LEBESGUE DIFFERENTIATION

ANUSH TSERUNYAN

Throughout, we work in the Lebesgue measure space $\left(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d), \lambda\right)$, where $\mathcal{L}(\mathbb{R}^d)$ is the σ-algebra of Lebesgue-measurable sets and λ is the Lebesgue measure.

For $r > 0$ and $x \in \mathbb{R}^d$, let $B_r(x)$ denote the open ball of radius r centered at x.

1. The spaces L^0 and L^1_{loc}

The space L^0 of measurable functions. By a *measurable function* we mean an $\mathcal{L}(\mathbb{R}^d)$-measurable function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ and we let $L^0(\mathbb{R}^d)$ (or just L^0) denote the vector space of all measurable functions modulo the usual equivalence relation $=_{\text{a.e.}}$ of a.e. equality.

There are two natural notions of convergence (topologies) on L^0: *a.e. convergence* and *convergence in measure*. As we know, these two notions are related, but neither implies the other. However, convergence in measure of a sequence implies a.e. convergence of a subsequence, which in many situations can be boosted to a.e. convergence of the entire sequence (recall Problem 3 of Midterm 2).

Convergence in measure is captured by the following family of norm-like maps: for each $\alpha > 0$ and $f \in L^0$,

$$\|f\|_\alpha^* := \alpha \cdot \lambda\Big(\Delta_\alpha(f)\Big),$$

where $\Delta_\alpha(f) := \Big\{x \in \mathbb{R}^d : |f(x)| > \alpha\Big\}$. Note that $\|f\|_\alpha^*$ can be infinite.

Observation 1 (Chebyshev's inequality). *For any $\alpha > 0$ and any $f \in L^0$, $\|f\|_\alpha^* \leqslant \|f\|_1$.*

The space L^1_{loc} of locally integrable functions. Differentiation at a point $x \in \mathbb{R}^d$ is a local notion as it only depends on the values of the function in some open neighborhood of x, so, in this context, it is natural to work with a class of functions defined via a local property, as opposed to global (e.g. being in $L^1(\mathbb{R}^d)$).

Definition 2. Call a measurable function f *locally integrable* if it is integrable on every bounded subset of $\mathbb{R}^d$, i.e. for any $r > 0$, $f \mathbb{1}_{B_r(x)} \in L^1$.

Denote by $L^1_{\mathrm{loc}}(\mathbb{R}^d)$ the vector space of all locally integrable functions modulo the equivalence relation $=_{\text{a.e.}}$. Obviously, $L^1 \subseteq L^1_{\mathrm{loc}}$. Moreover, letting $C(\mathbb{R}^d)$ denote the space of continuous functions on $\mathbb{R}^d$, we have $C(\mathbb{R}^d) \subseteq L^1_{\mathrm{loc}}$ because each $f \in C(\mathbb{R}^d)$ is bounded on every ball.

The natural notion of convergence in L^1_{loc} is the *local L^1-convergence*, namely,

$$f_n \to_{L^1_{\mathrm{loc}}} f :\Leftrightarrow \forall r > 0 \left(f_n \mathbb{1}_{B_r(0)}\right) \to_{L^1} \left(f \mathbb{1}_{B_r(0)}\right).$$

Clearly, $f_n \to_{L^1} f_n$ implies $f_n \to_{L^1_{\mathrm{loc}}} f$.

Date: August 27, 2018.

2. Averaging operators

For $r \geqslant 0$, $f \in L^1_{\mathrm{loc}}$, and $x \in \mathbb{R}^d$, we define

$$A_r f(x) := \frac{1}{\lambda\Big(B_r(x)\Big)} \int_{B_r(x)} f \, d\lambda = \frac{1}{\lambda\Big(B_1(0)\Big)} \frac{1}{r^d} \int_{B_r(x)} f \, d\lambda.$$

We refer to $A_r f(x)$ as the *average of f at x within radius r*.

Proposition 3. *Let $r > 0$.*

(3.a) *A_r is a linear transformation from L^1_{loc} to $C(\mathbb{R}^d)$. Moreover, for any $f \in L^1_{\mathrm{loc}}$, the map $(r,x) \mapsto A_r f(x)$ is continuous as a function $(0,\infty) \times \mathbb{R}^d \to \mathbb{R}$.*

(3.b) *A_r is continuous with respect to L^1_{loc}-convergence on L^1_{loc} and pointwise convergence on $C(\mathbb{R}^d)$, i.e. if $f_n \to_{L^1_{\mathrm{loc}}} f$, then $A_r f_n \to A_r f$ pointwise.*

Proof. Part (3.a) follows by the absolute continuity of $f \mathbb{1}_B$ in L^1-norm, where B is any ball (bounded set). Part (3.b) is obvious because $f_n \to_{L^1_{\mathrm{loc}}} f$ implies $\int_{B(x,r)} f_n d\lambda \to \int_{B(x,r)} f d\lambda$. $\square$

3. Vitali covers

Lemma 4 (Vitali)**.** *Any finite collection $\mathcal{C}$ of open balls in $\mathbb{R}^d$ admits a subcollection $\mathcal{C}' \subseteq \mathcal{C}$ of pairwise disjoint balls, whose union is a constant proportion of the total union, more precisely:*

$$\lambda\Big(\bigsqcup \mathcal{C}'\Big) \geqslant 3^{-d} \lambda\Big(\bigcup \mathcal{C}\Big).$$

Proof. Let B_0 be a ball in $\mathcal{C}$ with maximum radius. If there is no ball in $\mathcal{C}$ disjoint from B_0, stop; otherwise let B_1 be a ball of maximum radius among those balls in $\mathcal{C}$ that are disjoint from B_0. Continue. If there is no ball in $\mathcal{C}$ disjoint from $B_0 \sqcup B_1$, stop. Otherwise, let B_2 be a ball of maximum radius among those balls in $\mathcal{C}$ that are disjoint from $B_0 \sqcup B_1$. And so on and so forth. Because $\mathcal{C}$ is finite, this process terminates, i.e. there is $n \in \mathbb{N}$ such that there is no ball in $\mathcal{C}$ disjoint from $\bigsqcup_{i<n} B_i$ and we put $\mathcal{C}' := \{B_i : i < n\}$.

To check that $\mathcal{C}'$ is as desired, it is enough to show that if all balls in $\mathcal{C}'$ are made 3 times bigger, than they cover $\bigcup \mathcal{C}$. More precisely, letting B_i' be the ball with the same center as B_i but with radius 3 times bigger than that of B_i, we want to show that $\bigcup_{i<n} B_i' \supseteq \bigcup \mathcal{C}$.

To this end, fix a ball $B \in \mathcal{C}$ and suppose it is not one of the balls in $\mathcal{C}'$, so it must intersect $\bigsqcup_{i<n} B_i$ (otherwise, the above process wouldn't stop at n). Let i be the least such that $B \cap B_i \neq 0$. By the choice of B_i at the step i of the above algorithm, the radius of B is not bigger than that of B_i, so $B_i' \supseteq B$ and we are done. $\square$

Lemma 5 (Vitali Covering)**.** *Let a measurable $A \subseteq \mathbb{R}^d$ be covered by a (possibly infinite) collection $\mathcal{C}$ of open balls. For every $0 \leqslant a < \lambda(A)$, there is a finite subcollection $\mathcal{C}_a \subseteq \mathcal{C}$ of pairwise disjoint balls such that*

$$\lambda\Big(\bigsqcup \mathcal{C}_a\Big) \geqslant 3^{-d} a.$$

Proof. By the inner regularity of the Lebesgue measure (Problem 4(c) of Midterm 1), there is a compact set $K \subseteq A$ with $\lambda(K) \geqslant a$. $\mathcal{C}$ is an open cover of K, so there is a finite subcover $\mathcal{C}_K \subseteq \mathcal{C}$ of K. Apply Lemma 4 to $\mathcal{C}_K$ to obtain a pairwise disjoint subcollection $\mathcal{C}_a \subseteq \mathcal{C}_K$ as desired. $\square$

The rest of this section is devoted to a proof of a stronger form of the last lemma, which is not used anywhere in this note, but is very useful in general. We start with an easy observation.

Lemma 6 (Outer measure via open covers)**.** *For $A \subseteq \mathbb{R}^d$, $\lambda^*(A) < r$ if and only if there is an open set $U \supseteq A$ with $\lambda(U) < r$.*

Proof. $\Leftarrow$: Follows from the fact that any open set U is a countable disjoint union of boxes. To see the latter, recall that open boxes form a bases for $\mathbb{R}^d$ and write U as a countable union of open boxes $U = \bigcup_n B_n$. Make these disjoint by replacing B_n with $B_n' := B_n \setminus \bigcup_{i<n} B_i$ and observe that B_n' itself is a disjoint union of finitely many boxes.
$\Rightarrow$: Let $\{B_n\}$ be a countable cover of A with boxes witnessing $\lambda^*(A) < r$, i.e. $s := \sum_n \lambda(B_n) < r$. Replacing each box B_n by an open box $B_n' \supseteq B_n$ with $\lambda(B_n') < \lambda(B_n) + 2^{-(n+2)}(r-s)$, we see that $U := \bigcup_n B_n'$ is open, contains A, and has measure at most $s + \sum_n 2^{-(n+2)}(r-s) < r$. □

Definition 7. A collection $\mathcal{C}$ of open balls in $\mathbb{R}^d$ is called a *Vitali cover* of a set $A \subseteq \mathbb{R}^d$ if for each $x \in A$ there are arbitrarily small balls in $\mathcal{C}$ containing x.

Lemma 8 (Strong Vitali Covering)**.** *Let $\mathcal{C}$ be a Vitali cover of a (not necessarily measurable) set $A \subseteq \mathbb{R}^d$ with $\lambda^*(A) < \infty$. For every $\varepsilon > 0$, there is a pairwise disjoint finite subcollection $\mathcal{C}_\varepsilon \subseteq \mathcal{C}$ such that $\lambda^*(A \setminus \bigcup \mathcal{C}_\varepsilon) < \varepsilon$. Equivalently (by Lemma 6), there is an open set $U \supseteq A \setminus \bigcup \mathcal{C}_\varepsilon$ with $\lambda(U) < \varepsilon$.*

Proof. By Lemma 6, there is an open set $V \supseteq A$ with finite measure. The subcollection of $\mathcal{C} \cap \mathscr{P}(V)$ is still a Vitali cover of A, so replacing $\mathcal{C}$ by it, we may assume that $\mathcal{C} \subseteq \mathscr{P}(V)$.

We recursively define a sequence $(B_n)_{n\in\mathbb{N}} \in \mathcal{C}$ by letting $B_n \in \mathcal{C}$ be disjoint from $\bigcup_{i<n} B_i$ and of radius at least half of what's available, i.e.

$$\frac{1}{2} \sup \left\{ \text{raduis}(B) : B \in \mathcal{C} \text{ disjoint from } \bigcup_{i<n} B_i \right\}.$$

If for some n, such a B_n does not exist, i.e. the above supremum is 0, then we are done by taking $\mathcal{C}_\varepsilon := \{B_0, \dots, B_{n-1}\}$. Thus, suppose that the above procedure defined an infinite sequence $(B_n)_{n\in\mathbb{N}} \subseteq \mathcal{C}$ of necessarily pairwise disjoint sets. Because $\sum_n \lambda(B_n) = \lambda(\bigsqcup_n B_n) \leqslant \lambda(V) < \infty$, there is $N \in \mathbb{N}$ such that $\sum_{n \geqslant N} \lambda(B_n) < \frac{\varepsilon}{5^d}$. We show that $\mathcal{C}_\varepsilon := \{B_0, \dots, B_{N-1}\}$ is as desired. Because a boundary of a ball is null, it is enough to show that the outer measure of $A' := A \setminus \bigcup_{n<N} \overline{B_n}$ is less than ε, which we do by building an open set $U \supseteq A'$ with $\lambda(U) < \varepsilon$ and using Lemma 6. To this end, let

$$\mathcal{C}' := \left\{ B \in \mathcal{C} : B \cap \bigcup_{n<N} \overline{B_n} = \emptyset \right\}$$

and $U := \bigcup \mathcal{C}'$. Note that $U \supseteq A'$ since for each $x \in A'$ there is a ball $B \in \mathcal{C}'$ containing x because $\bigcup_{n<N} \overline{B_n}$ is closed and $\mathcal{C}$ is a Vitali cover.

Claim. $U \subseteq \bigcup_{n \geqslant N} \tilde{B}_n$, where $\tilde{B}_n$ is the ball with the same center as B_n but with radius 5 times that of B_n.

Proof of Claim. Each $B \in \mathcal{C}'$ must intersect a B_n for some $n \geqslant N$: indeed, because the measures of the B_n are summable, there is $m \geqslant N$ with $\lambda(B_m) < \frac{1}{2}\lambda(B)$, so we had no good reason to not include B in our sequence unless it intersected one of the B_n for $n < m$.

Letting $n \in \mathbb{N}$ be the least such that $B \cap B_n \neq \emptyset$, the choice of B_n in the recursive construction ensures that $\text{radius}(B) \leqslant 2\text{radius}(B_n)$. Thus, B is contained in $\tilde{B}_n$. ⊠

Thus, $\lambda(U) \leqslant \sum_{n \geqslant N} \lambda\left(\tilde{B}_n\right) = 5^d \sum_{n \geqslant N} \lambda(B_n) < \varepsilon$. □

4. Hardy–Littlewood maximal function

Heuristically speaking, in order to prove the Lebesgue Differentiation Theorem, we need to approximate an L^1_{loc}-function f with a continuous function g so that $|A_r f(x) - A_r g(x)|$ is small uniformly in r. We achieve this below by proving a modified version of (3.b), where L^1_{loc} is replaced by L^1 and pointwise convergence by convergence in measure.

The uniformity in r is captured by the following operator: for $f \in L^1_{\text{loc}}$ and $x \in \mathbb{R}^d$, put

$$\bar{A} f(x) := \sup_{r \leqslant 1} A_r |f|(x).$$

The function $\bar{A} f$ is known as the *Hardy–Littlewood maximal function*. Being a sup of measurable functions, $\bar{A} f$ is measurable, so $\bar{A}$ is a map from L^1_{loc} to L^0. We would like to show that it is a Lipschitz map with respect to the norm-like functions $\|\cdot\|^*_\alpha$ (uniformly in α).

Theorem 9 (The Maximal Theorem)**.** *The map $\bar{A} : L^1 \to L^0$ is 3^d-Lipschitz, i.e. $\forall f \in L^1$,*

$$\sup_{\alpha \geqslant 0} \|\bar{A} f\|^*_\alpha \leqslant 3^d \, \|f\|_1.$$

Proof. Fix $\alpha > 0$ and $f \in L^1$, and let $S := \Delta_\alpha(|f|) = \left\{x \in \mathbb{R}^d : |f(x)| > \alpha\right\}$. By definition, for each $x \in S$, there is a ball B_x centered at x of radius $\leqslant 1$ such that

$$\int_{B_x} |f| d\lambda > \alpha \cdot \lambda(B_x). \tag{10}$$

Thus, the balls B_x form a cover of S, so, for any fixed $s < \lambda(S)$, the Vitali Covering Lemma (Lemma 5) gives a finite subcollection $\mathcal{C}_s$ of pairwise disjoint balls with

$$\lambda\left(\bigsqcup \mathcal{C}_s\right) \geqslant 3^{-d} s.$$

Hence, we obtain

$$\begin{aligned}
\alpha \cdot s &\leqslant 3^d \cdot \alpha \cdot \lambda\left(\bigsqcup \mathcal{C}_s\right) \\
\left[\text{by the disjointness of } \mathcal{C}_s\right] \quad &= 3^d \cdot \sum_{B \in \mathcal{C}_s} \alpha \cdot \lambda(B) \\
\left[\text{by (10)}\right] \quad &\leqslant 3^d \sum_{B \in \mathcal{C}_s} \int_B |f| d\lambda \\
\left[\text{by the disjointness of } \mathcal{C}_s\right] \quad &= 3^d \int_{\bigsqcup \mathcal{C}_s} |f| d\lambda \leqslant 3^d \|f\|_1.
\end{aligned}$$

Since s is an arbitrary real less than $\lambda(S)$, we get $\|\bar{A} f\|^*_\alpha := \alpha \lambda(S) \leqslant 3^d \|f\|_1$. □

5. Lebesgue differentiation

Lemma 11 (Lebesgue differentiation for continuous functions). *For any $g \in C(\mathbb{R}^d)$, $A_r g(x) \to g(x)$ pointwise as $r \to 0^+$.*

Proof. For any $x \in \mathbb{R}^d$ and $r \leqslant 1$, $|A_r g(x) - g(x)| \leqslant \int_{B_r(x)} |g(y) - g(x)| d\lambda(y)$, so the continuity of g yields the conclusion. □

Theorem 12 (Lebesgue Differentiation Theorem). *For any $f \in L^1_{\text{loc}}$, $A_r f \to f$ a.e. as $r \to 0^+$.*

Proof. It is enough to show that this holds in every ball $B_N(0)$, $N \in \mathbb{N}$, so, replacing f with $f\mathbb{1}_{B_N(0)}$, we may assume that $f \in L^1$.

For each $\alpha \geqslant 0$, put

$$D_\alpha := \left\{ x \in \mathbb{R}^d : \limsup_{r \to 0} |A_r f(x) - f(x)| > \alpha \right\}.$$

What we need to show is that D_0 is null. But $D_0 = \bigcup_{n \geqslant 1} D_n$, so it is enough to show D_α is null for each $\alpha > 0$, so fix $\alpha > 0$, as well as $\varepsilon > 0$.

By the density of continuous functions in L^1, there is $g \in C(\mathbb{R}^d)$ with $\|f - g\|_1 < \varepsilon$. For each $x \in \mathbb{R}^d$,

$$\begin{aligned} \limsup_{r \to 0} |A_r f(x) - f(x)| &\leqslant \bar{A}(f-g)(x) + \limsup_{r \to 0} |A_r g(x) - g(x)| + |f-g|(x) \\ \Big[\text{by Lemma 11}\Big] \quad &= \bar{A}(f-g)(x) + |f-g|(x), \end{aligned}$$

so, for each $x \in D_\alpha$, $\bar{A}(f-g)(x) \geqslant \frac{\alpha}{2}$ or $|f-g|(x) \geqslant \frac{\alpha}{2}$, in other words,

$$D_\alpha \subseteq \Delta_{\frac{\alpha}{2}}\Big(\bar{A}(f-g)\Big) \cup \Delta_{\frac{\alpha}{2}}(f-g).$$

By Chebyshev's inequality $\lambda\Big(\Delta_{\frac{\alpha}{2}}(f-g)\Big) = \frac{2}{\alpha}\|f-g\|^*_{\frac{\alpha}{2}} \leqslant \frac{2}{\alpha}\|f-g\|_1 < \frac{2}{\alpha}\varepsilon$. Furthermore, by the Maximal Theorem, $\lambda\Big(\Delta_{\frac{\alpha}{2}}\Big(\bar{A}(f-g)\Big)\Big) = \frac{2}{\alpha}\|\bar{A}(f-g)\|^*_{\frac{\alpha}{2}} \leqslant \frac{3^d \cdot 2}{\alpha}\|f-g\|_1 < \frac{3^d \cdot 2}{\alpha}\varepsilon$. Thus, $\lambda(D_\alpha) \leqslant \frac{3^d \cdot 2}{\alpha}\varepsilon + \frac{2}{\alpha}\varepsilon$, but ε is arbitrary, so $\lambda(D_\alpha) = 0$. □

Corollary 13 (Stronger version). *Let $f \in L^1_{\text{loc}}$. For a.e. $x \in \mathbb{R}^d$,*

$$\lim_{r \to 0^+} \frac{1}{\lambda(B_r(x))} \int_{B_r(x)} |f(y) - f(x)| d\lambda(y) = 0.$$

Proof. Follows by applying Theorem 12 to $|f - q|$ for every rational $q \in \mathbb{Q}$ and using the density of $\mathbb{Q}$ in $\mathbb{R}$. □

6. Lebesgue density

We switch the notation $\mathbb{R}^d$ to $\mathbb{R}^n$ because the letter d is used for something else.

Definition 14. For a measurable set $S \subseteq \mathbb{R}^n$, define its *Lebesgue density function* $d_S : \mathbb{R}^n \to [0,1]$ by

$$d_S(x) := \lim_{r \to 0^+} \frac{\lambda\Big(S \cap B_r(x)\Big)}{\lambda\Big(B_r(x)\Big)}$$

if this limit exists, and leaving it undefined otherwise. Call $S' := \{x \in \mathbb{R}^n : d_S(x) = 1\}$ the set of density points of S.

The Lebesgue Differentiation Theorem applied to characteristic functions gives:

Corollary 15 (Lebesgue Density Theorem)**.** *For every measurable set $S \subseteq \mathbb{R}^n$, $d_S = \mathbb{1}_S$ a.e.*

For any sets $S_0, S_1 \subseteq \mathbb{R}^n$, write $S_0 =_\lambda S_1$ if $S_0 \triangle S_1$ is λ-null. Let $[S_0]_\lambda$ denote the $=_\lambda$-equivalence class of S_0. The last corollary provides a canonical choice of a measurable set out of each such class:

Corollary 16 (Selector for $=_\lambda$)**.** *For any measurable $S \subseteq \mathbb{R}^n$, $S' =_\lambda S$ and $S'' = S'$. In particular, the map $S \mapsto S'$ is a selector for $=_\lambda$ on measurable sets, i.e. for any measurable sets $S_0, S_1 \subseteq \mathbb{R}^n$,*

$$S_0 =_\lambda S_1 \iff S_0' = S_1'.$$

A further corollary of Corollary 15 is the following.

Corollary 17 (The 99% Lemma)**.** *Let $S \subseteq \mathbb{R}^n$ be a measurable set with $\lambda(S) > 0$. For every $\alpha \in (0,1)$* (e.g. $\alpha = 0.99$)*, there is a nonempty open ball $B \subseteq \mathbb{R}^n$ such that S occupies at least α fraction of B, i.e.*

$$\lambda(S \cap B) \geqslant \alpha \lambda(B).$$