

7.4. Singular functions

In this section (a, b) is a non-empty open interval in $\mathbb{R}$. We allow $a = -\infty$ and $b = \infty$.

7.4.1 Jump functions. A *jump function* is a function $F : (a, b) \to \mathbb{C}$ of the form

$$F(x) = c + \begin{cases} \sum_{x_n \ge 0} (g_n \chi_{(x_n,\infty)}(x) + h_n \chi_{[x_n,\infty)}(x)) & \text{if } x \ge 0, \\ -\sum_{x_n < 0} (g_n \chi_{(-\infty,x_n]}(x) + h_n \chi_{(-\infty,x_n)}(x)) & \text{if } x < 0 \end{cases}$$

where $c \in \mathbb{C}$, x_n is a sequence of pairwise distinct numbers in (a, b), and g_n, h_n are sequences of complex numbers such that $\sum_{x_n \in [c,d]} (|g_n| + |h_n|) < \infty$ whenever $a < c < d < b$. Loosely speaking one may say that a jump function is a function which only changes through jump discontinuities. Note that

$$F(x_n) - \lim_{x \uparrow x_n} F(x) = h_n \quad \text{and} \quad \lim_{x \downarrow x_n} F(x) - F(x_n) = g_n.$$

In particular, $h_n = 0$ if and only if F is left-continuous at x_n and $g_n = 0$ if and only if F is right-continuous at x_n. Every jump function is of locally bounded variation.

Theorem. If $f : (a, b) \to \mathbb{C}$ is a function of locally bounded variation, then there is a jump function $F : (a, b) \to \mathbb{C}$ such that $f - F$ is a continuous function of locally bounded variation. Moreover, f is the sum of a left-continuous function and a right-continuous function.

Sketch of proof. Denote the points of discontinuity of f by x_n, $n \in Z$ where $Z \subset \mathbb{Z}$ is an appropriate index set. Moreover, set $h_n = f(x_n) - f^-(x_n)$ and $g_n = f(x_n) - f^+(x_n)$. Define F as in (7.4.1). Then $f - F$ is a continuous function of locally bounded variation. Since F itself is the sum of a left-continuous function and a right-continuous function the same is true for f. □

7.4.2 Bounded variation and differentiability. A function of bounded variation is almost everywhere differentiable.

SKETCH OF PROOF. This was proved in 7.2.10 for left-continuous functions of bounded variation. A slight modification of that proof shows that right-continuous functions of bounded variation are also almost everywhere differentiable. Any function of bounded variation is a sum of a left-continuous one and a right-continuous one. □

7.4.3 Singular functions. Suppose F is a function of bounded variation. Then $F' = 0$ almost everywhere if and only if $\mu_F \perp \mathrm{m}$. A function whose derivative is zero almost everywhere is called a *singular function*. Continuous functions of bounded variation for which $\mu_F \perp \mathrm{m}$ are called *singular continuous function*. Obviously, jump functions are singular.

The most famous example of a singular continuous function is the Cantor function, also called the devil's staircase. It maps $[0, 1]$ onto $[0, 1]$, is non-decreasing and continuous and its derivative is zero almost everywhere. The set where the derivative is not zero is the famous Cantor set. Its image under the Cantor function is the whole interval $[0, 1]$. The graph of the Cantor function adorns the front of these notes.

7.4.4 Jump functions and discrete measures. Suppose F is a left-continuous function and either of bounded variation or else non-decreasing. Then μ_F is discrete if and only if F is a jump function.

7.4.5 Lebesgue decomposition of functions of bounded variation. If $F : (a, b) \to \mathbb{C}$ has finite variation on (a, b) then there are functions $F_{ac}, F_{sc}, F_d : (a, b) \to \mathbb{C}$ such that

$$F = F_{ac} + F_{sc} + F_d,$$

F_{ac} is absolutely continuous, F_{sc} is singular continuous, and F_d is a jump function. F_{ac}, F_{sc}, and F_d are unique except for additive constants. In particular, if F is an absolutely continuous and singular function, then F is constant.

7.5. Exercises

7.1. Determine the σ-algebra and the measure when F is the Heaviside function (which equals 0 on $(-\infty, 0]$ and 1 on $(0, \infty)$).

7.2. Show that $V_f(a, b) = |f(b) - f(a)|$ if f is monotone.

7.3. Show that the characteristic function of $\mathbb{Q}$ is not of bounded variation not even locally.

7.4. Show that the sine function is not of bounded variation but that it is of locally bounded variation.

7.5. Suppose that $f : (a, b) \to (\alpha, \beta)$ and $g : (\alpha, \beta) \to \mathbb{R}$ are absolutely continuous. Show that $f \circ g$ is absolutely continuous if and only if it is of bounded variation.

CHAPTER 8

Additional Topics

8.1. The substitution rule

8.1.1 Images of measures. Suppose $(X, \mathcal{M}, \mu)$ is a measure space, Y is a set, and $T : X \to Y$ is a function. Then $\Omega = \{E \subset Y : T^{-1}(E) \in \mathcal{M}\}$ is a σ-algebra in Y and $\tau : E \mapsto \mu(T^{-1}(E))$ is a measure defined on Ω. Of course, τ is a positive measure if μ is. If Y is a topological space and T is measurable, then $\mathcal{B}(Y) \subset \Omega$.

Theorem. If μ and τ are positive measures as above and if $g : Y \to \mathbb{C}$ (or $g : Y \to [0, \infty]$) is an integrable (or a positive measurable) function with respect to Ω, then $g \circ T$ is a measurable function with respect to $\mathcal{M}$. In this case

$$\int_\tau g = \int_\mu g \circ T.$$

If T is injective, the hypotheses that μ and τ are positive may be dropped.

Sketch of proof. For positive measures show this, in turn, for g being a characteristic function, a simple function, a positive function, and an integrable function.

The injectivity of T implies that the sets $E_j = \{Tx : x \in F_j\}$ are pairwise disjoint if the F_j are and that $T^{-1}(E_j) = F_j$. Given E and pairwise disjoint sets F_j so that $\bigcup_{j=1}^\infty F_j = T^{-1}(E)$ we get now

$$\sum_{j=1}^\infty |\mu(F_j)| = \sum_{j=1}^\infty |\tau(E_j)| \le |\tau|(\bigcup_{j=1}^\infty E_j) \le |\tau|(E)$$

and hence $|\mu|(T^{-1}(E)) \le |\tau|(E)$. The opposite inequality $|\tau|(E) \le |\mu|(T^{-1}(E))$ follows even without assuming that T is injective. Now observe that

$$\int_{|\mu|} (h\chi_E) \circ T = \int_{|\tau|} h\chi_E = \tau(E) = \mu(T^{-1}(E)) = \int_\mu \chi_E \circ T = \int_{|\mu|} k(\chi_E \circ T)$$

where $h = (\tau/|\tau|)$ and $k = (\mu/|\mu|)$. Thus $k = h \circ T$ and $\int_\tau g = \int_\mu g \circ T$. □

8.1.2 The area under the Gaussian bell curve.

$$\int_{\mathrm{m}} \mathrm{e}^{-x^2} = \sqrt{\pi}.$$

Sketch of proof. By Fubini's theorem we are done if we can show that

$$\int_{\mathrm{m}_2} \mathrm{e}^{-x^2-y^2} = \pi.$$

Now let $T : \mathbb{R}^2 \to [0, \infty) : (x, y) \mapsto \pi(x^2 + y^2)$ and use Theorem 8.1.1. □

8.1.3 Transformations of Lebesgue-Stieltjes measures. Suppose (a,b) and (α,β) are real intervals. If $F:(\alpha,\beta)\to\mathbb{C}$ is left-continuous and of bounded variation and $T:(a,b)\to(\alpha,\beta)$ is continuous and bijective (and hence strictly monotone), then $G=F\circ T:(a,b)\to\mathbb{C}$ is also left-continuous and of bounded variation.

Now let μ_F and μ_G denote the Lebesgue-Stieltjes measures associated with F and G. If $g\in L^1(|\mu_F|)$, then $g\circ T\in L^1(|\mu_G|)$ and

$$\int_{\mu_F} g = \pm\int_{\mu_G} g\circ T$$

where one has to choose the positive sign if T is strictly increasing and the negative if it is strictly decreasing.

Sketch of proof. Define $\tau(E)=\mu_G(T^{-1}(E))$ as in 8.1.1 and, for some $c\in(\alpha,\beta)$, set $H(y)=\tau([c,y))$ for $y\geq c$ and $H(y)=-\tau([y,c))$ for $y<c$ as in 7.2.5. Since T is bijective we get $H(y)=\pm(F^{\mp}(y)-F^{\mp}(c))$ which proves that $\tau=\pm\mu_F$. Now apply Theorem 8.1.1 □

8.1.4 The classical substitution rule. Suppose $T:(a,b)\to(\alpha,\beta)$ is absolutely continuous, strictly increasing and surjective. Then

$$\int_{\mathrm{m}} \chi_{(\alpha,\beta)} g = \int_{\mathrm{m}} \chi_{(a,b)} (g\circ T)T'$$

whenever $g\in L^1(\alpha,\beta)$.

Sketch of proof. In 8.1.3 choose F to be the identity on (α,β) so that $G=T$. According to 7.3.5 μ_T is absolutely continuous with respect to m. From 4.2.2 we obtain $\int_{\mathrm{m}}\chi_{(\alpha,\beta)}g=\int_{\mathrm{m}}\chi_{(a,b)}(g\circ T)h$ where h is the Radon-Nikodym derivative of μ_T with respect to m. Now apply 6.1.5. □

8.1.5 The substitution rule (general version). Suppose X and Y are topological spaces, μ and ν are positive measures defined on $\mathcal{B}(X)$ and $\mathcal{B}(Y)$, respectively, and $T:X\to Y$ is a surjective measurable function. We also require the following

(1) μ and ν are σ-finite, in fact there exist $A_n\in\mathcal{B}(X)$ such that $\bigcup_{n=1}^\infty A_n=X$ and $\mu(T^{-1}(T(A_n)))<\infty$.
(2) $\mu(E)=0$ implies $\nu(T(E))=0$ for all $E\in\mathcal{B}(X)$.

Then $(hg)\circ T\in L^1(\mu)$ whenever $g\in L^1(\nu)$. Moreover,

$$\int_\nu g = \int_\mu (hg)\circ T$$

where h is the Radon-Nikodym derivative of ν with respect to the measure $\tau: E\mapsto \mu(T^{-1}(E))$ introduced in 8.1.1.

Sketch of proof. Since T is surjective requirement (2) shows that $\nu\ll\tau$. Requirement (1) implies that τ is σ-finite. Hence, by 4.2.2, $\int_\nu g=\int_\tau gh$ for some measurable function $h:Y\to[0,\infty]$. Thus $\int_\nu g=\int_\tau gh=\int_\mu(gh)\circ T$ by 8.1.1. □

8.1.6 Linear Transformations in $\mathbb{R}^d$. Let T be a linear invertible transformation from $\mathbb{R}^d$ to $\mathbb{R}^d$ and g a Borel measurable function on $\mathbb{R}^d$. Then

$$\int_{\mathrm{m}_d} g = |\det(T)|\int_{\mathrm{m}_d} g\circ T$$

whenever $g : \mathbb{R}^d \to \mathbb{C}$ is integrable. This holds, in particular, for translations and rotations which have $\det(T) = 1$.

Sketch of proof. We begin with some basic facts from Linear Algebra. Denoting the canonical basis of $\mathbb{R}^d$ by e_1, ..., e_d define the transformations T_1, T_2, and T_3 by

(1) $T_1 e_1 = te_1$ where $t \neq 0$ and $T_1 e_k = e_k$ for $k \geq 2$,
(2) $T_2 e_1 = e_1 + e_2$ and $T_2 e_k = e_k$ for $k \geq 2$, and
(3) $T_3 e_k = e_{\pi(k)}$ where π is a permutation of $\{1, ..., d\}$.

It is well-known that every linear transformation T is a composition of these three types and that the determinant of T is the product of the determinants of the factors. We have $\det(T_1) = t$, $\det(T_2) = 1$, and $\det(T_3) = \pm 1$, depending on the parity of π. If Q is a rectangular box of volume v then $T_1(Q)$ and $T_3(Q)$ are again such boxes of volume $|t|v$ and v, respectively, while $T_2(Q)$ is a parallelepiped of volume v. Hence we get in summary, that $\mathrm{m}(T(Q)) = |\det T| \mathrm{m}(Q)$.

Now, if $\mathrm{m}(E) = 0$, then there is countable collection of disjoint open boxes whose total measure is arbitrarily small. It follows that $\mathrm{m}(T(E)) = 0$ and hence that the hypotheses of 8.1.5 are satisfied. In particular, $\mathrm{m} \ll \tau$ where $\tau = \mathrm{m}(T^{-1}(\cdot))$ and we need to determine the Radon-Nikodym derivative of m with respect to τ. Since $\mathrm{m}(E) = 0$ implies also $\mathrm{m}(T^{-1}(E)) = 0$ we have $\tau \ll \mathrm{m}$, too, so that, by 4.1.5, (m/τ) is the reciprocal of (τ/m). We compute the latter with the help of 6.1.5. For $a > 0$ let $C(x, a)$ be the cube $\times_{k=1}^d (x_k, x_k + a)$. The cubes $C(x, a)$ shrink nicely to x as a tends to 0. Since $\tau(C(x, a)) = \mathrm{m}(T^{-1}(C(x, a))) = \det(T^{-1})\mathrm{m}(C(x, a))$ we get $(\tau/\mathrm{m})(x) = \lim_{a \to 0} \det(T^{-1}) = 1/\det(T)$. □

8.1.7 Differentiable Transformations in $\mathbb{R}^d$. Let T be a map from an open set $V \subset \mathbb{R}^d$ to $\mathbb{R}^d$ and x_0 a point in V. If there is a linear transformation $A(x_0) : \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\lim_{x \to x_0} \frac{|T(x) - T(x_0) - A(x_0)(x - x_0)|}{|x - x_0|} = 0$$

then T is called differentiable at x_0 and $T'(x_0) = A(x_0)$ is called the derivative of T at x_0. In particular, if T is a linear transformation, then $A(x_0) = T$ for all $x_0 \in \mathbb{R}^d$.

If T is an injective transformation which is differentiable everywhere on V, then

$$\int_{\mathrm{m}_d} \chi_{T(V)} g = \int_{\mathrm{m}_d} \chi_V (g \circ T) |\det(T')|$$

whenever $g \in L^1(\mathrm{m}_d)$.

We shall not prove this theorem but compare with 8.1.4 in the case $d = 1$ and with 8.1.6 in the case where T is linear and invertible.

8.2. Convolutions

8.2.1 Convolutions. Let f and g be two complex-valued functions on $\mathbb{R}$. We define the *convolution* $f * g$ of f and g by

$$(f * g)(x) = \int_{\mathrm{m}(t)} f(x - t) g(t)$$

whenever the integral exists.

Define F by $F(x, t) = f(x - t)g(t)$. According to 2.1.2 F is Borel measurable, at least if f and g are. If f and g are merely Lebesgue measurable, then, by 2.1.8, there are Borel functions f_0 and g_0 which are equal to f and g, respectively, almost everywhere with respect to m. Let $F_0(x,t) = f_0(x-t)g_0(t)$ and note that $F = F_0$ almost everywhere with respect to m_2. Hence F is Lebesgue measurable.

Now suppose f, g, and h are Lebesgue-measurable. Then $f * g = g * f$ if either of these exists. Similarly $(f * g) * h = f * (g * h)$.

8.2.2 Young's inequality. If $f \in L^1(\mathbb{R})$, $g \in L^p(\mathbb{R})$, $1 \le p \le \infty$, then

$$\int_{\mathrm{m}(t)} |f(x-t)g(t)| < \infty$$

for almost every $x \in \mathbb{R}$, $f * g \in L^p(\mathbb{R})$ and

$$\|f * g\|_p \le \|f\|_1 \|g\|_p.$$

Sketch of proof. Define, as above, $F(x,t) = f(x-t)g(t)$. First assume $p = 1$. Since, by 8.1.4,

$$\int_{\mathrm{m}(t)} \int_{\mathrm{m}(x)} |F(x,t)| = \|g\|_1 \|f\|_1$$

we have $F \in L^1(\mathrm{m} \otimes \mathrm{m})$. An application of Fubini's theorem gives the result for $p = 1$.

If $p > 1$, write $|f(x-t)| = |f(x-t)|^{1/q} |f(x-t)|^{1/p}$ where q is the exponent conjugate to p. Then apply Hölder's inequality to reduce this case to the previous one. □

8.3. Exercises

8.1. Why does the formula in 8.1.4 involve $|T'|$ instead of the expected T' in the classical substitution rule?

APPENDIX A

Background

A.1. Topological and metric spaces

A.1.1 Topological spaces. Let X be a set. A subset τ of the power set $\mathcal{P}(X)$ of X is called a *topology* in X if it has the following three properties: (i) $\emptyset, X \in \tau$; (ii) if $\sigma \subset \tau$, then $\bigcup_{A\in\sigma} A \in \tau$; and (iii) if $A, B \in \tau$, then $A \cap B \in \tau$.

If τ is a topology in X, then (X, τ) (or simply X, if no confusion can arise) is called a *topological space*. The elements of τ are called *open sets*. Their complements are called *closed sets*.

A *neighborhood* of $x \in X$ is an open set containing x.

A.1.2 Base of a topology. A subset β of the power set $\mathcal{P}(X)$ of X is called a base of a topology, if (i) $\bigcup_{B\in\beta} B = X$ and (ii) for each $x \in B_1 \cap B_2$ there is an element $B \in \beta$ such that $x \in B \subset B_1 \cap B_2$ whenever $B_1, B_2 \in \beta$. The set of all unions of elements of β is then a topology. It is the smallest topology containing β.

$\beta \subset \tau$ is a base of a given topology τ, if for every $x \in X$ and every $U \in \tau$ which contains x there is a $V \in \beta$ such that $x \in V \subset U$.

A.1.3 Relative topology. Let (X, τ) be a topological space and Y a subset of X. Let $\tau' = \{U \cap Y : U \in \tau\}$. Then (Y, τ') is a topological space. τ' is called the *relative topology* on Y (inherited from (X, τ)).

A.1.4 Product topology. Let (X, τ) and (Y, σ) be topological spaces. The set $\{U \times V : U \text{ open in } X \text{ and } V \text{ open in } Y\}$ forms the base of a topology ρ in $X \times Y$. The topology ρ is called the *product topology* of τ and σ.

A.1.5 Metric and pseudo-metric spaces. Let X be a set. If the function $d : X \times X \to [0, \infty)$ satisfies

(1) $d(x, x) = 0$ for all $x \in X$,

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

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

it is a called a *pseudo-metric* on X and (X, d) (or simply X, if no confusion can arise) is called a *pseudo-metric space.*

If instead of property (1) we have

(1') $d(x, y) = 0$ if and only if $x = y$,

then d is a called a *metric* on X and (X, d) is called a *metric space.*

Every metric space is a pseudo-metric space and every pseudo-metric space is a topological space whose topology is generated by the base consisting of the *open balls* $B(x, r) = \{y \in X : d(x, y) < r\}$, $x \in X$, $r \ge 0$.

Suppose (X, d) is a pseudo-metric space. If we call x and y related if $d(x, y) = 0$ we obtain an equivalence relation. Denoting the equivalence class of x by $[x]$ one may show that $([x],[y]) \mapsto d(x,y)$ is well defined and determines a metric on the set of equivalence classes.

A.1.6 Sequences in pseudo-metric spaces. Let (X,d) be a pseudo-metric space. A sequence $x : \mathbb{N} \to X$ is said to *converge* to $x_0 \in X$, if for every positive ε there is a number N such that $d(x(n),x_0) < \varepsilon$ whenever $n > N$. The point x_0 is then called a *limit* of the sequence x. Limits in metric spaces are unique. In pseudo-metric spaces, however, they need not be unique.

A sequence $x : \mathbb{N} \to X$ is called a *Cauchy sequence* if for every positive ε there is a number N such that $d(x(n),x(m)) < \varepsilon$ whenever $n,m > N$.

A pseudo-metric space X is called *complete* if every Cauchy sequence in X converges.

A complete pseudo-metric space X containing the pseudo-metric space Y as a dense subset is called the *completion* of X.

A.2. Functional Analysis

A.2.1 Semi-normed, normed, and Banach spaces. Let X be a complex vector space. If the function $n : X \to [0,\infty)$ satisfies
(1) $n(\alpha x) = |\alpha| n(x)$ for all $\alpha \in \mathbb{C}$ and all $x \in X$,
(2) $n(x+y) \le n(x) + n(y)$ for all $x,y \in X$,
then it is called a *semi-norm* on X. Note that these properties imply $n(0) = 0$ and $n(x) \ge 0$ for all $x \in X$.

If a semi-norm also satisfies
(3) $n(x) = 0$ only if $x = 0$
then it is called a *norm* and (X,n) (or simply X, if no confusion can arise) is called a *normed vector space*.

Note that every vector space with a semi-norm n is a pseudo-metric space with the pseudo-metric $d(x,y) = n(x-y)$. If n is even a norm, then (X,d) is a metric space. A complete normed vector space is called a *Banach space*.

Suppose (X,n) is a vector space with a semi-norm n. If we call x and y related if $n(x-y) = 0$ we obtain an equivalence relation. Denoting the equivalence class of x by $[x]$ one may show that $[x] \mapsto n(x)$ is well defined and determines a norm on the set of equivalence classes (which is a vector space upon proper definition of addition and scalar multiplication). Of course, the set of these equivalence classes is also a metric space. We obtain the same metric space, if we first introduce the pseudo-metric space induced by n and then turn it into a metric space as in A.1.5.

A.2.2 Inner product and Hilbert spaces. Let X be a complex vector space. If the function $p : X \times X \to \mathbb{C}$ satisfies
(1) $p(x,x) > 0$ for all $0 \ne x \in X$,
(2) $p(\alpha x + \beta y, z) = \alpha p(x,z) + \beta p(y,z)$ for all $x,y,z \in X$ and all $\alpha,\beta \in \mathbb{C}$,
(3) $p(x,y) = \overline{p(y,x)}$ for all $x,y \in X$,
then it is called an *inner product* on X and (X,p) is called an *inner product space*.

Every inner product space is a normed space with the norm $n(x) = \sqrt{p(x,x)}$ and hence a metric space. A complete inner product space is called a *Hilbert space*.

A.2.3 Linear functionals. A *linear functional* is a complex-valued function ϕ defined on a complex vector space V satisfying $\phi(\alpha f + \beta g) = \alpha\phi(f) + \beta\phi(g)$ for all $f,g \in V$ and all $\alpha,\beta \in \mathbb{C}$. A linear functional ϕ on a normed vector space V is called *bounded*, if there is a constant C such that $|\phi(x)| \leq C\|x\|$ for all $x \in V$. The set of all bounded linear functionals on a normed vector space V is called the *dual* of V.

A.2.4 Riesz' representation theorem for linear functionals in a Hilbert space. This is one of the major facts about Hilbert spaces.

THEOREM. Let H be a complex Hilbert space and L a bounded linear functional on H. Then there exists a unique $y \in H$ such that $Lx = \langle x, y\rangle$ for all $x \in H$. In fact, there is a one-to-one correspondence between the elements of H and the bounded linear functionals on H.

Glossary

left-continuous: A function f is called *left-continuous*, if $f(x) = \lim_{t \uparrow x} f(t)$.

pairwise disjoint: The elements of a collection of sets are called *pairwise disjoint*, if the intersection of any two distinct sets taken from the collection is empty.

right-continuous: A function f is called *right-continuous*, if $f(x) = \lim_{t \downarrow x} f(t)$.

totally ordered set: A set is called *totally ordered* when it is equipped with a binary relation $\prec$ satisfying (i) $a \prec b$ and $b \prec a$ imply $a = b$, (ii) $a \prec b$ and $b \prec c$ imply $a \prec c$, and (iii) either $a \prec b$ or $b \prec a$.

List of special symbols

Index

Bibliography

[1] Christer Bennewitz. Spectral theory in Hilbert space. Lecture Notes, 2008.

[2] Gerald B. Folland. *Real analysis*. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, second edition, 1999. Modern techniques and their applications, A Wiley-Interscience Publication.

[3] Russell A. Gordon. *The integrals of Lebesgue, Denjoy, Perron, and Henstock*, volume 4 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1994.

[4] Ernst Henze. *Einführung in die Maßtheorie*. Bibliographisches Institut, Mannheim, second edition, 1985.

[5] Edwin Hewitt and Karl Stromberg. *Real and abstract analysis. A modern treatment of the theory of functions of a real variable*. Springer-Verlag, New York, 1965.

[6] A. N. Kolmogorov and S. V. Fomīn. *Introductory real analysis*. Dover Publications, Inc., New York, 1975. Translated from the second Russian edition and edited by Richard A. Silverman, Corrected reprinting.

[7] Frigyes Riesz and Béla Sz.-Nagy. *Functional analysis*. Dover Books on Advanced Mathematics. Dover Publications, Inc., New York, 1990. Translated from the second French edition by Leo F. Boron, Reprint of the 1955 original.

[8] Walter Rudin. *Principles of mathematical analysis*. McGraw-Hill Book Co., New York-Auckland-Düsseldorf, third edition, 1976. International Series in Pure and Applied Mathematics.

[9] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.