MAT205a, Fall 2019 Part III: Differentiation

Lecture 7, Following Folland, ch 3.1, 3.2

1. Signed measures

1.1. **Definition and examples.** We now extend the notion of measure, and consider non necessarily positive functions on subsets of X.

Definition 1.1. *Suppose that $(X, \mathcal{M})$ is a measurable spaces, a function $\nu : \mathcal{M} \to [-\infty, +\infty]$ is called a* **signed measure** *if $\nu(\emptyset) = 0$, ν talkes at most one of the values $-\infty$ and $+\infty$, and ν is countably additive. If $\{E_j\}$ is a sequence of disjoint sets in $\mathcal{M}$ then $\nu(\cup_j E_j) = \sum_j \nu(E_j)$.*

Example 1.1. If μ_1 and μ_2 are measures, and one of them is finite, then $\mu_1 - \mu_2$ is a signed measure.

Example 1.2. If f is a measurable function on $(X, \mathcal{M})$, μ is a measure on $\mathcal{M}$ and $\int f d\mu$ makes sense. Then $\nu(E) = \int_E f d\mu$ is a signed measure.

The following properties of measures can be easily extended to signed measures.

Proposition 1.1. *Let ν be a signed measure on $(X, \mathcal{M})$.*
If $\{E_j\}$ is an increasing sequence, $E_j \in \mathcal{M}, E_j \subset E_{j+1}$ then $\nu(\cup_j E_j) = \lim_{j\to\infty} \nu(E_j)$.
If $\{E_j\}$ is a decreasing sequence, $E_j \in \mathcal{M}$, $E_j \supset E_{j+1}$ and $\nu(E_1)$ is finite, then $\nu(\cap_j E_j) = \lim_{j\to\infty} \nu(E_j)$.

1.2. **Hahn decomposition.** We will show that any signed measure is a difference of two measures (as in Example 1.1 above).

Definition 1.2. *A set $E \in \mathcal{M}$ is said to be* **positive** *for a signed measure ν if for any $F \subset E$ such that $F \in \mathcal{M}$ we have $\nu(F) \geq 0$.*
Similarly $C \in \mathcal{M}$ is called **negative** *if for any $B \subset C$, $B \in \mathcal{M}$ we have $\nu(B) \neq 0$.*

It is clear that any measurable subset of a positive set is positive and countable union of positive sets is also positive. The empty set is both positive and negative.

Theorem 1.1 (Hahn Decomposition). *Let ν be a signed measure on $(X, \mathcal{M})$. Then there exist a positive set P and a negative set N for ν such that $P \cup N = X$ and $P \cap N = \emptyset$.*

Proof. We assume that ν does not take value $+\infty$. Let $m = \sup\{\nu(E) : E \in \mathcal{M}\}$. Then there exits a sequence of sets $E_j \in \mathcal{M}$ such that $\nu(E_j) \to m$.

We consider algebras $\mathcal{A}_n$ generated by the sets $E_1, ..., E_n$. Let now $F_n \in \mathcal{A}_n$ be such that $\nu(F_n) = \max_{F\in\mathcal{A}_n} \nu(F)$, then $\nu(E_n) \leq \nu(F_n) \leq m$. For any $D \in \mathcal{A}_n$ such that $D \cap F_n = \emptyset$ we know that $\nu(D) \leq 0$ and for any $C \in \mathcal{A}_n$ such that $C \subset F_n$ we have $\nu(C) \geq 0$. Let now $B_m = \cap_m^\infty F_n$. Then $\nu(B_m) = \lim_{n\to\infty} \nu(\cap_m^n F_k)$. We have $D_{m,n} = \cap_m^{n-1} F_k \setminus F_n \in \mathcal{A}_n$ and $\nu(D_{m,n}) \leq 0$ since $D_{m,n} \cap F_n = \emptyset$. Thus $\nu(\cap_m^n F_k) \geq \nu(\cap_m^{n-1} F_k) \geq \nu(F_m)$ and $\nu(B_m) \geq \nu(F_m)$. Now consider the set $P = \cup_m B_m$, then $\nu(P) = m$.

If $E \subset P$ then $\nu(P \setminus E) = \nu(P) - \nu(E)$ and thus $\nu(E) \geq 0$ and if $F \subset P^c$ then $\nu(F \cup P) = \nu(F) + \nu(P)$ and $\nu(F) \leq 0$. It means that P is a positive set and its complement, $N = X \setminus P$ is negative. □

The decomposition given by the theorem is called a Hahn decomposition, it is usually not unique, but if $X = P \cup N = P' \cup N'$ where $P \cap N = P' \cap N' = \emptyset$ and P, P' are positive sets and N, N' are negative, then $\nu(N \Delta N') = \nu(P \Delta P') = 0$. To see this, note that for example $N \setminus N' \subset N$ and $N \setminus N' \subset P'$ thus $N \setminus N'$ is both positive and negative set, therefore $\nu(N \setminus N') = 0$, moreover if $C \subset N \setminus N'$ then $\nu(C) = 0$.

1.3. **Jordan decomposition.**

Definition 1.3. *We say that two (positive) measures μ_1 and μ_2 on $(X, \mathcal{M})$ are* **mutually singular** *if there exist sets E_1 and E_2 in $\mathcal{M}$ such that $\mu_1(E_2) = 0$, $\mu_2(E_1) = 0$, $E_1 \cup E_2 = X$ and $E_1 \cap E_2 = \emptyset$.*

If two measures are mutually singular, we write $\mu_1 \perp \mu_2$. For example the Lebesgue measure m on $(\mathbb{R}, \mathcal{B})$ and the Dirac measure δ_0 are mutually singular.

We say that two signed measures λ_1 and λ_2 are mutually singular if there exist E_1 and E_2 in $\mathcal{M}$, $E_1 \cap E_2 = \emptyset$, $E_1 \cup E_2 = X$ such that for any $F_1 \subset E_1$, $F_1 \in \mathcal{M}$ we have $\lambda_2(F_1) = 0$ and similarly, if $F_2 \subset E_2$, $F_2 \in \mathcal{M}$, then $\lambda_1(F_2) = 0$.

Now we are ready to show that each signed measure is the difference of two (positive) measures.

Theorem 1.2. *Let ν be a signed measure on $(X, \mathcal{M})$ then there exist mutually singular (positive) measures ν_1 and ν_2 such that $\nu = \nu_1 - \nu_2$, such decomposition is unique.*

Proof. Let $X = N \cup P$ be a Hahn decomposition, define $\nu_1(E) = \nu(E \cap P)$ and $\nu_2(E) = -\nu(E \cap N)$. Clearly, ν_1 and ν_2 are positive measures and $N \cap P = \emptyset$, so the measures are mutually singular.

Suppose that $\nu = \mu_1 - \mu_2$ is another decomposition where μ_1 and μ_2 are mutually singular, $\mu_1(N') = 0$ and $\mu_2(P') = 0$. Then P' is a positive set for ν and N' is negative set. As we saw above then $\nu(P \Delta P') = \nu(N \Delta N') = 0$ moreover, $\nu(E) = 0$ when $E \subset P \Delta p'$ or $E \subset N \Delta N'$. We have

$$\mu_1(E) = \mu_1(E \cap P') = \nu(E \cap P') = \nu(E \cap P) = \nu_1(E),$$

for any $E \in \mathcal{M}$. Thus $\mu_1 = \nu_1$ and $\mu_2 = \nu_2$. □

We say that ν_1 and ν_2 constructed in the last Theorem are the positive and negative parts of the signed measure ν, and we denote them ν^+ and ν^-. The decomposition $\nu = \nu^- - \nu^-$ is called the Jordan decomposition. The positive measure $|\nu| = \nu^+ + \nu^-$ is called the **total variation** of ν.

2. Absolutely continuous measures

2.1. Definition and examples.

Definition 2.1. *Let ν be a signed measure on $(X, \mathcal{M})$ and μ be a positive measure on $(X, \mathcal{M})$. We say that ν is* **absolutely continuous** *with respect to μ and write $\nu \ll \mu$ if $\nu(E) = 0$ for any $E \in \mathcal{M}$ such that $\mu(E) = 0$.*

Example 2.1. Let f be a measurable function on $(X, \mathcal{M})$ such that $\int f d\mu$ makes sense, define $\nu(E) = \int_E f d\mu$. then $\nu \ll \mu$.

If μ is a measure and ν is a signed measure such that $\nu \perp \mu$ and $\nu \ll \mu$ then $\nu = 0$. If ν is a signed measure and μ is a measure then $\nu \ll \mu$ if and only if $|\nu| \ll \mu$.

Proposition 2.1. *Let ν be a finite signed measure and let μ be a measure on $(X, \mathcal{M})$. Then $\nu \ll \mu$ if and only if for any $\varepsilon > 0$ there exists $\delta > 0$ such that $|\nu(E)| < \varepsilon$ when $\mu(E) < \delta$.*

Proof. We may assume that ν is positive, replacing ν by $|\nu|$. It is clear that if for any $\varepsilon > 0$ there exists δ as in the proposition, then for $\mu(E) = 0$ we have $\nu(E) = 0$.

To show the inverse implication, assume that $\nu \ll \mu$ and for a given $\varepsilon > 0$ there is a sequence of sets E_n such that $\nu(E_n) > \varepsilon$ and $\mu(E_n) < 2^{-n}$. Now let $E = \cap_n \cup_n^\infty E_k$, then

$$\mu(E) = \lim_{n\to\infty} \mu(\cup_n^\infty E_k) \leq \lim_{n\to\infty} \sum_n^\infty \mu(E_k) = 0.$$

On the other hand, $\nu(E) = \lim_{n\to\infty} \nu(\cup_n^\infty E_k) \geq \limsup_{n\to\infty} \nu(E_n) \geq \varepsilon$. It contradicts to the assumption that $\nu \ll \mu$. □

Corollary 2.1. *If $f \in L^1(\mu)$ then for any $\varepsilon > 0$ there exists δ such that $\left|\int_E f d\mu\right| < \varepsilon$ when $\mu(E) < \delta$.*

It follows from he proposition above, considering $\nu(E) = \int_E f d\mu$.

2.2. **Lebesgue decomposition.** We fix a measure μ on $(X, \mathcal{M})$ and decompose another signed measure ν into two parts, one is singular to μ and another is absolutely continuous with respect to μ. We assume that measures are σ-finite.

Lemma 2.1. *Assume that μ and ν are finite measures on $(X, \mathcal{M})$. Then either $\nu \perp \mu$, or there exist $\varepsilon > 0$ and a set $E \in \mathcal{M}$ such that $\mu(E) > 0$ and for any $F \subset E$, $F \in \mathcal{M}$, $\nu(F) \geq \varepsilon\mu(F)$.*

Proof. For each positive integer m consider the signed measure $\nu - m^{-1}\mu$ and let $X = P_m \cup N_m$ be a Hahn decomposition for ν_m. Further let $N = \cap_m N_m$, then $N \subset N_m$ and $\nu - m^{-1}\mu$ is negative on N. Therefore $\nu(N) \leq m^{-1}\mu(N)$ and $\nu(N) = 0$. We have $N^c = \cup_m P_m = P$. If $\mu(P) = 0$ then $\nu \perp \mu$. Otherwise there exists P_m such that $\mu(P_m) > 0$ but on P_m the measure $\nu - m^{-1}\mu$ is positive and for $F \subset P_m$, $F \in \mathcal{M}$, we have $\nu(F) \geq m^{-1}\mu(F)$. $\square$

Theorem 2.1 (Lebesgue decomposition)**.** *Let ν be a σ-finite signed measure on $(X, \mathcal{M})$ and μ be a σ-finite measure on $(X, \mathcal{M})$. Then there exist unique measures ν_a and ν_s such that $\nu = \nu_a + \nu_s$, $\nu_a \ll \mu$ and $\nu_s \perp \mu$.*

Proof. It is enough to consider the case when μ and ν are finite and ν is positive. We consider the measure λ defined on $\mathcal{M}$ by

$$\lambda(E) = \begin{cases} +\infty, & \text{if } \mu(E) > 0, \\ -\nu, & \text{otherwise} \end{cases}.$$

We think about λ as $\infty \cdot \mu - \nu$. Let $X = P \cup N$ be a Hahn decomposition for λ. Then $\mu(N) = 0$. We define $\nu_s(E) = \nu(E \cap N)$ and $\nu_a(E) = \nu(E \cap P)$. Then it is clear that $\nu = \nu_a + \nu_s$ and $\nu_s \perp \mu$. Assume that $\mu(C) = 0$ then $\lambda(D) \leq 0$ for any $D \subset C$. Thus $\nu(C \cap P) = 0$ and $\nu_a(C) = 0$, so $\nu_a \ll \mu$.

To show that the decomposition is unique, assume that $\nu = \nu_a + \nu_s = \nu'_a + \nu'_s$. Then consider $\gamma = \nu_a - \nu'_a = \nu_s - \nu'_s$, $\gamma \perp \mu$ and $\gamma \ll \mu$. Therefore $\gamma = 0$. $\square$

2.3. **Radon-Nikodym derivative.** Now we describe the measures absolutely continuous with respect to the given one.

Theorem 2.2 (Radon-Nikodym)**.** *Let μ and ν be σ-finite measures on $(X, \mathcal{M})$ (μ is positive and ν is signed), such that $\nu \ll \mu$. Then there exists a function $f : X \to [-\infty, \infty]$ such that $\int f d\mu$ makes sense and $\nu(E) = \int_E f\, d\mu$ for any $E \in \mathcal{M}$.*

Proof. We assume that μ and ν are finite positive measures. Let

$$\mathcal{F} = \{f : X \to [0, +\infty], \text{ measurable}, \int_E f\, d\mu \leq \nu(E) \text{ for any } E \in \mathcal{M}\}.$$

Clearly $0 \in \mathcal{F}$. If $f, g \in \mathcal{F}$ then $h = \max\{f, g\} \in \mathcal{F}$. Indeed, if $A = \{f > g\}$ then

$$\int_E h d\mu = \int_{E\cap A} f\, d\mu + \int_{E\setminus A} g\, d\mu \leq \nu(E \cap A) + \nu(E \setminus A) = \nu(E).$$

Next, let $a = \sup\{\int f d\mu, f \in \mathcal{F}\}$. There exists a sequence $\{f_n\} \subset \mathcal{F}$ such that $\lim_{n\to\infty} \int f d\mu = a$. By replacing f_n by $\max\{f_1, ..., f_n\}$, we may assume that f_n is an increasing sequence. Let $f = \lim_n f_n$, by the monotone convergence theorem, $f \in \mathcal{F}$ and $\int f\, d\mu = a$. Remark that $a = \int f d\mu \leq \nu(X) < \infty$.

Define the measure $\lambda(E) = \nu(E) - \int_E f\, d\mu$, by the definition of f, $\lambda(E) \geq 0$ and $\lambda \ll \mu$. If λ is not a zero measure, then there exists a set F such that $\mu(F) > 0$ and $\lambda > \varepsilon\mu$ on F. Let $g = f + \varepsilon\chi_F$, then $g \in \mathcal{F}$ and $\int g\, d\mu = \int f\, d\mu + \varepsilon\mu(F) > a$. □

MAT205a, Fall 2019 Part III: Differentiation

Lectures 8-9, Following Folland, ch 2.6, 3.4

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

3.1. **Regularity of the Lebesgue measure.** Let $m = m_n$ be the Lebesgue measure on the Lebesgue σ-algebra $\mathcal{L}_n$ of subsets of $\mathbb{R}^n$.

Theorem 3.1. *If $E \in \mathcal{L}_n$ then*

$$m(E) = \inf\{m(U) : E \subset U, U \text{ open}\} = \sup\{m(K) : K \subset E, K \text{ compact}\}.$$

Proof. First, there exists a sequence of rectangles $\{R_n\}$ such that $E \subset \cup_n R_n$ and $\sum_n m(R_n) \leq m(E) + \varepsilon$. For each rectangular set $R_j = E_{1,j} \times E_{2,j} \times ... \times E_{n,j}$, using the regularity of the one-dimensional Lebesgue measure, we can find an open set $U_j = U_{1,j} \times ... \times U_{n,j}$ such that $R_j \subset U_j$ and $m(U_j) < m(r_j) - \varepsilon 2^{-j}$. Now let $U = \cup_j U_j$, then $E \subset U$ and $m(U) < m(E) + 2\varepsilon$.

If E is bounded, we have $E \subset B$ for some open ball B. Consider $F = B \setminus E$. Take $\varepsilon > 0$, then here exists an open set O such that $F \subset O$ and $m(O) < m(F) + \varepsilon$. Let $K = \bar{B} \setminus O$, then K is compact and $K \subset E$. Furthermore, $m(E) - m(K) = m(O) - m(F) < \varepsilon$.

For arbitrarily E we have $m(E) = \sup_R m(E \cap B_R)$, and applying the result for each bounded set $E \cap B_r$ we get that $m(E) = \sup\{m(K) : K \subset E, K \text{ is compact}\}$. $\square$

3.2. **Lebesgue measure under linear transformations.** We note that the one-dimensional Lebesgue integral satisfies the following properties

$$\int_{\mathbb{R}} f(x+a)\, dm(x) = \int_{\mathbb{R}} f(x)\, dx, \quad \int_{\mathbb{R}} f(cx) dm(x) = |c|^{-1} \int_{\mathbb{R}} f(x)\, dm(x),$$

where f is integrable, $a \in \mathbb{R}$ and $c \in \mathbb{R}$, $c \neq 0$. The second property holds for characteristic functions of half-open intervals and then for simple functions and by the monotone convergence theorem for all $f \geq 0$ and finally, for all integrable f. We want to extend this properties to the Lebesgue integral in higher dimensions.

Theorem 3.2. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible linear transformation. If f is a non-negative measurable function, then $f \circ T$ is also measurable and*

$$\int_{\mathbb{R}^n} f \circ T(x)\, dm_n(x) = |\det T|^{-1} \int_{\mathbb{R}^n} f(x) dm_n(x).$$

Proof. First of all if f is Borel measurable, then so is $f \circ T$ since $(f \circ T)^{-1}(E) = T^{-1}(f^{-1}(E))$ and an invertible linear transformation transforms Borel sets into Borel sets. Any invertible linear transformation is a composition of transformations of the type: (1) $(x_1, x_2, ..., x_n) \mapsto (cx_1, x_2, ..., x_n)$, (2) $(x_1, x_2, ..., x_n) \mapsto (x_2, x_1, ..., x_n)$ (3)

$(x_1, x_2, ..., x_n) \mapsto (x_1 + cx_2, x - 2, ..., x_n)$, which correspond to elementary operations on matrices. It is enough to show the the formula above holds for every elementary transformation of these types. For type (1) it follows immediately from the one-dimensional formula and the Tonelli theorem, for (2) it is obvious after the application of the Tonelli theorem. For the third type, using the Tonelli theorem we reduce the identity to

$$\int f(x_1 + cx_2)\, dm(x_1) = \int f(x_1)\, dm(x_1).$$

Therefore, we obtain the change of variables formula for the integrals of Borel measurable functions. In particular for any Borel set E we have $m(T(E)) = |\det T| m(E)$. Then a set of zero measure is transformed by T^{-1} into a set of zero measure. Now we know that if A is Lebesgue measurable, then $T^{-1}(A)$ is also Lebesgue measurable. Thus we can extend the formula to Lebesgue measurable functions. □

Corollary 3.1. *The Lebesgue measure on $\mathbb{R}^n$ is invariant under rotations and reflections.*

More general change of variables formula is also true for Lebesgue integrals. Let $G : \Omega \to \mathbb{R}^n$ be a one-to-one map between Ω and $G(\Omega)$ such that G has continues derivatives and the matrix of first partial derivatives DG is invervible at each point. Then for any $f \in L^1(G(\Omega), m_n)$ we have

$$\int_{G(\Omega)} f\, dm_n = \int_{\Omega} f \circ G |\det DG|\, dm_n.$$

4. Differentiation on Euclidean space

4.1. **Vitali covering lemma.** We first prove a finite version of the lemma.

Lemma 4.1. *Suppose that $\{B_j\}_{j \in J}$ is a finite collection of balls. There exists a sub-collection $\{B_k\}_{k \in K}$, $K \subset J$ of disjoint balls such that $\bigcup_k 3B_k \supset \bigcup_j B_j$, in particular $\sum_k m(B_k) \geq 3^{-n} m(\cup_j B_j)$.*

Proof. We start by choosing the largest ball B_1' in the (finite) family. Then all balls intersecting B_1' are contained in $3B_1'$. Suppose that we already chose balls $B_1', B_2', ..., B_N'$. Consider the balls which does not intersect the chosen ones and if there are any, then take the next ball B_{N+1}' to be the largest of the remaining ones.

The balls we choosing are disjoint and all balls that are not chosen are contained in $3B$ for some of the balls B is our sub-collection. □

Now we formulate a version for infinite family of balls, many variations of this one can be found in various textbooks.

Lemma 4.2 (Vitali). *Let $U = \cup_{\alpha \in A} B_\alpha$, where B_α are open balls. If $c < m(U)$ then there exits a finite sub-collection of disjoint balls $\{B_k\}$ such that $\sum m(B_k) \geq 3^{-n} c$.*

Proof. Since $c < m(U)$ there exists a compact set $K \subset U$ such that $c < m(K)$. Using the compactness, we take a finite sub-cover of the cover $K \subset \bigcup B_\alpha$ and then use the previous lemma on this sub-cover. We obtain a disjoint finite collection of balls such that $\sum_k m(B_k) \geq 3^{-n} m(K) > 3^{-n} c$. □

4.2. **Maximal function.** We say that a measurable function f on $\mathbb{R}^n$ is in L^1_{loc} if f is integrable (with respect to the Lebesgue measure) over any bounded ball.

Definition 4.1. *Let $f \in L^1_{loc}$, the Hardy-Littlwood* **maximal function** *of f is defined by*

$$Mf(x) = \sup_{r>o} \frac{1}{m(B(r,x))} \int_{B(x,r)} |f(x)| \, dm_n.$$

We will show below that $Mf(x) \geq f(x)$ a.e. A very important property of the maximal function is given by the next theorem.

Theorem 4.1. *There exists a constant C such that for any $f \in L^1(\mathbb{R}^n)$ and any $a > 0$*

$$m(\{Mf > a\}) \leq Ca^{-1} \int |f| \, dm.$$

Proof. Let $E = \{Mf > a\}$. Then for any $x \in E$ there is a ball B_x centered at x such that $\int_{B_x} |f| \, dm \geq am(B_x)$. We take $c < m(E)$ and apply the Vitaly covering lemma, there is a disjoint sub-collection $\{B_k\}$ such that $\sum_k m(B_k) \geq 3^{-n} c$. On the other hand,

$$\sum_k m(B_k) \leq a^{-1} \sum_k \int_{B_k} |f| \, dm.$$

Since $\{B_k\}$ are disjoint, we conclude

$$c \leq 3^n a^{-1} \int |f| dm.$$

Then $m(E) \leq 3^n a^{-1} \|f\|_1$. □

4.3. **First differentiation theorem.** Assume that $f \in L^1_{loc}(\mathbb{R}^n)$ and we can compute $\nu(E) = \int_E f \, dm$ for bounded measurable sets E. Our question is how to reconstruct the function f from the measure ν, or we may assume that $\nu \ll m$ and ask how to fund the Radon-Nikodym derivative.

Theorem 4.2. *If $f \in L^1_{loc}$ then*

$$\lim_{r \to 0} \frac{1}{m(B_r(x))} \int_{B_r(x)} f(y) \, dm(y) = f(x)$$

for almost every $x \in \mathbb{R}^n$*.*

We can rewrite the limit using the measure ν as $\lim_{r\to 0} \nu(B_r(x))(m(B_r(x)))^{-1}$.

Proof. The theorem holds for continuous functions. To show that it is true for a function f it is enough to show that is holds for $f\chi_{\{|x|<N\}}$ for any N. Thus we may assume that $f \in L^1(\mathbb{R}^n)$. The function f can be approximated by simple functions and a simple function can be approximated by a continuous one, so there exists a continuous function g such that $\int |f-g| < \varepsilon$. Let us denote

$$A_r h(x) = \frac{1}{m(B_r(x))} \int_{B_r(x)} h(y)\, dm(y).$$

Then $A_r f = A_r g + A_r(f-g)$, using the Hardy-Littlewood maximal function we then see that

$$\limsup_{r\to 0} |A_r f(x) - f(x)| = \limsup_{r\to 0} (A_r|f-g|(x)) + \lim_{r\to 0} A_r g(x) - g(x)) + |g(x) - f(x)| \leq M(f-g)(x) + |g-f|(x).$$

Therefore

$$m(\{\limsup_{r\to 0} |A_r f(x) - f(x)| > \delta\}) \leq m(\{M|f-g| > \delta/2\}) + m(\{|f-g| > \delta/2\}).$$

We estimate the first term applying the previous theorem and the second term by integrating the inequality $|f-g| > \delta/2\chi_E$, where $E = \{|f-g| > \delta/2\}$. We get

$$m(\{|A_r f(x) - f(x)| > \delta\}) \leq (2C+1)\varepsilon\delta^{-1}.$$

Since the inequality holds for any $\varepsilon > 0$ we conclude that

$$m(\{\limsup_{r\to 0} |A_r f(x) - f(x)| > \delta\}) = 0,$$

now we take the union of such sets with $\delta_n = n^{-1}$ and obtain that

$$m(\{\limsup_{r\to 0} |A_r(f)(x) - f(x)| > 0\}) = 0.$$

□

In the prove above (and at least once before), we applied the so called Chebyshev inequality (or Markov inequality), if $f \in L^1(\mu)$ then

$$\mu(\{|f| > a\}) \leq a^{-1}\|f\|_1.$$

4.4. **Lebesgue differentiation theorem.** We consider the case $n = 1$. Let $f \in L^1_{loc}(\mathbb{R})$, we define

$$F(x) = \begin{cases} \int_{[0,x]} f(y)\, dm(y), & x \geq 0 \\ -\int_{[x,0]} f(y)\, dm(y), & x < 0. \end{cases}$$

Then F is a continuous function and we just proved that

$$\lim_{t \to 0} \frac{F(x+t) - F(x-t)}{2t} = f(x)$$

for almost every $x \in \mathbb{R}$. One can prove in a similar way that F is differentiable almost everywhere and $F' = f$ almost everywhere, it follows also form our next theorem.

Returning to the general dimension, we will prove a bit stronger result.

Definition 4.2. *Let $f \in L^1_{loc}(\mathbb{R}^n)$, we say that x is a Lebesgue point for f if*

$$\lim_{r \to 0} \frac{1}{m(B_r(x))} \int_{B_r(x)} |f(y) - f(x)|\, dm(y) = 0.$$

We denote by L_f the set of all Lebesgue points of the function f.

Theorem 4.3. *Let $f \in L^1_{loc}(\mathbb{R}^n)$ then almost every $x \in \mathbb{R}^n$ is a Lebesgue point for f, i.e., $m(\mathbb{R}^n \setminus L_f) = 0$.*

Proof. The proof is similar to the one given above, we denote by

$$A^*_r f(x) = \frac{1}{m(B_r(x)} \int_{B_r(x))} |f(y) - f(x)|\, dm(y).$$

Then for a continuous function g we have $\limsup_{r \to 0} A^*_r g(x) = 0$. Furthermore, $A^*_r f(x) \leq A^*_r(f-g)(x) + A^*_r(g)(x)$, then $\limsup_{r\to 0} A^*_r f(x) \leq \limsup_{r \to 0} A^*_r(f-g)(x)$. We finish the prove as above, noticing that

$$A^*_r(f-g)(x) \leq M(f-g)(x) + |(f-g)(x)|.$$

□

Corollary 4.1 (Density points)**.** *Let $E \subset \mathbb{R}^n$ be a Lebesgue measurable set. Define*

$$d^+_E(x) = \limsup_{r \to 0} \frac{m(E \cap B_r(x))}{m(B_r(x))}, \quad d^-_E(x) = \liminf_{r \to 0} \frac{m(E \cap B_r(x))}{m(B_r(x))}.$$

Then for almost every $x \in E$ we have $d^-_E(x) = 1$ and for almost every $x \in E^c$, $d^+_E(x) = 0$.

We apply the Lebesgue differentiation theorem to the function $f = \chi_E$.

Theorem 4.4. *Let ν be a regular Borel measure on $\mathbb{R}^n$ and m be the Lebesgue measure. Then for m-a.e. $x \in \mathbb{R}^n$ there is a limit*

$$\lim_{r\to 0} \frac{\nu(B_r(x))}{m(B_r(x))} = f(x),$$

where f is the Radon-Nykodim derivative of the absolutely continuous part of ν with respect to the Lebesgue measure.

Proof. Let $\nu = \nu_a + \nu_s$ be the Lebesgue decomposition of ν. We already know from Theorem 4.2 that the limit $\lim_{r\to 0} \nu_a(B_r(x))(m(B_r(x)))^{-1}$ exists almost everywhere and is equal to $f(x)$, where $\nu_s(E) = \int_E f(x)dx$. We need to show that for the singular part

$$\lim_{r\to 0} \frac{\nu_s(B_r(x))}{m(B_r(x))} = 0, \tag{1}$$

almost everywhere.

Let $X = A \cup B$ be a decomposition of X such that $A \cap B = \emptyset$, $m(B) = 0$ and $|\nu_S|(A) = 0$. We fix k and consider the set

$$F_k = \left\{ x \in A : \limsup_{r\to 0} \frac{|\nu_s|(B_r(x))}{m(B_r(x))} > \frac{1}{k} \right\}.$$

It is enough to show that $m(F_k) = 0$ for each k, then $f = \cup_k F_k$ has measure zero and on the complement of F we know that (1) holds. We take $\varepsilon > 0$ and consider an open set $U \supset A$ such that $|\nu_s|(U) < \varepsilon$. Now for each $x \in F_k$ there exists a ball $B_r(x)$ such that $|\nu_s|(B_r(x)) > k^{-1}m(B_r(x))$. We use the Vitaly covering lemma to choose disjoint balls $\{B_j\}$ form this family such that $\sum_j m(B_j) \geq 3^{-n}m(F_k)$. We have also $\sum_j m(B_j) < k\sum_j |\nu_s|(B_j) \leq k|\nu_s|(U) < k\varepsilon$. Thus $m(F_k) \leq 3^n k\varepsilon$ for any $\varepsilon > 0$ and therefore $m(F_k) = 0$. $\square$

MAT205a, Fall 2019 Part III: Differentiation

Lecture 10, Following Folland, ch 3.5

5. Functions of bounded variation

5.1. **Differentiation of regular measures on the real line.** We now study differentiation on the real line, almost everywhere will be always in this section with respect to the Lebesgue measure. Suppose that μ_F is a Lebesgue-Stieltjes measure on $\mathbb{R}$. It follows from the previous section that $\lim_{r\to 0}\mu_F(x+r,x-r)/2r$ exists for almost every $x\in\mathbb{R}$, denote it by $f(x)$, then $\mu_F = m_s + m_a$, where m_s is singular with respect to the Lebesgue measure, $m_a \ll m$ and the Radon-Nikodym derivative of m_a with respect to m is f. By the last theorem in the previous lecture, we know that

$$\lim_{r\to 0}\frac{\mu_s(x-r,x+r)}{2r} = 0$$

then we also have

$$\lim_{r\to 0}\frac{\mu_s(x,x+r)}{r} = \lim_{r\to 0}\frac{\mu_s(x-r,x)}{r} = 0.$$

For the absolutely continuous part, the Lebesgue differentiation theorem gives

$$\lim_{r\to 0}\frac{1}{2r}\int_{x-r}^{x+r}|f(t)-f(x)|\,dt = 0$$

almost everywhere and for x for which the limit above is zero, we have

$$\lim_{r\to 0}\frac{1}{r}\int_{x}^{x+r} f(t)dt = \lim_{r\to 0}\frac{1}{r}\int_{x-r}^{x} f(t)\,dt = f(x).$$

By adding those identities up, we obtain that almost everywhere

$$\lim_{r\to 0}\frac{\mu_F((x,x+r])}{r} = \lim_{r\to 0}\frac{\mu_F((x-r,x])}{r} = f(x).$$

5.2. **Monotone functions.** First we show that any monotone function is almost everywhere differentiable. Remind that a function f is called differentiable at a point $x\in\mathbb{R}$ if there is a limit $\lim_{t\to 0}\frac{f(x+t)-f(x)}{t}$.

Theorem 5.1. *Suppose that $F:\mathbb{R}\to\mathbb{R}$ is a monotone function. Then the set of points where F is discontinuous is countable. Moreover, F is differentiable almost everywhere.*

Proof. Since F is monotone there are limits $\lim_{y\to x-}F(y) = F(x-)$ and $\lim_{y\to x+}F(y) = F(x+)$. Suppose that F is discontinuous at some point x, then $F(x+)\neq F(x-)$ and there is a rational number $q(x)$ between $F(x-)$ and $F(x+)$. Those numbers for all such x are distinct. If F is for example increasing and $x<y$ then there exists t such that $x<t<y$ and $q(x)<F(x+)\le F(t)\le F(y-)<q(y)$. Now since the set of

rational numbers is countable, we conclude that the number of points of discontinuity is at most countable.

To prove the second statement, we assume that F is increasing (otherwise take the function $-F$) and define a new function $G(x) = F(x+)$, then G is also increasing and right continuous, moreover $F = G$ everywhere except at most countable set of points, where F is discontinuous. We consider the Lebesgue-Stieltjes measure μ_G. We know that $G(x+h) - G(x) = \mu_G((x, x+h])$ and by our observation above, G is almost everywhere differentiable. let now $H = G - F$, we want to show that $= H' = 0$ a.e. Since F was increasing, $G(x) = F(x+) \geq F(x)$ and $H \geq 0$ and differs from zero only at countably many points. We can write $H = \sum c_j \chi_{x_j}$, where $c_j > 0$ and each sum $\sum_{-N<|x_j|<N} c_j$ is finite. We consider the set $E = \{x_j\}_j$, it is a countable set, let $\nu = \sum c_j \delta_{x_j}$ this is a Radon measure which is singular with respect to the Lebesgue measure. Applying the Theorem 4.4 once again, we see that for almost all x

$$\frac{H(x+t) - H(x-t)}{2t} = \lim_{t\to 0} \frac{\nu((x-t, x+t])}{2t} = 0,$$

thus $H'(x) = 0$ almost everywhere. It implies that F is differentiable almost everywhere and $F' = G'$ a.e. □

5.3. **Total variation.** We define the total variation of a function over an interval, for a smooth function F the quantity we want to capture is the integral of $|F'|$ over this interval. General definition is the following.

Definition 5.1. *Let $F : \mathbb{R} \to \mathbb{R}$, for $[a, b] \subset \mathbb{R}$ the total variation of F over $[a, b]$ is defined as*

$$T_F([a, b]) = \sup \left\{ \sum_1^n |F(x_j) - F(x_{j-1})| : a = x_0 < ... < x_n = b \right\}.$$

We say that F is a function of bounded variation on $[a, b]$ and write $F \in BV([a, b])$ if $T_F([a, b]) < \infty$.

We also define

$$T_F(x) = \sup \left\{ \sum_1^n |F(x_j) - F(x_{j-1})| : -\infty < x_0 < ... < x_n = x \right\},$$

and call T_F the total variation function. Clearly it is increasing, we say that F has a bounded variation on $\mathbb{R}$ if $\sup_x T_F(x) = \lim_{x\to\infty} T_F(x) < \infty$, and we write $F \in BV$.

Example 5.1.
1. If F is a monotone bounded function, then $F \in BV$, $T_F(x) = |F(x) - \lim_{y\to-\infty} F(y)|$.
2. If $F \in C^1$ and $\int_{-\infty}^{\infty} |F'(t)dt < \infty$ then $F \in BV$.
3. $F(x) = \sin(X)$ has bounded variation over any bounded interval, but not over $\mathbb{R}$.

4. $F(x) = \sin(x^{-1}$ has infinite variation over $[0, 1]$.
5. If F_1 and F_2 have bounded variation, then so does $aF_1 + bF_2$.
6. If $F \in BV$ then F is bounded.

Lemma 5.1. *If $F \in BV$ then the functions $T_F + F$ and $T_F - F$ are increasing.*

Proof. Let $\varepsilon > 0$ and $x < y$, we choose points $x_0 < ... < x_n = x$ such that $T_F(x) < \sum_1^n |F(x_j) - F(x_{j-1})| + \varepsilon$. Then $T_F(y) \geq \sum_1^n |F(x_j) - F(x_{j-1})| + |F(x) - F(y)| \geq T_F(x) - \varepsilon + F(x) - F(y)$. Taking the limit as $\varepsilon \to 0$ we conclude that $T_F(y) + F(y) \geq T_F(x) + F(x)$. A similar argument shows that $T_F - F$ is increasing. □

Now we can characterize all functions of bounded variation.

Theorem 5.2. *A function F has bounded variation if and only if it can be written as the difference of two bounded increasing functions.*

Proof. We already know that a bounded monotone function has bounded variation and that the difference of two functions with bounded variation has bounded variation.

It remains to show that if $F \in BV$ then $F = F_1 - F_2$, where F_1 and F_2 are bounded and increasing. We define $F_1 = (T_F + F)/2$ and $F_2 = (T_F - F)/2$. Then F_1 and F_2 are increasing by the Lemma above and they are bounded since T_F and F are bounded. □

Corollary 5.1. *If $F \in BV$ then*
(a) the limits $F(x\pm) = \lim_{y \to x\pm} F(y)$ exist for all $x \in \mathbb{R}$,
(b) the set of points where F is discontinuous is countable,
(c) F is differentiable almost everywhere.

Proof. We know that all these properties hold for monotone functions, then they also hold for the difference of two increasing functions. □

5.4. **Functions of Bounded variation and Lebesgue-Stieltjes measures.** We know that each increasing right continuous function F corresponds to a Radon measure on $\mathbb{R}$. If F is bounded then the measure is finite. We want to describe a class of functions of bounded variation that correspond to signed Radon measures.

Lemma 5.2. *If $F \in BV$ and F is right continuous then T_F is also right continuous.*

Proof. As we know T_F is an increasing function and for any $x \in \mathbb{R}$ there is a limit $T_F(x+) = \lim_{y \to x+} T_F(y)$. We fix $x \in \mathbb{R}$ and assume that T_F is not right continuous at x, then $a = T_F(x+) - T_F(x) > 0$. Let $\varepsilon > 0$ such that $\varepsilon < a/4$. Since F is right continuous at x, there exists $\delta > 0$ such that $|F(x + t) - F(x)| < \varepsilon$ when $0 < t < \delta$, we may also assume that $|T_F(x+t) - T_F(x+)| < \varepsilon$ when $0 < t < \delta$. We choose $t < \delta$ and a partition $x = x_0 < ... < x_n = x + t$ such that

$$\sum_1^n |F(x_j) - F(x_{j-1})| > a - \varepsilon.$$

Since $x_1 \leq x_n = x+t$, we know that $|F(x_1) - F(x_0)| < \varepsilon$ and $\sum_2^n |F(x_j) - F(x_{j-1})| > a - 2\varepsilon$. Now we can find a partition of (x, x_1) such that $x = y_0 < ... < y_k = x_1$ and $\sum_1^k |F(y_i) - F(y_{i-1})| > a - \varepsilon$. Then we get

$$a + \varepsilon > T_F(x+) - T_F(x) + T_F(x+t) - T_F(x+) =$$
$$T_F(x+t) - T_F(x) \geq \sum_1^k |F(y_i) - F(y_{i-1}| + \sum_2^n |F(x_j) - F(x_{j-1}| \geq 2a - 3\varepsilon.$$

This contradicts to our choice of ε. Thus the assumption that T_F is not right continuous at x was wrong. □

The lemma implies that if $F \in BV$ is right continuous then $F = (T_F + F)/2 - (T_F - F)/2$ and there are two Lebesgue-Stieltjes measures μ_1 and μ_2 such that $F(b) - F(a) = \mu_1((a, b]) - \mu_2((a, b])$. We denote the signed measure $\mu_1 - \mu_2$ by μ_F.

Theorem 5.3. *Suppose that $F \in BV$, F is right continuous then $F' \in L^1$. If $\mu_F \perp m$ then $F' = 0$ a.e. and if $\mu_F \ll m$ then $F(y) - F(x) = \int_x^y F'(t)dm(t)$.*

Proof. We know that F' exits almost everywhere and at point x where the derivative exists it is equal to

$$F'(x) = \lim t \to 0 \frac{F(x+t) - F(x-t)}{2t} = \lim_{t \to 0} \frac{\mu_F((x-t, x+t])}{m((x-t, x+t])},$$

by the Lebesgue differentiation theorem the limit is the Radon-Nikodym derivative of μ_F with respect to m. □

5.5. **Absolutely continuous functions.** We want to reformulate the condition $\mu - F \ll m$ in terms of the function F.

Definition 5.2. *We say that a function $F : [a, b] \to \mathbb{R}$ is absolutely continuous if for every $\varepsilon > 0$ there is $\delta > 0$ such that if (x_j, y_j) are disjoint intervals in $[a, b]$ with $\sum_j (y_j - x_j) < \delta$ then $\sum |F(y_j) - F(x_j)| < \varepsilon$.*

Example 5.2. If $f \in C^1$ and F' is bounded then F is absolutely continuous, since $\sum |F(y_j) - F(x_j)| \leq \max |F'| \sum (y_j - x_j)$.

We remark also that an absolutely continuous function is uniformly continuous.

Lemma 5.3. *If F is absolutely continuous on a bounded interval $[a, b]$ then it has bounded variation over this interval.*

Proof. By the definition of the absolute continuity, there is $\delta > 0$ such that if the sum of the length of the intervals $\{(x_j, y_j)\}$ is less then δ, then $\sum_j |F(y_j) - F(x_j)| < 1$. Now let $a = t_0 < t_1 < ... < t_N < b$ be a partition, we may refine this partition such that it can be decompose into a family of sets of intervals with the total length less that δ on the number of families in the decomposition is bounded by C/δ. □

Theorem 5.4 (Fundamental Theorem of Calculus)**.** *Let $F : [a, b] \to \mathbb{R}$ be a function. The following statements are equivalent:*
(a) F is absolutely continuous on $[a, b]$,
(b) $F(x) - F(a) = \int_a^x f(t)\, dt$ for some function $f \in L^1(m)$,
(c) F' exists almost everywhere and $F(x) - F(a) = \int_a^x F'(t)\, dt$.

Proof. We should show that μ_F is absolutely continuous with respect to the Lebesgue measure, then the theorem follows from the previous results. □