ANALYSIS I & II

VERSION: December 8, 2021

ARMIN SCHIKORRA

Contents

In Analysis
there are no theorems
only proofs

These lecture notes take great inspiration from the lecture notes by Michael Struwe (Analysis III, German), as well as by Piotr Hajłasz (Analysis I). We will also follow the presentations in Evans-Gariepy [Evans and Gariepy, 2015] (measure theory), Grafakos [Grafakos, 2014] (Fourier Analysis) and wikipedia. Sometimes we follow those sources verbatim.

Pictures that were not taken from above mentioned sources or wikipedia are usually made with geogebra.

References

[Adams and Fournier, 2003] Adams, R. A. and Fournier, J. J. F. (2003). *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition.

[Bandle, 2017] Bandle, C. (2017). Dido's problem and its impact on modern mathematics. *Notices Amer. Math. Soc.*, 64(9):980–984.

[Brezis, 2011] Brezis, H. (2011). *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York.

[Clason, 2020] Clason, C. ([2020] ©2020). *Introduction to functional analysis*. Compact Textbooks in Mathematics. Birkhäuser/Springer, Cham.

[Evans, 2010] Evans, L. C. (2010). *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition.

[Evans and Gariepy, 2015] Evans, L. C. and Gariepy, R. F. (2015). *Measure theory and fine properties of functions*. Textbooks in Mathematics. CRC Press, Boca Raton, FL, revised edition.

[Gilbarg and Trudinger, 2001] Gilbarg, D. and Trudinger, N. S. (2001). *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin. Reprint of the 1998 edition.

[Grafakos, 2014] Grafakos, L. (2014). *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics*. Springer, New York, third edition.

[Hajłasz and Liu, 2010] Hajłasz, P. and Liu, Z. (2010). A compact embedding of a Sobolev space is equivalent to an embedding into a better space. *Proc. Amer. Math. Soc.*, 138(9):3257–3266.

[Leoni, 2017] Leoni, G. (2017). *A first course in Sobolev spaces*, volume 181 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition.

[Lieb and Loss, 2001] Lieb, E. H. and Loss, M. (2001). *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition.

[Maz'ya, 2011] Maz'ya, V. (2011). *Sobolev spaces with applications to elliptic partial differential equations*, volume 342 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, augmented edition.

[Schikorra et al., 2017] Schikorra, A., Spector, D., and Van Schaftingen, J. (2017). An L^1-type estimate for Riesz potentials. *Rev. Mat. Iberoam.*, 33(1):291–303.

[Shor, 1997] Shor, P. W. (1997). Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509.

[Simon, 1996] Simon, L. (1996). *Theorems on regularity and singularity of energy minimizing maps*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel. Based on lecture notes by Norbert Hungerbühler.

[Sokal, 2011] Sokal, A. D. (2011). A really simple elementary proof of the uniform boundedness theorem. *Amer. Math. Monthly*, 118(5):450–452.

[Stein and Shakarchi, 2003] Stein, E. M. and Shakarchi, R. (2003). *Fourier analysis*, volume 1 of *Princeton Lectures in Analysis*. Princeton University Press, Princeton, NJ. An introduction.

[Talenti, 1976] Talenti, G. (1976). Best constant in Sobolev inequality. *Ann. Mat. Pura Appl. (4)*, 110:353–372.

[Ziemer, 1989] Ziemer, W. P. (1989). *Weakly differentiable functions*, volume 120 of *Graduate Texts in Mathematics*. Springer-Verlag, New York. Sobolev spaces and functions of bounded variation.

Index

Part 1. Analysis I: Measure Theory

1. Measures, σ-Algebras

A measure is a way to measure (hence the name!) volumes. So for some set X it should be a map

$$\mu : 2^X \to [0, \infty]$$

that to a subset $A \subset X$ assigns the volume $\mu(A)$. Here 2^X denotes the potential set, i.e. the collection of subsets of X.

What would we want from a volume in $\mathbb{R}^n$? Well it seems to be a reasonable assumption to axiomatically assume the following

- For any $A \subset \mathbb{R}^n$ we have $\mu(A) \in [0, \infty]$
- (Invariance under translation and rotation) For any set $A \subset \mathbb{R}^n$, any rotation $P \in O(n)$ and any vector $x \in \mathbb{R}^n$ we have $\mu(x + OA) = \mu(A)$ where we denote

$$x + OA := \{x + Oa \in \mathbb{R}^n : \quad a \in A\}$$

- For any $A, B \subset \mathbb{R}^n$ disjoint we have $\mu(A \cup B) = \mu(A) + \mu(B)$

As reasonable as that sounds, there are two problems here:

- For $n \geq 3$ the only map $\mu : 2^{\mathbb{R}^n} \to [0, \infty]$ that satisfies our axiom is constant (Hausdorff, 1914)
- For $n = 1, 2$ there are indeed nonconstant maps $\mu : 2^{\mathbb{R}^n} \to [0, \infty]$ that satisfy the above axioms, however even if we fix $\mu([0, 1]^n) := 1$ there is more than one possibility for such a μ (Banach 1923).
- the whole business about disjoint sets is really tricky, as illustrated by the *Banach-Tarski-Paradoxon* (1924):

  Let $n \geq 3$, A and B be bounded sets with $int(A)$ and $int(B) \neq \emptyset$. Then there exist finitely many $(x_i)_{i=1}^N \subset \mathbb{R}^n$, $(O_i)_{i=1}^N \subset O(n)$ and parwise disjoint sets $(C_i)_{i=1}^N$ so that $(x_i + O_i C_i)_{i=1}^N$ are pairwise disjoint and

$$A = \bigcup_{i=1}^N C_i, \quad \text{and} \quad B = \bigcup_{i=1}^N (x_i + O_i C_i) . \tag{1.1}$$

  That is we can deconstruct any set A in $\mathbb{R}^n$ into disjoint sets, move them around (without any scaling!) and obtain another completely different set B - see Figure 1.1.

This is crazy, so the axiomatic definition of a reasonable volume in $\mathbb{R}^n$ has failed, and we are back to square one.

FIGURE 1.1. A ball can be decomposed into a finite number of disjoint sets and then reassembled into two balls identical to the original

So instead of defining a volume in $\mathbb{R}^n$ axiomatically, let us generally define what a reasonable notion of a volume should satisfy. Later we will then construct the Lebesgue measure that has most of the desired properties on $\mathbb{R}^n$.

Clearly $\mu(\emptyset) = 0$ is a reasonable assumption. Ideally we would also like $\mu(A \cup B) = \mu(A) \cup \mu(B)$ – but this will be a surprisingly tricky, confusing, and paradox assumption, so let us settle for the following notion

Definition 1.1. Let X be any set and 2^X the potential set of X. A map $\mu : 2^X \to [0, \infty]$ is a *measure* on X if we have

(1) $\mu(\emptyset) = 0$
(2) $\mu(A) \leq \sum_{k=1}^{\infty} \mu(A_k)$ whenever $A, A_k \subset X$, $k \in \mathbb{N}$ and $A \subset \bigcup_{k\in\mathbb{N}} A_k$

Remark 1.2. Condition (1) and (2) implies *monotonicity*,

$$\mu(A) \leq \mu(B) \quad \forall A \subset B.$$

(simply set $A_1 := B$ and $A_k := \emptyset$ for $k \geq 2$).

In particular we could equivalently replace (2) above by *σ-subadditivity*, namely

$$\mu(\bigcup_{k=1}^{\infty} A_k) \leq \sum_{k=1}^{\infty} \mu(A_k).$$

Remark 1.3.
- A word of warning: we will use here the notion of an *outer measure* that is defined on all of 2^X, not only on its σ-algebra of measurable sets.
- In particular we have $\mu(A \cup B) \leq \mu(A) + \mu(B)$ for any set $A, B \subset X$. However, in general, we cannot hope for all disjoint sets A and B hope that $\mu(A \cup B) = \mu(A) + \mu(B)$ (see above), this will lead to the notion of *non-measurable sets*.

Example 1.4 (Jordan content).
- The outer *Jordan content* $J_*(E)$ of a set $E \subset \mathbb{R}^n$ is defined as follows.

  For a product of bounded cubes $C = [a_1, b_1) \times [a_2, b_2) \times \ldots \times [a_n, b_n)$ we set

$$\mathrm{vol}(C) := (b_1 - a_1) \cdot (b_2 - a_2) \cdot \ldots (b_n - a_n).$$

$$J^*(E) := \inf \left\{ \sum_{i=1}^{N} \mathrm{vol}(C_i) \quad \text{for some } N \in \mathbb{N}, \text{ and cubes } (C_i)_{i=1}^N \text{ such that } E \subset \bigcup_{i=1}^{N} C_i \right\}$$

Here we follow the convention that $\inf \emptyset = +\infty$.

$J^\varepsilon(\cdot)$ is *not* a measure: take any enumeratotion of $\mathbb{Q} \cap [0,1] = \{q_1, \ldots, q_n, \ldots\}$. Set $A_k := \{q_k\}$ and $A := \bigcup_{k=1}^\infty A_k = [0,1] \cap \mathbb{Q}$. If $(C_i)_{i=1}^N$ is a finite cover of $[0,1] \cap \mathbb{Q}$ then $\bigcup_i \overline{C_i} \supset [0,1]$[1], so $J^*(A) = 1$. However $J^*(A_k) = 0$ for each k, we have $J^*(A) \not\leq \sum_{k=1}^\infty J^*(A_k)$.

However J^* satisfies finite additivity,

$$\mu(A \cup B) \leq \mu(A) + \mu(B),$$

i.e.

$$\mu(A) \leq \sum_{k=1}^N \mu(A_k) \quad \text{whenever } A, A_k \subset X,\ k \in \{1, \ldots, N\},\ N \in \mathbb{N}, \text{ and } A \subset \bigcup_{k \in \mathbb{N}} A_k.$$

Such a map $J^\varepsilon : 2^X \to [0, \infty)$ is called a *content*.

- The countable version of the outer Jordan content, is called the *Lebesgue outer measure*

$$m^*(E) := \inf \left\{ \sum_{i=1}^\infty \mathrm{vol}(C_i) \quad \text{for some , and cubes } (C_i)_{i=1}^\infty \text{ such that } E \subset \bigcup_{i=1}^\infty C_i \right\} \tag{1.2}$$

It is again clear that $m^*(\emptyset) = 0$. Let now $A \subset \bigcup_{k=1}^n A_k$. We may assume that $m^*(A_k) < \infty$ otherwise there is nothing to show. Fix $\varepsilon > 0$. For each k we can pick $(C_{k;i})_{i=1}^\infty$ such that $\bigcup_{i=1}^\infty C_{k;i} \supset A_k$ and

$$\sum_{i=1}^\infty \mathrm{vol}(C_{k,i}) \leq m^*(A_k) + \frac{\varepsilon}{2^k}.$$

Now $\bigcup_{k,i \in \mathbb{N}} C_{k,i} \supset A$ and thus

$$m^*(A) \leq \sum_{k,i \in \mathbb{N}} \mathrm{vol}(C_{k,i}) \leq \sum_{k=1}^\infty m^*(A_k) + \sum_{k=1}^\infty \frac{\varepsilon}{2^k}$$

That is, we have shown that for any $\varepsilon > 0$,

$$m^*(A) \leq \sum_{k=1}^\infty m^*(A_k) + \varepsilon$$

Taking $\varepsilon \to 0$ we conclude that $m^*(A) \leq \sum_{k=1}^\infty m^*(A_k)$ – that is $m^*(A)$ is indeed a measure.

Later the Lebesgue measure $\mathcal{L}^n$ will coincide with $m^*(A)$.

- The *inner Jordan content*,

$$J_*(E) := \sup \left\{ \sum_{i=1}^N \mathrm{vol}(C_i) \quad \text{for some } N \in \mathbb{N}, \text{ and cubes } (C_i)_{i=1}^N \text{ such that } \bigcup_{i=1}^N C_i \subset E \right\}$$

Here we follow the convention that $\sup \emptyset = 0$.

[1]Indeed, take $r \in [0,1]$ then there exists q_k converging to r, q_k belongs infinitely often to the same interval, so $r \in \overline{C_i}$ for some i

Still $J_*(\cdot)$ is not a measure. Take $A_1 := [0,1]\backslash\mathbb{Q}$ and for $i \geq 2$ we set $A_i = \{q_i\}$ for $\{q_2, \dots, \} = \mathbb{Q} \cap [0,1]$ any enumeration of $\mathbb{Q} \cap [0,1]$. Since A_1 has empty interior we have $J_*(A_1) = 0$. Similarly, $J_*(A_i) = 0$ for $i \geq 2$. However $A := \bigcup_{i=1}^{\infty} A_i = [0,1]$ satisfies $J_*([0,1]) = 1$. So we have $J_*(A) \not\leq \sum_{i=1}^{n} J_*(A_i)$.

- If we simply make the innter Jordan content countable, i.e. if we set

$$\tilde{J}_*(E) := \sup\left\{\sum_{i=1}^{\infty} \mathrm{vol}(C_i) \quad \text{for cubes } (C_i)_{i=1}^{N} \text{ such that } \bigcup_{i=1}^{\infty} C_i \subset E\right\}$$

we run into the same problem as for J_*, namely $J_*([0,1]\backslash\mathbb{Q}) = 0$. So $\tilde{J}_*(E)$ is still not a measure.

Example 1.5 (Counting measure)**.** Let X be any set. Then $\#2^X \to \mathbb{N} \cup \{0\}$ defined by

$$\#A := \text{number of elements in } A,$$

is a measure, called the *counting measure*.

Exercise 1.6. *Let X be a metric space and $\mu : 2^X \to [0,\infty]$ a measure. Let $A \subset X$ then the measure $\mu \llcorner A$: $2^X \to [0,\infty]$ given by*

$$(\mu \llcorner A)(B) := \mu(A \cap B)$$

is a measure.

1.1. **Example: Hausdorff measure.** Let (X, d) be a metric space.

Definition 1.7. The s-dimensional *Hausdorff measure*, $s > 0$ is defined as follows.

Let $\delta \in (0,\infty]$, then for any $A \subset X$ we define

$$\mathcal{H}^s_\delta(A) := \alpha(s) \inf\left\{\sum_{k=1}^{\infty} r_k^s : \quad A \subset \bigcup_{k=1}^{\infty} B(x_k, r_k), \quad r_k \in (0,\delta)\right\}.$$

Here $B(x_k, r_k)$ are open balls with radius r centered at x_k, i.e.

$$B(x_k, r_k) := \{y \in X : d(x_k, y) < r_k\}.$$

Moreover[2]

$$\alpha(s) := \frac{\pi^{\frac{s}{2}}}{\Gamma(\frac{s}{2}+1)}.$$

where Γ is the Γ-function.

Now observe that $\delta \mapsto \mathcal{H}^s_\delta(A)$ is monotonce decreasing. So we can write

$$\mathcal{H}^s(A) := \lim_{\delta \to 0^+} \mathcal{H}^s_\delta(A) \equiv \sup_{\delta > 0} \mathcal{H}^s_\delta(A) \in [0,\infty].$$

Often one writes $\mathcal{H}^0(A) := \#A$, the *counting measure*.

$\mathcal{H}^s_\infty$ is called the *Hausdorff content*.

[2]Warning: Some authors set $\alpha(s) := 1$. The main reason to not do that is so that $\mathcal{H}^n = \mathcal{L}^n$ in $\mathbb{R}^n$

Remark 1.8. • Observe that while $\mathcal{H}^s_\delta(A) < \infty$ whenever $s > 0$, $\delta > 0$ and A is any bounded set, as $\delta \to 0$ $\mathcal{H}^s(A)$ will be infinite whenever s is smaller than the "dimension of A" (a notion we will define more carefully below).

Lemma 1.9. *$\mathcal{H}^s$ is a measure in $\mathbb{R}^n$.*

Proof. One can show similar to the argument for m^* that $\mathcal{H}^s_\delta(\cdot)$ is a measure for each $\delta > 0$.

We clearly have $\mathcal{H}^s(\emptyset) = 0$. Moreover, since $\mathcal{H}^s_\delta$ is a measure for any $\delta > 0$, we have for any $A \subset \bigcup_{k=1}^\infty A_k$,

$$\mathcal{H}^s_\delta(A) \leq \sum_{k=1}^\infty \mathcal{H}^s_\delta(A_k) \leq \sum_{k=1}^\infty \mathcal{H}^s(A_k).$$

Taking the supremum over δ in this inequality we have σ-additivity for $\mathcal{H}^s$.

$$\mathcal{H}^s(A) \leq \sum_{k=1}^\infty \mathcal{H}^s(A_k).$$

□

Exercise 1.10. *Show that*

$$\mathcal{H}^0_\delta(\mathbb{Q}) \xrightarrow{\delta \to 0} \infty.$$

Lemma 1.11. *$\mathcal{H}^s$ is a* metric outer measure *that means that if $A, B \subset (X, d)$ satisfy*

$$d(A, B) := \inf_{a \in A, b \in B} d(a, b) > 0.$$

then

$$\mathcal{H}^s(A \cup B) = \mathcal{H}^s(A) + \mathcal{H}^s(B).$$

Proof. This is relatively easy to see. Take $\delta < \frac{d(A,B)}{3}$, then since any covering $B(x, r)$ with $r < \delta$ cannot contain points of both A or B at the same time, we have that $\mathcal{H}^s_\delta$ is a metric outer measure. Taking $\delta \to 0^+$ we obtain that $\mathcal{H}^s$ is a metric outer measure. □

Remark 1.12. One can, and we will in Corollary 1.78, show that the n-dimensional Hausdorff measure in $\mathbb{R}^n$ coincides with the Lebesgue measure $\mathcal{L}^n$, i.e.

$$\mathcal{L}^n(A) = \mathcal{H}^n(A).$$

Exercise 1.13. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be (uniformly) Lipschitz continuous that is*

$$|f(x) - f(y)| \leq L|x - y| \quad \forall x, y \in \mathbb{R}^n$$

Then for any set Ω and any $s \geq 0$,

$$\mathcal{H}^s(f(\Omega)) \leq C(L)\mathcal{H}^s(\Omega)$$

where $C(L)$ is a constant only depending on L.

Exercise 1.14. *Show that*

- $\mathcal{H}^1 = \mathcal{L}^1$ *in* $\mathbb{R}$

- $\mathcal{H}^s(\lambda A) = \lambda^s \mathcal{H}^s(A)$ *for all* $\lambda > 0$, *where* $\lambda A = \{\lambda x : \quad x \in A\}$.
- $\mathcal{H}^s(LA) = \mathcal{H}^s(A)$ *whenever* $L : \mathbb{R}^n \to \mathbb{R}^n$ *is an affine isometry, i.e. if* $Lx = Ax + b$ *for* $A \in O(n)$ *and* $b \in \mathbb{R}^n$ *constant.*

Exercise 1.15. *Let* $U \subset \mathbb{R}^n$ *be any non-empty open set. Then* $\mathcal{H}^s(U) = \infty$ *for all* $s < n$.

Exercise 1.16 (translation and rotation invariant)**.** *Let* $A \subset \mathbb{R}^n$ *and* $s \in (0, \infty)$. *Show the following*

(1) If $p \in \mathbb{R}^n$ *then* $\mathcal{H}^s(p + A) = \mathcal{H}^s(A)$.
(2) If $O \in O(n)$ *(i.e.* $O \in \mathbb{R}^{n\times n}$ *and* $O^t O = I$*) then* $\mathcal{H}^s(OA) = \mathcal{H}^s(A)$.
(3) If $A \subset \mathbb{R}^\ell \times \{0\}$ *for* $0 < \ell < n$ *and* $\pi : (x_1, \dots, x_n) := (x_1, \dots, x_\ell)$ *is the projection from* $\mathbb{R}^n = \mathbb{R}^\ell \times \mathbb{R}^{n-\ell}$ *to* $\mathbb{R}^\ell$, *then* $\mathcal{H}^s_{\mathbb{R}^n}(A) = \mathcal{H}^s_{\mathbb{R}^\ell}(\pi(A))$.

Lemma 1.17. *Let* $0 \le s < t < \infty$.

(1) If $\mathcal{H}^s(A) < \infty$ *then* $\mathcal{H}^t(A) = 0$
(2) If $\mathcal{H}^t(A) > 0$ *then* $\mathcal{H}^s(A) = \infty$.

Proof. Indeed, whenever $r_k \le \delta$ and $(B(x_k, r_k))_{k\in\mathbb{N}}$ cover A we have

$$\mathcal{H}^t_\delta(A) \le \alpha(t) \sum_{k=1}^\infty r_k^t \le \alpha(t)\delta^{t-s} \sum_{k=1}^\infty r_k^s.$$

Taking the infimum over any such covering $B(x_k, r_k)$ of A we find

$$\mathcal{H}^t_\delta(A) \le \frac{\alpha(t)}{\alpha(s)} \delta^{t-s} \mathcal{H}^s_\delta(A).$$

Taking $\lim_{\delta\to 0}$ on both sides we obtain

$$\mathcal{H}^t(A) \le \frac{\alpha(t)}{\alpha(s)} 0 \cdot \mathcal{H}^s(A).$$

This implies that if $\mathcal{H}^t(A) > 0$ then necessarily $\mathcal{H}^s(A) = \infty$, and if $\mathcal{H}^s(A) < \infty$ then $\mathcal{H}^t(A) = 0$. □

Example 1.18. If $k \in \mathbb{N}$ it is conceivable that $\mathcal{H}^k$ measures something of "dimension k". For example assume that $C = [0, 1]^2 \times \{0\} \subset \mathbb{R}^3$ is a 2D-square of sidelength 1. We need $\approx \frac{1}{\delta^2}$ many balls to cover C. Then

$$\mathcal{H}^s_\delta(C) \le \alpha(s) \frac{1}{\delta^2} \delta^s.$$

So if $s > 2$ we see that $\mathcal{H}^s(C) \le \lim_{\delta\to 0} \delta^{s-2} = 0$. That is C has no s-volume for $s > 2$.

For $s = 2$ one can argue that covering uniformly by balls of radius δ is optimal and thus we have

$$0 < \mathcal{H}^2(C) < \infty.$$

In particular $\mathcal{H}^s(C) = \infty$ for any $s < 2$.

(this argument is easy to generalize to a ℓ-dimensional manifold in $\mathbb{R}^N$)

Indeed, with the Hausdorff measure we can define a dimension

Definition 1.19. The *Hausdorff dimension* is defined as

$$\dim_{\mathcal{H}} A := \inf \{ s \geq 0 : \quad \mathcal{H}^s(A) = 0 \} .$$

If $\mathcal{H}^s(A) > 0$ for all $s > 0$ then $\dim_{\mathcal{H}}(E) := \infty$.

Lemma 1.20. *Let C be a set in a metric space and let $s \geq 0$*

(1) If $\mathcal{H}^s(C) = 0$ then $\dim_{\mathcal{H}}(E) \leq s$.
(2) If $\mathcal{H}^s(C) > 0$ then $\dim_{\mathcal{H}}(E) \geq s$.
(3) If $0 < \mathcal{H}^s(C) < \infty$ then $\dim_{\mathcal{H}}(E) = s$.
(4) If $\mathcal{H}^s_\infty(C) > 0$ and $\mathcal{H}^s(C) < \infty$ then $\dim_{\mathcal{H}}(E) = s$.

Proof. This follows from Lemma 1.17 and the definition of Hausdorff measure.

(1) follows from the definition of the Hausdorff measure as infimum. then $\dim_{\mathcal{H}}(E) \leq s$.
(2) If $\mathcal{H}^s(C) > 0$ then by Lemma 1.17 $\mathcal{H}^t(C) = \infty$ for all $t < s$. Again from the definition it is clear that $\dim_{\mathcal{H}}(E) \geq s$.
(3) This is a consequence of the two above statements.
(4) Follows from the statement before since $\mathcal{H}^s_\infty(C) \leq \mathcal{H}^s(C)$

□

Exercise 1.21 (Hausdorff dimension under Lipschitz and Hölder maps)**.** *Let (X, d_x) and (Y, d_Y) be two metric spaces and let $f : X \to Y$. Assume that $A \subset X$ has Hausdorff-dimension $\dim_{\mathcal{H}}(A) = s$.*

(1) If f is uniformly Lipschitz continuous, i.e. for some $L > 0$,

$$d_Y(f(x), f(y)) \leq L\, d(x, y) \quad \forall x, y \in X$$

then $\dim_{\mathcal{H}}(f(A)) \leq s$.
(2) Give an example where $\dim_{\mathcal{H}}(A) < s$
(3) Assume f is uniformly Hölder continuous, i.e. for some $L > 0$ and $\alpha > 0$

$$d_Y(f(x), f(y)) \leq L\, d(x, y)^\alpha \quad \forall x, y \in X$$

What can we say about the Hausdorff dimension of $f(A) \subset Y$?

Cf Exercise 1.13.

Example 1.22. The Cantorset is defined as follows.

$$C_0 := [0, 1]$$

Let $C_0 := [0, 1]$. In the k-th step we construct C_k by removing of each interval the open middle interval of size 3^{-n}. For exmaple

$$C_1 := [0, \frac{1}{3}] \cup [\frac{1}{3}, 1].$$

FIGURE 1.2. The cantor set

See Figure 1.2.

Set $C := \bigcap_{k=1}^{\infty} C_k$. Observe that C is closed and bounded, so compact.

Lemma 1.23. $\dim_{\mathcal{H}}(C) = \frac{\log 2}{\log 3}$.

Proof. For each $k \in \mathbb{N}$ we have $C \subset C_k$. Observe that C_k consists of 2^k disjoint intervals each of diameter 3^{-k} (i.e. radius $\frac{1}{2}3^{-k}$). Thus for any $\delta > 0$ and for any $k \gg 1$ so that $\frac{1}{2}3^{-k} < \delta$ we have

$$\mathcal{H}^s_\delta(C) \leq \alpha(s) \sum_{\ell=1}^{2^k} (\frac{1}{2}3^{-k})^s = 2^{-s}\left(\frac{2}{3^s}\right)^k \xrightarrow{k\to\infty} \alpha(s) \begin{cases} 2^{-s} & s = \frac{\log 2}{\log 3} \\ 0 & s > \frac{\log 2}{\log 3} \\ \infty & s > \frac{\log 2}{\log 3} \end{cases}$$

In particular we have

$$\mathcal{H}^s(C) = 0 \quad \forall s > \frac{\log 2}{\log 3}.$$

So from the definition of the Hausdorff dimension we get

$$\dim_{\mathcal{H}} C \leq \frac{\log 2}{\log 3}.$$

Now we need to show the <u>other direction</u>. From now on set $s := \frac{\log 2}{\log 3}$. Let $(B(x_i, r_i))_{i=1}^{\infty}$ be any covering of C. We claim that

$$\sum_{i=1}^{\infty} r_i^s \geq \frac{1}{2^s 4}. \tag{1.3}$$

Once we have (1.3) we are done, because (1.3) implies

$$\mathcal{H}^s\infty(C) \geq \frac{1}{2^s 4}.$$

In particular (recall that $s = \frac{\log 2}{\log 3}$) we have $\infty > \mathcal{H}^s(C) \geq \mathcal{H}^s_\infty(C) > 0$.

Let us make some notation. Denote by A_k the intervals of C_k, i.e. A_k consists of pairwise disjoint, closed intervals in $\mathbb{R}$ such that $C_k = \bigcup_{I \in A_k} I$. E.g.

$$C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1], \quad A_1 = \{[0, \frac{1}{3}], [\frac{2}{3}, 1]\}.$$

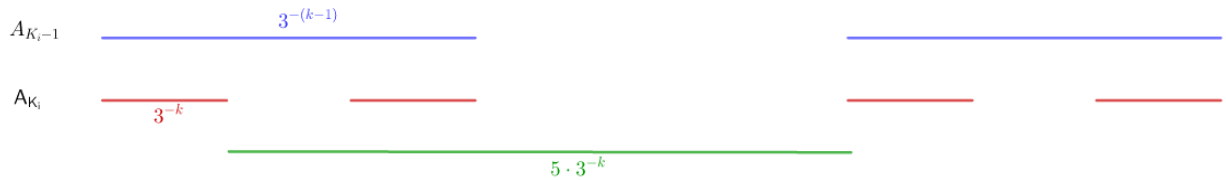

FIGURE 1.3. If a ball intersects three intervals of A_{K_i} its *diameter* is at least $5 \cdot 3^{-K_i}$

Proof of (1.3) Since C is compact, we may assume that there finitely many, w.l.o.g. the first N balls $(B(x_i, r_i))_{i=1}^N$ already cover C. We may assume that each $r_i < \frac{1}{2}$, otherwise (1.3) is obvious.

Fix $i \in \{1, \dots, N\}$.

Let $K_i \in \mathbb{N} \cup \{0\}$ so that

$$2r_i \in [3^{-K_i-1}, 3^{-K_i}).$$

Now we consider the construction step C_{K_i}. Each ball $B(x_i, r_i)$ has nonempty intersection with at most 2 intervals of C_{K_i}. Indeed, otherwise its *diameter* would be at least $5 \cdot 3^{-K_i}$, see Figure 1.3.

But then $B(x_i, r_i)$ has nonempty intersection with at most $2 \cdot 2^{j-K_i}$ intervals of C_j for any $j \geq K_i$. Since $s = \frac{\log 2}{\log 3}$ we have

$$2 \cdot 2^{j-K_i} = 2^{j+1} 2^{-K_i} = 2^{j+1} 3^{-K_i s} \leq 2^{j+1} 3^s (2r_i)^s = 2^{j+2} (2r_i)^s.$$

Set now $K := \max_{\{i=1,\dots,N\}} K_i$.

Then for any $i \in \{1, \dots, N\}$ each of the balls $B(x_i, r_i)$ has nonempty intersection with at most $2^{K+2}(2r_i)^s$ many intervals of A_K.

So if we set Γ_i to be the number of intervals in A_K that intersect $B_{r_i}(x_i)$ we have $\Gamma_i \leq 2^{K+2}(2r_i)^s$ and thus

$$\begin{aligned} \sum_{i=1}^N \Gamma_i (3^{-K})^s \leq & \sum_{i=1}^N \underbrace{(3^{-K})^s}_{=2^{-K}} 2^{K+2} (2r_i)^s \\ = & 4\ 2^s \sum_{i=1}^N (r_i)^s \end{aligned} \tag{1.4}$$

Now for each $x \in C$ there is exactly one interval I in A_K such that $x \in I$. Since $(B_{r_i}(x_i))_{i=1}^N$ covers all of C we have the following: for each interval I in A_K there exists some $i \in \{1, \dots, N\}$ such that $B_{r_i}(x_i) \cap I \neq \emptyset$. That is,

$$\sum_{i=1}^N \Gamma_i \geq \text{number of intervals in } A_K = 2^K.$$

FIGURE 1.4. The fat cantor set for $a = \frac{1}{4}$, see Example 1.24

Thus,

$$\sum_{i=1}^{N} \Gamma_i 3^{-Ks} \geq 2^K 3^{-Ks} = 1. \tag{1.5}$$

Together, (1.4) and (1.5) imply (1.3). □

Example 1.24. The Smith–Volterra–Cantor set, aka fat cantor set is defined as follows.

Let $C_0 := [0, 1]$. In the k-th step we construct C_k by removing of each interval the open middle interval of size a^n. That is

$$C_1 = [0, \frac{1-a}{2}] \cup [\frac{1+a}{2}, 1].$$

$$C_2 = [0, \frac{1-a}{4} - \frac{a^2}{2}] \cup [\frac{1-a}{4} + \frac{a^2}{2}, \frac{1-a}{2}] \cup [\frac{1+a}{2}, \frac{1+\frac{1+a}{2}}{2} - \frac{a^2}{2}] \cup [\frac{1+\frac{1+a}{2}}{2} + \frac{a^2}{2}, 1].$$

Cf. Figure 1.4.

Set $C := \bigcap_{k=1}^{\infty} C_k$. For $a = \frac{1}{3}$ this is the typical *Cantor set*. For $a = \frac{1}{4}$ this is the *Fat Cantor set*.

Exercise 1.25. *The fat Cantor above set has positive $\mathcal{H}^1$-measure.*

1.2. **Measurable sets.** As we have discussed, our definition of measure does not include the "natural" condition that $\mu(B) = \mu(B \cap A) + \mu(B \backslash A)$ for all $A, B \subset X$ – because this "natural" condition leads to incompatibility such as the Banach-Tarski Paradoxon.

So we will denote the class of sets $A \subset 2^X$ where we have the above "natural" condition as the σ-algebra of measurable sets.

Definition 1.26 (Carathéodory)**.** Let μ be a measure on X.

$A \subset X$ is called *μ-measurable* if

$$\mu(B) = \mu(A \cap B) + \mu(B \backslash A) \quad \text{for any } B \subset X$$

Remark 1.27. By additivity of the measure, measurability is equivalent to

$$\mu(B) \geq \mu(A \cap B) + \mu(B \backslash A) \quad \text{for any } B \subset X$$

Exercise 1.28. *Let $X \neq \emptyset$ be any set*

- *and assume $\mu(\emptyset) = 0$ and $\mu(A) = 1$ for any $A \neq \emptyset$. Then A is μ-measurable if and only if $A = \emptyset$ or $A = X$.*
- *If $\nu = \#$ the counting measure then any set A is ν-measurable.*

Clearly, whatever choice of measure we have, $\emptyset$ and X are measurable sets. We also have

$$(A_i)_{i=1}^N \text{ are measurable} \quad \Rightarrow \quad \bigcup_{i=1}^N A_i \text{ is measurable} \tag{1.6}$$

Proof of (1.6). We proof this by induction. Clearly this holds for $N = 1$. So to conclude (1.6) we only need to show:

If A_1, A_2 are μ-measurable, then so is $A_1 \cup A_2$.

So assume A_1 and A_2 are μ-measurable and $B \subset X$.

$$\begin{aligned}\mu(B) =& \mu(B \setminus A_1) + \mu(B \cap A_1)\\ =& \mu\left((B \setminus A_1) \cap A_2\right) + \mu\left((B \setminus A_1) \setminus A_2\right)\\ &+ \mu\left((B \cap A_1) \cap A_2\right) + \mu\left((B \cap A_1) \setminus A_2\right)\\ \geq& \mu\left(B \setminus (A_1 \cup A_2)\right) + \mu(B \cap (A_1 \cup A_2))\end{aligned}$$

In the last step we have used that

$$\mu\left((B \setminus A_1) \cap A_2\right) + \mu\left((B \cap A_1) \cap A_2\right) + \mu\left((B \cap A_1) \setminus A_2\right) \geq \mu\left(B \setminus (A_1 \cup A_2)\right),$$

by sublinearity and the fact that

$$B \setminus (A_1 \cup A_2) = \left((B \setminus A_1) \cap A_2\right) \cup \left((B \cap A_1) \cap A_2\right) \cup \left((B \cap A_1) \setminus A_2\right).$$

By Remark 1.27 we have that $(A_1 \cup A_2)$ is also measurable. □

We have much more than that:

Lemma 1.29. *Let X be a set and μ be a measure on X.*

The collection $\mathcal{A} \subset 2^X$ of μ-measurable functions

$$\mathcal{A} := \{A \subset X : \quad A \text{ is } \mu\text{-measurable}\}$$

is a σ-algebra, that is

(1) $X \in \mathcal{A}$
(2) $A \in \mathcal{A}$ implies that $X \setminus A \in \mathcal{A}$
(3) If $(A_i)_{i=1}^\infty \subset \mathcal{A}$ then $\bigcup_{i=1}^\infty A_i \in \mathcal{A}$.[3]

In particular

[3]This is the σ in σ-algebra, σ means for countably many. If we only had for any $N \in \mathbb{N}$: $(A_i)_{i=1}^N \subset \mathcal{A}$ then $\bigcup_{i=1}^N A_i \in \mathcal{A}$, $\mathcal{A}$ would be merely an Algebra (no σ!)

- $\emptyset \in \mathcal{A}$
- *if* $(A_i)_{i=1}^{\infty} \subset \mathcal{A}$ *then* $\bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$

Proof. (1) For any $B \subset X$: since $B \cap X = B$ and $B \backslash X = \emptyset$ we have

$$\mu(B) = \mu(B) + \mu(\emptyset) = \mu(B \cap X) + \mu(B \setminus X).$$

(2) Assume that $A \in \mathcal{A}$. Set $\tilde{A} := X \backslash A$. For any $B \subset X$ we have

$$\tilde{A} \cap B = (X \backslash A) \cap B = B \backslash A,$$

and

$$B \setminus \tilde{A} = B \setminus (X \backslash A) = B \cap A.$$

Since A is measurable we then have

$$\mu(B \cap \tilde{A}) + \mu(B \setminus \tilde{A}) = \mu(B \setminus A) + \mu(B \cap A) = \mu(B).$$

(3) Let $(A_i)_{i \in \mathbb{N}} \subset \mathcal{A}$. Set $A := \bigcup_{i=1}^{\infty} A_i$.

Without loss of generality we have that $A_i \cap A_j = \emptyset$ for $i \neq j$. Indeed, otherwise we set $\tilde{A}_1 := A_1$ and $\tilde{A}_k := A_k \backslash \bigcup_{i=1}^{k-1} A_i$. By the previously proven properties and (1.6) each $\tilde{A}_k$ belongs to $\mathcal{A}$ and we have $A = \bigcup_{k=1}^{\infty} \tilde{A}_k$ – so we could work with $\tilde{A}_k$ instead of A_k.

We have by measurability of each A_k and since A_N and $\bigcup_{k=1}^{N-1} A_k$ are disjoint,

$$\begin{aligned}\mu(B \cap \bigcup_{k=1}^{N} A_k) =& \mu(B \cap \left(\bigcup_{k=1}^{N} A_k\right) \cap A_N) + \mu(\left(B \cap \bigcup_{k=1}^{N} A_k\right) \setminus A_N) \\ =& \mu(B \cap A_N) + \mu(B \cap \bigcup_{k=1}^{N-1} A_k)\end{aligned}$$

Repeating this computation $N-1$ times we obtain

$$\mu(B \cap \bigcup_{k=1}^{N} A_k) = \sum_{k=1}^{N} \mu(B \cap A_k). \tag{1.7}$$

By (1.6) and the monotonicity of μ, Remark 1.2, we then have

$$\mu(B) = \mu(B \cap \bigcup_{k=1}^{N} A_k) + \mu(B \setminus \bigcup_{k=1}^{N} A_k) \geq \sum_{k=1}^{N} \mu(B \cap A_k) + \mu(B \setminus \bigcup_{k=1}^{\infty} A_k)$$

This holds for any N, so we obtain

$$\mu(B) \geq \sum_{k=1}^{\infty} \mu(B \cap A_k) + \mu(B \setminus \bigcup_{k=1}^{\infty} A_k)$$

By the σ-subadditivity of μ we then have

$$\mu(B) \geq \mu(B \cap \bigcup_{k=1}^{\infty} A_k) + \mu(B \setminus \bigcup_{k=1}^{\infty} A_k)$$

In view of Remark 1.27 this implies measurability of $\bigcup_{k=1}^{\infty} A_k$.

□

Definition 1.30. Let $\mathcal{C} \subset 2^X$ any nonempty family of subsets of X, then

$$\sigma(\mathcal{C})$$

denotes the *σ-Algebra generated by $\mathcal{C}$*, namely the smallest σ-algebra containing $\mathcal{C}$.

Exercise 1.31.

- *$\{\emptyset, X\}$ is a σ-algebra of X*
- *2^X is a σ-algebra of X*
- *Let (X, d) be a metric space. Denote $\mathcal{O} \subset 2^X$ the family of all open sets.*

  Let $\mathcal{F}$ be the family of σ-Algebras that contain all open sets. That is, $\mathcal{A} \subset 2^X$ belongs to $\mathcal{F}$ if and only if $\mathcal{A}$ is a σ-Algebra, and any open set $O \in \mathcal{O}$ belongs to $\mathcal{A}$, i.e. $O \in \mathcal{A}$.

  Define

  $$\mathcal{B} := \bigcap \{\mathcal{A} : \quad \mathcal{A} \in \mathcal{F}\}.$$

  Show that (a) $\mathcal{F}$ is nonempty, (b) $\mathcal{B}$ is a σ-algebra and (c) $\mathcal{B}$ is the smallest σ-Algebra containing all open sets, i.e. show that $\mathcal{B} = \sigma(\mathcal{O})$.

  $\mathcal{B}$ is called the Borel σ-Algebra and a set $B \in \mathcal{B}$ is called a Borel set.

Definition 1.32. If $\mu : 2^X \to [0, \infty]$ is a measure on X, and Σ is the σ-algebra of μ-measurable sets, then once calls (X, Σ, μ) a *measure space*.

Some author choose to define measures only on their σ-algebra Σ of measurable sets, and call our definition of a measure an outer measure.

There is a reason for restricting μ only to act on measurable sets – a measure μ acts in a very intuitive way on its measurable sets!

Theorem 1.33. *Let (X, Σ, μ) be a measure space.*

Let $(A_k)_{k \in \mathbb{N}} \subset \Sigma$ (i.e. each A_k is measurable). Then we have

(1) If $A_k \cap A_\ell = \emptyset$ for $k \neq \ell$ we have

(σ-Additivity)
$$\mu(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mu(A_k)$$

(2) If $A_1 \subset A_2 \subset \ldots \subset A_k \subset \ldots$ then

$$\mu(\bigcup_{k=1}^{\infty} A_k) = \lim_{k \to \infty} \mu(A_k).$$

(3) If $\mu(A_1) < \infty$ and $A_1 \supset A_2 \supset \ldots \supset A_k \supset \ldots$ then

$$\mu(\bigcap_{k=1}^{\infty} A_k) = \lim_{k \to \infty} \mu(A_k).$$

Proof. (1) Above in (1.7) we computed (take $B = X$) that for finitely many pairwise disjoint sets

$$\mu(\bigcup_{k=1}^{N} A_k) = \sum_{k=1}^{N} \mu(A_k).$$

By monotonicity

$$\mu(\bigcup_{k=1}^{\infty} A_k) \geq \mu(\bigcup_{k=1}^{N} A_k) = \sum_{k=1}^{N} \mu(A_k)$$

Taking $N \to \infty$ we find

$$\mu(\bigcup_{k=1}^{\infty} A_k) \geq \sum_{k=1}^{\infty} \mu(A_k).$$

By σ-subadditivity,

$$\mu(\bigcup_{k=1}^{\infty} A_k) \geq \sum_{k=1}^{\infty} \mu(A_k) \geq \mu(\bigcup_{k=1}^{\infty} A_k).$$

The number on left- and right-hand side are the same so we have

$$\sum_{k=1}^{\infty} \mu(A_k) = \mu(\bigcup_{k=1}^{\infty} A_k).$$

(2) Let $\tilde{A}_1 := A_1$ and $\tilde{A}_k := A_k \backslash A_{k-1}$. Then $(\tilde{A}_k)_{k=1}^{\infty}$ is pairwise disjoint, and each $\tilde{A}_k$ is measurable. So

$$\begin{aligned}
\mu(\bigcup_{k=1}^{\infty} A_k) =& \mu(\bigcup_{k=1}^{\infty} \tilde{A}_k) = \sum_{k=1}^{\infty} \mu(\tilde{A}_k) \\
&= \lim_{N\to\infty} \sum_{k=1}^{N} \mu(\tilde{A}_k) \\
&= \lim_{N\to\infty} \mu(\bigcup_{k=1}^{N} \tilde{A}_k) \\
&= \lim_{N\to\infty} \mu(A_N).
\end{aligned}$$

(3) Set $\tilde{A}_k := A_1 \backslash A_k$, $k \in \mathbb{N}$. Then $\emptyset = \tilde{A}_1 \subset \tilde{A}_2 \subset \ldots$. Moreover we have

$$\mu(A_1) = \mu(\tilde{A}_k) + \mu(A_k), \quad k \in \mathbb{N}.$$

By the above argument (observe that $\mu(A_k) \leq \mu(A_1) < \infty$)

$$\begin{aligned}\mu(A_1) - \lim_{k\to\infty} \mu(A_k) = \lim_{k\to\infty} \mu(\tilde{A}_k) \\ =&\mu(\bigcup_{k=1}^{\infty} \tilde{A}_k) \\ =&\mu(A_1 \backslash \bigcap_{k=1}^{\infty} A_k) \\ =&\mu(A_1) - \mu(\bigcap_{k=1}^{\infty} A_k).\end{aligned}$$

Since $\mu(A_1) < \infty$ we can conclude.

□

Example 1.34.

- There is no way we can assume (or should hope for) that for uncountable unions we have (even sub-)additivity: For example $\mathbb{R} = \bigcup_{x\in\mathbb{R}}\{x\}$. The Lebesgue measure would satisfy $\mathcal{L}^1\{x\} = 0$ for all x, so
$$\mathcal{L}^1(\mathbb{R}) \neq \lim_{k\to\infty} \mathcal{L}^1\{x\} = 0.$$
- The assumption $\mu(A_1) < \infty$ in Theorem 1.33(3) is necessary.
  Let $X = \mathbb{N}$ and μ the counting measure. Set for $k \in \mathbb{N}$
$$A_k = \{k, k+1, k+2, \ldots\}.$$
  Then $A_k \supset A_{k+1}$, but $\mu(A_k) = \infty$ for all $k \in \mathbb{N}$. However $\bigcap_{k\in\mathbb{N}} A_k = \emptyset$, so
$$0 = \mu(\bigcap_{k\in\mathbb{N}} A_k) \neq \lim_{k\to\infty} \mu(A_k) = \infty.$$

Lastly, from Advanced Calculus we are aware of sets of measure zero (there it was that Riemann-integrable functions are continuous outside a set of *Lebesgue*-measure zero)

Definition 1.35 (Zero sets)**.** A set $N \subset X$ is called a set of *μ-measure zero* if $\mu(N) = 0$. We also say N is a *μ-zeroset*.

A property $P(x)$ holds *μ-a.e.* in X if $P(x)$ holds in $X\backslash N$ where N is a μ-zeroset.

Theorem 1.36. *Let $N \subset X$ be a μ-zero set. Then N is measurable.*

Proof. Let $B \subset X$, by monotonicity we have $\mu(B \cap N) \leq \mu(N) = 0$. Moreover $\mu(B\backslash N) \leq \mu(B)$. So we have

$$\mu(B \cap N) + \mu(B\backslash N) \leq \mu(N) + \mu(B) = \mu(B).$$

This implies measurability. □

Exercise 1.37. *Let X be a set and μ be a measure on X.*

(1) Let $(N_i)_{i\in\mathbb{N}}$ be sets of μ-measure zero. Show that $\bigcup_{i\in\mathbb{N}} N_i$ is a set of μ-measure zero.

(2) Let N be a μ-zeroset. Show that any $A \subset N$ is a μ-zeroset.
(3) Show that is a property $P(x)$ holds for μ-a.e. x and a property $Q(y)$ holds for μ-a.e. y then $Q(x)$ and $P(x)$ hold (simultaneously) for a.e. x.

1.3. **Construction of Measures: Carathéodory-Hahn Extension Theorem.** While we already have constructed in (1.2) the Lebesgue (outer) measure, we are still interested in finding a more axiomatic approach to construct the Lebesgue measure.

The idea is to define a *pre-measure* on some sets (like cubes!) and build a measure out of that.

Definition 1.38 (Algebra)**.** Let X be a set and $\mathcal{A} \subset 2^X$. $\mathcal{A}$ is called an *algebra* if

(1) $X \in \mathcal{A}$
(2) $A \in \mathcal{A}$ implies that $X \backslash A \in \mathcal{A}$
(3) If $A_1, A_2 \in \mathcal{A}$ then $A_1 \cup A_2 \in \mathcal{A}$.

Definition 1.39 (Pre-measure)**.** Let X be a set, $\mathcal{A} \subset 2^X$ an algebra (not necessarily a σ-algebra!). A map $\lambda : \mathcal{A} \to [0, \infty]$ is a *pre-measure*, if

(1) $\lambda(\emptyset) = 0$
(2) $\lambda(A) = \sum_{k=1}^{\infty} \lambda(A_k)$ for any $A \in \mathcal{A}$ such that $A = \bigcup_{k=1}^{\infty} A_k$ for some parwise disjoint $(A_k)_{k=1}^{\infty} \subset \mathcal{A}$.

A pre-measure is called *σ-finite*, if $X = \bigcup_{k=1}^{\infty} S_k$ with $S_k \in \mathcal{A}$ and $\lambda(S_k) < \infty$ for each k.

Exercise 1.40. *Let $\lambda : \mathcal{A} \to [0, \infty]$ be a premeasure*

(1) Assume $A \subset B$ with $A, B \in \mathcal{A}$ then $\lambda(A) \leq \lambda(B)$.

Example 1.41. An interval is the set (a, b) or $[a, b]$ or $(a, b]$ or $[a, b)$, where $-\infty < a \leq b < \infty$ (we allow $\pm\infty$ if the set is open).

A *block* Q in $\mathbb{R}^n$ is the cartesian product of n intervals $Q = I_1 \times I_2 \times \ldots \times I_n$.

A *figure* is made out of finitely many blocks

$$\mathcal{A} := \left\{ A \subset \mathbb{R}^n : A = \bigcup_{i=1}^{N} Q_i \quad \text{some } N \in \mathbb{N},\ Q_i \text{ blocks with pairwise disjoint interior} \right\}.$$

<u>Exercise: $\mathcal{A}$ is an Algebra</u>

The volume of a block $Q = I_1 \times \ldots \times I_n$ is given by the n-dimensional volume, i.e.

$$\mathrm{vol}(Q) = \Pi_{i=1}^{n} |I_i|,$$

where as usual, $|[a, b]| = |(a, b)| = |[a, b)| = |(a, b]| := |b - a|$.

Indeed, vol defines now a premeasure on $\mathcal{A}$:

Whenever $A \in \mathcal{A}$, i.e. $A = \bigcup_{i=1}^N Q_i$ for pairwise disjoint blocks Q_i then

$$\mathrm{vol}(A) := \sum_{i=1}^N \mathrm{vol}(Q_i).$$

One can check that this is independent of the specific choice of Q_i. I.e. if also $A = \bigcup_{j=1}^{\tilde{N}} \tilde{Q}_j$ for another cobination of pairwise disjoint blocks $\tilde{Q}_j$ then

$$\sum_{i=1}^N \mathrm{vol}(Q_i) = \sum_{j=1}^{\tilde{N}} \mathrm{vol}(\tilde{Q}_j).$$

Indeed, now vol defines a *pre-measure* on $\mathcal{A}$. vol is also σ-finite, simply take $S_k := [-k, k]^n$.

The important thing is that a premeasure extends (more or less uniquely) into a real measure. This is called *Carathéodory–Hahn extension*.

Theorem 1.42 (Carathéodory–Hahn extension)**.** *Let X be a set, and $\mathcal{A} \subset 2^X$ an algebra with a premeasure $\lambda : \mathcal{A} \to [0, \infty]$.*

For $A \subset X$ the Carathéodory–Hahn extension *$\mu : 2^X \to [0, \infty]$ is defined as*

$$\mu(A) := \inf \left\{ \sum_{k=1}^\infty \lambda(A_k) : \quad A \subset \bigcup_{k=1}^\infty A_k, \quad A_k \in \mathcal{A} \right\}$$

Then

(1) $\mu : 2^X \to [0, \infty]$ is a measure on X,
(2) $\mu(A) = \lambda(A)$ for all $A \in \mathcal{A}$,
(3) Any $A \in \mathcal{A}$ is μ-measurable.

Compare this to the definition of the Lebesgue measure, (1.2).

Proof of Theorem 1.42. (1) Clearly $\mu : 2^X \to [0, \infty]$ is well-defined (observe that $X \in \mathcal{A}$, so that $\mu(B) \le \mu(X)$ for all $B \subset X$, and $\mu(\emptyset) = 0$.

Now let $B \subset \bigcup_{k=1}^\infty B_k$. Take any $\varepsilon > 0$. By definition of the infimum, for each B_k there exist some $(A_{k;\ell})_{\ell=1}^\infty \subset \mathcal{A}$ such that $B_k \subset \bigcup_{\ell=1}^\infty A_{k;\ell}$ and

$$\sum_{\ell=1}^\infty \lambda(A_{k;\ell}) \le \mu(B_k) + 2^{-k}\varepsilon.$$

Clearly, $A \subset \bigcup_{k,\ell=1}^{\infty} A_{k;\ell}$ so we have

$$\begin{aligned}\mu(A) &\leq \sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty} \lambda(A_k;\ell)\\ &\leq \sum_{k=1}^{\infty}\Big(\mu(B_k) + 2^{-k}\varepsilon\Big)\\ &\leq \left(\sum_{k=1}^{\infty}\mu(B_k)\right) + \varepsilon.\end{aligned}$$

This holds for any $\varepsilon > 0$, so letting $\varepsilon \to 0$ we find

$$\mu(A) \leq \left(\sum_{k=1}^{\infty}\mu(B_k)\right) + \varepsilon.$$

Thus, μ is σ-subaddive, and thus μ is a measure.

(2) For $A \in \mathcal{A}$ we clearly have $\mu(A) \leq \lambda(A)$.

Now we show $\lambda(A) \leq \mu(A)$. We may assume that $\mu(A) < \infty$ otherwise there is nothing to show.

Take any $(A_k)_{k=1}^{\infty} \subset \mathcal{A}$ such that $\bigcup_{k=1}^{\infty} A_k \supset A$. By the usual argument we may assume (without loosing that $A_k \in \mathcal{A}$) that $A_k \cap A_j = \emptyset$ for all $k \neq j$. Set $\tilde{A}_k := A_k \cap A$, $k \in \mathbb{N}$. Then $(\tilde{A}_k)_{k=1}^{\infty} \in \mathcal{A}$ are pairwise disjoint, so since λ is premeasure

$$\lambda(A) = \sum_{k=1}^{\infty}\lambda(\tilde{A}_k) \leq \sum_{k=1}^{\infty}\lambda(A_k).$$

Here we also used Exercise 1.40 (monotonicity of λ) and $\tilde{A}_k \subset A_k$.

Taking the infimum over all covers $(A_k)_{k\in\mathbb{N}}$ as above we obtain that

$$\lambda(A) \leq \mu(A),$$

as claimed.

(3) Let $A \in \mathcal{A}$ and $B \subset X$.

Fix $\varepsilon > 0$ arbitrary. By the definition of μ, there exist $(B_k)_{k=1}^{\infty} \subset \mathcal{A}$ such that $B \subset \bigcup_{k=1}^{\infty} B_k$ and

$$\sum_{k=1}^{\infty}\mu(B_k) \geq \mu(B) \geq \sum_{k=1}^{\infty}\mu(B_k) - \varepsilon.$$

Since $B_k \in \mathcal{A}$ we have $\mu(B_k) = \lambda(B_k)$. Since $A, B_k \in \mathcal{A}$ and $\mathcal{A}$ is an algebra we have $B \cap A_k$ and $B_k \backslash A \in \mathcal{A}$. These are disjoint sets, and since λ is a pre-measure we have

$$\mu(B_k) = \lambda(B_k) = \lambda(B_k \cap A) + \lambda(B_k \backslash A).$$

So we have

$$\mu(B) \geq \sum_{k=1}^{\infty}\lambda(B_k \cap A) + \sum_{k=1}^{\infty}\lambda(B_k \backslash A) - \varepsilon.$$

Now observe that $B \cap A \subset \bigcup_{k\in\mathbb{N}} B_k \cap A$ and $B \backslash A \subset \bigcup_{k=1}^{\infty} B_k \backslash A$ (and both coverings belong to $\mathcal{A}$). By the definition of μ we thus find

$$\sum_{k=1}^{\infty} \lambda(B_k \cap A) \geq \mu(B \cap A)$$

and

$$\sum_{k=1}^{\infty} \lambda(B_k \setminus A) \geq \mu(B \backslash A)$$

Together we arrive at

$$\mu(B) \geq \mu(B \cap A) + \mu(B \backslash A).$$

That is, A is μ-measurable.

□

Definition 1.43 (Lebesgue measure)**.** The Lebesgue measure $\mathcal{L}^n$ in $\mathbb{R}^n$ is the Caratheodory-Hahn extension of λ in Example 1.41. Compare this with (1.2).

If the pre-measure is *additionally σ-finite*, then the Carathéodory-Hahn extension μ is essentially unique (on the sets we care about: the measurable sets).

Theorem 1.44 (Uniqueness)**.** *Let $\lambda : \mathcal{A} \to [0, \infty]$ as in Theorem 1.42 be additionally σ-finite and denote by μ the Carathéodory-Hahn-extension.*

Whenever $\tilde{\mu} : 2^X \to [0, \infty]$ is another measure such that

$$\tilde{\mu}(A) = \lambda(A) \quad \textit{for all } A \in \mathcal{A},$$

then indeed

$$\tilde{\mu}(A) = \mu(A) \quad \textit{for all } \mu\textit{-measurable } A$$

Proof. (1) We have $\tilde{\mu}(A) \leq \mu(A)$ for all $A \subset X$. Indeed, let $A \subset \bigcup_{k\in\mathbb{N}} A_k$ for $A_k \in \mathcal{A}$. Then by σ-subadditivity of $\tilde{\mu}$,

$$\tilde{\mu}(A) \leq \sum_{k=1}^{\infty} \tilde{\mu}(A_k) = \sum_{k=1}^{\infty} \lambda(A_k).$$

Taking the infimum over all such covers $(A_k)_{k=1}^{\infty} \subset \mathcal{A}$ of A we obtain

$$\tilde{\mu}(A) \leq \mu(A) \quad \forall A \subset X \tag{1.8}$$

(2) Let Σ be the σ-algebra of μ-measurable sets.
We now show:
$\tilde{\mu}(A) = \mu(A)$ for all $A \subset \Sigma$ such that there is $S \in \mathcal{A}$ with $\lambda(S) < \infty$, and $S \supset A$.
So fix such A and S.
Then

$$\tilde{\mu}(S \backslash A) \leq \tilde{\mu}(S) \overset{S\in\mathcal{A}}{=} \lambda(S) < \infty.$$

Consequently, since $A \in \Sigma$ and $S \in \mathcal{A}$,

$$\tilde{\mu}(A) + \tilde{\mu}(S\backslash A) \overset{(1.8)}{\leq} \mu(A) + \mu(S\backslash A) \overset{A\in\Sigma}{=} \mu(S) \overset{A\in\mathcal{A}}{=} \tilde{\mu}(S) \leq \tilde{\mu}(A) + \tilde{\mu}(S\backslash A).$$

So we have equality everywhere, which leads to

$$\tilde{\mu}(A) + \tilde{\mu}(S\backslash A) = \mu(A) + \mu(S\backslash A)$$

With (1) we conclude

$$\tilde{\mu}(A) + \tilde{\mu}(S\backslash A) \geq \mu(A) + \tilde{\mu}(S\backslash A)$$

and thus $\tilde{\mu}(A) \geq \mu(A)$ (here we use that $\tilde{\mu}(S\backslash A) < \infty$). Again with (1) we have shown that

$$\mu(A) = \tilde{\mu}(A) \quad \forall A \in \Sigma : \quad \text{s.t.} \, \exists \, S \in \mathcal{A} : \; A \subset S, \; \lambda(S) < \infty. \tag{1.9}$$

(3) $\underline{\tilde{\mu}(A) = \mu(A) \text{ for } \textit{all } A \subset \Sigma}$

In comparison to (2) we need to remove the restriction $A \subset S$ for some $\lambda(S) < \infty$.

We write $X = \bigcup_{k=1}^{\infty} S_k$ with $(S_k)_{k\in\mathbb{N}}$ pairwise disjoint, $S_k \in \mathcal{A}$, $\lambda(S_k) < \infty$ for all $k \in \mathbb{N}$.

Set $A_k := A \cap S_k$, which are pairwise disjoint sets that all belong to Σ. From (1.9) we have for any $m \in \mathbb{N}$

$$\tilde{\mu}(\bigcup_{k=1}^{m} A_k) = \mu(\bigcup_{k=1}^{m} A_k).$$

Thus, by monotonicity of $\tilde{\mu}$, and since each $A_k \in \Sigma$, we have

$$\tilde{\mu}(A) \geq \limsup_{m\to\infty} \tilde{\mu}(\bigcup_{k=1}^{m} A_k) = \limsup_{m\to\infty} \mu(\bigcup_{k=1}^{m} A_k) \overset{A_k\in\Sigma}{=} \sum_{k=1}^{\infty} \mu(A_k).$$

From Theorem 1.33 (1) we obtain

$$\sum_{k=1}^{\infty} \mu(A_k) = \mu(\bigcup_{k=1}^{\infty} A_k).$$

Thus we have shown for any $A \in \Sigma$,

$$\tilde{\mu}(A) \geq \mu(A),$$

and we have equality in view of (1.8).

□

Remark 1.45. Observe that in Theorem 1.44 it is *not* said that any μ-measurable A was also $\tilde{\mu}$-measurable

Example 1.46. In general μ and $\tilde{\mu}$ in Theorem 1.44 might be different for non-measurable sets.

Let $X = [0, 1]$, $\mathcal{A} = \{\emptyset, X\}$, and set

$$\lambda(\emptyset) := 0, \quad \lambda([0, 1]) := 1.$$

Then $\lambda : \mathcal{A} \to \mathbb{R}$ is a pre-measure.

One can check that the Caratheodory-Hahn extension of λ is given by

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

If on the other hand we consider $\tilde{\mu}$ the Lebesgue measure on $[0,1]$, i.e. the Caratheodory-Hahn extension of $\tilde{\lambda} : \mathcal{A} \to [0,\infty]$ where $\tilde{A}$ are the figures from Example 1.41 and $\tilde{\lambda}$ is the volume as defined. Then $\tilde{\mu}$ coincides with μ on $\mathcal{A}$ but not in $\tilde{A}$, because, e.g.

$$\frac{1}{2} = \lambda([0,1/2]) = \tilde{\mu}([0,1/2]) \neq \mu([0,1/2]) = 1.$$

1.4. **Classes of Measures.**

Definition 1.47. Let X be a metric space and μ be a measure on X. μ is called a *metric measure* if

$$\mu(E \cup F) = \mu(E) + \mu(F)$$

whenever $\operatorname{dist}(E,F) > 0$

Exercise 1.48. *The Hausdorff measure $\mathcal{H}^s$ on a metric space X is metric measures for any $s \geq 0$.*

Exercise 1.49. *The Lebesgue measure on $\mathbb{R}^n$ is a metric measure.*

Definition 1.50. Let X be a metric space and μ be a measure on X.

- Let $\mathcal{B} \subset 2^X$ be the smallest σ-algebra that contains all open sets of $\mathbb{R}^n$, that is[4]

  $$\mathcal{B} := \bigcap \left\{ \mathcal{A} \subset 2^X : \quad \mathcal{A} \text{ is } \sigma\text{-algebra, all open sets belong to } \mathcal{A} \right\}.$$

  (Cf. Exercise 1.31). Any set $A \in \mathcal{B}$ is called a *Borel set* and $\mathcal{B}$ is called the *Borel σ-algebra*.
- A measure μ on X for which (at least) all Borel-sets are μ-measurable is called a *Borel measure*.

Exercise 1.51. *If $f : X \to Y$ is homeomorphism then $f(A)$ is Borel if and only if A is Borel*

Theorem 1.52. *If μ is a metric measure on a metric space X then μ is a Borel measure, i.e. all open sets are μ-measurable. In particular Lebesgue and Hausdorff measure are Borel measures.*

[4] it is an easy exercise to show that this indeed defines a σ-algebra. Observe in particular that 2^X is a σ-algebra which contains all open sets so the right-hand side is not an intersection of empty sets.

Proof. It suffice to show that all open sets in X are μ-measurable, since then the set of measurable sets (which is a σ-algebra) must contain the Borel sets $\mathcal{B}$.

Let $G \subset X$ be open and $A \subset X$ arbitrary. We need to show

$$\mu(A) \geq \mu(A \cap G) + \mu(A \backslash G).$$

If $\mu(A) = \infty$ this is obvious, so from now on assume $\mu(A) < \infty$.

For $k \in \mathbb{N}$ define

$$G_k := \left\{ x \in G : \quad \operatorname{dist}(x, X \backslash G) > \frac{1}{k} \right\}$$

Then $\operatorname{dist}(G_k, X \backslash G) \geq \frac{1}{k} > 0$.

We have by monotonicity,

$$\mu(A \cap G) + \mu(A \backslash G) \leq \mu(A \cap G_k) + \mu(A \backslash G) + \mu(A \cap (G \backslash G_k)).$$

Since μ is a metric measure and $\operatorname{dist}(A \cap G_k, A \backslash G) > 0$ we conclude

$$\mu(A \cap G) + \mu(A \backslash G) \leq \mu(A) + \mu(A \cap (G \backslash G_k)).$$

So the statement is proven once we show

$$\lim_{k \to \infty} \mu(A \cap (G \backslash G_k)) = 0. \tag{1.10}$$

To see (1.10) let

$$D_k := G_{k+1} \backslash G_k = \left\{ x \in G : \quad \operatorname{dist}(x, X \backslash G) \in (\frac{1}{k+1}, \frac{1}{k}] \right\}.$$

We then have (here we use that G is open)

$$G \backslash G_k = \bigcup_{i=k}^{\infty} D_i \quad \text{so } \mu(A \cap (G \setminus G_k)) \leq \sum_{i=k}^{\infty} \mu(A \cap D_i). \tag{1.11}$$

and whenever $i + 2 \leq j$ we have

$$\operatorname{dist}(D_i, D_j) \geq \frac{1}{i+1} - \frac{1}{j} > 0.$$

Since μ is a metric measure we can sum up even and odd D_i's i.e.

$$\sum_{i=1}^{k} \mu(A \cap D_{2i-1}) = \mu(A \cap \bigcup_{i=1}^{k} D_{2i-1}) \leq \mu(A)$$

and

$$\sum_{i=1}^{k} \mu(A \cap D_{2i}) = \mu(A \cap \bigcup_{i=1}^{k} D_{2i}) \leq \mu(A).$$

In particular we have

$$\sum_{i=1}^{\infty} \mu(A \cap D_i) \leq 2\mu(A) < \infty.$$

From (1.11),

$$\mu(A \cap (G \backslash G_k)) \leq \sum_{i=k}^{\infty} \mu(A \cap D_i)$$

and since the series on the right-hand side converges

$$\mu(A \cap (G \backslash G_k)) \xrightarrow{k \to \infty} 0.$$

That is (1.10) is established and we can conclude. □

Theorem 1.53. *Let X be a metric space and μ any Borel measure. Suppose that X is a union of countably many open sets of finite measure. Then for all Borel sets $E \subset X$*

$$\mu(E) = \inf_{U \supset E; U \text{ open}} \mu(U) = \sup_{C \subset E; C \text{ closed}} \mu(C).$$

The first part of Theorem 1.53 is a direct consequence of

Proposition 1.54. *Let X be a metric space and μ any Borel measure. Suppose that X is a union of countably many open sets of finite measure. If we define $\tilde{\mu} : 2^X \to [0, \infty]$ as*

$$\tilde{\mu}(E) := \inf_{U \supset E; U \text{ open}} \mu(U) \quad E \subset X \tag{1.12}$$

then $\tilde{\mu}$ is a metric measure and we have

$$\tilde{\mu}(E) = \mu(E) \quad \forall \text{Borel sets } E \subset X.$$

The Hausdorff measure $\mathcal{H}^s$ (which is metric and thus Borel) shows that we cannot skip the assumption that X finite union of countably many open sets of finite measure.

Indeed if $s < n$ for any nonempty open set $\mathcal{H}^s(U) = \infty$, so (1.12) is certainly false.

Proof of Proposition 1.54. It is easy to see that $\tilde{\mu}$ is a measure (exercise).

We will show, it is a *metric* measure: let $E, F \subset X$ such that $\delta := \operatorname{dist}(E, F) > 0$. Set

$$V_E := \bigcup_{x \in E} B(x, \frac{1}{3}\delta),$$

and

$$V_F := \bigcup_{x \in F} B(x, \frac{1}{3}\delta).$$

V_E and V_F are then disjoint and open. Fix $\varepsilon > 0$, let $U \supset E \cup F$ be open such that

$$\mu(U) \leq \tilde{\mu}(E \cup F) + \varepsilon.$$

Since $V_E \cap U$ and $V_F \cap U$ are open they are μ-measurable and since they are moreover disjoint

$$\mu((V_E \cap U) \cup (V_F \cap U)) = \mu(V_E \cap U) + \mu(V_F \cap U).$$

Since $E = E \cap V_E \subset U \cap V_E$ and $F = F \cap V_F \subset U \cap V_F$ we have

$$\tilde{\mu}(E) + \tilde{\mu}(F) \leq \mu(V_E \cap U) + \mu(V_F \cap U) = \mu((V_E \cap U) \cup (V_F \cap U)) \leq \mu(U) \leq \tilde{\mu}(E \cup F) + \varepsilon.$$

Taking $\varepsilon \to 0$ we obtain that $\tilde{\mu}(E) + \tilde{\mu}(F) \leq \tilde{\mu}(E \cup F)$ which shows that $\tilde{\mu}$ is metric.

Consequently, in view of Theorem 1.52, $\tilde{\mu}$ is Borel.

It remains to show $\tilde{\mu}(E) = \mu(E)$ for all Borel sets. Clearly,

$$\tilde{\mu}(E) \geq \mu(E) \quad \forall E \subset X.$$

and

$$\tilde{\mu}(E) = \mu(E) \quad \forall E \subset X \text{ open}.$$

By assumption we can write $X = \bigcup_{k=1}^{\infty} X_k$ with X_k open and $\mu(X_k) < \infty$. We also may assume that $X_k \subset X_{k+1}$.

Then we have for all set $E \subset X$

$$\mu(X_n \backslash E) \leq \tilde{\mu}(X_n \backslash E)$$

and

$$\mu(X_n \cap E) \leq \tilde{\mu}(X_n \backslash E).$$

Now if E is Borel, we have

$$\mu(X_n \backslash E) = \tilde{\mu}(X_n \backslash E) \quad \text{and} \quad \mu(X_n \cap E) = \tilde{\mu}(X_n \backslash E). \tag{1.13}$$

Indeed, if not, either $\mu(X_n \backslash E) < \tilde{\mu}(X_n \backslash E)$ or $\mu(X_n \cap E) < \tilde{\mu}(X_n \backslash E)$, which leads to

$$\mu(X_n) = \mu(E \cap X_n) + \mu(X_n \backslash E_n) < \tilde{\mu}(E \cap X_n) + \tilde{\mu}(X_n \backslash E_n) = \tilde{\mu}(X_n) = \mu(X_n)$$

the second to last equation is the $\tilde{\mu}$-measurability of E (since it is a Borel set and $\tilde{\mu}$ is a Borel measure), the last equation uses that X_n is open. The above estimate is impossible (since $\mu(X_n), \tilde{\mu}(X_n) < \infty$), so (1.13) is established.

So in particular we have for any Borel set E, $\mu(X_n \cap E) = \tilde{\mu}(X_n \cap E)$. In view of Theorem 1.33 (recall: both μ and $\tilde{\mu}$ are Borel!)

$$\tilde{\mu}(E) = \lim_{n \to \infty} \tilde{\mu}(X_n \cap E) = \lim_{n \to \infty} \mu(X_n \cap E) = \mu(E).$$

That is, we have shown for any Borel set E,

$$\mu(E) = \tilde{\mu}(E) := \inf_{U \supset E; U \text{ open}} \mu(U). \tag{1.14}$$

□

Proof of Theorem 1.53 last part. Having from Proposition 1.54 (1.14), it remains to prove that for Borel sets E

$$\mu(E) = \sup_{C \subset E; C \text{ closed}} \mu(C).$$

We apply (1.14) to $X_n \backslash E$ and find an open set U_n such that

$$X_n \backslash E \subset U_n$$

and

$$\mu(U_n \backslash (X_n \backslash E)) = \mu(U_n) - \mu(X_n \backslash E) < \frac{\varepsilon}{2^n}.$$

The set $U := \bigcup_{n=1}^{\infty} U_n$ is open and $C = X \backslash U \subset E$ is closed. Now it suffices to observe that

$$E \backslash C = E \cap \bigcup_{n=1}^{\infty} U_n \subset \bigcup_{n=1}^{\infty} G_n \backslash (U_n \backslash E)$$

and hence $\mu(E \backslash C) = \mu(E) - \mu(C) < \varepsilon$. Thus, $\mu(C) \leq \mu(E) \leq \mu(C) + \varepsilon$. Taking $\varepsilon \to 0$ the proof is complete. □

Corollary 1.55. *Let X be a metric space and μ, ν Borel measure. Suppose that X is a union of countably many open sets of finite measure. If ν and μ coincide for open sets, namely if*

$$\nu(U) = \mu(U) \quad \forall \textit{open set } U \subset X,$$

then

$$\nu(E) = \mu(E) \quad \forall \textit{Borel sets } E.$$

Proof. Twice applying Theorem 1.52, we find that for any Borel set E

$$\mu(E) = \inf_{U \supset E; U \text{ open}} \mu(U) = \inf_{U \supset E; U \text{ open}} \nu(U) = \nu(E).$$

□

On $\mathbb{R}^n$ we can simplify Corollary 1.55

Theorem 1.56 (Borel measures on $\mathbb{R}^n$ that coincide on rectangles)**.** *Let μ and ν be two finite Borel measures on $\mathbb{R}^n$ such that*

$$\mu(R) = \nu(R)$$

for all closed rectangles R of the form

$$R := \{x \in \mathbb{R}^n : \quad a_i \leq x_i \leq b_i, \quad i = 1, \ldots, n\}$$

where $-\infty \leq a_i \leq b_i \leq \infty$, $i = 1, \ldots, n$.

Then

$$\mu(B) = \nu(B)$$

for all Borel sets $B \subset \mathbb{R}^n$

For the proof of Theorem 1.56 we need a essentially combinatorial observation, called the *π-λ Theorem*. π and λ-systems are families of sets which are invariant under less operations than the σ-Algebras.

Definition 1.57. (1) A nonempty family $\mathcal{P} \subset 2^X$ is called a *π-system* if it is closed under (finitely many) intersections, i.e.

$$A, B \in \mathcal{P} \quad \text{implies} \quad A \cap B \in \mathcal{P}.$$

(2) A family of subsets $\mathcal{L} \subset 2^X$ is called a *λ-system* if
- $X \in \mathcal{L}$
- $A, B \in \mathcal{L}$ *and $B \subset A$* implies $A \setminus B \in \mathcal{L}$

- if $A_k \in \mathcal{L}$ and $A_k \subset A_{k+1}$ *for* $k = 1, \dots,$ then
$$\bigcup_{k=1}^{\infty} A_k \in \mathcal{L}.$$

Exercise 1.58. *Show that if $\mathcal{P}$ is a λ-system and a π-system, then it is a σ-Algebra.*

Clearly any σ-Algebra is also a π-system and a λ-system.

Theorem 1.59 (π-λ Theorem)**.** *If $\mathcal{P}$ is a π-system and $\mathcal{L}$ is a λ-system with*
$$\mathcal{P} \subset \mathcal{L}$$
then[5]
$$\sigma(\mathcal{P}) \subset \mathcal{L}$$

Proof. Define $\mathcal{S}$ to tbe the intersection of all λ-systems $\mathcal{L}'$ containing $\mathcal{P}$
$$\mathcal{S} := \bigcap_{\mathcal{L}' \supset \mathcal{P}} \mathcal{L}',$$
We first claim that <u>$\mathcal{S}$ is a π-system</u>.

Indeed let $A, B \in \mathcal{S}$. We must show $A \cap B \in \mathcal{S}$. Define
$$\mathcal{A} := \{C \subset X : \quad A \cap C \in \mathcal{S}\}.$$
Since $\mathcal{S}$ is a λ-system, it follows that $\mathcal{A}$ is a λ-system. Therefore $\mathcal{S} \subset \mathcal{A}$. But then since $B \in \mathcal{S}$ we see that $A \cap B \in \mathcal{S}$.

Next we claim that <u>$\mathcal{S}$ is a σ-algebra</u>

Indeed, we only need to show that $\mathcal{S}$ is a λ-system (cf. Exercise 1.58).

Since $X \in \mathcal{S}$, $\emptyset = X \setminus X \in \mathcal{S}$. Also $A \in \mathcal{S}$ implies $X \setminus A \in \mathcal{S}$. So $\mathcal{S}$ is closed under complements and under finite intersections, and thus under finite unions.

If $A_1, A_2 \in \mathcal{S}$ then $B_n := \bigcup_{k=1}^{n} A_k \in \mathcal{S}$. Since $\mathcal{S}$ is a λ-system we conclude that $\bigcup_{k=1}^{\infty} A_k = \bigcup_{n=1}^{\infty} B_n \in \mathcal{S}$. Thus $\mathcal{S}$ is a σ-algebra.

Since $\mathcal{S} \supset \mathcal{P}$ is a σ-algebra it follows that
$$\sigma(\mathcal{P}) \subset \mathcal{S} \subset \mathcal{L}.$$
□

Proof of Theorem 1.56. Let
$$\mathcal{P} := \{A \subset \mathbb{R}^n : \quad \text{for some } n \in \mathbb{N},\ A = \bigcap_{k=1}^{n} R_k \text{ where } (R_k)_{k=1}^{n} \text{ are a rectangles}\},$$

[5]Recall that $\sigma(\mathcal{P})$ is the σ-Algebra generated by $\mathcal{P}$

which a π-system. We also set

$$\mathcal{L} := \{B \subset \mathbb{R}^n : \quad B \text{ Borel:} \quad \mu(B) = \nu(B)\}.$$

While it is not so clear that $\mathcal{L}$ is a σ-Algebra, one can check it is a λ-system. Also $\mathcal{P} \subset \mathcal{L}$ by assumption (observe that each R_k is μ and σ-measurable).

By Theorem 1.59 we have $\sigma(\mathcal{P}) \subset \mathcal{L}$. Since $\sigma(\mathcal{P})$ contains the Borel sets (any open set can be written as countable union of rectangles, see Lemma 1.75 below), any Borel set B satisfies $B \in \mathcal{L}$ and thus $\mu(B) = \nu(B)$, and we can conclude. □

Definition 1.60 (Borel Regular measure)**.** A Borel measure μ is *Borel regular*, if for any $A \subset \mathbb{R}^n$ there exists some Borel set $B \supset A$ such that $\mu(A) = \mu(B)$.

Corollary 1.61. *Let X be a metric space and μ any Borel measure. Suppose that X is a union of countably many open sets of finite measure. If we define $\tilde{\mu} : 2^X \to [0, \infty]$ as*

$$\tilde{\mu}(E) := \inf_{U \supset E; U \text{ open}} \mu(U) \quad E \subset X$$

then $\tilde{\mu}$ is a metric, Borel-regular measure that coincides with μ on Borel sets.

Proof. In view of Proposition 1.54 we only need to show that $\tilde{\mu}$ is Borel-regular.

Indeed, let $E \subset X$ with $\mu(E) < \infty$ (otherwise $\mu(E) = \mu(X) = \infty$) be arbitary. Let $(U_k)_{k=1}^{\infty}$ be open sets so that $U_k \supset E$ and

$$\mu(U_k) - \frac{1}{k} \leq \tilde{\mu}(E).$$

Set $U := \bigcap_{k=1}^{\infty} U_k$. Then we have $E \subset U$ and thus $\tilde{\mu}(E) \leq \tilde{\mu}(U)$. On the other hand we have

$$\tilde{\mu}(U) = \tilde{\mu}(\bigcap_{k=1}^{\infty} U_k) \leq \mu(U_k) \leq \tilde{\mu}(E) + \frac{1}{k}.$$

This holds for any $k \in \mathbb{N}$ so letting $k \to \infty$ we find

$$\tilde{\mu}(U) \leq \tilde{\mu}(E).$$

So $\tilde{\mu}(U) = \tilde{\mu}(E)$, and since as an intersection of countably many open sets U is a Borel-set, we can conclude. □

Example 1.62. The Lebesgue measure $\mathcal{L}^n$ (as defined in Definition 1.43) is Borel regular.

Proof. In view of Corollary 1.61 it suffices to show that

$$\mathcal{L}^n(A) = \inf \{\mathcal{L}^n(G) : \quad G \subset \mathbb{R}^n \text{ open and } G \supset A\}.$$

$\leq$ is obvious by monotonicity. For $\geq$ assume $\mathcal{L}^n(A) < \infty$, let $\varepsilon > 0$ and take blocks $(Q_i)_{i=1}^{\infty}$ such that

$$\sum_{i=1}^{\infty} \text{vol}(Q_i) \leq \mathcal{L}^n(A) + \varepsilon.$$

To each block Q_i we can choose an open block $Q_i^o \supset Q_i$ such that

$$\mathrm{vol}(Q_i^o) \leq \mathrm{vol}(Q_i) + 2^{-i}\varepsilon.$$

Set $G := \bigcup_{i=1}^{\infty}(Q_i)^o \supset A$. Then

$$\begin{aligned}\mathcal{L}^n(G) \leq & \sum_{i=1}^{\infty} \mathrm{vol}(Q_i^o) \leq \sum_{i=1}^{\infty} \mathrm{vol}(Q_i) + \sum_{i=1}^{\infty} 2^{-i}\varepsilon \\ \leq & \mathcal{L}^n(A) + 2\varepsilon.\end{aligned}$$

That is we have shown

$$\mathcal{L}^n(A) + 2\varepsilon \geq \inf\left\{\mathcal{L}^n(G) : \quad G \subset \mathbb{R}^n \text{ open and } G \supset A\right\}.$$

which holds for any $\varepsilon > 0$, letting $\varepsilon \to 0$ we conclude. □

Example 1.63. Let X be a metric space. For any $s \geq 0$, $\mathcal{H}^s(X)$ is Borel-regular.

Proof. In this case we cannot use Corollary 1.61, since we cannot assume that we can cover X with countably many finite measure sets.

If $s = 0$ the claim is easy. Take any $E \subset X$. If $\mathcal{H}^0(E) = \infty$, then $\mathcal{H}^0(E) = \mathcal{H}^0(X) = \infty$. If $\mathcal{H}^0(E) < \infty$ then E contains finitely many points, thus E is closed and thus a Borel set.

Now let $s > 0$ and $E \subset X$. If $\mathcal{H}^s(E) = \infty$ then again $\mathcal{H}^s(E) = \mathcal{H}^s(X) = \infty$ and we conclude. So assume $\mathcal{H}^s(E) < \infty$. In particular $\mathcal{H}^s_\delta(E) < \infty$ for all $\delta > 0$.

So for each ℓ there exists a covering of E by open balls $(B(r_{k;\ell}))_{k=1}^{\infty}$ with radius $r_{k;\ell} \leq \frac{1}{\ell}$ such that

$$\mathcal{H}^s_{\frac{1}{\ell}}(\bigcup_{k=1}^{\infty} B(r_{k;\ell})) \leq \mathcal{H}^s_{\frac{1}{\ell}}(E) + \frac{1}{\ell} \leq \mathcal{H}^s(E) + \frac{1}{\ell}. \tag{1.15}$$

In the last step we used again $\mathcal{H}^s(A) = \sup_{\delta>0} \mathcal{H}^s_\delta(A)$.

Then $G := \bigcap_{\ell=1}^{\infty} \bigcup_{k=1}^{\infty} B(r_{k;\ell})$ is a Borel set, and we have

$$\mathcal{H}^s(G) = \lim_{\ell\to\infty} \mathcal{H}^s_{\frac{2}{\ell}}(G) \leq \limsup_{\ell\to\infty} \mathcal{H}^s_{\frac{2}{\ell}}(\bigcup_{k=1}^{\infty} B(r_{k;\ell})) \overset{(1.15)}{\leq} \mathcal{H}^s(E) + 0$$

Since on the other hand $G \supset E$ we have

$$\mathcal{H}^s(G) = \mathcal{H}^s(E).$$

We can conclude. □

Definition 1.64. Let X be a metric space and $\mu : 2^X \to [0, \infty]$ a Borel regular measure. μ is called a *Radon measure* if $\mu(K) < \infty$ for all compact sets.

Example 1.65.
- $\mathcal{L}^n$ is a Radon measure on $\mathbb{R}^n$
- $\mathcal{H}^s$ is not a Radon measure in $\mathbb{R}^n$ whenever $s < n$.

Example 1.66. Let X is a locally compact and separable metric space[6] and μ is a Borel measure such that $\mu(K) < \infty$ for all K compact.

Then X is the union of countably many open sets of finite measure and thus μ coincides with a Radon measure $\tilde{\mu}$ on all Borel sets; cf. Corollary 1.61.

Exercise 1.67. *Let X be a metric space and μ any Radon measure. Suppose that X is a union of countably many open sets of finite μ-measure. Let $A \subset X$ be μ-measurable. Show that $\mu \llcorner A$ is a Radon measure.*

Theorem 1.68. *Let X be a metric space and μ any Borel-regular measure. Suppose that X is a union of countably many open sets of finite μ-measure.*

(1) If μ is Borel-regular, then for any $A \subset X$ we have
$$\mu(A) = \inf_{A \subset G,\ G \text{ open}} \mu(G)$$
(2) If $X = \mathbb{R}^n$ and μ is a Radon measure, for any μ-measurable $A \subset X$ we have
$$\mu(A) = \sup_{F \subset A,\ F \text{ compact}} \mu(F).$$

Proof. (1) Let $A \subset X$. By monotonicity we always have
$$\mu(A) \leq \inf_{A \subset G,\ G \text{ open}} \mu(G).$$
For the other inequality, since μ is by assumption Borel-regular we find a Borel set $E \supset A$ such that $\mu(A) = \mu(E)$. By Theorem 1.53 we have
$$\mu(A) = \mu(E) = \inf_{E \subset G,\ G \text{ open}} \mu(G) \geq \inf_{A \subset G,\ G \text{ open}} \mu(G)$$
(2) Follows from part (1), see Exercise 1.69.

□

Exercise 1.69. *Show Theorem 1.68 (2).*

Use the argument as in the proof of Theorem 1.53 last part and Theorem 1.68(1). Observe we need measurability of A because we need to use that $\mu(U \backslash (X_n \backslash A)) = \mu(U) - \mu(X_n \backslash A)$, which is the measurability of $X_n \backslash A$.

Theorem 1.70. *Let X be a metric space μ a Radon measure on X, and assume that X is a union of countably many open sets of finite μ-measure. Then the following are equivalent for $A \subset X$.*

(1) A is μ-measurable
(2) For every $\varepsilon > 0$ there is an open set $G \subset X$ such that $A \subset G$ and $\mu(G \backslash A) < \varepsilon$

[6]Separable means, a countable set is dense. Locally compact means that ever point has a neighborhood whose closure is compact. $\mathbb{R}^n$ is locally compact and separable, but also $\mathbb{R}^n \setminus 0$ is locally compact and separable

(3) There is a G_δ-set, namely a set $H = \bigcap_{i=1}^\infty G_i$ for some open sets $G_i \subset \mathbb{R}^n$, such that $A \subset H$ and $\mu(H \backslash A) = 0$
(4) For every $\varepsilon > 0$ there is a closed set F such that $F \subset A$ and $\mu(A \backslash F) < \varepsilon$
(5) There is a F_σ-set, namely a set $M = \bigcup_{i=1}^\infty F_i$ for some closed set $F_i \subset \mathbb{R}^n$, such that $M \subset A$ and $\mu(A \backslash M) = 0$
(6) For every $\varepsilon > 0$ there is an open set G and a closed set F such that $F \subset A \subset G$ and $\mu(G \backslash F) < \varepsilon$.
(7) $A = B \cup N$ where B is a Borel set and N is a μ-zero-set.

Proof. By assumption we have $X = \bigcup_{k=1}^\infty X_k$ with X_k open and $\mu(X_k) < \infty$.

(1) $\Rightarrow$ (2): Let $\varepsilon > 0$. Set $A_k := A \cap X_k$ then $\mu(A_k) < \infty$ and by Theorem 1.68 there exists for each k an open set $G_k \supset A_k$ such that

$$\mu(G_k) \leq \mu(A_k) + \frac{\varepsilon}{2^k}.$$

Since A_k is measurable and $G_k \supset A_k$ we have

$$\mu(G_k \backslash A_k) = \mu(G_k) - \mu(A_k) \leq \frac{\varepsilon}{2^k}.$$

Hence $A \subset G := \bigcup_{k=1}^\infty G_k$ and we have

$$\mu(G \setminus A) \leq \sum_{k=1}^\infty \mu(G_k \setminus A) \leq \sum_{k=1}^\infty \mu(G_k \setminus A_k) \leq \sum_{k=1}^\infty \frac{\varepsilon}{2^k} = \varepsilon.$$

Letting $\varepsilon \to 0$ we conclude. (2) $\Rightarrow$ (3). We set $H := \bigcap_{i=1}^\infty U_i$ where U_i are open sets such that $A \subset U_i$ and $\mu(U_i \backslash A) < \frac{1}{i}$. Then $A \subset H$ and

$$\mu(H \setminus A) \leq \mu(U_i \setminus A) \leq \frac{1}{i} \xrightarrow{i \to \infty} 0.$$

(3) $\Rightarrow$ (1) Let $A \subset H$ for some G_δ-set H with $\mu(H \setminus A) = 0$. Then for any $B \subset X$ we have by monotonicity

$$\mu(B \cap A) \leq \mu(B \cap H)$$

and since $B \cap H \subset (B \cap A) \cup (H \backslash A)$,

$$\mu(B \cap H) \leq \mu(B \cap A) + \mu(H \backslash A) = \mu(B \cap A).$$

thus we actually have

$$\mu(B \cap H) = \mu(B \cap A).$$

Similarly,

$$\mu(B \setminus H) \leq \mu(B \setminus A)$$

and since $B \setminus A \subset (B \setminus H) \cup (H \setminus A)$,

$$\mu(B \setminus A) \leq \mu(B \setminus H) + \mu(H \setminus A) = \mu(B \setminus H).$$

Thus

$$\mu(B \setminus H) = \mu(B \setminus A).$$

Now since any G_δ-set is a Borel set and μ is in particular a Borel measure we have that H is μ-measurable. Consequently

$$\mu(B) = \mu(B \cap H) + \mu(B \setminus H) = \mu(B \cap A) + \mu(B \setminus A).$$

Thus A is μ-measurable.

(1) $\Leftrightarrow$ (4) $\Leftrightarrow$ (5) this equivalence follows from the equivalence of the conditions (1), (2), and (3) applied to $X \setminus A$.

(1) $\Rightarrow$ (6) If A is μ-measurable, the existence of the sets F and G follows from the conditions (2) and (4).

(6) $\Rightarrow$ (3) Take closed and open sets F_i, G_i such that $F_i \subset A \subset G_i$, $\mu(G_i \setminus F_i) < \frac{1}{i}$. Then the set $H := \bigcap_{i=1}^{\infty} G_i$ is G_δ and $\mu(H \setminus A) = 0$.

(7) $\Rightarrow$ (1) If $A = B \cup N$ where B is a Borel set and N is a μ-zero-set, then clearly A is it is the union of two μ-measurable sets.

(5) $\Rightarrow$ (7) We have $A = M \cup A \setminus M$, where M is a F_σ-set, in particular a Borel set. and $A \setminus M$ is a μ-zero set. □

From Theorem 1.56 we obtain the following uniqueness result

Theorem 1.71 (Radon measures on $\mathbb{R}^n$ that coincide on rectangles)**.** *Let μ and ν be two Radon measures on $\mathbb{R}^n$ such that*

$$\mu(R) = \nu(R)$$

for all closed rectangles R of the form

$$R := \{x \in \mathbb{R}^n : \quad a_i \le x_i \le b_i, \quad i = 1, \dots, n\}$$

where $-\infty \le a_i \le b_i \le \infty$, $i = 1, \dots, n$.

Then

$$\mu(A) = \nu(A) \quad \forall A \subset \mathbb{R}^n.$$

Proof. By Theorem 1.56 we have

$$\mu(K) = \nu(K)$$

for all compact set K (observe that then the measure is w.l.o.g. finite). By Theorem 1.68(2) we obtain that

$$\mu(A) = \nu(A)$$

for all Borel sets (which are both μ- and ν-measurable).

If $A \subset \mathbb{R}^n$ is any set, then there exists two Borel sets $B_\mu \supset A$ and $B_\nu \supset A$ such that

$$\mu(A) = \mu(B_\mu)$$

and

$$\nu(A) = \nu(B_\nu).$$

Observe that we have

$$\mu(A) \leq \mu(B_\nu \cap B_\mu) = \nu(B_\nu \cap B_\mu) \leq \nu(B_\nu) = \nu(A).$$

Similarly we find $\nu(A) \leq \mu(A)$. That is $\mu(A) = \nu(A)$ and we can conclude. □

Proposition 1.72. *Let μ be a Radon measure on $\mathbb{R}^n$. Then for each $x \in \mathbb{R}^n$ there are at most countably many $r > 0$ such that*

$$\mu(\partial B(x,r)) > 0.$$

Exercise 1.73. *Construct a Radon measure μ on $\mathbb{R}^n$ such that for countably many $r > 0$,*

$$\mu(\partial B(x,r)) > 0.$$

Proof of Proposition 1.72. For simplicity let us assume that $x = 0$.

Now consider

$$f(r) := \mu(\overline{B(r)}).$$

Clearly $f : (0,\infty) \to [0,\infty)$ is increasing, so f has at most countably many points of discontinuity.

Moreover we observe that

$$\mu(B(r)) = \lim_{\tilde{r} \to r^-} f(\tilde{r}).$$

Indeed, this follows since for any increasing sequence $r_1 < r_2 < \ldots < r$ with $\lim r_i = r$, so by Theorem 1.33(2)

$$\mu(B(r)) = \mu(\bigcup_{i=1}^{\infty} \overline{B(r_i)}) = \lim_{i\to\infty} \mu(\overline{B(r_i)}) = \lim_{i\to\infty} f(r_i).$$

Next we claim that whenever r is a point of continuity for f then $\mu(\partial B(r)) = 0$. Indeed, since $\overline{B(r)}$ and $\partial B(r)$ are compact and $B(r)$ open (thus all are measurable),

$$0 \leq \mu(\partial B(r)) = \mu(\overline{B(r)}) - \mu(B(r)) = f(r) - \lim_{\tilde{r}\to r^-} f(\tilde{r}).$$

So if r is a point of continuity, then $\lim_{\tilde{r}\to r^-} f(\tilde{r}) = f(r)$ and thus

$$\mu(\partial B(r)) = 0.$$

□

Similarly we have

Exercise 1.74. *Let μ be a Radon measure on $\mathbb{R}^n$ and $g \in C^0(\mathbb{R}^n, \mathbb{R})$ such that $g \equiv 0$ outside a compact set K. Then there exist at most countable $r \in \mathbb{R}$ such that*

$$\mu(g^{-1}(r)) > 0.$$

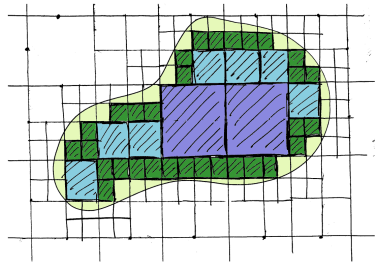

FIGURE 1.5. An open set can be represented by a countable union of dyadic cubes.

1.5. **More on the Lebesgue measure.**

Lemma 1.75 (Dyadic decomposition of open sets)**.** *Any open set $\Omega \subset \mathbb{R}^n$ can be written as*

$$\Omega = \bigcup_{i=1}^{\infty} Q_i$$

where Q_i are closed cubes with pairwise disjoint interior and each cube has sidelength 2^{-k} for some $k \in \mathbb{N}$.

Proof. See Figure 1.5 for a picture proof. We consider *dyadic cubes* of sidelength $2^{-\ell}$, $\ell \in \mathbb{N} \cup \{0\}$ which cover $\mathbb{R}^n$:

$$\mathbb{R}^n = \bigcup_{a \in 2^\ell \mathbb{Z}^n} a + [0, 2^{-\ell}]^n.$$

Let

$$\mathcal{Q}_\ell := \left\{ Q = a + [0, 2^{-\ell}]^n : \quad a \in 2^\ell \mathbb{Z}^n \text{ and } Q \subset \Omega \setminus \left(\bigcup_{\tilde{Q} \in \mathcal{Q}_{\ell-1}} \tilde{Q} \right)^o \right\}.$$

Each $\mathcal{Q}_\ell$ is a countable family of cubes, so $\mathcal{Q} := \bigcup_{\ell=0}^{\infty} \mathcal{Q}_\ell$ is a countable family of cubes.

By construction we have

$$\bigcup_{Q \in \mathcal{Q}} Q \subset \Omega.$$

Observe also that by construction two cubes in $\mathcal{Q}$ intersect only along their boundary.

Now assume by contradiction that $\bigcup_{Q \in \mathcal{Q}} Q \neq \Omega$, i.e.

$$\Omega \setminus \bigcup_{Q \in \mathcal{Q}} Q \supset \{x_0\}.$$

Observe that any point $x \in \bigcup_{Q \in \mathcal{Q}} Q$ belongs to at most 2^n many cubes $Q \in \mathcal{Q}$.

We conclude that $\bigcup_{Q \in \mathcal{Q}} Q$ is a closed set, even though it may be a countable union of closed cubes. Indeed assume $(x_i)_{i \in \mathbb{N}} \subset \bigcup_{Q \in \mathcal{Q}} Q$ converges to some $\bar{x}$. Then all but finitely many x_i belong to the same cube $\tilde{Q}$ which is closed, so $\bar{x} \in \tilde{Q}$.

By the argument from above, $\Omega \setminus \bigcup_{Q\in\mathcal{Q}} Q$ is open thus there exists a small ball $B(x_0,\rho) \subset \Omega \setminus \bigcup_{Q\in\mathcal{Q}} Q \supset \{x_0\}$. Let k_0 such that $2^{-k_0} < \frac{\rho}{1000}$. Then we can find a tiny cube $\bar{Q} = a + [0, 2^{-k_0}] \subset B(x_0,\rho)$ (each component of x_0 lies between some $[2^{-k_0(\gamma+1)}, 2^{-k_0(\gamma)}]$). Contradiction because that cube would belong to $\tilde{Q}$. □

Exercise 1.76. *Show that any open set $U \subset \mathbb{R}^n$ can be written as a countable union of open cubes.*

The Lebesgue measure is essentially built from blocks, the dyadic decomposition shows that the Lebesgue measure is then the only natural volume notion on $\mathbb{R}^n$: any other translation invariant volume notion is the same.

Theorem 1.77 (Uniqueness of the Lebesgue measure)**.** *If μ is a Borel-measure on $\mathbb{R}^n$ such that*

- *$\mu(a+E) = \mu(E)$ for all $a \in \mathbb{R}^n$ and all Borel-sets $E \subset \mathbb{R}^n$ (translation invariance)*
- *$\mu([0,1]^n) = 1$*

then $\mu(E) = \mathcal{L}^n(E)$ for all Borel-sets E.

If μ is moreover Borel-regular (and thus a Radon measure) then $\mu(E) = \mathcal{L}^n(E)$ for all $E \subset \mathbb{R}^n$.

Proof. Consider an open cube $Q_k = (0, 2^{-k})^n$ where k is a positive integer. There are 2^{kn} pairwise disjoint cubes contained in the unit cube $[0,1]^n$, each being a translation of Q_k. Since the measure μ is invariant under translation we have

$$\mu(Q_k) \leq 2^{-kn}.$$

Now we can use a similar argument to cover ∂Q by translations small cubes in such a way that the sum of μ-measures of the cubes covering ∂Q goes to zero as $k \to \infty$. Thus $\mu(\partial Q) = 0$ for any cube Q.

Now $[0,1]^n$ can be covered by 2^{kn} cubes which are translations of $\overline{Q_k}$. So,

$$1 = \mu([0,1]^n) \leq 2^{kn}\mu(\overline{Q_k}).$$

Thus, $\mu(\overline{Q_k}) \geq 2^{-kn}$ and $\mu(Q_k) \leq 2^{-kn}$. Since $\mu(\partial Q_k) = 0$ we find that $\mu(\overline{Q_k}) = \mu(Q_k) = 2^{-kn}$ for all k.

In conclusion: any cube Q of sidelength 2^{-k} satisfies

$$\mu(Q) = \mathcal{L}^n(Q).$$

By dyadic decomposition, Lemma 1.75, any open set E can be written as $E = \bigcup_{i=1}^{\infty} Q_i$ for cubes which are translations of Q_k for some k, and with pairwise disjoint interior. Thus

$$\mu(E) = \sum_i \mu(Q_i) = \sum_i \mathcal{L}^n(Q_i) = \mathcal{L}^n(E).$$

In view of Theorem 1.53 applied to μ and the Lebesgue measure we conclude that $\mu(E) = \mathcal{L}^n(E)$ for all Borel sets E.

If μ is Borel regular then the additional claim follows from Theorem 1.71. $\square$

Since the n-Hausdorf measure $\mathcal{H}^n$ a translation invariant, Exercise 1.16, Borel measure, we obtain that it essentially coincides with the Lebesgue measure.

Corollary 1.78. *Let $n \in \mathbb{N}$. Then n-th Hausdorff measure and Lebesgue measure coincide in $\mathbb{R}^n$ up to a multiplicative factor C_n for all Borel sets. That is*

$$\mathcal{H}^n(E) = \mathcal{L}^n(E) \quad \forall E \subset \mathbb{R}^n.$$

Proof. From Example 1.63, Exercise 1.16, Theorem 1.77 we have $\mathcal{H}^n(E) = C\mathcal{L}^n(E)$ for some constant C. The fact that $C = 1$ one can check in by testing for $E = B(0,1)$. $\square$

Proposition 1.79. *A set $E \subset \mathbb{R}^n$ has Lebesgue measure zero if and only if for every $\varepsilon > 0$ there is a family of balls $(B(x_i, r_i))_{i=1}^{\infty}$ such that*

$$E \subset \bigcup_{i=1}^{\infty} B(x_i, r_i)$$

and

$$\sum_{i=1}^{\infty} r_i^n < \varepsilon.$$

Proof. $\Leftarrow$. Assume

$$E \subset \bigcup_{i=1}^{\infty} B(x_i, r_i) \tag{1.16}$$

and

$$\sum_{i=1}^{\infty} r_i^n < \varepsilon. \tag{1.17}$$

Obserserve that if $Q(x_i, 2r_i)$ denotes the cube with sidelength $2r_i$ centered at x_i then $B(x_i, r_i) \subset Q(x_i, 2r_i)$ and thus by the definition of the Lebesgue measure, Definition 1.43,

$$\mathcal{L}^n(B(x_i, r_i)) \leq \mathrm{vol}(Q(x_i, 2r_i)) = (4r_i)^n.$$

Consequently, by monotonicity

$$\mathcal{L}^n(E) \leq \sum_{i=1}^{\infty} 4^n r_i^n \leq 4^n \varepsilon.$$

Thus if for every $\varepsilon > 0$ there is a family of balls $(B(x_i, r_i))_{i=1}^{\infty}$ such that (1.16) and (1.16) holds, we have

$$\mathcal{L}^n(E) = 0.$$

That is E is a $\mathcal{L}^n$-zero set.

$\underline{\Rightarrow}$: Assume $\mathcal{L}^n(E) = 0$.

By definition of the Lebesgue measure, Definition 1.43, for any $\varepsilon > 0$ there exist $(A_k)_{k=1}^{\infty}$ figures which each can decomposed into blocks of pairwise disjoint interior $A_k = \bigcup_{i=1}^{N_k} Q_{k;i}$, such that

$$E \subset \bigcup_{k=1}^{\infty} A_k \equiv \bigcup_{k=1}^{\infty} Q_{k;i}$$

and

$$\sum_{k=1}^{\infty} \mathrm{vol}(A_k) \equiv \sum_{k=1}^{\infty} \sum_{i=1}^{N_k} \mathrm{vol}(Q_{k;i}) < \varepsilon.$$

Denote by $L_{k;i}$ the sidelength of the cube $Q_{k;i}$, and by $x_{k;i}$ the center of the cube $Q_{k;i}$. Then $Q_{k;i} \subset B(x_{k;i}, L_{k;i})$ and

$$\mathrm{vol}(Q_{k;i}) = c_n (L_{k;i})^n$$

for some dimensional constant c_n. Thus,

$$E \subset \bigcup_{k}^{\infty} \bigcup_{i=1}^{N_k} B(x_{k;i}, L_{k;i})$$

and

$$\sum_{k=1}^{\infty} \sum_{i=1}^{\infty} (L_{k;i})^n = \frac{1}{c_n} \sum_{k=1}^{\infty} \sum_{i=1}^{\infty} \mathrm{vol}(Q_{k;i}) < \frac{\varepsilon}{c_n}.$$

□

Exercise 1.80. *Use Proposition 1.79 to show that*

(1) if $f : \mathbb{R}^n \to \mathbb{R}^n$ is uniformly continuous, then for any $A \subset \mathbb{R}^n$ with $\mathcal{L}^n(A) = 0$ we also have $\mathcal{L}^n(f(A)) = 0$. This property of f is called the Lusin property.
(2) Whenever $\Sigma \subset \mathbb{R}^n$ with $\mathcal{L}^n(\Sigma) = 0$ then $\mathbb{R}^n \setminus \Sigma$ is dense in $\mathbb{R}^n$, i.e. $\overline{\mathbb{R}^n \setminus \Sigma} = \mathbb{R}^n$.

So we know the Lebesgue measure is translation invariant (and that its pretty much the only measure that does that). But we have not verified that this means that it is rotation invariant. Namely what is the $\mathcal{L}^n$ measure of a cube Q rotated? The following result shows (in particular) that rotations do not change the measure.

Theorem 1.81. *Let $Lx := Ax + b$ be an affine linear map $\mathbb{R}^n \to \mathbb{R}^n$ with a non-degenerate matrix $A \in \mathbb{R}^{n \times n}$, i.e. $\det(A) \neq 0$, and $b \in \mathbb{R}^n$. Then*

(1) $E \subset \mathbb{R}^n$ is a Borel set if and only if $L(E) \subset \mathbb{R}^n$ is a Borel set.
(2) E is $\mathcal{L}^n$-measurable if and only if $L(E)$ is $\mathcal{L}^n$-measurable
(3) We have $\mathcal{L}^n(L(\Omega)) = |\det(L)|(\Omega)$ for all $\Omega \subset \mathbb{R}^n$.

Proof. (1) L is a homeomorphism, so $E \subset \mathbb{R}^n$ is a Borel set if and only if $L(E) \subset \mathbb{R}^n$ is a Borel set, see Exercise 1.51.

(2) In view of Theorem 1.70(7), a set A is $\mathcal{L}^n$-measurable if and only if $A = B \cup N$ where B is a Borel set and N is a zero set. Since L and L^{-1} are both Lipschitz maps they map zero sets into zero sets, Exercise 1.13 and Corollary 1.78. Since L and L^{-1} also preserve the Borel property of set we get the claim.

(3) By Theorem 1.68 applied to the Lebesgue measure $\mathcal{L}^n$ it suffices to show

$$\mathcal{L}^n(L(\Omega)) = \det(L)(\Omega) \text{ for all open } \Omega \subset \mathbb{R}^n$$

Let

$$\mu(A) := \mathcal{L}^n(L(A)) \quad A \in \mathbb{R}^n.$$

Then μ is a Borel measure, and it is still translation invariant. So we apply Theorem 1.77 to $\tilde{\mu} := \frac{1}{a}\mu$ where $a := \mu([0,1]^n)$ and have that

$$\mathcal{L}^n(L(\Omega)) = \mathcal{L}^n(L([0,1]^n))\,\mathcal{L}^n(\Omega) \quad \text{for all open sets } \Omega.$$

So it remains to show that

$$\mathcal{L}^n(A([0,1]^n)) = |\det(A)| \quad \forall A \in GL(n).$$

So consider $f : GL(n) \to (0,\infty)$, $f(A) := \mathcal{L}^n(A([0,1]^n))$. We observe two properties: first,

$$f(AB) = f(A)f(B) \quad \forall A, B \in GL(n). \tag{1.18}$$

Indeed, we have

$$f(AB) = \mathcal{L}^n(AB([0,1]^n)) = f(A)\mathcal{L}^n(B([0,1]^n)) = f(A)f(B)\mathcal{L}^n(([0,1]^n)).$$

Secondly, for the identity matrix $I \in \mathbb{R}^{n\times n}$

$$f(\lambda I) = \lambda^n, \quad \forall \lambda > 0. \tag{1.19}$$

This is immediate from the definition of the pre-measure of the Lebesgue measure.

We conclude with the following Lemma, Lemma 1.82

□

Lemma 1.82. *Let $f : GL(n) \to (0,\infty)$ satisfy (1.18) and (1.19). Then*

$$f(A) = |\det(A)|.$$

Proof. Let $A_i(s)$ be the diagonal matrix with $-s$ on the ith place and s on all other places on the diagonal. Since $A_i(s)^2 = s^2 I$ we have $f(A_i(s))^2 = f(A_i(s)^2) = s^{2n}$ and hence $f(A_i(s)) = |s^n| = |\det A_i(s)|$. For $k \neq \ell$ let $B_{k\ell}(s) = (a_{ij})_{ij}$ be the matrix such that $a_{k\ell} = s$, $a_{ii} = 1$, $i = 1, 2, \ldots, n$ and all other entries equal zero. Multiplication by the matrix $B_{kl}(s)$ from the right (left) is equivalent to adding kth column (ℓth row) multiplied by s to ℓth column (kth row).

It is well known from linear algebra (and easy to prove) that applying such operations to any nonsingular matrix A it can be transformed to a matrix of the form tI or $A_n(t)$. Since multiplication by $B_{k\ell}(s)$ does not change determinant, $t = |\det A|^{1/n}$. It remains to prove that $f(B_{kl}(s)) = 1$. Since $B_{kl}(-s) = A_k(1)\,B_{kl}(s)\,A_k(1)$ we have that $f(B_{kl}(-s)) =$

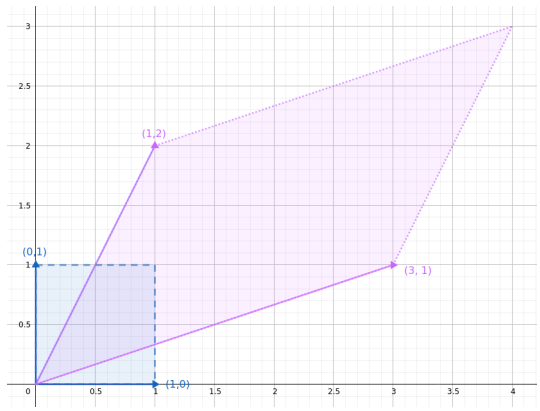

FIGURE 1.6. Let $a_{11} = 3$, $a_{12} = 1$, $a_{21} = 1$ and $a_{22} = 2$ then $[0,1]^2$ (blue) is transformed into $A[0,1]^2$ (purple). The purple area is $\det(A) = 5$.

$f(B_{kl}(s))$. On the other hand $B_{k\ell}(s)B_{k\ell}(-s) = I$ and hence $f(B_{kl}(s))^2 = 1$, so $f(B_{kl}(s)) = 1$. The proof of Lemma 1.82 and hence the proof of Theorem 1.81 is complete. □

Example 1.83. Let $A \in \mathbb{R}^{n\times n}$, and let $I = [0,1]^n$. Set

$$AI := \{Ax, x \in I\}.$$

Then the volume (in the elementary geometrical sense) of AI is

$$|AI| = |\det(A)||I|.$$

Exercise 1.84. *Let $Lx := Ax + b$ be an affine linear map $\mathbb{R}^n \to \mathbb{R}^n$ for any matrix $A \in \mathbb{R}^{n\times n}$ and any $b \in \mathbb{R}^n$. Show the following:*

(1) If E is $\mathcal{L}^n$-measurable then $L(E)$ is $\mathcal{L}^n$-measurable
(2) We have $\mathcal{L}^n(L(\Omega)) = |\det(L)|(\Omega)$ for all $\Omega \subset \mathbb{R}^n$.

Hint: If $\det(A) \neq 0$ this follows from Theorem 1.81. If $\det(A) = 0$, what is $L(E)$ or $L(\Omega)$? (compare to the Hausdorff measure $\mathcal{H}^s$ for $s = \operatorname{rank} A$)

Corollary 1.85. *The Lebesgue measure $\mathcal{L}^n$ is translation and rotationally invariant on $\mathbb{R}^n$. Namely if $A \subset \mathbb{R}^n$ then $\mathcal{L}^n(A) = \mathcal{L}^n(\Phi(A))$ where for some $x_0 \in \mathbb{R}^n$ and $R \in O(n)$ we have*

$$\Phi(x) := x_0 + Rx$$

then $\mathcal{L}^n(\Phi(A)) = \mathcal{L}^n(A)$.

1.6. **Nonmeasurable sets.** Corollary 1.85 combined with the Banach-Tarski paradoxon, (1.1) and Figure 1.1, implies the existence of nonmeasurable sets (w.r.t. $\mathcal{L}^n$) in $\mathbb{R}^n$: Any set A and B such that for some pairwise disjoint and *measurable* C_i they are represented by

$$A = \bigcup_{i=1}^{N} C_i, \quad \text{and} \quad B = \bigcup_{i=1}^{N} (x_i + O_i C_i).$$

(where $O_i \in O(n)$ is a rotation and x_i a point, and $(x_i + O_i C_i)_{i=1}^N$ are pairwise disjoint) would satisfy

$$\mathcal{L}^n(A) = \sum_{i=1}^N \mathcal{L}^n(C_i) = \sum_{i=1}^N \mathcal{L}^n(x_i + O_i C_i) = \mathcal{L}^n(B),$$

so if $A = [0,1]^n$ and $B = [0,2]^n$ this would mean the impossible $\mathcal{L}^n(A) = \mathcal{L}^n(B)$ – so one of the C_i must be nonmeasurable.

We will not prove the Banach-Tarski-paradox, but instead show the $\mathbb{R}^1$-version.Observe below that the existence of nonmeasurable sets requires the the axiom of choice (and suitable invariances of the underlying measure).

Theorem 1.86 (Vitali)**.** *Let $\mu : 2^{\mathbb{R}} \to \mathbb{R}$ be translation invariant*[7]*, i.e.*

$$\mu(x + A) = \mu(A) \quad \forall x \in \mathbb{R}, A \subset \mathbb{R}.$$

If moreover $\mu([0,1]) \in (0,\infty)$ then there exists a Vitali-set $A \subset [0,1]$ that is not μ-measurable.

Proof. Construction (Vitali) Fix $\xi \in \mathbb{R}\backslash\mathbb{Q}$, and set

$$G_\xi := \{k + \ell\xi; \quad k, \ell \in \mathbb{Z}\}$$

We use G_ξ to define an equivalence relation $\sim$ on $\mathbb{R}$.

$$x \sim y :\Leftrightarrow x - y \in G_\xi.$$

For $x \in \mathbb{R}$ denote by $[x]$ the set

$$[x] := x + G_\xi = \{y \in \mathbb{R} : \quad y = x + k + \ell\xi \quad k, \ell \in \mathbb{Z}\}.$$

Let $A \subset \mathbb{R}$ be a set such that for each class $[x]$ there exists exactly one element $y \in A \cap [x]$. The set A exists by the *axiom of choice*: if we set

$$X := \{[x] \subset \mathbb{R} : \quad x \in \mathbb{R}\}$$

then the axiom of choice says there exists a choice function $f : X \to \mathbb{R}$ such that $f([x]) \in [x]$ for all $[x] \in X$. Then $A := f(X)$.

Without loss of generality, $A \subset [0,1]$. Indeed if we can adapt the choice function f above such that

$$\tilde{f}([x]) := f([x]) - k,$$

where $k \in \mathbb{Z}$ is chosen such that $f([x]) \in [k, k+1)$.

The Vitali-set A is not μ-measurable

Since $\xi \notin \mathbb{Q}$ the set $G_\xi \cap [-1,1]$ contains infinitely many points. Indeed: if $k + \ell\xi = k' + \ell'\xi$ then $k = k'$ and $\ell = \ell'$ (otherwise $\xi = \frac{k-k'}{\ell'-\ell} \in \mathbb{Q}$). Thus we can find infinitely many different $k + \ell\xi \in [-1,1]$ (choosing $k = -\lfloor \ell\xi \rfloor$ or similar)

[7] yes, if μ is moreover Borel then it is pretty much the Lebesgue measure, Theorem 1.77, but the main point of this theorem is: invariances mean non-measurable sets

Also, G_ξ is clearly countable we can write $G_\xi \cap [-1,1] = \bigcup_{k=1}^\infty g_k$ where $(g_k)_{k=1}^\infty$ are pairwise different points. Set $A_k := g_k + A \subset [-1,2]$, $k \in \mathbb{N}$.

We claim

$$A_k \cap A_\ell = \emptyset \quad k \neq \ell. \tag{1.20}$$

Indeed, assume that $x \in A_k \cap A_\ell$ then

$$x - g_k, x - g_\ell \in A.$$

Observe that $x - g_k \sim x - g_\ell$, that is $[x] = [x - g_k] = [x - g_\ell]$. But by definition of A there is exactly one element of $[x - g_k] = [x - g_\ell]$ in A, which implies that $x - g_k = x - g_\ell$, that is $g_k = g_\ell$, that is $k = \ell$. This establishes (1.20).

Next we claim that

$$[0,1] \subset \bigcup_{k=1}^\infty A_k. \tag{1.21}$$

Indeed, let $y \in [0,1]$. By the construction of A there must be exactly one $x \in A \cap [y]$. In particular $y - x \in G_\xi$. Since $x, y \in [0,1]$ we have $y - x \in [-1,1] \cap G_\xi$, and thus there must be some $g_k = y - x$. Consequently

$$y = x + g_k \subset A + g_k = A_k.$$

This establishes (1.21).

Now assume A is μ-measurable, i.e.

$$\mu(B) = \mu(A \cap B) + \mu(B \setminus A) \quad \forall B \subset \mathbb{R}.$$

Then by translation invariance, so is A_k, indeed for any $B \subset \mathbb{R}$,

$$\begin{aligned}
\mu(B) =& \mu(B - g_k) = \mu(A \cap (B - g_k)) + \mu((B - g_k) \setminus A) \\
=& \mu(g_k + A \cap (B - g_k)) + \mu(g_k + (B - g_k) \setminus A) \\
=& \mu((A + g_k) \cap B) + \mu(B \setminus (A + g_k)) \\
=& \mu(A_k \cap B) + \mu(B \setminus A_k).
\end{aligned}$$

Then we have by (1.20) and σ-additivity of disjoint measurable sets, Theorem 1.33,

$$\sum_{k=1}^\infty \mu(A_k) = \mu(\bigcup_{k=1}^\infty A_k).$$

Since $[0,1] \subset \bigcup A_k \subset [-1,2]$ we have again by translation invariance

$$\mu([0,1]) \leq \sum_{k=1}^\infty \mu(A_k) \leq \mu([0,3]) \leq \mu([0,1]) + \mu([1,2]) + \mu([2,3]) = 3\mu([0,1]).$$

On the other hand, again by translation invariance

$$\mu(A_k) = \mu(A).$$

That is if A was measurable we'd have

$$\mu([0,1]) \leq \sum_{k=1}^{\infty} \mu(A) \leq 3\mu([0,1]).$$

This is only possible if $\mu([0,1]) = \infty$ or $\mu([0,1]) = 0$, both are ruled out by assumption. $\square$

So in general even for very 'reasonable' measures μ (in the sense that they measure indeed something like area, volume, length etc.) we cannot really hope to avoid the existence of non-measurable sets. The way to deal with that fact is to shun non-measurable sets, and only work with measurable sets.

This strategy will propagate throughout this course: we only care about functions that stay (in a reasonable way) within the category of measurable sets.

2. Measurable functions

Our goal is integration, and for this purpose it is convenient with functions that can be infinite (think of $\frac{1}{|x|^2}$). We will use the notation

$$\bar{\mathbb{R}} := \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}.$$

Definition 2.1. Let (X, Σ, μ) be a measure space. We say that $f : X \to \overline{\mathbb{R}}$ is a *measurable function*, if f^{-1} maps open sets of $\mathbb{R}$ into μ-measurable sets (where we consider $\{+\infty\}$, $\{-\infty\}$ as open sets). That is

$$f^{-1}(U) \in \Sigma \quad \forall \text{open sets } U \subset \mathbb{R} \text{ and } U = \{+\infty\} \text{ and } U = \{-\infty\}.$$

More generally if Ω is μ-measurable then $f : \Omega \to Y$ is a measurable function if $f^{-1}(U)$ is μ-measurable for any open set $U \subset Y$ (and $U = \{+\infty\}$ and $U = \{-\infty\}$)

For consistency we observe

Lemma 2.2. *Let (X, Σ, μ) be a measure space $\Omega \subset X$ μ-measurable and $f : \Omega \to \overline{\mathbb{R}}$. Recall the notation of the measure $\mu \llcorner \Omega$,*

$$\mu \llcorner \Omega(A) := \mu(\Omega \cap A).$$

Denote the $\mu \llcorner \Omega$-measurable sets by Σ_Ω. Then

$$\Sigma_\Omega := \{A \subset \Omega : \quad A \ \mu\text{-measurable}\} = \{\Omega \cap B : \quad B \ \mu\text{-measurable}\}$$

And in particular the following are equivalent

(1) f is μ-measurable with respect to (X, Σ, μ)
(2) f is $\mu \llcorner \Omega$-measurable with respect to $(\Omega, \Sigma, \mu \llcorner \Omega)$.

Exercise 2.3. *Let X be some set, μ a measure on X. Assume $f : X \to \overline{\mathbb{R}}$.*

Show that

$$\mathcal{G} := \Big\{B \subset \mathbb{R} : f^{-1}(B) : \quad \textit{is } \mu\textit{-measurable}\Big\} \subset 2^{\mathbb{R}}$$

is a σ-Algebra in $\mathbb{R}$.

Lemma 2.4. *Let (X, Σ, μ) be a measure space $\Omega \subset X$ μ-measurable and $f : \Omega \to \overline{\mathbb{R}}$.*

The following are equivalent

(1) $f^{-1}(U)$ is μ-measurable for any open set $U \subset \mathbb{R}$
(2) $f^{-1}(B)$ is μ-measurable for any Borel set $B \subset \mathbb{R}$
(3) $f^{-1}((-\infty, a])$, $f^{-1}((-\infty, a))$, $f^{-1}((-\infty, a])$ $f^{-1}([a, \infty))$ are μ-measurable for any $a \in \mathbb{R}$.

Proof. (1) $\Rightarrow$ (2) Observe that

$$\mathcal{G} := \Big\{B \subset \mathbb{R} : f^{-1}(B) : \quad \text{is } \mu\text{-measurable}\Big\}$$

is a σ-Algebra (Exercise 2.3) which in view of (1) contains the open sets. So $\mathcal{G}$ contains the Borel sets.

(2) $\Rightarrow$ (1) and (2) $\Rightarrow$ (3) are obvious since any open set and any interval is a Borel set.

(3) $\Leftrightarrow$ (2) follows the same way as above, since the Borel σ-algebra are generated by the infinite intervals. □

We can (usually we don't want to) extend a bit the notion of measurability to a *topological space* as target.

Definition 2.5 (Topological space)**.** A *topological space* is a set X and a collection $\tau \subset 2^X$ of "open sets" which satisfies the following axioms

- $\emptyset \in \tau$ and $X \in \tau$.
- Let I be a (finite or infinite) index set and let $A_i \in \tau$ for all $i \in I$. Then $\bigcup_{i \in I} A_i \in \tau$.
- Let $N \in \mathbb{N}$ and $A_i \in \tau$ for $i \in \{1, \dots, N\}$ then

$$\bigcap_{i=1}^{N} A_i \in \tau.$$

τ is called a *topology*.

For two topological spaces (X, τ_X) and (Y, τ_Y) a map $f : X \to Y$ is *continuous* if f^{-1} maps open sets to open sets, i.e. $f^{-1}(U) \in \tau_X$ for any $U \in \tau_Y$.

If (X, d) is a metric space then the collection of open sets (w.r.t. the metric) form a topology.

Continuity between two metric spaces (w.r.t. topology) is the same as continuity w.r.t metric structure.

Remark 2.6. When talking about measurability of a map $f : \Omega \to \overline{\mathbb{R}}$ we actually consider $\overline{\mathbb{R}}$ as a *topological space*, where the open sets $A \subset \overline{\mathbb{R}}$ are

- the open sets in $\mathbb{R}$
- the "open neighborhoods" of $+\infty$ and $-\infty$, i.e. $[-\infty, a)$ and $(a, \infty]$ for $a \in \mathbb{R}$

and unions thereof.

Then $f : \Omega \to \overline{\mathbb{R}}$ is μ-measurable if and only if $f^{-1}(U)$ is μ-measurable whenever $U \subset \overline{\mathbb{R}}$ is open.

In particular we have that $f : \Omega \to \overline{\mathbb{R}}$ is *measurable if and only if $f^{-1}([-\infty, a))$ and $f^{-1}((a, +\infty])$ are measurable for all $a \in \mathbb{R}$.*

This leads to the following definition:

Definition 2.7. Let (X, Σ, μ) be a measure space, $\Omega \subset X$ measurable, and (Y, τ) a topological space. A map $f : \Omega \to Y$ is called measurable if $f^{-1}(U)$ is measurable for any open set $U \subset Y$ (i.e. for any set $U \in \tau$).

Example 2.8. Let (X, d) be a metric space and μ be a Borel-measure. Then any continuous function $f : X \to \mathbb{R}$ is μ-measurable.

Proof. $f^{-1}(+\infty) = f^{-1}(-\infty) = \emptyset$ which is clearly μ-measurable. Moreover since f is continuous $f^{-1}(U)$ is open whenever U is open, so $f^{-1}(U)$ is μ-measurable since μ is Borel. □

Exercise 2.9. *Let (X, Σ, μ) be a measure space, Y and Z topological spaces. Assume $f : X \to Y$ is a measurable function and $g : Y \to Z$ is continuous. Then $g \circ f : X \to Z$ is measurable.*

Example 2.10. Let (X, Σ, μ) be a measure space and Y a topological space. If the functions $u_1, u_2, \ldots, u_n : X \to \mathbb{R}$ are measurable and $\Phi : \mathbb{R}^n \to Y$ is continuous, then the function

$$h(x) = \Phi(u_1(x), u_2(x), \ldots, u_n(x)) : X \to Y$$

is μ-measurable.

Proof. We only discuss the case $n = 2$, and write $u_1 = u$ and $u_2 = v$.

Set $f(x) := (u(x), v(x))$. In view of Exercise 2.9 it suffices to prove that f is measurable. We observe that if $R = (a, b) \times (c, d)$ is an open rectangle, then

$$f^{-1}(R) = u^{-1}((a, b)) \cap u^{-1}((c, d)) \text{is measurable.}$$

Since any open set $U \subset \mathbb{R}^2$ can be written as a countable union of open rectangles (Exercise 1.76))

$$U = \bigcup_{i=1}^{\infty} R_i$$

we have

$$f^{-1}(U) = \bigcup_{i=1}^{\infty} f^{-1}(R_i),$$

so $f^{-1}(U)$ is μ-measurable. □

Exercise 2.11. *(1) If $f = (f^1, f^2, \dots, f^n) : X \to \mathbb{R}^n$ where $f^1, \dots, f^n : X \to \mathbb{R}$ are μ-measurable, then f is μ-measurable (take $\Phi(f^1, \dots, f^n) := (f^1, \dots, f^n)^t$).*

(2) If $f : X \to \mathbb{R}^n$ is μ-measurable, then its components $f^i : X \to \mathbb{R}$ are μ-measurable, $i = 1, \dots, n$. (take $\Phi_i(f) := f^i$).

Also $|f|$ is μ-measurable (take $\Phi(f) := |f|$).

(3) If $f, g : X \to \mathbb{R}$ are μ-measurable then so are $f + g$, $f - g$, $-f$, and $fg : X \to \mathbb{R}$ [8]

Example 2.12. (1) A set $E \subset X$ is μ-measurable if and only if its *characteristic function* χ_E

$$\chi_E(x) := \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{otherwise} \end{cases}$$

is μ-measurable. Indeed

$$(\chi_E)^{-1}(U) = \begin{cases} \emptyset & \text{if } 0 \notin U \text{ and } 1 \notin U \\ E & \text{if } 0 \notin U \text{ and } 1 \in U \\ X \backslash E & \text{if } 0 \in U \text{ and } 1 \notin U \\ X & \text{if } 0 \in U \text{ and } 1 \in U \end{cases}$$

(2) There are non-$\mathcal{L}^1$-measurable functions $f : \mathbb{R} \to \mathbb{R}$. Indeed take the Vitali-Set V from Theorem 1.86. Then $f := \chi_V$ is not measurable.

Theorem 2.13. *Let $f_k : \Omega \to \overline{\mathbb{R}}$ μ-measurable, $k \in \mathbb{N}$. Then*

$$F_1(x) := \inf_k f_k(x)$$

$$F_2(x) := \sup_k f_k(x)$$

$$F_3(x) := \liminf_{k \to \infty} f_k(x)$$

$$F_4(x) := \limsup_{k \to \infty} f_k(x)$$

are all μ-measurable.

[8] observe that product is pointwise defined since $f(x), g(x) \neq \pm\infty$, so we have no problem with $0 \cdot \infty$

Proof. Observe that

$$(\inf_{k\in\mathbb{N}} f_k)^{-1}([-\infty, a)) = \bigcup_{k=1}^{\infty} f_k^{-1}([-\infty, a)),$$

$$(\inf_{k\in\mathbb{N}} f_k)^{-1}((a, \infty]) = \bigcap_{k=1}^{\infty} f_k^{-1}((a, \infty]),$$

so F_1 is measurable. Since

$$\sup_k f_k(x) = -\inf_k(-f_k)(x)$$

we see that F_2 is measurable. For F_3 and F_4 observe

$$\liminf_{k\to\infty} f_k(x) = \lim_{\ell\to\infty}\left(\inf_{k\geq\ell} f_k(x)\right) = \sup_{\ell\in\mathbb{N}}\left(\inf_{k\geq\ell} f_k(x)\right),$$

and similarly

$$\limsup_{k\to\infty} f_k = \inf_{\ell\in\mathbb{N}}\left(\sup_{k\geq\ell} f_k\right).$$

□

The following theorem says that any μ-measurable function can be approximated by *simple functions*. Simple functions (sometimes called *step functions*) are functions

$$f(x) = \sum_{i=1}^{n} \lambda_i \chi_{A_i}$$

where $\lambda_i \neq \lambda_j$ for $i \neq j$ and A_i are disjoint measurable sets.

Theorem 2.14. *Let $f : (X, d) \to [0, \infty]$ be a μ-measurable function. Then there are μ-measurable sets $A_k \subset X$, for all $k \in \mathbb{N}$ such that*[9]

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

Proof. Let

$$A_1 := \{x \in X : \quad f(x) \geq 1\} = f^{-1}([1, \infty]).$$

Since f is μ-measurable, A_1 is μ-measurable. Now define inductively the μ-measurable sets

$$A_k := \left\{x \in X : \quad f(x) \geq \frac{1}{k} + \sum_{j=1}^{k-1} \frac{1}{j}\chi_{A_j}(x)\right\} = \left(f - \sum_{j=1}^{k-1} \frac{1}{j}\chi_{A_j}\right)^{-1}([\frac{1}{k}, \infty]), \quad k = 2, 3, \ldots.$$

Now let $x \in X$. We first show

$$f(x) \geq \sum_{k=1}^{\infty} \frac{1}{k}\chi_{A_k}(x) \quad \forall x \in X. \tag{2.1}$$

[9] observe that the sum of the right hand side either converges absolutely or is infinite

If $\underline{\#\{k \in \mathbb{N} : \ x \in A_k\} = \infty}$ then $f(x) \geq \sum_{j=1}^{k} \frac{1}{j}\chi_{A_j}(x)$ for infinitely many k, and thus

$$f(x) \geq \sum_{k=1}^{\infty} \frac{1}{k}\chi_{A_k}(x).$$

If $\underline{1 \leq \#\{k \in \mathbb{N} : \quad x \in A_k\} < \infty}$ then set $k_0 := \max\{k \in \mathbb{N} : \quad x \in A_k\} \in \mathbb{N}$. Since $x \in A_{k_0}$ and $x \notin A_k$ for $k > k_0$ we have

$$f(x) \geq \frac{1}{k_0} + \sum_{j=1}^{k_0-1} \frac{1}{j}\chi_{A_j}(x) = \frac{1}{k_0}\underbrace{\chi_{A_{k_0}}(x)}_{=1} + \sum_{j=1}^{k_0-1} \frac{1}{j}\chi_{A_j}(x) + \sum_{j=k_0}^{\infty} \frac{1}{j}\underbrace{\chi_{A_j}(x)}_{=0} = \sum_{j=1}^{\infty} \frac{1}{j}\chi_{A_j}(x).$$

If $\underline{\{k \in \mathbb{N} : \quad x \in A_k\} = \emptyset}$ then

$$f(x) \geq 0 = \sum_{k=1}^{\infty} \frac{1}{k}\underbrace{\chi_{A_k}(x)}_{=0}.$$

(2.1) is now established.

We can conclude once we show that also

$$f(x) \leq \sum_{k=1}^{\infty} \frac{1}{k}\chi_{A_k}(x) \quad \forall x \in X. \tag{2.2}$$

Let $x \in X$. If $\underline{f(x) = \infty}$ then $x \in A_k$ for all k, so

$$f(x) = \infty = \sum_{k=1}^{\infty} \frac{1}{k} = \sum_{k=1}^{\infty} \frac{1}{k}\underbrace{\chi_{A_k}(x)}_{=1}.$$

If $\underline{f(x) = 0}$ (2.2) is obvious since the right-hand side is nonnegative.

If $\underline{0 < f(x) < \infty}$, then we may assume that $x \notin A_k$ for infinitely many $k \in \mathbb{N}$. Indeed, otherwise there exists some k_0 such that $x \in A_k$ for all $k \geq k_0$ and thus

$$\sum_{k=1}^{\infty} \frac{1}{k}\chi_{A_k}(x) \geq \sum_{k=k_0}^{\infty} \frac{1}{k}\underbrace{\chi_{A_k}(x)}_{=1} = \sum_{k=k_0}^{\infty} \frac{1}{k} = \infty \geq f(x),$$

and (2.2) is established.

So the only remaining case is that $f(x) \in (0, \infty)$ and $x \notin A_k$ for infinitely many $k \in \mathbb{N}$. Now take an increasing sequence $k_i \to \infty$ with $x \notin A_{k_i}$ for each i. With the definition of A_k we then have

$$f(x) \leq \frac{1}{k_i} + \sum_{j=1}^{k_i-1} \frac{1}{j}\chi_{A_j}(x) \quad \forall i.$$

That is,

$$f(x) \leq \limsup_{k\to\infty} \left(\frac{1}{k} + \sum_{j=1}^{k-1} \frac{1}{j}\chi_{A_j}(x) \right) = \sum_{j=1}^{\infty} \frac{1}{j}\chi_{A_j}(x)$$

(2.2) is now established. □

Another way to approximate by simple functions is

Exercise 2.15. *Let $f : X \to [0, \infty)$ be μ-measurable and bounded*

$$\sup_{x \in X} f(x) < \infty$$

Set

$$s_n(x) := \begin{cases} n & \text{if } f(x) \geq n \\ \frac{k}{2^k} & \text{if } \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n} \leq n. \end{cases}$$

Show that s_n converges uniformly to f.

So we can approximate nonnegative measurable function $f : X \to [0, \infty]$ by *step functions*,

$$f(x) = \lim_{L \to \infty} \sum_{k=1}^{L} \frac{1}{k} \chi_{A_k}(x).$$

- Observe that this approximation is monotone increasing.
- Splitting a μ-measurable function $f : X \to \bar{\mathbb{R}}$ into $f = f_+ - f_-$ and we can approximate any measurable function by step-functions.

Exercise 2.16. *Let X be a metric space. A function $f : X \to \bar{R}$ is* upper semicontinuous

$$\limsup_{x \to x_0} f(x) \leq f(x_0) \quad \forall \text{for all } x_0 \in X$$

A function is lower semicontinuous *if*

$$\liminf_{x \to x_0} f(x) \geq f(x_0) \quad \text{for all } x_0 \in X.$$

(1) *Show that upper semicontinuity is equivalent to saying*

$$f^{-1}([a, \infty]) \text{is closed} \quad \forall a \in \mathbb{R}.$$

and this in turn is equivalent to

$$f^{-1}([-\infty, a)) \text{is open} \quad \forall a \in \mathbb{R}.$$

(2) *Show that any step function $f(x) = \sum_{k=1}^{K} \chi_{A_k}$ is upper semicontinuous if A_k are closed*
(3) *Show that any step function $f(x) = \sum_{k=1}^{K} \chi_{A_k}$ is lower semicontinuous if A_k are open*
(4) *Show that if μ is Borel measure then any upper or lower semicontinuous function is μ-measurable.*

3. Integration

In integration theory we will follow the notation that $0 \cdot \infty = 0$.

Definition 3.1. Let X be a metric space and μ a measure.

- Let $f : X \to [0, \infty)$ be a (nonnegative!) simple function of the form
$$f(x) = \sum_{i=1}^{n} \alpha_i \chi_{A_i}(x)$$
where $\alpha_i \neq \alpha_j$ for $i \neq j$ and A_i are pairwise disjoint μ-measurable sets.
For any μ-measurable set $\Omega \subset X$ we define
$$\int_\Omega f \, d\mu := \sum_{i=1}^{n} \alpha_i \mu(A_i \cap \Omega).$$
(Recall that if $\mu(A) = \infty$ we still, by assumption, set $0 \cdot \mu(A) = 0$. And if $\alpha = \infty$ but $\mu(A_i \cap \Omega) = 0$ then still $\alpha\mu(A_i \cap \Omega) := 0$.)
- It is easy to see that if f is a simple function represented by two different sums
$$f(x) = \sum_{i=1}^{n} \alpha_i \chi_{A_i}(x) = \sum_{j=1}^{m} \beta_j \chi_{B_j}(x)$$
then
$$\sum_{i=1}^{n} \alpha_i \mu(A_i \cap \Omega) = \sum_{j=1}^{m} \beta_j \mu(B_i \cap \Omega),$$
so the integral notion is well-defined.
- Clearly if $f(x) \leq g(x)$ for all x and both f and g are simple functions, then
$$\int_\Omega f d\mu \leq \int_\Omega g d\mu.$$
- Thus, if $f : X \to [0, \infty)$ is a (nonnegative!) simple function as above then
$$\int_\Omega f \, d\mu = \sup_s \int_\Omega s d\mu$$
where the supremum is taken over all simple functions s such that $0 \leq s \leq f$.
- So we can extend the above notion to all nonnegative measurable functions. If $g : X \to [0, \infty]$ is a μ-measurable function and $\Omega \subset X$ is μ-measurable we define the *Lebesgue integral* of g over Ω by
$$\int_\Omega g d\mu := \sup_s \int_\Omega s d\mu$$
where the supremum is taken over all simple functions s such that $0 \leq s \leq g$ a.e.

It is worth to note that this defines the Lebesgue integral for all nonnegative measurable functions (it may just be infinite). This is reminiscient of the lower Darboux sum approximation of the Riemann integral, the main difference here being that we use step functions not on blocks but on measurable sets.

We observe the following properties

Exercise 3.2.
- *If* $0 \leq f \leq g$ *then* $\int_E f d\mu \leq \int_E g d\mu$
- *If* $A \subset B$ *and* $f \geq 0$ *then* $\int_A f d\mu \leq \int_B f d\mu$.
- *If* $f(x) = 0$ *for all* $x \in E$ *then* $\int_E f d\mu = 0$, *even if* $\mu(E) = \infty$.
- *If* $\mu(E) = 0$ *then* $\int_E f d\mu = 0$ *even if* $f(x) = \infty$ *for all* $x \in E$.
- *if* $f \geq 0$ *then* $\int_E d\mu = \int_X \chi_E f \, d\mu$.

The following is an easy observation

Exercise 3.3. *Let* g *and* g *be nonnegative,* μ*-measurable simple functions on* X *and* $\lambda \geq 0$ *a constant. Then*

$$\int_X (f+g) d\mu = \int_X f d\mu + \int_X g d\mu$$

and

$$\int_X \lambda f d\mu = \lambda \int_X f d\mu.$$

It is easy to conclude homogeneity

Exercise 3.4 (homogeneity of the integral)**.** *Let* $f : X \to [0, \infty]$ *be* μ*-measurable and* $\lambda \geq 0$. *Show, without using the Lebesgue monotonce convergence theorem, that*

$$\int_X \lambda f d\mu = \lambda \int_X f d\mu.$$

So we have reason to believe the integral is linear, i.e.

$$\int_X \lambda f + g d\mu = \lambda \int_X f d\mu + \int_X g d\mu.$$

We will prove this below, Corollary 3.7, with a technique that can be generalized to one of the most fundamental theorems of integration theory, the *Lebesgue monotone convergence theorem*

Theorem 3.5 (Lebesgue monotone convergence theorem)**.** *Let* $(f_n)_n$ *be a sequence of* μ*-measurable functions on* X *such that*

- $0 \leq f_1 \leq f_2 \leq \ldots \leq \infty$ *for almost every*[10] $x \in X$
- $\lim_{n\to\infty} f_n(x) = f(x) \in [0, \infty]$ *for almost every* $x \in X$.

Then f *is* μ*-measurable and*

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu.$$

Before we can prove Theorem 3.5 we need the following

[10] recall Definition 1.35

Lemma 3.6. *Let f be a nonnegative μ-measurable simple functions on X. Denote by Σ the μ-measurable sets. The map $\tilde{\nu} : \Sigma \to [0, \infty]$ given by*

$$\tilde{\nu}(E) := \int_E f \, d\mu \quad \forall E \ \mu\text{-measurable}$$

is a premeasure, cf. Definition 1.39.

Its Carathéodory-Hahn Extension ν of $\tilde{\nu}$, cf. Theorem 1.42, in symbols

$$\nu := f \llcorner \mu,$$

is called the concatenation of f and μ.

In particular any μ-measurable set E is $f \llcorner \mu$-measurable.

Proof. Since Σ is a σ-Algebra, it is is in particular an algebra.

To confirm that $\tilde{\nu}$ is a premeasure, let $A = \bigcup_{k=1}^{\infty} A_k$, where $(A_k)_{k=1}^{\infty}$ are pairwise disjoint μ-measurable functions.

We need to show

$$\tilde{\nu}(A) = \sum_{k=1}^{\infty} \tilde{\nu}(A_k).$$

Let f be given as

$$f = \sum_{i=1}^{n} \alpha_i \chi_{B_i}.$$

Then

$$\tilde{\nu}(A_k) = \int_{A_k} f d\mu = \sum_{i=1}^{n} \alpha_i \mu(B_i \cap A_k).$$

For each $i \in \{1, \ldots, n\}$, the collection $(A_k \cap B_i)_{k \in \mathbb{N}}$ consists of measurable and pairwise disjoint sets. By Theorem 1.33(1) we find

$$\sum_{k=1}^{\infty} \tilde{\nu}(A_k) = \sum_{k=1}^{\infty} \sum_{i=1}^{n} \alpha_i \mu(B_i \cap A_k) = \sum_{i=1}^{n} \alpha_i \mu(B_i \cap A) = \int_A f d\mu = \tilde{\nu}(A).$$

This proves that $\tilde{\nu}$ is a pre-measure. □

Proof of Theorem 3.5. f is measurable in view of Theorem 2.13. We can replace the "almost every" in the assumption by "every", by changing f_i and f on zero sets (observe that a countable union of zeroset is a zeroset).

By monotonicity of $\int d\mu$ the sequence

$$\alpha_n := \int_X f_n d\mu \in [0, \infty]$$

is monotonically increasing and hence it has a limit $\alpha = \lim_{n\to\infty} \int_X f_n d\mu \in [0,\infty]$. Since $f_n(x) \le f(x)$ everywhere we have

$$\alpha = \lim_{n\to\infty} \int_X f_n d\mu \le \int_X f d\mu. \tag{3.1}$$

To conclude we need to show "$\ge$" in the above inequality.

Let $0 \le s \le f$ be a simple function. For a fixed constant $0 < c < 1$ we define

$$E_n := \{x \in X : \quad f_n(x) \ge c\, s(x)\}, \quad n \in \mathbb{N}.$$

Since f_n is increasing, $E_1 \subset E_2 \subset \dots$. Since $c < 1$ and $f_n(x) \xrightarrow{n\to\infty} f(x)$ for each x we have that $\bigcup_n E_n = X$.

Since the E_n are μ-measurable sets ($f_n(\cdot) - cs(\cdot)$ is measurable!) we have

$$\int_X f_n d\mu \ge \int_{E_n} f_n d\mu \ge c \int_{E_n} s d\mu = c s \llcorner \mu(E_n). \tag{3.2}$$

Now we have by Lemma 3.6 that $s\llcorner\mu$ is a measure and each E_n is $s\llcorner\mu$-measurable. Since $E_n \subset E_{n+1}$ we find by Theorem 1.33

$$s\llcorner\mu(E_n) \xrightarrow{n\to\infty} s\llcorner\mu(\bigcup_{n=1}^{\infty} E_n) = s\llcorner\mu(X) = \int_X s d\mu.$$

Taking the limit in (3.2) we have

$$\alpha = \lim_{n\to\infty} \int_X f_n d\mu \ge c \int_X s d\mu.$$

Recall that this holds whenever s is a simple function with $s \le f$. Taking the supremum over such s we conclude

$$\alpha \ge c \int_X f d\mu.$$

This holds for any $c \in (0,1)$, taking the limit as $c \to 1$ we find

$$\alpha \ge \int_X f d\mu.$$

Combining this with (3.1) we conclude. □

Corollary 3.7 (Linearity of the integral)**.** *Let $f, g : X \to [0,\infty]$ μ-measurable and $\lambda, \sigma \ge 0$. Then*

$$\int_X \lambda f + \sigma g d\mu = \lambda \int_X f d\mu + \sigma \int_X g d\mu.$$

Proof. We only show

$$\int_X f + g d\mu = \int_X f d\mu + \int_X g d\mu.$$

the general case then follows from this and Exercise 3.4.

Let s_i and t_i be *monotonically increasing* sequence of simple nonnegative functions such that $s_i \xrightarrow{i\to\infty} f$ and $t_i \xrightarrow{i\to\infty} f_2$ (existence follows from Theorem 2.14). Then $s_i + t_i \xrightarrow{i\to\infty} f + g$.

From Lebesgue monotone convergence theorem, Theorem 3.5 and Exercise 3.3

$$\int_X f + g d\mu = \lim_{i\to\infty} \int_X (s_i + t_i) d\mu = \lim_{i\to\infty} \int_X s_i d\mu + \int_X t_i d\mu = \int_X f d\mu + \int_X g d\mu$$

□

Indeed, we have more than Corollary 3.7 we can take out infinite sums (recall that all functions considered are *nonnegative*, so series always absolute converge (or are infinity))

Corollary 3.8. *Let $f_n : X \to [0, \infty]$ be a sequence of μ-measurable functions and set for $x \in X$*

$$f(x) := \sum_{n=1}^{\infty} f_n(x) \in [0, \infty].$$

Then f is measurable and we have

$$\int_X f d\mu = \sum_{i=1}^{\infty} \int_X f_n d\mu.$$

Proof. By induction and Corollary 3.7 we have

$$\sum_{i=1}^{N} \int_X f_n d\mu = \int_X \sum_{i=1}^{N} f_n d\mu$$

Now observe that f_n are nonnegative so $\sum_{i=1}^{N} f_n$ is a monotonce sequence in N with $\lim_{N\to\infty} \sum_{i=1}^{N} f_n = \sum_{i=1}^{\infty} f_n$. So by Lebesgue monotone convergence we conclude that

$$\sum_{i=1}^{\infty} \int_X f_n d\mu = \lim_{N\to\infty} \sum_{i=1}^{N} \int_X f_n d\mu = \lim_{N\to\infty} \int_X \sum_{i=1}^{N} f_n d\mu = \int_X \sum_{i=1}^{\infty} f_n d\mu.$$

□

Then next important consequence is called *Fatou's lemma*, which essentially tells us *lower semi-continuity of the integral under pointwise convergence*

Corollary 3.9 (Fatou's Lemma)**.** *Let $f_n : X \to [0, \infty]$ be a sequence of μ-measurable functions, then*

$$\int_X \left(\liminf_{n\to\infty} f_n \right) d\mu \leq \liminf_{n\to\infty} \int_X f_n d\mu.$$

Proof. Let $g_n := \inf\{f_n, f_{n+1}, \ldots\}$. Then $g_n \leq f_n$ and hence

$$\int_X g_n d\mu \leq \int_X f_n d\mu$$

Since $0 \leq g_1 \leq g_2 \leq \ldots$, all the functions g_n are measurable, Theorem 2.13. Taking the $\liminf$ in the above inequality we have

$$\liminf_{n\to\infty} \int_X g_n d\mu \leq \liminf_{n\to\infty} \int_X f_n d\mu.$$

By Theorem 3.5

$$\liminf_{n\to\infty} \int_X g_n d\mu = \int_X \lim_{n\to\infty} g_n d\mu.$$

We conclude since $\lim_{n\to\infty} g_n = \liminf_{k\to\infty} f_k$. □

Exercise 3.10. *Construct a sequence of measurable functions $f_n : \mathbb{R} \to [0, \infty]$ such that the inequality in Fatou's Lemma, Corollary 3.9, is sharp.*

Hint: the conditions in Theorem 3.26 below must be violated.

We can also extend Lemma 3.6

Theorem 3.11. *Let f be a nonnegative μ-measurable function on X. Denote by Σ the μ-measurable sets. The map $\tilde{\nu} : \Sigma \to [0, \infty]$ given by*

$$\tilde{\nu}(E) := \int_E f\, d\mu \quad \forall E \ \mu\text{-measurable}$$

is a premeasure, cf. Definition 1.39.

As before, we denote by $f \llcorner \mu$ its Carathéodory-Hahn Extension (concatenation of f and μ).

Moreover for every measurable function $g : X \to [0, \infty]$ we have

$$\int_X g\, d\,(f \llcorner \mu) = \int_X g f\, d\mu. \tag{3.3}$$

Proof. As in the proof of Lemma 3.6 we only need to show that whenever $E_1, E_2, \ldots \in \Sigma$ are disjoint μ-measurable sets and $E = \bigcup_{n=1}^{\infty} E_n$ then

$$f \llcorner \mu(E) = \sum_n f \llcorner \mu(E_n).$$

Now observe that

$$\chi_E f(x) = \sum_{n=1}^{\infty} \chi_{E_n} f(x) \quad \forall x \in X.$$

Thus

$$(f \llcorner \mu)(E) = \int_X \chi_E f d\mu = \int_X \sum_{n=1}^{\infty} \chi_{E_n} f.$$

By Lebesgue Monotone Convergence theorem, Theorem 3.5, we can take out the sum, and have

$$(f \llcorner \mu)(E) = \sum_{n=1}^{\infty} \int_X \chi_{E_n} f = \sum_{n=1}^{\infty} (f \llcorner \mu)(E_n).$$

For (3.3) observe that it holds by definition if $g = \chi_E$ whenever E is μ-measurable. Since any nonnegative μ-measurable function can be approximated by simple functions, Theorem 2.14, we conclude the (3.3) again from Lebesgue Monotone Convergence theorem, Theorem 3.5. □

3.1. **L^p-spaces and Lebesgue dominated convergence theorem.** We now want to discuss not only nonnegative measurable functions $f : X \to [0, \infty]$, but want to integrate general measurable functions $f : X \to \bar{\mathbb{R}}$.

Definition 3.12. Let (X, Σ, μ) be a measure space and $f : X \to \bar{\mathbb{R}}$.

- We say that f is *μ-integrable* if $|f| : X \to [0, \infty]$[11] if
$$\|f\|_{L^1(X,\mu)} := \int_X |f| < \infty.$$
If f is integrable we write
$$\int_X f d\mu := \int_X f_+ d\mu - \int_X f_- d\mu$$
where as usual $f_+ := \max\{f, 0\}$ and $f_- = -\min\{f, 0\}$. Observe that f integrable assures us that $\int_X f_+ d\mu < \infty$ and $\int_X f_- d\mu < \infty$ so there is no issue with $\infty - \infty$.
- More generally for $p \in (1, \infty)$ set
$$\|f\|_{L^p(X,\mu)} := \left(\int_X |f(x)|^p \, dx \right)^{\frac{1}{p}}.$$
If $\|f\|_{L^p(X,\mu)} < \infty$ we say that $f \in L^p(X, \mu)$.
- We also have the *L^2-scalar product* or the *L^2-pairing*
$$\langle f, g \rangle := \int_X f(x) \, g(x) \, d\mu$$
Let us remark (although we make no substantial use of this) that for *complex* functions $f, g : X \to \mathbb{C}$ the scalar product is
$$\langle f, g \rangle := \int_X f(x) \, \overline{g(x)} \, d\mu$$
where $\overline{g}$ is the complex conjugation.
Clearly
$$\langle f, f \rangle = \|f\|^2_{L^2(X,\mu)}.$$
- For $p = \infty$ we define the *essential supremum* (often still denoted by sup)
$$\|f\|_{L^\infty(X,d\mu)} := \sup_X |f| := \inf \{\Lambda \in [0, \infty] : \quad \mu\{x \in X : |f(x)| > \Lambda\} = 0\}$$
In words, $\sup_X |f|$ is the smallest number Λ such that the superlevel set of $\{f > \Lambda\}$ has zero measure.
- If $X \subset \mathbb{R}^n$ and μ is the Lebesgue measure we drop the μ and simpy write $L^p(X)$.

Exercise 3.13. *Let $p \in [1, \infty]$, $f, g \in L^p(X, \mu)$ and $\lambda \in \mathbb{R}$. Show that*

(1) (Homogeneity) $\|\lambda f\|_{L^p(X,\mu)} = |\lambda| \|f\|_{L^p(X,\mu)}$,
(2) (Minkowski-inequality) $\|f + g\|_{L^p(X,\mu)} \leq \|f\|_{L^p(X,\mu)} + \|g\|_{L^p(X,\mu)}$,

[11] which is μ-measurable

(3) *(Hölder-inequality)* $\|fg\|_{L^p(X,\mu)} \leq \|f\|_{L^q(X,\mu)}\|g\|_{L^r(X,\mu)}$, *where* $q, r \in [1,\infty]$ *are such that*

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

(Here $\frac{1}{\infty} := 0$*).*

In particular,

$$\|f\|_{L^p(X,\mu)} \leq \mu(X)^{\frac{1}{r}}\|f\|_{L^q(X,\mu)}$$

(4) *(generalized Hölder's inequality)* $\|f_1 \cdot \ldots \cdot f_n\|_{L^p(X,\mu)} \leq \|f_1\|_{L^{q_1}(X,\mu)} \cdot \ldots \|f_n\|_{L^{q_n}(X,\mu)}$, *where* $q_i, p \in [1,\infty]$ *are such that*

$$\frac{1}{p} = \sum_{i=1}^{n} \frac{1}{q_i}$$

(5) *(Jensen) Let* $\mu(X) < \infty$, $f : X \to (-\infty, \infty)$ *and* $\varphi : \mathbb{R} \to \mathbb{R}$ *convex, then*

$$\varphi\left(\mu(X)^{-1}\int_X f(x)d\mu\right) \leq \mu(X)^{-1}\int_X \varphi(f(x))d\mu$$

Hint: Adv.Calc or Wikipedia

Exercise 3.14. *Let* $p \in [1,\infty)$. *Assume* $(f_k)_{k\in\mathbb{N}} \subset L^p(X,\mu)$ *with*

$$\sup_{k\in\mathbb{N}} \|f_k\|_{L^p(X,\mu)} < \infty.$$

Assume moreover there exists $f : X \to \bar{\mathbb{R}}$ *such that* $f(x) = \lim_{k\to\infty} f_k(x)$ *for* μ*-a.e.* $x \in X$.

Using Fatou's Lemma, Corollary 3.9, show that $f \in L^p(X,\mu)$ *and we have*

$$\|f\|_{L^p(X,\mu)} \leq \liminf_{k\to\infty} \|f_k\|_{L^p(X,\mu)}.$$

Exercise 3.15. (1) *If* f *is continuous and* μ *is the Lebesgue measure on an open set* Ω *show that the essential supremum coincides with the usual supremum. (Show also this is not the case if* Ω *contains e.g. isolated points)*

(2) *Give an example of measurable* f *where essential supremum and supremum is not the same.*

Exercise 3.16. *Assume* $\mu(X) < \infty$ *and* $f : X \to \mathbb{R}$ *is* μ*-measurable. Show that*

$$\lim_{p\to\infty}\left(\mu(X)^{-1}\int_X |f|^p d\mu\right)^{\frac{1}{p}} = \|f\|_{L^\infty(X)}.$$

Exercise 3.17. *We have*

(1)

$$|f(x)| \leq \|f\|_{L^\infty(\mu)} \quad \mu\text{-a.e. } x.$$

(2) *If* $|f(x)| \leq \Lambda$ *for* μ*-a.e.* x, *then* $\|f\|_{L^\infty} \leq \Lambda$.

Exercise 3.18. *Let $1 \le p \le r \le q \le \infty$ and assume $f \in L^p(X)$ and $f \in L^q(X)$. Then $f \in L^r(X)$ and we have*

$$\|f\|^r_{L^r(X)} \le \|f\|^q_{L^q(X)} + \|f\|^p_{L^p(X)}.$$

Hint: Let $A := \{x : f(x) \le 1\}$. Then $f(x) = f(x)\chi_A(x) + f(x)\chi_{X\setminus A}(x)$

So the collection of μ-measurable functions $f : X \to \bar{\mathbb{R}}$ such that $f \in L^p(X, \mu)$ is a *linear space*: if $f, g : X \to \bar{\mathbb{R}}$ satisfy $\|f\|_{L^p(X,\mu)}$, $\|g\|_{L^p(X,\mu)} < \infty$ then for any $\lambda_1, \lambda_2 \in \mathbb{R}$ we have $\|\lambda_1 f + \lambda_2 g\|_{L^p(X,\mu)} < \infty$.

$\|\cdot\|_{L^p(X,\mu)}$ is a *pseudonorm* on this linear space. Recall that for a linear space L a map $\|\cdot\| : L \to [0, \infty)$ is a *norm* iff

(1) $\|f\| = 0$ if and only if $f = 0$
(2) $\|f + g\| \le \|f\| + \|g\|$ for all $f, g \in L$
(3) $\|\lambda f\| = |\lambda|\|f\|$ for all $f \in L$, $\lambda \in \mathbb{R}$.

If the first property (1) fails, i.e. if there are $f \in L$ such that $\|f\| = 0$ but $f \ne 0$, then $\|\cdot\|$ is a *pseudonorm*.

We don't want to work with pseudonorms. (Don't worry, this looks more complicated than it is). We will instead work with classes of functions. From now on we say that two μ-measurable functions $f, g : X \to \bar{\mathbb{R}}$ are equal

$$f = g \quad \Leftrightarrow f(x) = g(x) \quad \mu\text{-a.e..}$$

(check: this is an equivalence relation).

In particular we will often define a function only in $X\setminus N$ where $\mu(N) = 0$.

The integral does not see differences on zero-sets.

Exercise 3.19. *Let $f, g : X \to \bar{\mathbb{R}}$ be μ-measurable and $f = g$ in the above sense. Then*

$$\int_X f d\mu = \int_X g d\mu$$

in the following sense:

- *$f \in L^1(X)$ if and only if $g \in L^1(X)$.*
- *If $f \in L^1(X)$ (or equivalently $g \in L^1(X)$) then*

$$\int_X f d\mu = \int_X g d\mu.$$

Then we define

$$L^p(X, \mu) := \left\{ f : X \to \bar{\mathbb{R}} : \mu\text{-measurable and } \|f\|_{L^p(X,\mu)} < \infty \right\} / =$$

That is: an element in $L^p(X, \mu)$ is not a single function, but a class of functions (two functions belong to the same class if they only differ on a zero set). We often brush over

this fact by saying f is a function (but actually meaning f's class). We don't even need to define a function everywhere, defining it outside a zero-set is enough. A specific function f in a class $[f]$ is called a *representative*.

While most of the time we don't bother whether we talk about a specific representative f or its class $[f]$ – sometimes we do care. For example if f is a continuous representative of its class $[f]$. The good news is that there cannot be two different continuous representatives of a class $[f]$, Exercise 4.51 for the Lebesgue measure.

Example 3.20.

- The function
$$f(x) = \frac{1}{|x|}$$
is $\mathcal{L}^n$-measurable in $\mathbb{R}^n$ and L^p-integrable in $B(0,1) \subset \mathbb{R}^n$ if $p < n$.

  Clearly the $f(0)$ is not defined, but $\{0\}$ is a $\mathcal{L}^n$-zeroset, so it does not matter. A proper representative of $\frac{1}{|x|}$ could be
$$g(x) := \begin{cases} \frac{1}{|x|} & x \neq 0 \\ 0 & x = 0 \end{cases}$$
or
$$g(x) := \begin{cases} \frac{1}{|x|} & x \neq 0 \\ \infty & x = 0 \end{cases}$$
or
$$g(x) := \begin{cases} \frac{1}{|x|} & x \neq 0 \\ \infty & x = 0 \\ 736 & x = (7, 0, 0, \dots, 0). \end{cases}$$
- Let $f \in C^0(\Omega)$ for some bounded open set Ω. Take any set A with $\mu(A) = 0$, and set
$$g := f + \chi_A.$$
Then $g = f$ μ-a.e. – that is not every representative of a class containing a continuous representative is continuous. Indeed, quite the opposite: there is at most one continuous representative in each class if μ is the Lebesgue measure, Exercise 4.51.

Lemma 3.21. *Assume $f \in L^p(X, \mu)$ for some $p \in [1, \infty]$. Then (for any representative of f), $|f(x)| < \infty$ μ-a.e.*

Proof. Set $A := \{x : |f(x)| = \infty\}$. Then we have (assume $p < \infty$, exercise for $p = \infty$!)
$$\mu(A) \cdot \infty = \int_A |f(x)|^p d\mu \leq \int_X |f(x)|^p d\mu = \|f\|^p_{L^p(X,\mu)} < \infty.$$
The only way this is possible is if $\mu(A) = 0$. □

Lemma 3.22. *(1) Suppose $f : X \to [0, \infty]$ is measurable and E is measurable. Then $\int_E f d\mu = 0$ if and only if $f = 0$ a.e. in E.*

(2) Suppose $f \in L^1(X,\mu)$ and $\int_E f d\mu = 0$ for all E measurable. Then $f = 0$ a.e. in X

Proof. (1) Set $A_n := \{x \in E : f(x) \geq \frac{1}{n}\}$. Then[12]

$$\frac{1}{n}\mu(A_n) = \int_{A_n} \frac{1}{n} d\mu \leq \int_{A_n} f d\mu \overset{f\geq 0}{\leq} \int_E f d\mu = 0.$$

Thus $\mu(A_n) = 0$ for all $n \in \mathbb{N}$. Therefore the set $\{x \in E : f(x) > 0\} = \bigcup_{n=1}^{\infty} A_n$ has measure zero.

(2) Define $E = \{x \in X : f(x) \geq 0\}$. Then

$$0 = \int_E f d\mu.$$

By part (1) we conclude that $f = 0$ a.e. in E. Arguing the same way for $X \backslash E = \{f(x) < 0\}$ we conclude.

□

Exercise 3.23. *For the classes of functions $f \in L^p(X,\mu)$, $\|\cdot\|_{L^p}$ is indeed a norm. (Hint: Use Lemma 3.22)*

Definition 3.24. Let $p \in [1,\infty]$ and $f : X \to \mathbb{R}^N$ μ-measurable. We say $f \in L^p(X,\mu,\mathbb{R}^n)$ (meaning there is a class of which f is a representative) if and only if $|f| \in L^p(X,\mu)$. We can also define the integral componentwise

$$\int_X f \; d\mu = \left(\int_X f^1 \; d\mu, \ldots, \int_X f^n \; d\mu\right).$$

Lemma 3.25. *If $f \in L^1(X,\mu,\mathbb{R}^n)$ then*

$$\left|\int_X f \, d\mu\right| \leq \int_X |f| d\mu.$$

Proof. Since $\int_X f \, d\mu \in \mathbb{R}^N$ there exists[13] a vector $v \in \mathbb{R}^N$ with $|v| = 1$ such that

$$|\int_X f| = \langle v, \int_X f d\mu\rangle = \sum_{i=1}^n v^i \int_X f^i d\mu.$$

So we have

$$|\int_X f| = \int_X \sum_{i=1}^n v^i f^i d\mu = \int_X \langle v, f\rangle d\mu$$

Now pointwise (Cauchy-Schwarz)

$$|\langle v, f\rangle| \leq |f|,$$

which implies the claim. □

[12]The first inequality will later be called the Chebycheff inequality, Lemma 3.49

[13]if $w \in \mathbb{R}^n$ is a vector then $v := \frac{w}{|w|}$ (if $|w| \neq 0$) or v with $|v| = 1$ if $w = 0$) satisfies $\langle v, w\rangle = |w|$

Theorem 3.26 (Lebesgue dominated convergence theorem)**.** *Suppose $(f_n)_{n\in\mathbb{N}}$ is a sequence of μ-measurable functions $f_n : X \to \overline{\mathbb{R}}$ and $f : X \to \overline{\mathbb{R}}$ a function with $f(x) = \lim_{n\to\infty} f_n(x)$ for μ-almost every $x \in X$*[14]*.*

If there exists $g \in L^1(X, \mu)$ such that

$$|f_n(x)| \le g(x) \quad \mu\text{-a.e. } x \in X,\ \forall n \in \mathbb{N}$$

then $f \in L^1(\mu)$ and

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

in particular

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu.$$

Proof. By setting f_n, f, g to be zero on a zeroset we can assume that all the "almost everywhere" above can be replaced by everywhere.

f as a pointwise limit of measurable functions f_n is measurable, and clearly $|f| \le g$ everywhere in X. Thus $\int_X |f| d\mu \le \int_X g d\mu < \infty$, i.e. $f \in L^1(X, \mu)$.

Moreover we have $2g - |f - f_n| \ge 0$, so we can apply Fatou's lemma, Corollary 3.9,

$$\begin{aligned}
\int_X 2g d\mu &= \int_X \lim_{n\to\infty} (2g - |f - f_n|)\, d\mu \\
&\le \liminf_{n\to\infty} \int_X (2g - |f - f_n|)\, d\mu \\
&= \int_X 2g d\mu + \liminf_{n\to\infty} \left(- \int_X |f - f_n| d\mu \right) \\
&= \int_X 2g d\mu - \limsup_{n\to\infty} \left(\int_X |f - f_n| d\mu \right)
\end{aligned}$$

The integral $\int_X 2g d\mu < \infty$. Subtracting it form both sides of the above inequality we find

$$\limsup_{n\to\infty} \left(\int_X |f - f_n| d\mu \right) \le 0.$$

Since the integral is nonnegative we conclude

$$\lim_{n\to\infty} \|f - f_n\|_{L^1(X,\mu)} = \lim_{n\to\infty} \left(\int_X |f - f_n| d\mu \right) = 0.$$

The last claim follows from Lemma 3.25,

$$\left| \int_X f_n d\mu - \int_X f d\mu \right| \le \int_X |f_n - f|\, d\mu \xrightarrow{n\to\infty} 0.$$

□

The next theorem, while it is now not extremely difficult anymore, is one of the main advantages of the Lebesgue integral: it leads to *complete* L^p-spaces.

[14] observe without additional assumptions there is no hope of having $\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu$, Exercise 3.10

Theorem 3.27 (L^p is complete)**.** *Let $(f_k)_{k\in\mathbb{N}}$ be a Cauchy sequence in $L^p(X,\mu)$, that is*

$$\forall \varepsilon > 0 \quad \exists K \in \mathbb{N}: \quad \|f_k - f_\ell\|_{L^p} < \varepsilon \quad \forall k, \ell \geq K.$$

Then there exists $f \in L^p(X,\mu)$ such that

$$\|f_k - f\|_{L^p(X,\mu)} \xrightarrow{k\to\infty} 0.$$

Proof. By assumption, for each k there exists a number $\varphi(k)$ such that

$$\|f_\ell - f_j\|_{L^p} < 2^{-k} \quad \forall \ell, j \geq \varphi(k).$$

W.l.o.g. $\varphi(k) \leq \varphi(k+1)$.

We claim that

$$f(x) := \lim_{k\to\infty} f_{\varphi(k)}(x)$$

exists μ-a.e., and that this limit is actually an L^p-limit. To show that we write

$$f_{\varphi(k)} = \sum_{j=1}^{k-1} \left(f_{\varphi(j+1)} - f_{\varphi(j)}\right) + f_{\varphi(1)}.$$

Now define

$$G(x) := \sum_{j=1}^{\infty} \left|f_{\varphi(j+1)}(x) - f_{\varphi(j)}(x)\right|$$

As a limit of μ-measurable functions, G is μ-measurable.

Either by monotonce convergence theorem (if $p < \infty$ or by hand if $p = \infty$) we have

$$\|G\|_{L^p(X,\mu)} \leq \sum_{j=1}^{\infty} \|f_{\varphi(j+1)} - f_{\varphi(j)}\|_{L^p(X,\mu)} \leq \sum_{j=1}^{\infty} 2^{-j} < \infty.$$

That is $G \in L^p(X,\mu)$. Thus G is μ-a.e. finite, Lemma 3.21. Consequently, we have absolute convergence of the series

$$F(x) := \sum_{j=1}^{\infty} \left(f_{\varphi(j+1)}(x) - f_{\varphi(j)}(x)\right) \in \mathbb{R} \quad \mu\text{-a.e.},$$

so F is μ-measurable as limit of μ-measurable functions. And we also have

$$|F(x)| \leq G(x) \quad \mu\text{-a.e.}.$$

In particular, $\|F\|_{L^p(X,\mu)} < \infty$. Since the series is μ-a.e. absolute convergence we can set for μ-a.e. x

$$f(x) := \lim_{k\to\infty} f_{\varphi(k)}(x) = \lim_{k\to\infty} \sum_{j=1}^{k-1} \left(f_{\varphi(j+1)(x)} - f_{\varphi(j)}\right)(x) + f_{\varphi(1)}(x) = F(x)) + f_{\varphi(1)}(x),$$

which exists and is a finite number μ-a.e.. As a sum of L^p-functions $f \in L^p(X,\mu)$.

Using again the telescoping sum we have

$$|f - f_{\varphi(\ell)}| = \sum_{k=\ell}^{\infty} |f_{\varphi(k+1)} - f_{\varphi(k)}| \leq F \quad \forall \ell \in \mathbb{N}$$

Then we can apply dominated convergence, Theorem 3.26, and have

$$\lim_{\ell\to\infty} \|f - f_{\varphi(\ell)}\|_{L^p(X)} = \| \lim_{\ell\to\infty} \left(f - f_{\varphi(\ell)}\right) \|_{L^p(X)} = 0.$$

That is $(f_{\varphi(\ell)})_\ell$ converges in L^p to f.

This is only a subsequence, but with the Cauchy-condition we can conclude. Fix $\varepsilon > 0$, then there must be some K such that

$$\|f_k - f_\ell\|_{L^p} \leq \varepsilon \quad \forall k, \ell \geq K.$$

Now let $k \geq K$ and pick any ℓ such that $\varphi(\ell) \geq K$. Then

$$\|f_k - f\|_{L^p} \leq \|f_k - f_{\varphi(\ell)}\|_{L^p} + \|f_{\varphi(\ell)} - f\|_{L^p} \leq \varepsilon + \|f_{\varphi(\ell)} - f\|_{L^p}.$$

Taking $\ell \to \infty$ we find

$$\|f_k - f\|_{L^p} \leq \varepsilon \quad \forall k \geq K.$$

Thus f_k converges to f in L^p. □

Theorem 3.28. *Let X be a locally compact metric space and μ a Radon measure on X. Then the class of compactly supported continuous functions $C_c(X)$ is dense in $L^p(\mu)$ for all $1 \leq p < \infty$.*

Exercise 3.29. *Show that Theorem 3.28 is false for $p = \infty$ (hint: L^∞ is uniform convergence. What do we know about continuity under uniform convergence?)*

For the proof of Theorem 3.28 we need the following Lemma:

Lemma 3.30. *Let S be the class of finite, measurable, simple functions s on X such that*

$$\mu(\{x : s(x) \neq 0\}) < \infty.$$

If $1 \leq p < \infty$, then S is dense in $L^p(\mu)$.

Proof. Clearly $S \subset L^p(\mu)$. If $f \in L^p(\mu)$ and $f \geq 0$, let s_n be a sequence of simple functions such that $0 \leq s_n \leq f$, $s_n \to f$ pointwise. Since $s_n \in L^p$ it easily follows that $s_n \in S$. Now inequality $0 \leq |f - s_n|^p \leq f^p$ and the dominated convergence theorem implies that $s_n \to f$ in L^p. In the general case we write $f = (u^+ - u^-)$ and apply the above argument to each of the functions u^+ , u^- separately. □

Proof of Theorem 3.28. According to Lemma 3.30 it suffices to prove that the characteristic function of a set of finite measure can be approximated in L^p by compactly supported continuous functions. Let E be a measurable set of finite measure. Given $\varepsilon > 0$ let $K \subset E$ be a compact set such that $\mu(E \setminus K) < (\varepsilon/2)^p$ – this is possible by Theorem 1.68. Let U be an open set such that $K \subset U$, $\overline{U}$ is compact and $\mu(U \setminus K) < (\varepsilon/2)^p$. Finally let

$\varphi \in C_c(X)$ be such that $\operatorname{supp}\varphi \subset U$, $0 \le \varphi \le 1$, $\varphi(x) = 1$ for $x \in K$ – for example if we set $d := \frac{1}{2}\operatorname{dist}(K, X \setminus U)$ we can choose $\varphi(x) := \frac{(d - \operatorname{dist}(x,K))_+}{d}$. We then have

$$\begin{aligned}\|\chi_E - \varphi\|_{L^p} = \left(\int_{X\setminus K} |\chi_E - \varphi|^p d\mu\right)^{\frac{1}{p}} \le \left(\int_{X\setminus K} |\chi_E|^p d\mu\right)^{\frac{1}{p}} + \left(\int_{X\setminus K} |\varphi|^p d\mu\right)^{\frac{1}{p}} \\ \le \mu(E \setminus K)^{1/p} + \mu(U \setminus K)^{1/p} < \varepsilon.\end{aligned}$$

□

If $X = \mathbb{R}^n$ (or any smooth manifold) it is easy to change $C_c^0(X)$ to $C_c^\infty(X)$ in Theorem 3.28, by choosing a suitable smoother version of φ (see also Exercise 4.34).

Definition 3.31. A set $A \subset X$ for a metric space X is called *dense* if for any $x \in X$ and any $\varepsilon > 0$ there exists $a \in A$ with $d(a, x) < \varepsilon$.

A metric space X is *separable* if there is a *countable* dense set.

Theorem 3.32. *Let $\Omega \subset \mathbb{R}^n$ be open, μ a nonzero-Radon measure, and $1 \le p < \infty$, then*

(1) $L^p(\Omega, \mu)$ is separable
(2) $C_c^0(\Omega)$ is dense in $L^p(\Omega, \mu)$, where

$$C_c^0(\Omega) := \{f \in C^0(\Omega) : \quad \operatorname{supp} f \subset \Omega\},$$

and $\operatorname{supp} f$ denotes the support of f,

$$\operatorname{supp} f = \overline{\{x \in \Omega : f(x) \neq 0\}}.$$

(3) $C_c^\infty(\Omega)$ is dense in $L^p(\Omega, \mu)$, where

$$C_c^\infty(\Omega) := \{f \in C^\infty(\Omega) : \quad \operatorname{supp} f \subset \Omega\}.$$

Exercise 3.33. *Show Theorem 3.32.*

Hint: for (1) use Theorem 3.28 and Stone-Weierstrass. (2) and (3) can be proven similarly to Theorem 3.28.

By Stone-Weierstrass $C^0(\overline{\Omega})$ is a separable space if Ω is a bounded open set, moreover $C^0(\overline{\Omega})$ is a closed subset of $L^\infty(\Omega)$ with Lebesgue measure. It is a strict subset, because $L^\infty(\Omega)$ is not separable (so C_c^0 is not dense, and the Theorem 3.32 does not hold for $p = \infty$).

Exercise 3.34. *Let $\Omega = [0, 1]$, $\mu = \mathcal{L}^1$. Set $f_t := \chi_{[0,t]}$, $0 < t \le 1$. Show that*

(1) $\|f_t - f_s\|_{L^\infty} = 1$ whenever $s \neq t$.
(2) $(f_t)_{t\in[0,1]}$ is uncountable
(3) Thus L^∞ is not separable.

3.2. **Lebesgue integral vs Riemann integral.** How do we actually compute the Lebesgue integral? Well, if we really want to compute the Lebesgue integral chances are we can just use the Riemann integral. They coincide on Riemann-integrable functions.

Theorem 3.35. *Let $f : -\infty < a < b < \infty \to \mathbb{R}$ be Riemann-integrable (in the classical Darboux sense. No improper Riemann integrability allowed!). Then f is Lebesgue integrable, i.e. $f \in L^1([a,b], \mathcal{L}^1)$ and we have*

$$\mathrm{Ri} - \int_{[a,b]} f(x)dx = \int_{[a,b]} f(x)d\mathcal{L}^1(x).$$

Proof. W.l.o.g. $a = 0$ and $b = 1$. Clearly Riemann integral and Lebesgue integral coincide on constant functions.

Since f is Riemann-integrable it is bounded, and by adding a constant to f (changing the integrals equally), we may assume w.l.o.g. that f is nonnegative.

The Riemann-Lebesgue theorem states that a function f is Riemann integrable if and only if f is bounded, and f is continuous outside of a set N of zero Lebesgue measure.

Then $f : [0,1]\backslash N \to [0,\infty)$ is continuous and thus $\mathcal{L}^1$-measurable (because f^{-1} maps open sets to open sets). But then $f : [0,1] \to [0,\infty)$ is also measurable since

$$f^{-1}(A) = \left(f\Big|_{[0,1]\backslash N}\right)^{-1}(A) \cup \underbrace{(f^{-1}(A) \cap N)}_{\text{zeroset}}.$$

Thus f is measurable and nonnegative, and since it is also bounded it is indeed $\mathcal{L}^1$-integrable with

$$\int_{[0,1]} f(x)d\mathcal{L}^1(x) \leq \sup_{[0,1]} f \ \mathcal{L}^1([0,1]) < \infty.$$

We still need to show that the two integrals coincide.

Fix $\varepsilon > 0$. By the Darboux sum definition there exists a partition $0 = x_0 < x_1 < \ldots < x_N = 1$ such that for $m_i := \inf_{z\in[x_{i-1},x_i]} f(z)$ and $M_i := \sup_{z\in[x_{i-1},x_i]} f(z)$ we have

$$\sum_{i=1}^{N}(x_i - x_{i-1})M_i - \varepsilon \leq \mathrm{Ri}\int_{[0,1]} f(x)dx \leq \sum_{i=1}^{N}(x_i - x_{i-1})m_i + \varepsilon$$

Set

$$g(x) := m_i \quad \text{where } i \text{ is such that } x \in (x_{i-1}, x_i)$$

and

$$G(x) := M_i \quad \text{where } i \text{ is such that } x \in (x_{i-1}, x_i)$$

(observe g_i and G_i are a.e. defined only, but that is enough).

Since g_i and G_i are step functions, we have

$$\int_{[0,1]} gd\mathcal{L}^1(x) = \sum_{i=1}^{N} \mathcal{L}^1\left((x_{i-1}, x_i)\right) m_i = \sum_{i=1}^{N}(x_i - x_{i-1})m_i$$

and

$$\int_{[0,1]} G d\mathcal{L}^1(x) = \sum_{i=1}^{N} \mathcal{L}^1\left((x_{i-1}, x_i)\right) M_i = \sum_{i=1}^{N} (x_i - x_{i-1}) M_i.$$

That is we have

$$\int_{[0,1]} G d\mathcal{L}^1(x) - \varepsilon \leq \mathrm{Ri} \int_{[0,1]} f(x)dx \leq \int_{[0,1]} g d\mathcal{L}^1(x) + \varepsilon$$

Now observe that by definition $g \leq f \leq G$ a.e. . So we end up with

$$\int_{[0,1]} f d\mathcal{L}^1(x) - \varepsilon \leq \mathrm{Ri} \int_{[0,1]} f(x)dx \leq \int_{[0,1]} f d\mathcal{L}^1(x) + \varepsilon$$

This implies that

$$\left| \int_{[0,1]} f d\mathcal{L}^1(x) - \mathrm{Ri} \int_{[0,1]} f(x)dx \right| < \varepsilon.$$

This holds for any $\varepsilon > 0$ and by letting $\varepsilon \to 0$ we obtain

$$\int_{[0,1]} f d\mathcal{L}^1(x) = \mathrm{Ri} \int_{[0,1]} f(x)dx.$$

□

So we conclude that for many situations the Lebesgue-integral is just the Riemann integral. But the Lebesgue integral allows often for more functions (they have usually no practical meaning, but are theoretically important).

Example 3.36. The Dirichlet function

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

is not Riemann-integrable, however since $f(x) = 0$ $\mathcal{L}^1$-a.e. (exercise!) we have that f is Lebesgue integrable and $\int f(x) d\mathcal{L}^1 = 0$.

Since the n-dimensional Riemann integral is just (by Fubini's theorem) several 1-dimension integrals one can extend this theorem to any dimensions and domains as long as they are Riemann-measurable domains (once we have the Fubini theorem for measures, Section 4).

One can also conclude that the improper integrals for *nonnegative* functions coincide (by dominated convergence theorem for the Lebesgue integral and definition for the Riemann integral).

However this is not true for general improper integrals:

Exercise 3.37. *Show that* $\mathrm{Ri} - \int_{\mathbb{R}} \frac{\sin(x)}{x}\, dx$ *is finite, but show that* $\frac{\sin(x)}{x}$ *is not Lebesgue-integrable.*

So we conclude that Lebesgue integral and Riemann integral are pretty much the same from a practical point of view (i.e. "I want to compute the integral" – indeed the Riemann integral is easier to compute e.g. numerically).

The Lebesgue integral however are better theoretical spaces, the main advantage being that many limits of integrable functions are integrable (under mild assumptions – because even the *a.e.*-limit stays measurable), whereas for Riemann-integral you need something like uniform convergence. So one of the main reasons of doing all of this is the completeness above Theorem 3.27. Another reason is that we have much more variety for measures (and thus integrals), we can integrate w.r.t Hausdorff measure etc.

3.3. **Theorems of Lusin and Egorov.** We have seen in Theorem 2.14 that measurable functions are "infinite step functions". But more is true. Measurable functions are continuous on large parts of their domain (Lusin). Moroever pointwise convergence of measurable functions is uniform convergence on a large set (Egorov).

Theorem 3.38 (Egorov)**.** *Let X be a metric space and μ any Radon measure. Suppose that $\mu(X) < \infty$.*

Let $f_k, f : X \to \bar{\mathbb{R}}$ μ-measurable and μ-a.e. finite[15]*, i.e. $|f(x)| < \infty$ for μ-a.e. $x \in X$.*

Moreover for μ-a.e. $x \in X$ assume

$$\lim_{k\to\infty} f_k(x) = f(x).$$

Then for any $\delta > 0$ there exists a compact set $F \subset X$ such that

$$\mu(X \backslash F) < \delta$$

and we have uniform convergence $f_k\Big|_F \xrightarrow{k\to\infty} f\Big|_F$, namely

$$\sup_{x\in F} |f_k(x) - f(x)| \xrightarrow{k\to\infty} 0.$$

Remark 3.39.
- Taking $f_k := \chi_{B(0,k)}$ which converges pointwise to $f \equiv 1$ in $\mathbb{R}^n$, however there is no compact set $\mathcal{L}^n(\mathbb{R}^n \backslash F) < \infty$. So the assumption $\mu(X) < \infty$ is in general needed necessary.
- The result of Theorem 3.38 is not possible for $\delta = 0$ (i.e. uniform convergence outside of a zero-set may not be true), take the usual $f_k(x) = x^k$ in $[0, 1]$ which does converge pointwise but not uniformly in $[0, 1]$, nor any dense subset of $[0, 1]$, Exercise 1.80.

[15]Alternatively we could assume f_k or f in $L^1(X, \mu)$. Indeed, if $f \in L^1(X, \mu)$ then f is μ-a.e. finite, Lemma 3.21.

Proof of Theorem 3.38. Fix $\delta > 0$. For each $i, j \in \mathbb{N}$ set

$$C_{i,j} := \bigcup_{k=j}^{\infty} \{x \in X : |f_k(x) - f(x)| > 2^{-i}\}.$$

Observe that pointwise convergence f_k to f a.e. $\mu(\bigcap_j C_{i,j}) = 0$ for each i. Indeed, if $x \in \bigcup_{k=j}^{\infty} \{z \in X : |f_k(z) - f(z)| > 2^{-i}\}$ for every j, then there are infinitely many k such that $x \in \{z \in X : |f_k(z) - f(z)| > 2^{-i}\}$, so there is a subsequence $(f_{k_\ell})_{\ell \in \mathbb{N}}$ such that

$$|f_{k_\ell}(x) - f(x)| \geq 2^{-i}.$$

But by a.e. convergence of f_k this means that x belongs to a zero set.

The pointwise limit of measurable functions is measurable, so f and also $C_{i,j}$ is measurable, Theorem 2.13.

Moroever we have $C_{i,j+1} \subset C_{i,j}$.

Since $\mu(X) < \infty$ we can apply Theorem 1.33 (3) and have

$$\lim_{j \to \infty} \mu(C_{i,j}) = \mu(\bigcap_{j=1}^{\infty} C_{i,j}) = 0.$$

In particular for any $i \in \mathbb{N}$ there exists $N(i)$ such that

$$\mu(C_{i,N(i)}) < \delta 2^{-i-1}.$$

Set $A := X \backslash \bigcup_{i=1}^{\infty} C_{i,N(i)}$ then we have

$$\mu(X \backslash A) \leq \sum_{i=1}^{\infty} \mu(C_{i,N(i)}) < \frac{\delta}{2}.$$

Also, by definition of $C(i, N(i))$ we have

$$\sup_{x \in A} |f_k(x) - f(x)| \leq \sup_{x \notin C(i,N(i))} |f_k(x) - f(x)| \leq 2^{-i}, \quad \forall k \geq N(i).$$

That is: f_k converges uniformly to f in A!

Since A is measurable and $\mu(A) \leq \mu(X) < \infty$, we apply Theorem 1.68 and find some compact $F \subset X$ such that

$$\mu(A) \leq \mu(F) + \frac{\delta}{2},$$

i.e. (as a Radon measure, μ is Borel-measure, so F is μ-measurable)

$$\mu(A \backslash F) \leq \frac{\delta}{4}.$$

In particular we have

$$\mu(X \backslash F) \leq \mu(X \backslash A) + \mu(A \backslash F) < \delta.$$

Since f_k uniformly converges to f in A we can conclude. □

Theorem 3.40 (Lusin)**.** *Let X be a metric space and μ any Radon measure. Suppose that $\mu(X) < \infty$.*

Let $f : X \to \bar{\mathbb{R}}$ be a μ-measurable and μ-a.e. finite, i.e. $|f(x)| < \infty$ μ-a.e.

Then there exists $\delta > 0$ and $F \subset X$ compact such that

- $\mu(X \backslash F) < \delta$
- $f\Big|_F : F \to \mathbb{R}$ *is continuous.*

Remark 3.41. Observe that

$$f\Big|_F : F \to \mathbb{R} \quad \text{is continuous}$$

and

$$f : X \to \mathbb{R} \quad \text{is continuous in F}$$

are *not* the same.

For example if we take $X = [0,1]$, $f := \chi_{[0,1]\backslash\mathbb{Q}}$ then f is nowhere continuous in $[0,1]$. However $f : \Big|_{[0,1]\backslash\mathbb{Q}}$ is constant and thus continuous as map $[0,1]\backslash\mathbb{Q} \to \mathbb{R}$.

In particular this example also shows (for the Lebesgue measure) that Theorem 3.40 may not hold for $\delta = 0$ (i.e. outside of a zero-set) because in general one cannot find a *compact* F for $\delta = 0$.

Proof of Theorem 3.40. Step 1 We show the statement for simple functions

$$g := \sum_{i=1}^{I} b_i \chi_{B_i},$$

where B_i are measurable, pairwise disjoint sets with $X = \bigcup_{i=1}^{I} B_i$, and $b_i \in \mathbb{R}$.

Fix $\delta > 0$. By Theorem 1.68 for each B_i (since $\mu(B_i) < \infty$) there exists a compact $F_i \subset B_i$ such that

$$\mu(B_i) \leq \mu(F_i) + \delta 2^{-i}.$$

In particular F_i is μ-measurable since F_i is a Borel set and μ is Borel measure. Thus,

$$\mu(B_i \backslash F_i) < \delta 2^{-i} \quad 1 \leq i \leq I.$$

Since $(B_i)_i$ are pairwise disjoint, so are the sets $(F_i)_i$. Since each F_i is compact this implies $\operatorname{dist}(F_i, F_j) > 0$ for each $i \neq j$. So if we set

$$F := \bigcup_{i=1}^{I} F_i \subset X$$

then we have that $g\Big|_F$ is locally constant (namely: for each $x_0 \in F$ there exists a radius $\rho = \rho(x_0) > 0$ such that $g\Big|_{F\cap B(x_0,\rho)}$ is constant). This implies $g\Big|_F$ is continuous as a map $F \to \mathbb{R}$ (and thus in particular bounded).

Moreover,

$$\mu(X\backslash F) = \mu(\bigcup_{i=1}^{I}(B_i\backslash F_i)) \leq \sum_{i=1}^{I}\mu(B_i\backslash F_i) < \delta.$$

Step 2 Since $|f(x)| < \infty$ for μ-a.e. x, up to changing f on a μ-zero set (which does not influence the result, because we can always pass to a smaller (compact) set if needed to get rid of the zero-set). Now if we write $f = f^+ - f^-$ and apply Theorem 2.14 to f_+ and f_- then we have

$$f(x) = \sum_{\ell=1}^{\infty}\frac{1}{\ell}\chi_{A_\ell^+}(x) - \sum_{\ell=1}^{\infty}\frac{1}{\ell}\chi_{A_\ell^-}(x) \quad \text{everywhere in } X$$

In particular by writing

$$f_k(x) = \sum_{\ell=k}^{\infty}\frac{1}{\ell}\chi_{A_\ell^+}(x) - \sum_{\ell=1}^{k}\frac{1}{\ell}\chi_{A_\ell^-}(x).$$

Since these are finitely many sets, we can make them disjoint, that is we can write

$$f_k(x) = \sum_{i=1}^{I_k} b_i\chi_{B_i}$$

where $(B_i)_{i=1}^{\infty}$ are pairwise disjoint, and by adding $X\backslash\bigcup_{i=1}^{I_k}B_i$ (which still has finite measure) we have that f_k satisfies the conditions of Step 1.

So fix again $\delta > 0$, apply Step 1 to f_k and we find compact sets $F_k \subset X$ such that for each $k \in \mathbb{N}$.

$$\mu(X\backslash F_k) < \delta 2^{-k-1}$$

and

$$f_k\Big|_{F_k} : F_k \to \mathbb{R} \quad \text{is continuous}$$

We furthermore find a compact set $F_0 \subset X$ from Egorov's theorem, Theorem 3.38, such that

$$\mu(X\backslash F_0) < \frac{\delta}{2}$$

and

$$\sup_{x\in F_0}|f_k(x) - f(x)| \xrightarrow{k\to\infty} 0.$$

Now set $F := \bigcap_{k=0}^{\infty}F_k \subset X$. Then F is compact and

$$\mu(X\backslash F) \leq \mu(\bigcup_{k=0}^{\infty}(X\backslash F_k)) \leq \sum_{k=0}^{\infty}\mu(X\backslash F_k) < \delta.$$

Since $F \subset F_0$ we have uniform convergence of $f_k\Big|_F$ to $f\Big|_F$. Since each $f_k\Big|_F$ is continuous, its uniform limit $f\Big|_F : F \to \mathbb{R}$ is continuous. $\square$

Corollary 3.42 (Lusin in Euclidean space)**.** *Let $E \subset \mathbb{R}^n$ be Lebesgue-measurable. Then a function $f : E \to \mathbb{R}$ (i.e. f is pointwise finite!) is Lebesgue measurable if and only for any $\varepsilon > 0$ there is a closed*[16] *set $F \subset E$ such that $f\Big|_F : F \to \mathbb{R}$ is continuous and $\mathcal{L}^n(E \backslash F) < \varepsilon$.*

Proof. $\Rightarrow$ Fix $\delta \in (0,1)$.

Let $A_0 := \overline{B(0,1)}$ and $A_k := \overline{B(0,k\delta)} \backslash B(0, k(1-2^{-k-2})\delta)$, $k \geq 1$.

Observe that this way

$$\mathcal{L}^n(A_k \cap A_\ell) \begin{cases} = 0 & |\ell - k| \geq 2 \\ \leq C(n) 2^{-k} k \delta^n & |\ell - k| = 1. \end{cases}$$

Apply Theorem 3.40 to each A_k. We then find a compact $\tilde{F}_k \subset A_k$ such that $\mathcal{L}^n(A_k \backslash F_k) < 2^{-k}\delta$ and $f\Big|_{\tilde{F}_k}$ is continuous. Setting $F_k := \tilde{F}_k \cap \overline{B(0, k\delta(1-2^{-k-1})}$ we have that the (F_k) are pairwise disjoint compact sets, so $f\Big|_{\bigcup_k F_k}$ is continuous and still we have

$$\mathcal{L}^n(A_k \backslash F_k) \leq \mathcal{L}^n(A_k \backslash \tilde{F}_k) + \mathcal{L}^n(\overline{B(0,k\delta)} \backslash \overline{B(0, k(1-2^{-k-1}\delta)}) \leq C(n)\delta 2^{-k}(1+k).$$

In particular $\mathbb{R}^n = \bigcup_{k=1}^\infty A_k$ so if we set $\tilde{F} := \bigcup_{k=1}^\infty F_k$ then $f\Big|_{\tilde{F}}$ is continuous and

$$\mathcal{L}^n(\mathbb{R}^n \backslash \tilde{F}) \leq \sum_{k=1}^\infty \mathcal{L}^n(A_k \backslash F_k) \leq C(n)\delta.$$

However $\tilde{F}$ may not be closed. But in view of Theorem 1.68 there exists an open set $G \supset \mathbb{R}^n \backslash \tilde{F}$ such that

$$\mathcal{L}^n(G) \leq 2\mathcal{L}^n(\mathbb{R}^n \backslash \tilde{F}).$$

Set $F := \mathbb{R}^n \backslash G$ then F is closed and $F \subset \tilde{F}$. So $f\Big|_F$ is continuous and

$$\mathcal{L}^n(\mathbb{R}^n \backslash F) = \mathcal{L}^n(G) \leq 2C(n)\delta.$$

If we chose δ so that $2C(n)\delta < \varepsilon$ we can conclude.

$\Leftarrow$ f is measurable iff $E_a := \{x : f(x) \geq \alpha\}$ is measurable for each α, (cf. Lemma 2.4, and observe that f is pointwise finite)

[16] it clearly can't be compact. Namely $\mathcal{L}^n(\mathbb{R}^n \backslash F) = \infty$ for any bounded set F

Let F be a closed set such that $f\Big|_F$ is continuous and $\mathcal{L}^n(E\backslash F) < \varepsilon$. Then the set

$$F'_a := \{x \in F : \quad f\Big|_F(x) \geq a\} = E_a \cap F$$

is closed (this follows from the continuity of f). Thus

$$\mathcal{L}^n(E_a\backslash F'_a) = \mathcal{L}^n(E_a\backslash F) \leq \mathcal{L}^n(E\backslash F) < \varepsilon.$$

Since this argument works for any $\varepsilon > 0$, from Theorem 1.70(3) we obtain that E_a is measurable.

□

3.4. **Convergence in measure.** For sequences of μ-measurable functions $f_k : X \to \mathbb{R}$, we have learned about pointwise convergence

$$f_k(x) \xrightarrow{k\to\infty} f(x) \quad \forall x$$

slighly weaker, μ-almost everywhere convergence

$$f_k(x) \xrightarrow{k\to\infty} f(x) \quad \mu\text{-a.e. } x,$$

then (way stronger) than a.e. uniform convergence

$$\|f_k - f\|_{L^\infty(X,\mu)} = \sup_{x\in X} |f_k(x) - f(x)| \xrightarrow{k\to\infty} 0.$$

Recall that from now on we consider sup to be the essential sup $\operatorname{ess\,sup}$.

From the definition of L^p-spaces we also have L^p-convergence,

$$\|f_k - f\|_{L^p(X,\mu)} = \left(\int_X |f_k(x) - f(x)|^p d\mu\right)^{\frac{1}{p}} \xrightarrow{k\to\infty} 0.$$

Uniform convergence is then very similar to L^∞-convergence if the functions in question are continuous and the domains are open sets.

Now we introduce a weaker notion of convergence than pointwise, namely *convergence in measure*.

Definition 3.43. Let (X, μ) be a metric measure space. We say that a sequence of μ-measurable functions $(f_k)_{k\in\mathbb{N}}$, $f_k : X \to \mathbb{R}$ *converges in measure* to $f : X \to \mathbb{R}$ is for every $\varepsilon > 0$

$$\lim_{n\to\infty} \mu\left(\{x \in X : \ |f_n(x) - f(x)| > \varepsilon\}\right) = 0.$$

Some people (us including) will write sometimes $f_n \xrightarrow{\mu} f$. As usual we assume that the functions f_n and f are defined a.e.

Theorem 3.44 (Lebesgue)**.** *Assume $\mu(X) < \infty$ and a sequence of measurable functions $f_n : X \to \mathbb{R}$ converges to $f : X \to \mathbb{R}$ μ-almost everywhere. Then $f_n \xrightarrow{\mu} f$.*

Proof. Fix any $\varepsilon > 0$. Set

$$E_i := \{x \in X : |f_i(x) - f(x)| \geq \varepsilon\}.$$

The sequence of sets $A_n := \bigcup_{i=n}^{\infty} E_i$ is decreasing sequence of measurable sets, all with finite measure, and hence, Theorem 1.33,

$$\mu\left(\bigcup_{i=n}^{\infty} E_i\right) \xrightarrow{n\to\infty} \mu\left(\bigcap_{n=1}^{\infty}\bigcup_{i=n}^{\infty} E_i\right) = 0.$$

The last set has measure zero, because if $x \in \bigcup_{i=n}^{\infty} E_i$ for every n, then x belongs to infinitely many E_i's, so there is a subsequence f_{n_i} such that

$$|f_{n_i}(x) - f(x)| \geq \varepsilon$$

which is only true on a set of measure zero, by the a.e. convergence.

Hence

$$\mu(\{x : |f_n(x) - f(x)| \geq \varepsilon\}) = \mu(E_n) \leq \mu(\bigcup_{i=n}^{\infty} E_i) \xrightarrow{n\to\infty} 0,$$

which proves convergence in measure. □

The converse of Theorem 3.44 does not hold (see Exercise 3.47 below), but it holds up to subsequence.

Theorem 3.45 (Riesz)**.** *If $f_n \xrightarrow{\mu} f$ then there is a subsequence $(f_{n_i})_{i\in\mathbb{N}}$ such that $f_{n_i} \to f$ a.e.*

Observe that here $\mu(X) = \infty$ is permissible.

Proof. For every $i \in \mathbb{N}$ there exist n_i such that

$$\mu\left(\{x : \quad |f_{n_i}(x) - f(x)| \geq \frac{1}{i}\}\right) \leq 2^{-i}.$$

We can assume that $n_1 < n_2 < n_3$. Let

$$F_k := X \setminus \bigcup_{i=k}^{\infty} \{x : |f_{n_i}(x) - f(x)| \geq \frac{1}{i}\}$$

and

$$F := \bigcup_{k=1}^{\infty} F_k.$$

Then

$$\mu(X \setminus F_k) \leq \sum_{i=k}^{\infty} 2^{-i} = 2^{1-k}.$$

and thus

$$\mu(X \setminus F) \leq \limsup_{k\to\infty} \mu(X \setminus F_k) \leq \limsup_{k\to\infty} 2^{1-k} = 0.$$

We claim that

$$f_{n_i}(x) \xrightarrow{i\to\infty} f(x) \quad \text{for any } x \in F. \tag{3.4}$$

Since $X \setminus F$ is a μ-zeroset we can conclude once we have proven this.

It suffices to prove (3.4) for each F_k, and actually we will show that for each F_k we have uniform convergence.

Indeed, if $x \in F_k$ then x then $x \notin \bigcup_{i=k}^{\infty}\{x : |f_{n_i}(x) - f(x)| \geq \frac{1}{i}\}$. That is

$$|f_{n_j}(x) - f(x)| \leq \frac{1}{j} \quad \forall j \geq k, \quad \forall x \in F_k$$

This is uniform convergence in F_k. □

Remark 3.46. We actually proved not only convergence a.e. but also uniform convergence on subsets of X whose complement has arbitrary small measure. Note that Lebesgue's theorem Theorem 3.44 combined with this stronger conclusion of Riesz' theorem Theorem 3.45 implies Egorov's theorem for sequences of real valued functions, Theorem 3.38.

One does indeed need to pass to a subsequence in Theorem 3.45, as the following example shows.

Exercise 3.47. *Let $E_k \subset [0,1]$ be defined as follows.*

$$E_k := \begin{cases} [0, \frac{1}{i}] & \text{if } k = 2i-1 \\ [\frac{1}{i}, 1] & \text{if } k = 2i. \end{cases}$$

Set

$$f_k := \chi_{E_k}.$$

(1) Show w.r.t Lebesgue measure that f_k converges in measure to f (what is f?)
(2) Show that f_k does not pointwise a.e. converge to f.

Hint: *(1) is false* **ToDo** *, don't do it.*

Exercise 3.48. *Let $f(x) := \chi_{[0,1]}$. For each $n \in \mathbb{N}$ there exists exactly one $m \in \mathbb{N}$ and $k \in \{0, \dots, 2^m - 1\}$ such that $n = 2^m + k$. Set*

$$f_n(x) := f(2^m x - k).$$

Show

- *f_n converges to 0 in $L^2([0,1])$*
- *$f_n(x)$ does not converge to 0 for any $x \in (0,1)$.*

3.5. **L^p-convergence and weak L^p.** Next we want to consider convergence in L^p-norm. For this we first show

Lemma 3.49 (Chebyshev's inequality)**.** *Let $f : X \to \mathbb{R}$ be μ-integrable. Then for any $\lambda > 0$, any $p \in [1, \infty)$*

$$\mu(\{x : f(x) > \lambda\}) \leq \frac{1}{\lambda^p} \int_{\{x:f(x)>\lambda\}} |f(x)|^p \, dx,$$

and in particular

$$\mu(\{x : f(x) > \lambda\}) \leq \frac{1}{\lambda^p} \|f\|^p_{L^p(X)}.$$

Proof. Since f is measurable the set $\{x : f(x) > \lambda\}$ is measurable and thus

$$\begin{aligned}\mu(\{x : f(x) > \lambda\}) =& \int_{\{x:f(x)>\lambda\}} d\mu \\ =& \lambda^{-p} \int_{\{x:f(x)>\lambda\}} \lambda^p d\mu \\ \leq& \lambda^{-p} \int_{\{x:f(x)>\lambda\}} |f|^p d\mu.\end{aligned}$$

□

As a slight digression, Chebyshev's inequality leads to a space very similar to L^p, but weaker – hence called *weak L^p-space*, denoted by $L^{(p,\infty)}$.

Definition 3.50 (weak L^p)**.** We say that a μ-measurable $f : X \to \mathbb{R}$ belongs to weak L^p, $f \in L^{(p,\infty)}(X, \mu)$ if there exists $\Lambda \geq 0$ such that for all $\lambda > 0$

$$\mu(\{x : f(x) > \lambda\}) \leq \frac{1}{\lambda^p} \Lambda^p.$$

The minimal value of Λ such that the above inequality holds is denoted by $\|f\|_{L^{(p,\infty)}(X,\mu)}$.

$$\|f\|_{L^{(p,\infty)}(X,\mu)} := \sup_{\lambda>0} \lambda \left(\mu(\{x : f(x) > \lambda\})\right)^{\frac{1}{p}}$$

One can check that $\|f\|_{L^{(p,\infty)}(X,\mu)}$ is not a true norm (triangle inequality is only true up to a multiplicative constant).

Chebyshev's inequality implies that $\|f\|_{L^{(p,\infty)}} \leq \|f\|_{L^p}$. The other direction is not true.

Exercise 3.51. *Show that for $\sigma \in (0, n]$ the function $f(x) = |x|^{-\sigma}$ belongs to $L^{(\frac{n}{\sigma},\infty)}(\mathbb{R}^n)$. Show that $f \notin L^q(\mathbb{R}^n)$ for any $q \in [1, \infty]$.*

Weak L^p-spaces become very important in Harmonic Analysis, they can also be generalized to so-called Lorentz spaces $L^{(p,q)}$.

As a corollary of Chebyshev inequality we get

Theorem 3.52. *Let $p \in [1, \infty]$ and assume that $f_k \xrightarrow{k\to\infty} f \in L^p(X, \mu)$, i.e.*

$$\|f_k - f\|_{L^p(X,\mu)} \xrightarrow{k\to\infty} 0.$$

Then

- $f_k \xrightarrow{\mu} f$
- *there exists a subsequence $k_i \to \infty$such that $f_{k_i} \xrightarrow{i\to\infty} f$ μ-a.e.*

Proof of Theorem 3.52. The second claim follows from the first one from Riesz theorem, Theorem 3.45. For the first claim observe that by Lemma 3.49

$$\mu\{|f_k - f| > \varepsilon\} \leq \varepsilon^{-p}\|f_k - f\|_{L^p}^p \xrightarrow{k\to\infty} 0.$$

□

Here is another criterion for convergence

Theorem 3.53. *Let $p \in [1, \infty)$ and assume $f, f_k \in L^p(X, \mu)$ for some metric space X. Then the following are equivalent*

(1) $\|f_k - f\|_{L^p(X)} \xrightarrow{k\to\infty} 0$
(2) $f_k \xrightarrow{\mu} f$ *and* $\|f_k\|_{L^p(X)} \xrightarrow{k\to\infty} \|f\|_{L^p(X)}$.

Exercise 3.54. *Show Theorem 3.53 is false for $p = \infty$ (hint: continuous functions)*

Proof of Theorem 3.53. <u>(1) ⇒ (2)</u>: The convergence in measure follows from Theorem 3.52. Moreover we have by reverse triangle inequality

$$\Big|\|f_k\|_{L^p(X)} - \|f\|_{L^p(X)}\Big| \leq \|f_k - f\|_{L^p(X)} \xrightarrow{k\to\infty} 0.$$

<u>(2) ⇒ (1)</u>: First assume that $f_k(x) \xrightarrow{k\to\infty} f(x)$ for μ-a.e. x.

Observe that by convexity (here we need $1 \leq p < \infty$)

$$2^{p-1}|f_k|^p + 2^{p-1}|f|^p - |f_k - f|^p \geq 0.$$

Moreover, by pointwise convergence

$$\liminf_{k\to\infty} 2^{p-1}|f_k(x)|^p + 2^{p-1}|f(x)|^p - |f_k(x) - f|^p = 2 * 2^{p-1}|f(x)| \quad \mu\text{-a.e. } x$$

By assumption (2) we have

$$\liminf_{k\to\infty} \int_\Omega |f_k|^p = \int_\Omega |f|^p,$$

so by Fatou's Lemma, Corollary 3.9,

$$\begin{aligned}
&2^p \int_\Omega |f|^p d\mu - \limsup_{k\to\infty} \int_\Omega |f_k - f|^p d\mu \\
=&\liminf_{k\to\infty} \int_\Omega \left(2^{p-1}|f_k|^p + 2^{p-1}|f|^p - |f_k - f|^p\right) d\mu \\
\geq&\int_\Omega \liminf_{k\to\infty} \left(2^{p-1}|f_k|^p + 2^{p-1}|f|^p - |f_k - f|^p\right) d\mu \\
=&2^p \int_\Omega |f|^p d\mu
\end{aligned}$$

Subtracting the (finite!) $2^p \int_\Omega |f|^p d\mu$ from both sides of this inequality we find

$$\limsup_{k\to\infty} \|f_k - f\|^p_{L^p(\Omega,\mu)} = \limsup_{k\to\infty} \int_\Omega |f_k - f|^p d\mu = 0.$$

This proves (1) under the assumption that $f_k(x) \xrightarrow{k\to\infty} f(x)$ for μ-a.e. x and $\|f_k\|_{L^p(X)} \xrightarrow{k\to\infty} \|f\|_{L^p(X)}$.

Assume now $f_k \xrightarrow{\mu} f$ and $\|f_k\|_{L^p(X)} \xrightarrow{k\to\infty} \|f\|_{L^p(X)}$. Assume to the contrary that there exists some subsequence f_{k_i} such that

$$\|f_{k_i} - f\|_{L^p(X)} > \varepsilon \quad \forall i \in \mathbb{N}. \tag{3.5}$$

By Theorem 3.45 we can find a subsubsequence $(f_{k_{i_j}})_j$ such that $f_{k_{i_j}}(x) \xrightarrow{j\to\infty} f(x)$ for μ-a.e. x. But then we have by the above arguments that

$$\|f_{k_{i_j}} - f\|_{L^p(X)} \xrightarrow{j\to\infty} 0.$$

This contradicts (3.5). □

We will later discuss another way to decide when the convergence in measure implies L^p-convergence in Theorem 3.60 below, but first we need to talk about absolute continuity.

3.6. **Absolute continuity.** Let $f : X \to [0,\infty)$ μ-integrable. For μ-measurable $A \subset \Omega$ set

$$\nu(A) := \mu \llcorner f = \int_A f d\mu.$$

Cf. Theorem 3.11. Observe that then $\mu(A) = 0$ implies $\nu(A) = 0$ – an effect which we call *absolute continuity*.

Definition 3.55. Let μ, ν be measures on X such that

- any μ-measurable set is also ν-measurable
- if $\mu(A) = 0$ then $\nu(A) = 0$.

Then we say that ν is *absolutely continuous* with respect to μ, and write $\nu \ll \mu$.

Lemma 3.56. *Let μ and ν be Radon measures on $\mathbb{R}^n$*[17]

Then

(1) *Assume that for all $A \subset \mathbb{R}^n$ it holds that if $\mu(A) = 0$ then also $\nu(A) = 0$. Then any μ-measurable set is also ν-measurable.*
(2) *In particular, $\mu \ll \nu$ if and only if for all $A \subset \mathbb{R}^n$ if $\mu(A) = 0$ then also $\nu(A) = 0$.*

Proof. We use Theorem 1.70: any μ-measurable set A can be written as $A = B \cup N$ where B is Borell and $\mu(N) = 0$. So $\nu(N) = 0$ and consequently $A = B \cup N$ is ν-measurable.

□

The notion of continuity is justified by the following (important on its own) *absolute continuity of the integral*, and Exercise 3.58.

Theorem 3.57. *Let $f \in L^1(\mu)$. Then for every $\varepsilon > 0$ there is $\delta > 0$ such that $\int_E |f| d\mu < \varepsilon$ whenever E is measurable and $\mu(E) < \delta$.*

Proof. We argue by contradiction. If the claim was false, there would exist $\varepsilon_0 > 0$ and μ-measurable sets E_n such that $\mu(E_n) < 2^{-n}$ but $\int_{E_n} |f| d\mu \geq \varepsilon_0$.

The sequence of sets $A_k := \bigcup_{n=k}^{\infty} E_n$ is decreasing and

$$\mu(\bigcap_{k=1}^{\infty} A_k) = 0,$$

since

$$\mu(\bigcap_{k=1}^{\infty} A_k) \leq \mu(A_n) \leq \sum_{i=n}^{\infty} \mu(E_i) \leq 2^{1-n} \xrightarrow{n\to\infty} 0.$$

Now

$$|f| \llcorner \mu(E) = \int_E |f| d\mu$$

is a measure, Theorem 3.11, so we have

$$\int_{A_k} |f| d\mu = |f| \llcorner \mu(A_k) \xrightarrow{k\to\infty} |f| \llcorner \mu(\bigcap_{k=1}^{\infty} A_k) = \int_{\bigcap_{k=1}^{\infty} A_k} |f| d\mu = 0,$$

since $\mu(\bigcap_{k=1}^{\infty} A_k) = 0$.

On the other hand from the choice of E_n we have

$$\int_{A_k} |f| d\mu \geq \int_{E_k} |f| d\mu \geq \varepsilon_0,$$

a contradiction. □

More generally we can reformulate Theorem 3.57 in the following way

[17] easily extendable to more general sets that satisfy the assumptions of Theorem 1.70

Exercise 3.58. *Let $\mu, \nu : 2^X \to [0, \infty]$ be two measures with $\nu \ll \mu$ and assume $\nu(X) < \infty$. Then*

$$\forall \varepsilon > 0, \quad \exists \delta > 0 : \quad \forall A \subset X \ \mu\text{-measurable:} \mu(A) < \delta \Rightarrow \nu(A) < \varepsilon.$$

3.7. **Vitali's convergence theorem.** We know that if f_k converges to f in L^p then the convergence is also in measure, Theorem 3.52. Vitali's convergence theorem is a characterization when the other direction is true, i.e. when convergence in measure (or in view of Theorem 3.44 a.e. convergence if on a sets of finite measure) implies L^p-convergence.

Definition 3.59. A family $\mathcal{F}$ of μ-integrable functions $f : X \to \bar{\mathbb{R}}$ is said to have *uniformly absolutely continuous integrals* if

$$\forall \varepsilon > 0 \, \exists \delta > 0 : \quad \forall f \in \mathcal{F} : \quad \forall A \subset X : \ \mu\text{-measurable} : \quad \mu(A) < \delta \Rightarrow \int_A |f| d\mu < \varepsilon.$$

(Compare this notion with Arzela-Ascoli!)

Theorem 3.60 (Vitali convergence theorem)**.** *Let X be a metric space and μ a Radon measure. Assume $\mu(X) < \infty$ and let $(f_k)_{k \in \mathbb{N}}$ be a sequence in $L^1(X, d\mu)$ and $f \in L^1(X, d\mu)$. Then the following are equivalent*

(1) $f_k \xrightarrow{\mu} f$ and $(f_k)_{k \in \mathbb{N}}$ has uniformly absolutely continuous integrals
(2) $\lim_{k \to \infty} \int_X |f_k - f| d\mu = 0$.

Example 3.61. The assumption $\mu(X) < \infty$ is needed in general. E.g. take $\mu = \mathcal{L}^n$, $X = \mathbb{R}^n$, $f_k = k^{-n} \chi_{B(0,k)}$. Then $f_k \xrightarrow{\mu} 0$ (because the convergence is a.e.). $(f_k)_k$ has uniformly absolutely continuous integrals, but $\int_{\mathbb{R}^n} |f_k| d\mu = c > 0$ with a constant $c = \mathcal{L}^n(B(0,1))$ independent of k, but $\int 0 = 0$.

Proof of Theorem 3.60. <u>(2) $\Rightarrow$ (1):</u> If $\|f_k - f\|_{L^1} \xrightarrow{k \to \infty} 0$ then $f_k \xrightarrow{\mu} f$ by Theorem 3.52. To get the uniform absolute continuous integral property fix $\varepsilon > 0$.

In view of L^1-convergence there exists a $K \in \mathbb{N}$ such that

$$\|f_k - f\|_{L^1(X)} \leq \frac{\varepsilon}{2} \quad \forall k \geq K.$$

Moreover, by absolute continuty of the integral, Theorem 3.57, there exists a $\delta > 0$ such that any μ-measurable set $A \subset X$ with $\mu(A) < \delta$ satisfies

$$\max_{k \in \{1, \ldots, K\}} \int_A |f_k| + \int_A |f| < \frac{\varepsilon}{2}.$$

Since

$$\int_A |f_k| \leq \int_A |f_k - f| + \int_A |f| d\mu \leq \|f_k - f\|_{L^1(X)} + \int_A |f| d\mu,$$

we find that for all $k \geq K$ we have

$$\int_A |f_k| < \varepsilon.$$

The same holds also for $k \in \{1, \dots, K\}$, so we have shown the uniform absolute continuity of $(f_k)_k$.

<u>(1) $\Rightarrow$ (2)</u>: Let $f_k \xrightarrow{\mu} f$ with uniform absolute continuous integrals.

Assume by contradiction that

$$\limsup_{k\to\infty} \int_\Omega |f_k - f| d\mu > 0.$$

By passing to a subsequence we can assume w.l.o.g.

$$\liminf_{k\to\infty} \int_\Omega |f_k - f| d\mu > 0. \tag{3.6}$$

Passing yet to another subsequence, applying Theorem 3.45, we may assume that

$$f_k \xrightarrow{k\to\infty} f \quad \text{a.e. in } X.$$

Fix $\varepsilon > 0$. Since f_k are uniformly absolutely continuous and f is integrable we can find $\delta > 0$ such that for any μ-measurable $A \subset X$ with $\mu(A) < \delta$ we have

$$\int_A |f| d\mu < \varepsilon, \quad \int_A |f_k| d\mu < \varepsilon \quad \forall k \in \mathbb{N}.$$

For this δ we can apply Egorov's theorem, Theorem 3.38, and find a compact set $F \subset X$ with $\mu(X \backslash F) < \delta$ and

$$\sup_{x\in F} |f_k(x) - f(x)| \xrightarrow{k\to\infty} 0.$$

In particular there must be $K \in \mathbb{N}$ such that

$$\sup_{k\geq K} \sup_{x\in F} |f_k(x) - f(x)| < \frac{\varepsilon}{2\mu(X)}$$

Consequently, for any $k \geq k_0$ we have

$$\begin{aligned}\int_X |f_k - f| d\mu &= \int_{X\backslash F} |f_k - f| d\mu + \int_F |f_k - f| d\mu \\ &\leq \int_{X\backslash F} |f_k| + |f| d\mu + \int_F \sup_{k\geq K} \sup_{x\in F} |f_k - f| d\mu \\ &\leq \varepsilon + \varepsilon + \varepsilon,\end{aligned}$$

the first two estimates are because of absolute continuity, the last one because of uniform convergence.

In particular we conclude that

$$\liminf_{k\to\infty} \int_X |f_k - f| d\mu \leq 3\varepsilon,$$

since $\varepsilon > 0$ was arbitary this is a contradiction to (3.6). □

Exercise 3.62. *Formulate and prove a Vitali-type condition for L^p-convergence.*

In the calculus of variations Vitali's theorem can be used for minimization.

Definition 3.63. Let $\Omega \subset \mathbb{R}^n$. A map $F = F(x, y) : \Omega \times \mathbb{R} \to \mathbb{R}$ is a *Carathéodory-function* if

- for all $y \in \mathbb{R}$ the function $x \mapsto F(x, y)$ is μ-measurable
- for μ-a.e. $x \in \Omega$ the function $y \mapsto F(x, y)$ is continuous.

Exercise 3.64. *If F is a Carathéodory function and $u : \Omega \to \mathbb{R}$ is μ-measurable then $f(x) := F(x, u(x))$ is μ-measurable.*

Example 3.65. Let $\Omega \subset \mathbb{R}^n$ with $\mu(\Omega) < \infty$, $F = F(x, y) : \Omega \times \mathbb{R} \to \mathbb{R}$ a *Carathéodory-function* as above. Assume moreover that F has a *p-growth* for $p \in [1, \infty)$, namely

$$|F(x, y)| \leq C(1 + |y|^p) \quad \mu\text{-a.e. } x \in \Omega,\ y \in \mathbb{R}.$$

Let furthermore $u, u_k : \Omega \to \mathbb{R}$ be μ-measurable functions with

$$\int_\Omega |u|^p d\mu < \infty$$

$$\int_\Omega |u_k|^p d\mu < \infty \quad \forall k$$

$$\int_\Omega |u_k - u|^p d\mu \xrightarrow{k\to\infty} 0.$$

Then

$$\int_\Omega F(x, u_k(x)) d\mu \xrightarrow{k\to\infty} \int_\Omega F(x, u(x)) d\mu$$

(This is much easier if F is also Lipschitz or Hölder continuous in y but we assume only continuity (and $u_k(\Omega)$ is not a k-uniformly bounded set!) We also cannot use dominated convergence, since we do not have a dominating function.

Solution. Set

$$f(x) = F(x, u(x)), \quad f_k(x) := F(x, u_k(x)).$$

In view of Exercise 3.64 f and f_k are μ-measurable.

Also

$$\int_\Omega |f_k(x)| d\mu \leq C \int_\Omega (1 + |u_k(x)|^p) d\mu < \infty,$$

so f_k (and similarly f) are μ-integrable.

We want to argue with Theorem 3.60 and thus need to check its assumptions. $\underline{f_k \xrightarrow{\mu} f}$.

Assume this is not the case, then there exists some $c > 0$ such that

$$\limsup_{k\to\infty} \mu(\{x : |f_k - f| > \varepsilon\}) \geq c > 0$$

By the definition of lim sup there must be a subsequence f_{k_i} such that

$$\liminf_{i\to\infty} \mu(\{x : |f_{k_i} - f| > \varepsilon\}) \geq c > 0. \tag{3.7}$$

We will find that this leads to a contradiction.

Since u_{k_i} converges in L^p to u by Chebyshev inequality, more precisely by Theorem 3.52, we have $u_{k_i} \xrightarrow{\mu} u$.

Since $u_{k_i} \xrightarrow{\mu} u$ there exists a further subsequence $u_{k_{i_j}}$ such that $u_{k_{i_j}} \xrightarrow{j\to\infty} u$ μ-a.e. in Ω, see Theorem 3.45. Since F is continuous in the second entry we see that $f_{k_{i_j}} \xrightarrow{j\to\infty} f$ μ-a.e.

Since μ-a.e. convergence implies convergence in measure, Theorem 3.44 we found that $f_{k_{i_j}} \xrightarrow{\mu} f$ – but this implies

$$\lim_{j\to\infty} \mu(\{x : |f_{k_{i_j}} - f| > \varepsilon\}) = 0,$$

a contradiction to (3.7).

So indeed $f_k \xrightarrow{\mu} f$.

<u>$\int |f_k| d\mu$ is uniformly absolutely continuous</u>

Fix $\varepsilon > 0$. Since u_k L^p-converges to u by Vitali's theorem, Theorem 3.60, we have that $\int |u_k|^p$ is uniformly absolutely continuous, i.e. there exists $\delta > 0$ (and w.l.o.g. $\delta < \frac{\varepsilon}{2}$) such that for all μ-measurable sets A

$$\mu(A) < \infty \Rightarrow \int_A |u|^p d\mu + \sup_k \int_A |u_k|^p d\mu < \frac{\varepsilon}{2}.$$

Then we have

$$\int_A f_k d\mu \leq \int_A 1 + |u_k|^p d\mu = \mu(A) + \int_A |u_k|^p d\mu < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This proves uniform absolute continuity for $\int_A f_k d\mu$.

□

4. Product Measures, Multiple Integrals – Fubini's theorem

Fubini's theorem is essentially saying that if we want to integrate on a cube $[0,1]^2$ then we can write this as an integral on $[0,1] \times [0,1]$,

$$\int_{[0,1]^2} f(x,y)dxdy = \int_{[0,1]} \int_{[0,1]} f(x,y)dxdy.$$

Since we now work with more abstract measures, we first need to discuss what is dx dy...

Recall that the cartesian product of two spaces X and Y is given by

$$X \times Y = \{(x,y) : \quad x \in X : \quad y \in Y\}.$$

When measuring a set $A \times B$ it is easy to think that we somehow should multiply $\mu(A)\nu(B)$. But not all sets $S \subset X \times Y$ are of the form $A \times B$. What do we do? We take the best cover!

Definition 4.1 (Product measures)**.** Let $\mu : 2^X \to [0,\infty]$ and $\nu : 2^Y \to [0,\infty]$ be two measures. The *product measure* $\mu \times \nu : 2^{X\times Y} \to [0,\infty]$ for $S \subset X \times Y$ is defined as

$$(\mu\times\nu)(S) := \inf\left\{\sum_{i=1}^{\infty}\mu(A_i)\nu(B_i) : \quad S \subset \bigcup_{i=1}^{\infty} A_i \times B_i; \quad A_i \subset X \ \mu\text{-measurable}, \ B_i \subset Y \ \nu\text{-measurable}\right\}$$

In the (most relevant) case of Lebesgue measure one can see now easily that $\mathcal{L}^k \times \mathcal{L}^\ell$ on $\mathbb{R}^k \times \mathbb{R}^\ell$ is indeeed $\mathcal{L}^{k+\ell}$ on $\mathbb{R}^{k+\ell}$.

Theorem 4.2 (Fubini's theorem)**.** *Let μ, ν be Radon measures on metric spaces X and Y.*

(1) If $A \subset X$ is μ-measurable, $B \subset Y$ is ν-measurable then $A \times B$ is $\mu \times \nu$-measurable and we have
$$(\mu \times \nu)(A \times B) = \mu(A)\nu(B).$$
(2) Let $S \subset X \times Y$ be $\mu \times \nu$-measurable and $\mu \times \nu(S) < \infty$. Define for $y \in Y$ the set $S_y \in X$ as
$$S_y := \{x : \quad (x,y) \in S\}$$
then S_y is μ-measurable for ν-a.e. y. Moreover the map
$$y \mapsto \mu(S_y) = \int_X \chi_S(x,y)d\mu(x)$$
is ν-integrable and
$$(\mu \times \nu)(S) = \int_Y \mu(S_y)d\nu(y) = \int_Y \left(\int_X \chi_S(x,y)d\mu(x)\right) d\nu(y).$$
Similarly for S_x.
(3) $\mu \times \nu$ is a Radon measure
(4) Is $f : X \times Y \to \bar{\mathbb{R}}$ $\mu \times \nu$-integrable then
$$y \mapsto \int_X f(x,y)d\mu(x)$$
is integrable w.r.t ν, and
$$x \mapsto \int_Y f(x,y)d\nu(y)$$
is integrable w.r.t. μ and we have
$$\int_{X\times Y} f \ d(\mu \times \nu) = \int_X \left(\int_Y f(x,y) \ d\nu(y)\right) d\mu(x) = \int_Y \left(\int_X f(x,y) \ d\mu(x)\right) d\nu(y).$$

This is a very technical statement, and we shall not go through the proof here (see [Evans and Gariepy, 2015, Theorem 1.22] for this). More instructive is to look at warnings:

Example 4.3.

- A set $S \subset X \times Y$ such that
$$x \mapsto \chi_S(x,y) \quad \text{is } \mu\text{-measurable for } \nu\text{-a.e. } y$$
and
$$y \mapsto \mu(S_y) \equiv \int_X \chi_S(x,y)d\mu(x) \quad \text{is } \nu\text{-measurable}$$

may *not* be measurable.

Indeed let $X = Y = \mathbb{R}$ and $\mu = \nu = \mathcal{L}^1$ and take $A \subset [0,1]$ the non-measurable Vitali set. Set

$$S := \{(x,y) : \quad |x - \chi_A(y)| < \frac{1}{2}, \quad y \in [0,1]\}.$$

We have S_y either $|x| < \frac{1}{2}$ or $|x-1| < \frac{1}{2}$ (both measurable), so S_y is μ-measurable for *all* y with $\mu(S_y) \equiv 1$.

However when $x > \frac{1}{2}$ then $S_x = A$. So S_x is not measurable for $\mathcal{L}^1$-a.e. x.

If S was $\mathcal{L}^2$-measurable then both S_x and S_y would need to be measurable (by Fubini's theorem).

- Similarly (take for example $f := \chi_S$ for the S from above) there is no reason that if $\int f d\mu$ exists and is integrable that then $f : X \times Y \to \mathbb{R}$ is integrable (or even measurable).

A slight variant of Fubini's theorem is the following *Tonelli's theorem*

Theorem 4.4 (Tonelli)**.** *Let $f : X \times Y \to [0, \infty]$ be $\mu \times \nu$-measurable, and assume that the iterated integral*

$$\int_X \left(\int_Y f(x,y) d\nu(y) \right) d\mu(x)$$

exists. Then f is $\nu \times \mu$ integrable and we have

$$\int_{X \times Y} f d\mu \times \nu = \int_X \left(\int_Y f(x,y) d\nu(y) \right) d\mu(x) = \int_Y \left(\int_X f(x,y) d\mu(x) \right) d\nu(y)$$

For the Lebesgue measure and $f \in L^1(\mathbb{R}^n, \mathcal{L}^n)$ we often use as an application from Fubini's theorem

$$\int_{\mathbb{R}^n} f(x) d\mathcal{L}^n(x) = \int_{\mathbb{R}} \dots \int_{\mathbb{R}} f(x_1, \dots, x_n) d\mathcal{L}^1 x_1 \, d\mathcal{L}^1 x_2 \dots d\mathcal{L}^1 x_n$$

This uses

Theorem 4.5. *The Lebesgue measure $\mathcal{L}^n$ on $\mathbb{R}^n$ satisfies*

$$\mathcal{L}^n = \mathcal{L}^k \times \mathcal{L}^{n-k} \quad \forall \kappa \in \{0, \dots, n\}$$

In particular

$$\mathcal{L}^n = \mathcal{L}^1 \times \mathcal{L}^1 \times \dots \times \mathcal{L}^1.$$

Proof. Both measures coincide on rectangles R (by construction), so for the application of Theorem 1.71 we only need that both measures are indeed Radon measures, which follows from Fubini's theorem Theorem 4.2. □

A consequence of Fubini's theorem

Proposition 4.6 (Slicing)**.** *Assume $f \in L^1((0,R)^2)$ (take a representative and fix it). Then there exists $r \in (0,R)$ such that $x \mapsto f(x,r)$ belongs to $L^1((0,R))$ and moreover*

$$\int_{(0,R)} |f(x,r)|dx \leq \frac{1}{R}\int_{(0,R)^2} |f(x,y)|d(x,y).$$

More precisely, consider for $\lambda > 1$ the set $Y_\lambda \subset \mathbb{R}$

$$Y_\lambda := \left\{ r \in (0,R) : \int_{(0,R)} |f(x,r)|dx \leq \frac{\lambda}{R}\int_{(0,R)^2} |f(x,y)|d(x,y) \right\}.$$

then $\mathcal{L}^1(Y_\lambda) > (1-\frac{1}{\lambda})R$.

Proof. If

$$\int_{(0,R)^2} |f(x,y)|d(x,y) = 0$$

from Fubini's theorem we have

$$\int_{(0,R)} |f(x,r)|dx = 0 \quad \mathcal{L}^1\text{-a.e. } r \in (0,R),$$

and we can conclude.

So assume from now on

$$\int_{(0,R)^2} |f(x,y)|d(x,y) > 0.$$

Let

$$X_\lambda = \left\{ r \in (0,R) : \int_{(0,R)} |f(x,r)|dx > \frac{\lambda}{R}\int_{(0,R)^2} |f(x,y)|d(x,y) \right\}.$$

By Fubini's theorem X_λ is $\mathcal{L}^1$-measurable and $\mathcal{L}^1(X_\lambda) = \mathcal{L}^1((0,R)) - \mathcal{L}^1(Y_\lambda)$. For any $r \in X_\lambda$ we have

$$\frac{\lambda}{R}\int_{(0,R)^2} |f(x,y)|d(x,y) \leq \int_{(0,R)} |f(x,r)|dx$$

So integrating this in r we find

$$\mathcal{L}^1(X_\lambda)\frac{\lambda}{R}\int_{(0,R)^2} |f(x,y)|d(x,y) \leq \int_{X_\lambda}\int_{(0,R)} |f(x,r)|dxdr \leq \int_{(0,R)^2} |f(x,y)|dxdy$$

Dividing the integral on both sides we find

$$\mathcal{L}^1(X_\lambda) \leq \frac{R}{\lambda}.$$

Consequently, for any $\lambda > 1$

$$\mathcal{L}^1(Y_\lambda) \geq \mathcal{L}^1((0,R)) - \mathcal{L}^1(X_\lambda) = R - \frac{R}{\lambda} > 0$$

□

The measure estimate $\mathcal{L}^1(Y_\lambda) > (1-\frac{1}{\lambda})R$ in Proposition 4.6 is useful to show that several slicing properties hold on the same set simultaneously.

Exercise 4.7. *Show that there exists a uniform $\Lambda > 0$ such that the following holds:*

Take any (representative of) $f, g \in L^1((0,1)^2)$. There exists an $\mathcal{L}^1$-measurable set $Y \subset (0,1)$ with $\mathcal{L}^1(Y) > 0$ (depending on f and g) such that for each $r \in Y$ we have

(1) $x \mapsto f(x,r)$ belongs to $L^1((0,1))$
(2) $x \mapsto g(x,r)$ belongs to $L^1((0,1))$
(3) $\int_{(0,1)} |f(x,r)|dx \leq \Lambda \int_{(0,1)^2} |f(x,y)|d(x,y)$.
(4) $\int_{(0,1)} |g(x,r)|dx \leq \Lambda \int_{(0,1)^2} |f(x,y)|d(x,y)$.

Hint: *Use Proposition 4.6 for both f and g and obtain Y_f and Y_g where the above statements hold. To show that $\mathcal{L}^1(Y_f \cap Y_g) > 0$ use Exercise 4.8*

Exercise 4.8. *Let (X, μ) be a measure space with $\mu(X) < \infty$. Assume $A, B \subset X$ are μ-measurable with $\mu(A) > \frac{2}{3}\mu(X)$ and $\mu(B) > \frac{2}{3}\mu(X)$. Show that $\mu(A \cap B) \geq \frac{1}{6}\mu(X)$.*

Exercise 4.9. *Let $f \in L^1((0,1)^2)$ then for $\mathcal{L}^1$-a.e. $r \in (0,1)$ we have $f(r, \cdot) \in L^1(0,1)$.*

Hint: *Apply Proposition 4.6 to*

$$Y_k := \left\{ r \in (0,1) : \int_{(0,1)} |f(x,r)|dx \leq 2^k \int_{(0,1)^2} |f(x,y)|d(x,y) \right\}$$

and show $(0,1) \setminus \bigcup_k Y_k$ is a zero-set.

4.1. Application: Interpolation between L^p-spaces – Marcienkiewicz interpolation theorem.

If $f \in L^{p_1}(\mathbb{R}^n) \cap L^{p_2}(\mathbb{R}^n)$ then $f \in L^p(\mathbb{R}^n)$ for $p \in (p_1, p_2)$, cf. Exercise 3.18.

So what if we have some information about a map T acting on $L^{p_1}(\mathbb{R}^n)$ and acting on $L^{p_2}(\mathbb{R}^n)$ – do we know something about how the maps acts on $L^p(\mathbb{R}^n)$. Under certain conditions yes.

As a first step observe that any L^p-map can be decomposed into an L^{p_1}-map and L^{p_2}-map.

Lemma 4.10. *Assume $1 \leq p_1 \leq p \leq p_2 \leq \infty$ and assume $f \in L^p(\mathbb{R}^n)$.*

For each fixed $\lambda > 0$ there exists f_1, f_2 with the following conditions

- *$f_i \in L^{p_i}(\mathbb{R}^n)$, $i = 1, 2$.*
- *$f = f_1 + f_2$ a.e. in $\mathbb{R}^n$.*

Moreover we have

$$\|f_i\|_{L^{p_1}(\mathbb{R}^n)} \leq \lambda^{1-\frac{p}{p_i}} \|f\|_{L^p(\mathbb{R}^n)}^{\frac{p}{p_i}} \quad i = 1, 2.$$

f_i can be chosen explicitely:

$$f_1(x) := f(x)\chi_{|f|>\lambda} = \begin{cases} f(x) & \text{if } |f(x)| > \lambda \\ 0 & \text{if } |f(x)| \leq \lambda \end{cases}$$

and

$$f_2(x) := f(x)\chi_{|f|\le\lambda} = \begin{cases} f(x) & \text{if } |f(x)| \le \lambda \\ 0 & \text{if } |f(x)| > \lambda \end{cases}$$

Proof. We may assume $p_1 < p < p_2$ otherwise there is nothing to show.

Observe that since $\{f > \lambda\}$ etc. are measurable sets, so f_1 and f_2 are measurable functions, and we have $f = f_1 + f_2$ a.e..

It remains to compute (and here is where we use $p_1 < p < p_2$)

$$\begin{aligned}
\|f_1\|_{L^{p_1}(\mathbb{R}^n)} &= \left(\int_{\{|f|>\lambda\}} |f(x)|^{p_1} dx\right)^{\frac{1}{p_1}} \\
&= \lambda^{1-\frac{p}{p_1}} \left(\int_{\{|f|>\lambda\}} \lambda^{p-p_1} |f(x)|^{p_1} dx\right)^{\frac{1}{p_1}} \\
&\le \lambda^{1-\frac{p}{p_1}} \left(\int_{\{|f|>\lambda\}} |f(x)|^{p} dx\right)^{\frac{1}{p_1}} \\
&\le \lambda^{1-\frac{p}{p_1}} \left(\int_{\mathbb{R}^n} |f(x)|^{p} dx\right)^{\frac{1}{p_1}} \\
&= \lambda^{1-\frac{p}{p_1}} \|f\|_{L^p(\mathbb{R}^n)}^{\frac{p}{p_1}}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\|f_1\|_{L^{p_2}(\mathbb{R}^n)} &= \left(\int_{\{|f|\le\lambda\}} |f(x)|^{p_2} dx\right)^{\frac{1}{p_2}} \\
&\le \left(\int_{\{|f|\le\lambda\}} \lambda^{p_2-p} |f(x)|^{p} dx\right)^{\frac{1}{p_2}} \\
&\le \lambda^{1-\frac{p}{p_2}} \left(\int_{\mathbb{R}^n} |f(x)|^{p} dx\right)^{\frac{1}{p_2}} \\
&= \lambda^{1-\frac{p}{p_2}} \|f\|_{L^p(\mathbb{R}^n)}^{\frac{p}{p_2}}.
\end{aligned}$$

□

Now assume T is a linear operator from $L^{p_i}(\mathbb{R}^n)$ to $L^{q_i}(\mathbb{R}^n)$ for $i = 1, 2$. That is, for $i = 1, 2$ if $f, g \in L^{p_i}(\mathbb{R}^n)$ and $\lambda, \mu \in \mathbb{R}$ then

$$T(\lambda f + \mu g) = \lambda T f + \mu T g.$$

(this definition assumes that for $f \in L^{p_1} \cap L^{p_2}(\mathbb{R}^n)$ the Tf is well-defined).

Then if $p \in (p_1, p_2)$ the operator T naturally is defined for $f \in L^p(\mathbb{R}^n)$. Namely if we split $f = f_1 + f_2$ where $f_i \in L^{p_i}(\mathbb{R}^n)$ (we can always do that in view of Lemma 4.10) we set

$$Tf := Tf_1 + Tf_2.$$

Now assume that $f = \tilde{f}_1 + \tilde{f}_2$ is another decomposition but still $\tilde{f}_i \in L^{p_i}(\mathbb{R}^n)$, $i = 1, 2$. To make the operator well-defined we need to ensure that

$$T\tilde{f}_1 + T\tilde{f}_2 = Tf_1 + Tf_2 \quad a.e. \text{ in } \mathbb{R}^n. \tag{4.1}$$

Equivalently, we need to show

$$Tf_1 - T\tilde{f}_1 = T\tilde{f}_2 - Tf_2.$$

Observe that

$$f_1 - \tilde{f}_1 = f - f_2 - (f - \tilde{f}_2) = \tilde{f}_2 - f_2.$$

In particular, $f_1 - \tilde{f}_1 \in L^{p_1} \cap L^{p_2}$.

So, by linearity,

$$Tf_1 - T\tilde{f}_1 = T(f_1 - \tilde{f}_1) = T(\tilde{f}_2 - f_2) = T\tilde{f}_2 - Tf_2.$$

This establishes (4.1), that is Tf is well defined for each $f \in L^p(\mathbb{R}^n)$ as long as $p \in (p_1, p_2)$. There is more, we get also an estimate on $\|Tf\|_{L^p(\mathbb{R}^n)}$.

Theorem 4.11 (Marcienkiewicz Interpolation Theorem ("diagonal"))**.** *Let $1 \le p_1 < p_2 \le \infty$*[18]*.*

Assume that T is a bounded linear operator from $L^{p_i}(\mathbb{R}^n)$ to $L^{p_i}(\mathbb{R}^n)$ for $i = 1, 2$. That is, for $i = 1, 2$

- *There exists $\Lambda_i > 0$ such that for any $f \in L^{p_i}(\mathbb{R}^n)$ we have $Tf \in L^{p_i}(\mathbb{R}^n)$ and moreover*
$$\|Tf\|_{L^{p_i}(\mathbb{R}^n)} \le \Lambda_i \|f\|_{L^{p_i}(\mathbb{R}^n)}.$$
- *if $f, g \in L^{p_i}(\mathbb{R}^n)$ and $\lambda, \mu \in \mathbb{R}$ then*
$$T(\lambda f + \mu g) = \lambda Tf + \mu Tg.$$

Then for each $p \in (p_1, p_2)$ the operator T is a bounded linear operator from $L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$, and we have

$$\|Tf\|_{L^p(\mathbb{R}^n)} \le \Lambda_p \|f\|_{L^p(\mathbb{R}^n)}$$

where $\Lambda_q$is a constant depending only on Λ_1 and Λ_2 and θ.

Remark 4.12.
- The constant Λ_p here is not sharp in general, a technique called *complex interpolation* gives a sharper estimate
- We don't need linearity for T, only sublinearity, i.e. $Tf \le Tf_1 + Tf_2$.
- Of course this works in much more generality.

[18] diagonal refers to the fact that $q_i = p_i$

Before we come to the proof of Theorem 4.11 we record a crucial tool, an amazing application of Fubini's theorem.

Proposition 4.13. *Fix $p \in [1, \infty)$ and $\Omega \subset \mathbb{R}^n$ be an open set.*

Let $f : \Omega \to \overline{\mathbb{R}}$ be $\mathcal{L}^n$-measurable. Then $f \in L^p(\Omega)$ if and only if

$$\lambda \mapsto \lambda^{p-1}\mathcal{L}^n\left(\{x \in \Omega : |f(x)| > \lambda\}\right) \quad \text{in } L^1((0, \infty)).$$

Moreover, in either case

$$\int_\Omega |f(x)|^p dx = p \int_{\lambda=0}^\infty \lambda^{p-1}\mathcal{L}^n\left(\{x \in \Omega : |f(x)| > \lambda\}\right) d\lambda.$$

Proof. Since $|f|$ is $\mathcal{L}^n$-measurable, also

$$\lambda \mapsto \lambda^{p-1}\mathcal{L}^n\left(\{x \in \Omega : |f(x)| > \lambda\}\right)$$

is $\mathcal{L}^1$-measurable (exercise).

So all we have to obtain is the identity.

$$\begin{aligned} &p \int_{\lambda=0}^\infty \lambda^{p-1}\mathcal{L}^n\left(\{x \in \Omega : |f(x)| > \lambda\}\right) d\lambda \\ =&p \int_{\lambda=0}^\infty \lambda^{p-1} \int_\Omega \chi_{\{|f(\cdot)|>\lambda\}}(x) dx\, d\lambda \end{aligned}$$

Now we use Fubini's theorem (rather: Tonelli), Theorem 4.4. One easily checks that $\lambda \times x \mapsto \lambda^{p-1}\chi_{\{|f(\cdot)|>\lambda\}}(x)$ is $\mathcal{L}^1 \times \mathcal{L}^n$-measurable (and nonnegative). So,

$$\begin{aligned} =&p \int_\Omega \int_{\lambda=0}^\infty \lambda^{p-1}\chi_{\{|f(\cdot)|>\lambda\}}(x) d\lambda\, dx \\ =&p \int_\Omega \int_{\lambda=0}^{|f(x)|} \lambda^{p-1} d\lambda\, dx \\ =&\int_\Omega \lambda^p \Big|_{\lambda=0}^{|f(x)|} dx \\ =&\int_\Omega |f(x)|^p\, dx. \end{aligned}$$

□

Proof of Theorem 4.11. Observe that $p \in (1, \infty)$.

Fix $f \in L^p(\mathbb{R}^n)$. For a fixed $\lambda > 0$ we apply Lemma 4.10 and split $f = f_1 + f_2$. Then $Tf = Tf_1 + Tf_2$ which is measurable.

Since $|Tf| \leq |Tf_1| + |Tf_2|$,

$$\{|Tf| > \lambda\} \subset \{|Tf_1| > \frac{\lambda}{2}\} \cup \{|Tf_2| > \frac{\lambda}{2}\}$$

so

$$\mathcal{L}^n\left(\{|Tf|>\lambda\}\right) \leq \mathcal{L}^n\left(\{|Tf_1|>\frac{\lambda}{2}\}\right) + \mathcal{L}^n\left(\{|Tf_2|>\frac{\lambda}{2}\}\right).$$

By Chebyshev inequality, Lemma 3.49, for $i=1,2$ and boundedness of T,

$$\begin{aligned}\mathcal{L}^n\left(\{|Tf_i|>\frac{\lambda}{2}\}\right) &\leq \left(\frac{\lambda}{2}\right)^{-p_i}\|Tf_i\|^{p_i}_{L^{p_i}(\mathbb{R}^n)}\\ &\leq \left(\frac{\lambda}{2}\right)^{-p_i}\|f_i\|^{p_i}_{L^{p_i}(\mathbb{R}^n)}\end{aligned}$$

By the definition of f_i we have

$$\left(\frac{\lambda}{2}\right)^{-p_1}\|f_1\|^{p_1}_{L^{p_1}(\mathbb{R}^n)} = \left(\frac{\lambda}{2}\right)^{-p_1}\int_{\{|f(\cdot)|>\lambda\}}|f(x)|^{p_1}dx$$

$$\left(\frac{\lambda}{2}\right)^{-p_2}\|f_2\|^{p_2}_{L^{p_2}(\mathbb{R}^n)} = \left(\frac{\lambda}{2}\right)^{-p_2}\int_{\{|f(\cdot)|\leq\lambda\}}|f(x)|^{p_2}dx$$

And thus

$$\mathcal{L}^n\left(\{|Tf|>\lambda\}\right) \leq \left(\frac{\lambda}{2}\right)^{-p_1}\int_{\{|f(\cdot)|>\lambda\}}|f(x)|^{p_1}dx + \left(\frac{\lambda}{2}\right)^{-p_2}\int_{\{|f(\cdot)|\leq\lambda\}}|f(x)|^{p_2}dx$$

This estimate holds for any $\lambda>0$. In view of Proposition 4.13 we then have

$$\begin{aligned}\|Tf\|^p_{L^p(\mathbb{R}^n)} \leq& p\int_{\lambda=0}^{\infty}\lambda^{p-1}\mathcal{L}^n\left(\{x\in\Omega : |Tf(x)|>\lambda\}\right)d\lambda\\ \leq& C\int_{\lambda=0}^{\infty}\lambda^{p-1}\left(\lambda^{-p_1}\int_{\{|f(\cdot)|>\lambda\}}|f(x)|^{p_1}dx + \lambda^{-p_2}\int_{\{|f(\cdot)|\leq\lambda\}}|f(x)|^{p_2}dx\right)\,d\lambda\end{aligned}$$

Again by Fubini

$$\begin{aligned}\|Tf\|^p_{L^p(\mathbb{R}^n)} \leq& C\int_{\mathbb{R}^n}|f(x)|^{p_1}\left(\int_{\lambda=0}^{|f(x)|}\lambda^{p-1}\lambda^{-p_1}d\lambda\right)dx\\ &+ C\int_{\mathbb{R}^n}|f(x)|^{p_2}\left(\int_{\lambda=|f(x)|}^{\infty}\lambda^{p-1}\lambda^{-p_2}d\lambda\right)dx\end{aligned}$$

Since $p_1<p<p_2$ the integrals in λ converge and we have

$$\begin{aligned}\|Tf\|^p_{L^p(\mathbb{R}^n)} \leq& C\int_{\mathbb{R}^n}|f(x)|^{p_1}\left(|f(x)|^p|f(x)|^{-p_1}\right)dx\\ &+ C\int_{\mathbb{R}^n}|f(x)|^{p_2}\left(|f(x)|^p|f(x)|^{-p_2}d\lambda\right)dx\\ =&\tilde{C}\int_{\mathbb{R}^n}|f(x)|^p dx\\ =&\tilde{C}\|f\|^p_{L^p(\mathbb{R}^n)}.\end{aligned}$$

We can conclude. □

Exercise 4.14. *Prove Theorem 4.11 under the weakened assumptions*

$$\|Tf\|_{L^{p_i}(\mathbb{R}^n)} \leq \Lambda_i \|f\|_{L^{p_i,\infty}(\mathbb{R}^n)}.$$

(Cf. Definition 3.50)

The ideas of the Marcienkiewicz-interpolation theorem can be vastly generalized. As particular version we record (without proof) the following "off-diagonal" version of Theorem 4.11. "off-diagonal" means that the L^p-spaces in the domain and target may not be the same. For a proof see [Grafakos, 2014, §1.4.4].

Theorem 4.15 (Marcienkiewicz Interpolation Theorem (off-diagonal))**.** *Let $1 \leq p_1 < p_2 \leq \infty$ and $1 \leq q_1 < q_2 \leq \infty$.*

Assume that T is a bounded linear operator from $L^{p_i}(\mathbb{R}^n)$ to $L^{q_i}(\mathbb{R}^n)$ for $i = 1, 2$. That is, for $i = 1, 2$

- *There exists $\Lambda_i > 0$ such that for any $f \in L^{p_i}(\mathbb{R}^n)$ we have $Tf \in L^{q_i}(\mathbb{R}^n)$ and moreover*
$$\|Tf\|_{L^{q_i}(\mathbb{R}^n)} \leq \Lambda_i \|f\|_{L^{p_i}(\mathbb{R}^n)}.$$
- *if $f, g \in L^{p_i}(\mathbb{R}^n)$ and $\lambda, \mu \in \mathbb{R}$ then*
$$T(\lambda f + \mu g) = \lambda Tf + \mu Tg.$$

Then for each $p \in (p_1, p_2)$, i.e. whenever $\theta \in (0, 1)$ and

$$p = (1 - \theta)p_1 + \theta p_2,$$

then for

$$q := (1 - \theta)q_1 + \theta q_2$$

we that T is a bounded linear operator from $L^p(\mathbb{R}^n) \to L^q(\mathbb{R}^n)$, and we have

$$\|Tf\|_{L^q(\mathbb{R}^n)} \leq \Lambda_q \|f\|_{L^p(\mathbb{R}^n)}$$

where Λ_q is a constant depending only on Λ_1 and Λ_2 and θ.

4.2. **Application: convolution.** Let us start with an observation that since the Lebesgue measure in invariant under translations and under the mapping $x \mapsto -x$ for any $f \in L^1(\mathbb{R}^n)$ and any $y \in \mathbb{R}^n$ we have

$$\int_{\mathbb{R}^n} u(x)dx = \int_{\mathbb{R}^n} u(x+y)dx = \int_{\mathbb{R}^n} u(-x)dx.$$

Also, since for every measurable set E and any $t > 0$ the set $tE = \{tE : x \in E\}$ has measure $\mathcal{L}^n(tE) = t^n \mathcal{L}^n(E)$, Theorem 1.81, we conclude that the Lebesgue integral has the following scaling property

$$\int_{\mathbb{R}^n} u(x/t)dx = t^n \int_{\mathbb{R}^n} u(x)dx \quad \forall t > 0.$$

Observe that in the case of the Riemann integral the above equalities are direct consequences of the change of variables formula. We will prove a corresponding change of

variables formula for the Lebesgue integral later, but as for the proof of the above equalities we do not have to refer to the general change of variables formula as they follow directly from the properties of the Lebesgue measure mentioned above.

Definition 4.16. For measurable functions f and g on $\mathbb{R}^n$ we define the *convolution* by

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

The convolution plays a crucial role in Analysis, in particular in Partial Differential equations,

- convolutions can be used to "mollify functions" (approximating non-differentiable functions by differentiable ones, we will discuss this below)
- representing solutions to linear differential equations, e.g. for $n \geq 3$

  $$\Delta u(x) = f(x) \quad \text{in } \mathbb{R}^n$$

  is (under suitable assumptions on f and u equivalent) to

  $$u(x) = c \int_{\mathbb{R}^n} |x-y|^{2-n} f(y)\, dy.$$

  (where c is a suitable constant)

To start with our analysis, the first question is under what conditions the convolution is well defined. If $f \in L^1(\mathbb{R}^n)$ and g is bounded, measurable and vanishes outside a bounded set, then the function $y \mapsto f(x-y)g(y)$ is integrable, so $(f*g)(x)$ is well defined and finite for every $x \in \mathbb{R}^n$. If $f, g \in L^1(\mathbb{R}^n)$, then it can happen that for a given x the function $y \mapsto f(x-y)g(y)$ is not integrable and hence $(f*g)(x)$ is not defined. However as a powerful application of the Fubini theorem we can prove the following surprising result.

Theorem 4.17. *If $f, g \in L^1(\mathbb{R}^n)$ then for a.e. $x \in \mathbb{R}^n$ the function $y \mapsto f(x-y)g(y)$ is integrable and hence $f * g(x)$ exists. Moreover $f * g \in L^1(\mathbb{R}^n)$ and*

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

Proof. The function $|f(x-y)g(y)|$ as a function of a variable $(x,y) \in \mathbb{R}^{2n}$ is measurable (because $(x,y) \mapsto f(x-y)$ and $(x,y) \mapsto g(y)$ are both $\mathcal{L}^2$-measurable). Hence Fubini's theorem (rather: Tonelli Theorem 4.4), implies

$$\begin{aligned}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(x-y)g(y)| dx\, dy &= \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} |f(x-y)g(y)| dx\right) dy \\ &= \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} |f(x-y)| dx\right) |g(y)| dy = \|f\|_{L^1(\mathbb{R}^n)} \|g\|_{L^1(\mathbb{R}^n)}.\end{aligned}$$

Therefore

$$\|f*g\|_{L^1(\mathbb{R}^n)} \leq \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(x-y)g(y)| dx\, dy \leq \|f\|_{L^1(\mathbb{R}^n)} \|g\|_{L^1(\mathbb{R}^n)}.$$

□

Theorem 4.17 is a special case of *Young's convolution inequality*

Theorem 4.18 (Young's convolution inequality)**.** *Let $p, q, r \in [1, \infty]$ satisfy the following relation*

$$1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}$$

(where $\frac{1}{\infty} = 0$).

If $f \in L^p(\mathbb{R}^n)$, $g \in L^q(\mathbb{R}^n)$ then $f * g \in L^r(\mathbb{R}^n)$ and

$$\|f * g\|_{L^r(\mathbb{R}^n)} \leq \|f\|_{L^p(\mathbb{R}^n)} \|g\|_{L^q(\mathbb{R}^n)}.$$

Exercise 4.19. *Proof the statement for $p = \infty$ or $q = \infty$ and $r = \infty$ (what does each imply for the other coefficients?).*

Proof. We consider only the case $1 < p, q, r < \infty$.

Set $s := \frac{p}{1-\frac{p}{r}}$, $t = \frac{q}{1-\frac{q}{r}}$. Observe that then

$$\frac{1}{r} + \frac{1}{s} + \frac{1}{t} = \frac{1}{r} + \left(\frac{1}{p} - \frac{1}{r}\right) + \left(\frac{1}{q} - \frac{1}{r}\right) = 1.$$

We write

$$|f(x-y)g(y)| = \underbrace{|f(x-y)|^{1-\frac{p}{r}}}_{\in L^s(dy)} \underbrace{|g(y)|^{1-\frac{q}{r}}}_{\in L^t(dy)} \underbrace{|f(x-y)|^{\frac{p}{r}} |g(y)|^{\frac{q}{r}}}_{L^r(dy)}.$$

From (generalized) Hölder's inequality we then have

$$\begin{aligned}\int_{\mathbb{R}^n} |f(x-y)g(y)|\, dy \leq &\left(\int_{\mathbb{R}^n} |f(x-y)|^p dy\right)^{\frac{1-\frac{p}{r}}{p}} \left(\int_{\mathbb{R}^n} |g(y)|^q dy\right)^{\frac{1-\frac{q}{r}}{q}} \left(\int_{\mathbb{R}^n} |f(x-y)|^p |g(y)|^q dy\right)^{\frac{1}{r}} \\ =&\|f\|_{L^p(\mathbb{R}^n)}^{1-\frac{p}{r}} \|g\|_{L^q(\mathbb{R}^n)}^{1-\frac{q}{r}} \left(|f|^p * |g|^q\right)^{\frac{1}{r}}\end{aligned}$$

Now from the L^1-case,

$$\begin{aligned}\| \left(|f|^p * |g|^q\right)^{\frac{1}{r}} \|_{L^r(\mathbb{R}^n)} =&\||f|^p * |g|^q\|_{L^1(\mathbb{R}^n)}^{\frac{1}{r}} \leq \left(\||f|^p\|_{L^1(\mathbb{R}^n)} \||g|^q\|_{L^1(\mathbb{R}^n)}\right)^{\frac{1}{r}} \\ =&\|f\|_{L^p(\mathbb{R}^n)}^{\frac{p}{r}} \|g\|_{L^q(\mathbb{R}^n)}^{\frac{q}{r}}\end{aligned}$$

Plugging all this together we find

$$\begin{aligned}\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)g(y)|\, dy\right)^r dx &\leq \left(\|f\|_{L^p(\mathbb{R}^n)}^{1-\frac{p}{r}} \|g\|_{L^q(\mathbb{R}^n)}^{1-\frac{q}{r}} \| \left(|f|^p * |g|^q\right)^{\frac{1}{r}} \|_{L^r(\mathbb{R}^n)}\right)^r \\ &\leq \left(\|f\|_{L^p(\mathbb{R}^n)}^{1-\frac{p}{r}} \|g\|_{L^q(\mathbb{R}^n)}^{r-\frac{q}{r}} \|f\|_{L^p(\mathbb{R}^n)}^{\frac{p}{r}} \|g\|_{L^q(\mathbb{R}^n)}^{\frac{q}{r}}\right)^r \\ &= \left(\|f\|_{L^p(\mathbb{R}^n)} \|g\|_{L^q(\mathbb{R}^n)}\right)^r\end{aligned}$$

Thus,

$$\|f * g\|_{L^r(\mathbb{R}^n)}^r \leq \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)g(y)|\, dy\right)^r dx \leq \left(\|f\|_{L^p(\mathbb{R}^n)} \|g\|_{L^q(\mathbb{R}^n)}\right)^r.$$

Taking this inequality to the power $()^{\frac{1}{r}}$ we conclude. $\square$

Now let us look at algebraic properties of the convolution

Exercise 4.20. *$(L^1(\mathbb{R}^n), *)$ is a commutative algebra. Namely, if $\alpha, \beta \in \mathbb{R}$, $f, g, h \in L^1(\mathbb{R}^n)$ then*

(1) $f * g = g * f$
(2) $f * (g * h) = (f * g) * h$
(3) $f * (\alpha g + \beta h) = \alpha f * g + \beta f * h$.

Hint: *Fubini*

Let us also record

Exercise 4.21. *Let $f, g \in L^1(\mathbb{R}^n)$ and $h \in C_c^\infty(\mathbb{R}^n)$ then*

$$\int_{\mathbb{R}^n} (f * g)\ h = \int_{\mathbb{R}^n} f\ (\bar{g} * h)$$

where

$$\bar{g}(x) := g(-x).$$

Hint: *Fubini*

So, nice algebraic structure, nice! But now lets move towards serious business, namely "mollification".

For the "mollification aspect" we need a further definition

Definition 4.22. (1) We say that $f \in L^p_{loc}(\mathbb{R}^n)$ if for any $x \in \mathbb{R}^n$ there exists some ball $B(x, r)$ such that $f \in L^p(B(x, r))$.

(2) For $g \in C^0(\mathbb{R}^n)$ the *support* is "where the function lives", more precisely

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

A continuous function g is said to have *compact support* if $\operatorname{supp} g$ is compact.

(3) This can be generalized for $\mathcal{L}^n$-measurable $g : \mathbb{R}^n \to \overline{\mathbb{R}}$.

Fix such a g, and define the family of (measurable!) $\mathcal{S}$ as follows

$$S \in \mathcal{S} \quad :\Leftrightarrow S \text{ is closed and } g(x) = 0 \text{ for } \mathcal{L}^n\text{-a.e. } x \in \mathbb{R}^n \setminus S$$

$$\operatorname{supp} g := \bigcap_{S \in \mathcal{S}} S.$$

So a $\mathcal{L}^n$-measurable function is said to have *compact support* if and only if $\operatorname{supp} g$ is compact.

Exercise 4.23. *Show that whenever $f \in L^p_{loc}(\mathbb{R}^n)$ then $f \in L^q_{loc}(\mathbb{R}^n)$ for any $1 \leq q \leq p$*

Hint: *Hölder's inequality.*

Exercise 4.24. *Show that the two definitions of support in Definition 4.22 coincide for continuous functions.*

If $f \in L^1_{loc}(\mathbb{R}^n)$ and g is bounded, measurable and vanishes outside a bounded set, then the function $y \mapsto f(x-y)g(y)$ is integrable, so $(f*g)(x)$ is well defined and finite for every $x \in \mathbb{R}^n$. If g is better, then $f*g$ is as good as g.

Theorem 4.25. *If $f \in L^1_{loc}(\mathbb{R}^n)$ and g is continuous with compact support then*

(1) $f*g$ is continuous on $\mathbb{R}^n$
(2) If additionally $g \in C^k(\mathbb{R}^n)$ then $f*g \in C^k(\mathbb{R}^n)$, $k \in \mathbb{N}$, and we have

$$\partial_\alpha(f*g)(x) = (f*\partial_\alpha g)(x) \quad \alpha \text{ multiindex: } |\alpha| \le k.$$

Proof. (1) Since g is bounded and has compact support, $(f*g)(x)$ is well defined and finite for all $x \in \mathbb{R}^n$. Fix $x \in \mathbb{R}^n$. The function $y \mapsto f(y)g(x-y)$ vanishes outside a sufficiently large ball B (because g as compact support). Let $x_n \to x$. Then there is a ball B (perhaps larger than the one above), so that all the functions

$$y \mapsto f(y)g(x_n - y)$$

vanish outside B. Hence

$$|f(y)g(x_n-y)| \le \|g\|_{L^\infty} |f(y)| \chi_B(y) \in L^1(\mathbb{R}^n)$$

and the dominated convergence theorem, Theorem 3.26, yields

$$(f*g)(x_n) = \int_{\mathbb{R}^n} f(y)g(x_n-y)dy \to \int_{\mathbb{R}^n} f(y)g(x-y)dy = (f*g)(x)$$

which proves continuity of $f*g$.

(2) Pick $i \in \{1,\dots,n\}$ and $|h| \le 1$ and fix $x \in \mathbb{R}^n$

We have (observe all terms exists since $g \in C^1(\mathbb{R}^n)$ and $\partial_i g$ still has compact support)

$$\begin{aligned}&(f*g)(x+he_i) - (f*g)(x) - (f*\partial_i g)(x)h \\ =&\int_{\mathbb{R}^n} f(y)\left(g(x+he_i-y) - g(x-y) - \partial_i g(x-y)h\right)dy\end{aligned}$$

Observe that if $\operatorname{supp} g$ is compact then $\operatorname{supp} g(z-\cdot)$ is still compact (for a fixed z). It is easy to check that the following set is compact

$$K_x := \bigcup_{|z|\le 1} \operatorname{supp} g(x+z-\cdot).$$

Moreover,

$$g(x+he_i-y) - g(x-y) - \partial_i g(x-y)h = 0 \quad y \notin K_x.$$

Since g is continuously differentiable on $\mathbb{R}^n$ we have

$$|g(x+he_i-y) - g(x-y) - \partial_i g(x-y)h| \le C_x|h| \quad \forall y \in \mathbb{R}^n, |h| \le 1.$$

and moreover by differentiability for every y

$$\frac{|g(x+he_i-y)-g(x-y)-\partial_i g(x-y)h|}{|h|} \xrightarrow{|h|\to 0} 0.$$

Thus

$$\frac{(f*g)(x+he_i)-(f*g)(x)-(f*\partial_i g)(x)h}{|h|}$$
$$=\int_{\mathbb{R}^n} f(y)\frac{g(x+he_i-y)-g(x-y)-\partial_i g(x-y)h}{|h|}dy$$

and

$$f(y)\frac{g(x+he_i-y)-g(x-y)-\partial_i g(x-y)h}{|h|} \xrightarrow{|h|\to 0} 0\text{pointwise for a.e. } y$$

and

$$|f(y)\frac{g(x+he_i-y)-g(x-y)-\partial_i g(x-y)h}{|h|}| \le C|f(y)|$$

We conclude by the dominated convergence theorem that

$$\frac{(f*g)(x+he_i)-(f*g)(x)-(f*\partial_i g)(x)h}{|h|} \xrightarrow{|h|\to 0} 0.$$

Thus $f*g$ has a partial derivative at every x. That by itself is not enough to conclude differentiability! However, since morevoer $f*\partial_i g$ is continuous (by (1)) we can conclude that $f*g$ is continuously differentiable everywhere. Repeating this argument for higher derivatives gives the claim.

□

Let $\eta \in C_c^\infty(B(0,1))$ be a function such that $\eta \ge 0$ and $\int_{\mathbb{R}^n}\eta(x)=1$. E.g. take

$$\tilde{\eta}(x) := \begin{cases} e^{\frac{1}{|x|^2-1}} & |x|\le 1 \\ 0 & |x|\ge 1 \end{cases} \tag{4.2}$$

One can show that $\tilde{\eta}\in C_c^\infty(B(0,1))$, and we set $\eta := \left(\int_{\mathbb{R}^n}\tilde{\eta}\right)^{-1}\tilde{\eta}$. η is often called a *bump function* (because thats what it is) or a *mollifier* (reason: below).

For $\varepsilon>0$ we set

$$\eta_\varepsilon(x) := \varepsilon^{-n}\eta(x/\varepsilon)$$

Exercise 4.26. *Show that* $\operatorname{supp}\eta_\varepsilon \subset B(0,\varepsilon)$, $\eta_\varepsilon \ge 0$ *and* $\int_{\mathbb{R}^n}\eta_\varepsilon = 1$.

For $f\in L^1_{loc}(\mathbb{R}^n)$ we set

$$f_\varepsilon := f*\eta_\varepsilon.$$

In view of Theorem 4.25, $f_\varepsilon \in C^\infty(\mathbb{R}^n)$. This is called the *mollification* of f.

Observe that as $\varepsilon \to 0$ we have in a very handwaving sense

$$\eta_\varepsilon(x) \xrightarrow{\varepsilon\to 0} \begin{cases} 0 & x \neq 0 \\ \infty \cdot \text{“}\eta(0)'' & x = 0 \end{cases}$$

More precisely we have for any measurable $A \subset \mathbb{R}^n$

$$\mathcal{L}^n \llcorner \eta_\varepsilon \xrightarrow{*} \eta(0)\delta_0$$

where δ_0 is the *Dirac measure*

$$\delta_0(A) = \begin{cases} 1 & 0 \in A \\ 0 & \text{otherwise.} \end{cases}$$

Above, (weak*) convergence is understood in the following sense:

Lemma 4.27. *For any $f \in C^0(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ we have*

$$\int_{\mathbb{R}^n} f d\left(\mathcal{L}^n \llcorner \eta_\varepsilon\right) \xrightarrow{\varepsilon\to 0} \int_{\mathbb{R}^n} f d\delta_0 = f(0).$$

Proof. We have

$$\int_{\mathbb{R}^n} f d\left(\mathcal{L}^n \llcorner \eta_\varepsilon\right) = \int_{\mathbb{R}^n} f(x)\eta_\varepsilon(x)\, d\mathcal{L}^n.$$

So all we need to show is that

$$\int_{\mathbb{R}^n} f(x)\eta_\varepsilon(x)\, d\mathcal{L}^n \xrightarrow{\varepsilon\to 0} f(0).$$

Now observe since $\int \eta_\varepsilon = 1$ we have

$$\int_{\mathbb{R}^n} f(x)\eta_\varepsilon(x)dx - f(0) = \int_{\mathbb{R}^n} \left(f(x) - f(0)\right)\eta_\varepsilon(x)dx.$$

Since f is continuous at 0, for any $\delta > 0$ there exists a radius $r > 0$ such that $|f(x) - f(0)| \leq \delta$ for all $x \in B(0, r)$. If we take $\varepsilon < r$ we then have

$$\left|\int_{\mathbb{R}^n} \left(f(x) - f(0)\right)\eta_\varepsilon(x)dx\right| \leq \delta \int_{\mathbb{R}^n} \eta_\varepsilon(x)dx = \delta.$$

This holds for any δ and $\varepsilon < r(\delta)$ so we have

$$\lim_{\varepsilon\to 0}\left|\int_{\mathbb{R}^n} f(x)\eta_\varepsilon(x)dx - f(0)\right| = 0.$$

□

What does this mean for $f_\varepsilon(x) = f * \eta_\varepsilon(x) = \int f(x-y)\eta_\varepsilon(y)$? It converges to $f(x)$. And it does so quite nicely.

Theorem 4.28. *Let $f : \mathbb{R}^n \to \mathbb{R}$.*

(1) If f is continuous then f_ε converges uniformly to f on compact sets as $\varepsilon \to 0$, i.e. for all K compact

$$\|f - f_\varepsilon\|_{L^\infty(K)} \xrightarrow{\varepsilon\to 0} 0.$$

(2) *If $f \in C^k(\mathbb{R}^n)$ then $D^\alpha f_\varepsilon$ converges uniformly to $D^\alpha f$ on compact sets for any $|\alpha| \le r$ as $\varepsilon \to 0$, i.e. for all K compact*

$$\max_{|\alpha|\le k} \|D^\alpha f - D^\alpha f_\varepsilon\|_{L^\infty(K)} \xrightarrow{\varepsilon\to 0} 0.$$

Proof. (1) Uniform continuity of f on compact sets implies that for any compact set K

$$\sup_{|y|<\varepsilon} \sup_{x\in K} |f(x) - f(x-y)| \to 0 \quad \text{as } \varepsilon \to 0.$$

Since $\int_{\mathbb{R}^n} \eta_\varepsilon(y)dy = 1$ we have $\int_{\mathbb{R}^n} f(x)\eta_\varepsilon(y)dy = f(x)$ and hence for any compact set K,

$$\begin{aligned}\sup_{x\in K} |f(x) - f_\varepsilon(x)| &= \sup_{x\in K} \left| \int_{\mathbb{R}^n} (f(x) - f(x-y))\eta_\varepsilon(y)dy \right| \\ &\le \sup_{x\in K} \int_{B(0,\varepsilon)} |f(x) - f(x-y)|\ \eta_\varepsilon(y)dy \to 0 \quad \text{as } \varepsilon \to 0.\end{aligned}$$

(2) Follows from (1) by an induction argument and the fact tat $D^\alpha(f * \eta_\varepsilon) = (D^\alpha f) * \eta_\varepsilon$.

□

Exercise 4.29. *(1) If $f \in C^\alpha$ for $\alpha \in (0,1]$ show that f_ε converges to f in C^β for any $\beta < \alpha$, i.e.*

$$\sup_{x\ne y} \frac{|f(x) - f_\varepsilon(x) - f(y) - f_\varepsilon(y)|}{|x-y|^\beta} \xrightarrow{\varepsilon\to 0} 0.$$

(2) Show that this may not be true for $\beta = \alpha = 1$ (it is also not necessarily true for $\beta = \alpha < 1$)

Hint: Lipschitz functions may not be everywhere differentiable

Now we want to get convergence also in $L^p(\mathbb{R}^n)$.

Lemma 4.30. *If $f \in L^p(\mathbb{R}^n)$, $1 \le p \le \infty$ then $f_\varepsilon \in L^p(\mathbb{R}^n)$ and $\|f_\varepsilon\|_{L^p(\mathbb{R}^n)} \le \|f\|_{L^p(\mathbb{R}^n)}$.*

Proof. Follows from Theorem 4.18:

$$\|f_\varepsilon\|_{L^p(\mathbb{R}^n)} = \|f * \eta_\varepsilon\|_{L^p(\mathbb{R}^n)} \le \|f\|_{L^p(\mathbb{R}^n)} \underbrace{\|\eta_\varepsilon\|_{L^1(\mathbb{R}^n)}}_{=1}.$$

□

Theorem 4.31. *If $f \in L^p(\mathbb{R}^n)$ and $1 \le p < \infty$ then $f_\varepsilon \in L^p(\mathbb{R}^n)$ and $\|f - f_\varepsilon\|_{L^p(\mathbb{R}^n)} \xrightarrow{\varepsilon\to 0} 0$.*

For the proof of Theorem 4.31 we need the following lemma

Lemma 4.32. *If $f \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$ then*

$$\lim_{y\to 0} \int_{\mathbb{R}^n} |f(x+y) - f(x)|^p dx = 0.$$

Proof. For $y \in \mathbb{R}^n$ let $\tau_y : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$ be the translation operator defined by

$$(\tau_y f)(x) = f(x+y) \quad \text{for } f \in L^p(\mathbb{R}^n).$$

The lemma claims that τ_y is continuous, i.e. $\|\tau_y f - f\|_{L^p} \xrightarrow{y\to 0} 0$ as $y \to 0$. Given $\varepsilon > 0$ let g be a compactly supported continuous function such that $\|f-g\|_{L^p} < \varepsilon/3$, see Theorem 3.28. Then

$$\begin{aligned}\|\tau_y f - f\|_{L^p} \leq & \|\tau_y f - \tau_y g\|_{L^p} + \|\tau_y g - g\|_{L^p} + \|f-g\|_{L^p} = 2\|f-g\|_{L^p} + \|\tau_y g - g\|_{L^p} \\ < & 2\varepsilon/3 + \|\tau_y g - g\|_{L^p}\end{aligned}$$

Since $\tau_y g \to g$ converges uniformly, and $K := \bigcup_{|y|\leq 1} \operatorname{supp} \tau_y g$ is bounded, there is $\delta > 0$ such that

$$\|\tau_y g - g\|_{L^p(\mathbb{R}^n)} = \|\tau_y g - g\|_{L^p(K)} \overset{\text{Hölder}}{\leq} C(K)\, \|\tau_y g - g\|_{L^\infty(\mathbb{R}^n)} < \varepsilon/3 \quad \text{for } |y| < \delta.$$

Hence,

$$\|\tau_y f - f\|_{L^p} < \frac{\varepsilon}{3} \quad \text{for } |y| < \delta,$$

which proves the lemma. □

Proof of Theorem 4.31. Assume first that $1 < p < \infty$. Hölder's inequality and Fubini's theorem yield

$$\begin{aligned}\|f - f_\varepsilon\|_{L^p} &= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} (f(x) - f(x-y))\eta_\varepsilon(y)dy \right|^p dx \\ &\leq \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x) - f(x-y)|\eta_\varepsilon(y)^{1/p}\eta_\varepsilon(y)^{1/q} dy \right)^p dx \\ &\leq \int_{\mathbb{R}^n} (|f(x) - f(x-y)|^p \eta_\varepsilon(y)dy) \underbrace{\left(\int_{\mathbb{R}^n} \eta_\varepsilon(y)dy \right)^{\frac{p}{q}}}_{=1} dx \\ &= \int_{\mathbb{R}^n} \int_{B(0,\varepsilon)} |f(x) - f(x-y)|^p \eta_\varepsilon(y)dydx \\ &= \int_{B(0,\varepsilon)} \left(\int_{\mathbb{R}^n} |f(x) - f(x-y)|^p dx \right) \eta_\varepsilon(y)\, dy \xrightarrow{\varepsilon\to 0} 0.\end{aligned}$$

In the last step we used Lemma 4.32. If $p = 1$, then the above argument simplifies, because we do not have to use Hölder's inequality. □

Corollary 4.33. *Let $1 \leq p < \infty$. If $f \in L^p(\Omega)$ for an open set Ω then for any open compactly contained $\Omega' \subset\subset \Omega$ (i.e. Ω' is open, $\overline{\Omega'}$ is compact and $\overline{\Omega'} \subset \Omega$) we have $f_\varepsilon \in C^\infty(\overline{\Omega'})$ for all small $\varepsilon \ll 1$ and $\|f_\varepsilon - f\|_{L^p(\Omega')} \xrightarrow{\varepsilon\to 0} 0$.*

Proof. Set $d := \frac{\operatorname{dist}(\Omega', \partial\Omega)}{100} > 0$. Let $\bar{\eta} \in C_c^\infty(B(\Omega', 50d), [0,\infty))$ be a bump function with $\bar{\eta} \equiv 1$ in $B(\Omega', 20d)$[19]

[19]Here: $B(A,r) := \{x \in \mathbb{R}^n : \operatorname{dist}(x, A) < r\}$

Set

$$\bar{f} := \bar{\eta} f.$$

Clearly $\bar{f} \in L^p(\mathbb{R}^n)$, and by Theorem 4.31 we have that $\bar{f} \in C^\infty(\mathbb{R}^n)$ and

$$\|\bar{f}_\varepsilon - \bar{f}\|_{L^p(\mathbb{R}^n)} \xrightarrow{\varepsilon\to 0} = 0.$$

In particular $\bar{f}_\varepsilon \in C^\infty(\overline{\Omega'})$ and (using that $\bar{\eta} \equiv 1$ in Ω')

$$\|\bar{f}_\varepsilon - f\|_{L^p(\Omega')} = \|\bar{f}_\varepsilon - \bar{f}\|_{L^p(\Omega')} \xrightarrow{\varepsilon\to 0} = 0.$$

We now claim that for $\varepsilon < d$ we have $\bar{f}_\varepsilon = f_\varepsilon$ for $x \in \Omega'$. Indeed if $x \in \Omega'$ then denoting the mollification kernel for f_ε by $\eta \in C_c^\infty(B(0,1))$, $f_\varepsilon = f * \varepsilon^{-n}\eta(\cdot/\varepsilon)$ then $\operatorname{supp} \varepsilon^{-n}\eta((x-\cdot)/\varepsilon) \subset B(x,\varepsilon) \subset B(\Omega', d)$. Consequently, for any $x \in \Omega'$

$$\begin{aligned} f_\varepsilon(x) =& \varepsilon^{-n} \int_{\mathbb{R}^n} f(y)\eta((x-y)/\varepsilon)dy = \varepsilon^{-n} \int_{B(\Omega',d)} f(y)\eta((x-y)/\varepsilon)dy \\ =& \varepsilon^{-n} \int_{B(\Omega',d)} \bar{\eta}(y) f(y)\eta((x-y)/\varepsilon)dy = \bar{f}_\varepsilon(x). \end{aligned}$$

Thus

$$\|f_\varepsilon - f\|_{L^p(\Omega')} = \|\bar{f}_\varepsilon - f\|_{L^p(\Omega')} \xrightarrow{\varepsilon\to 0} = 0.$$

The proof is complete. □

One can use convolution to obtain another proof of density, see Theorem 3.32(3). See Lemma 4.38 for a version for any open set $\Omega \subset \mathbb{R}^n$.

Exercise 4.34. *Use convolution to give another proof of Theorem 3.32(3) on $\mathbb{R}^n$:*

Namely show that $C_c^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for all $1 \le p < \infty$. That is for any $f \in L^p(\mathbb{R}^n)$ there exists $(g_k)_{k\in\mathbb{N}} \in C_c^\infty(\mathbb{R}^n)$ such that

$$\|g_k - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

(1) Approximate f by convolution with some $f_\varepsilon \in C^\infty(\mathbb{R}^n)$ (but there is no reason that $f_\varepsilon \in C_c^\infty(\mathbb{R}^n)$)

(2) Take $\bar{\eta} \in C_c^\infty(B(0,2))$, $\bar{\eta} \equiv 1$ in $B(0,1)$ and $\|\bar{\eta}\|_{L^\infty(\mathbb{R}^n)} \le 1$ a bump function and set

$$g_k(x) := \bar{\eta}(x/k)\, f_{\frac{1}{k}}(x).$$

(3) Show that

$$\|\bar{\eta}(x/k)\left(f_{1/k} - f\right)\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

(4) Show (Hint: dominated convergence!)

$$\|(1-\bar{\eta}(x/k))f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

(5) Conclude that

$$\|g_k - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

Exercise 4.35. *Show that Exercise 4.34 is false for $p = \infty$*

(Hint: continuous functions)

Exercise 4.36. *Let $p \in [1, \infty)$. Show that for any $f \in L^p(\mathbb{R}^n)$ with $f \geq 0$ a.e. there exists $f_k \in C_c^\infty(\mathbb{R}^n)$, $f_k \geq 0$ everywhere such that*

$$\|f_k - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

Proposition 4.37. *Let $p \in [1, \infty)$ and $f : \mathbb{R}^n \to \mathbb{R}$ be any function. Then the following are equivalent:*

(1) $f \in L^p(\mathbb{R}^n)$
(2) There exists $f_k \in C_c^\infty(\mathbb{R}^n)$ with $f_k \to f$ a.e. and

$$\sup_k \|f_k\|_{L^p(\mathbb{R}^n)} < \infty.$$

In either case we have

$$\|f\|_{L^p(\mathbb{R}^n)} = \inf_{(f_k)_k} \liminf_{k\to\infty} \|f_k\|_{L^p(\mathbb{R}^n)} \tag{4.3}$$

where the infimum is taken over all sequences $(f_k)_k$ as in (2).

Proof. $\Rightarrow$ This follows from Theorem 4.31. Since $f \in L^p(\mathbb{R}^n)$ we find approximations $f_k \in C_c^\infty(\mathbb{R}^n)$ and

$$\|f_k\|_{L^p(\mathbb{R}^n)} \leq \|f - f_k\|_{L^p(\mathbb{R}^n)} + \|f\|_{L^p(\mathbb{R}^n)}.$$

We also get one part of the inequality

$$\inf_{(\tilde{f}_k)_k} \liminf_{k\to\infty} \|\tilde{f}_k\|_{L^p(\mathbb{R}^n)} \leq \liminf_{k\to\infty} \|f_k\|_{L^p(\mathbb{R}^n)} \leq \liminf_{k\to\infty} \|f - f_k\|_{L^p(\mathbb{R}^n)} + \liminf_{k\to\infty} \|f\|_{L^p(\mathbb{R}^n)} = 0 + \|f\|_{L^p(\mathbb{R}^n)}.$$

$\Leftarrow$ Let f_k be as in (2). In particular f_k are $\mathcal{L}^n$-measurable, so as a pointwise limit f is $\mathcal{L}^n$-measurable, Theorem 2.13. In particular $|f|$ is $\mathcal{L}^n$-measurable.

We can apply Fatou's lemma, Corollary 3.9, to $|f_k|^p$ and by pointwise convergence

$$\int_{\mathbb{R}^n} |f|^p = \int_{\mathbb{R}^n} \lim_{k\to\infty} |f_k|^p \leq \liminf_{k\to\infty} \int_{\mathbb{R}^n} |f_k|^p.$$

Since $\int_{\mathbb{R}^n} |f|^p < \infty$ we have $f \in L^p(\mathbb{R}^n)$. The above inequality holds for any approximating sequence $(f_k)_k$ so we get the remaining part of (4.3), namely

$$\|f\|_{L^p(\mathbb{R}^n)} \leq \inf_{(f_k)_k} \liminf_{k\to\infty} \|f_k\|_{L^p(\mathbb{R}^n)}$$

□

One can also get another proof of Theorem 3.32(3) on any Ω.

Lemma 4.38. *Let $\Omega \subset \mathbb{R}^n$ open and $f \in L^p(\Omega)$. Then there exist $f_k \in C_c^\infty(\Omega)$ such that*

$$\|f_k - f\|_{L^p(\Omega)} \xrightarrow{k\to\infty} 0.$$

Proof. Extend f by zero outside of Ω. We first show that there exist $g_k \in L^p(\mathbb{R}^n)$ with $\operatorname{supp} g_k \subset \Omega$ and

$$\|f - g_k\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

Any open set can be decomposed into a countable union of closed dyadic cubes $\bigcup_{i=1}^{\infty} Q_i$ with pairwise disjoint interior. Let $g_k := \chi_{\bigcup_{i=1}^{k} Q_i} f$. By the dominated convergence theorem $\|g_k - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0$. Observe that $\operatorname{supp} g_k \subset \bigcup_{i=1}^{k} Q_i$ which is a closed subset inside Ω. Since Ω is open $\delta := \operatorname{dist}(\operatorname{supp} g_k, \mathbb{R}^n \backslash \Omega) > 0$.

Now set $f_{k,\sigma} := \eta_\sigma * g_k$, for $\sigma < \delta$. Fix $\varepsilon > 0$. Then there exists $\sigma > 0$ so small such that

$$\|\eta_\sigma * f - f\|_{L^p(\mathbb{R}^n)} \leq \frac{\varepsilon}{2}.$$

Also we can take $k \in \mathbb{N}$ (independent of σ actually) such that

$$\|\eta_\sigma * g_k - \eta_\sigma * f\| \leq \|g_k - f\| < \frac{\varepsilon}{2}.$$

Then we have $f_{k,\sigma} \in C_c^\infty(\Omega)$ and

$$\|f_{k,\sigma} - f\|_{L^p(\mathbb{R}^n)} \leq \|\eta_\sigma * g_k - \eta_\sigma * f\| + \|\eta_\sigma * f - f\| < \varepsilon.$$

□

Exercise 4.39 (Censored Mollification)**.** *Take three radii $0 < r < \rho < R$ and assume $u \in C^0(\overline{B(0,R)})$. Show that there is an approximation $u_k \in C^0(\overline{B(0,R)})$ with*

$$\|u_k - u\|_{L^\infty(B(0,R))} \xrightarrow{k\to\infty} 0$$

that satisfies the following conditions for all $k \in \mathbb{N}$

- *$u_k \equiv u$ in $B(0,R) \setminus B(0,\rho)$*
- *$u_k \in C^\infty(B(0,r))$*

For this use the following definition of a "censored" mollification

$$u_\delta(x) := \int_{\mathbb{R}^n} \eta(z) u(x + \delta\theta(x) z) dz,$$

for some choice of $\theta \in C_c^\infty(B(0,\rho),[0,1])$ and $\theta \equiv 1$ in $B(0,r)$, and a typical bump function $\eta \in C_c^\infty(B(0,2))$, $\eta \equiv 1$ in $B(0,1)$ and $\int \eta = 1$.

4.3. **A first glimpse on Sobolev spaces.** From the approximation Exercise 4.34 we have the following equivalent definition of $L^p(\mathbb{R}^n)$.

Proposition 4.40. *Let $p \in [1,\infty)$.*

Consider the space $\mathcal{L}^p(\mathbb{R}^n)$ as all $f \in C_c^\infty(\mathbb{R}^n)$ with the norm $\|f\|_{L^p(\mathbb{R}^n)}$. Every norm induces a metric $d_{\mathcal{L}^p}(f,g) := \|f - g\|_{L^p(\mathbb{R}^n)}$. Denote by $\tilde{L}^p(\mathbb{R}^n)$ the metric completion of $\mathcal{L}^p(\mathbb{R}^n)$ under the metric $d_{\mathcal{L}^p}(f,g)$.

Then $\tilde{L}^p(\mathbb{R}^n) = L^p(\mathbb{R}^n)$ in the following sense:

The uniformly continuous linear functional $\mathrm{id} : \mathcal{L}^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$ *given by* $\mathrm{id}\, f := f$ *extends to a (unique) isometric isomorphism* $\tilde{L}^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$.

Proof. Since (by definition) $\mathcal{L}^p(\mathbb{R}^n)$ is dense in $\tilde{L}^p(\mathbb{R}^n)$ and id is clearly Lipschitz continuous with Lipschitz constant 1 (and thus uniformly continuous) id extends uniquely to a continuous map $\mathrm{id} : \tilde{L}^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$, given by $\mathrm{id}\, f := \lim_{k\to\infty} \mathrm{id}\, f_k$ where f_k is any sequence in $\mathcal{L}^p(\mathbb{R}^n)$ converging to f.

Clearly id is injective and isometric. It only needs to be shown that it is onto. So let $f \in L^p(\mathbb{R}^n)$ then by Exercise 4.34 there exists $f_k \in L^p(\mathbb{R}^n)$ with $\|f_k - f\|_{L^p} \to 0$. But then f_k is a Cauchy sequence also in $\mathcal{L}^p$, and thus $\mathrm{id}\, f_k = f_k$ converges on both sides. □

Exercise 4.41. *Let* $p \in [1, \infty]$. *Show that* $(C_c^\infty(\mathbb{R}^n), \|\cdot\|_{W^{1,p}(\mathbb{R}^n)})$ *is a normed space where*

$$\|f\|_{W^{1,p}(\mathbb{R}^n)} := \|f\|_{L^p(\mathbb{R}^n)} + \|\nabla f\|_{L^p(\mathbb{R}^n)}.$$

This leads to a (first, there are several) definition of *Sobolev spaces* $W^{1,p}(\mathbb{R}^n)$.

Definition 4.42. Let $p \in [1, \infty)$. Consider $C_c^\infty(\mathbb{R}^n)$ equipped with the norm

$$\|f\|_{W^{1,p}(\mathbb{R}^n)} := \|f\|_{L^p(\mathbb{R}^n)} + \|\nabla f\|_{L^p(\mathbb{R}^n)}.$$

We denote $W^{1,p}(\mathbb{R}^n)$ as the metric completion of $C_c^\infty(\mathbb{R}^n)$ under this norm.

So what kind of functions are in this Sobolev space $W^{1,p}(\mathbb{R}^n)$? Since the $\|f\|_{L^p(\mathbb{R}^n)} \leq \|f\|_{W^{1,p}(\mathbb{R}^n)}$ we see that $W^{1,p}(\mathbb{R}^n)$ must be a subspace of $L^p(\mathbb{R}^n)$, i.e. a first equivalent notion is

Definition 4.43. Let $p \in [1, \infty)$. $f \in W^{1,p}(\mathbb{R}^n)$ if and only if $f \in L^p(\mathbb{R}^n)$ and there exists $f_k \in C_c^\infty(\mathbb{R}^n)$ such that f_k is a $W^{1,p}(\mathbb{R}^n)$-Cauchy sequence and

$$\|f_k - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

Alright, so what does that mean: $W^{1,p}$ consists of L^p-functions that satisfy an additional condition: their *distributional derivative* belongs also to L^p. That needs some explanation.

Definition 4.44. A *distribution* is a linear map $T : C_c^\infty(\mathbb{R}^n) \to \mathbb{R}$ (it should also be continuous, but we discuss this later, **??**).

The (distributional) derivative of a distribution T is given by

$$\partial_\alpha T(\varphi) := (-1)^\alpha T(\partial_\alpha \varphi), \quad \varphi \in C_c^\infty(\mathbb{R}^n).$$

Every L^1_{loc}*-function* f *is (equivalent) to a distribution:*

Usually we think of functions $f : \mathbb{R}^n \to \mathbb{R}$ as function evaluated at some point x.

If we know for all $x \in \mathbb{R}^n$ *what* $f(x)$ *looks like, then we know what* f *is*. More precisely, *we can test when to functions are the same:* $f = g$ *if* $f(x) = g(x)$ *for all* $x \in \mathbb{R}^n$

But we can evaluate functions differently, with other test-functions. For example, for $\varphi \in C_c^\infty(\mathbb{R}^n)$, we can consider f as *distribution*, i.e.

$$f[\varphi] := \langle f, \varphi \rangle := \int f(x)\varphi(x)dx$$

Still: *if we know for all $\varphi \in C_c^\infty(\mathbb{R}^n)$ what $\langle f, \varphi \rangle$ looks like, then we know what f is.*

More precisely *we can test when to functions are the same: $f = g$ if $\langle f, \varphi \rangle = \langle g, \varphi \rangle$ for all $\varphi \in C_c^\infty(\mathbb{R}^n)$*

Lemma 4.45 (Fundamental lemma of the Calculus of Variations)**.** *Let $f, g \in L^1_{loc}(\mathbb{R}^n)$ and assume that*

$$\langle f, \varphi \rangle = \langle g, \varphi \rangle \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n)$$

then $f = g$ almost everywhere.

In particular if

$$\langle f, \varphi \rangle = 0 \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n),$$

then $f = 0$ almost everywhere.

Proof. Since

$$\langle f, \varphi \rangle = \langle g, \varphi \rangle \quad \Leftrightarrow \quad \langle f - g, \varphi \rangle = 0,$$

we only need to show that if

$$\langle f, \varphi \rangle = 0 \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n),$$

then $f = 0$ a.e.

<u>First assume that f is continuous around a point $x_0 \in \mathbb{R}^n$</u>. If $f(x_0) \neq 0$ then w.l.o.g. $f(x_0) > 0$. By continuity there exists a small ball $B(x_0, r)$ such that $f(x) \geq \frac{1}{2} f(x_0)$ for all $x \in B(x_0, r)$.

Now let φ be a bump function in $B(x_0, r)$, i.e. $\varphi \in C_c^\infty(B(x_0, r))$, $\varphi \geq 0$ everywhere and $\inf_{B(x_0,r/2)} \varphi > 0$; e.g. $\varphi = \tilde{\varphi}((x - x_0)/r)$ where $\tilde{\varphi}$ is from (4.2).

Then

$$0 = \langle f, \varphi \rangle = \int_{B(x_0,r)} f\varphi \geq \frac{1}{2} f(x_0) \int_{B(x_0,r)} \varphi > 0,$$

a contradiction. So $f(x_0) = 0$. This holds for all $x_0 \in \mathbb{R}^n$ where f is continuous.

<u>General case:</u> In the general case we cannot argue with continuity – but we will argue with convolution.

Let η be a typical bump function, $\eta \in C_c^\infty(B(0, 1))$, $\eta \geq 0$, $\int \eta = 1$ and set $\eta_\varepsilon := \varepsilon^{-n}\eta(\cdot/\varepsilon)$. Fix $x \in \mathbb{R}^n$. Then the assumption imply

$$f_\varepsilon(x) = \langle f, \eta_\varepsilon(x - \cdot) \rangle = 0.$$

That is $f_\varepsilon(x) = 0$ for all $x \in \mathbb{R}^n$. On the other hand we have by Corollary 4.33,

$$\|f\|_{L^1(K)} = \|0 - f\|_{L^1(K)} = \|f_\varepsilon - f\|_{L^1(K)} \xrightarrow{\varepsilon \to 0} = 0.$$

That is $f = 0$ $\mathcal{L}^n$-a.e. in K, and since K was arbitary $f = 0$ in a.e. in $\mathbb{R}^n$. □

Alright, so we can describe L^1_{loc}-functions as distributions. And all distributions have distributional derivatives.

What is the relation to Sobolev space? Here it is.

Theorem 4.46. *Let $p \in [1, \infty)$ and $f \in L^p(\mathbb{R}^n)$. The following are equivalent.*

(1) $f \in W^{1,p}(\mathbb{R}^n)$
(2) for each $\alpha \in \{1, \dots, n\}$ there exists $g_\alpha \in L^p(\mathbb{R}^n)$ such that the distributional derivative $\partial_\alpha f$ coincides with g_α in the following sense:

$$\partial^\alpha f[\varphi] \equiv -\langle f, \partial^\alpha \varphi \rangle = \int_R \mathbb{R}^n g_\alpha \varphi.$$

We will simply write this as $\partial_\alpha f = g_\alpha$.

Proof. $\Rightarrow$ Let $f \in W^{1,p}(\mathbb{R}^n)$. By definition there exists $f_k \in C_c^\infty(\mathbb{R}^n)$ which is a $W^{1,p}$-Cauchy sequence converging in $L^p(\mathbb{R}^n)$ to f.

Let $g_{\alpha,k} := \partial_\alpha f_k$. These are L^p-Cauchy sequences! So they have a limit $g_\alpha \in L^p(\mathbb{R}^n)$.

Now let $\varphi \in C_c^\infty(\mathbb{R}^n)$. We have from the integration by parts formula (for Riemann Integrals!)

$$\int_{\mathbb{R}^n} \partial_\alpha f_k \varphi = -\int_{\mathbb{R}^n} f_k \partial_\alpha \varphi$$

Taking the limit $k \to \infty$ we find

$$\int_{\mathbb{R}^n} g_\alpha \varphi = -\int_{\mathbb{R}^n} f \partial_\alpha \varphi = \partial_\alpha f[\varphi].$$

That proves the first direction.

$\Leftarrow$ Now assume that

$$\int_{\mathbb{R}^n} g_\alpha \varphi = -\int_{\mathbb{R}^n} f \partial_\alpha \varphi \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$

Since g_α and f are L^p-maps we can approximate them with convolutions

$$g_{\alpha,\varepsilon} := g_\alpha * \eta_\varepsilon$$

and

$$f_\varepsilon := f * \eta_\varepsilon.$$

We may assume that $\eta(x) = \eta(-x)$, and then we have by Exercise 4.21

$$\int_{\mathbb{R}^n} (g_\alpha * \eta_\varepsilon)\varphi = \int_{\mathbb{R}^n} g_\alpha \, (\varphi * \eta_\varepsilon)$$

Since $(\varphi * \eta_\varepsilon) \in C_c^\infty(\mathbb{R}^n)$ we have

$$= -\int_{\mathbb{R}^n} f\, \partial_\alpha\, (\varphi * \eta_\varepsilon) = \int_{\mathbb{R}^n} \partial_\alpha (f * \eta_\varepsilon)\, \varphi.$$

That is, we have

$$\langle g_\alpha * \eta_\varepsilon, \varphi \rangle = \langle \partial_\alpha (f * \eta_\varepsilon), \varphi \rangle.$$

This holds for all $\varphi \in C_c^\infty(\mathbb{R}^n)$, so by the fundamental theorem of Calculus of Variations, Lemma 4.45,

$$g_\alpha * \eta_\varepsilon = \partial_\alpha (f * \eta_\varepsilon).$$

Thus we have

$$\|\partial_\alpha (f * \eta_\varepsilon) - g_\alpha\|_{L^p(\mathbb{R}^n)} = \|g_\alpha * \eta_\varepsilon - g_\alpha\|_{L^p(\mathbb{R}^n)} \xrightarrow{\varepsilon \to 0} 0.$$

That is $f * \eta_\varepsilon$ is a $W^{1,p}$-Cauchy sequence that converges to f.

However $f * \eta_\varepsilon \in C^\infty(\mathbb{R}^n)$, not in $C_c^\infty(\mathbb{R}^n)$. But this can be remedied easily.

Let $\chi \in C_c^\infty(B(0,2),[0,1])$, $\chi \equiv 1$ in $B(0,1)$ (this can be done similar to the bump functions). Set

$$f_k(x) := \chi(x/k)\ (f * \eta_{1/k})(x).$$

We have

$$\|f_k - f\|_{L^p(\mathbb{R}^n)} \leq \|(1 - \chi(x/k)) f\|_{L^p(\mathbb{R}^n)} + \|\chi(x/k) \left(f - (f * \eta_{1/k})\right)\|_{L^p(\mathbb{R}^n)}.$$

The first term tends to zero by dominated convergence, the second by the L^p-convergence of the convolution.

$$\begin{aligned}\|\partial_\alpha f_k - g_\alpha\|_{L^p(\mathbb{R}^n)} \leq & \|(1 - \chi(x/k)) g_\alpha\|_{L^p(\mathbb{R}^n)} + \|\chi(x/k) \left(g_\alpha - \partial_\alpha (f * \eta_{1/k})\right)\|_{L^p(\mathbb{R}^n)} \\ & + \|\partial_\alpha (\chi(x/k))\|_{L^\infty} \|(f * \eta_{1/k})(x) - f\|_{L^p(\mathbb{R}^n)}\end{aligned}$$

Observe that $|\partial_\alpha \chi(x/k)| \leq C/k \xrightarrow{k \to \infty} 0$. So everything converges.

□

So for now we take with us: Sobolev functions $f \in W^{1,p}(\mathbb{R}^n)$ are maps $f \in L^p(\mathbb{R}^n)$ such that $\partial_\alpha f \in L^p(\mathbb{R}^n)$ (where ∂_α denotes the distributional derivative).

Let us stress that distributional derivatives are not the same as a.e. derivatives. "a.e. derivatives are not a good object, since they "forget" things.

Example 4.47. Let h be the *Heaviside function*

$$H(x) := \begin{cases} 1 & x > 0 \\ 0 & x \leq 0. \end{cases}$$

Then clearly for almost all $x \in \mathbb{R}$ we have $H'(x)$ exists and $H'(x) = 0$ a.e. – however H is not a constant map.

What is the distributional derivative? By definition for $\varphi \in C_c^\infty(\mathbb{R})$,

$$H'[\varphi] := -\int_{\mathbb{R}} H(x)\varphi'(x)dx = -\int_0^\infty \varphi'(x)dx = \varphi(0) - \varphi(\infty) = \varphi(0).$$

(Here we can use the Riemann integral since φ is continuous).

What does that tell us? Well if we denote δ_0 again the Dirac delta, then

$$\varphi(0) = \int_{\mathbb{R}} \varphi(x) d\delta_0(x).$$

That is we have

$$H' = \delta_0$$

in distributional sense.

Proposition 4.48. *Let $1 \le p < \infty$ and assume that $f \in W^{1,p}(\mathbb{R}^n)$ satisfies $\partial_\alpha f = 0$ for all $\alpha = 1, \dots, n$. Then f is constant a.e., that is for some $c \in \mathbb{R}$ we have*

$$f(x) = c \quad a.e.\ x \in \mathbb{R}^n$$

(actually $c = 0$ since any constant function in $L^p(\mathbb{R}^n)$ is zero).

Proof. By definition there must be $f_k \in C_c^\infty(\mathbb{R}^n)$ such that

$$\|f_k - f\|_{L^p(\mathbb{R}^n)} + \|Df_k\|_{L^p(\mathbb{R}^n)} < \frac{1}{k}$$

By the classical fundamental theorem we have

$$f_k(x) - f_k(y) = \int_0^1 \sum_{\alpha=1}^n \partial_\alpha f_k(x + t(y-x))(y-x)^\alpha \, dt.$$

Integrating this inequality in both x and y in $B(0,R)$ (for some fixed R) we have

$$\begin{aligned}
&\int_{B(0,R)} \int_{B(0,R)} |f_k(x) - f_k(y)| dx dy \\
\le & 2R \left(\int_0^{\frac{1}{2}} \int_{B(0,R)} \int_{B(0,R)} |Df_k(x + t(y-x))| dx dy dt + \int_{\frac{1}{2}}^1 \int_{B(0,R)} \int_{B(0,R)} |Df_k(x + t(y-x))| dy \, dx dt \right)
\end{aligned}$$

Now for $t \in (0, \frac{1}{2})$ by the substitution rule (we use also the convexity of $B(0,R)$) and Hölder's inequality

$$\int_{B(0,R)} |Df_k(x+t(y-x))| dx \le \frac{1}{(1-t)^n} \int_{B(0,R)} |Df_k(z))| dz \le 2^n \left(\int_{B(0,R)} |Df_k(z))|^p dz \right)^{\frac{1}{p}} (\mathcal{L}^n(B(0,R)))^{1-\frac{1}{p}}$$

Doing a similar argument for the $\int_{\frac{1}{2}}^1$-integral we obtain

$$\int_{B(0,R)} \int_{B(0,R)} |f_k(x) - f_k(y)| dx dy \le C(R) \|Df_k\|_{L^p(B(0,R))}.$$

Now since f_k converges to f in $L^p(\mathbb{R}^n)$ we have that f_k converges to f in $L^1(B(0,R))$ (Hölder inequality again!), so by taking the limit we find

$$\int_{B(0,R)}\int_{B(0,R)} |f(x)-f(y)|dxdy \leq \lim_{k\to\infty} C(R)\|Df_k\|_{L^p(B(0,R))} = 0.$$

Consequently, $|f(x)-f(y)|=0$ for $\mathcal{L}^{2n}$-a.e. (x,y). By Fubini's theorem (e.g. the version in Proposition 4.6)there must be some $x \in B(0,R)$ such that $f(y)=f(x)$ for $\mathcal{L}^n$-a.e. $y \in B(0,R)$. Setting $c := f(x)$ we find that $f(y)=c$ a.e., that is f is constant in $B(0,R)$. This holds for any $R>0$ so $f=c$ a.e. in $\mathbb{R}^n$. □

Exercise 4.49. *Use Proposition 4.48 to show that $\chi_{[0,1]} \notin W^{1,p}(\mathbb{R})$ for any $p \in [1,\infty)$, where as usual*

$$\chi_{[0,1]}(x) = \begin{cases} 1 & x \in [0,1], \\ 0 & x \in \mathbb{R}\setminus[0,1]. \end{cases}$$

More precisely show that

(1) $\chi_{[0,1]} \in L^p(\mathbb{R})$ for any $p \in [1,\infty]$
(2) If there was $f \in L^1_{loc}(\mathbb{R})$ such that

$$\chi_{[0,1]}[-\varphi'] = -\int f(x)\varphi \quad \forall \varphi \in C_c^\infty(\mathbb{R})$$

then $f(x)=0$ for $\mathcal{L}^1$-a.e. $x \in \mathbb{R}$ (Hint: Fundamental theorem: Lemma 4.45)
(3) Conclude with Proposition 4.48.

We will treat Sobolev functions in way more detail later (next semester), let us just give one more application of the approximation.

Lemma 4.50. *Let $f \in W^{1,1}(\mathbb{R})$ (1 dimension only!) then f is continuous. That is, there exists $\bar{f}:\mathbb{R}\to\mathbb{R}$ continuous such that $f=\bar{f}$ a.e.*

Proof. Let $f_k \in C_c^\infty(\mathbb{R})$ be the approximation of f. By the fundamental theorem we have

$$f_k(x)-f_k(y) = \int_x^y f_k'(z)dz.$$

That is

$$|f_k(x)-f_k(y)| \leq \int_{[x,y]} |f_k'(z)|dz.$$

Restricting ourselves to a ball $[-R,R]$ we can apply Vitali's convergence theorem, Theorem 3.60: since f_k' converges in $L^1[-R,R]$ to f' we have that the integral is uniformly absolutely continuous. That is, for any $\varepsilon>0$ there exists a $\delta>0$ such that for all $x,y \in [-R,R]$ with $|x-y|<\delta$ we have

$$\int_{[x,y]} |f_k'(z)|dz < \varepsilon.$$

Thus for any $\varepsilon>0$ there exists a $\delta>0$ such that

$$|f_k(x)-f_k(y)| < \varepsilon \quad x,y \in [-R,R]: \quad |x-y|<\delta.$$

This is uniformly equicontinuity of $(f_k)_{k=1}^{\infty} : [-R, R] \to \mathbb{R}$. If only we had uniform boundedness, then we could use Arzela-Ascoli!

But don't despair. Since f_k converges to f in $L^1([-R, R])$, by Theorem 3.52 there exists a subsequence f_{k_i} which converges a.e. to f in $[-R, R]$. Let $x_0 \in [-R, R]$ be a point such that $f_{k_i}(x_0) \to f(x_0)$. Then we have from the equicontinuity

$$|f_{k_i}(x) - f_{k_i}(x_0)| \leq C,$$

or rather

$$|f_{k_i}(x)| \leq C + \sup_i |f_{k_i}(x_0)| < \infty \quad \forall x \in [-R, R].$$

This is uniform boundedness, so by Arzela-Ascoli, taking yet another subsequence if necessary we conclude that $f_{k_{i_j}}$ uniformly converges to a continuous limit function $g : [-R, R] \to \mathbb{R}$. Since on the other hand $f_{k_{i_j}}$ converges a.e. to f we have that $f = g$ a.e. in $B(0, R)$.

Since the *continuous representative*[20] of a Lebesgue-integrable function is unique, Exercise 4.51, we can let $R \to \infty$ to find a continuous representative of f. □

Exercise 4.51. *Let $f \in L^1_{loc}(\Omega)$ for some open set $\Omega \subset \mathbb{R}^n$ (and the Lebesgue measure). Assume there exist $g, h : \Omega \to \mathbb{R}$ which are continuous and $f = g$ a.e. and $f = h$ a.e. Show that $g = h$ everywhere.*

Exercise 4.52. *Use Lemma 4.50 to essentially reprove Exercise 4.52: show that $\chi_{[0,1]} \notin W^{1,p}(\mathbb{R})$ for any $p \in [1, \infty)$, where as usual*

$$\chi_{[0,1]}(x) = \begin{cases} 1 & x \in [0, 1], \\ 0 & x \in \mathbb{R} \setminus [0, 1]. \end{cases}$$

5. Differentiation of Radon measures - Radon-Nikodym Theorem on $\mathbb{R}^n$

5.1. **Preparations: Besicovitch Covering theorem.** If a set A is covered by closed balls, we would sometimes like it to be covered by countably many *disjoint* closed balls (e.g. so we can sum of $\mu(B)$). One very useful theorem is the

Theorem 5.1 (Besicovitch Covering theorem)**.** *There exists a constant N_n depending only on the dimension n with the following property.*

If $\mathcal{F}$ is a family of closed balls $B(x, r)$, $r > 0$ in $\mathbb{R}^n$ with finite maximal radius, i.e.

$$\sup \{\operatorname{diam} B : \quad B \in \mathcal{F}\} < \infty,$$

and if A is the set of centers of balls in $\mathcal{F}$, then there exist N_n countable families $\mathcal{G}_1, \ldots, \mathcal{G}_{N_n}$ of balls where each family $\mathcal{G}_i \subset \mathcal{F}$ consists of pairwise disjoint balls in $\mathcal{F}$ such that

$$A \subset \bigcup_{i=1}^{N_n} \bigcup_{B \in \mathcal{G}_i} B.$$

[20] more on this later: see Definition 5.17

Proof. The proof is lengthy and technical and we skip it here. See [Evans and Gariepy, 2015, Theorem 1.27]. □

As an application of Besicovitch covering theorem we have

Theorem 5.2. *Let μ be a Borel measure on $\mathbb{R}^n$ and $\mathcal{F}$ any collection of closed balls $\overline{B(a,r)}$ for $a \in \mathbb{R}^n$ and $r \in (0,\infty)$.*

Let A denote the set of centers of the balls in $\mathcal{F}$ and assume $\mu(A) < \infty$.

Moreover assume that for each $a \in A$

$$\inf\{r : \quad \overline{B(a,r)} \in \mathcal{F}\} = 0.$$

Then for each open set $U \subset \mathbb{R}^n$ there exists a countable collection $\mathcal{G}$ of disjoint balls in $\mathcal{F}$ such that

$$\bigcup_{B \in \mathcal{G}} B \subset U$$

and

$$\mu\left((A \cap U) \setminus \bigcup_{B \in \mathcal{G}} B\right) = 0.$$

(observe there are no assumptions on the measurability of A).

Proof. See [Evans and Gariepy, 2015, Theorem 1.28]. □

5.2. **The Radon-Nikodym Theorem.** In this section[21] let ν and μ be *always* two Radon measures on $\mathbb{R}^n$. *Extremely formally* we could hope to write

$$\mu(A) = \int_A d\mu = \int_A \frac{d\mu}{d\nu} d\nu. \tag{5.1}$$

And then (as we do in calculus), we could interpret $\frac{d\mu}{d\nu}$ as the derivative of μ in direction ν

$$\frac{d\mu}{d\nu} = D_\mu \nu.$$

In some sense (5.1) is the fundamental theorem of calculus for measures.

Definition 5.3. For each point $x \in \mathbb{R}^n$ set

$$\overline{D}_\mu \nu(x) := \begin{cases} \limsup_{r \to 0} \frac{\nu(\overline{B(x,r)})}{\mu(\overline{B(x,r)})} & \text{if } \mu(\overline{B(x,r)}) > 0 \text{ for all } r > 0 \\ +\infty & \text{if } \mu(\overline{B(x,r)}) = 0 \text{ for some } r > 0 \end{cases}$$

and

$$\underline{D}_\mu \nu(x) := \begin{cases} \liminf_{r \to 0} \frac{\nu(\overline{B(x,r)})}{\mu(\overline{B(x,r)})} & \text{if } \mu(\overline{B(x,r)}) > 0 \text{ for all } r > 0 \\ +\infty & \text{if } \mu(\overline{B(x,r)}) = 0 \text{ for some } r > 0 \end{cases}$$

[21] We follow to a substantial extend [Evans and Gariepy, 2015].

If $\overline{D}_\mu\nu(x) = \underline{D}_\mu\nu(x) < \infty$ we say that ν is *differentiable with respect to μ*. We also call $D_\mu\nu$ the *density* of ν with respect to μ. [22]

Example 5.4.
- If $\nu(B) := \lambda\mu(B)$ then $D_\mu\nu(x) = \lambda$.
- We will later see (Theorem 5.15, as a consequence of the Radon-Nikodym theorem Theorem 5.13) that if for a μ-integrable function f

$$\nu = f \llcorner \mu$$

i.e. for μ-measurable A,

$$\nu(A) := \int_A f d\mu$$

then

$$D_\mu\nu(x) = f(x) \quad \mu - \text{a.e.}.$$

This is easy to see for $\mu = \mathcal{L}^n$ and f continuous in x:
Then

$$\nu(\overline{B(x,r)}) = \int_{\overline{B(x,r)}} f d\mu = \mu(\overline{B(x,r)})f(x) + \int_{\overline{B(x,r)}} (f - f(x))\, d\mu.$$

By continuity for any $\varepsilon > 0$ there exists $r > 0$ such that $\sup_{z\in B(x,r)} |f(z) - f(x)| < \varepsilon$, so we have

$$\left|\nu(\overline{B(x,r)}) - \mu(\overline{B(x,r)})f(x)\right| \le \varepsilon\mu(\overline{B(x,r)})$$

So we have

$$\left|\frac{\nu(\overline{B(x,r)})}{\mu(\overline{B(x,r)})} - f(x)\right| \xrightarrow{r\to 0} 0.$$

Our first goal is to understand when $D_\mu\nu$ exists and when (5.1) holds.

For this the following observation is useful

Lemma 5.5. *Let μ be a Radon measure on $\mathbb{R}^n$ and ν be any other measure on $\mathbb{R}^n$ (not necessarily Borel regular or anything).*

Assume that $A \subset \mathbb{R}^n$ is a (possibly not measurable) set such that

$$\nu(A \cap V) \le \mu(A \cap V) \quad \textit{for all bounded Borel sets } V \subset \mathbb{R}^n$$

Then

$$\nu(A) \le \mu(A).$$

[22] This indeed has some features of a derivative, because we could believe that for "a.e." x, $\nu(\overline{B(x,0)}) = \mu(\overline{B(x,0)}) = 0$ (where $B(x,0) = \{x\}$) and then write

$$\liminf_{r\to 0} \frac{\nu(\overline{B(x,r)})}{\mu(\overline{B(x,r)})} = \liminf_{r\to 0} \frac{\nu(\overline{B(x,r)}) - \nu(\overline{B(x,0)})}{\mu(\overline{B(x,r)}) - \mu(\overline{B(x,0)})}$$

Proof. If $\mu(A) = \infty$ there is nothing to show. So assume $\mu(A) < \infty$. Then there must be an open set $U \supset A$ such that

$$\mu(U) \leq \mu(A) + \varepsilon.$$

We can write $\mathbb{R}^n = \bigcup_{i=1}^{\infty} V_i$ where each V_i is a bounded Borel set and $V_i \cap V_j = \emptyset$ and thus $\mu(V_i) < \infty$ for $i \neq j$ (e.g. take partially open and closed cubes).

We then have

$$\nu(A) \leq \sum_{i=1}^{\infty} \nu(A \cap V_i) \leq \sum_{i=1}^{\infty} \mu(A \cap V_i) \leq \sum_{i=1}^{\infty} \mu(U \cap V_i) = \mu(U).$$

In the last step we used the measurability of $U \cap V_i$ (but A may not be measurable!) Thus,

$$\nu(A) \leq \mu(A) + \varepsilon.$$

This holds for any $\varepsilon > 0$ and we can conclude by letting $\varepsilon \to 0$. □

Lemma 5.6. *Let μ, ν be Radon measures on $\mathbb{R}^n$. Fix $0 < \alpha < \infty$. Let $A \subset \mathbb{R}^n$*

(1) If $A \subset \{x \in \mathbb{R}^n : \quad \underline{D}_\mu \nu(x) \leq \alpha\}$ then $\nu(A) \leq \alpha\mu(A)$
(2) If $A \subset \{x \in \mathbb{R}^n : \quad \overline{D}_\mu \nu(x) \geq \alpha\}$ then $\nu(A) \geq \alpha\mu(A)$

Observe that we don't assume A to be ν or μ measurable!

Proof. (1) Fix $\varepsilon > 0$ and assume $A \subset \{x \in \mathbb{R}^n : \quad \underline{D}_\mu \nu(x) \leq \alpha\}$.
Let $V \subset \mathbb{R}^N$ be any bounded Borel set (in particular $\mu(V), \nu(V) < \infty$).
We are going to show

$$\nu(A \cap V) \leq (\alpha + \varepsilon)\mu(A \cap V). \tag{5.2}$$

Since this holds for any bounded Borel set V, in view of Lemma 5.5, we find ($(\alpha+\varepsilon)\mu$ is still a Radon measure!)

$$\nu(A) \leq (\alpha + \varepsilon)\mu(A).$$

Letting $\varepsilon \to 0$ we can then conclude.

Take any bounded open set $U \subset \mathbb{R}^n$ with $A \cap V \subset U$ (this exists, because V is bounded). Then $\nu(U) < \infty$.

We define a family of balls:

$$\mathcal{F} := \left\{\overline{B(a,r)} : \quad \text{where } a \in A \cap V, r > 0 : \ \overline{B(a,r)} \subset U, \nu(\overline{B(a,r)}) \leq (\alpha + \varepsilon)\mu(\overline{B(a,r)})\right\}$$

Since for each $a \in A$ we have

$$\liminf_{r \to 0} \frac{\nu(\overline{B(a,r)})}{\mu(\overline{B(a,r)})} \leq \alpha$$

we find that in particular for each $a \in A \cap V$

$$\inf\{r > 0 : \quad \overline{B(a,r)} \in \mathcal{F}\} = 0.$$

By Theorem 5.2 there is a countable subfamily $\mathcal{G} \subset \mathcal{F}$ of pairwise disjoint balls such that

$$\bigcup_{B\in\mathcal{G}} B \subset U$$

and

$$\nu((A\cap V)\cap U \setminus \bigcup_{B\in\mathcal{G}} B) = \nu((A\cap V)\setminus \bigcup_{B\in\mathcal{G}} B) = 0$$

Then, for each $B \in \mathcal{G} \subset \mathcal{F}$

$$\nu(B) \le (\alpha+\varepsilon)\mu(B)$$

and summing this up over $B \in \mathcal{G}$ (observe the B are all measurable as closed balls) we have

$$\nu(A\cap V) = \nu(A\cap V\cap U) \le \sum_{B\in\mathcal{G}} \nu(B)+0 \le \sum_{B\in\mathcal{G}} (\alpha+\varepsilon)\mu(B)+0 = (\alpha+\varepsilon)\mu(\bigcup_{B\in\mathcal{G}} B) \le (\alpha+\varepsilon)\mu(U).$$

We have this inequality for any bounded open set $U \subset \mathbb{R}^n$ with $A\cap V \subset U$. Taking the infimum over all such U we find in view of Theorem 1.68 Equation (5.2) holds and we can conclude.

(2) (similar)

□

Theorem 5.7 (Differentiating measures)**.** *Let μ and ν be Radon measures on $\mathbb{R}^n$. Then*

(1) $D_\mu\nu(x)$ exists and is finite for μ-a.e.[23] *$x \in \mathbb{R}^n$ and*
(2) $D_\mu\nu$ is μ-measurable.

We may assume that $\mu(\mathbb{R}^n)$ and $\nu(\mathbb{R}^n) < \infty$. Otherwise we consider $\mu \llcorner K$ and $\nu \llcorner K$ for compact sets and show (1) and (2) in open sets $U \subset K$ with $\operatorname{dist}(U, K) > 0$. Then we can argue by exhaustion.

We split the proof of Theorem 5.7 into different parts.

Lemma 5.8. *Assumptions as in Theorem 5.7. Then $D_\mu\nu$ exists and is finite μ-a.e.*

Proof. Let $A := \{x \in \mathbb{R}^n : \overline{D}_\mu\nu(x) = \infty\}$. Then for each $\alpha > 0$

$$A \subset \{x \in \mathbb{R}^n : \overline{D}_\mu\nu(x) \ge \alpha\},$$

and we can apply Lemma 5.6(2) to obtain

$$\mu(A) \le \frac{1}{\alpha}\nu(A).$$

If $\nu(A) < \infty$ we can let $\alpha \to \infty$ to conclude $\mu(A) = 0$. If $\nu(A) = \infty$ then since $\nu(A \cap (B(0,R))) < \infty$ (ν is Radon) we find $\mu(A \cap B(0,R)) = 0$ for all $R > 0$ and thus again $\mu(A) = 0$.

[23] this μ-a.e. takes care of pathological examples. E.g. assume $\mu(\mathbb{R}^n) = 0$ then $D_\mu\nu \equiv \infty$, but still $D_\mu\nu(x) = 0$ for μ-a.e. $x \in \mathbb{R}^n$ since $\mu(\mathbb{R}^n) = 0$

That is

$$\overline{D}_\mu \nu(x) < \infty \quad \mu\text{-a.e.}$$

Now let $0 < a < b$ and set

$$R(a,b) := \{x \in \mathbb{R}^n : \underline{D}_\mu \nu(x) < a < b < \overline{D}_\mu \nu(x) < \infty\}.$$

We apply again Lemma 5.6 and have

$$b\mu(R(a,b)) \le \nu(R(a,b)) \le a\mu(R(a,b)).$$

Since $b > a$ and (w.l.o.g. otherwise again intersect with balls) $\mu(R(a,b)) < \infty$ we conclude $\mu(R(a,b)) = 0$. Thus

$$\left\{x \in \mathbb{R}^n : \quad \underline{D}_\mu \nu(x) < \overline{D}_\mu \nu(x) < \infty\right\} = \bigcup_{a \in \mathbb{Q}} \bigcup_{b \in \mathbb{Q} \cap (a,\infty)} R(a,b).$$

The right-hand side is a countable union of μ-zerosets, so

$$\mu\left(\left\{x \in \mathbb{R}^n : \quad \underline{D}_\mu \nu(x) < \overline{D}_\mu \nu(x) < \infty\right\}\right) = 0.$$

Since we always have $\underline{D}_\mu \nu(x) \ge \overline{D}_\mu \nu(x)$ we conclude that

$$\mu\left(\left\{x \in \mathbb{R}^n : \quad \underline{D}_\mu \nu(x) \ne \overline{D}_\mu \nu(x)\right\}\right) = 0.$$

That is $\underline{D}_\mu \nu(x) \ne \overline{D}_\mu \nu(x) < \infty$ for μ-a.e. x. □

One ingredient is the upper semicontinuity of $x \mapsto \mu(B(x,r))$.

Lemma 5.9. *Let μ be a Radon measure, $x \in \mathbb{R}^n$ and $r > 0$. Then*

$$\limsup_{y \to x} \mu(\overline{B(y,r)}) \le \mu(\overline{B(x,r)})$$

Proof. Since $\overline{B(x,r)}$ is Borel, it is measurable and $\chi_{\overline{B(x,r)}}$ is μ-integrable since $\overline{B(x,r)}$ is compact.

Let $y_k \to y$ and set $f_k := \chi_{\overline{B(y_k,r)}}$ and $f := \chi_{\overline{B(x,r)}}$. While it is not so clear whether $f_k \to f$ everywhere we certainly have for any $z \in \mathbb{R}^n$

$$\limsup_{k \to \infty} f_k(z) \le f(z) \tag{5.3}$$

Indeed, the only case we need to consider is $z \in \mathbb{R}^n$ with $f(z) = 0$, otherwise the inequality is obvious. But then $z \in \mathbb{R}^n \backslash \overline{B(x,r)}$, and thus $\operatorname{dist}(z, \overline{B(x,r)}) > \delta$ for some $\delta > 0$. Now if $|y_k - x| < \frac{\delta}{2}$ then $z \notin B(y_k, r)$, so since $y_k \to x$ we have $f_k(z) = 0$ for all but finitely many $k \in \mathbb{N}$. So (5.3) is established.

From (5.3) we conclude

$$\liminf_{k \to \infty} (1 - f_k) \ge (1 - f).$$

Applying Fatou's lemma Corollary 3.9 on $X = B(x, 2r)$

$$\int_{B(x,2r)} (1-f)d\mu \leq \int_{B(x,2r)} \liminf_{k\to\infty}(1-f_k)d\mu \leq \liminf_{k\to\infty} \int_{B(x,2r)} (1-f_k)d\mu$$

That is

$$\underbrace{\mu(B(x,2r))}_{<\infty} - \mu(B(x,r)) \leq \liminf_{k\to\infty} \left(\mu(B(x,2r)) - \mu(B(y_k,r))\right).$$

Subtracting the constant $\mu(B(x,2r))$ from both sides we conclude. □

Lemma 5.10. *Assumptions as in Theorem 5.7. Then $D_\mu\nu$ is μ-measurable.*

Proof. By Lemma 5.9, the functions $x \mapsto \mu(\overline{B(x,r)})$ and $x \mapsto \nu(\overline{B(x,r)})$ are upper semi-continuous. In view of Exercise 2.16 both functions are then μ-measurable. Consequently for each $r > 0$

$$f_r(x) := \begin{cases} \frac{\nu(\overline{B(x,r)})}{\mu(\overline{B(x,r)})} & \text{if } \mu(\overline{B(x,r)}) > 0 \\ +\infty & \text{if } \mu(\overline{B(x,r)}) = 0 \end{cases}$$

is μ-measurable. Since by Lemma 5.8 $D_\mu\nu(x) = \lim_{r\to 0} f_r(x)$ for μ-a.e. we have by Theorem 2.13 that $D_\mu\nu$ is μ-measurable. □

Exercise 5.11. *Let μ, ν_1, ν_2 be Radon measures. Show that*

$$D_\mu(\nu_1 + \nu_2) = D_\mu\nu_1 + D_\mu\nu_2 \quad \mu\text{-a.e.}$$

We recall the definition of absolute continuity from Definition 3.55 (see also Lemma 3.56) and somewhat (as we shall see) define the opposite: mutually singular.

Definition 5.12. Assume μ and ν are Borel measures on $\mathbb{R}^n$.

(1) The measure ν is *absolutely continuous* with respect to μ, in symbols $\nu \ll \mu$, if $\mu(A) = 0$ implies $\nu(A) = 0$ for all $A \subset \mathbb{R}^n$.
(2) The measures ν and μ are *mutually singular*, in symbols $\nu \perp \mu$ if there exists a Borel set $B \subset \mathbb{R}^n$ such that

$$\mu(\mathbb{R}^n \setminus B) = \nu(B) = 0.$$

Here is the fundamental theorem for measures, also called the *Radon-Nikodym Theorem*.

Theorem 5.13. *Let ν, μ be Radon measures on $\mathbb{R}^n$ with $\nu \ll \mu$.*

Then for all μ-measurable sets A we have

$$\nu(A) = \int_A D_\mu\nu \; d\mu.$$

Proof. Let $A \subset \mathbb{R}^n$ be μ-measurable. Then, by Lemma 3.56, A is also ν-measurable.

Set

$$Z := \{x \in \mathbb{R}^n : \quad D_\mu\nu(x) = 0\}$$

and

$$I := \{x \in \mathbb{R}^n : \quad D_\mu \nu(x) = +\infty\}$$

By Theorem 5.7 $\mu(I) = 0$ and thus since $\nu \ll \mu$ we have $\nu(I) = 0$. Also $\nu(I) = 0$ by Lemma 5.6. In particular

$$\nu(Z) = 0 = \int_Z D_\mu \nu \; d\mu$$
$$\nu(I) = 0 = \int_I D_\mu \nu \; d\mu.$$

Now let A be μ-measurable and fix $t \in (1, \infty)$ and set for $m \in \mathbb{Z}$

$$A_m := A \cap \{x \in \mathbb{R}^n : \quad t^m \leq D_\mu \nu(x) < t^{m+1}\}.$$

Then A_m is also μ-measurable, and thus ν-measurable, Lemma 3.56, and we have

$$t^m \mu(A_m) \leq \int_{\{x:\, t^m \leq D_\mu \nu(x)\}} t^m d\mu \leq \int_{A_m} D_\mu \nu d\mu$$

Since

$$A \setminus \bigcup_{m \in \mathbb{Z}} A_m \subset Z \cup I \cup \{x \in \mathbb{R}^n : \quad \overline{D}_\mu \nu(x) \neq \underline{D}_\mu \nu(x)\}$$

we have that $\mu(A \setminus \bigcup_{m \in \mathbb{Z}} A_m) = 0$ and thus ($(A_m)_{m \in \mathbb{Z}}$ are pairwise disjoint!)

$$\sum_{m \in \mathbb{Z}} t^m \mu(A_m) \leq \sum_m \int_{A_m} D_\mu \nu d\mu = \int_A D_\mu \nu \; d\mu.$$

Since $\nu \ll \mu$ we also have $\nu(A \setminus \bigcup_{m \in \mathbb{Z}} A_m) = 0$ and thus

$$\nu(A) = \sum_{m \in \mathbb{Z}} \nu(A_m).$$

With the help of Lemma 5.6 we find

$$\nu(A) = \sum_{m \in \mathbb{Z}} \nu(A_m) \leq \sum_m t^{m+1} \mu(A_m) \leq t \int_A D_\mu \nu \; d\mu$$

and similarly

$$\nu(A) = \sum_{m \in \mathbb{Z}} \nu(A_m) \geq \sum_m t^m \mu(A_m) = t^{-1} \sum_m t^{m+1} \mu(A_m) \geq t^{-1} \int_A D_\mu \nu \; d\mu.$$

That is for any $t > 1$ we have

$$t^{-1} \int_A D_\mu \nu \; d\mu \leq \nu(A) \leq t \int_A D_\mu \nu \; d\mu.$$

Letting $t \to 1$ we conclude. □

If $\nu \not\ll \mu$ we can get also the following refinement

Theorem 5.14 (Lebesgue Decomposition Theorem)**.** *Let ν and μ be Radon measures on $\mathbb{R}^n$.*

(1) Then we can decompose
$$\nu = \nu_{ac} + \nu_s$$
where the absolutely continuous part $\nu_{ac} \ll \mu$ *and the* singular part $\nu_s \perp \mu$.

(2) Furthermore,
$$D_\mu \nu = D_\mu \nu_{ac}, \quad D_\mu \nu_s = 0 \quad \mu\text{-a.e.}$$
and consequently
$$\nu(A) = \int_A D_\mu \nu d\mu + \nu_s(A)$$
for each Borel set $A \subset \mathbb{R}^n$[24].

Proof. Again assume $\mu(\mathbb{R}^n), \nu(\mathbb{R}^n) < \infty$ or argue with compact exhaustion.

We begin by constructing the singular measure. For this we are going to find a suitable Borel set B and set
$$\nu_{ac} := \nu \llcorner B, \quad \nu_s := \nu \llcorner (\mathbb{R}^n \setminus B).$$
To find B define $\mathcal{E}$ the set of "good candidates", i.e. Borel sets A where $\mu(\mathbb{R}^n \setminus A) = 0$, i.e.
$$\mathcal{E} := \{A \subset \mathbb{R}^n : \quad A \text{ Borel}, \mu(\mathbb{R}^n \setminus A) = 0\}.$$
We are going to choose B such that
$$\nu(B) = \min_{A \in \mathcal{E}} \nu(A).$$
But for this we have to show that such a B exists. Certainly, $\mathcal{E} \neq \emptyset$ so there must be some $B_k \in \mathcal{E}$ with
$$\nu(B_k) \leq \inf_{A \in \mathcal{E}} \nu(A) + \frac{1}{k} \quad k = 1, \ldots$$
Write $B := \bigcap_{k=1}^\infty B_k$. Then
$$\mu(\mathbb{R}^n \setminus B) \leq \sum_{k=1}^\infty \mu(\mathbb{R}^n \setminus B_k) = 0.$$
That is $B \in \mathcal{E}$ and thus
$$\nu(B) = \inf_{A \in \mathcal{E}} \nu(A). \tag{5.4}$$
Define
$$\nu_{ac} := \nu \llcorner B$$
and
$$\nu_s := \nu \llcorner (\mathbb{R}^n \setminus B).$$
Both, ν_{ac} and ν_s are Radon measures, see Exercise 1.67.

We now show that $\nu_{ac} \ll \mu$. Assume $A \subset \mathbb{R}^n$ with $\mu(A) = 0$. Since μ is a Borel regular measure we may assume that A is Borel (otherwise we pass to $A \subset \tilde{A}$ with $\mu(\tilde{A}) = \mu(A)$). Then $\nu_{ac}(A) = \nu(A \cap B)$. Observe that
$$\mu(\mathbb{R}^n \setminus (B \setminus A)) \leq \mu(\mathbb{R}^n \setminus B) + \mu(A) = 0,$$

[24] so A is μ and ν-measurable

so $B \setminus A \in \mathcal{E}$. Then by (5.4) (observe both A and B are Borel, so ν-measurable)

$$\nu(B) \leq \nu(B \setminus A) = \nu(B) - \nu(A \cap B).$$

Thus $\nu(A \cap B) = 0$, i.e. $\nu_{ac}(A) = 0$. I.e., we have established $\nu_{ac} \ll \mu$.

Now fix $\alpha > 0$ and set

$$C_\alpha := \{x \in B : \quad D_\mu \nu_s(x) \geq \alpha\}$$

In view of Lemma 5.6 we have

$$\alpha\mu(C_\alpha) \leq \nu_s(C_\alpha) = \nu(C_\alpha \cap (\mathbb{R}^n \setminus B)) = \nu(\emptyset) = 0.$$

That is $\mu(C_\alpha) = 0$ for all $\alpha > 0$, that is $D_\mu \nu_s(x) = 0$ for μ-a.e. $x \in B$ (and since $D_\mu \nu_s(x) = D_\mu(\nu \llcorner B)$ we have that $D_\mu \nu_s(x) = 0$ for μ-a.e. $x \in \mathbb{R}^n \setminus B$ by Theorem 5.7(1). So $D_\mu \nu_s(x) = 0$ for a.e. $x \in \mathbb{R}^n$, and thus (with the help of Exercise 5.11)

$$D_\mu \nu_{ac} = D_\mu \nu \quad \mu\text{-a.e.} \tag{5.5}$$

By the first Radon-Nikodym theorem, Theorem 5.13 we then have for any Borel set (thus ν_{ac}-measurable)

$$\nu_{ac}(A) = \int_A D_\mu \nu_{ac} d\mu \overset{(5.5)}{=} \int_A D_\mu \nu d\mu$$

Thus, by the definition of ν_s and ν_{ac}, for any Borel measure

$$\nu(A) = \nu_{ac}(A) + \nu_s(A) = \int_A D_\mu \nu d\mu + \nu_s(A)$$

□

5.3. **Lebesgue differentiation theorem.** The *integral average* of f over a measurable set E of finite, strictly positive measure will be denoted by

$$f_E \equiv \fint_E d\mu := \mu(E)^{-1} \int_E f d\mu$$

Theorem 5.15. *Let μ be a Radon measure on $\mathbb{R}^n$ and $f \in L^1_{loc}(\mathbb{R}^n, \mu)$. Then for μ-a.e. $x \in \mathbb{R}^n$*

$$\lim_{r \to 0} \fint_{B(x,r)} f \, d\mu = f(x)$$

Exercise 5.16. *Show Theorem 5.15 for continuous f (using only continuity, no deep theorem).*

Proof of Theorem 5.15. For μ-measurable sets E set

$$\nu_+(E) := \int_E f_+ d\mu$$

and

$$\nu_-(E) := \int_E f_- d\mu.$$

Since $f \in L^1_{loc}(\mathbb{R}^n)$ we see that $\nu_\pm$ extend to Radon measures, and $\nu_\pm \ll \mu$.

By the Radon-Nikdoym theorem, Theorem 5.13

$$\nu_+(A) = \int_A D_\mu \nu_+ d\mu = \int_A f_+ d\mu \quad \text{for all } \mu\text{-measurable } A \subset \mathbb{R}^n$$
$$\nu_-(A) = \int_A D_\mu \nu_- d\mu = \int_A f_- d\mu \quad \text{for all } \mu\text{-measurable } A \subset \mathbb{R}^n$$

Since ν_+, ν_- are Radon measures, we see that $D_\mu \nu_\pm$ are $L^1_{loc}(d\mu)$-functions, so by Lemma 3.22(2), $f_+ = D_\mu \nu_+$ and $f_- = D_\mu \nu_-$ μ-a.e. in $\mathbb{R}^n$.

By the definition of the Radon-Nikodym derivative we have

$$\begin{aligned} \fint_{B(x,r)} f d\mu =& \frac{\nu_+(B(x,r))}{\mu(B(x,r))} - \frac{\nu_-(B(x,r))}{\mu(B(x,r))} \\ \xrightarrow{r\to 0^+} & D_\mu \nu^+(x) - D_\mu \nu^-(x) = f^+(x) - f^-(x) = f(x), \end{aligned}$$

which holds for μ-a.e. x in $\mathbb{R}^n$ – we can pass to $\overline{B(x,r)}$ by Proposition 1.72. □

Definition 5.17. Let $f \in L^1_{loc}(\mu)$.

The *precise representative* of f is the (pointwise!, not a.e., not a class) defined function

$$f^*(x) := \begin{cases} \lim_{r\to 0^+} \fint_{B(x,r)} f\, d\mu & \text{when the limit exists} \\ 0 & \text{otherwise} \end{cases}$$

By Theorem 5.15 the limit exists μ-a.e.

We say that $x \in \mathbb{R}^n$ is a *Lebesgue point* of f if

$$\lim_{r\to 0} \fint_{B(x,r)} |f(y) - f^*(x)| d\mu(y) = 0.$$

Theorem 5.18. *Let μ be a Radon measure and $f \in L^1_{loc}(\mu)$. Then μ-a.e. point $x \in \mathbb{R}^n$ is a Lebesgue point of f*

Proof. By Theorem 5.15 for any $c \in \mathbb{R}$ there exists a set $N_c \subset \mathbb{R}$ with $\mu(N_c) = 0$ such that

$$\lim_{r\to 0} \fint_{B(x,r)} |f(y) - c| d\mu(y) = \lim_{r\to 0} \fint_{B(x,r)} |f^*(y) - c| d\mu(y) = |f^*(x) - c| \quad \text{for all } x \in \mathbb{R}^n \setminus N_c.$$

Let

$$M := \{x : |f^*(x)| = \infty\}$$

which is a zeroset, since $f^*(x) = f(x)$ μ-a.e. and $f \in L^1_{loc}$.

Let

$$N := M \cup \bigcup_{c\in\mathbb{Q}} N_c.$$

Since $\mathbb{Q}$ is countable, $\mu(N) = 0$. Let $x \in \mathbb{R}^n \setminus N$ and let $c_n \in \mathbb{Q}$ such that $c_n \xrightarrow{n\to\infty} f^*(x)$. Then

$$\fint_{B(x,r)} |f(y) - f^*(x)| d\mu(y) \le \fint_{B(x,r)} |f(y) - c_n| d\mu(y) + |c_n - f^*(x)| \xrightarrow{r\to 0^+} |f^*(x) - c_n| + |c_n - f^*(x)|,$$

that is

$$\limsup_{r\to 0^+} \fint_{B(x,r)} |f(y) - f^*(x)| d\mu(y) \le |f^*(x) - c_n| + |c_n - f^*(x)|$$

This holds for any $n \in \mathbb{N}$ and letting $n \to \infty$ we have

$$\limsup_{r\to 0^+} \fint_{B(x,r)} |f(y) - f^*(x)| d\mu(y) = 0,$$

and thus

$$\lim_{r\to 0^+} \fint_{B(x,r)} |f(y) - f^*(x)| d\mu(y) = 0.$$

We can conclude since this holds for any $x \in \mathbb{R}^n \setminus N$, where N satisfies $\mu(N) = 0$. □

Sometimes one wants to work with cubes or ellipsoids shrinking to zero instead of balls. This is possible, as long as they are regular (and μ is doubling). We will restrict here our attention to the Lebesgue measure.

Definition 5.19. We say that a family $\mathcal{F}$ of measurable sets in $\mathbb{R}^n$ is *regular* at $x \in \mathbb{R}^n$ if there is a constant $C > 0$ such that for every $S \in \mathcal{F}$ there is a ball $B(x, r_S)$ such that

$$S \subset B(x, r_S), \quad \mathcal{L}^n(B(x, r_S)) \le C\mathcal{L}^n(S),$$

and for every $\varepsilon > 0$ there is a set $S \in F$ with $\mathcal{L}^n(S) < \varepsilon$.

Theorem 5.20. *$f \in L^1_{loc}(\mathbb{R}^n, \mathcal{L}^n)$. If $x \in \mathbb{R}^n$ is a Lebesgue point of f and $\mathcal{F}$ is a regular family at $x \in \mathbb{R}^n$, then*

$$\lim_{S\in F, \mathcal{L}^n(S)\to 0} \fint_S f d\mathcal{L}^n = f(x).$$

Proof. For $S \in \mathcal{F}$ let r_S be defined as above. Observe that if $\mathcal{L}^n(S) \to 0$, then $r_S \to 0$. We have

$$\begin{aligned}
\left|\fint_S f d\mathcal{L}^n - f(x)\right| \le & \fint_S |f(y) - f(x)| d\mathcal{L}^n(y) \\
\le & \mathcal{L}^n(S)^{-1} \int_{B(x,r_s)} |f(y) - f(x)| d\mathcal{L}^n(y) \\
= & \frac{\mathcal{L}^n(B(x, r_S))}{\mathcal{L}^n(S)} \fint_{B(x,r_S)} |f(y) - f(x)| d\mathcal{L}^n(y) \\
\le & C \fint_{B(x,r_S)} |f(y) - f(x)| d\mathcal{L}^n(y) \xrightarrow{\mathcal{L}^n(S)\to 0} 0.
\end{aligned}$$

□

Corollary 5.21. *If $f \in L^1_{loc}(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$ is a Lebesgue point of f, then for any sequence of cubes Q_i such that $x \in Q_i$, $\operatorname{diam} Q_i \to 0$ we have*

$$\lim_{i\to\infty} \fint_{Q_i} f d\mathcal{L}^n = f(x).$$

Corollary 5.22. *Let $F(x) = \int_a^x f(t)dt$ where $f \in L^1(a, b)$. Then $F'(x) = f(x)$ for a.e. $x \in (a, b)$.*

Proof. We have

$$\frac{F(x+h)-F(x)}{h} = \fint_x^{x+h} f(t)dt \xrightarrow{h\to 0} f(x)$$

whenever x is a Lebesgue point of f. □

Exercise 5.23. *Show the following: if $A \subset \mathbb{R}^n$ is a $\mathcal{L}^n$-measurable set, then for almost all $x \in A$*

$$\lim_{r\to 0^+} \frac{|B(x,r)\cap A|}{|B(x,r)|} = 1$$

and for almost all $x \in \mathbb{R}^n \setminus A$

$$\lim_{r\to 0^+} \frac{|B(x,r)\cap A|}{|B(x,r)|} = 0.$$

Hint: $f := \chi_A$

Definition 5.24. Let $A \subset \mathbb{R}^n$ be a measurable set. We say that $x \in \mathbb{R}^n$ is a *density point* of A if

$$\lim_{r\to 0} \frac{|A\cap B(x,r)|}{|B(x,r)|} = 1.$$

Thus the above theorem says that almost every point of a measurable set $A \subset \mathbb{R}^n$ is a density point.

5.4. **Signed (pre-)measures – Hahn decomposition theorem.** We want to apply now and then the Radon-Nikodym theorem to measures which can be positive and negative. For example

$$\mu \llcorner f(A) := \int_A f d\mu \quad A\ \mu\text{-measurable}$$

where $f \in L^1(X, d\mu)$ but f may be positive or negative.

This leads to the so-called *signed measures*. Since e.g. $\mu \llcorner f$ is only defined on measurable sets, we will actually work with pre-measures.

Definition 5.25. Let Σ be a *σ-algebra* and $\mu : \Sigma \to [-\infty, \infty]$. μ is called a *signed premeasure* if

(1) $\mu(\emptyset) = 0$
(2) μ attains at most one of the values $\pm\infty$.
(3) μ is countably additive, i.e. if $(E_k)_{k\in\mathbb{N}}$ are disjoint sets in Σ then

$$\mu(\bigcup E_k) = \sum_{k\in\mathbb{N}} \mu(E_k)$$

Example 5.26.
- If μ_1 and μ_2 are (positive) measures and one of them is finite then $\mu := \mu_1 - \mu_2$ is a signed premeasure on the intersection of the measurable sets of μ_1 and μ_2.

- If ν is a (positive) measure and $f \in L^1(\nu)$ then
$$\mu(E) := \nu \llcorner f = \int_E f d\nu$$
is a signed premeasure on the set of ν-measurable sets. Indeed, if $f \in L^1$ then $f_+, f_- \in L^1$, so
$$\mu(E) = \int_E f^+ d\nu - \int_E f^- d\nu$$
and each of the terms is a (positive) measure.

Theorem 5.27 (Hahn's decomposition theorem)**.** *Let $\mu : \Sigma \to \bar{\mathbb{R}}$ be a signed premeasure on X. Then there are two disjoint sets $X^+, X^- \subset \Sigma$ such that*
$$X = X^+ \cup X^-, \quad X^+ \cap X^- = \emptyset,$$
such that
$$\mu_+ = \mu \llcorner X^+, \quad \mu_- := -\mu \llcorner X_-$$
then $\mu = \mu_+ - \mu_-$ and $\mu_+, \mu_- : \Sigma \to [0, \infty]$ are nonnegative premeasures on X.

Exercise 5.28. *Prove that the decomposition $X = X_+ \cup X-$ is unique up to sets of μ-measure zero, i.e. if $X = \tilde{X}_+ \cup \tilde{X}_-$ is another decomposition, then*
$$\mu(X_+ \setminus \tilde{X}_+) = \mu(\tilde{X}_+ \setminus X_+) = \mu(X_- \setminus \tilde{X}_-) = \mu(\tilde{X}_- \setminus X_-) = 0$$

Example 5.29. Let $\mu := f \llcorner \nu$, for $f \in L^1(\nu)$, i.e.
$$\mu(E) = \int_E f d\nu$$
for ν-measurable f. Then μ is a signed premeasure and
$$\mu(E) = \int_E f^+ d\nu - \int_E f^- d\nu$$
Hence we can take
$$X^+ = \{x : f(x) \geq 0\}$$
$$X^- = \{x : f(x) < 0\}$$
but we can also take
$$\tilde{X}^+ = \{x : f(x) > 0\}$$
$$\tilde{X}^- = \{x : f(x) \leq 0\}$$
In general $X^+ \neq \tilde{X}^+$, and $X^- \neq \tilde{X}^-$. But the sets differ by the set where $f = 0$ and this set has μ-measure zero,
$$\mu(\{x : f(x) = 0\}) = \int_{\{f=0\}} f d\nu = 0.$$

For the proof of Theorem 5.27 we need the notion of positive sets.

Definition 5.30. Let $\mu : \Sigma \to \bar{\mathbb{R}}$ be a signed premeasure. A set $E \in \Sigma$ is called *positive* if

- $\mu(E) > 0$

- $\mu(A) \geq 0$ for any $A \in \Sigma$, $A \subset E$.

Example 5.31. If $E = E_1 \cup E_2$ and $E_1 \cap E_2 = \emptyset$, $\mu(E_1) = 7$, $\mu(E_2) = -3$ then $\mu(E) = 4 > 0$, but E is not positive.

Lemma 5.32. *If $\mu : \Sigma \to \bar{\mathbb{R}}$ is a signed measure then every measurable set E such that $0 < \mu(E) < \infty$ contains a positive set A with $\mu(A) > 0$.*

Proof. If E itself is positive we can take $A = E$.

Otherwise there is $B \subset E$ with $\mu(E) < 0$. Let $n_1 \in \mathbb{N}$ be the smallest positive integer such that there is $B_1 \subset E$ with

$$\mu(B_1) \leq -\frac{1}{n_1}.$$

Observe that $A_1 := E \setminus B_1$ satisfies

Also observe that $\mu(A_1) < \infty$, indeed if $\mu(A_1) = \infty$ then $\mu(B_1) = -\infty$, but then μ attains both $+\infty$ and $-\infty$ which is ruled out. That is $0 < \mu(A_1) < \infty$.

So either A_1 is positive and we take A_1, or we can repeat this construction. Namely in the next step we take $n_2 \in \mathbb{N}$ the smallest positive integer such that there is $B_2 \subset A_1 = E \setminus B_1$ with

$$\mu(B_2) \leq -\frac{1}{n_2}.$$

As above, either $A_2 = E \setminus (B_1 \cup B_2)$ is positive, then we can take $A = A_2$, if not we continue.

If $A_m = E \setminus \bigcup_{j=1}^m B_j$ is positive for some m, we take $A = A_m$ (and have $0 < \mu(A_m) < \infty$), otherwise we obtain an infinite sequence B_j and we set

$$A := E \setminus \bigcup_{j=1}^{\infty} B_j \in \Sigma$$

We claim that $0 < \mu(A) < \infty$ and A is positive.

Firstly, $\mu(E) = \mu(A) + \mu(E \setminus A)$. Since $\mu(E) \in (0, \infty)$ we conclude that if $\mu(A) = +\infty$ then $\mu(E \setminus A) = -\infty$ which is not allowed for a signed measure. So $\mu(A) < \infty$, and we have (observe: $(B_j)_j$ and A are by definition pairwise disjoint!)

$$0 < \mu(E) = \mu(A) + \sum_{j=1}^{\infty} \mu(B_j) \leq \mu(A) - \sum_{j=1}^{\infty} \frac{1}{n_j}.$$

That is

$$\infty > \mu(A) > \sum_{j=1}^{\infty} \frac{1}{n_j} > 0.$$

It remains to prove that A is positive. From the above inequality we conclude that in particular the series on the right-hand side must converge, that is $\lim_{j\to\infty} n_j = \infty$.

If $C \in \Sigma$ and $C \subset A = E \setminus \bigcup_{j=1}^{\infty} B_j$ then for every $m \in \mathbb{N}$ we have $C \subset E \setminus \bigcup_{j=1}^{m} B_j = A_m$. Since n_{m+1} is the smallest positive integer such that there is a set $\tilde{B} \subset A_m$ with $\mu(\tilde{B}) \leq -\frac{1}{n_{m+1}}$ we must have

$$\mu(C) > -\frac{1}{n_{m+1}-1} \xrightarrow{m\to\infty} 0.$$

That is $\mu(C) \geq 0$. Thus A is positive. □

Proof of Theorem 5.27. W.l.o.g. assume that μ does not attain the value $+\infty$, otherwise we work with $-\mu$.

Let

$$M := \sup\left\{\mu(\tilde{A}) : \tilde{A} \in \Sigma,\ \tilde{A} \text{ positive}\right\} \in (-\infty, \infty].$$

Then there is a sequence of positive sets

$$A_1 \subset A_2 \subset A_3 \subset \dots$$

with

$$\mu(A_i) \xrightarrow{i\to\infty} M.$$

Set $A := \bigcup_{i=1}^{\infty} A_i$ then

- A is positive
- $\mu(A) = M < \infty$

It remains to prove that $X \setminus A$ is negative (then we take $X_+ = A$, $X_- = X \setminus A$).

Assume by contradiction that $X \setminus A$ is not negative, there is $E \subset X \setminus A$ with $0 < \mu(E) < \infty$. Hence Lemma 5.32 implies that there is a positive subset $C \subset E$ of positive measure. Now $A \cup C$ is positive and

$$\mu(A \cup C) = \mu(A) + \mu(C) > M$$

which is an obvious contradiction. □

5.5. **Riesz representation theorem.** We will apply now the Radon-Nikodym theorem to classify all continuous linear functionals on $L^p(\mathbb{R}^n)$ for $1 \leq p < \infty$. Let us start first with some general facts about continuous linear mappings between normed spaces.

Theorem 5.33. *Let $L : X \to Y$ be a linear mapping between normed spaces. Then the following conditinos are equivalent*

(1) L is continuous;
(2) L is continuous at 0;
(3) L is bounded, i.e. there is $C > 0$ such that

$$\|Lx\| \leq C\|x\| \quad \text{for all } x \in X \tag{5.6}$$

(4) L is uniformly Lipschitz continuous, i.e. there is $C > 0$ such that $\|Lx - Ly\| \leq C\|x - y\|$ for all $x \in X$.

Remark 5.34. Formally we should use different symbols to denote norms in spaces X and Y. Since it will always be clear in which space we take the norm it is not too dangerous to use the same symbol $\|\cdot\|$ to denote apparently different norms in X and Y.

Proof. The implication (1)⇒(2) is obvious.

(2)⇒(3). Suppose L is continuous at 0 but not bounded. Then there is a sequence $x_n \in X$ such that

$$\|Lx_n\| \geq n\|x_n\|.$$

In particular $\|x_n\| \neq 0$, so

$$0 \leftarrow \left\| L\frac{x_n}{n\|x_n\|} \right\| \geq 1$$

which is a contradiction[25].

(3)⇒(1). Let $x_n \to x$. Then

$$\|Lx - Lx_n\| = \|L(x - x_n)\| \leq C\|x - x_n\| \to 0$$

and hence $Lx_n \to Lx$ which proves continuity of L.

(3) ⇔ (4) is obvious, simply take $y = 0$ (so $Ly = 0$) and use linearity of L. □

Definition 5.35. The number $\|L\| = sup_{\|x\|\leq 1}\|Lx\|$ is called the norm of L. It is often called the *operatornorm*

Exercise 5.36. *Let X and Y be two normed vector spaces. Denote by*

$$L(X,Y) := \{T : X \to Y : \quad T \text{ linear and continuous}\}$$

the space of linear continouts (aka bounded) maps from X to Y.

For two linear continuous operators $T, S : X \to Y$ and $\lambda, \mu \in \mathbb{R}$ we set

$$(\lambda T + \mu S) : X \to Y$$

defined via

$$(\lambda T + \mu S)\, x := \lambda Tx + \mu Sx.$$

(1) Show that for $T, S \in L(X,Y)$ we also have $\lambda T + \mu S \in L(X,Y)$.
(2) Show that the operatornorm $\|\cdot\|$ from Lemma 5.37 is indeed a norm on $L(X,Y)$.

Lemma 5.37. *If $L : X \to Y$ is a linear continuous operator then the norm $\|L\|$ is the smallest number C for which the inequality (5.6) is satisfied.*

That is

$$\|Lx\| \leq \|L\|\,\|x\| \quad \forall x \in X$$

and for each $C \geq 0$ if

$$\|Lx\| \leq C\,\|x\| \quad \forall x \in X$$

[25]Convergence to 0 follows from the continuity of L at 0.

then $C \geq \|L\|$.

Proof. Clearly, if $x = 0$ then

$$\|Lx\| = 0 \leq \|L\|\|0\|.$$

If $x \geq 0$ then

$$\|Lx\| = \|L\frac{x}{\|x\|}\| \, \|x\| \leq \|L\|\|x\|.$$

Now assume that $\|Lx\| \leq C\|x\|$ for all $x \in X$.

If $\|x\| \leq 1$, then $\|Lx\| \leq C\|x\| \leq C$ and hence $\|L\| = sup_{\|x\|\leq 1}\|Lx\| \leq C$. □

Now we want to consider linear bounded *functionals*, which are continuous linear maps $L : X \to \mathbb{R}$.

Example 5.38. If $\frac{1}{p} + \frac{1}{q} = 1$, $1 < p, q < \infty$ are Hölder conjugate, and $g \in L^q(X)$, then

$$L_g(f) := \int_X fgd\mu$$

defines a bounded linear functional on $L^p(\mu)$ and $|Lg| \leq \|g\|_{L^q}$. Indeed, Hölder's inequality yields

$$|L_g f| = \left|\int_X fgd\mu\right| \leq \|g\|_{L^q}\|f\|_{L^p}.$$

Similarly if $g \in L^\infty(\mu)$, then

$$L_g f = \int_X fgd\mu$$

defines a bounded linear functional on $L^1(\mu)$ and $\|L^g\| \leq \|g\|_{L^\infty}$.

The following theorem states that these are the only linear functionals on $L^p(\mathbb{R}^n)$ as long as $1 \leq p < \infty$ (similar statement hold for reasonable μ).

Theorem 5.39 (Riesz representation theorem)**.** *Let* $1 \leq p < \infty$ *and* $L : L^p(\mathbb{R}^n) \to \mathbb{R}$ *be a bounded linear functional.*

$$|L[f]| \leq \|T\|\|f\|_{L^p(\mathbb{R}^n)} \tag{5.7}$$

Denoting by $p' = \frac{p}{p-1}$ *the Hölder conjugate there exists* $g \in L^{p'}(\mathbb{R}^n)$ *such that*

$$L[f] = \int_{\mathbb{R}^n} fg \quad \forall f \in L^p(\mathbb{R}^n).$$

and

$$\|g\|_{L^{p'}(\mathbb{R}^n)} = \|T\|. \tag{5.8}$$

The function g *is unique in the sense that if there are* g_1 *and* g_2 *satisfying the above then* $g_1 = g_2$ *a.e..*

Remark 5.40. • The set of bounded linear functionals on $L^p(\mathbb{R}^n)$ is usually denoted by $(L^p(\mathbb{R}^n))^*$. It is a vector space (we will discuss more properties in Functional Analysis) and called the *dual space* to $L^p(\mathbb{R}^n)$. What Theorem 5.39 says is that there exists an isometric isomorphism from the space of bounded linear functionals L on $L^p(\mathbb{R}^n)$ with $L^{p'}(\mathbb{R}^n)$, i.e.

$$(L^p(\mathbb{R}^n))^* \cong L^{p'}(\mathbb{R}^n).$$

Often one says that *the dual space to $L^p(\mathbb{R}^n)$, i.e. $(L^p(\mathbb{R}^n))^*$, is $L^{p'}(\mathbb{R}^n)$*. What is important here (but often this is not so explicitely clear for other spaces) is the *pairing*, i.e. in which sense this "identity" holds. We observe this is the L^2-scalar products (even if L^p-functions do not in general belong to L^2. Any element in $L \in (L^p(\mathbb{R}^n))^*$ can be identified with some $g \in L^{p'}(\mathbb{R}^n)$ via the *L^2-pairing*

$$L(f) = \langle f, g\rangle.$$

Proof of Theorem 5.39. Uniqueness Assume that there are g_1 and g_2 that both satisfy the claims, then

$$\int (g_1 - g_2) f = 0 \quad \forall f \in L^p(\mathbb{R}^n).$$

In particular

$$\int (g_1 - g_2)\varphi = 0 \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$

By the fundamental theorem of the calculus of variations, Lemma 4.45 we have $g_1 = g_2$ a.e.

Existence Let $K \subset \mathbb{R}^n$ be compact with $\mathcal{L}^n(\partial K) = 0$ and consider for $\mathcal{L}^n$-measurable $A \subset \mathbb{R}^n$

$$\nu(A) := L[\chi_{A\cap K}].$$

Observe that $|\nu(A)| \le \|\chi_K\|_{L^p(\mathbb{R}^n)} < \infty$.

We claim that ν is a *signed* pre-measure on the σ-Algebra of $\mathcal{L}^n$-measurable sets. For this let $(A_k)_{k=1}^\infty$ be disjoint $\mathcal{L}^n$-measurable sets and set $A := \bigcup_k A_k$. Then for all $x \in \mathbb{R}^n$ we have

$$f(x) := \chi_{A\cap K}(x) = \lim_{k\to\infty} f_k(x)$$

where $f_k := \sum_{i=1}^k \chi_{A_i\cap K}$. Since K is compact, $f \in L^p(\mathbb{R}^n)$. Moreover, since the $(A_i)_i$ are pairwise disjoint, so by Lebesgue monotone convergence, Corollary 3.8,

$$\|f_k\|^p_{L^p(\mathbb{R}^n)} = \|\sum_{i=1}^k \chi_{A_i\cap K}\|_{L^1(\mathbb{R}^n)} \xrightarrow{k\to\infty} \|f\|_{L^1(\mathbb{R}^n)} = \|f\|^p_{L^p(\mathbb{R}^n)}.$$

Sine $p < \infty$[26], in view of Theorem 3.53, we then have

$$\|f_k - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0 \tag{5.9}$$

[26]here we use this crucially

and thus

$$\nu(\bigcup_{i=1}^{\infty} A_i) = L(f) = \lim_{k\to\infty} L(f_k) = \sum_{i=1}^{\infty} \nu(A_i).$$

That is ν is a signed premeasure.

By Hahn decomposition theorem, Theorem 5.27, we can split ν into two pre-measures ν_+ and ν_-,

$$\nu = \nu_+ - \nu_-.$$

Then ν_+, ν_- both extend to a Radon measure, which by (5.7) satisfies $\nu_\pm \ll \mathcal{L}^n$. Indeed, since ν is a signed premeasure and $|\nu(\mathbb{R}^n)| < \infty$ we have that $\nu_+(\mathbb{R}^n), \nu_-(\mathbb{R}^n) < \infty$ (because otherwise they would both be infinite which would contradict $|\nu(\mathbb{R}^n)| < \infty$).

By Radon-Nikodym theorem for $g_K := \chi_K D_{\mathcal{L}^n}\nu_+ - \chi_K D_{\mathcal{L}^n}\nu_-$ which is $\mathcal{L}^n$-a.e. finite and $\mathcal{L}^n$-integrable (because each ν_+ and ν_- are finite measures), we have

$$\nu(A) = \nu(A \cap K) = \int_{\mathbb{R}^n} \chi_A g_K(x)\, d\mathcal{L}^n(x).$$

In particular we have $g_K(x) \in L^1(\mathbb{R}^N)$, and for all $A \subset \mathbb{R}^n$ $\mathcal{L}^n$-measurable

$$L(\chi_{A\cap K}) = \int_{\mathbb{R}^n} \chi_A g_K(x)\, d\mathcal{L}^n(x).$$

Now let $f \in L^\infty(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$, $f \geq 0$ *a.e.* in $\mathbb{R}^n$ then there exists an approximation

$$f_k(x) = \sum_{i=1}^{\infty} a_k \chi_{A_k}(x)$$

such that $f_k \to f$ a.e. in $\mathbb{R}^n$. In particular

$$|f(x)|^p = \lim_{k\to\infty} |f_k(x)|^p \quad \text{a.e. in } \mathbb{R}^n.$$

By the monotone convergence theorem we conclude that $\lim_{k\to\infty} \|f_k\|_{L^p(\mathbb{R}^n)} = \|f\|_{L^p(\mathbb{R}^n)} < \infty$. Here we used Exercise 3.18. Thus from Theorem 3.53 we find that $\|f_k - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0$.

We then have

$$L(\chi_K f) = L\left(\chi_K\left(f - f_k\right)\right) + L\left(\chi_K f_k\right) = L(\chi_K(f - f_k)) + \int_{\mathbb{R}^n} f_k g_K \, d\mathcal{L}^n.$$

Since L is uniformly Lipschitz and $f - f_k \xrightarrow{k\to\infty} 0$ in $L^p(\mathbb{R}^n)$ we have $L(\chi_K(f - f_k)) \xrightarrow{k\to\infty} 0$.

Also, by the monotone convergence theorem

$$\begin{aligned}\lim_{k\to\infty} \int_{\mathbb{R}^n} f_k g_K \, d\mathcal{L}^n &= \lim_{k\to\infty} \int_{\mathbb{R}^n} f_k (g_K)_+ \, d\mathcal{L}^n - \lim_{k\to\infty} \int_{\mathbb{R}^n} f_k (g_K)_- \, d\mathcal{L}^n \\ &= \int_{\mathbb{R}^n} f (g_K)_+ \, d\mathcal{L}^n + \int_{\mathbb{R}^n} f (g_K)_+ \, d\mathcal{L}^n = \int_{\mathbb{R}^n} f g_K \, d\mathcal{L}^n.\end{aligned}$$

In the last step we have used that $\|f g_K\|_{L^1(\mathbb{R}^n)} \leq \|f\|_{L^\infty(\mathbb{R}^n)} \|g_K\|_{L^1(\mathbb{R}^n)} < \infty$.

Thus, we have shown

$$L(\chi_K f) = \int_{\mathbb{R}^n} f g_K \, d\mathcal{L}^n \quad \forall f \in L^\infty(\mathbb{R}^n) \cap L^1(\mathbb{R}^n),\ f \geq 0.$$

Splitting $f = f_+ + f_-$ we conclude that we don't need the sign condition on f.

$$L(\chi_K f) = \int_{\mathbb{R}^n} f g_K \, d\mathcal{L}^n \quad \forall f \in L^\infty(\mathbb{R}^n) \cap L^1(\mathbb{R}^n).$$

For now we have $g_K \in L^1(\mathbb{R}^n)$ only. Next we need to show $\|g_K\|_{L^{p'}(\mathbb{R}^n)} \leq \|T\|$.

If $p = 1$ and thus $p' = \infty$ let $x \in \mathbb{R}^n \backslash \partial K$ be any Lebesgue point of g_K. If we choose $f := \chi_{B(x,r)}$ for $r < \operatorname{dist}(x, \partial K)$

$$\left| \int_{B(x,r)} g_K \, d\mathcal{L}^n \right| = \left| L\chi_{K \cap B(x,r)} \right| \leq \|T\| \mathcal{L}^n(K \cap B(x,r)).$$

By the assumptions on r if $x \in K$ then $\mathcal{L}^n(K \cap B(x,r)) = \mathcal{L}^n(B(x,r))$. If $x \notin K$ then $\mathcal{L}^n(K \cap B(x,r)) = 0$. In either case by the Lebesgue differentiation theorem we can let $r \to 0$ and have

$$|g_K(x)| \leq \|T\|$$

which holds for any Lebesgue point $x \in \mathbb{R}^n \backslash \partial K$. Since $\mathcal{L}^n(\partial K) = 0$ we have $g_k(x) \leq \|T\|$ for $\mathcal{L}^n$-a.e. $x \in \mathbb{R}^n$, and thus

$$\|g_K\|_{L^\infty} \leq \|T\|.$$

Now assume $p \in (1, \infty)$. We argue *by duality*.

Formally, this is easy: For $p' = \frac{p}{p-1} \in (1, \infty)$ we have

$$\|g_K\|_{L^{p'}}^{p'} = \int_{\mathbb{R}^n} g_K \, g_K |g_K|^{p'-2} = T[|g_K|^{p'-2} g_K] \leq \|T\| \||g_K|^{p'-2} g_K\|_{L^p} = \|T\| \|g_K\|_{L^{p(p'-1)}}^{p'-1}$$

so that (since $p(p'-1) = p'$, we can divide on both sides $\|g_K\|_{L^{p(p'-1)}}^{p'-1} \equiv \|g_K\|_{L^{p'}}^{p'-1}$)) we have found

$$\|g_K\|_{L^{p'}} \leq \|T\|.$$

Essentially that will be our argument, but the above only works if we already know that $|g_K\|_{L^{p'}}^{p'-1}) < \infty$ (which is pretty much what we want to show). The above type of argument is called an *a priori estimate*. To make this precise we need mollification (a technique which is often sweeped under the rug by saying *by approximation* or *by density*):

Let $f \in L^1(\mathbb{R}^n)$ with $f = 0$ a.e. outside a compact set $\tilde{K}$. For $\varepsilon > 0$ we denote by f_ε the usual convolution $f_\varepsilon = f * \varepsilon^{-n} \eta(\cdot/\varepsilon)$ (with symmetric kernel η). Then by Exercise 4.21

$$L(\chi_K f_\varepsilon) = \int_{\mathbb{R}^n} f_\varepsilon \, g_K \, d\mathcal{L}^n = \int_{\mathbb{R}^n} f (g_K)_\varepsilon \, d\mathcal{L}^n.$$

Observe that since $\operatorname{supp} g_K \subset K$ we have that $(g_K)_\varepsilon \in C_c^\infty(\mathbb{R}^n)$.

Then we may pick (here we use that $p \in (1, \infty)$)

$$f(x) := \begin{cases} |(g_K)_\varepsilon(x)|^{\frac{p}{p-1}-1} & \text{if } (g_K)_\varepsilon(x) \geq 0 \\ -|(g_K)_\varepsilon(x)|^{\frac{p}{p-1}-1} & \text{if } (g_K)_\varepsilon(x) < 0 \end{cases} \tag{5.10}$$

(you can check this is measurable). Then we have $f \in L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$. Indeed, this follows from (observe: $\frac{p}{p-1} - 1 > 0$)

$$|f(x)| \leq |(g_K)_\varepsilon(x)|^{\frac{p}{p-1}-1} \leq \|(g_K)_\varepsilon(x)\|_{L^\infty}^{\frac{p}{p-1}-1},$$

so $f \in L^\infty(\mathbb{R}^n)$, but also f has compact support. Thus we have

$$\begin{aligned} \int_{\mathbb{R}^n} |(g_K)_\varepsilon|^{p'} &= \int_{\mathbb{R}^n} f(x)\,(g_K)_\varepsilon \\ &= L(\chi_K f) \leq \|T\| \underbrace{\|\,((g_K)_\varepsilon)^{\frac{p}{p-1}-1}\,\|_{L^p(\mathbb{R}^n)}}_{<\infty}. \end{aligned}$$

A short computation yields

$$(p'-1)p = \frac{p}{p-1} = p'$$

and thus

$$\|\,((g_K)_\varepsilon)^{\frac{p}{p-1}-1}\,\|_{L^p(\mathbb{R}^n)} = \|(g_K)_\varepsilon\|_{L^{p'}(\mathbb{R}^n)}^{\frac{p'}{p}}$$

We conclude that

$$\|(g_K)_\varepsilon\|_{L^{p'}(\mathbb{R}^n)}^{p'} = \int_{\mathbb{R}^n} |(g_K)_\varepsilon|^{p'} \leq \|T\| \|(g_K)_\varepsilon\|_{L^p(\mathbb{R}^n)}^{\frac{p'}{p}}$$

Observing that $p' - \frac{p'}{p} = p'(1 - \frac{1}{p}) = \frac{p'}{p'} = 1$ we have shown that

$$\|(g_K)_\varepsilon\|_{L^{p'}(\mathbb{R}^n)} \leq \|T\|.$$

This holds for all $\varepsilon > 0$. Since $g_K \in L^1(\mathbb{R}^N)$ we have that $(g_K)_\varepsilon$ converges to g_K in L^1 as $\varepsilon \to 0$, which means for some $\varepsilon_i \to 0$ the functions $(g_K)_{\varepsilon_i}$ converge a.e. to g_K. From Exercise 3.14 we finally obtain

$$\|g_K\|_{L^{p'}(\mathbb{R}^n)} \leq \|T\|.$$

<u>Removing the K</u>. Now from the construction of g_K we see that if $K \subset K'$ both compact with ∂K and $\partial K'$ both zero sets, then $g_K = g_{K'}$ a.e. in K and $|g_K| \leq |g_{K'}|$ a.e. in $\mathbb{R}^n$.

So $g(x) := g_K(x)$ where K is a large compact set containing x makes g well defined and we have

$$\|\chi_K g\|_{L^{p'}(\mathbb{R}^N)} = \|g_K\|_{L^{p'}(\mathbb{R}^N)} \leq \|T\| \quad \forall K \subset \mathbb{R}^n \text{ compact.}$$

By monotonce convergence theorem we have

$$\|g\|_{L^{p'}(\mathbb{R}^n)} = \|\,|g|\,\|_{L^{p'}(\mathbb{R}^n)} = \lim_{k\to\infty} \|\chi_{B(0,k)}\,|g|\,\|_{L^{p'}(\mathbb{R}^n)} \leq \|T\|$$

and we have

$$L(f) = \int_{\mathbb{R}^n} gf \; d\mathcal{L}^n(\mathbb{R}^n) \quad \forall f \in L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n): \quad \text{supp}\, f \text{ compact}.$$

(simply take K to contain $\text{supp}\, f$).

It remains to show that

$$L(f) = \int_{\mathbb{R}^n} gf \quad \forall f \in L^p(\mathbb{R}^n). \tag{5.11}$$

We do this *by density* (here we use that $p < \infty$). Any $f \in L^p(\mathbb{R}^n)$ can be approximated by $f_k \in C_c^\infty(\mathbb{R}^n)$. Then we have

$$\begin{aligned} L(f) =& L(f - f_k) + L(f_k) = L(f - f_k) + \int_{\mathbb{R}^n} gf_k \\ =& L(f - f_k) + \int_{\mathbb{R}^n} gf + \int_{\mathbb{R}^n} g(f_k - f). \end{aligned}$$

Now by boundedness of L

$$|L(f - f_k)| \leq \|T\| \|f - f_k\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

Moreover by Hölder's inequality

$$\left| \int_{\mathbb{R}^n} g(f_k - f) \right| \leq \|g\|_{L^p(\mathbb{R}^n)} \|f_k - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

So we have established (5.11).

The last thing to establish is (5.8). We have already

$$\|g\|_{L^{p'}} \leq \|T\|.$$

On the other hand by Hölder's inequality

$$|T(f)| = \left| \int fg \right| \leq \|f\|_{L^p(\mathbb{R}^n)} \|g\|_{L^{p'}(\mathbb{R}^n)} \quad \forall f \in L^p(\mathbb{R}^n).$$

In view of Lemma 5.37 we find $\|g\| \leq \|T\|$. □

Remark 5.41. Theorem 5.39 is not true for $p = \infty$, since (5.9) may fail for $p = \infty$ (cf. Exercise 3.54). Indeed any linear functional L on $L^\infty(\mathbb{R}^n)$ can be written as an integration $L(f) = \int f d\mu$ where μ is a *finite* additive signed measures, which are absolutely continuous with respect to $\mathcal{L}^n$. The Radon-Nikodym theorem does not apply for *finite* additive measures, but for *infinitely* additive measures.

See also **??**.

Exercise 5.42. *Use Theorem 5.39 and Exercise 4.34 to prove the following slight (but useful!) generalization of Riesz representation theorem.*

Assume $1 \leq p \leq \infty$ *and assume* $L : C_c^\infty(\mathbb{R}^n) \to \mathbb{R}$ *is a linear map that satisfies for some* $\Lambda \geq 0$

$$|L\varphi| \leq \Lambda \|\varphi\|_{L^p(\mathbb{R}^n)} \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$

Denoting by $p' = \frac{p}{p-1}$ *the Hölder conjugate there exists* $g \in L^{p'}(\mathbb{R}^n)$ *such that*

$$L[f] = \int_{\mathbb{R}^n} fg \quad \forall f \in C_c^\infty(\mathbb{R}^n).$$

and

$$\|g\|_{L^{p'}(\mathbb{R}^n)} = \|T\|.$$

The function g *is unique in the sense that if there are* g_1 *and* g_2 *satisfying the above then* $g_1 = g_2$ *a.e..*

Hint: Show that the following functional $\tilde{L} : L^p(\mathbb{R}^n) \to \mathbb{R}$ *is well-defined, linear, and bounded. For* $f \in L^p(\mathbb{R}^n)$ *take any approximation* $f_k \in C_c^\infty(\mathbb{R}^n)$, $f_k \xrightarrow{k\to\infty} f$ *in* $L^p(\mathbb{R}^n)$. *Then we define*

$$\tilde{L}f := \lim_{k\to\infty} Lf_k.$$

(You have to check that $\tilde{L}(f) \in \mathbb{R}$ *and that this value is independent of the choice of approximation!)*

Exercise 5.43 (Minkowski's integral inequality)**.** *Let* $f \in C_c^\infty(\mathbb{R}^{k+\ell})$ *and assume that* $1 < p < \infty$. *We denote variables in* $\mathbb{R}^{k+\ell}$ *by* (x, y), $x \in \mathbb{R}^k$, $y \in \mathbb{R}^\ell$. *Show the following*

$$\left(\int_{\mathbb{R}^k} \left(\int_{\mathbb{R}^\ell} |f(x,y)|dy\right)^p dx\right)^{\frac{1}{p}} \leq \int_{\mathbb{R}^\ell} \left(\int_{\mathbb{R}^k} |f(x,y)|^p dx\right)^{\frac{1}{p}} dy$$

Use the duality characterization of L^p *as in* (5.10)*, and Fubini's theorem.*

Let us also record (without proof) the following version of the Riesz representation theorem that classify the set of Radon measures wiht linear functions on

$$C_c(\mathbb{R}^n) := \{f \in C^0(\mathbb{R}^n) : \quad \operatorname{supp} f \text{ bounded}\}.$$

Theorem 5.44 (Riesz Representation Theorem)**.** *Let*

$$L : C_c(\mathbb{R}^n; \mathbb{R}^m) \to \mathbb{R}$$

be a linear functional satisfying

$$\sup\{L(f) : \quad f \in C_c(\mathbb{R}^n; \mathbb{R}^m),\ |f| \leq 1, \operatorname{supp} f \subset K\} < \infty$$

for every compact set $K \subset \mathbb{R}^n$.

Then there exists a Radon measure μ *on* $\mathbb{R}^n$ *and a* μ*-measurable function* $\sigma : \mathbb{R}^n \to \mathbb{R}^m$ *such that*

$$|\sigma(x)| = 1 \quad \mu\text{-a.e. } x$$

and

$$L(f) = \int_{\mathbb{R}^n} f \cdot \sigma d\mu \quad \forall f \in C_c(\mathbb{R}^n, \mathbb{R}^m).$$

For a proof see Cf. [Evans and Gariepy, 2015, Theorem 1.38]

6. Transformation Rule

In this section we discuss the transformation rule, and will assume always $\mu = \mathcal{L}^n$.

The change of variables formula is a general version of the substitution rule. If $\Phi' \neq 0$ we have

$$\int_{[a,b]} f(\Phi(x))|\Phi'(x)|dx = \int_{\Phi([a,b])} f(x)dx.$$

Why do we write $|\Phi'(x)|$ not $\Phi'(x)$ (as we learned in Calculus)?

Assume that $\Phi(x) = -x$ then for $a < b$ (i.e. $-a > -b$),

$$\int_{[a,b]} f(-x)dx = \int_a^b f(-x)dx = -\int_{-a}^{-b} f(x)dx = \int_{-b}^{-a} f(x)dx = \int_{[-b,-a]} f(x) \tag{6.1}$$

I.e. *orientation* is accounted for differently with the notation $\int_a^b$ than with $\int_{[a,b]}$.

The one-dimensional $\Phi'(x)$ becomes $\det(D\Phi(x))$ because (as we have also seen in Theorem 1.81), this is the change of volume of the unit cube under the linear map $D\Phi(x)$, cf. Figure 6.1 and Figure 1.6.

Definition 6.1. Let $X, Y \subset \mathbb{R}^n$ be open.

$\Phi : X \to Y$ is a C^k-diffeomorphism if

- $\Phi : X \to Y$ is bijective
- Φ, Φ^{-1} are C^k

One can show that this implies

$$\det(D\Phi(x)) \neq 0 \quad \text{in } X.$$

Theorem 6.2 (Change of variables formula/transformation rule)**.** *Let $\Phi : X \to \Phi(X)$ be a C^1-diffeomorphism between open and bounded sets X, $\Phi(X) \subset \mathbb{R}^n$.*

(1) for any measurable $\Omega \subset X$ we have

$$\mathcal{L}^n(\Phi(\Omega)) = \int_\Omega |\det(D\Phi(x))|dx.$$

(2) $f : Y \to \bar{R}$ is integrable in Y if and only if $f \circ \Phi|\det(D\Phi)|$ is integrable in X and if that is the case,

$$\int_Y f dx = \int_X f \circ \Phi|\det(D\Phi(x))|dx.$$

or equivalently

$$\int_{\Phi(X)} f dx = \int_X f \circ \Phi|\det(D\Phi(x))|dx.$$

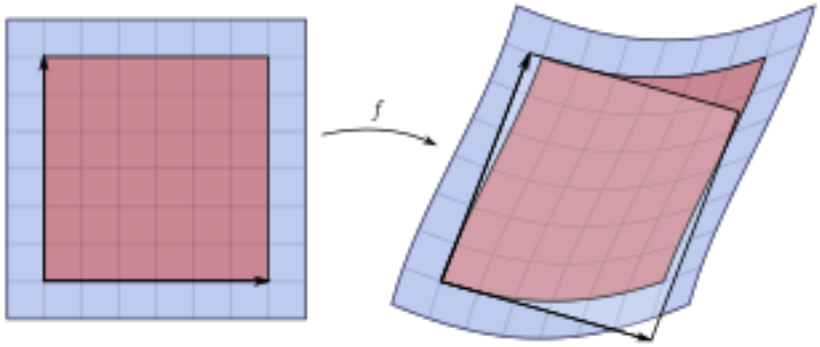

FIGURE 6.1. A nonlinear map $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ sends a small square (left, in red) to a distorted parallelogram (right, in red). The Jacobian at a point gives the best linear approximation of the distorted parallelogram near that point (right, in translucent white), and the Jacobian determinant gives the ratio of the area of the approximating parallelogram to that of the original square. By Blacklemon67 - Own work, CC BY-SA 3.0, wikipedia

Lemma 6.3. *Under the assumptions of Theorem 6.2 for any $\Omega \subset X$ measurable*

$$\mathcal{L}^n(\Phi(\Omega)) = \int_\Omega |\det(D\Phi(x))| dx.$$

Proof. Set

$$\mu(A) := \mathcal{L}^n(\Phi(A \cap X)) \quad A \text{ is } \mathcal{L}^n\text{-measurable}$$

Since Φ is a C^1-diffeomorphism this induces a Radon measure and $\mu \ll \mathcal{L}^n$, cf. Exercise 1.80, Exercise 1.13 and Corollary 1.78.

By Radon-Nikdoym theorem, Theorem 5.13, we have

$$\mu(A) = \int_A D_\mu \mathcal{L}^n \, d\mathcal{L}^n.$$

and

$$D_\mu \mathcal{L}^n(x) = \lim_{r \downarrow 0} \frac{\mu(B(x,r))}{\mathcal{L}^n(B(x,r))} \quad \text{for } \mathcal{L}^n - a.e.\ x.$$

So fix $x \in X$ (by exhaustion argument using the monotone convergence theorem we may assume $\operatorname{dist}(x, \partial X) > 0$) and $r \ll 1$. We then have from Taylor's theorem for each $z \in B(x,r)$, since $\Phi \in C^1$

$$|\Phi(z) - \Phi(x) - D\Phi(x)(z-x)| \le o(r).$$

Thus, since $D\Phi(x)$ is invertible matrix,

$$\Phi(B(x,r)) \subset \Phi(x) + \underbrace{D\Phi(x)}_{\in \mathbb{R}^{n\times n}}(B(x,r) - x) + B_{o(r)}(0) = \Phi(x) + D\Phi(x)\left((B(x,r) - x) + D\Phi(x)^{-1} B_{o(r)}(0)\right)$$

Consequently by translation invariance of the Lebesgue measure and Theorem 1.81.

$$\begin{aligned}\mu(B(x,r)) =& \mathcal{L}^n(\Phi(B(x,r))) \leq \mathcal{L}^n\left(D\Phi(x)\left((B(x,r)-x)+D\Phi(x)^{-1}B_{o(r)}(0)\right)\right)\\ =&|\det(D\Phi(x))|\,\mathcal{L}^n\left((B(x,r)-x)+D\Phi(x)^{-1}B_{o(r)}(0)\right)\\ =&|\det(D\Phi(x))|\,\mathcal{L}^n\left(B(x,r)+D\Phi(x)^{-1}B_{o(r)}(0)\right)\\ \leq&|\det(D\Phi(x))|\mathcal{L}^n(B(x,r+o(r))).\end{aligned}$$

Thus

$$\frac{\mu(B(x,r))}{\mathcal{L}^n(B(x,r))} \leq |\det(D\Phi(x))|\frac{\mathcal{L}^n(B(x,r+o(r)))}{\mathcal{L}^n(B(x,r))}.$$

Now $\mathcal{L}^n(B(x,R)) = c_n R^n$, so we have

$$\frac{\mu(B(x,r))}{\mathcal{L}^n(B(x,r))} = |\det(D\Phi(x))|\left(\frac{r+o(r)}{r}\right)^n \xrightarrow{r\to 0^+} |\det(D\Phi(x))|.$$

So

$$D_\mu\mathcal{L}^n(x) = |\det(D\Phi(x))| \quad \text{for } \mathcal{L}^n\text{-a.e. } x$$

and we can conclude. □

Lemma 6.4. *Under the assumptions of Theorem 6.2 assume that $f : Y \to \bar{R}$ is measurable. Then*

$$\int_Y f d\mathcal{L}^n = \int_X f\circ\Phi|\det(D\Phi(x))|dx.$$

Proof. First we assume $f \geq 0$. Since Φ is a diffeomorphism (and in particular maps open sets into open sets) $f\circ\Phi$ is $\mathcal{L}^n$-measurable, and so is $f\circ\Phi|\det D\Phi|$.

We approximate f by simple functions

$$f_k(x) := \sum_{i=1}^{k} a_i\chi_{A_i}(x)$$

wher $a_i \geq 0$ and $A_i \subset Y$ are measurable, and

$$f_k(x) \to f(x) \quad \text{for every } x \in Y\text{, and monotonely.}$$

Set $B_i := \Phi^{-1}(A_i)$ (which is measurable) and $B_i \subset X$.

Then

$$\begin{aligned}f_k\circ\Phi(x) =& \sum_{i=1}^{k} a_i\chi_{A_i}\circ\Phi(x)\\ =& \sum_{i=1}^{k} a_i\chi_{B_i}(x)\end{aligned}$$

So with Lemma 6.3,

$$\begin{aligned}\int_Y f_k dx &= \sum_{i=1}^k a_i \mathcal{L}^n(A_i)\\ &= \sum_{i=1}^k a_i \mathcal{L}^n(\Phi(B_i))\\ &= \sum_{i=1}^k a_i \int_{B_i} |\det(D\Phi)|\\ &= \int_X f_k \circ \Phi \int_{B_i} |\det(D\Phi)|\end{aligned}$$

We conclude the case $f \geq 0$ by using monotone convergence theorem.

For general f we set $f = f_+ - f_-$ and use twice the above argument. □

We will not go into applications here (the polar coordinates and related coordinate changes one being the most prominent ones and we did those in Advanced Calculus).

6.1. **Area formula and integration on manifolds.** Let $\mathcal{M} \subset \mathbb{R}^N$ be a (subset of an) n-dimensional manifold. For $f : \mathcal{M} \to \mathbb{R}$ we can then define

$$\int_{\mathcal{M}} f d\mathcal{H}^n := \int_{\mathbb{R}^N} f d\left(\mathcal{M} \llcorner \mathcal{H}^n\right),$$

because f is defined on $(\mathcal{M} \llcorner \mathcal{H}^n)$-a.e. point. In Advanced Calculus I we talked about another way of defining the integral on $\mathcal{M}$:

Assume that there is $\theta_0 \in \mathcal{M}$ and an open neighborhood $U \subset \mathbb{R}^n$ such that $\Phi : U \to \mathcal{M}$ is a parametrization of a neighborhood $\theta_0 \in \partial\Omega$ then we defined

$$\int_{\mathcal{M}\cap\Phi(U)} f(\theta) d\theta := \int_U f(\Phi(z)) |\mathrm{Jac}(\Phi)(z)|\, dz.$$

Here we defined the *Jacobian* $|\mathrm{Jac}(\Phi)(z)|$ as follows

$$|\mathrm{Jac}(\Phi)(z)| := \sqrt{\Big|\det_{\mathbb{R}^{n\times n}} (\underbrace{D\Phi^t(z)}_{\mathbb{R}^{n\times N}} \underbrace{D\Phi(z)}_{\mathbb{R}^{N\times n}})\Big|} \equiv \sqrt{\Big|\det_{\mathbb{R}^{n\times n}} \left((\langle \partial_\alpha \Phi(z), \partial_\beta \Phi(z)\rangle_{\mathbb{R}^N})_{\alpha,\beta=1,\dots,n}\right)\Big|}$$

We want to show that this coincides with the Hausdorff-measure integral. This is usually referred to as the *area formula*.

Theorem 6.5 (Area formula)**.** *Let $\Phi \in C^1(\mathbb{R}^n, \mathbb{R}^N)$ where $n \leq N$. For each $f \in L^1(\mathbb{R}^n)$ and each $\mathcal{L}^n$-measurable subset $A \subset \mathbb{R}^n$ such that $\Phi : A \to \Phi(A)$ is one-to-one*

$$\int_{\Phi(\mathbb{R}^n)\subset\mathbb{R}^N} f \circ \Phi^{-1}(y) d\mathcal{H}^n(y) := \int_{\mathbb{R}^n} f(z) |\mathrm{Jac}(\Phi)(z)|\, d\mathcal{L}^n(z).$$

Remark 6.6.
- There exist generalizations for Φ which are not one-to-one.

- Having $f : \mathbb{R}^n \to \mathbb{R}$ instead of $f : \mathcal{M} \to \mathbb{R}$ avoids dealing with integrability. But this can be done with density arguments

Proof. By the usual approximation arguments (see the transformation rule) it suffices to prove the statement for $f = \chi_A$, so we need to prove for $\mathcal{L}^n$-measurable sets A,

$$\mathcal{H}^n(\Phi(A)) = \int_A |\mathrm{Jac}(\Phi)(z)|\, d\mathcal{L}^n(z).$$

Observe that since Φ is Lipschitz and $\mathcal{H}^n = \mathcal{L}^n$ in $\mathbb{R}^n$ we have that

$$\nu(A) := \mathcal{H}^n(\Phi(A)) \quad A \subset \mathbb{R}^n \text{ is } \mathcal{L}^n\text{-measurable}$$

extends to a Radon measure in $\mathbb{R}^n$. By Radon-Nikodyn we then have

$$\mathcal{H}^n(\Phi(A)) = \nu(A) = \int_A \frac{d\nu}{d\mathcal{L}^n}(z) d\mathcal{L}^n(z).$$

where for μ-a.e. $z \in \mathbb{R}^n$,

$$\frac{d\nu}{d\mathcal{L}^n}(z) = \lim_{r\to 0^+} \frac{\mathcal{H}^n(\Phi(B(z,r)))}{\mathcal{L}^n(B(z,r))}.$$

As in the argument for the transformation rule, since Φ is smooth, $\mathcal{H}^n$ is translation invariant, the claim follows by Tailor's theorem once we can show

$$\mathcal{H}^n(D\Phi(x)A) = \mathrm{Jac}(\Phi)(x)\mathcal{L}^n(A).$$

This is the content of the following Lemma 6.7. □

Lemma 6.7. *Let $L : \mathbb{R}^n \to \mathbb{R}^N$ be a linear map, i.e. $L \in \mathbb{R}^{N\times n}$. Then*

$$\mathcal{H}^n(L(A)) = \sqrt{\det(L^t L)}\mathcal{L}^n(A).$$

Proof. By polar decomposition (cf. [Evans and Gariepy, 2015, Theorem 3.5] there exists a matrix $O \in \mathbb{R}^{N\times n}$ and $S \in \mathbb{R}^{n\times n}$ such that $O \in O(n, N)$, i.e. $O^t O = I_{n\times n}$, i.e.

$$\langle Ox, Oy\rangle_{\mathbb{R}^N} = \langle x, y\rangle_{\mathbb{R}^n} \quad \forall x, y \in \mathbb{R}^n,$$

and

$$S^t = S,$$

and we have

$$L = OS.$$

We then have

$$\det(L^t L) = \det(S^t O^t O S) = \det(S^t S) = (\det(S))^2,$$

that is

$$\sqrt{\det(L^t L)} = |\det(S)|.$$

The conditions on O simply mean that

$$O = (o_1, \dots, o_n) \in \mathbb{R}^{N\times n}$$

for orthonormal vectors $o_i \in \mathbb{R}^N$, $i = 1, \dots, n$. We can extend $(o_i)_{i=1}^n$ to a orthonormal basis $(o_i)_{i=1}^N$ of $\mathbb{R}^N$. Set

$$\tilde{O} := (o_1, \dots, o_n, o_{n+1}, \dots, o_N) \in \mathbb{R}^{N \times N}.$$

Then $\tilde{O} \in O(n)$, i.e. $\tilde{O}^t \tilde{O} = I_{N \times N}$. By Exercise 1.16

$$\mathcal{H}^n(LA) = \mathcal{H}^n(\tilde{O}^t O S A).$$

Now

$$\tilde{O}^t O = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & \ddots & 0 \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \\ 0 & 0 & \dots & 0 \end{pmatrix} = \begin{pmatrix} I_{n \times n} \\ 0_{N-n \times n} \end{pmatrix} \in \mathbb{R}^{N \times n}$$

In particular $\tilde{O}^t L A \subset \mathbb{R}^n \times \{0\}$. So again by Exercise 1.16

$$\mathcal{H}^n(LA) = \mathcal{H}^n(\underbrace{SA}_{\subset \mathbb{R}^n}).$$

Since $\mathcal{H}^n = \mathcal{L}^n$ in $\mathbb{R}^n$ we have by the transformation rule on $\mathbb{R}^n$,

$$\mathcal{H}^n(LA) = \mathcal{L}^n(SA) = |\det(S)| \mathcal{L}^n(A) = \sqrt{\det(L^t L)} \mathcal{L}^n(A).$$

□

Email address: armin@pitt.edu