Brief lecture notes on measure theory

GEORGIOS PSARADAKIS
University of Mannheim

April 10, 2018[1]

[1]last update: December 31, 2020

Contents

Preface

This set of notes grew from the graduate course "*Advanced Analysis*" I taught at the Mathematics Institute of the University of Mannheim (Spring 2017, 2018, 2019 and Fall 2020). The aim of the course was to present the analytic techniques required to prove various functional inequalities. The lectures started with basic knowledge of real analysis (measure and integral) and then it went into some advanced topics in analysis such as L^p spaces, symmetric-decreasing rearrangement of functions, the Fourier transform, distributions and Sobolev spaces. In parallel, the proofs of inequalities such as Riesz's rearrangement inequality, Young's inequality, Hardy-Littlewood-Sobolev inequality and logarithmic Sobolev inequality, were given. The skeleton of the course was the first half of the second edition of *Analysis* by E. H. Lieb and M. Loss. The first chapter of this book includes a very brief introduction to measure theory. For instance, outer/inner regularity of the Lebesgue measure and Lusin's theorem are included in the exercises of the first chapter. Having taught only some basic facts about the Lebesgue integral (sometimes without proofs), students found difficult to study the book itself and were restricted to the classroom notes.

The aim of these notes is to quickly cover possible measure theoretic gaps of the students knowledge so that to appreciate the remarkable book by Lieb and Loss. In order not to deviate from the proving-inequalities-purpose of the course, we present: (i) a special proof of the Euclidean isoperimetric inequality due to H. Hadwiger and D. Ohmann, that essentially uses only the outer/inner regularity of Lebesgue measure, (ii) an elementary version of the coarea formula which together with the isoperimetric inequality leads to the L^1-Sobolev inequality.

I start with a section of material students are expected to know from undergraduate analysis. The main part of sections 2, 4 and 5 is a more explanatory presentation of particular subsections from the first two chapters of *Measure theory and fine properties of functions* by L. C. Evans and R. F. Gariepy. For the presentation and proof of the isoperimetric inequality through the Brunn-Minkowski inequality of section 3, I followed the book *Geometric measure theory* by H. Federer. The coarea formula of section 6 is from *Sobolev spaces* by V. G. Maz'ya. Section 7 is a brief account on convex and Lipschitzian functions. For these notes and throughout the course I have consulted several times the books found at the end of these notes.

I want to thank Veniamin Gvozdik and Paul Nikolaev for spotting several misprints. Paul

has provided me with the nice proof of Lemma 7.2.8. Any further misprints or comments are very welcome at

`psaradakis at outlook dot com`

Chapter 1

Prerequisites

1.1 Subsets of the Euclidean space

1.1.1 Abstract sets and sequences of abstract sets

Let X be a nonempty set and 2^X be the set of all subsets of X.

1. Writing A^c for the *complement* of $A \subseteq X$, that is $A^c := X \setminus A$, there holds

$$\Big(\bigcup_{A \in \mathscr{F}} A\Big)^c = \bigcap_{A \in \mathscr{F}} A^c \quad \text{and} \quad \Big(\bigcap_{A \in \mathscr{F}} A\Big)^c = \bigcup_{A \in \mathscr{F}} A^c,$$

for any family of sets $\mathscr{F} \subseteq 2^X$.

2. A sequence $\{A_k \subseteq X\}_{k \in \mathbb{N}}$ is said to *non-decrease* (to $\bigcup_{k \in \mathbb{N}} A_k$) if $A_k \subseteq A_{k+1}$ for all $k \in \mathbb{N}$, and to *non-increase* (to $\bigcap_{k \in \mathbb{N}} A_k$) if $A_k \supseteq A_{k+1}$ for all $k \in \mathbb{N}$.

3. For a sequence $\{A_k \subseteq X\}_{k \in \mathbb{N}}$ and $j \in \mathbb{N}$, let

$$\overline{M}_j := \bigcup_{k \geq j} A_k \quad \text{and} \quad \underline{M}_j := \bigcap_{k \geq j} A_k.$$

Then $\{\overline{M}_j\}_{j \in \mathbb{N}}$ is non-increasing, $\{\underline{M}_j\}_{j \in \mathbb{N}}$ is non-decreasing and we set

$$\limsup_{k \to \infty} A_k := \bigcap_{j \in \mathbb{N}} \overline{M}_j \quad \text{and} \quad \liminf_{k \to \infty} A_k := \bigcup_{j \in \mathbb{N}} \underline{M}_j.$$

1.1.2 Subsets of the real line

1. For a nonempty subset A of the *extended real line* $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$, let

(i) $\sup A$ to be the *least upper bound* of A,

(ii) $\inf A$ to be the *greatest lower bound* of A.

We have the following "variational" characterizations

(i) $$\mathbb{R} \ni M = \sup A \Leftrightarrow \begin{cases} x \leq M \;\; \forall x \in A, & \text{(upper bound)} \\ \forall\, \varepsilon > 0,\; \exists\, x_\varepsilon \in A \text{ such that } x_\varepsilon > M - \varepsilon. & \text{(least one)} \end{cases}$$

Taking $\varepsilon = 1/k$, $k \in \mathbb{N}$, we obtain a sequence $\{x_k \in A\}$ such that $x_k \to M$.

(ii) $$\mathbb{R} \ni m = \inf A \Leftrightarrow \begin{cases} x \geq m \;\; \forall x \in A, & \text{(lower bound)} \\ \forall\, \varepsilon > 0,\; \exists\, x_\varepsilon \in A \text{ such that } x_\varepsilon < m + \varepsilon. & \text{(greatest one)} \end{cases}$$

Taking $\varepsilon = 1/k$, $k \in \mathbb{N}$, we obtain a sequence $\{x_k \in A\}$ such that $x_k \to m$.

Remark 1.1.1. Initially we will refer to this characterization when we use it. After some point we will use it without any comment.

2. For a sequence $\{a_k \in \mathbb{R}\}_{k\in\mathbb{N}}$ and $j \in \mathbb{N}$, let

$$\overline{m}_j := \sup_{k \geq j} a_k \quad \text{and} \quad \underline{m}_j := \inf_{k \geq j} a_k.$$

Then $\{\overline{m}_j\}_{j\in\mathbb{N}}$ is non-increasing, $\{\underline{m}_j\}_{j\in\mathbb{N}}$ is non-decreasing and we set

$$\limsup_{k\to\infty} a_k := \lim_{j\to\infty} \overline{m}_j \quad \text{and} \quad \liminf_{k\to\infty} a_k := \lim_{j\to\infty} \underline{m}_j.$$

When finite, $\limsup_{k\to\infty} a_k$ and $\liminf_{k\to\infty} a_k$ correspond respectively to the largest and smallest subsequential limit points of $\{a_k\}_{k\in\mathbb{N}}$. Thus,

$$a_k \to a \in \bar{\mathbb{R}} \iff \limsup_{k\to\infty} a_k = \liminf_{k\to\infty} a_k = a.$$

Exercise 1.1.2. Prove $\liminf_{k\to\infty}(-a_k) = -\limsup_{k\to\infty} a_k$.

Exercise 1.1.3. Suppose there exists $K \in \mathbb{N}$ such that $a_k \geq b_k$ for all $k \in \mathbb{N}$ with $k \geq K$. Prove $\liminf_{k\to\infty} a_k \geq \liminf_{k\to\infty} b_k$.

1.1.3 Subsets of the n-dimensional Euclidean space

We assume the student is familiar with elementary metric space properties of $\mathbb{R}^n$. Notions like *Cauchy sequence*, *density* and *separability* will be reviewed when needed in the classroom. For our purposes in these notes, we only need to say that for a point $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ we use $|x|$ to denote its Euclidean distance to the origin; that is,

$$|x| := \sqrt{\sum_{i=1}^{n} x_i^2},$$

and recall the fundamental notion of a limit point of a sequence of points in $\mathbb{R}^n$:

Definition 1.1.4. A point $x \in \mathbb{R}^n$ is called *limit point of the sequence* $\{x_k \in \mathbb{R}^n\}_{k \in \mathbb{N}}$, whenever $|x_k - x| \to 0$ as $k \to \infty$.

1. Diameter, distance between sets and open balls

Definition 1.1.5. (i) The *diameter* of a nonempty $A \subseteq \mathbb{R}^n$ is given by

$$\operatorname{diam}(A) := \sup \{ |x - y| \mid x, y \in A \}.$$

(ii) The *distance* between nonempty $A_1, A_2 \subseteq \mathbb{R}^n$ is given by

$$\operatorname{dist}(A_1, A_2) := \inf \{ |x - y| \mid x \in A_1,\ y \in A_2 \}.$$

(iii) The *open ball* of radius $r > 0$ having centre at $x \in \mathbb{R}^n$ is denoted by $B_r(x)$; that is,

$$B_r(x) := \{ y \in \mathbb{R}^n \mid |y - x| < r \}.$$

2. Interior of a set, open/closed sets, limit point of a set, closure and boundary of a set

Definition 1.1.6. Let A be a nonempty subset of $\mathbb{R}^n$. Then

(i) $x \in A$ is called *interior point* of A if $B_r(x) \subseteq A$ for some $r > 0$,

(ii) the set of all interior points of A is called *interior* of A and is denoted by A°,

(iii) A is called *open* if $A^\circ = A$ and *closed* if its complement is open,[1]

(iv) $x \in \mathbb{R}^n$ is called *limit point of the set* A if it is the limit point of a sequence in A,

[1] $\emptyset$ is open by convention

(v) $x \in A$ is called *isolated point* of A if it is not the limit of any nontrivial sequence in A,

(vi) the union of A and all its limit points is called the *closure* of A and is denoted by $\bar{A}$,

(vii) the set $\bar{A} \setminus A^\circ$ is called the *boundary* of A and is denoted by ∂A.

Exercise 1.1.7. Prove we have $\operatorname{diam}(\bar{C}) = \operatorname{diam}(C)$ for any $C \subset \mathbb{R}^n$.

Theorem 1.1.8. (i) $\bar{B}_r(x) = \{y \in \mathbb{R}^n \mid |y-x| \le r\}$.

(ii) *$A \subset \mathbb{R}^n$ is closed if and only if $A = \bar{A}$.*

(iii) *Given $A \subset \mathbb{R}^n$ we have that $\bar{A}$ is the smallest closed set containing A.*

(iv) *$x \in A$ is an isolated point of A if and only if $B_r(x) \cap A = \{x\}$ for some $r > 0$.*

Theorem 1.1.9. (i) *The union (intersection) of any number of open (closed) sets is open (closed).*

(ii) *The intersection (union) of a finite number of open (closed) sets is open (closed).*

3. Bounded and compact sets

Definition 1.1.10. $E \subset \mathbb{R}^n$ is called *bounded* if $E \subset B_r(0)$ for some $r > 0$.

Exercise 1.1.11. (i) Prove that for any $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, there holds

$$|x_j| \le |x| \le \sum_{i=1}^{n} |x_i|, \quad j = 1, \ldots, n.$$

(ii) Write down a proof of the *triangle inequality* in $\mathbb{R}^n$:

$$\big||x| - |y|\big| \le |x+y| \le |x| + |y| \quad \forall\, x, y \in \mathbb{R}^n.$$

(iii) Prove $B_r(x_0)$ is an open subset of $\mathbb{R}^n$. Prove it is also bounded.

(iv) Prove $E \subset \mathbb{R}^n$ is bounded if and only if for each $i \in \{1, \ldots, n\}$, the i-th coordinates of its points form a bounded subset of $\mathbb{R}$.

Definition 1.1.12. $K \subset \mathbb{R}^n$ is called *compact* if every open cover of K has a finite subcover.

Theorem 1.1.13. *$K \subset \mathbb{R}^n$ is compact if and only if*

(i) *it is closed and bounded,*

(ii) *every sequence of points of K has a subsequence converging in K.*

Exercise 1.1.14. Prove the distance of a compact set to a disjoint closed set is positive.

Definition 1.1.15. We say U_1 *is compactly contained in* U_2 and write this as $U_1 \Subset U_2$, whenever U_1 and U_2 are open subsets of $\mathbb{R}^n$ such that $\bar{U_1} \subset U_2$ and $\bar{U_1}$ is bounded.

Remark 1.1.16. From exercise 1.1.14 we deduce that given a compact subset K of an open set $U \subseteq \mathbb{R}^n$, then $\delta := \text{dist}\{K, U^c\} \in (0, \infty]$. For any $\varepsilon \in (0, \delta)$ we define

$$K_\varepsilon := \{x \in \mathbb{R}^n \mid \text{dist}(x, K) \leq \varepsilon\}.$$

Then K_ε is again a compact subset of U with $K \subset (K_\varepsilon)^\circ$.

As a consequence, given $U_1 \Subset U_2$, by taking $K = \bar{U_1}$ and $U = U_2$ in the above remark, it readily follows that

Proposition 1.1.17. *Given $U_1 \Subset U_2$, there exists V such that $U_1 \Subset V \Subset U_2$.*

4. Connected and convex sets

Definition 1.1.18. A nonempty $A \subsetneq \mathbb{R}^n$ is called *connected* if it cannot be included in the disjoint union of two nonempty open sets, each having nonempty intersection with A.

Remark 1.1.19. If A is open, connectedness of A implies that any two points of A can be connected by a polygonal curve which is contained in A.

Definition 1.1.20. A nonempty $A \subsetneq \mathbb{R}^n$ is called *convex* if the line segment connecting any two of its points is contained in A.

5. Intervals

Definition 1.1.21. For $n \in \mathbb{N}$, an n-dimensional *interval* I is a subset of $\mathbb{R}^n$ of the form

$$\begin{aligned} I &= \{x = (x_1, \ldots, x_n) \mid \alpha_k \leq x_k \leq \beta_k,\ k = 1, \ldots, n\} \\ &\equiv [\alpha_1, \beta_1] \times [\alpha_2, \beta_2] \times \ldots \times [\alpha_n, \beta_n], \end{aligned}$$

where $\alpha_k < \beta_k$ for all $k = 1, \ldots, n$. If the edge lengths $\beta_k - \alpha_k$ are all equal, then I is called a *cube*. For an *open interval/cube* take strict inequalities in the above form. Two intervals are *non-overlapping* whenever their interiors are disjoint.

The *volume of an interval* equals the volume of its interior and is given by

$$\text{Vol}(I) = \prod_{k=1}^{n} (\beta_k - \alpha_k).$$

Theorem 1.1.22. *Every open set in $\mathbb{R}$ can be written as a countable union of disjoint open intervals. Every open set in $\mathbb{R}^n$, $n \in \mathbb{N}$, can be written as a countable union of non-overlapping intervals.*

Exercise 1.1.23. For compact $X \subset \mathbb{R}^n$, prove there exists a non-increasing sequence $\{X_k \subset \mathbb{R}^n\}_{k\in\mathbb{N}}$ with $X = \bigcap_{k\in\mathbb{N}} X_k$ and each X_k consists of a finite union of non-overlapping intervals.

Exercise 1.1.24. Show that an open set in $\mathbb{R}^n$, $n \in \mathbb{N}\setminus\{1\}$, is not necessarily a countable union of disjoint open sets by explaining why the open unit disk in the plane cannot be written as a countable disjoint union of open squares.

Disjointness can be achieved in the following way:

Definition 1.1.25. Let $k \in \mathbb{Z}$. The n-dimensional cubes of the form

$$\Big[\frac{j_1}{2^k}, \frac{j_1+1}{2^k}\Big) \times \dots \times \Big[\frac{j_n}{2^k}, \frac{j_n+1}{2^k}\Big), \quad (j_1, \dots, j_n) \in \mathbb{Z}^n,$$

are called *(right) half-open cubes of side length* 2^{-k}, or simply *dyadic cubes*.

Theorem 1.1.26. *Every open set in $\mathbb{R}^n$, $n \in \mathbb{N}$, can be written as a countable union of disjoint dyadic cubes.*

1.2 Functions

1. Pre-image properties

For any function $f : X \to Y$ and any $A, \{A_k\}_{k\in\mathbb{N}}$ subsets of Y, there holds

$$f^{-1}(A^c) = \big(f^{-1}(A)\big)^c \;\text{ and }\; f^{-1}\Big(\bigcup_{k\in\mathbb{N}} A_k\Big) = \bigcup_{k\in\mathbb{N}} \big(f^{-1}(A_k)\big).$$

2. Continuity

We denote by E an arbitrary nonempty subset of $\mathbb{R}^n$ and by $f : E \to \bar{\mathbb{R}}$ a given function.

For any limit point x_0 of E we consider the punctured balls

$$B_r^*(x_0) := B_r(x_0) \setminus \{x_0\}, \quad r > 0.$$

We set

$$M_r(x_0) := \sup_{B_r^*(x_0)\cap E} f(x) \quad\text{and}\quad m_r(x_0) := \inf_{B_r^*(x_0)\cap E} f(x).$$

As r decreases, $\{M_r(x_0)\}_{r>0}$ is non-increasing, $\{m_r(x_0)\}_{r>0}$ is non-decreasing and we set

$$\limsup_{E\ni x\to x_0} f(x) := \lim_{r\to 0} M_r(x_0) \quad\text{and}\quad \liminf_{E\ni x\to x_0} f(x) := \lim_{r\to 0} m_r(x_0).$$

Definition 1.2.1. Let $x_0 \in E$ be a limit point of E.

(i) f is *continuous at* x_0 if $\forall\ \varepsilon > 0$, $\exists\ \delta > 0$ s.t. $|f(x) - f(x_0)| < \varepsilon$ whenever $|x - x_0| < \delta$, $x \in E$.

(ii) f is *upper semi-continuous at* x_0 if $\limsup_{E \ni x \to x_0} f(x) \leq f(x_0)$.

(iii) f is *lower semi-continuous at* x_0 if $\liminf_{E \ni x \to x_0} f(x) \geq f(x_0)$.

Theorem 1.2.2. *f is continuous at x_0 if and only if $|f(x_0)| < \infty$ and is both upper semi-continuous and lower semi-continuous at x_0.*

Theorem 1.2.3. *f is continuous at a non-isolated point of $x_0 \in E$ if and only if it is sequentially continuous at x_0; that is, for any sequence $\{x_k \in E\}_{k \in \mathbb{N}}$ with $x_k \to x$ we have $f(x_k) \to f(x)$.*

Definition 1.2.4. (i) f is *continuous in E* if f is continuous at any limit point of E which lies in E.

(ii) f is *uniformly continuous in E* if $\forall\ \varepsilon > 0$, $\exists\ \delta > 0$ s.t. $|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \delta$, $x, y \in E$.

Theorem 1.2.5. *Let E be compact and f be continuous in E. Then the following are true:*

(i) *f is bounded in E.*

(ii) *f attains its supremum and infimum in E.*

(iii) *f is uniformly continuous in E.*

Lemma 1.2.6. (Urysohn's lemma - elementary form). *Let $U \subseteq \mathbb{R}^n$ be open and $K \subset U$ be compact. There exists a real valued function $g \in C_c(\mathbb{R}^n)$; that is, g is continuous in $\mathbb{R}^n$ and vanishes outside some compact subset of $\mathbb{R}^n$, such that*

(i) $0 \leq g(x) \leq 1$ *for all* $x \in \mathbb{R}^n$,

(ii) $g(x) = 1$ *for all* $x \in K$,

(iii) $\operatorname{supp}(g) := \overline{\{x \in \mathbb{R}^n \mid g(x) \neq 0\}} \subset U$.

The above lemma is essentially a corollary of proposition 1.1.17. Indeed, the hypotheses $K^\circ \Subset U$ implies the existence of W such that $K^\circ \Subset W \Subset U$. The function defined by

$$g(x) := \frac{\operatorname{dist}(x, W^c)}{\operatorname{dist}(x, K) + \operatorname{dist}(x, W^c)}, \quad x \in \mathbb{R}^n,$$

readily satisfies (i), (ii) while $\operatorname{supp}(g) = \bar{W} \subset U$. The only unclear at the moment requirement is the continuity of g. However, this follows from the fact that the distance to any nonempty closed subset of $\mathbb{R}^n$ is a Lipschitz function (see example 7.2.3-(i) and the remark before).

3. Differentiability

We denote by E an arbitrary nonempty subset of $\mathbb{R}^n$ and by $f : E \to \mathbb{R}^m$ a given function.

Definition 1.2.7. Let $x_0 \in E$ be an interior point of E. f is *differentiable at* x_0 if there exist a matrix $M_{x_0} = [\alpha_{ij}]_{m\times n}$ such that

$$\lim_{\mathbb{R}^n \ni h\to 0} \frac{|f(x_0+h) - f(x_0) - M_{x_0}h|}{|h|} = 0.$$

The matrix M_{x_0} is called the *derivative of* f *at* x_0 and is denoted by $Df(x_0)$. In case $m = 1$ we write $\nabla f(x_0)$ instead of $Df(x_0)$ and call this the *gradient vector of* f *at* x_0.

Remark 1.2.8. Using mollification we can strengthen the continuity of g in lemma 1.2.6 to read $g \in C_c^\infty(\mathbb{R}^n)$; that is, g is infinitely many times differentiable with respect to any variable and has compact support. Simply, for any $\varepsilon \in (0, \mathrm{dist}\{K, U^c\})$, consider the function

$$g(x) = \eta_\varepsilon * \chi_{K_\varepsilon}(x) = \int_{K_\varepsilon} \eta_\varepsilon(x-y)\mathrm{d}y \qquad x \in U_\varepsilon,$$

where

$$K_\varepsilon := \{x \in \mathbb{R}^n \bigm| \mathrm{dist}(x,K) \le \varepsilon\}, \qquad U_\varepsilon := \{x \in \mathbb{R}^n \bigm| \mathrm{dist}(x,U) > \varepsilon\},$$

and $\eta_\varepsilon : \mathbb{R}^n \to [0,\infty)$ is the standard mollifier given by $\eta_\varepsilon(x) := \varepsilon^{-n}\eta(x/\varepsilon)$, $x \in \mathbb{R}^n$, where

$$\eta(x) := \begin{cases} c\exp\{(|x|^2-1)^{-1}\} & |x| < 1, \\ 0 & |x| \ge 1, \end{cases} \tag{1.2.1}$$

with the constant $c > 0$ being such that $\int_{\mathbb{R}^n} \eta(x)\,\mathrm{d}x = 1$.

Theorem 1.2.9. *The gradient vector of a* $C^1(\Omega)$ *scalar function* f *at a point* x_0 *of any non empty level surface* $\mathfrak{S}_\lambda = \{x \in \Omega \bigm| f(x) = \lambda\}$, $\lambda \in \mathbb{R}$, *is perpendicular to the tangent at* x_0 *of any* C^1*-curve that lies on* $\mathfrak{S}_\lambda$ *and goes through* x_0. *In other words,* $\nabla f(x_0)$ *is perpendicular to the surface* $\mathfrak{S}_\lambda$ *at* $x_0 \in \mathfrak{S}_\lambda$.

proof. Suppose $\vec{\gamma}(t) \in \mathfrak{S}_\lambda$ for all $t \in I := (\alpha,\beta) \subseteq \mathbb{R}$ and $t_0 \in I$ is such that $\vec{\gamma}(t_0) = x_0 \in \mathfrak{S}_\lambda$. Then

$$f\big(\vec{\gamma}(t)\big) = \lambda \qquad \text{for all } t \in I.$$

Differentiating with respect to t we readily get by the chain rule that

$$\nabla f\big(\vec{\gamma}(t)\big) \cdot \frac{\mathrm{d}\vec{\gamma}}{\mathrm{d}t}(t) = 0 \qquad \text{for all } t \in I.$$

For $t = t_0$ this gives $\nabla f(x_0) \cdot \frac{\mathrm{d}\vec{\gamma}}{\mathrm{d}t}(t_0) = 0$, which implies the result since $\frac{\mathrm{d}\vec{\gamma}}{\mathrm{d}t}(t_0)$ is the tangent vector to the curve at its point $\vec{\gamma}(t_0)$. ■

4. The Riemann integral

Let I be an interval in $\mathbb{R}^n$ and $f : I \to \mathbb{R}$ be a bounded function. Decomposing I into a finite collection $\mathscr{I}$ of non-overlapping intervals $\{I_k\}_k$, $k = 1,...,N$, we define

$$\|\mathscr{I}\| := \max_{k\in\{1,...,N\}} \operatorname{diam}(I_k).$$

Selecting $x_k \in I_k$, $k = 1,...,N$, we define

$$R_{\mathscr{I}}(f) := \sum_{k=1}^{N} f(x_k)\operatorname{Vol}(I_k).$$

We further define

$$U_{\mathscr{I}}(f) := \sum_{k=1}^{N} \sup_{x\in I_k} f(x)\operatorname{Vol}(I_k) \quad \text{and} \quad L_{\mathscr{I}}(f) := \sum_{k=1}^{N} \inf_{x\in I_k} f(x)\operatorname{Vol}(I_k).$$

We say the *Riemann integral of f on I exists* whenever the limit $\lim_{\|\mathscr{I}\|\to 0} R_{\mathscr{I}}$ exists. If the Riemann integral of f on I exists we write

$$\int_I f = \lim_{\|\mathscr{I}\|\to 0} R_{\mathscr{I}}(f).$$

Theorem 1.2.10. (i) *The Riemann integral of f on I exists if and only if*

$$\inf_{\mathscr{I}} U_{\mathscr{I}}(f) = \sup_{\mathscr{I}} U_{\mathscr{I}}(f).$$

(ii) *If f is continuous in I then the Riemann integral of f on I exists.*

Exercise 1.2.11. Let $f : [0,1] \to \mathbb{R}$ be a continuous function. Prove

$$f(0) = \lim_{n\to\infty} \int_0^1 ne^{-nx} f(x)\mathrm{d}x \quad \text{and} \quad f(1) = \lim_{n\to\infty} \int_0^1 nx^n f(x)\mathrm{d}x.$$

1.3 Volume of unit ball in $\mathbb{R}^n$

The *gamma function* is given by

$$\Gamma(s) := \int_0^\infty e^{-x} x^{s-1}\mathrm{d}x, \quad s > 0.$$

Through the gamma function we define the number

$$\omega_s := \frac{\pi^{s/2}}{\Gamma(1+s/2)}, \quad s \geq 0.$$

If $s \in \mathbb{N}$ then ω_s is precisely the volume of the unit ball of $\mathbb{R}^s$. Also the volume of a ball of radius $r > 0$ is in this case given by

$$\mathrm{Vol}(B_r(x_0)) = \omega_s r^s.$$

Exercise 1.3.1. Prove the following

(i) $\int_0^\infty e^{-t^2} \mathrm{d}t = \sqrt{\pi}/2$.

(ii) $\Gamma(1/2) = \sqrt{\pi}$.

(iii) $\Gamma(s+1) = s\Gamma(s)$ for all $s > 0$.

(iv) For any $s \in \mathbb{N}$ there holds $\Gamma(s) = (s-1)!$ and $\Gamma(s+1/2) = 1 \cdot 3 \cdot \ldots \cdot (2s-1) \cdot \sqrt{\pi}/2^s$.

(v) For all $\alpha > 2$ and $r > 0$, there holds $\int_0^\infty s^{-\alpha/2} e^{-r^2/(4s)} \mathrm{d}s = (2/r)^{\alpha-2}\Gamma(\alpha/2-1)$.

Chapter 2

Measure

Throughout, X stands for a nonempty set.

2.1 Measures

Definition 2.1.1 (measure). A mapping $\mu : 2^X \to [0,\infty]$ is called *measure*[1] provided

(i) $\mu(\emptyset) = 0$, and

(ii) that is, $\mu(A) \leq \sum_{k\in\mathbb{N}} \mu(A_k)$ whenever $A \subseteq \bigcup_{k\in\mathbb{N}} A_k$.

Remark 2.1.2. The two properties readily imply the *monotonicity* of μ:

$$A \subseteq B \implies \mu(A) \leq \mu(B).$$

Exercise 2.1.3. Let μ be a measure on X and suppose $\{A_k \subseteq X\}_{k\in\mathbb{N}}$ satisfies $\sum_{k\in\mathbb{N}} \mu(A_k) < \infty$. Prove that $\mu(\limsup_{k\to\infty} A_k) = 0$.

Example 2.1.4. Important examples of measures are:

(i) *Dirac's delta measure concentrated to a given $x \in X$*; that is the measure on X defined by

$$\delta_x(A) := \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases} \big(= \chi_A(x)\big), \qquad A \subseteq X.$$

Proof. Let $A \subseteq \bigcup_{k\in\mathbb{N}} A_k$. If $x \notin \bigcup_{k\in\mathbb{N}} A_k$ then $\delta_x(A) \leq \delta_x(\bigcup_{k\in\mathbb{N}} A_k) = 0 = \sum_{k\in\mathbb{N}} \delta_x(A_k)$. If $x \in \bigcup_{k\in\mathbb{N}} A_k$ then x is in A_{k_0} for some $k_0 \in \mathbb{N}$, which implies $\delta_x(A) \leq \delta_x(\bigcup_{k\in\mathbb{N}} A_k) = 1 = \delta_x(A_{k_0}) \leq \sum_{k\in\mathbb{N}} \delta_x(A_k)$. ■

[1] *outer measure* in other texts

(ii) The *Lebesgue measure* on $\mathbb{R}^n$ (denoted throughout by $\mathscr{L}^n$), defined for any $A \subseteq \mathbb{R}^n$ by

$$\mathscr{L}^n(A) := \inf\Big\{ \sum_{k\in\mathbb{N}} \mathrm{Vol}(I_k) \;\Big|\; \{I_k\}_{k\in\mathbb{N}} \text{ are intervals and } A \subseteq \bigcup_{k\in\mathbb{N}} I_k \Big\}.$$

Proof. Let $A \subseteq \bigcup_{k\in\mathbb{N}} A_k$ with $\mathscr{L}^n(A_k) < \infty$ for all $k \in \mathbb{N}$ and fix $\varepsilon > 0$. Given $k \in \mathbb{N}$, by definition of $\mathscr{L}^n(A_k)$ and the variational characterization of inf, we can choose intervals $\{I_j^{(k)}\}_{j\in\mathbb{N}}$ such that $A_k \subseteq \bigcup_{j\in\mathbb{N}} I_j^{(k)}$ and

$$\sum_{j\in\mathbb{N}} \mathrm{Vol}(I_j^{(k)}) < \mathscr{L}^n(A_k) + \frac{\varepsilon}{2^k}.$$

Since $A \subseteq \bigcup_{k\in\mathbb{N}} \bigcup_{j\in\mathbb{N}} I_j^{(k)}$, the definition of $\mathscr{L}^n(A)$ implies

$$\mathscr{L}^n(A) \le \sum_{k\in\mathbb{N}} \sum_{j\in\mathbb{N}} \mathrm{Vol}(I_j^{(k)}).$$

Combining the two estimates, $\mathscr{L}^n(A) < \sum_{k\in\mathbb{N}} \mathscr{L}^n(A_k) + \varepsilon$, which implies the subadditivity since $\varepsilon > 0$ is arbitrary. ■

(iii) The *s-dimensional Hausdorff measure* on $\mathbb{R}^n$ (denoted throughout by $\mathscr{H}^s$), defined for any $s \ge 0$ and any $A \subseteq \mathbb{R}^n$ by

$$\mathscr{H}^s(A) := \lim_{\delta\to 0^+} \mathscr{H}^s_\delta(A),$$

where (see §1.3 for the definition of ω_s)

$$\mathscr{H}^s_\delta(A) := \inf\Big\{ \sum_{j\in\mathbb{N}} \omega_s \Big(\frac{\mathrm{diam}\, C_j}{2}\Big)^s \;\Big|\; A \subseteq \bigcup_{j\in\mathbb{N}} C_j,\ \mathrm{diam}(C_j) \le \delta \Big\}.$$

Remark 2.1.5. Observe that for $\delta_1 < \delta_2$, all coverings of A by sets with diameters no more than δ_1 are included in the set of coverings of A by sets with diameters no more than δ_2. Thus $\mathscr{H}^s_\delta$ increases as δ decreases. This justifies writing “lim” in the above definition and also

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

Remark 2.1.6. For $s = 0$ we deduce that $\mathscr{H}^0$ is just the counting measure; that is, $\mathscr{H}^0(A)$ equals the number of elements in A when A is finite, and infinity otherwise. This follows from the fact that $\omega_0 = 1$, hence $\mathscr{H}^0\big(\{a\}\big) = 1$ for any $a \in \mathbb{R}^n$.

Proof. We show first that $\mathscr{H}^s_\delta$ is a measure. Let $A \subseteq \bigcup_{k\in\mathbb{N}} A_k$ with $\mathscr{H}^s_\delta(A_k) < \infty$ for all $k \in \mathbb{N}$ and fix $\varepsilon > 0$. Given $k \in \mathbb{N}$, by definition of $\mathscr{H}^s_\delta(A_k)$ and the variational characterization of inf, we can choose $\{C_j^{(k)}\}_{j\in\mathbb{N}}$ such that $A_k \subseteq \bigcup_{j\in\mathbb{N}} C_j^{(k)}$, $\operatorname{diam}(C_j^{(k)}) \le \delta$ for all $j \in \mathbb{N}$ and

$$\sum_{j\in\mathbb{N}} \omega_s \Big(\frac{\operatorname{diam} C_j^{(k)}}{2}\Big)^s < \mathscr{H}^s_\delta(A_k) + \frac{\varepsilon}{2^k}.$$

Since $A \subseteq \bigcup_{k\in\mathbb{N}} \bigcup_{j\in\mathbb{N}} C_j^{(k)}$ and $\operatorname{diam}(C_j^{(k)}) \le \delta$ for all $k, j \in \mathbb{N}$, the definition of $\mathscr{H}^s_\delta(A)$ implies

$$\mathscr{H}^s_\delta(A) \le \sum_{k\in\mathbb{N}} \sum_{j\in\mathbb{N}} \omega_s \Big(\frac{\operatorname{diam} C_j^{(k)}}{2}\Big)^s.$$

Combining the two estimates, $\mathscr{H}^s_\delta(A) < \sum_{k\in\mathbb{N}} \mathscr{H}^s_\delta(A_k) + \varepsilon$, which implies the subadditivity since $\varepsilon > 0$ is arbitrary. □

We show next that $\mathscr{H}^s$ is a measure. Since $\mathscr{H}^s_\delta$ is a measure, if $A \subseteq \bigcup_{k\in\mathbb{N}} A_k$ then

$$\mathscr{H}^s_\delta(A) \le \sum_{k\in\mathbb{N}} \mathscr{H}^s_\delta(A_k).$$

By remark 2.1.5 we have $\mathscr{H}^s_\delta(A_k) \le \mathscr{H}^s(A_k)$ for all $k \in \mathbb{N}$, thus

$$\mathscr{H}^s_\delta(A) \le \sum_{k\in\mathbb{N}} \mathscr{H}^s(A_k),$$

and the subadditivity follows by taking the limit $\delta \to 0$. ■

(iv) The *restriction* $\mu|_A$ of a measure μ on X to an arbitrary $A \subset X$; that is,

$$\mu|_A(B) := \mu(B \cap A), \qquad B \subseteq X.$$

Proof. Let $B \subseteq \bigcup_{k\in\mathbb{N}} B_k$ with $\mu|_A(B_k) < \infty$ for all $k \in \mathbb{N}$ and note

$$\Big(\bigcup_{k\in\mathbb{N}} B_k\Big) \cap A = \bigcup_{k\in\mathbb{N}} (B_k \cap A). \tag{2.1.1}$$

Using the monotonicity of μ, then (2.1.1) and finally the subadditivity of μ,

$$\mu|_A(B) \le \mu\Big(\big(\bigcup_{k\in\mathbb{N}} B_k\big) \cap A\Big) \le \sum_{k\in\mathbb{N}} \mu(B_k \cap A) = \sum_{k\in\mathbb{N}} \mu|_A(B_k).$$

Thus $\mu|_A$ is also subadditive. ■

Exercise 2.1.7. Prove for each one of the measures given in the above example, that its value on the empty set is zero.

Exercise 2.1.8. Prove $\mathscr{L}^n\big(\{a\}\big) = 0$ for $a \in \mathbb{R}^n$. Prove that for any countable set E in $\mathbb{R}^n$ we have $\mathscr{L}^n(E) = 0$.

Exercise 2.1.9. Prove that the Lebesgue measure of an interval equals its volume.

Exercise 2.1.10. Prove that the Lebesgue measure of a hyperplane is zero.

Exercise 2.1.11. We set $\lambda X := \{\lambda x \mid x \in X\}$ whenever $\lambda > 0$ and $X \subseteq \mathbb{R}^n$. Prove that $\mathscr{L}^n(\lambda X) = \lambda^n \mathscr{L}^n(X)$.

2.2 Measurable sets

Let μ be a measure on X.

Definition 2.2.1 (μ-measurable set). $A \subseteq X$ is called *μ-measurable* if

$$\mu(B) \geq \mu(B \cap A) + \mu(B \setminus A) \qquad \forall\, B \subseteq X,\ \mu(B) < \infty.$$

Remark 2.2.2. The above inequality is trivial when $\mu(B) = \infty$. Also, since $B = (B \cap A) \cup (B \setminus A)$ we get by the subadditivity of μ that $\mu(B) \leq \mu(B \cap A) + \mu(B \setminus A)$. In particular, μ-measurability of $A \subseteq X$ is equivalent with

$$\mu(B) = \mu(A \cap B) + \mu(B \setminus A) \qquad \forall\, B \subseteq X.$$

Exercise 2.2.3. Let $X \neq \emptyset$ and define the following function on subsets A of X

$$\mu(A) = \begin{cases} 1 & \text{if } A \neq \emptyset, \\ 0 & \text{if } A = \emptyset. \end{cases}$$

Prove that μ is a measure on X and find all μ-measurable subsets of X.

Theorem 2.2.4. (i) *A is μ-measurable $\Leftrightarrow$ A^c is μ-measurable.*

(ii) *$\emptyset, X$ and sets of μ-measure 0 are μ-measurable.*

(iii) *Any μ-measurable set is a also $\mu|_A$-measurable for any $A \subset X$.*

(iv) *A finite union (or intersection) of μ-measurable sets is μ-measurable.*

Proof. The first assertion readily follows from the definition of μ-measurability and so is the μ-measurability of $\emptyset$ and X. Let $\mu(A) = 0$ and $B \subseteq X$ with $\mu(B) < \infty$. Since $B \cap A \subseteq A$ and $B \setminus A \subseteq B$, we get from the monotonicity of μ that $\mu(B \cap A) = 0$ and $\mu(B \setminus A) \leq \mu(B)$, respectively. Add these to get $\mu(B \cap A) + \mu(B \setminus A) \leq \mu(B)$ as required. To prove the third assertion let $C \subseteq X$ be μ-measurable and $B \subseteq X$ with $\mu|_A(B) < \infty$. It is enough to show $\mu|_A(B \cap C) + \mu|_A(B \setminus C) \leq \mu|_A(B)$. Since $(B \setminus C) \cap A = (B \cap A) \setminus C$, this is written as $\mu\big((B \cap A) \cap C\big) + \mu\big((B \cap A) \setminus C\big) \leq \mu(B \cap A)$, which is implied by the μ-measurability of C. For the final assertion, let $B \subseteq X$ with $\mu(B) < \infty$ and pick $A_1, A_2 \subseteq X$. If A_1 is μ-measurable then

$$\mu(B) = \mu(B \cap A_1) + \mu(B \setminus A_1).$$

If A_2 is μ-measurable then

$$\mu(B \setminus A_1) = \mu\big((B \setminus A_1) \cap A_2\big) + \mu\big((B \setminus A_1) \setminus A_2\big).$$

Thus if they are both μ-measurable we have

$$\mu(B) = \mu(B \cap A_1) + \mu\big((B \setminus A_1) \cap A_2\big) + \mu\big((B \setminus A_1) \setminus A_2\big).$$

Noting $B \cap (A_1 \cup A_2) = (B \cap A_1) \cup \big((B \setminus A_1) \cap A_2\big)$, the subadditivity of μ gives

$$\mu\big(B \cap (A_1 \cup A_2)\big) \leq \mu(B \cap A_1) + \mu\big((B \setminus A_1) \cap A_2\big),$$

and inserting this in the previous equality,

$$\mu(B) \geq \mu\big(B \cap (A_1 \cup A_2)\big) + \mu\big((B \setminus A_1) \setminus A_2\big).$$

Observing $(B \setminus A_1) \setminus A_2 = B \setminus (A_1 \cup A_2)$, this last inequality says $A_1 \cup A_2$ is μ-measurable. By induction, the union of finitely many μ-measurable sets is μ-measurable. Taking complements we get using (i) that the intersection of finitely many μ-measurable sets is μ-measurable. ■

Theorem 2.2.5. *Let $\{A_k \subset X\}_{k \in \mathbb{N}}$ be a sequence of μ-measurable sets.*

(i) *If $\{A_k\}_{k \in \mathbb{N}}$ are disjoint, then $\mu\big(\bigcup_{k \in \mathbb{N}} A_k\big) = \sum_{k \in \mathbb{N}} \mu(A_k)$.*

(ii) *If $\{A_k\}_{k \in \mathbb{N}}$ is non-decreasing, then $\lim_{k \to \infty} \mu(A_k) = \mu\big(\bigcup_{k \in \mathbb{N}} A_k\big)$.*

(iii) *If $\{A_k\}_{k \in \mathbb{N}}$ is non-increasing and $\mu(A_1) < \infty$, then $\lim_{k \to \infty} \mu(A_k) = \mu\big(\bigcap_{k \in \mathbb{N}} A_k\big)$.*

(iv) *The sets $\bigcup_{k \in \mathbb{N}} A_k$ and $\bigcap_{k \in \mathbb{N}} A_k$ are μ-measurable.*

Proof. (i) Let $\{A_k\}_{k\in\mathbb{N}}$ be a sequence of μ-measurable disjoint sets. By the μ-measurability of A_{j+1}, $j \in \mathbb{N}$,

$$\mu\big(\bigcup_{k=1}^{j+1} A_k\big) = \mu\Big(\big(\bigcup_{k=1}^{j+1} A_k\big) \cap A_{j+1}\Big) + \mu\Big(\big(\bigcup_{k=1}^{j+1} A_k\big) \setminus A_{j+1}\Big) \tag{2.2.1}$$

$$= \mu(A_{j+1}) + \mu\big(\bigcup_{k=1}^{j} A_k\big).$$

Repeating this formula in the last summand, we arrive at

$$\mu\big(\bigcup_{k=1}^{j+1} A_k\big) = \mu(A_{j+1}) + \mu(A_j) + \mu\big(\bigcup_{k=1}^{j-1} A_k\big) = \dots = \sum_{k=1}^{j+1} \mu(A_k) \qquad \forall\, j \in \mathbb{N}.$$

This readily implies by the monotonicity of μ that

$$\sum_{k=1}^{j+1} \mu(A_k) \le \mu\big(\bigcup_{k\in\mathbb{N}} A_k\big) \qquad \forall\, j \in \mathbb{N}.$$

The proof is completed by taking the limit as $j \to \infty$. Note $\mu(\bigcup_{k\in\mathbb{N}} A_k) \le \sum_{k\in\mathbb{N}} \mu(A_k)$ is also true by the subadditivity of μ. □

(ii) Let $\{A_k\}_{k\in\mathbb{N}}$ be a non-decreasing sequence of μ-measurable sets. For any $k \in \mathbb{N}$ write A_k as a disjoint union of a finite number of μ-measurable sets (because of Theorem 2.2.4-(i) and (iv)) as follows

$$A_k = A_1 \cup \big(\bigcup_{j=1}^{k} (A_{j+1} \setminus A_j)\big).$$

Using (i),

$$\mu(A_k) = \mu(A_1) + \sum_{j=1}^{k} \mu(A_{j+1} \setminus A_j),$$

and taking the limit as $k \to \infty$,

$$\lim_{k\to\infty} \mu(A_k) = \mu(A_1) + \sum_{j=1}^{\infty} \mu(A_{j+1} \setminus A_j).$$

But again because of (i), $\mu(A_1) + \sum_{j=1}^{\infty} \mu(A_{j+1} \setminus A_j) = \mu(\bigcup_{k\in\mathbb{N}} A_k)$. □

(iii) Let $\{A_k\}_{k\in\mathbb{N}}$ be a non-increasing sequence of μ-measurable sets. Then $\{A_1 \setminus A_k\}_{k\in\mathbb{N}}$ is a non-decreasing sequence of μ-measurable sets (because of Theorem 2.2.4-(i) and (iv)) and from (ii) we get

$$\lim_{k\to\infty} \mu(A_1 \setminus A_k) = \mu\big(\bigcup_{k\in\mathbb{N}} (A_1 \setminus A_k)\big). \tag{2.2.2}$$

Because of (i), $\mu(A_1) = \mu(A_1 \setminus A_k) + \mu(A_k)$, $k \in \mathbb{N}$, and since $\mu(A_1) < \infty$ we write $\mu(A_1 \setminus A_k) = \mu(A_1) - \mu(A_k)$, $k \in \mathbb{N}$. Taking the limit,

$$\lim_{k\to\infty} \mu(A_1 \setminus A_k) = \mu(A_1) - \lim_{k\to\infty} \mu(A_k). \tag{2.2.3}$$

On the other hand, noting $\bigcup_{k\in\mathbb{N}}(A_1 \setminus A_k) = A_1 \setminus \bigcap_{k\in\mathbb{N}} A_k$ and since $A_1 = \big(A_1 \setminus \bigcap_{k\in\mathbb{N}}(A_k)\big) \cup \big(\bigcap_{k\in\mathbb{N}} A_k\big)$, the subadditivity of μ gives $\mu(A_1) \le \mu\big(\bigcup_{k\in\mathbb{N}}(A_1 \setminus A_k)\big) + \mu\big(\bigcap_{k\in\mathbb{N}} A_k\big)$.[2] Since $\mu(A_1) < \infty$ we arrive at

$$\mu\Big(\bigcup_{k\in\mathbb{N}}(A_1 \setminus A_k)\Big) \ge \mu(A_1) - \mu\Big(\bigcap_{k\in\mathbb{N}} A_k\Big). \tag{2.2.4}$$

Inserting (2.2.3) and (2.2.4) in (2.2.2) we end up with

$$\lim_{k\to\infty} \mu(A_k) \le \mu\Big(\bigcap_{k\in\mathbb{N}} A_k\Big).$$

The reverse inequality is also true because $\mu\big(\bigcap_{k\in\mathbb{N}} A_k\big) \le \mu(A_k)$ for all $k \in \mathbb{N}$. □

(iv) To prove next that $\bigcup_{k\in\mathbb{N}} A_k$ is μ-measurable, let $B \subseteq X$ with $\mu(B) < \infty$. By definition of the restriction measure we have

$$\mu\Big(B \cap \big(\bigcup_{k\in\mathbb{N}} A_k\big)\Big) + \mu\Big(B \setminus \big(\bigcup_{k\in\mathbb{N}} A_k\big)\Big) = \mu|_B\Big(\bigcup_{k\in\mathbb{N}} A_k\Big) + \mu|_B\Big(\bigcap_{k\in\mathbb{N}} A_k^c\Big). \tag{2.2.5}$$

Defining the sequence of sets $\{B_j\}_{j\in\mathbb{N}}$ by $B_j := \bigcup_{k=1}^{j} A_k$, $j \in \mathbb{N}$, it is trivial that $\bigcup_{k\in\mathbb{N}} A_k = \bigcup_{k\in\mathbb{N}} B_k$ and thus also $\bigcap_{k\in\mathbb{N}} A_k^c = \bigcap_{k\in\mathbb{N}} B_k^c$. The advantage is $\{B_j\}_{j\in\mathbb{N}}$ is a non-decreasing sequence of sets, and thus by (ii) and (iii) we get respectively

$$\mu|_B\Big(\bigcup_{k\in\mathbb{N}} A_k\Big) = \lim_{k\to\infty} \mu|_B(B_k) \quad \text{and} \quad \mu|_B\Big(\bigcap_{k\in\mathbb{N}} A_k^c\Big) = \lim_{k\to\infty} \mu|_B(B_k^c),$$

provided B_j, $j \in \mathbb{N}$, are $\mu|_B$-measurable. By theorem 2.2.4-(iii) it is enough to establish B_j, $j \in \mathbb{N}$, are μ-measurable which is however true since each B_j is a finite union of μ-measurable sets (see Theorem 2.2.4-(iv)). Taking into account that both limits exist ($\mu(B) < \infty$), equation (2.2.5) is the same as

$$\mu\Big(B \cap \big(\bigcup_{k\in\mathbb{N}} A_k\big)\Big) + \mu\Big(B \setminus \big(\bigcup_{k\in\mathbb{N}} A_k\big)\Big) = \lim_{k\to\infty} \big[\mu|_B(B_k) + \mu|_B(B_k^c)\big].$$

Since $B_k \cap B_k^c = \emptyset$, we get from (i) that $\mu|_B(B_k) + \mu|_B(B_k^c) = \mu|_B(X) = \mu(B)$. Taking complements we see the intersection of countably many μ-measurable sets is also μ-measurable. ■

[2] we cannot use equality in place of inequality here (see (i)) because we don't know yet that $\bigcap_{k\in\mathbb{N}} A_k$ is μ-measurable

Definition 2.2.6 (**σ-algebra).** A collection of subsets $\mathscr{A} \subseteq 2^X$ is called a *σ-algebra* provided

(i) $\emptyset,\ X \in \mathscr{A}$,

(ii) $A \in \mathscr{A} \Rightarrow X \setminus A \in \mathscr{A}$,

(iii) $\{A_k \in \mathscr{A}\}_{k\in\mathbb{N}} \Rightarrow \bigcup_{k\in\mathbb{N}} A_k \in \mathscr{A}$.

Example 2.2.7. (i) the collection of all μ-measurable sets,

(ii) the *Borel σ-algebra* of $\mathbb{R}^n$; that is the smallest σ-algebra of $\mathbb{R}^n$ containing the open subsets of $\mathbb{R}^n$.[3] A set in the Borel σ-algebra will be called a *Borel set*.

2.3 Regularity of measures

Definition 2.3.1 (**regular measure).** A measure μ on X is called *regular* if for each $C \subseteq X$ there exists μ-measurable set B such that $C \subseteq B$ and $\mu(B) \leq \mu(C)$ (so $\mu(C) = \mu(B)$ by the monotonicity of measures).

Exercise 2.3.2. For regular μ prove that if $\{A_k\}_{k\in\mathbb{N}}$ is a nondecreasing sequence of subsets of X, then $\lim_{k\to\infty} \mu(A_k) = \mu\big(\bigcup_{k\in\mathbb{N}} A_k\big)$.

Definition 2.3.3 (**Borel, Borel regular and Radon measures).** Let μ be a measure on $\mathbb{R}^n$.

(i) μ is called *Borel* if each Borel set is μ-measurable,

(ii) μ is *Borel regular* if: **(a)** μ is Borel and **(b)** for each $C \subseteq \mathbb{R}^n$ there exists Borel set $B \supseteq C$ with $\mu(B) \leq \mu(C)$ (hence $\mu(C) = \mu(B)$ by the monotonicity of measures),

(iii) μ is called *Radon* if μ is Borel regular and $\mu(K) < \infty$ for all compact $K \subset \mathbb{R}^n$.

Exercise 2.3.4. Prove first that

$$\mathscr{L}^n(A) = \inf\{\mathscr{L}^n(G) \mid \text{open } G \supseteq A\} \qquad \forall\, A \subseteq \mathbb{R}^n.$$

Show next that for each $C \subseteq \mathbb{R}^n$, there exists Borel set $B \supseteq C$ with $\mathscr{L}^n(B) \leq \mathscr{L}^n(C)$ (this does not imply $\mathscr{L}^n$ is Borel regular because we don't know yet it is Borel).

Theorem 2.3.5. *Let μ be a measure on $\mathbb{R}^n$.*

(i) *If μ is Borel then $\mu|_A$ is Borel.*

[3] one can more generally consider any topological space X in place of $\mathbb{R}^n$ and the same applies for definition 2.3.3

(ii) *If μ is Borel regular and $A \subset \mathbb{R}^n$ is a Borel set, then $\mu|_A$ is Borel regular.*

(iii) *If μ is Borel regular and $A \subset \mathbb{R}^n$ is a μ-measurable set with $\mu(A) < \infty$, then $\mu|_A$ is Radon.*

Proof. (i) By definition of μ being Borel, any Borel set is μ-measurable. From theorem 2.2.4-(iii) we know that any μ-measurable set is $\mu|_A$-measurable. Thus any Borel set is $\mu|_A$-measurable and hence $\mu|_A$ is Borel. □

(ii) From (i) we readily have that $\mu|_A$ is a Borel measure. Now let $C \subseteq \mathbb{R}^n$. We want to find a Borel set D such that $C \subseteq D$ and $\mu|_A(D) \leq \mu|_A(C)$. By definition of μ being Borel regular, given $C \subseteq \mathbb{R}^n$, there exists Borel set B such that $C \cap A \subseteq B$ and $\mu|_A(C) = \mu(B)$. Thus, it suffices to find a Borel set D with $C \subseteq D$ and $\mu|_A(D) \leq \mu(B)$. The set $D := B \cup A^c$ is clearly Borel and also $C \subseteq D$. Furthermore, $D \cap A = B \cap A$ and we get $\mu|_A(D) = \mu(B \cap A) \leq \mu(B)$. □

(iii) By definition of μ being Borel regular, there exists Borel set A' such that $A \subseteq A'$ and $\mu(A') = \mu(A)$. Since A is μ-measurable we have $\mu(A') = \mu(A) + \mu(A' \setminus A)$. Since $\mu(A) < \infty$ we conclude $\mu(A' \setminus A) = 0$. Now given $C \subseteq \mathbb{R}^n$ we have $\mu\big((C \cap A') \setminus A\big) = 0$ and using the μ-measurability of A once more,

$$\mu|_{A'}(C) = \mu\big((C \cap A') \cap A\big) + \mu\big((C \cap A') \setminus A\big) = \mu|_A(C).$$

This shows $\mu|_{A'}$ and $\mu|_A$ agree on any set and since A' is a Borel set we conclude from (ii) that $\mu|_A$ is Borel regular. Finally note $\mu|_A(\mathbb{R}^n) = \mu(A) < \infty$; that is μ is a *finite measure*. In particular $\mu|_A(K) < \infty$ for any compact $K \subset \mathbb{R}^n$. ■

Lemma 2.3.6. *Let μ be a Borel measure on $\mathbb{R}^n$ and let B be a Borel set.*

(i) *If $\mu(B) < \infty$, then $\forall\, \varepsilon > 0$, $\exists$ closed $C \subseteq B$ with $\mu(B \setminus C) < \varepsilon$.*

(ii) *If μ is a Radon measure, then $\forall\, \varepsilon > 0$, $\exists$ open $U \supseteq B$ with $\mu(U \setminus B) < \varepsilon$.*

Proof. (i) It is enough to show $B \in \mathscr{F}$ with

$$\mathscr{F} := \big\{A \subseteq \mathbb{R}^n \;\big|\; A \text{ is } \mu\text{-measurable and } \forall\, \varepsilon > 0,\ \exists \text{ closed } C \subseteq A \text{ with } \mu|_B(A \setminus C) < \varepsilon\big\}.$$

Indeed, that would imply the existence of a closed set C with $C \subseteq B$ and $\mu|_B(B \setminus C) < \varepsilon$, which in turn implies (i) because $\mu|_B(B \setminus C) = \mu(B \setminus C)$. We next prove $B \in \mathscr{F}$ by showing

$$\mathscr{G} := \{A \in \mathscr{F} \;\big|\; A^c \in \mathscr{F}\}$$

is a σ-algebra containing all open sets. This would imply it contains all Borel sets, hence B in particular.

We start by listing the required properties of $\mathscr{F}$.

(a) $\mathscr{F}$ contains all closed sets (this follows readily from the definition of $\mathscr{F}$).

(b) If $\{A_k \in \mathscr{F}\}_{k\in\mathbb{N}}$ then $A := \bigcap_{k\in\mathbb{N}} A_k \in \mathscr{F}$.

Proof. Let $\varepsilon > 0$. Since $A_k \in \mathscr{F}$ for all $k \in \mathbb{N}$, there exist closed sets $\{C_k\}_{k\in\mathbb{N}}$ with $C_k \subseteq A_k$ and $\mu|_B(A_k \setminus C_k) < \varepsilon 2^{-k}$, $k \in \mathbb{N}$. By theorem 1.1.9 we have $C := \bigcap_{k\in\mathbb{N}} C_k$ is closed and since $A \setminus C \subseteq \bigcup_{k\in\mathbb{N}}(A_k \setminus C_k)$[4],

$$\mu|_B(A \setminus C) \le \sum_{k\in\mathbb{N}} \mu|_B(A_k \setminus C_k) < \varepsilon.$$

Thus $A \in \mathscr{F}$. $\square$

(c) If $\{A_k \in \mathscr{F}\}_{k\in\mathbb{N}}$ then $A := \bigcup_{k\in\mathbb{N}} A_k \in \mathscr{F}$.

Proof. Let ε and C_k as above. Setting $C := \bigcup_{k\in\mathbb{N}} C_k$ we have $A \setminus C \subseteq \bigcup_{k\in\mathbb{N}}(A_k \setminus C_k)$[5] which gives

$$\mu|_B(A \setminus C) \le \sum_{k\in\mathbb{N}} \mu|_B(A_k \setminus C_k) < \varepsilon.$$

The sequence $A \setminus (\bigcup_{k=1}^{m} C_k)$ is non-increasing and so the above inequality together with theorem 2.2.5-(iii) imply $\lim_{m\to\infty} \mu|_B\big(A \setminus \big(\bigcup_{k=1}^{m} C_k\big)\big) < \varepsilon$. In particular $\mu|_B\big(A \setminus \big(\bigcup_{k=1}^{m_0} C_k\big)\big) < \varepsilon$ for some $m_0 \in \mathbb{N}$. By theorem 1.1.9 we get $\bigcup_{k=1}^{m_0} C_k$ is also closed and thus $A \in \mathscr{F}$. $\square$

(d) $\mathscr{F}$ contains all open sets.

Proof. According to theorem 1.1.22, any open set in $\mathbb{R}^n$ is a (countable) union of intervals, which are closed by definition. But as observed in (a), $\mathscr{F}$ contains all closed sets and thus (c) implies $\mathscr{F}$ contains all open sets. $\square$

We may now proceed to check the collection of subsets $\mathscr{G}$ is a σ-algebra containing all open sets. Since $\mathscr{F}$ contains all open and all closed sets, $\mathscr{G}$ contains all open sets. Consequently it contains $\emptyset$ and $\mathbb{R}^n$ and by it's definition, it contains the complement of each of it's elements. Let $\{A_k \in \mathscr{G}\}_{k\in\mathbb{N}}$. Then $A_k \in \mathscr{F}$, $k \in \mathbb{N}$, and also $A_k^c \in \mathscr{F}$, $k \in \mathbb{N}$, which imply because of (c) and (b), respectively, that $\bigcup_{k\in\mathbb{N}} A_k \in \mathscr{F}$ and $\big(\bigcup_{k\in\mathbb{N}} A_k\big)^c = \bigcap_{k\in\mathbb{N}} A_k^c \in \mathscr{F}$. $\square$

(ii) The sequence $\{B_k(0) \setminus B\}_{k\in\mathbb{N}}$ is comprised of Borel sets with

$$\mu(B_k(0) \setminus B) \le \mu(\bar{B}_k(0)) < \infty \qquad \forall\, k \in \mathbb{N}.$$

[4] Clearly, $A \setminus C_k \subseteq A_k \setminus C_k$ for all $k \in \mathbb{N}$, hence $\bigcup_{k\in\mathbb{N}}(A \setminus C_k) \subseteq \bigcup_{k\in\mathbb{N}}(A_k \setminus C_k)$. Now observe that $\bigcup_{k\in\mathbb{N}}(A \setminus C_k) = A \setminus C$.

[5] Clearly, $A_k \setminus C \subseteq A_k \setminus C_k$ for all $k \in \mathbb{N}$, hence $\bigcup_{k\in\mathbb{N}}(A_k \setminus C) \subseteq \bigcup_{k\in\mathbb{N}}(A_k \setminus C_k)$. Now observe that $\bigcup_{k\in\mathbb{N}}(A_k \setminus C) = A \setminus C$.

Applying (i) we get closed sets $\{C_k \subseteq B_k(0)\setminus B\}_{k\in\mathbb{N}}$ such that $\mu\big((B_k(0)\setminus B)\setminus C_k\big) < \varepsilon 2^{-k}$, $k\in\mathbb{N}$. Set

$$U := \bigcup_{k\in\mathbb{N}} (B_k(0)\setminus C_k).$$

Since $B \subseteq C_k^c$ [6] for all $k\in\mathbb{N}$, we have $B = \bigcup_{k\in\mathbb{N}}(B_k(0)\cap B) \subseteq \bigcup_{k\in\mathbb{N}}(B_k(0)\cap C_k^c) = U$. Furthermore, observing

$$U\setminus B = \Big(\bigcup_{k\in\mathbb{N}}(B_k(0)\setminus C_k)\Big)\setminus B = \bigcup_{k\in\mathbb{N}}\big((B_k(0)\setminus C_k)\setminus B\big) = \bigcup_{k\in\mathbb{N}}\big((B_k(0)\setminus B)\setminus C_k\big),$$

we have

$$\mu(U\setminus B) \le \sum_{k\in\mathbb{N}} \mu\big((B_k(0)\setminus B)\setminus C_k\big) < \varepsilon.$$

Finally notice U is open by it's definition. ■

Theorem 2.3.7 (outer regularity). *If μ is a Radon measure on $\mathbb{R}^n$ and $A\subseteq\mathbb{R}^n$, then*

$$\mu(A) = \inf\{\mu(U) \mid \textit{open } U \supseteq A\}.$$

Proof. If $\mu(A)=\infty$, by the monotonicity of μ we have $\mu(U)=\infty$ for any open $U\supseteq A$ and so the above infimum is also ∞. Assume next $\mu(A)<\infty$ and let $\varepsilon>0$. If in addition A is a Borel set, then by the second part of the previous lemma we find open $U\supseteq A$ with $\mu(U\setminus A)<\varepsilon$. The μ-measurability of A together with $\mu(A)<\infty$ imply $\mu(U)-\mu(U\cap A)<\varepsilon$ and so $\mu(U)<\mu(A)+\varepsilon$. For arbitrary A we know there exists a Borel set B such that $A\subseteq B$, $\mu(B)=\mu(A)$ and, by the above argument, $\mu(B)=\inf\{\mu(U)\mid \text{open } U\supseteq B\}$. Altogether,

$$\mu(A) = \inf\{\mu(U)\mid \text{open } U\supseteq B\} \ge \inf\{\mu(U)\mid \text{open } U\supseteq A\}.$$

The reverse inequality holds true as well by the monotonicity of μ. ■

Theorem 2.3.8 (inner regularity). *If μ is a Radon measure on $\mathbb{R}^n$ and $A\subseteq\mathbb{R}^n$ is μ-measurable, then*

$$\mu(A) = \sup\{\mu(K)\mid \textit{compact } K\subseteq A\}.$$

Proof. Assume first $\mu(A)<\infty$ and let $\varepsilon>0$. By theorem 2.3.5-(iii) we know $\mu|_A$ is a Radon measure. Applying theorem 2.3.7 for $\mu|_A$ on the set A^c (observe $\mu|_A(A^c)=0$) we find open $U\supseteq A^c$ with $\mu|_A(U)<\varepsilon$. Then $C:=U^c$ is closed with $C\subseteq A$. Since also $U\cap A = A\setminus C$, there holds

$$\varepsilon > \mu(A\setminus C) = \mu(A)-\mu(C),$$

[6] Taking complements in $C_k\subseteq B_k(0)\setminus B$, we get $B_k^c(0)\cup B\subseteq C_k^c$. In particular, $B\subseteq C_k^c$.

by the μ-measurability of C. Hence,

$$\mu(A) = \sup\{\mu(C) \mid \text{closed } C \subseteq A\}. \tag{2.3.1}$$

Next we show (2.3.1) is valid also when $\mu(A) = \infty$. Define $D_k := B_k(0) \setminus B_{k-1}(0)$, $k \in \mathbb{N}$, and write A as a disjoint union of sets: $A = \bigcup_{k\in\mathbb{N}}(D_k \cap A)$. The subadditivity of μ implies

$$\sum_{k\in\mathbb{N}} \mu(D_k \cap A) \geq \mu(A) = \infty. \tag{2.3.2}$$

Since μ is Radon, each $D_k \cap A$ is of finite μ-measure[7] and we may apply the above argument to get closed C_k, $k \in \mathbb{N}$, such that $C_k \subseteq D_k \cap A$ with $\mu(C_k) \geq \mu(D_k \cap A) - 2^{-k}$, $k \in \mathbb{N}$. Clearly, $\bigcup_{k\in\mathbb{N}} C_k \subseteq A$. Since $\{C_k\}_{k\in\mathbb{N}}$ is a sequence of disjoint μ-measurable sets, using theorem 2.2.5-(i) we have

$$\mu\Big(\bigcup_{k\in\mathbb{N}} C_k\Big) = \sum_{k\in\mathbb{N}} \mu(C_k) \geq \sum_{k\in\mathbb{N}} \big(\mu(D_k \cap A) - 2^{-k}\big).$$

The last sum is ∞ by (2.3.2). Hence by theorem 2.2.5-(ii) we conclude $\lim_{m\to\infty} \mu\big(\bigcup_{k=1}^m C_k\big) = \infty$. But theorem 1.1.9 says $\big\{\bigcup_{k=1}^m C_k\big\}_{m\in\mathbb{N}}$ is comprised of closed sets and thus (2.3.1) is true also if $\mu(A) = \infty$. To complete the proof of the theorem, it suffices to show

$$\sup\{\mu(C) \mid \text{closed } C \subseteq A\} \leq \sup\{\mu(K) \mid \text{compact } K \subseteq A\},$$

the reverse inequality being trivially true. Let C be a closed subset of A. By theorem 2.2.5-(ii) we get

$$\mu(C) = \lim_{m\to\infty} \mu\big(C \cap \bar{B}_m(0)\big) \leq \sup\{\mu(K) \mid \text{compact } K \subseteq A\},$$

where in the last inequality we used $C \cap \bar{B}_m(0)$, $m \in \mathbb{N}$, are compact subsets of A. ■

Exercise 2.3.9. Let $0 < \lambda < 1$ and suppose $X \subset \mathbb{R}$ satisfies $0 < \mathscr{L}^1(X) < \infty$. Prove there exists an open interval I such that $\mathscr{L}^1(X \cap I) \geq \lambda \mathscr{L}^1(I)$.

Lemma 2.3.10. *Let $\{E_t \subset \mathbb{R}^n\}_{t\in\mathscr{I}}$ be a family of disjoint Borel sets and μ be a Radon measure on $\mathbb{R}^n$. Then $\mu(E_t) > 0$ for at most countably many $t \in \mathscr{I}$.*

Proof. This is self-evident in case $\mathscr{I}$ is at most countable. Suppose next $\mathscr{I}$ is uncountable. We write (why?)

$$\big\{t \in \mathscr{I} \mid \mu(E_t) > 0\big\} = \bigcup_{k\in\mathbb{N}} I_k, \tag{2.3.3}$$

[7] $\mu(D_k \cap A) \leq \mu(\bar{B}_k(0)) < \infty$

where

$$I_k := \Big\{ t \in \mathscr{I} \;\Big|\; \mu\big(E_t \cap B_k(0)\big) > \frac{1}{k} \Big\}.$$

Then for any finite set $J \subset I_k$, $k \in \mathbb{N}$, we have

$$\begin{aligned} \mu\big(B_k(0)\big) &\geq \mu\Big(\bigcup_{t \in \mathscr{I}} E_t \cap B_k(0) \Big) \\ &\geq \mu\Big(\bigcup_{t \in J} E_t \cap B_k(0) \Big) \\ &= \sum_{t \in J} \mu\big(E_t \cap B_k(0)\big) \\ &\geq \frac{\mathscr{H}^0(J)}{k}. \end{aligned}$$

Hence, for any $k \in \mathbb{N}$ the finite number $k\mu\big(B_k(0)\big)$ bounds $\mathscr{H}^0(J)$ independently of the choice of the finite set of indices $J \subset I_k$. This implies for any $k \in \mathbb{N}$, the set I_k is itself finite. Consequently, (2.3.3) implies $\big\{ t \in \mathscr{I} \;\big|\; \mu(E_t) > 0 \big\}$ is at most countable. ■

2.4 Metric measures and Carathéodory's criterion

Definition 2.4.1 (metric measure). A measure μ on $\mathbb{R}^n$ is called a *metric measure* if

$$\mu(A \cup B) \geq \mu(A) + \mu(B) \text{ whenever } A, B \subset \mathbb{R}^n \text{ with } \operatorname{dist}(A,B) > 0 \text{ and } \mu(A \cup B) < \infty.$$

Remark 2.4.2. The above inequality is trivial when $\mu(A \cup B) < \infty$. Also, by the subadditivity property of measures, the reverse inequality is true as well. Hence we have the following equivalent condition that a measure μ has to satisfy in order to be a metric measure:

$$\mu(A \cup B) = \mu(A) + \mu(B) \text{ whenever } A, B \subset \mathbb{R}^n \text{ with } \operatorname{dist}(A,B) > 0.$$

Example 2.4.3. (i) The Lebesgue measure on $\mathbb{R}^n$.

Proof. Let $A, B \subset \mathbb{R}^n$ with $\operatorname{dist}(A,B) > 0$ and $\mathscr{L}^n(A \cup B) < \infty$. Let $\varepsilon > 0$ and pick a covering of $A \cup B$ by intervals $\{I_k\}_{k \in \mathbb{N}}$, such that

$$\sum_{k \in \mathbb{N}} \operatorname{Vol}(I_k) < \mathscr{L}^n(A \cup B) + \varepsilon.$$

By possibly dividing I_k we may further assume $\operatorname{diam}(I_k) < \operatorname{dist}(A,B)$ for all $k \in \mathbb{N}$. Hence the covering $\{I_k\}_{k \in \mathbb{N}}$ splits in two subsequences $\{I_k^A\}_{k \in \mathbb{N}}, \{I_k^B\}_{k \in \mathbb{N}}$, the first of which covers A and the second covers B. We have

$$\mathscr{L}^n(A) + \mathscr{L}^n(B) \leq \sum_{k \in \mathbb{N}} \operatorname{Vol}(I_k^A) + \sum_{k \in \mathbb{N}} \operatorname{Vol}(I_k^B) = \sum_{k \in \mathbb{N}} \operatorname{Vol}(I_k).$$

Coupling these two inequalities, $\mathscr{L}^n(A)+\mathscr{L}^n(B)<\mathscr{L}^n(A\cup B)+\varepsilon$, which implies the claim since $\varepsilon>0$ is arbitrary. ■

(ii) The s-dimensional Hausdorff measure on $\mathbb{R}^n$.

Proof. Let $A,B\subset\mathbb{R}^n$ with $\operatorname{dist}(A,B)>0$ and $\mathscr{H}^s(A\cup B)<\infty$. Let $\varepsilon>0$. By remark 2.1.5 we know $\mathscr{H}^s_\delta(A\cup B)\uparrow\mathscr{H}^s(A\cup B)$ as $\delta\downarrow 0$. We pick a covering of $A\cup B$ by sets $\{C_k\}_{k\in\mathbb{N}}$ with $\operatorname{diam}(C_k)\le\delta<\operatorname{dist}(A,B)$, such that

$$\sum_{k\in\mathbb{N}}\omega_s\Big(\frac{\operatorname{diam}C_k}{2}\Big)^s<\mathscr{H}^s_\delta(A\cup B)+\varepsilon\le\mathscr{H}^s(A\cup B)+\varepsilon.$$

Let $\mathscr{A}$ be the set of all sets of $\{C_k\}_{k\in\mathbb{N}}$ having nonempty intersection with A; that is,

$$\mathscr{A}:=\{C_k\mid C_k\cap A\neq\emptyset\}.$$

Correspondingly set

$$\mathscr{B}:=\{C_k\mid C_k\cap B\neq\emptyset\}.$$

Clearly, $A\subseteq\bigcup_{C_k\in\mathscr{A}}C_k$ and $B\subseteq\bigcup_{C_k\in\mathscr{B}}C_k$, while the assumption $\delta<\operatorname{dist}(A,B)$ gives that any set in $\mathscr{A}$ is disjoint with any set from $\mathscr{B}$. We have

$$\begin{aligned}\mathscr{H}^s_\delta(A)+\mathscr{H}^s_\delta(B)&\le\sum_{C_k\in\mathscr{A}}\omega_s\Big(\frac{\operatorname{diam}C_k}{2}\Big)^s+\sum_{C_k\in\mathscr{B}}\omega_s\Big(\frac{\operatorname{diam}C_k}{2}\Big)^s\\&=\sum_{k\in\mathbb{N}}\omega_s\Big(\frac{\operatorname{diam}C_k}{2}\Big)^s.\end{aligned}$$

Coupling these two inequalities, $\mathscr{H}^s_\delta(A)+\mathscr{H}^s_\delta(B)<\mathscr{H}^s(A\cup B)+\varepsilon$. Letting $\delta\to 0$, $\mathscr{H}^s(A)+\mathscr{H}^s(B)<\mathscr{H}^s(A\cup B)+\varepsilon$ and the claim follows since $\varepsilon>0$ is arbitrary. ■

Theorem 2.4.4 (Carathéodory). *If μ is a metric measure on $\mathbb{R}^n$, then μ is a Borel measure.*

Proof. It is enough to prove any closed set is μ-measurable. Let $C\subseteq\mathbb{R}^n$ be closed. It is enough to prove $\mu(A)\ge\mu(A\cap C)+\mu(A\setminus C)$ for any $A\subset\mathbb{R}^n$ with $\mu(A)<\infty$. To this end we set

$$C_m:=\{x\in\mathbb{R}^n\mid\operatorname{dist}(x,C)\le 1/m\},\qquad m\in\mathbb{N}.$$

Let $A\subseteq\mathbb{R}^n$ with $\mu(A)<\infty$. Since $\operatorname{dist}(A\setminus C_m,A\cap C)\ge 1/m>0$ for all $m\in\mathbb{N}$, by the hypotheses of the theorem we have

$$\mu(A\setminus C_m)+\mu(A\cap C)=\mu\big((A\setminus C_m)\cup(A\cap C)\big)\le\mu(A).$$

It remains to show $\lim_{m\to\infty}\mu(A\setminus C_m)=\mu(A\setminus C)$. Set

$$R_k:=\{x\in A\mid 1/(k+1)<\operatorname{dist}(x,C)\le 1/k\},\qquad k\in\mathbb{N}.$$

Then $A \setminus C = (A \setminus C_m) \cup \left(\bigcup_{k=m}^{\infty} R_k\right)$ and by the subadditivity of μ,

$$\mu(A \setminus C_m) \leq \mu(A \setminus C) \leq \mu(A \setminus C_m) + \sum_{k=m}^{\infty} \mu(R_k).$$

In turn, for $\lim_{m\to\infty} \sum_{k=m}^{\infty} \mu(R_k) = 0$ to hold true, it is enough that the series $\sum_{k\in\mathbb{N}} \mu(R_k)$ converges. But $\operatorname{dist}(R_i, R_j) > 0$ for $j \geq i+2$ and using successively the hypotheses of the theorem,

$$\sum_{k=1}^{m} \mu(R_{2k}) = \mu\big(\bigcup_{k=1}^{m} R_{2k}\big) \leq \mu(A) \qquad \forall\, m \in \mathbb{N},$$

and also

$$\sum_{k=0}^{m} \mu(R_{2k+1}) = \mu\big(\bigcup_{k=0}^{m} R_{2k+1}\big) \leq \mu(A) \qquad \forall\, m \in \mathbb{N}.$$

These readily imply $\sum_{k\in\mathbb{N}} \mu(R_k) \leq 2\mu(A)$. ■

Remark 2.4.5. By exercise 2.3.4 and by example 2.4.3-(i) we obtain $\mathscr{L}^n$ is Borel regular. Since it is clearly finite on compact sets, we conclude $\mathscr{L}^n$ is Radon.

Exercise 2.4.6. Prove $\mathscr{H}^s$, $s \geq 0$, is Borel regular.

Remark 2.4.7. $\mathscr{H}^s$ is not a Radon measure for $0 \leq s < n$. Using the Brunn-Minkowski inequality of the next section one can prove the *isodiametric inequality*; that is, *among all bounded sets of* $\mathbb{R}^n$ *having the same fixed diameter, it is the ball that maximizes the volume*, or what is the same, *for all* $X \subseteq \mathbb{R}^n$ *there holds*

$$\mathscr{L}^n(X) \leq \mathscr{L}^n(B_X).$$

where B_X *is defined to be any ball with radius equal to half the diameter of* X. Using the isodiametric inequality one can show that $\mathscr{H}^n = \mathscr{L}^n$. These facts will be included in a possible future addendum to this notes.

Chapter 3

Brunn-Minkowski and isoperimetric inequalities

Given $\emptyset \neq X \subseteq \mathbb{R}^n$ and $\varepsilon > 0$ we define the set

$$X^\varepsilon := \{x \in \mathbb{R}^n \mid \operatorname{dist}(x, X) < \varepsilon\},$$

with $X^\varepsilon := \mathbb{R}^n$ in case $X = \mathbb{R}^n$. We start with a version of the *isoperimetric inequality* not involving the notion of perimeter.

Theorem 3.0.1 (isoperimetric inequality). *Given any $\mathscr{L}^n$-measurable $X \subset \mathbb{R}^n$ satisfying $0 < \mathscr{L}^n(X) < \infty$, consider any ball B_{r_X} with the same Lebesgue measure; that is,*

$$\mathscr{L}^n(X) = \omega_n r_X^n. \tag{3.0.1}$$

Then for any $\varepsilon > 0$ there holds

$$\mathscr{L}^n(X^\varepsilon) \geq \mathscr{L}^n(B_{r_X}^\varepsilon).$$

Definition 3.0.2 (Minkowski sum). The *Minkowski sum* of nonempty $X, Y \subseteq \mathbb{R}^n$ is given by

$$X + Y := \{x + y \mid x \in X,\ y \in Y\}.$$

Remark 3.0.3. Throughout we use the simple fact that

$$X_1 \subseteq Y_1 \quad \text{and} \quad X_2 \subseteq Y_2 \implies X_1 + X_2 \subseteq Y_1 + Y_2.$$

Observing $X^\varepsilon \supseteq X + B_\varepsilon(0)$[1] and $B^\varepsilon_{r_X} = B_{r_X+\varepsilon}$[2], the isoperimetric inequality would be readily implied by the following inequality

$$\begin{aligned}\Big(\mathscr{L}^n\big(X+B_\varepsilon(0)\big)\Big)^{1/n} &\geq \omega_n^{1/n}(r_X+\varepsilon) \\ &= \big(\mathscr{L}^n(X)\big)^{1/n} + \Big(\mathscr{L}^n\big(B_\varepsilon(0)\big)\Big)^{1/n}.\end{aligned} \tag{3.0.2}$$

This shows the isoperimetric inequality is a special case of the following inequality

Theorem 3.0.4 (Brunn-Minkowski inequality). *For any non-empty and $\mathscr{L}^n$-measurable $X, Y \subseteq \mathbb{R}^n$, there holds*

$$\big(\mathscr{L}^n(X+Y)\big)^{1/n} \geq \big(\mathscr{L}^n(X)\big)^{1/n} + \big(\mathscr{L}^n(Y)\big)^{1/n}. \tag{3.0.3}$$

Now define the $n-1$ *dimensional lower Minkowski content* of a bounded $X \subset \mathbb{R}^n$ by[3]

$$\mathrm{Per}(X) := \liminf_{\varepsilon\downarrow 0} \frac{\mathscr{L}^n\big(X+B_\varepsilon(0)\big) - \mathscr{L}^n(X)}{\varepsilon}.$$

If $X \subset \mathbb{R}^n$ is bounded, noting $B_{r_X+\varepsilon} \supseteq B_{r_x} + B_\varepsilon(0)$ (combine the last two footnotes taking $X = B_{r_X}$), we further obtain from (3.0.2) and (3.0.1)

$$\frac{\mathscr{L}^n\big(X+B_\varepsilon(0)\big) - \mathscr{L}^n(X)}{\varepsilon} \geq \frac{\mathscr{L}^n\big(B_{r_X}+B_\varepsilon(0)\big) - \mathscr{L}^n(B_{r_X})}{\varepsilon}.$$

Passing to the lim inf (see Exercise 1.1.3) we conclude

$$\mathrm{Per}(X) \geq \mathrm{Per}(B_{r_X}), \tag{3.0.4}$$

which asserts

> *among all bounded sets of $\mathbb{R}^n$ having the same fixed volume, it is the ball that minimizes the perimeter,*

[1]Given $y \in X + B_\varepsilon(0)$ we know there exists $x \in X$ and $z \in B_\varepsilon(0)$ such that $y = x + z$. Hence, $\mathrm{dist}(y,X) \leq |y-x| = |z| < \varepsilon \Rightarrow y \in X^\varepsilon$. As an exercise, show these two sets are in fact equal whenever X is closed.

[2]It is enough to prove $B^\varepsilon_r(0) = B_{r+\varepsilon}(0)$. Given $x \in B^\varepsilon_r(0)$ we know $\mathrm{dist}\big(x, B_r(0)\big) < \varepsilon$. For any $z \in B_r(0)$ we have $|x| \leq |x-z| + |z| < |x-z| + r$. Hence $|x| \leq \mathrm{dist}\big(x, B_r(0)\big) + r < \varepsilon + r \Rightarrow x \in B_{r+\varepsilon}(0)$. Conversely, given $x \in B_{r+\varepsilon}(0)$ we know $|x| < r + \varepsilon$, so if $x \in B_{r+\varepsilon}(0) \setminus B_r(0)$ then $\mathrm{dist}(x, B_r(0)) = |x| - r < \varepsilon$. Hence $B_{r+\varepsilon}(0) \setminus B_r(0) \subset B^\varepsilon_r(0)$. But also $B_r(0) \subset B^\varepsilon_r(0)$ and so $B_{r+\varepsilon}(0) \subset B^\varepsilon_r(0)$.

[3]The smoother the boundary of the set X is, the $n-1$ dimensional lower Minkowski content coincides with its perimeter (in a weak notion which is out of the scope of these notes), $\mathscr{H}^{n-1}(\partial X)$ or even the surface area of ∂X from calculus. This is why here (by abuse of notation) we denote it by Per.

or, equivalently,

> *among all bounded sets of $\mathbb{R}^n$ having the same fixed perimeter, it is the ball that maximizes the volume.*

This last formulation is the classical isoperimetric statement. Finally we notice the perimeter of B_{r_X} can be explicitly computed. Indeed,

$$\mathrm{Per}(B_{r_X}) = \liminf_{\varepsilon\downarrow 0} \frac{\omega_n(r_X+\varepsilon)^n - \omega_n r_X^n}{\varepsilon} = n\omega_n r_X^{n-1},$$

and using once more (3.0.1) on the right, we deduce after coupling with (3.0.4),

$$\mathrm{Per}(X) \geq n\omega_n^{1/n}\big(\mathscr{L}^n(X)\big)^{1-1/n}.$$

This is the classical expression of the isoperimetric inequality involving the notion of perimeter.

Proof of the Brunn-Minkowski inequality. The proof splits in three steps.

Step 1 - The simplest case. Suppose first X and Y are intervals with sides lengths $\{x_1,...,x_n\}$, $\{y_1,...,y_n\}$ respectively. Then $X+Y$ is an interval with sides lengths $\{x_1+y_1,...,x_n+y_n\}$ and the Brunn-Minkowski inequality reads

$$\Big(\prod_{i=1}^n (x_i+y_i)\Big)^{1/n} \geq \Big(\prod_{i=1}^n x_i\Big)^{1/n} + \Big(\prod_{i=1}^n y_i\Big)^{1/n},$$

where we have used exercise 2.1.9. Writing this as

$$1 \geq \Big(\prod_{i=1}^n \frac{x_i}{x_i+y_i}\Big)^{1/n} + \Big(\prod_{i=1}^n \frac{y_i}{x_i+y_i}\Big)^{1/n},$$

we see it is a direct consequence of the *arithmetic-geometric mean inequality*, which we recall next. □

Lemma 3.0.5 (arithmetic-geometric mean inequality). [4] *For positive real numbers $a_1,...,a_n$,*

$$\Big(\prod_{i=1}^n a_i\Big)^{1/n} \leq \frac{1}{n}\sum_{i=1}^n a_i.$$

[4] it's proof will be an easy exercise (see remark 7.1.4) as soon as we learn about convex functions

Step 2 - Reductions. From remark 2.4.5 we know $\mathscr{L}^n$ is Radon and this implies the inner regularity (theorem 2.3.8) of $\mathscr{L}^n$. Suppose X and Y are $\mathscr{L}^n$-measurable and pick compact sets K_X, K_Y such that $K_X \subseteq X$ and $K_Y \subseteq Y$. Then $K_X + K_Y \subseteq X + Y$ and assuming the Brunn-Minkowski inequality holds true for all compact sets, we get from the monotonicity of $\mathscr{L}^n$

$$\left(\mathscr{L}^n(X+Y)\right)^{1/n} \geq \left(\mathscr{L}^n(K_X)\right)^{1/n} + \left(\mathscr{L}^n(K_Y)\right)^{1/n}.$$

Taking the supremum over all compact $K_X \subseteq X$ and over all $K_Y \subseteq Y$, we deduce (3.0.3). Thus (3.0.3) will be true for all $\mathscr{L}^n$-measurable X and Y, if it is true for all compact X and Y. Let X and Y be compact. Recalling exercise 1.1.23, we construct two non-increasing sequences $\{X_k \supseteq X\}_{k\in\mathbb{N}}$ and $\{Y_k \supseteq Y\}_{k\in\mathbb{N}}$ with $X = \bigcap_{k\in\mathbb{N}} X_k$, $Y = \bigcap_{k\in\mathbb{N}} Y_k$ and such that each X_k and Y_k, $k \in \mathbb{N}$, consists of a finite union of non-overlapping intervals. Suppose

$$\left(\mathscr{L}^n(X_k+Y_k)\right)^{1/n} \geq \left(\mathscr{L}^n(X_k)\right)^{1/n} + \left(\mathscr{L}^n(Y_k)\right)^{1/n}.$$

Since $X_k \supseteq X$ and $Y_k \supseteq Y$, from the monotonicity of $\mathscr{L}^n$ we arrive at

$$\left(\mathscr{L}^n(X_k+Y_k)\right)^{1/n} \geq \left(\mathscr{L}^n(X)\right)^{1/n} + \left(\mathscr{L}^n(Y)\right)^{1/n}.$$

Taking the limit $k \to \infty$ we use theorem 2.2.5-(*iii*) to arrive to (3.0.3). Indeed, noting $\{X_k + Y_k \supseteq X+Y\}_{k\in\mathbb{N}}$ is a non-increasing sequence of sets with $X+Y = \bigcap_{k\in\mathbb{N}}(X_k+Y_k)$,[5] we deduce $\mathscr{L}^n(X+Y) = \lim_{k\to\infty} \mathscr{L}^n(X_k+Y_k)$. □

Step 3 - The case where X and Y are finite unions of non-overlapping intervals. Let X and Y be finite unions of non-overlapping intervals. The proof in this case follows by induction on the total number of intervals $N(X,Y)$ that constitute X and Y (we will denote simply by $N(X)$ the case where $X = Y$). If $N(X,Y) = 2$ then (3.0.3) has been proved in the first step. Let $N(X,Y) = m \geq 3$ and suppose that (3.0.3) is true when $N(X,Y) \leq m-1$. Since $m \geq 3$ we assume that X consists of at least two non-overlapping intervals I_1, I_2. Then there exists $i \in \{1,...,n\}$ and $s \in \mathbb{R}$ such that the hyperplane $\{x = (x_1,...,x_n) \in \mathbb{R}^n \mid x_i = s\}$ separates them. From now on, for simplicity, we denote a hyperplane of this form by $\{x_i = s\}$ and by $\{x_i \geq s\}$, $\{x_i \leq s\}$ the two "half spaces" produced by it. We define the sets

$$X^+ := X \cap \{x_i \geq s\} \quad \text{and} \quad X^- := X \cap \{x_i \leq s\}.$$

Then X^+ and X^- are non-overlapping sets comprised of non-overlapping intervals and the number of intervals in each one is strictly less than $m-1$ (it can be at most $m-2$ in case Y is

[5]If $\alpha \in \bigcap_{k\in\mathbb{N}}(X_k+Y_k)$, then $\forall\, k \in \mathbb{N}$, $\exists\, x_k \in X_k$ and $\exists\, y_k \in Y_k$ such that $x_k + y_k = \alpha$. But $\{X_k\}_{k\in\mathbb{N}}$ is non-increasing, hence $x_k \in X_1$ for all $k \in \mathbb{N}$ and since X_1 is compact, a subsequence of x_k converges to $x \in X_1$. Since $x_k \in X_k$ for all $k \in \mathbb{N}$ and $\{X_k\}_{k\in\mathbb{N}}$ is non-increasing to X, we deduce $x \in X_k$ for all $k \in \mathbb{N}$, that is $x \in X$. In the same way, we get a a subsequence of y_k converging to $y \in Y$. Taking the subsequential limit in $x_k + y_k = \alpha$, we get $x + y = \alpha$. Thus $\alpha \in X+Y$. The reverse inclusion is immediate since $X+Y \subseteq X_k+Y_k$ for all $k \in \mathbb{N}$.

comprised of only one interval). Set

$$\lambda := \frac{\mathscr{L}^n(X^+)}{\mathscr{L}^n(X)}.$$

Then $\lambda \in (0,1)$ and since $X = X^+ \cup X^-$, we have $\mathscr{L}^n(X) = \mathscr{L}^n(X^+) + \mathscr{L}^n(X^-)$ by exercise 2.1.10. Thus

$$\mathscr{L}^n(X^-) = (1-\lambda)\mathscr{L}^n(X). \tag{3.0.5}$$

Pick a parallel hyperplane $\{x_i = t\}$ such that if $Y^+ := Y \cap \{x_i \geq t\}$ and $Y^- := Y \cap \{x_i \leq t\}$, then

$$\frac{\mathscr{L}^n(Y^+)}{\mathscr{L}^n(Y)} = \lambda.^6 \tag{3.0.6}$$

Since $X^\pm \subset X$ and $Y^\pm \subset Y$, we get $X^\pm + Y^\pm \subset X + Y$ which gives $(X^+ + Y^+) \cup (X^- + Y^-) \subset X + Y$. Thus

$$\mathscr{L}^n(X+Y) \geq \mathscr{L}^n\big((X^+ + Y^+) \cup (X^- + Y^-)\big).$$

But $X^+ + Y^+ \subseteq \{x_i \geq s+t\}$ and $X^- + Y^- \subseteq \{x_i \leq s+t\}$, which implies $X^+ + Y^+$ and $X^- + Y^-$ are non-overlapping. Using exercise 2.1.10 the above inequality becomes

$$\mathscr{L}^n(X+Y) \geq \mathscr{L}^n(X^+ + Y^+) + \mathscr{L}^n(X^- + Y^-).$$

Since $N(X^+ + Y^+),\ N(X^- + Y^-) \leq m-1$, the induction hypothesis further gives

$$\mathscr{L}^n(X+Y) \geq \Big[\big(\mathscr{L}^n(X^+)\big)^{1/n} + \big(\mathscr{L}^n(Y^+)\big)^{1/n}\Big]^n + \Big[\big(\mathscr{L}^n(X^-)\big)^{1/n} + \big(\mathscr{L}^n(Y^-)\big)^{1/n}\Big]^n.$$

As we worked to get (3.0.5), we may use (3.0.6) to get $\mathscr{L}^n(Y^-) = (1-\lambda)\mathscr{L}^n(Y)$. This together with (3.0.6), (3.0.5) and the definition of λ when inserted in the last inequality imply

$$\begin{aligned}\mathscr{L}^n(X+Y) &\geq \lambda\Big[\big(\mathscr{L}^n(X)\big)^{1/n} + \big(\mathscr{L}^n(Y)\big)^{1/n}\Big]^n + (1-\lambda)\Big[\big(\mathscr{L}^n(X)\big)^{1/n} + \big(\mathscr{L}^n(Y)\big)^{1/n}\Big]^n \\ &= \Big[\big(\mathscr{L}^n(X)\big)^{1/n} + \big(\mathscr{L}^n(Y)\big)^{1/n}\Big]^n,\end{aligned}$$

which in turn gives (3.0.3). ■

---

[6] Note that the quotient on the left is a continuous function of $t \in \mathbb{R}$, that takes all values of $[0,1]$.

Chapter 4

Integral

4.1 Measurable functions

Unless otherwise stated, μ stands for a measure on a set X.

Definition 4.1.1 (μ-measurable function). [1] $f : X \to \mathbb{R}^m$ is called *μ-measurable* if

$$U \subseteq \mathbb{R}^m \text{ is open } \Rightarrow f^{-1}(U) \text{ is } \mu\text{-measurable.}$$

Definition 4.1.2 (Borel measurable function). [2] Let μ be a measure on $\mathbb{R}^n$. A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is called *Borel measurable* if

$$U \subseteq \mathbb{R}^m \text{ is open } \Rightarrow f^{-1}(U) \text{ is a Borel set.}$$

Example 4.1.3. If $f : \mathbb{R}^n \to \mathbb{R}^m$ is continuous then it is Borel measurable. This is because the pre-image $f^{-1}(U)$ of an open set $U \subseteq \mathbb{R}^m$ through a continuous function f is again open.

Theorem 4.1.4. (i) *If $f : X \to \mathbb{R}^m$ is μ-measurable, then $f^{-1}(B)$ is μ-measurable for each Borel set $B \subseteq \mathbb{R}^m$.*

(ii) *$f : X \to \bar{\mathbb{R}}$ is μ-measurable if and only if $f^{-1}\big([-\infty, \alpha)\big)$ is μ-measurable for each $\alpha \in \mathbb{R}$.*

(iii) *If $f : X \to \mathbb{R}^n$ and $g : X \to \mathbb{R}^m$ are μ-measurable, then $(f,g) : X \to \mathbb{R}^{n+m}$ is μ-measurable.*

[1] you can replace $\mathbb{R}^m$ in this definition by any topological space Y

[2] you can replace $\mathbb{R}^n, \mathbb{R}^m$ in this definition by any topological spaces X, Y

Proof. (i) By assumption the set

$$\mathscr{A} := \{A \subseteq \mathbb{R}^m \mid f^{-1}(A) \text{ is } \mu\text{-measurable}\},$$

contains all open sets of $\mathbb{R}^m$. If we show it is a σ-algebra then by definition of the Borel σ-algebra it will contain also all Borel sets. To this end notice first that $\emptyset, \mathbb{R}^m \in \mathscr{A}$. Also, if $A \in \mathscr{A}$ then by a property of the pre-image we have $f^{-1}(A^c) = \big(f^{-1}(A)\big)^c$, which is μ-measurable. Finally, if $\{A_k \in \mathscr{A}\}_{k\in\mathbb{N}}$, then again by a property of the pre-image we have $f^{-1}\big(\bigcup_{k\in\mathbb{N}} A_k\big) = \bigcup_{k\in\mathbb{N}} f^{-1}(A_k)$ which is μ-measurable.

(ii) Similarly, the set

$$\mathscr{A} := \{A \subseteq \bar{\mathbb{R}} \mid f^{-1}(A) \text{ is } \mu\text{-measurable}\},$$

is a σ-algebra containing $[-\infty, \alpha)$ for each $\alpha \in \mathbb{R}$.

(iii) One proves the set

$$\mathscr{A} := \{A \subseteq \mathbb{R}^{n+m} \mid h^{-1}(A) \text{ is } \mu\text{-measurable}\}, \quad \text{with } h := (f,g),$$

is a σ-algebra containing $U \times V$ whenever $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$ are open. ■

Exercise 4.1.5. Let $f,g : \mathbb{R}^n \to \mathbb{R}^n$ be Borel measurable. Prove $f \circ g$ is Borel measurable.

Theorem 4.1.6. (i) *If $f,g : X \to \bar{\mathbb{R}}$ are μ-measurable, then so are*

$$f+g, \ fg, \ |f|, \ \min\{f,g\}, \ \max\{f,g\} \ \textit{and } f/g \ (\textit{provided } g \neq 0 \textit{ in } X).$$

(ii) *If $\{f_k : X \to \bar{\mathbb{R}}\}_{k\in\mathbb{N}}$ are μ-measurable, then so are*

$$\inf_{k\in\mathbb{N}} f_k, \ \sup_{k\in\mathbb{N}} f_k, \ \liminf_{k\to\infty} f_k \ \textit{and} \ \limsup_{k\to\infty} f_k.$$

Exercise 4.1.7. Prove the above theorem.

Exercise 4.1.8. Let $E \subset \mathbb{R}^n$ be $\mathscr{L}^n$-measurable with $\mathscr{L}^n(E) < \infty$ and let $f : E \to \mathbb{R}$ be an $\mathscr{L}^n$-measurable function. Define $\mu_{E,f} : \mathbb{R} \to [0,\infty)$ by

$$\mu_{E,f}(s) := \mathscr{L}^n\big(\{x \in E \mid f(x) > s\}\big).$$

Prove the following assertions

(i) $\mu_{E,f}$ is bounded, non-increasing, $\mu_{E,f}(-\infty) = \mathscr{L}^n(E)$ and $\mu_{E,f}(\infty) = 0$.

(ii) $\mu_{E,f}(s) - \mu_{E,f}(t) = \mathscr{L}^n\big(\{x \in E \mid s < f(x) \leq t\}\big)$ whenever $s < t$.

(iii) $\lim_{s\to t^+} \mu_{E,f}(s) = \mu_{E,f}(t)$ (thus $\mu_{E,f}$ is right-continuous).

(iv) $\lim_{s\to t^-} \mu_{E,f}(s) = \mathscr{L}^n\big(\{x\in E \mid f(x)\geq t\}\big) = \mu_{E,f}(t) + \mathscr{L}^n\big(\{x\in E \mid f(x)=t\}\big)$ (thus $\mu_{E,f}$ is continuous at $t\in\mathbb{R}$ if and only if $\mathscr{L}^n\big(\{x\in E \mid f(x)=t\}\big)=0$).

(v) $\mu_{E,f}$ is constant in the interval (s,t) if and only if $\mathscr{L}^n\big(\{x\in E \mid s<f(x)<t\}\big)=0$.

Theorem 4.1.9. *If $f: X\to[0,\infty]$ is μ-measurable, there exist μ-measurable sets $\{A_k\subset X\}_{k\in\mathbb{N}}$ such that*

$$f(x) = \sum_{k\in\mathbb{N}} \frac{1}{k}\chi_{A_k}(x) \quad \forall\, x\in X.$$

Proof. We define the sets

$$A_1 := \{x\in X \mid f(x)\geq 1\} \quad\text{and}\quad A_k := \Big\{x\in X \mid f(x)\geq \frac{1}{k} + \sum_{j=1}^{k-1}\frac{1}{j}\chi_{A_j}(x)\Big\},\ \ k\in\mathbb{N}\setminus\{1\}.$$

First we observe

$$f(x)\geq \chi_{A_1}(x) \quad \forall\, x\in X.$$

Indeed, this is self evident if $x\in A_1$ and trivially true by the nonnegativity of f in case $x\in X\setminus A_1$. Next we claim

$$f(x)\geq \chi_{A_1}(x) + \frac{1}{2}\chi_{A_2}(x) \quad \forall\, x\in X. \tag{4.1.1}$$

Indeed, noting

$$A_2 = \{x\in X \mid f(x)\geq 3/2 \text{ if } x\in A_1 \text{ and } f(x)\geq 1/2 \text{ if } x\in X\setminus A_1\},$$

we see again that (4.1.1) is self evident if $x\in A_1\cap A_2$ and trivially true by the nonnegativity of f in case $x\in X\setminus(A_1\cap A_2)$. If $x\in A_2\setminus A_1$ then $1/2\leq f(x)<1$ and this gives $f(x)\geq 1/2$ which is (4.1.1) in this case. Finally if $x\in A_1\setminus A_2$ then $1\leq f(x)<3/2$ and this gives $f(x)\geq 1$ which is (4.1.1) in this case. In the same fashion we can prove inductively that for any $m\in\mathbb{N}$

$$f(x)\geq \sum_{j=1}^{m}\frac{1}{j}\chi_{A_j}(x) \quad \forall\, x\in X.$$

Letting $m\to\infty$ we obtain

$$f(x)\geq \sum_{k\in\mathbb{N}}\frac{1}{k}\chi_{A_k}(x) \quad \forall\, x\in X. \tag{4.1.2}$$

To show the equality we observe first that in case $f(x)=\infty$ we have $x\in A_k$ for all $k\in\mathbb{N}$. Thus $\sum_{k\in\mathbb{N}}(1/k)\chi_{A_k}(x) = \sum_{k\in\mathbb{N}}(1/k) = \infty = f(x)$ in this case. If $f(x)<\infty$, then there exists $k_x\in\mathbb{N}$ such that $x\notin A_k$ for all $k\geq k_x$. This means

$$f(x) < \frac{1}{k} + \sum_{j=1}^{k-1}\frac{1}{j}\chi_{A_j}(x) \quad \forall\, k\geq k_x,$$

which coupled with (4.1.2) gives

$$0 \leq f(x) - \sum_{j=1}^{k-1} \frac{1}{j} \chi_{A_j}(x) < \frac{1}{k} \quad \forall\, k \geq k_x. \tag{4.1.3}$$

The proof follows by letting $k \to \infty$. ■

4.2 Lusin and Egoroff theorems

Theorem 4.2.1 (Lusin). *Let μ be a Borel regular measure on $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^m$ be μ-measurable. Assume $A \subset \mathbb{R}^n$ is μ-measurable with $\mu(A) < \infty$. Then*

$$\forall\, \varepsilon > 0,\ \exists \text{ compact } K \subset A \text{ such that } \mu(A \setminus K) < \varepsilon \text{ and } f|_K \text{ is continuous.}$$

Proof. By theorem 1.1.26, given $i \in \mathbb{N}$, we may consider a partition of $\mathbb{R}^m$ comprised of dyadic cubes having diameter less than $1/i$. Actually, given $i \in \mathbb{N}$, any countable family $\{B_{i,j}\}_{j\in\mathbb{N}}$ comprised of disjoint Borel sets such that $\mathbb{R}^m = \bigcup_{j\in\mathbb{N}} B_{i,j}$ and $\operatorname{diam}(B_{i,j}) < 1/i$ will work in what follows. So pick such a partition and set

$$A_{i,j} := A \cap f^{-1}(B_{i,j}), \quad i, j \in \mathbb{N}.$$

These sets enjoy the following properties

(P_1) $\{A_{i,j}\}_{i,j\in\mathbb{N}}$ are μ-measurable. Indeed, since $\{B_{i,j}\}_{i,j\in\mathbb{N}}$ are Borel and f is μ-measurable, theorem 4.1.4-(i) implies $\{f^{-1}(B_{i,j})\}_{i,j\in\mathbb{N}}$ are μ-measurable. Since A is μ-measurable we conclude that $\{A_{i,j}\}_{i,j\in\mathbb{N}}$ are μ-measurable.

(P_2) $A = \bigcup_{j\in\mathbb{N}} A_{i,j}$ for any $i \in \mathbb{N}$. By definition and standard properties of the pre-image

$$\bigcup_{j\in\mathbb{N}} A_{i,j} = A \cap \Big(\bigcup_{j\in\mathbb{N}} f^{-1}(B_{i,j}) \Big) = A \cap f^{-1}\Big(\bigcup_{j\in\mathbb{N}} B_{i,j} \Big) = A \cap f^{-1}(\mathbb{R}^m) = A \cap \mathbb{R}^n = A.$$

(P_3) $\{A_{i,j}\}_{j\in\mathbb{N}}$ are disjoint for any $i \in \mathbb{N}$. Indeed, $x \in A_{i,j}$ if and only if $x \in A$ and there exists $y \in B_{i,j}$ such that $y = f(x)$. Assuming $x \in A_{i,j} \cap A_{i,k}$ with $j \neq k$, we get $y \in B_{i,j}$ and $y' \in B_{i,k}$ such that $y = f(x) = y'$. This contradicts the disjointness of $\{B_{i,j}\}_{j\in\mathbb{N}}$.

Hence, out of a partition $\{B_{i,j}\}_{j\in\mathbb{N}}$ of the target space $\mathbb{R}^m$, we have constructed a partition $\{A_{i,j}\}_{j\in\mathbb{N}}$ of $A \subseteq \mathbb{R}^n$ comprised of μ-measurable sets.

Now let $\varepsilon > 0$. Theorem 2.3.5-(iii) assures that $\mu|_A$ is Radon and because of (P_1) we may apply inner regularity (theorem 2.3.8) to get compact sets $\{K_{i,j} \subseteq A_{i,j}\}_{i,j\in\mathbb{N}}$ with $\mu|_A(A_{i,j} \setminus$

$K_{i,j}) < \varepsilon 2^{-i-j}$. Using (P_2) we observe $A \setminus \bigcup_{j\in\mathbb{N}} K_{i,j} = \bigcup_{j\in\mathbb{N}} \big(A_{i,j} \setminus \bigcup_{j\in\mathbb{N}} K_{i,j}\big) \subseteq \bigcup_{j\in\mathbb{N}}(A_{i,j} \setminus K_{i,j})$ and so

$$\begin{aligned}\mu\big(A \setminus \bigcup_{j\in\mathbb{N}} K_{i,j}\big) &= \mu|_A\big(A \setminus \bigcup_{j\in\mathbb{N}} K_{i,j}\big) \\ &\le \mu|_A\big(\bigcup_{j\in\mathbb{N}} (A_{i,j} \setminus K_{i,j})\big) < \varepsilon 2^{-i},\end{aligned}$$

by the subadditivity of measures. Noting $\{A \setminus \bigcup_{j=1}^{N} K_{i,j}\}_{N\in\mathbb{N}}$ non-increases to $A \setminus \bigcup_{j\in\mathbb{N}} K_{i,j}$, we use theorem 2.2.5-(ii) to deduce further from the above estimate

$$\lim_{N\to\infty} \mu\big(A \setminus \bigcup_{j=1}^{N} K_{i,j}\big) < \varepsilon 2^{-i} \quad \forall\, i \in \mathbb{N}.$$

This implies for each $i \in \mathbb{N}$ there exists $N_i \in \mathbb{N}$ such that

$$\mu\big(A \setminus \bigcup_{j=1}^{N_i} K_{i,j}\big) < \varepsilon 2^{-i}. \tag{4.2.1}$$

Next we define the sets

$$D_i := \bigcup_{j=1}^{N_i} K_{i,j}, \quad i \in \mathbb{N}.$$

Being a finite union of compact sets, each D_i is compact. For each $i \in \mathbb{N}$ we pick $\{b_{i,j} \in B_{i,j}\}_{j\in\mathbb{N}}$ and define functions $g_i : D_i \to \mathbb{R}^m$ by

$$g_i(x) := \sum_{j=1}^{N_i} b_{i,j}\chi_{K_{i,j}}(x), \quad x \in D_i.$$

(P_3) together with $\{K_{i,j} \subseteq A_{i,j}\}_{i,j\in\mathbb{N}}$ imply $\{K_{i,j}\}_{i,j\in\mathbb{N}}$ are disjoint and thus, each being compact, of positive distance apart (use exercise 1.1.14). This means each g_i is continuous. Also, if $x \in D_i$, then $x \in K_{i,j} \subseteq A_{i,j}$ for some $j \in \{1,...,N_i\}$. Thus $f(x) \in B_{i,j}$ for some $j \in \{1,...,N_i\}$. Since $g_i(x) = b_{i,j} \in B_{i,j}$ for this j, we obtain

$$|f(x) - g_i(x)| \le \operatorname{diam}(B_{i,j}) < \frac{1}{i} \quad \forall\, x \in D_i. \tag{4.2.2}$$

Finally we define the set

$$K := \bigcap_{i\in\mathbb{N}} D_i.$$

Then K is a compact subset of A and $A \setminus K = A \cap \left(\bigcup_{i\in\mathbb{N}} D_i^c\right) = \bigcup_{i\in\mathbb{N}}(A \cap D_i^c)$. Subadditivity of measures gives

$$\mu(A \setminus K) \le \sum_{i\in\mathbb{N}} \mu(A \setminus D_i) < \varepsilon,$$

by the definition of D_i and (4.2.1). Also, if $x \in K$ then $x \in D_i$ for all $i \in \mathbb{N}$ and because of (4.2.2)

$$\sup_{x\in K} |f(x) - g_i(x)| < \frac{1}{i}.$$

Thus $\{g_i : D_i \to \mathbb{R}^m\}_{i\in\mathbb{N}}$ is a sequence of continuous functions which converges uniformly to f on $K = \bigcap_{i\in\mathbb{N}} D_i$. Standard analysis implies the limit function f is continuous in K.[3] ■

Notation 4.2.2. Let μ be a measure on a set X. The expression "μ-a.e. in A" where $A \subseteq X$ means "almost everywhere in A with respect to the measure μ"; that is, "for all $x \in A \setminus N$ where $\mu(N) = 0$".

Theorem 4.2.3 (Egoroff). *Let μ be a measure on X and suppose $\{f_k \mid X \to \mathbb{R}^m\}_{k\in\mathbb{N}}$ are μ-measurable. Let $A \subset X$ be a μ-measurable set such that $\mu(A) < \infty$ and $f_k \to f$ μ-a.e. in A. Then*

$$\forall\, \varepsilon > 0,\ \exists\ \mu\text{-measurable } B \subset A \text{ such that } \mu(A \setminus B) < \varepsilon \text{ and } f_k \to f \text{ uniformly in } B.$$

Proof. Let $N \subseteq A$ be the set of μ-measure 0 out of which $f_k \to f$. For each $i \in \mathbb{N}$ we define

$$C_{i,j} := \bigcup_{k\ge j} \left\{x \in X \mid |f_k(x) - f(x)| > 2^{-i}\right\}, \quad j \in \mathbb{N}.$$

Then $\{A \cap C_{i,j}\}_{j\in\mathbb{N}}$ are μ-measurable (by theorem 4.1.6-(ii)) and non-increase to $\bigcap_{j\in\mathbb{N}} \left(A \cap C_{i,j}\right)$. But $f_k \to f$ μ-a.e. in A implies $\bigcap_{j\in\mathbb{N}} \left(A \cap C_{i,j}\right) \subseteq N$. Taking into account the fact that $\mu(A \cap C_{i,1}) \le \mu(A) < \infty$, we apply theorem 2.2.5-(iii) to get

$$\lim_{j\to\infty} \mu(A \cap C_{i,j}) \le \mu(N) = 0 \quad \forall\, i \in \mathbb{N}.$$

Hence, for each $i \in \mathbb{N}$ there exists $N_i \in \mathbb{N}$ such that $\mu(A \cap C_{i,N_i}) < \varepsilon 2^{-i}$.

We define next the set

$$B := A \setminus \bigcup_{i\in\mathbb{N}} C_{i,N_i}.$$

Observing that $A \setminus B = A \cap \left(A^c \cup \left(\bigcup_{i\in\mathbb{N}} C_{i,N_i}\right)\right) = A \cap \left(\bigcup_{i\in\mathbb{N}} C_{i,N_i}\right) = \bigcup_{i\in\mathbb{N}}(A \cap C_{i,N_i})$ we get

$$\mu(A \setminus B) \le \sum_{i\in\mathbb{N}} \mu(A \cap C_{i,N_i}) < \varepsilon.$$

Moreover, since for each $i \in \mathbb{N}$ we have $x \in B \Rightarrow x \notin A \cap C_{i,N_i}$, we get by the definition of the $C_{i,j}$'s that for each $i \in \mathbb{N}$, any $x \in B$ and all $k \ge N_i$, there holds $|f_k(x) - f(x)| \le 2^{-i}$. ■

[3] $|f(x) - f(x_0)| \le |f(x) - g_i(x)| + |g_i(x) - g_i(x_0)| + |g_i(x_0) - f(x_0)| \le |g_i(x) - g_i(x_0)| + 2\max_{x\in K} |f(x) - g_i(x)|\ \forall x, x_0 \in K$

4.3 Integration of measurable functions

Notation 4.3.1. For a function $g : X \to \bar{\mathbb{R}}$ we set

$$g_+ := \max\{0, g\} \quad \text{and} \quad g_- := \max\{0, -g\}.$$

Observe these are nonnegative functions and that the following decompositions are valid

$$g = g_+ - g_- \quad \text{and} \quad |g| = g_+ + g_-.$$

Definition 4.3.2 (simple function). A function $g : X \to \bar{\mathbb{R}}$ is called *simple function* if the image of g is countable.

Remark 4.3.3. The above definition implies that if $g : X \to \bar{\mathbb{R}}$ is simple, then there exist disjoint $\{A_k \subseteq X\}_{k\in\mathbb{N}}$ and $\{\alpha_k \in \bar{\mathbb{R}}\}_{k\in\mathbb{N}}$ such that

$$g = \sum_{k\in\mathbb{N}} \alpha_k \chi_{A_k} \quad \text{and} \quad \bigcup_{k\in\mathbb{N}} A_k = X.$$

Consider now μ to be a measure on the set X. The sets $\{A_k\}_{k\in\mathbb{N}}$ in the above expression can be taken μ-measurable if g is known to be μ-measurable. Indeed, for each $k \in \mathbb{N}$ we have

$$\begin{aligned} A_k = g^{-1}\big(\{\alpha_k\}\big) &= g^{-1}\big([-\infty, \alpha_k]\big) \cap g^{-1}\big([\alpha_k, \infty)\big) \\ &= \bigcap_{m\in\mathbb{N}} g^{-1}\big([-\infty, \alpha_k + 1/m)\big) \cap \big\{g^{-1}\big([-\infty, \alpha_k)\big)\big\}^c, \end{aligned}$$

and thus by theorem 4.1.4-(ii) A_k is μ-measurable. Consequently, we have the useful fact[4]

Proposition 4.3.4. *Given a μ-measurable simple function $g : X \to [0, \infty]$, then there exist disjoint μ-measurable sets $\{A_k\}_{k\in\mathbb{N}}$ and $\{\alpha_k \in (0, \infty]\}_{k\in\mathbb{N}}$ such that*

$$g = \sum_{k\in\mathbb{N}} \alpha_k \chi_{A_k} \quad \textit{and} \quad \bigcup_{k\in\mathbb{N}} A_k \subseteq X.$$

Definition 4.3.5. (i) If $g : X \to [0, \infty]$ is a simple, μ-measurable function, we define its *integral* on X by

$$\int_X g \,\mathrm{d}\mu := \sum_{y\in[0,\infty]} y\mu\big(g^{-1}\{y\}\big).$$

Note that applying proposition 4.3.4 on g we get $\int_X g \,\mathrm{d}\mu = \sum_{k\in\mathbb{N}} \alpha_k \mu(A_k)$.

[4] compare with theorem 4.1.9

(ii) If $g : X \to \bar{\mathbb{R}}$ is a simple μ-measurable function and either $\int_X g_+ \,\mathrm{d}\mu < \infty$ or $\int_X g_- \,\mathrm{d}\mu < \infty$, we call g a *μ-integrable simple function* and define its integral on X by

$$\int_X g \,\mathrm{d}\mu := \int_X g_+ \,\mathrm{d}\mu - \int_X g_- \,\mathrm{d}\mu.$$

Definition 4.3.6. Let $f : X \to \bar{\mathbb{R}}$.

(i) The *upper integral* of f on X is defined by

$$\overline{\int}_X f \,\mathrm{d}\mu := \inf\Big\{ \int_X g \,\mathrm{d}\mu \;\Big|\; g \ \mu\text{-integrable, simple, } g \geq f \ \mu\text{-a.e. in } X \Big\}.$$

(ii) The *lower integral* of f on X is defined by

$$\underline{\int}_X f \,\mathrm{d}\mu := \sup\Big\{ \int_X g \,\mathrm{d}\mu \;\Big|\; g \ \mu\text{-integrable, simple, } g \leq f \ \mu\text{-a.e. in } X \Big\}.$$

Theorem 4.3.7. *Let $f, h : X \to \bar{\mathbb{R}}$.*

(i) $\underline{\int}_X f \,\mathrm{d}\mu = -\overline{\int}_X (-f)\mathrm{d}\mu$.

(ii) *If $f \geq 0$ μ-a.e. in X, then $\overline{\int}_X f \,\mathrm{d}\mu \geq 0$.*

(iii) *If $\alpha > 0$ then $\overline{\int}_X \alpha f \,\mathrm{d}\mu = \alpha \overline{\int}_X f \,\mathrm{d}\mu$.*

(iv) *If $\overline{\int}_X f \,\mathrm{d}\mu + \overline{\int}_X h \,\mathrm{d}\mu < \infty$, then $\overline{\int}_X (f+h)\mathrm{d}\mu \leq \overline{\int}_X f \,\mathrm{d}\mu + \overline{\int}_X h \,\mathrm{d}\mu$.*

(v) $\underline{\int}_X f \,\mathrm{d}\mu \leq \overline{\int}_X f \,\mathrm{d}\mu$.

Exercise 4.3.8. Prove the above theorem.

Definition 4.3.9 (μ-integrable function). A μ-measurable function $f : X \to \bar{\mathbb{R}}$ is called *μ-integrable* if $\overline{\int}_X f \,\mathrm{d}\mu \leq \underline{\int}_X f \,\mathrm{d}\mu$. By 4.3.7-(v) we have then $\overline{\int}_X f \,\mathrm{d}\mu = \underline{\int}_X f \,\mathrm{d}\mu$ and we write

$$\int_X f \,\mathrm{d}\mu = \overline{\int}_X f \,\mathrm{d}\mu = \underline{\int}_X f \,\mathrm{d}\mu.$$

Remark 4.3.10. For a μ-integrable simple function $g : X \to \bar{\mathbb{R}}$ we have

$$\int_X g \,\mathrm{d}\mu = \sum_{y \in \bar{\mathbb{R}}} y\mu\big(g^{-1}\{y\}\big).$$

Theorem 4.3.11. *Any μ-measurable $f : X \to [0,\infty]$ is μ-integrable with $\int_X f \,\mathrm{d}\mu \in [0,\infty]$.*

Proof. Suppose first that $\mu\big(\{x\in X \mid f(x)=\infty\}\big)>0$. For $t>0$ we use the simple function $t\chi_{\{x\in X \mid f(x)=\infty\}}$ in the definition of $\underline{\int}_X f\ \mathrm{d}\mu$ to obtain

$$\underline{\int}_X f\ \mathrm{d}\mu \geq t\mu\big(\{x\in X \mid f(x)=\infty\}\big) \text{ for any } t>0.$$

Thus $\underline{\int}_X f\ \mathrm{d}\mu=\infty$. Because of theorem 4.3.7-(v) we also get $\overline{\int}_X f\ \mathrm{d}\mu=\infty$. Hence f is μ-integrable and $\int_X f \mathrm{d}\mu=\infty$. Suppose now $\mu\big(\{x\in X \mid f(x)=\infty\}\big)=0$. Then $f(x)<\infty$ μ-a.e. in X and for $t>1$ we consider the disjoint μ-measurable sets

$$E_k:=\{x\in X \bigm| t^k\leq f(x)<t^{k+1}\},\quad k\in\mathbb{Z},$$

as well as the simple μ-measurable function

$$g:=\sum_{k\in\mathbb{Z}} t^k\chi_{E_k}.$$

Then $X\setminus\{f=0\}=\bigcup_{k\in\mathbb{Z}}E_k$ and assigning the value 0 to g in $\{f=0\}$, we have $g(x)\leq f(x)\leq tg(x)$ μ-a.e. $x\in X$. From the definitions of upper and lower integrals of f we deduce

$$\overline{\int}_X f\ \mathrm{d}\mu \leq \int_X tg(x)\ \mathrm{d}\mu = t\int_X g(x)\ \mathrm{d}\mu \leq t\underline{\int}_X f(x)\ \mathrm{d}\mu,$$

where we have used 4.3.7-(iii) to get the equality. This holds for any $t>1$ and so $\overline{\int}_X f\ \mathrm{d}\mu\leq \underline{\int}_X f(x)\ \mathrm{d}\mu$. ■

Theorem 4.3.12. *Let $f,h: X\to\bar{\mathbb{R}}$.*

(i) *If $\alpha\neq 0$ then $\int_X \alpha f\ \mathrm{d}\mu=\alpha\int_X f\ \mathrm{d}\mu$.*

(ii) *If $\int_X f\ \mathrm{d}\mu+\int_X h\ \mathrm{d}\mu\in\bar{\mathbb{R}}$, then $\int_X(f+h)\ \mathrm{d}\mu=\int_X f\ \mathrm{d}\mu+\int_X h\ \mathrm{d}\mu$.*

(iii) *If f is μ-integrable $|\int_X f\ \mathrm{d}\mu|\leq\int_X|f|\ \mathrm{d}\mu$.*

Exercise 4.3.13. Prove the above theorem.

Proposition 4.3.14. *Let μ be a Radon measure on $\mathbb{R}^n$ and suppose $A\subset\mathbb{R}^n$ is μ-measurable with $\mu(A)<\infty$. Then, given any $p>0$ we have*

$$\forall\ \varepsilon>0,\ \exists\ g\in C_c(\mathbb{R}^n) \textit{ such that } \int_{\mathbb{R}^n}|\chi_A-g|^p\ \mathrm{d}\mu<\varepsilon.$$

.

Proof. Given $\varepsilon > 0$ we employ the outer and inner regularity properties of Radon measures (theorems 2.3.7 and 2.3.8), to get open $U \supset A$ and compact $K \subset A$ such that

$$\mu(U) - \varepsilon < \mu(A) < \mu(K) + \varepsilon.$$

Hence $\mu(U \setminus K) < 2\varepsilon$. Further, from lemma 1.2.6 we get a continuous function $g : \mathbb{R}^n \to [0,1]$ with $g = 1$ on K and $\operatorname{supp}(g) \subset U$. It follows that $|\chi_A - g| = 0$ in $(U \setminus K)^c$ and $|\chi_A - g| \leq 1$ on $U \setminus K$. Altogether,

$$\int_{\mathbb{R}^n} |\chi_A - g|^p \, \mathrm{d}\mu = \int_{U \setminus K} |\chi_A - g|^p \, \mathrm{d}\mu \leq \mu(U \setminus K) \leq 2\varepsilon.$$

Rescaling ε the proof is complete. ■

Definition 4.3.15 (**μ-summable function).** $f : X \to \bar{\mathbb{R}}$ is called *μ-summable* if it is μ-integrable and $\int_X |f| \mathrm{d}\mu < \infty$.

Exercise 4.3.16. Prove that if $f : X \to \bar{\mathbb{R}}$ is μ-summable then

$$\mu\big(\{x \in X \mid |f(x)| \geq s\}\big) \leq \frac{1}{s} \int_X |f| \, \mathrm{d}\mu \qquad \text{whenever } s > 0.$$

Use this to prove that $\int_X |f| \mathrm{d}\mu = 0$ if and only if $f = 0$ μ-a.e. in X.

Chapter 5

The basic theorems of advanced analysis

5.1 Limit theorems

Theorem 5.1.1 (Fatou). *Let* $f_k : X \to [0,\infty]$, $k \in \mathbb{N}$ *be* μ*-measurable. Then*

$$\int_X \liminf_{k\to\infty} f_k \, \mathrm{d}\mu \le \liminf_{k\to\infty} \int_X f_k \, \mathrm{d}\mu.$$

Proof. We assume $\liminf_{k\to\infty} \int_X f_k \, \mathrm{d}\mu < \infty$, otherwise the thesis is immediate. Theorem 4.1.6-(ii) says $\liminf_{k\to\infty} f_k : X \to [0,\infty]$ is μ-measurable. Its nonnegativity implies through theorem 4.3.11 that it is also μ-integrable. In particular

$$\int_X \liminf_{k\to\infty} f_k \, \mathrm{d}\mu = \sup_{g\in\mathscr{A}} \int g \, \mathrm{d}\mu, \tag{5.1.1}$$

where

$$\mathscr{A} := \big\{ g \ \mu\text{-integrable, simple, } g \le \liminf_{k\to\infty} f_k \ \mu\text{-a.e. in } X \big\}.$$

Let $g \in \mathscr{A}$. By proposition 4.3.4 we may assume $g = \sum_{j\in\mathbb{N}} \alpha_j \chi_{A_j}$, for some $\{\alpha_j > 0\}_{j\in\mathbb{N}}$ and disjoint μ-measurable $\{A_j \subseteq X\}_{j\in\mathbb{N}}$ with $\bigcup_{j\in\mathbb{N}} A_j \subseteq X$.

Claim. Let $t \in (0,1)$. Then for each $j \in \mathbb{N}$ we have $A_j = \bigcup_{k\in\mathbb{N}} B_{j,k}$, where

$$B_{j,k} := A_j \cap \{x \in X \mid f_\ell(x) > t\alpha_j \ \ \forall\, \ell \ge k\}.$$

Proof of claim. Let $j \in \mathbb{N}$. We only prove $A_j \subseteq \bigcup_{k\in\mathbb{N}} B_{j,k}$, the reverse inclusion being self evident. Since $g \in \mathscr{A}$ we know $\alpha_j \le \liminf_{k\to\infty} f_k$ μ-a.e. in A_j. Hence

$$t\alpha_j < \lim_{l\to\infty} \inf_{k\ge\ell} f_k \quad \mu\text{-a.e. in } A_j.$$

This implies the existence of $k_0 \in \mathbb{N}$ such that $t\alpha_j < \inf_{k\geq\ell} f_k$ μ-a.e. in A_j, for all $\ell \geq k_0$, which in turn gives $t\alpha_j < f_\ell$ μ-a.e. in A_j, for all $\ell \geq k_0$. □

Using $\{A_j\}_{j\in\mathbb{N}}$ are disjoint with $\bigcup_{j\in\mathbb{N}} A_j \subseteq X$ and then $A_j \supseteq B_{j,k}$ for all $k \in \mathbb{N}$, we get for any $m \in \mathbb{N}$

$$\begin{aligned}\int_X f_k \, \mathrm{d}\mu &\geq \sum_{j=1}^{m} \int_{A_j} f_k \, \mathrm{d}\mu \\ &\geq \sum_{j=1}^{m} \int_{B_{j,k}} f_k \, \mathrm{d}\mu \geq t \sum_{j=1}^{m} \alpha_j \mu(B_{j,k}),\end{aligned}$$

the last inequality coming straight from the definition of $B_{j,k}$. Observing next $B_{j,k+1} \supseteq B_{j,k}$ for all $k \in \mathbb{N}$ and taking the $\liminf_{k\to\infty}$, we arrive at

$$\begin{aligned}\liminf_{k\to\infty} \int_X f_k \, \mathrm{d}\mu &\geq t \sum_{j=1}^{m} \alpha_j \lim_{k\to\infty} \mu(B_{j,k}) \\ &= t \sum_{j=1}^{m} \alpha_j \mu(\cup_{k\in\mathbb{N}} B_{j,k}) = t \sum_{j=1}^{m} \alpha_j \mu(A_j),\end{aligned}$$

the penultimate equality coming from theorem 2.2.5-(ii). Consequently, letting $m \to \infty$,

$$\liminf_{k\to\infty} \int_X f_k \, \mathrm{d}\mu \geq t \int_X g \, \mathrm{d}\mu \quad \forall\, t \in (0,1), \ \forall\, g \in \mathscr{A},$$

which implies

$$\liminf_{k\to\infty} \int_X f_k \, \mathrm{d}\mu \geq \sup_{g\in\mathscr{A}} \int_X g \, \mathrm{d}\mu.$$

Coupling this with (5.1.1) gives the desired estimate. ■

Theorem 5.1.2 (monotone convergence - Beppo Levi). *Let $f_k : X \to [0,\infty]$, $k \in \mathbb{N}$ be μ-measurable such that $f_k \leq f_{k+1}$ μ-a.e. in X, for all $k \in \mathbb{N}$. Then*

$$\int_X \lim_{k\to\infty} f_k \, \mathrm{d}\mu = \lim_{k\to\infty} \int_X f_k \, \mathrm{d}\mu.$$

Proof. Fatou's lemma readily gives

$$\int_X \lim_{k\to\infty} f_k \, \mathrm{d}\mu \leq \liminf_{k\to\infty} \int_X f_k \, \mathrm{d}\mu. \tag{5.1.2}$$

On the other hand, $f_k \leq \lim_{k\to\infty} f_k$ μ-a.e. in X for all $k \in \mathbb{N}$. This implies

$$\int_X f_k \, \mathrm{d}\mu \leq \int_X \lim_{k\to\infty} f_k \, \mathrm{d}\mu \quad \forall\, k \in \mathbb{N},$$

and so

$$\limsup_{k\to\infty}\int_X f_k \,\mathrm{d}\mu \le \int_X \lim_{k\to\infty} f_k \,\mathrm{d}\mu. \tag{5.1.3}$$

The proof follows now from (5.1.2) and (5.1.3). ■

Exercise 5.1.3. Let $f, f_k : X \to \bar{\mathbb{R}}$, $k \in \mathbb{N}$, be μ-summable and satisfy

$$\lim_{k\to\infty}\int_X |f_k - f| \,\mathrm{d}\mu = 0.$$

Prove there exists a subsequence $\{f_{l_k}\}_{k\in\mathbb{N}}$ converging to f μ-a.e. in X.

[Hint: Setting $\alpha_k := \int_X |f_k - f|\mathrm{d}\mu$, extract a subsequence $\{\alpha_{\ell_k}\}_{k\in\mathbb{N}}$ such that $\sum_{k\in\mathbb{N}}\alpha_{\ell_k} \le 1$. Apply then the monotone convergence theorem to the sequence $g_k := \sum_{j=1}^{k} |f_{\ell_j} - f|$.]

Theorem 5.1.4 (absolute continuity of integrals - Vitali). *If $f : X \to \bar{\mathbb{R}}$ is μ-summable then*

$$\forall\, \varepsilon > 0,\ \ \exists\, \delta > 0 \textit{ such that if } A \subseteq X \textit{ is } \mu\textit{-measurable with } \mu(A) < \delta, \textit{ then } \int_A |f| \,\mathrm{d}\mu < \varepsilon.$$

Proof. We define a sequence of functions $g_k : X \to [0,\infty]$, $k \in \mathbb{N}$, by truncating $|f|$ as follows

$$g_k(x) := \begin{cases} |f(x)| & \text{if } |f(x)| < k, \\ k & \text{if } |f(x)| \ge k. \end{cases}$$

The sequence $\{g_k\}_{k\in\mathbb{N}}$ enjoys the following properties

(i) $g_k \le \min\{k, |f|\}$ μ-a.e. in X, for all $k \in \mathbb{N}$,

(ii) $g_k \to |f|$ as $k \to \infty$, μ-a.e. in X,

(iii) $g_k \le g_{k+1}$ μ-a.e. in X, for all $k \in \mathbb{N}$.

Indeed, for (i) observe that $g_k \le k$ and $g_k \le |f|$ both μ-a.e. in X. (ii) is obvious. For (iii) notice in case $|f(x)| < k$ that $g_k(x) = |f(x)| = g_{k+1}(x)$, in case $k \le |f(x)| < k+1$ that $g_k(x) = k \le |f(x)| = g_{k+1}(x)$ and in case $|f(x)| \ge k+1$ that $g_k(x) = k < k+1 = g_{k+1}(x)$.

Because of (ii) and (iii) we get through the monotone convergence theorem

$$\lim_{k\to\infty}\int_X g_k \,\mathrm{d}\mu = \int_X |f| \,\mathrm{d}\mu;$$

that is, given $\varepsilon > 0$, there exists $k_\varepsilon \in \mathbb{N}$ such that

$$\left|\int_X g_k \,\mathrm{d}\mu - \int_X |f| \,\mathrm{d}\mu\right| < \varepsilon \quad \text{for all } k \ge k_\varepsilon.$$

But (i) implies the difference in the absolute value is non-positive and so

$$\int_X |f| \, \mathrm{d}\mu < \int_X g_{k_\varepsilon} \, \mathrm{d}\mu + \varepsilon.$$

Let $\delta := \varepsilon / k_\varepsilon$. For any μ-measurable $A \subseteq X$ with $\mu(A) < \delta$ we rewrite the above inequality

$$\int_A |f| \, \mathrm{d}\mu < \int_A g_{k_\varepsilon} \, \mathrm{d}\mu + \int_{A^c} (g_{k_\varepsilon} - |f|) \, \mathrm{d}\mu + \varepsilon.$$

Property (i) implies the second integral on the right is non-positive and also the first integral does not exceed $k_\varepsilon \mu(A)$. Hence

$$\int_A |f| \, \mathrm{d}\mu < k_\varepsilon \delta + \varepsilon = 2\varepsilon.$$

Rescaling ε the proof is complete. ■

Theorem 5.1.5 (dominated convergence - Lebesgue). *Let $f, f_k : X \to \bar{\mathbb{R}}$, $k \in \mathbb{N}$, be μ-measurable, $g : X \to [0, \infty]$ be μ-summable, satisfying*

(i) *$f_k \to f$ as $k \to \infty$, μ-a.e. in X,*

(ii) *$|f_k| \le g$ μ-a.e. in X, for all $k \in \mathbb{N}$.*

Then

$$\lim_{k \to \infty} \int_X |f_k - f| \, \mathrm{d}\mu = 0. \tag{5.1.4}$$

Proof. First notice that (ii) and (i) imply $|f| \le g$ μ-a.e. in X. This and (ii) again give through the triangle inequality

$$2g - |f_k - f| \ge 2g - |f_k| - |f| \ge 0 \;\; \mu\text{-a.e. in } X \quad \forall\, k \in \mathbb{N}.$$

Applying Fatou's lemma to the sequence defined by the left hand side gives

$$\int_X \liminf_{k \to \infty} \big(2g - |f_k - f|\big) \, \mathrm{d}\mu \le \liminf_{k \to \infty} \int_X \big(2g - |f_k - f|\big) \, \mathrm{d}\mu,$$

and using (i) once more together with the summability of g and exercise 1.1.2

$$\int_X 2g \, \mathrm{d}\mu \le \int_X 2g \, \mathrm{d}\mu - \limsup_{k \to \infty} \int_X |f_k - f| \, \mathrm{d}\mu.$$

Cancelling the common integral on both sides we get

$$\limsup_{k \to \infty} \int_X |f_k - f| \, \mathrm{d}\mu \le 0,$$

which yields (5.1.4). ■

Exercise 5.1.6. Prove

$$\lim_{k\to\infty}\int_0^1 \left(1-e^{-x^2/k}\right)x^{-1/2}\,\mathrm{d}x = 0.$$

Exercise 5.1.7. Let $f, f_k : X \to \bar{\mathbb{R}}$, $k \in \mathbb{N}$, be μ-measurable and $g, g_k : X \to [0,\infty]$ be μ-summable, satisfying

(i) $f_k \to f$ as $k \to \infty$, μ-a.e. in X,

(ii) $g_k \to g$ as $k \to \infty$, μ-a.e. in X,

(iii) $\lim_{k\to\infty}\int_X g_k\,\mathrm{d}\mu = \int_X g\,\mathrm{d}\mu$, and

(iv) $|f_k| \le g_k$ μ-a.e. in X, for all $k \in \mathbb{N}$.

Prove $\lim_{k\to\infty}\int_X |f_k - f|\,\mathrm{d}\mu = 0$.

Exercise 5.1.8. Compute the following limit

$$\lim_{k\to\infty}\int_0^\infty \frac{\sin(x/k)}{(1+x/k)^k}\,\mathrm{d}x.$$

Theorem 5.1.9 (missing term in Fatou's lemma - Brezis and Lieb). *Let $f, f_k : X \to \bar{\mathbb{R}}$, $k \in \mathbb{N}$, be μ-measurable satisfying*

(i) *$f_k \to f$ as $k \to \infty$, μ-a.e. in X, and*

(ii) *$M := \sup_{k\in\mathbb{N}}\int_X |f_k|^p\,\mathrm{d}\mu < \infty$ for some $p > 0$.*

Then

$$\lim_{k\to\infty}\int_X \left||f_k|^p - |f_k - f|^p - |f|^p\right|\,\mathrm{d}\mu = 0.$$

Remark 5.1.10. Assumption (i) implies through the triangle inequality that $|f_k| \to |f|$ as $k \to \infty$, μ-a.e. in X. So we further deduce $|f_k|^p \to |f|^p$ as $k \to \infty$, μ-a.e. in X. Applying now Fatou's lemma to $|f_k|^p$ gives

$$\int_X |f|^p\,\mathrm{d}\mu \le M. \tag{5.1.5}$$

Remark 5.1.11. The conclusion of the theorem implies (use theorem 4.3.12-(iv))

$$\lim_{k\to\infty}\int_X \left(|f_k|^p - |f_k - f|^p - |f|^p\right)\,\mathrm{d}\mu = 0.$$

This together with the above remark imply

$$\int_X |f|^p\,\mathrm{d}\mu = \lim_{k\to\infty}\int_X \left(|f_k|^p - |f_k - f|^p\right)\,\mathrm{d}\mu, \tag{5.1.6}$$

but we don't know if $\lim_{k\to\infty}\int_X |f_k|^p\,\mathrm{d}\mu = \int_X |f|^p\,\mathrm{d}\mu$ or $\lim_{k\to\infty}\int_X |f_k - f|^p\,\mathrm{d}\mu = 0$.

Remark 5.1.12. Rewriting (5.1.6) as

$$\int_X |f|^p \, \mathrm{d}\mu = \liminf_{k\to\infty} \Big(\int_X |f_k|^p \, \mathrm{d}\mu - \int_X |f_k - f|^p \, \mathrm{d}\mu \Big),$$

and comparing this with Fatou's lemma

$$\int_X |f|^p \, \mathrm{d}\mu \le \liminf_{k\to\infty} \int_X |f_k|^p \, \mathrm{d}\mu,$$

justifies the "missing term in Fatou's lemma".

Proof. It is enough to prove the following statement: *Let $g_k : X \to \bar{\mathbb{R}}$, $k \in \mathbb{N}$, be μ-measurable satisfying*

(I) $g_k \to 0$ *as* $k \to \infty$, μ*-a.e. in* X, *and*

(II) $\tilde{M} := \sup_{k\in\mathbb{N}} \int_X |g_k|^p \, \mathrm{d}\mu < \infty$ *for some* $p > 0$.

Then for any μ*-measurable* $g : X \to \bar{\mathbb{R}}$ *such that* $\int_X |g|^p \mathrm{d}\mu < \infty$, *there holds*

$$\lim_{k\to\infty} \int_X \big| |g_k + g|^p - |g_k|^p - |g|^p \big| \, \mathrm{d}\mu = 0. \tag{5.1.7}$$

Indeed, let f_k, f satisfy the hypotheses of the theorem and take $g_k := f_k - f$, $g := f$ in the above statement (the choice for g is eligible because of remark 5.1.10). Then the elementary inequality [1] $\big(|a| + |b|\big)^p \le 2^p\big(|a|^p + |b|^p\big)$ gives

$$\begin{aligned} \int_X |g_k|^p \, \mathrm{d}\mu &\le \int_X \big(|f_k| + |f|\big)^p \, \mathrm{d}\mu \\ &\le 2^p \Big(\int_X |f_k|^p \, \mathrm{d}\mu + \int_X |f|^p \, \mathrm{d}\mu \Big) \le 2^{p+1} M, \end{aligned}$$

by (ii) and (5.1.5). This establishes (II) with $\tilde{M} = 2^{p+1} M$.

Now let $\varepsilon > 0$. We define $\{G_{\varepsilon,k} : X \to \bar{\mathbb{R}}\}_{k\in\mathbb{N}}$, by

$$G_{\varepsilon,k} := \big| |g_k + g|^p - |g_k|^p - |g|^p \big| - \varepsilon |g_k|^p.$$

Claim.[2] $\forall\, \varepsilon > 0$, $\exists\, C_\varepsilon > 0$ such that $\big| |a+b|^p - |a|^p \big| \le \varepsilon |a|^p + C_\varepsilon |b|^p \;\; \forall\, a, b \in \mathbb{R}$.

Accepting this claim for the moment, we readily have

$$\begin{aligned} \big| |g_k + g|^p - |g_k|^p - |g|^p \big| &\le \big| |g_k + g|^p - |g_k|^p \big| + |g|^p \\ &\le \varepsilon |g_k|^p + C_\varepsilon |g|^p + |g|^p. \end{aligned}$$

[1] $\big(|a| + |b|\big)^p \le \big(2\max\{|a|, |b|\}\big)^p = 2^p \max\{|a|^p, |b|^p\} \le 2^p\big(|a|^p + |b|^p\big)$

[2] it's proof will be an easy exercise (see remark 7.1.5) as soon as we learn about convex functions

Thus $G_{\varepsilon,k} \leq (C_\varepsilon + 1)|g|^p$ and taking positive parts

$$\big(G_{\varepsilon,k}\big)_+ \leq (C_\varepsilon + 1)|g|^p.$$

The right hand side of this is μ-summable by assumption. In addition, $G_{\varepsilon,k} \to 0$ as $k \to \infty$, μ-a.e. in X and so

$$\big(G_{\varepsilon,k}\big)_+ \to 0 \text{ as } k \to \infty, \ \ \mu\text{-a.e. in } X.$$

Therefore the dominated convergence theorem applies to $\big(G_{\varepsilon,k}\big)_+$ to deduce

$$\lim_{k\to\infty} \int_X \big(G_{\varepsilon,k}\big)_+ \,\mathrm{d}\mu = 0. \tag{5.1.8}$$

Integrating on X with respect to μ the obvious inequality $G_{\varepsilon,k} \leq \big(G_{\varepsilon,k}\big)_+$, we have

$$\begin{aligned}\int_X \big||g_k + g|^p - |g_k|^p - |g|^p\big| \,\mathrm{d}\mu &\leq \varepsilon \int_X |g_k|^p \,\mathrm{d}\mu + \int_X \big(G_{\varepsilon,k}\big)_+ \,\mathrm{d}\mu \\ &\leq \varepsilon\tilde{M} + \int_X \big(G_{\varepsilon,k}\big)_+ \,\mathrm{d}\mu,\end{aligned} \tag{5.1.9}$$

where we used (II) to get to the last inequality. Taking the limit and using (5.1.8)

$$\limsup_{k\to\infty} \int_X \big||g_k + g|^p - |g_k|^p - |g|^p\big| \,\mathrm{d}\mu \leq \varepsilon\tilde{M}.$$

Since $\varepsilon > 0$ is arbitrary we deduce (5.1.7). ■

5.2 Product measures and Fubini's theorem

Let μ be a measure on a set X, and ν be measure on a set Y.

Definition 5.2.1 (product measure). The *product measure of* μ *and* ν is the measure $\mu \times \nu : 2^{X\times Y} \to [0,\infty]$ given by

$$(\mu \times \nu)(S) := \inf \sum_{k\in\mathbb{N}} \mu(A_k)\nu(B_k) \quad S \subseteq X \times Y,$$

where the infimum is taken over all collections of μ-measurable sets $A_k \subseteq X$ and ν-measurable sets $B_k \subseteq Y$, $k \in \mathbb{N}$, such that

$$S \subseteq \bigcup_{k\in\mathbb{N}} (A_k \times B_k).$$

Definition 5.2.2 (**σ-finite set**)**.** The set $A \subseteq X$ is called *σ-finite* with respect to μ if

$$A = \bigcup_{k\in\mathbb{N}} A_k, \quad \text{where } A_k \text{ are } \mu\text{-measurable and } \mu(A_k) < \infty \ \forall\, k \in \mathbb{N}.$$

Exercise 5.2.3. Prove that if $A \subseteq X$ is σ-finite with respect to μ, then there exist *disjoint* μ-measurable sets B_k, $k \in \mathbb{N}$, such that $A = \bigcup_{k\in\mathbb{N}} B_k$ and $\mu(B_k) < \infty$ for all $k \in \mathbb{N}$.

Definition 5.2.4 (**σ-finite function**)**.** The function $f : X \to \bar{\mathbb{R}}$ is called *σ-finite* with respect to μ if f is μ-measurable and the set $\{x \in X \mid f(x) \neq 0\}$ is σ-finite with respect to μ.

Theorem 5.2.5 (Fubini). (i) *$\mu \times \nu$ is a regular measure on $X \times Y$.*

(ii) *If $A \subseteq X$ is μ-measurable and $B \subseteq Y$ is ν-measurable, then $A \times B$ is $\mu \times \nu$-measurable and $(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$.*

(iii) *If $S \subseteq X \times Y$ is σ-finite with respect to $\mu \times \nu$, then $S_y := \{x \in X \mid (x,y) \in S\}$ is μ-measurable for ν-a.e. $y \in Y$, $S_x := \{y \in Y \mid (x,y) \in S\}$ is ν-measurable for μ-a.e. $x \in X$, $\mu(S_y)$ is ν-integrable and $\nu(S_x)$ is μ-integrable. Moreover,*

$$(\mu \times \nu)(S) = \int_Y \mu(S_y)\ \mathrm{d}\nu(y) = \int_X \nu(S_x)\ \mathrm{d}\mu(x).$$

(iv) *If $f : X \times Y \to \bar{\mathbb{R}}$ is $\mu \times \nu$-integrable and also σ-finite with respect to $\mu \times \nu$, then the mapping $y \mapsto \int_X f(x,y)\mathrm{d}\mu(x)$ is ν-integrable, the mapping $x \mapsto \int_Y f(x,y)\mathrm{d}\nu(y)$ is μ-integrable, and*

$$\int_{X\times Y} f\ \mathrm{d}(\mu \times \nu) = \int_Y \Big[\int_X f(x,y)\ \mathrm{d}\mu(x)\Big]\mathrm{d}\nu(y) = \int_X \Big[\int_Y f(x,y)\ \mathrm{d}\nu(y)\Big]\mathrm{d}\mu(x).$$

Proof. `to be added` ■

Exercise 5.2.6. Let $\alpha > 0$. Given an $\mathscr{L}^1$-summable function $f : (0,\alpha) \to \bar{\mathbb{R}}$, define $g : (0,\alpha) \to \bar{\mathbb{R}}$ by

$$g(x) := \int_x^\alpha \frac{f(t)}{t}\ \mathrm{d}t.$$

Prove g is also $\mathscr{L}^1$-summable and satisfies

$$\int_0^\alpha g(x)\ \mathrm{d}x = \int_0^\alpha f(x)\ \mathrm{d}x.$$

Exercise 5.2.7. The *convolution* of two $\mathscr{L}^n$-measurable functions f, g is defined by

$$(f * g)(x) := \int_{\mathbb{R}^n} f(x-y)g(y)\ \mathrm{d}\mathscr{L}^n(y),$$

for any $x \in \mathbb{R}^n$ such that the integral exists. Use Fubini's theorem and Hölder's inequality to prove the following version of *Young's inequality*:

If $f \in L^1(\mathbb{R}^n, \mathscr{L}^n)$ and $g \in L^p(\mathbb{R}^n, \mathscr{L}^n)$, $p \in [1, \infty]$, then $f * g \in L^p(\mathbb{R}^n, \mathscr{L}^n)$ with

$$\|f * g\|_{L^p(\mathbb{R}^n, \mathscr{L}^n)} \leq \|f\|_{L^1(\mathbb{R}^n, \mathscr{L}^n)} \|g\|_{L^p(\mathbb{R}^n, \mathscr{L}^n)}.$$

Chapter 6

Some representation formulas

Theorem 6.0.1 (layer cake representation formula). *Let ν be a Radon measure on $[0,\infty)$ and set $\phi(t) := \nu\big([0,t)\big)$. Let μ be a regular measure on X and suppose that X is σ-finite with respect to μ. For any μ-measurable function $f : X \to [0,\infty]$ we have*[1]

$$\int_X \phi\big(f(x)\big)\,\mathrm{d}\mu(x) = \int_0^\infty \mu\big(\{f > t\}\big)\,\mathrm{d}\nu(t). \tag{6.0.1}$$

If in addition the measure ν satisfies $\nu\big([0,t)\big) = \nu\big([0,t]\big)$ for any $t \geq 0$, then also

$$\int_X \phi\big(f(x)\big)\,\mathrm{d}\mu(x) = \int_0^\infty \mu\big(\{f \geq t\}\big)\,\mathrm{d}\nu(t). \tag{6.0.2}$$

In particular, if $\phi : [0,\infty] \to [0,\infty]$ is a continuously differentiable non-decreasing function with $\phi(0) = 0$, we get

$$\int_X \phi\big(f(x)\big)\,\mathrm{d}\mu(x) = \int_0^\infty \phi'(t)\mu\big(\{f > t\}\big)\,\mathrm{d}\mathscr{L}^1(t). \tag{6.0.3}$$

Proof. Starting from the left hand side and using the definition of ν we have

$$\begin{aligned}\int_X \phi\big(f(x)\big)\,\mathrm{d}\mu(x) &= \int_X \nu\big(\big[0,f(x)\big)\big)\,\mathrm{d}\mu(x) \\ &= \int_X \int_0^\infty \chi_{\left[0,f(x)\right)}(t)\,\mathrm{d}\nu(t)\,\mathrm{d}\mu(x).\end{aligned} \tag{6.0.4}$$

Fubini's theorem applies to give

$$\int_X \phi\big(f(x)\big)\,\mathrm{d}\mu(x) = \int_0^\infty \int_X \chi_{\left[0,f(x)\right)}(t)\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(t).$$

[1] Throughout we write for instance $\{f > t\}$ in place of $\{y \in X \mid f(y) > t\}$

Since $\chi_{[0,f(x))}(t) = \chi_{\{f>t\}}(x)$ for all $t \in [0,\infty)$, $x \in X$, we conclude

$$\begin{aligned}\int_X \phi\big(f(x)\big)\, \mathrm{d}\mu(x) &= \int_0^\infty \int_X \chi_{\{f>t\}}(x)\, \mathrm{d}\mu(x)\, \mathrm{d}\nu(t)\\ &= \int_0^\infty \mu\big(\{f>t\}\big)\, \mathrm{d}\nu(t).\end{aligned}$$

For the formula (6.0.2), the assumption $\nu\big([0,t)\big) = \nu\big([0,t]\big)$ for any $t \geq 0$ when applied to (6.0.4) gives

$$\int_X \phi\big(f(x)\big)\, \mathrm{d}\mu(x) = \int_X \nu\big(\big[0,f(x)\big]\big)\, \mathrm{d}\mu(x).$$

We conclude with the proof by following the same steps as before. Finally, to get formula (6.0.3) notice that

$$\begin{aligned}\int_X \phi\big(f(x)\big)\, \mathrm{d}\mu(x) &= \int_X \int_0^{f(x)} \phi'(t)\, \mathrm{d}t\, \mathrm{d}\mu(x)\\ &= \int_X \int_0^\infty \chi_{[0,f(x))}(t)\phi'(t)\, \mathrm{d}\mathscr{L}^1(t)\, \mathrm{d}\mu(x).\end{aligned}$$

Now proceed as above from the application of Fubini's theorem and on. ■

Remark 6.0.2. Under the assumptions for (6.0.2) to hold true, since $\{f>t\} \subseteq \{f \geq t\}$, monotonicity of measures, formulas (6.0.1), (6.0.2) and exercise 4.3.16 readily imply $\mu(\{f > t\}) = \mu(\{f \geq t\})$ for ν-a.e. $t \in [0,\infty)$; that is,

$$\mu(\{f=t\}) = 0 \text{ for } \nu\text{-a.e. } t \in [0,\infty).$$

Observe that in case μ is a Radon measure on $\mathbb{R}^n$ and f is Borel measurable then this follows from lemma 2.3.10. Indeed, the sets $E_t := \{f=t\}$, $t \in \mathscr{I} := [0,\infty)$ form a family of disjoint Borel sets. Hence lemma 2.3.10 asserts $\mu(E_t) > 0$ for at most countable many $t \in \mathscr{I}$; that is,

$$\mu(\{f=t\}) = 0 \text{ for all but countably many } t \in [0,\infty).$$

But the condition $\nu\big([0,t)\big) = \nu\big([0,t]\big)$ for any $t \geq 0$ implies that singletons are of ν-measure 0 and consequently so is any countable subset of $[0,\infty)$.

Example 6.0.3. Some further useful cases are:

(i) Take $X = \mathbb{R}^n$ and μ to be Dirac's delta measure δ_x, $x \in \mathbb{R}^n$. For $\nu = \mathscr{L}^1|_{[0,\infty)}$, the formulas above read: *For any $\mathscr{L}^n$-measurable function $f : \mathbb{R}^n \to [0,\infty]$ we have*

$$f(x) = \int_0^\infty \chi_{\{f>t\}}(x)\, \mathrm{d}\mathscr{L}^1(t) = \int_0^\infty \chi_{\{f\geq t\}}(x)\, \mathrm{d}\mathscr{L}^1(t) \quad \mathscr{L}^n\text{-a.e. in } \mathbb{R}^n. \tag{6.0.5}$$

(ii) Take $X = \mathbb{R}^n$ and $\mu = \mathscr{L}^n$. If $u : \mathbb{R}^n \to \mathbb{R}$ is $\mathscr{L}^n$-measurable, then $f := |u|$ is $\mathscr{L}^n$-measurable by Theorem 4.1.6. Hence, with these choices for X, μ, f and taking $\nu\big([0,t)\big) = t^p$, $p \geq 1$, we get from (6.0.3) the following representation formula for the integral of the absolute value to the p-th power

$$\int_{\mathbb{R}^n} |u|^p \, \mathrm{d}\mathscr{L}^n = p \int_0^\infty t^{p-1} \mathscr{L}^n\big(\{|u| > t\}\big) \, \mathrm{d}\mathscr{L}^1(t). \tag{6.0.6}$$

Definition 6.0.4. Let $p, q > 0$. The *Lorentz space* $L^{p,q}$ is the collection of all measurable functions f defined on $\mathbb{R}^n$, such that $[f]_{L^{p,q}} < \infty$, where

$$[f]_{L^{p,q}} := \left(\frac{q}{p} \int_0^\infty \big(f^*(t)\big)^q t^{q/p-1} \, \mathrm{d}t \right)^{1/q}.$$

Here, f^* denotes the *decreasing rearrangement* of f,

$$f^*(t) := \inf\{s \geq 0 \text{ s.t. } \mu_f(s) \leq t\},$$

where μ_f is the *distribution function* of f, i.e.

$$\mu_f(s) := \mathscr{L}^n\big(\{x \in \mathbb{R}^n \text{ s.t. } |f(x)| > s\}\big).$$

Exercise 6.0.5. Prove the following properties of Lorenz spaces:

(i) We have $[f]_{L^{p,p}} = \int_{\mathbb{R}^n} |f|^p \, \mathrm{d}\mathscr{L}^n$.

(ii) If χ_ω is the characteristic function of a set ω with finite Lebesgue measure, then $[\chi_\omega]_{L^{p,q}} = [\mathscr{L}^n(\omega)]^{1/p}$ for all $p, q > 0$.

(iii) Let $p > 0$ and $q_2 \geq q_1 > 0$. There holds $[f]_{L^{p,q_1}} \geq [f]_{L^{p,q_2}}$.

We prove next an analogous to (6.0.6) formula for $f = |\nabla u|$, where $u \in C_c^\infty(\mathbb{R}^n)$.[2] To establish it we will need the following proposition and its consequence, the divergence or Gauss-Green theorem. The interested student can find the proof in [EG]-§3.3.

Proposition 6.0.6. *The measure $\mathscr{H}^{n-1}$ of a sufficiently smooth bounded surface agrees with the surface's area as taught in vector calculus.*

A consequence is the following[3]

[2] This a very simple case of the *coarea formula*; see also §3.4.2 of [EvG], or Theorem 3.2.11 of [H].

[3] A nice proof of this fundamental calculus fact can be found in the Appendix of [W].

Theorem 6.0.7 (divergence theorem). *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with smooth boundary and $\vec{F} = (F_1, ..., F_n)$, $F_i \in C^1(\bar{\Omega})$ for any $i = 1, ..., n$, be a vector field. Then*

$$\int_\Omega \operatorname{div} \vec{F}(x)\ \mathrm{d}x = \int_{\partial\Omega} \vec{F}(x) \cdot \vec{\nu}_{\partial\Omega}(x)\ \mathrm{d}\mathscr{H}^{n-1}(x),$$

where $\vec{\nu}_{\partial\Omega}(x)$ is the outwards pointing unit normal at the point $x \in \partial\Omega$.

Theorem 6.0.8 (representation formula for the integral of the modulus of the gradient). *For any $u \in C_c^\infty(\mathbb{R}^n)$ we have*

$$\int_{\mathbb{R}^n} |\nabla u(x)|\ \mathrm{d}x = \int_0^\infty \mathscr{H}^{n-1}\big(\{|u| = t\}\big)\ \mathrm{d}\mathscr{L}^1(t), \tag{6.0.7}$$

where $\mathscr{H}^{n-1}$ is the $n-1$-dimensional Hausdorff measure.

Proof. Let $\varepsilon > 0$ and $u \in C_c^\infty(\mathbb{R}^n)$. Define also $\vec{U}_\varepsilon : \mathbb{R}^n \to \mathbb{R}^n$ by

$$\vec{U}_\varepsilon(x) := \frac{\nabla u(x)}{\sqrt{|\nabla u(x)|^2 + \varepsilon}}.$$

Noting that for any $t \geq 0$ we have $\{u^+ > t\} = \{u > t\}$, formula (6.0.5) implies

$$u^+(x) = \int_0^\infty \chi_{\{u>t\}}(x)\ \mathrm{d}\mathscr{L}^1(t) \quad \mathscr{L}^n\text{-a.e. in } \mathbb{R}^n.$$

Also, since for any $t \geq 0$ we have $\{u^- \geq t\} = \{u \leq -t\}$, the same formula gives

$$\begin{aligned} u^-(x) &= \int_0^\infty \chi_{\{u\leq -t\}}(x)\ \mathrm{d}\mathscr{L}^1(t) \\ &= \int_{-\infty}^0 \chi_{\{u\leq t\}}(x)\ \mathrm{d}\mathscr{L}^1(t) \\ &= -\int_{-\infty}^0 \big(\chi_{\{u>t\}}(x) - 1\big)\ \mathrm{d}\mathscr{L}^1(t) \quad \mathscr{L}^n\text{-a.e. in } \mathbb{R}^n. \end{aligned}$$

We then have

$$\begin{aligned} \int_{\mathbb{R}^n} u^+(x) \operatorname{div} \vec{U}_\varepsilon(x)\ \mathrm{d}\mathscr{L}^n(x) &= \int_{\mathbb{R}^n} \Big(\int_0^\infty \chi_{\{u>t\}}(x)\ \mathrm{d}\mathscr{L}^1(t)\Big) \operatorname{div} \vec{U}_\varepsilon(x)\ \mathrm{d}\mathscr{L}^n(x) \\ &= \int_0^\infty \int_{\{u>t\}} \operatorname{div} \vec{U}_\varepsilon(x)\ \mathrm{d}x\ \mathrm{d}\mathscr{L}^1(t), \end{aligned}$$

by the Fubini theorem. Correspondingly,

$$
\begin{aligned}
-\int_{\mathbb{R}^n} u^-(x)\operatorname{div}\vec{U}_\varepsilon(x)\,\mathrm{d}\mathscr{L}^n(x) &= \int_{\mathbb{R}^n}\Big(\int_{-\infty}^{0}\big(\chi_{\{u>t\}}(x)-1\big)\,\mathrm{d}\mathscr{L}^1(t)\Big)\operatorname{div}\vec{U}_\varepsilon(x)\,\mathrm{d}\mathscr{L}^n(x)\\
&= \int_{-\infty}^{0}\int_{\mathbb{R}^n}\big(\chi_{\{u>t\}}(x)-1\big)\operatorname{div}\vec{U}_\varepsilon(x)\,\mathrm{d}\mathscr{L}^n(x)\,\mathrm{d}\mathscr{L}^1(t)\\
&= \int_{-\infty}^{0}\Big(\int_{\{u>t\}}\operatorname{div}\vec{U}_\varepsilon(x)\,\mathrm{d}x-\int_{\mathbb{R}^n}\operatorname{div}\vec{U}_\varepsilon(x)\,\mathrm{d}x\Big)\,\mathrm{d}\mathscr{L}^1(t)\\
&= \int_{-\infty}^{0}\int_{\{u>t\}}\operatorname{div}\vec{U}_\varepsilon(x)\,\mathrm{d}x\,\mathrm{d}\mathscr{L}^1(t),
\end{aligned}
$$

where we have applied Fubini's theorem in the second equality and the divergence theorem in the last one (to show that $\int_{\mathbb{R}^n}\operatorname{div}\vec{U}_\varepsilon(x)\mathrm{d}x = 0$). Adding the last two equalities

$$
\int_{\mathbb{R}^n} u(x)\operatorname{div}\vec{U}_\varepsilon(x)\,\mathrm{d}x = \int_{-\infty}^{\infty}\int_{\{u>t\}}\operatorname{div}\vec{U}_\varepsilon(x)\,\mathrm{d}x\,\mathrm{d}\mathscr{L}^1(t),
$$

from which, after an integration by parts on the left hand side and applying the divergence theorem on the right hand side[4], we conclude

$$
-\int_{\mathbb{R}^n}\nabla u(x)\cdot\vec{U}_\varepsilon(x)\,\mathrm{d}x = \int_{-\infty}^{\infty}\int_{\{u=t\}}\vec{U}_\varepsilon(x)\cdot\vec{\nu}(x)\,\mathrm{d}\mathscr{H}^{n-1}(x)\,\mathrm{d}\mathscr{L}^1(t),
$$

where $\vec{\nu}(x)$ is the unit normal at a point x of the surface $\{u=t\}$ pointing in the direction where $u\le t$. From Theorem 1.2.9 we know $\nabla u(x)\cdot\vec{\nu}(x) = -|\nabla u(x)|$ on $\{u=t\}$ and so

$$
\vec{U}_\varepsilon(x)\cdot\vec{\nu}(x) = -\frac{|\nabla u(x)|}{\sqrt{|\nabla u(x)|^2+\varepsilon}}\qquad\text{on } \{u=t\}.
$$

Hence

$$
\int_{\mathbb{R}^n}\frac{|\nabla u(x)|^2}{\sqrt{|\nabla u(x)|^2+\varepsilon}}\,\mathrm{d}x = \int_{-\infty}^{\infty}\int_{\{u=t\}}\frac{|\nabla u(x)|}{\sqrt{|\nabla u(x)|^2+\varepsilon}}\,\mathrm{d}\mathscr{H}^{n-1}(x)\,\mathrm{d}\mathscr{L}^1(t).
$$

The results follows letting $\varepsilon\downarrow 0$ through the monotone convergence theorem applied in both sides. ■

Combining the above representation theorem with the isoperimetric inequality we get the sharp L^1-*Sobolev inequality*, or sharp *Gagliardo-Nirenberg inequality*; that is,

[4] we need to know that the level set $\{u=t\}$ is a smooth surface to apply it. This is a consequence of Sard's lemma which is going to be added in the next update of the notes (see the Appendix of [W])

Theorem 6.0.9. *For any $u \in C_c^\infty(\mathbb{R}^n)$, $n \in \mathbb{N} \setminus \{1\}$, we have*

$$\int_{\mathbb{R}^n} |\nabla u(x)| \, \mathrm{d}x \geq n\omega_n^{1/n} \Big(\int_{\mathbb{R}^n} |u(x)|^{n/(n-1)} \, \mathrm{d}x \Big)^{(n-1)/n}. \tag{6.0.8}$$

Proof. From the isoperimetric inequality we have

$$\mathscr{H}^{n-1}\big(\{|u| = t\}\big) \geq n\omega_n^{1/n} \big[\mathscr{L}^n\big(\{|u| > t\}\big)\big]^{(n-1)/n}.$$

Combining this with the representation formula of the last theorem,

$$\int_{\mathbb{R}^n} |\nabla u(x)| \, \mathrm{d}x \geq n\omega_n^{1/n} \int_0^\infty \big[\mathscr{L}^n\big(\{|u| > t\}\big)\big]^{(n-1)/n} \, \mathrm{d}\mathscr{L}^1(t).$$

It suffices to prove

$$\int_0^\infty \big[\mathscr{L}^n\big(\{|u| > t\}\big)\big]^{(n-1)/n} \, \mathrm{d}\mathscr{L}^1(t) \geq I^{(n-1)/n}, \tag{6.0.9}$$

where $I := \int_{\mathbb{R}^n} |u(x)|^{n/(n-1)} \, \mathrm{d}x$. To this end using (6.0.5) with $f = |u|$ and then Fubini's theorem we write

$$\begin{aligned} I &= \int_{\mathbb{R}^n} |u(x)| |u(x)|^{1/(n-1)} \, \mathrm{d}x \\ &= \int_{\mathbb{R}^n} \int_0^\infty \chi_{\{|u|>t\}}(x) \, \mathrm{d}\mathscr{L}^1(t) |u(x)|^{1/(n-1)} \, \mathrm{d}\mathscr{L}^n(x) \\ &= \int_0^\infty \int_{\mathbb{R}^n} \chi_{\{|u|>t\}}(x) |u(x)|^{1/(n-1)} \, \mathrm{d}\mathscr{L}^n(x) \, \mathrm{d}\mathscr{L}^1(t). \end{aligned}$$

Applying next Hölder's inequality with the conjugate exponents $n/(n-1)$ and n,

$$\begin{aligned} I &\leq \int_0^\infty \Big(\int_{\mathbb{R}^n} \big(\chi_{\{|u|>t\}}(x)\big)^{n/(n-1)} \, \mathrm{d}\mathscr{L}^n(x) \Big)^{(n-1)/n} I^{1/n} \, \mathrm{d}\mathscr{L}^1(t) \\ &= I^{1/n} \int_0^\infty \big[\mathscr{L}^n\big(\{|u| > t\}\big)\big]^{(n-1)/n} \, \mathrm{d}\mathscr{L}^1(t), \end{aligned}$$

from which (6.0.9) readily follows. ■

Chapter 7

Elementary convexity

7.1 Convex functions

Definition 7.1.1 (convex function). Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be convex. A function $f : \Omega \to \mathbb{R}$ is called *convex in* Ω if

$$f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda)f(y) \quad \forall\, \lambda \in [0,1],\ \forall\, x,\ y \in \Omega. \tag{7.1.1}$$

In case the reverse inequality holds true in (7.1.1), then f is called *concave*.

Remark 7.1.2. Let $-\infty \leq a < b \leq \infty$. The mean value theorem shows that if $f : (a,b) \to \mathbb{R}$ has nonnegative second derivative at any point of (a,b), then f is convex in (a,b).

Remark 7.1.3. Using induction we obtain for any $k \in \mathbb{N}$ the equivalent inequality

$$f\Big(\sum_{i=1}^{k} \lambda_i x_i\Big) \leq \sum_{i=1}^{k} \lambda_i f(x_i),$$

valid for all $\lambda_i \in [0,1]$, $i = 1,\ldots,k$, satisfying $\sum_{i=1}^{k} \lambda_i = 1$, and all $x_i \in \Omega$, $i = 1,\ldots,k$. For instance, for $k = 3$ we have

$$\begin{aligned}
f(\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3) &= f\left(\lambda_1 x_1 + (\lambda_2+\lambda_3)\Big(\frac{\lambda_2}{\lambda_2+\lambda_3}x_2 + \frac{\lambda_3}{\lambda_2+\lambda_3}x_3\Big)\right)\\
&\leq \lambda_1 f(x_1) + (\lambda_2+\lambda_3) f\Big(\frac{\lambda_2}{\lambda_2+\lambda_3}x_2 + \frac{\lambda_3}{\lambda_2+\lambda_3}x_3\Big)\\
&\leq \lambda_1 f(x_1) + (\lambda_2+\lambda_3)\Big(\frac{\lambda_2}{\lambda_2+\lambda_3}f(x_2) + \frac{\lambda_3}{\lambda_2+\lambda_3}f(x_3)\Big)\\
&\leq \lambda_1 f(x_1) + \lambda_2 f(x_2) + \lambda_3 f(x_3).
\end{aligned}$$

Remark 7.1.4. The arithmetic-geometric inequality of lemma 3.0.5 follows from the above remark by taking $k = n$, $\lambda_i = 1/n$, $x_i = \log a_i$ and $f(t) = e^t$, $t \in \mathbb{R}$.

Remark 7.1.5. To show the claim made in the proof of theorem 5.1.9; that is,

$$\forall\, \varepsilon > 0,\ \ \exists\, C_\varepsilon > 0 \text{ such that } \big||a+b|^p - |a|^p\big| \le \varepsilon|a|^p + C_\varepsilon|b|^p \ \ \forall\, a,b \in \mathbb{R},$$

we observe first the function $(0,\infty) \ni t \mapsto t^p$ is convex in case $p > 1$. Hence, for $A, B \in \mathbb{R}$ and $\lambda \in (0,1)$,

$$|A+B|^p \le (|A|+|B|)^p = \Big(\lambda\frac{|A|}{\lambda} + (1-\lambda)\frac{|B|}{1-\lambda}\Big)^p \le \lambda^{1-p}|A|^p + (1-\lambda)^{1-p}|B|^p.$$

Taking $A = a$, $B = b$ and $\lambda = (1+\varepsilon)^{1/(1-p)}$ we get

$$|a+b|^p - |a|^p \le \varepsilon|a|^p + C_\varepsilon|b|^p.$$

On the other hand, if $\varepsilon < 1$ then taking $A = a+b$, $B = -b$ and $\lambda = (1-\varepsilon)^{1/(p-1)}$ we get

$$|a|^p - |a+b|^p \le \varepsilon|a|^p + C_\varepsilon|b|^p. \tag{7.1.2}$$

Since (7.1.2) is also true if $\varepsilon \ge 1$, the claim follows. For $0 < p \le 1$ observe instead

$$\big(|a|+|b|\big)^p \le |a|^p + |b|^p.$$

Exercise 7.1.6. Let $A \subset \mathbb{R}^n$ and denote its *indicator function* by $\mathfrak{I}_A$; that is

$$\mathfrak{I}_A(x) := \begin{cases} 0 & \text{if } x \in A, \\ \infty & \text{if else.} \end{cases}$$

Prove this is a convex function if and only if A is a convex set.

Exercise 7.1.7. (i) Prove using the Brunn-Minkowski inequality of theorem 3.0.3, the function $\big(\mathscr{L}^n(\cdot)\big)^{1/n} : 2^{\mathbb{R}^n} \to [0,\infty]$ is concave with respect to the Minkowski addition; that is, for any $\lambda \in [0,1]$ and all $X, Y \subseteq \mathbb{R}^n$ there holds

$$\Big(\mathscr{L}^n\big(\lambda X + (1-\lambda)Y\big)\Big)^{1/n} \ge \lambda\big(\mathscr{L}^n(X)\big)^{1/n} + (1-\lambda)\big(\mathscr{L}^n(Y)\big)^{1/n}.$$

(ii) Prove $\mathscr{L}^n\big(\lambda X + (1-\lambda)Y\big) \ge \big(\mathscr{L}^n(X)\big)^\lambda \big(\mathscr{L}^n(Y)\big)^{1-\lambda}$.

(iii) Prove that if the inequality of (ii) is known to be true for all $\lambda \in [0,1]$ and all $X, Y \subseteq \mathbb{R}^n$, then it implies the Brunn-Minkowski inequality.

Exercise 7.1.8. Let $f : [0,\infty) \to [0,\infty)$ be concave. Prove

(i) $g : (0,\infty) \to [0,\infty)$ defined by $g(x) = f(x)/x$ is decreasing.

(ii) f is *sub-additive; that is* $f(x+y) \le f(x) + f(y)$ *for all* $x, y \ge 0$.

Example 7.1.9. Let $\Omega \subsetneq \mathbb{R}^n$ be nonempty and convex.

(i) If Ω is closed, then the distance function to the set Ω is convex in $\mathbb{R}^n$.

Proof. Denote by d_Ω the distance to the set Ω; that is,

$$d_\Omega(x) := \operatorname{dist}(x,\Omega) = \inf_{y\in\Omega} |x-y|, \quad x \in \mathbb{R}^n.$$

Let $x, y \in \mathbb{R}^n$ with $x \neq y$ and choose $\xi_x, \xi_y \in \Omega$ such that $d_\Omega(x) = |x-\xi_x|$ and $d_\Omega(y) = |y-\xi_y|$. Let $\lambda \in (0,1)$ and consider the point $\lambda x + (1-\lambda)y$. The convexity of Ω implies that $\lambda\xi_x + (1-\lambda)\xi_y \in \Omega$ and we have

$$\begin{aligned} d_\Omega\big(\lambda x + (1-\lambda)y\big) &\le \big|\lambda x + (1-\lambda)y - \big(\lambda\xi_x + (1-\lambda)\xi_y\big)\big| \\ &\le \lambda|x-\xi_x| + (1-\lambda)|y-\xi_y| \\ &\le \lambda d_\Omega(x) + (1-\lambda)d_\Omega(y), \end{aligned}$$

where we have used the triangle inequality to get to the middle line. ■

(ii) If Ω is open, then the distance function to $\partial\Omega$ is concave in Ω.

Proof. Denote by $d_{\partial\Omega}(x)$ the distance of $x \in \Omega$ to the set $\partial\Omega$; that is,

$$d_{\partial\Omega}(x) := \operatorname{dist}(x,\partial\Omega) = \inf_{y\in\partial\Omega} |x-y|, \quad x \in \Omega.$$

Let $x, y \in \Omega$ with $x \neq y$ and choose $\xi_x, \xi_y \in \partial\Omega$ such that $d_{\partial\Omega}(x) = |x-\xi_x| =: r_x$ and $d_{\partial\Omega}(y) = |y-\xi_y| =: r_y$. Let $\lambda \in (0,1)$.

Claim. The open ball with center $z_\lambda := \lambda x + (1-\lambda)y$ and radius $r_\lambda := \lambda r_x + (1-\lambda)r_y$ is contained in Ω.

Proof of claim. Given $z \in B_{r_\lambda}(z_\lambda)$ consider the points

$$z_x := x + r_x\frac{z-z_\lambda}{r_\lambda} \quad \text{and} \quad z_y := y + r_y\frac{z-z_\lambda}{r_\lambda}.$$

Then $z_x \in B_{r_x}(x)$, $z_y \in B_{r_y}(y)$ and also $\lambda z_x + (1-\lambda)z_y = z$. Thus z is in the Minkowski sum of the sets $\lambda B_{r_x}(x)$ and $(1-\lambda)B_{r_y}(y)$. Since $B_{r_x}(x), B_{r_y}(y) \subset \Omega$ and Ω is convex, we conclude $z \in \Omega$. □

Since $B_{r_\lambda}(z_\lambda) \subset \Omega$ we readily have

$$d_{\partial\Omega}(z_\lambda) \geq d_{\partial B_{r_\lambda}(z_\lambda)}(z_\lambda) = r_\lambda.$$

Thus $d_{\partial\Omega}\big(\lambda x + (1-\lambda)y\big) \geq \lambda d_{\partial\Omega}(x) + (1-\lambda)d_{\partial\Omega}(y)$ as required. ∎

Lemma 7.1.10. *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is convex. Then*

$$\forall\, x \in \mathbb{R},\ \exists\, m \in \mathbb{R} \ \text{ such that } \ f(y) \geq f(x) + m(y-x) \quad \forall\, y \in \mathbb{R}.$$

Proof. Let $0 < \lambda < \kappa$. Then

$$f(x+\lambda) = f\Big(\frac{\kappa-\lambda}{\kappa}x + \frac{\lambda}{\kappa}(x+\kappa)\Big) \leq \frac{\kappa-\lambda}{\kappa}f(x) + \frac{\lambda}{\kappa}f(x+\kappa).$$

This implies

$$\frac{f(x+\lambda)-f(x)}{\lambda} \leq \frac{f(x+\kappa)-f(x)}{\kappa}, \tag{7.1.3}$$

and so

$$f'(x+) := \lim_{\lambda\downarrow 0}\frac{f(x+\lambda)-f(x)}{\lambda} \leq \frac{f(y)-f(x)}{y-x} \quad \forall\, y > x. \tag{7.1.4}$$

Similarly,

$$f(x-\lambda) = f\Big(\frac{\kappa-\lambda}{\kappa}x + \frac{\lambda}{\kappa}(x-\kappa)\Big) \leq \frac{\kappa-\lambda}{\kappa}f(x) + \frac{\lambda}{\kappa}f(x-\kappa).$$

This implies

$$\frac{f(x-\lambda)-f(x)}{-\lambda} \geq \frac{f(x)-f(x-\kappa)}{\kappa}, \tag{7.1.5}$$

and so

$$f'(x-) := \lim_{\lambda\downarrow 0}\frac{f(x-\lambda)-f(x)}{-\lambda} \geq \frac{f(x)-f(y)}{x-y} \quad \forall\, y < x. \tag{7.1.6}$$

Also, for $\lambda > 0$ we have

$$f(x) = f\Big(\frac{x-\lambda}{2} + \frac{x+\lambda}{2}\Big) \leq \frac{1}{2}f(x-\lambda) + \frac{1}{2}f(x+\lambda).$$

This implies

$$\frac{f(x-\lambda)-f(x)}{-\lambda} \leq \frac{f(x+\lambda)-f(x)}{\lambda},$$

and so

$$f'(x-) \leq f'(x+) \quad \forall\, x \in \mathbb{R}. \tag{7.1.7}$$

Because of (7.1.7), given $x \in \mathbb{R}^n$ we may pick $m \in [f'(x-), f'(x+)]$. Apply then (7.1.4) and (7.1.6) to obtain the inequality of the lemma. ∎

A byproduct of estimates presented in the above proof is

Lemma 7.1.11. *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is convex. Then*

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} \leq \frac{f(x_3)-f(x_1)}{x_3-x_1} \leq \frac{f(x_3)-f(x_2)}{x_3-x_2} \quad \forall\, x_1 < x_2 < x_3.$$

Proof. From (7.1.3) setting $x = x_1$, $x+\lambda = x_2$ and $x+\kappa = x_3$ we deduce the inequality on the left. From (7.1.5) setting $x-\kappa = x_1$, $x-\lambda = x_2$ and $x = x_3$ we deduce the inequality on the right. ∎

Theorem 7.1.12 (Jensen's inequality). *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is convex. Let μ be a measure on a set X satisfying $0 < \mu(X) < \infty$. If $u : X \to \bar{\mathbb{R}}$ is μ-summable, then*

$$f\Big(\frac{1}{\mu(X)} \int_X u \,\mathrm{d}\mu\Big) \leq \frac{1}{\mu(X)} \int_X f(u) \,\mathrm{d}\mu.$$

Proof. Taking

$$x := \frac{1}{\mu(X)} \int_X u \,\mathrm{d}\mu, \qquad y := u(z), \;\; z \in X,$$

in the previous lemma, we get

$$f\big(u(z)\big) \geq f\Big(\frac{1}{\mu(X)} \int_X u \,\mathrm{d}\mu\Big) + m\Big(u(z) - \frac{1}{\mu(X)} \int_X u \,\mathrm{d}\mu\Big) \quad \mu\text{-a.e. } z \in X.$$

The proof now follows by taking mean values on X with respect to μ on both sides. ∎

Exercise 7.1.13. Let μ be a measure on the set X and $\mu(X) = 1$. Let $u : X \to [0,\infty]$ be μ-summable.

(i) If $\int_X u \,\mathrm{d}\mu = 1$, prove

$$\int_X u \log u \,\mathrm{d}\mu \geq 0.$$

(ii) Prove that for all $p \geq 1$ there holds

$$\frac{1}{\int_X \frac{1}{u^p} \,\mathrm{d}\mu} \leq \Big(\int_X u \,\mathrm{d}\mu\Big)^p.$$

(iii) Prove

$$\sqrt{1+\Big(\int_X u \,\mathrm{d}\mu\Big)^2} \leq \int_X \sqrt{1+u^2} \,\mathrm{d}\mu \leq 1 + \int_X u \,\mathrm{d}\mu.$$

7.2 Lipschitz functions

Definition 7.2.1 (Lipschitz function). Let $A \subset \mathbb{R}^n$. A function $f : A \to \mathbb{R}$ is called *Lipschitz continuous*, or simply *Lipschitz*, if there exists a constant $L > 0$ such that

$$|f(x) - f(y)| \leq L|x-y| \quad \forall\, x, y \in A.$$

Remark 7.2.2. A Lipschitz function is uniformly continuous in A.

Example 7.2.3. Two examples of Lipschitz functions are

(i) the distance function d_F to a nonempty closed subset $F \subsetneq \mathbb{R}^n$; that is,

$$d_F(x) := \inf_{y \in F} |x-y|, \quad x \in \mathbb{R}^n. \tag{7.2.1}$$

Proof. Let $x, y \in \mathbb{R}^n$ and choose $\xi_x \in F$ so that $d_F(x) := |x - \xi_x|$. Then

$$d_F(y) - d_F(x) \leq |y - \xi_x| - |x - \xi_x| \leq |y - x|.$$

Chose now $\xi_y \in F$ so that $d_F(y) := |y - \xi_y|$. Then in the same fashion

$$d_F(x) - d_F(y) \leq |x - y|.$$

These estimates show d_F is Lipschitz continuous with $L = 1$. ■

(ii) real functions in $C^1(\Omega)$ having uniformly bounded derivative in Ω, with Ω being any open and convex subset of $\mathbb{R}^n$.

Proof. Let $f : \Omega \to \mathbb{R}$, $f \in C^1(\Omega)$, for which there exists $M > 0$ such that

$$|\nabla f(x)| \leq M \quad \forall\, x \in \Omega.$$

Let $x, y \in \Omega$ and consider the straight line $\gamma(t) := x + t(y-x)$, $t \in [0,1]$. Applying the mean value theorem to the function $u : [0,1] \to \mathbb{R}$ defined by $u(t) := f\big(\gamma(t)\big)$, we get $t_0 \in (0,1)$ such that $u'(t_0) = u(1) - u(0)$. By the chain rule we deduce

$$Df\big(\gamma(t_0)\big) \cdot (y-x) = f(y) - f(x).$$

Coupling this with the bound on the derivative of f we conclude that f is Lipschitz with $L = M$. ■

Exercise 7.2.4. Let $\Omega \subsetneq \mathbb{R}^n$ be open and suppose that $f : \bar{\Omega} \to \mathbb{R}$ is Lipschitz with $f = 0$ on $\partial\Omega$. Prove that there exists a positive constant C such that

$$|f(x)| \leq C d_{\partial\Omega}(x) \quad \forall\, x \in \bar{\Omega}.$$

Definition 7.2.5 (locally Lipschitz function). Let $A \subset \mathbb{R}^n$. A function $f : A \to \mathbb{R}$ is called *locally Lipschitz continuous*, or simply *locally Lipschitz*, if for any compact set $K \subseteq A$ there exists a constant $C_K > 0$ such that

$$|f(x) - f(y)| \leq C_K|x - y| \quad \forall\, x,\, y \in K.$$

Remark 7.2.6. A locally Lipschitz function is continuous.

Example 7.2.7. Any convex function $f : \Omega \to \mathbb{R}$, where $\Omega \subseteq \mathbb{R}^n$ is open and convex, is locally Lipschitz.

Proof. The proof is done by induction on the space dimension $n \in \mathbb{N}$:

For $n = 1$, let $K \neq \emptyset$ be a compact subset of Ω. By the convexity of Ω and proposition 1.1.17 we can choose connected $X, Y \subset \Omega$ such that $K^\circ \Subset X \Subset Y \Subset \Omega$. Since $n = 1$ we have that $X = (b, c)$ and $Y = (a, d)$ where $a < b < c < d$. Thus for any $x, y \in K$ with $x < y$ we apply lemma 7.1.11 two times to get

$$C_1 := \frac{f(b) - f(a)}{b - a} \leq \frac{f(y) - f(x)}{y - x} \leq \frac{f(d) - f(c)}{d - c} =: C_2.$$

This in turn implies

$$\frac{|f(y) - f(x)|}{|y - x|} \leq \max\{|C_1|, |C_2|\} \quad \text{for all } x, y \in K,\ x \neq y.$$

Now assume the result is true for $n = k - 1$, $k \in \mathbb{N} \setminus \{1\}$. Let $f : \Omega \to \mathbb{R}$ be a convex function, where Ω is an open and convex subset of $\mathbb{R}^k$. Given a compact $K \subset \Omega$ we choose as above connected $X, Y \subset \Omega$ such that $K^\circ \Subset X \Subset Y \Subset \Omega$. Recalling exercise 1.1.23 we replace X and Y by $X' \supset X$ and $Y' \supset Y$ respectively, where each one of $\bar{X}'$ and $\bar{Y}'$ is a union of a finite number of nonoverlapping intervals of $\mathbb{R}^k$ and such that $K^\circ \Subset X' \Subset Y' \Subset \Omega$. The line segment connecting two given distinct points $x, y \in K$ is extended in the direction of x and hits $\partial X'$ at b and $\partial Y'$ at a. It is also extended in the direction of y and hits $\partial X'$ at c and $\partial Y'$ at d. The function f restricted on the part of this line which lies in Ω is a one dimensional convex function and as before we have the estimate

$$\frac{|f(y) - f(x)|}{|y - x|} \leq \max\left\{\frac{|f(b) - f(a)|}{|b - a|}, \frac{|f(d) - f(c)|}{|d - c|}\right\}.$$

The right hand side is always finite under the above construction. Indeed, the function

$$F(x, y) := \frac{|f(y) - f(x)|}{|y - x|} \quad x \in \partial X',\ y \in \partial Y',$$

is well defined because $\partial X'$ and $\partial Y'$ are of positive distance. Furthermore, $\partial X'$ (resp. $\partial Y'$) is a finite union of parts of the $k-1$-dimensional faces of those nonoverlapping intervals of X' (resp. Y') that contribute to $\partial X'$ (resp. $\partial Y'$). Clearly, f is convex on any such face and by the induction hypothesis and remark 7.2.6, f is continuous on any such face. Thus f is continuous on $\partial X'$ and on $\partial Y'$ which implies in turn that F is continuous on $\partial X' \times \partial Y'$. Since $\partial X' \times \partial Y'$ is compact, F attains its maximum there (theorem 1.2.5). ■

For instructive reasons we provide a second proof for the statement in the above example. It is based on the following fact[1]

Lemma 7.2.8. *Let Q be an interval in $\mathbb{R}^n$ and $y_1, \ldots, y_{2^n}$ the vertices. Then every point $x \in Q$ can be written as a convex linear combination of the vertices, i.e. there exist $\lambda_i \in [0,1]$, $i = 1, \ldots, 2^n$, such that*

$$x = \sum_{i=1}^{2^n} \lambda_i y_i, \qquad \sum_{i=1}^{2^n} \lambda_i = 1.$$

Proof. The proof is done by induction on the space dimension $n \in \mathbb{N}$:

For $n = 1$ the statement is trivial. Suppose now the claim holds true for $n-1$. Let $x = (x_1, \ldots, x_n) \in Q = [l_1^1, l_1^2] \times [l_2^1, l_2^2] \times \ldots \times [l_n^1, l_n^2]$. Further, we denote by A the set of all vertices having their last coordinate equal to l_n^1 and by B the set of all vertices having their last coordinate equal to l_n^2. Observe that A and B are disjoint, $\operatorname{card}(A) = \operatorname{card}(B) = 2^{n-1}$ and their union is the whole set of vertices of Q. Hence we can define a bijection

$$\varphi : \begin{cases} A \to B \\ (z_1, z_2, \ldots, z_{n-1}, l_n^1) \to (z_1, z_2, \ldots, z_{n-1}, l_n^2). \end{cases}$$

Naturally, we obtain for each $a_i \in A$, $i = 1, \ldots, 2^{n-1}$, a curve $\gamma_i(t) = t a_i + (1-t)\varphi(a_i)$ for $t \in [0,1]$. Furthermore, let $H = \{z = (z_1, \ldots, z_n) \in \mathbb{R}^n \mid z_n = x_n\}$ be a hyperplane. Since γ_i is continuous for each $i = 1, \ldots, 2^{n-1}$ and all curves are identical in the last coordinate, we obtain a t^* such that the last coordinate of $\gamma_i(t^*)$ is equal to x_n for all i. Next we observe the points $\{\gamma_1(t^*), \ldots, \gamma_{2^{n-1}}(t^*)\}$ lie in the hyperplane H and generate an interval in this hyperplane. By our induction hypotheses we can write the point $(x_1, \ldots, x_{n-1})$ as a convex linear combination of the points $\{\gamma_1(t^*), \ldots, \gamma_{2^{n-1}}(t^*)\}$ restricted to the hyperplane H. But since the linear combination is convex and $\{\gamma_1(t^*), \ldots, \gamma_{2^{n-1}}(t^*)\}$ lie in the hyperplane H, we are allowed to write $(x_1, \ldots, x_n)$ as a linear combination of $\{\gamma_1(t^*), \ldots, \gamma_{2^{n-1}}(t^*)\}$ in $\mathbb{R}^n$. Finally, observe that the points $\{\gamma_1(t^*), \ldots, \gamma_{2^{n-1}}(t^*)\}$ are convex linear combinations of our 2^n vertices. ■

Second proof for the statement in example 7.2.7. Let $Q_l(y_0) \subseteq \Omega$ be a cube of side length $l > 0$, center at y_0 and vertices $y_1, y_2, \ldots, y_{2^n}$. Set $V := \{y_1, y_2, \ldots, y_{2^n}\}$. If $x \in Q_l(y_0)$ then

[1] this is a primitive version of the Krein-Milman theorem in Functional Analysis; see [Br, Theorem 1.13]

from the previous lemma we know there exists $\lambda_i \in [0,1]$, $i = 1,\dots,2^n$, with $\sum_{k=1}^{2^n} \lambda_k = 1$ and $x = \sum_{k=1}^{2^n} \lambda_k y_k$. The convexity of f implies through exercise 7.1.3 that

$$f(x) \le \sum_{k=1}^{2^n} \lambda_k f(y_k) \le \max_V |f| < \infty.$$

Thus $M := \sup_{Q_l} f < \infty$. Since $y_0 = (1/2)x + (1/2)(2y_0 - x)$ and $2y_0 - x \in Q_l(y_0)$ (why?), applying the convexity property of f gives

$$f(y_0) \le \frac{1}{2} f(x) + \frac{1}{2} f(2y_0 - x) \le \frac{1}{2} f(x) + \frac{1}{2} M.$$

This gives $\inf_{Q_l} f \ge 2f(y_0) - M$. Hence f is bounded in Q_l. It follows that f is locally bounded in Ω (why?). Next let $\bar{B}_{3r}(x_0) \subset \Omega$ and x, y be two distinct points in $\bar{B}_r(x_0)$. Select $\mu > 0$ and $z \in \partial B_{2r}(x)$ so that

$$z - x = \mu(y - x).$$

Then $\mu = 2r/|y-x| > 1$ and $y = (1/\mu)z + (1 - 1/\mu)x$. Hence, the convexity of f gives

$$\begin{aligned} f(y) &\le (1/\mu) f(z) + (1 - 1/\mu) f(x) \\ &\le f(x) + (1/\mu)\big(|f(z)| + |f(x)|\big) \\ &= f(x) + \frac{|f(z)| + |f(x)|}{2r} |y - x| \\ &\le f(x) + L|y - x|, \end{aligned} \tag{7.2.2}$$

where $L = (1/r) \sup_{\bar{B}_{3r}} |f|$. Interchanging the roles of x and y we also have

$$f(x) \le f(y) + L|y - x|.$$

Hence f is Lipschitz in $\bar{B}_r$. It follows that f is locally Lipschitz in Ω (why?). ■

7.3 A characterization of convex functions

Theorem 7.3.1. *Let $\Omega \subseteq \mathbb{R}^n$ be nonempty and convex. A function $f : \Omega \to \mathbb{R}$ is convex if and only if it is continuous and satisfies the inequality in the definition of convexity for $\lambda = 1/2$ only* [2]*; that is,*

$$f\Big(\frac{x+y}{2}\Big) \le \frac{1}{2} f(x) + \frac{1}{2} f(y) \quad \forall\, x,\ y \in \Omega. \tag{7.3.1}$$

[2] or any other value of λ in the interval $(0,1)$

Proof. If f is convex then (7.3.1) is true. Also, by the above example and remark we get f is continuous. To prove the other direction let $x, y \in \Omega$ and define $\gamma : [0,1] \to \Omega$ by

$$\gamma(\lambda) := \lambda x + (1-\lambda)y, \quad \lambda \in [0,1].$$

Define the sequence of sets

$$D_k := \big\{\gamma(\lambda) \mid \lambda \in \{0/2^k, 1/2^k, ..., 2^k/2^k\}\big\}, \quad k \in \mathbb{N}.$$

Claim. There holds $f\big(\gamma(\lambda)\big) \leq \lambda f(x) + (1-\lambda)f(y)$ for all $\gamma(\lambda) \in D_k$.

Accepting the claim for a moment, we can use $\bigcup_{k\in\mathbb{N}} D_k$ is dense in $\{\gamma(\lambda) : \lambda \in [0,1]\}$ and the continuity of u to deduce (7.1.1). Indeed, for any $\lambda \in [0,1]$, there exists $\big\{\lambda_k \in \{0/2^k, 1/2^k, ..., 2^k/2^k\}\big\}_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} \lambda_k = \lambda$. Then $\lim_{k\to\infty} \gamma(\lambda_k) = \gamma(\lambda)$ and the continuity of f implies through theorem 1.2.3 that $f\big(\gamma(\lambda)\big) = \lim_{k\to\infty} f\big(\gamma(\lambda_k)\big)$. From the claim we further obtain

$$f\big(\gamma(\lambda)\big) \leq \lim_{k\to\infty} \big(\lambda_k f(x) + (1-\lambda_k)f(y)\big),$$

which implies (7.1.1).

Proof of claim. This is done by induction. For $k = 1$ we have

$$D_1 = \{\gamma(0) = x,\ \gamma(1/2) = (x+y)/2,\ \gamma(1) = y\}.$$

The claim is thus trivially true for $\lambda = 0, 1$ and also true for $\lambda = 1/2$ by the hypothesis (7.3.1). Assuming next that (7.3.1) is true for all $\gamma(\lambda) \in D_k$, we prove it is true for all $\gamma(\lambda) \in D_{k+1}$. To this end let $\gamma(\lambda) \in D_{k+1} \setminus D_k$. Setting

$$\lambda_- = \lambda - 1/2^{k+1} \quad \text{and} \quad \lambda_+ = \lambda + 1/2^{k+1},$$

we readily have $\gamma(\lambda_-), \gamma(\lambda_+) \in D_k$. Therefore, by the induction hypothesis,

$$f\big(\gamma(\lambda_-)\big) \leq \lambda_- f(x) + (1-\lambda_-)f(y),$$
$$f\big(\gamma(\lambda_+)\big) \leq \lambda_+ f(x) + (1-\lambda_+)f(y).$$

Adding these we get

$$\frac{1}{2} f\big(\gamma(\lambda_-)\big) + \frac{1}{2} f\big(\gamma(\lambda_+)\big) \leq \lambda f(x) + (1-\lambda)f(y),$$

and because of (7.3.1), we conclude

$$f\Big(\frac{\gamma(\lambda_-) + \gamma(\lambda_+)}{2}\Big) \leq \lambda f(x) + (1-\lambda)f(y).$$

The proof follows since $\gamma(\lambda_-) + \gamma(\lambda_+) = 2\gamma(\lambda)$. ■

Example 7.3.2. The function $f : \mathbb{R}^n \to \mathbb{R}$ given by $f(x) = |x|^2 - d_F^2(x)$, where d_F is the distance function to a nonempty closed subset $F \subsetneq \mathbb{R}^n$, is convex.

Proof. From example 7.2.3-(i) we know f is continuous, hence because of the above theorem we only have to prove $f(x/2+y/2) \leq f(x)/2 + f(y)/2$. To this end pick $z \in F$ such that

$$d_F\Big(\frac{x+y}{2}\Big) = \Big|z - \frac{x+y}{2}\Big|.$$

Then by expanding the square

$$f\Big(\frac{x+y}{2}\Big) = -|z|^2 + z \cdot (x+y).$$

On the other hand

$$\begin{aligned}\frac{f(x)+f(y)}{2} &= \frac{|x|^2+|y|^2-d_F^2(x)-d_F^2(y)}{2}\\ &\geq \frac{|x|^2+|y|^2-|x-z|^2-|y-z|^2}{2}\\ &= -|z|^2 + z\cdot(x+y),\end{aligned}$$

as required. ■

Bibliography

[Br] Brezis, H. *Functional analysis, Sobolev spaces and partial differential equations.* Universitext. Springer 2011.

[DB] DiBenedetto, E. *Real analysis.* 2nd ed. Birkhäuser Adv. Texts. Basler Lehrbücher. Birkhäuser 2016.

[EvG] Evans, L. C., Gariepy, R. F. *Measure theory and fine properties of funcions.* Stud. Adv. Math. CRC Press 1992.

[F] Federer, H. *Geometric measure theory.* Grundlehren Math. Wiss. **153**, Springer 1969.

[KrP] Krantz, S., Parks H. *The geometry of domains in space.* Birkhäuser Adv. Texts. Basler Lehrbücher. Birkhäuser 1999.

[LL] Lieb, E. H., Loss, M. *Analysis.* 2nd ed. Grad. Stud. Math. **14**. Amer. Math. Soc. 2001.

[Mg] Maggi, F. *Sets of finite perimeter and geometric variational problems. An introduction to geometric measure theory.* Cambridge Stud. Adv. Math. **135**. Cambridge Univ. Press 2012.

[Mz] Maz'ya, V. G. *Sobolev spaces.* Translated from Russian by T. Shaposhnikova. Springer Ser. Soviet Math., Springer 1985.

[Schn] Schneider, R. *Convex Bodies: The Brunn-Minkowski theory.* 2nd expanded ed. Encyclopedia Math. Appl. **151**. Cambridge Univ. Press 2014.

[V] Villani, C. *Topics in Optimal Transportation.* (2nd edition). Grad. Stud. Math. **58**. Amer. Math. Soc. 2016

[WhZ] Wheeden, R. L.; Zygmund, A. *Measure and integral. An introduction to real analysis.* Pure Appl. Math. **43**. Marcel Dekker 1977.

[W] Willem, M. *Functional analysis.* Fundamentals and applications. Cornerstones. Birkhäuser/Springer 2013.

[Z2] Ziemer, W. P. *Modern real analysis.* 2nd ed. with contributions by Monica Torres. Grad. Texts in Math. **278**. Springer 2017.