Lebesgue Differentiation

Swayam Chube

December 1, 2024

Throughout this article, we work in a fixed Euclidean space $\mathbb{R}^k$ equipped with the standard Lebesgue measure $m = m_k$.

§1 The Lebesgue Differentiation Theorem

Definition 1.1. Let μ be a complex Borel measure on $\mathbb{R}^k$. For $x \in \mathbb{R}^k$ and $r > 0$, let

$$(Q_r\mu)(x) = \frac{\mu(B(x,r))}{m(B(x,r))}.$$

Define the *symmetric derivative* of μ at x as

$$(D\mu)(x) = \lim_{r \to 0^+} (Q_r\mu)(x)$$

whenever this limit exists. Further, if $\mu \geqslant 0$, define the *maximal function* $M\mu : \mathbb{R}^k \to [0, \infty]$ as

$$(M\mu)(x) = \sup_{r>0} (Q_r\mu)(x).$$

For an arbitrary complex Borel measure μ on $\mathbb{R}^k$, *define* $M\mu := M|\mu|$.

Proposition 1.2. The maximal function $M\mu$ is lower semicontinuous, in particular, is measurable.

Proof. We may assume that $\mu \geqslant 0$. Let $\lambda > 0$ and set $U = \{x \in \mathbb{R}^k \colon (M\mu)(x) > \lambda\}$. Let $x \in U$. Then, there is an $r > 0$ such that

$$t = \frac{\mu(B(x,r))}{m(B(x,r))} > r.$$

Choose $\delta > 0$ such that

$$r^k < (r+\delta)^k < \frac{r^k t}{\lambda}.$$

If $|y - x| < \delta$, then $B(y, r+\delta) \supseteq B(x,r)$ whence

$$\frac{\mu(B(y,r+\delta))}{m(B(y,r+\delta))} \geqslant \frac{\mu(B(x,r))}{m(B(x,r))} \frac{m(B(x,r))}{m(B(y,r+\delta))} = t\frac{r^k}{(r+\delta)^k} > \lambda,$$

according to our choice of δ. Thus, $B(x,\delta) \subseteq U$ and the latter is open. ■

LEMMA 1.3 (VITALI). Let $W = \bigcup_{i=1}^{N} B(x_i, r_i)$ where $x_i \in \mathbb{R}^k$ and $r_i > 0$ for $1 \leqslant i \leqslant N$. Then, there is a subset $S \subseteq \{1, \ldots, N\}$ such that

(a) The balls $B(x_i, r_i)$ are pairwise disjoint for $i \in S$,

(b) $W \subseteq \bigcup_{i \in S} B(x_i, 3r_i)$, and hence

(c) $m(W) \leqslant 3^k \sum_{i \in S} m(B(x_i, r_i))$.

Proof. Without loss of generality, suppose $r_1 \geqslant \cdots \geqslant r_N$. Begin by setting $i_1 = 1$. Remove all balls $B(x_i, r_i)$ that intersect $B(x_{i_1}, r_{i_1})$. Note that if $B(x_i, r_i) \cap B(x_{i_1}, r_{i_1}) \neq \varnothing$, then choosing some y in the intersection, we have that for any $z \in B(x_i, r_i)$,

$$|z - x_{i_1}| \leqslant |z - y| + |y - x_{i_1}| < 2r_i + r_{i_1} \leqslant 3r_{i_1}$$

since $i > i_1$. That is, $B(x_i, r_i) \subseteq B(x_{i_1}, 3r_{i_1})$. Next, choose i_2 to be the smallest index larger than i_1 that hasn't been deleted and repeat this procedure. It is easy to see that the balls that remain satisfy the required conditions. ■

Henceforth, we use the shorthand $\{f > \lambda\}$ to denote the set $\{x \in \mathbb{R}^k \colon f(x) > \lambda\}$.

THEOREM 1.4. Let μ be a complex Borel measure on $\mathbb{R}^k$. For $\lambda > 0$,

$$m\{M\mu > \lambda\} \leqslant 3^k \lambda^{-1} \|\mu\|.$$

Proof. Let $U = \{M\mu > \lambda\}$ and let $K \subseteq U$ be a compact set. For each $x \in K$, there is an $r_x > 0$ such that $(Q_{r_x}\mu)(x) > \lambda$. Since K is compact, we can choose a finite subcover

$$K \subseteq \bigcup_{i=1}^{N} B(x_i, r_i),$$

where r_i is shorthand for r_{x_i}. Using Lemma 1.3, there is a subcollection $S \subseteq \{1, \ldots, N\}$ such that $K \subseteq \bigcup_{i \in S} B(x_i, 3r_i)$ and the balls $B(x_i, r_i)$ are pairwise disjoint. Thus,

$$m(K) \leqslant 3^k \sum_{i \in S} m(B(x_i, r_i)) < 3^k \lambda^{-1} \sum_{i \in S} \mu(B(x_i, r_i)) = 3^k \lambda^{-1} \mu \left(\bigcup_{i \in S} B(x_i, r_i) \right) \leqslant 3^k \lambda^{-1} \|\mu\|,$$

thereby completing the proof. ■

DEFINITION 1.5. Let $f \in L^1(\mathbb{R}^k)$ and let μ be the complex Borel measure on $\mathbb{R}^k$ given by $d\mu = f\, dm$. Define the *maximal function* of f as $(Mf)(x) = (M\mu)(x)$. Then,

$$(Mf)(x) = \sup_{r>0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y)|\, dm(y).$$

Definition 1.6. Let $f \in L^1(\mathbb{R}^k)$. A point $x \in \mathbb{R}^k$ is said to be a *Lebesgue point* of f if

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

Henceforth, we also use the shorthand $m(B_r)$ to denote $m(B(x,r))$ for any $x \in \mathbb{R}^k$, since the Lebesgue measure is translation invariant.

Theorem 1.7 (Lebesgue). If $f \in L^1(\mathbb{R}^k)$, then almost every $x \in \mathbb{R}^k$ is a Lebesgue point of f.

Proof. Define

$$(T_r f)(x) = \frac{1}{m(B_r)} \int_{B(x,r)} |f(y) - f(x)| \, dm(y),$$

and

$$(Tf)(x) = \limsup_{r \to 0^+} (T_r f)(x).$$

It suffices to show that $Tf = 0$ a.e. on $\mathbb{R}^k$.

Fix a positive integer n and choose $g \in C_c(\mathbb{R}^k)$ with $\|f - g\|_1 < \frac{1}{n}$. Set $h = f - g$. We have

$$\begin{aligned} (T_r h)(x) &= \frac{1}{m(B_r)} \int_{B(x,r)} |h(y) - h(x)| \, dm(y) \\ &\leqslant \frac{1}{m(B_r)} \int_{B(x,r)} |h(y)| \, dm(y) + h(x). \end{aligned}$$

Taking lim sup with $r \to 0^+$, we have

$$(Th)(x) \leqslant (Mh)(x) + |h(x)|.$$

Since $f = g + h$, we have

$$T_r f \leqslant T_r g + T_r h \implies Tf \leqslant Tg + Th.$$

Recall that g is continuous, and hence, $Tg = 0$. This gives us

$$Tf \leqslant Mh + |h|.$$

Let $y > 0$ be arbitrary. We have the obvious inclusion

$$\{Tf > 2y\} \subseteq \{Mh > y\} \cup \{|h| > y\} =: E(y, n).$$

Using Theorem 1.4, we have

$$m\{Tf > 2y\} \leqslant \frac{3^n}{y}|h| + \frac{1}{y}|h| \leqslant \frac{3^n + 1}{yn}.$$

Since the inequality on the right holds for all positive integers n, we have that $m\{Tf > 2y\} = 0$ for all $y > 0$. It follows that $m\{Tf > 0\} = 0$, thereby completing the proof. ■

Remark 1.8. If $x \in \mathbb{R}^k$ is a Lebesgue point of f, then it is easy to see that

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

Theorem 1.9. Suppose μ is a complex Borel measure on $\mathbb{R}^k$, and $\mu \ll m$. Let f be the Radon-Nikodym derivative of μ with respect to m. Then, $D\mu = f$ a.e. on $\mathbb{R}^k$, and hence,

$$\mu(E) = \int_E (D\mu)\, dm,$$

for all Borel sets $E \subseteq \mathbb{R}^k$.

Proof. At any Lebesgue point $x \in \mathbb{R}^k$ of f,

$$f(x) = \lim_{r\to 0^+} \frac{1}{m(B_r)} \int_{B(x,r)} f(y)\, dm(y) = \lim_{r\to 0^+} \frac{\mu(B(x,r))}{m(B(x,r))} = (D\mu)(x).$$

This completes the proof. ■

Definition 1.10. Let $x \in \mathbb{R}^k$. A sequence $(E_i)_{i\geqslant 1}$ of Borel sets in $\mathbb{R}^k$ is said to *shrink to x nicely* if there is a number $\alpha > 0$ and a sequence of balls $B(x, r_i)$ with $\lim_{i\to\infty} r_i = 0$ such that $E_i \subseteq B(x, r_i)$ and $m(E_i) \geqslant \alpha m(B(x, r_i))$.

Theorem 1.11. Associate to each $x \in \mathbb{R}^k$ a sequence $(E_i(x))_{i\geqslant 1}$ that shrinks to x nicely, and let $f \in L^1(\mathbb{R}^k)$. Then

$$f(x) = \lim_{i\to\infty} \frac{1}{m(E_i(x))} \int_{E_i(x)} f(y)\, dm(y)$$

at every Lebesgue point x of f, and hence, a.e. on $\mathbb{R}^k$.

Proof. If $x \in \mathbb{R}^k$ is a Lebesgue point of f, and $\alpha(x) > 0$ be such that $m(E_i(x)) \geqslant \alpha(x) m(B(x, r_i))$. Then

$$0 \leqslant \frac{1}{m(E_i(x))} \int_{E_i(x)} |f(y) - f(x)|\, dm(x) \leqslant \frac{1}{\alpha(x) B(x, r_i)} \int_{B(x,r_i)} |f(y) - f(x)|\, dm(y).$$

As $i \to \infty$, the right hand side goes to 0 and hence, so does the left. This completes the proof. ■

Theorem 1.12. Let $f \in L^1(\mathbb{R})$ and define $F : \mathbb{R} \to \mathbb{R}$ by

$$F(x) = \int_{-\infty}^{x} f(y)\, dm(y).$$

Then $F'(x) = f(x)$ at every Lebesgue point of f, and hence, a.e. on $\mathbb{R}$.

Proof. Let $x \in \mathbb{R}$ be a Lebesgue point. If $(\delta_i)_{i\geqslant 1}$ is a sequence of positive reals converging to 0, then set $E_i = (x, x + \delta_i)$. Due to the preceding result, we have

$$f(x) = \lim_{i\to\infty} \frac{1}{m(E_i)} \int_{E_i} f(y)\, dm(y) = \lim_{i\to\infty} \frac{F(x + \delta_i) - F(x)}{\delta_i}.$$

This completes the proof. ■

§2 THE FUNDAMENTAL THEOREM OF CALCULUS

DEFINITION 2.1. A function $f : I = [a,b] \to \mathbb{C}$ is said to be *absolutely continuous* on I, if for every $\varepsilon > 0$, there is a $\delta > 0$ such that

$$\sum_{i=1}^{n} |f(\beta_i) - f(\alpha_i)| < \varepsilon$$

for any disjoint collection of segments $(\alpha_1, \beta_1), \ldots, (\alpha_n, \beta_n)$ in I provided

$$\sum_{i=1}^{n} \beta_i - \alpha_i < \delta.$$

THEOREM 2.2. Let $f : I = [a,b] \to \mathbb{R}$ be a continuous, increasing function. The following are equivalent:

(a) f is AC on I.

(b) f maps sets of measure 0 to sets of measure 0.

(c) f is differentiable a.e. on I, $f' \in L^1$, and

$$f(x) - f(a) = \int_a^x f'(x)\, dm(x).$$

for all $a \leqslant x \leqslant b$.

Proof. Let $\mathfrak{M}$ denote the σ-algebra of Lebesgue measurable sets in $\mathbb{R}$.
$\underline{(a) \implies (b)}$ Let $E \subseteq I$ be such that $m(E) = 0$. We must show that $f(E) \in \mathfrak{M}$ and $m(f(E)) = 0$. We may suppose, without loss of generality, that $a, b \notin E$.

Let $\varepsilon > 0$. Since f is AC on I, there is a δ corresponding to this ε as in the definition of absolute continuity. There is an open set V such that $E \subseteq V \subseteq I$ and $m(V) < \delta$. We can write $V = \bigcup_i (\alpha_i, \beta_i)$ with $\sum_i (\beta_i - \alpha_i) < \delta$. For any finite collection J of the indexing set over which i runs,

$$\sum_{j \in J} |f(\beta_j) - f(\alpha_j)| < \varepsilon \implies \sum_i |f(\beta_i) - f(\alpha_i)| \leqslant \varepsilon.$$

Hence, $m(E) \leqslant m(f(V)) \leqslant \varepsilon$. Since this inequality holds for all $\varepsilon > 0$, $m(E) = 0$.
$\underline{(b) \implies (c)}$ Let $g : I \to \mathbb{R}$ be given by $g(x) = f(x) + x$. This is a strictly increasing function of x. We claim that g maps measure 0 sets to measure 0 sets. Suppose $E \subseteq I$ has measure 0. We would like to show that $g(E)$ has measure 0. We may assume further that $a, b \notin E$. Let $\varepsilon > 0$. There is an open set V containing $f(E)$ such that $m(V) < \varepsilon$. Note that $f^{-1}(V)$ is an open subset of I containing E. There is an open set U containing E and contained in V such that $m(U) < \varepsilon$.

Being an open set, U is a disjoint union of (countably many) disjoint intervals. Since f is an increasing function, the image of disjoint intervals is either disjoint or they have at most one point in common. If f maps an interval of length η to an interval of length

η', then g maps the aforementioned interval to one of length $\eta + \eta'$. Now, the sum of the lengths of the images of the intervals that constitute U under f is at most ε, and hence, the measure of $g(U)$ is at most $\varepsilon + m(U) < 2\varepsilon$. Consequently, $m(g(E)) \leqslant 2\varepsilon$. Since this inequality holds for all $\varepsilon > 0$, we see that $m(g(E)) = 0$.

We come back to our original line of proof. Let $E \subseteq I$ be measurable. Then, we can write $E = E_1 \cup E_0$, where $m(E_0) = 0$ and E_1 is an F_σ-set. Thus, E_1 is a countable union of compact sets and because g is continuous, so is $g(E_1)$. Since g maps measure 0 sets to measure 0, $m(g(E_0)) = 0$ and finally, since $g(E) = g(E_0) \cup g(E_1)$, we conclude that $g(E) \in \mathfrak{M}$.

Define a measure $\mu(E) = m(g(E))$ on I. It is also easy to see that μ is a nonnegative complex Borel measure on $\mathbb{R}$ that is absolutely continuous with respect to m. Let $h : I \to \mathbb{R}$ denote the Radon-Nikodym derivative, where $h \in L^1(I)$. We shall show that our required derivative of f is $h - 1$.

If $E = [a, x]$, then $g(E) = [g(a), g(x)]$, since the image must be a compact interval. Thus,

$$g(x) - g(a) = m(g(E)) = \mu(E) = \int_a^x h(y)\, dm(y),$$

whence

$$h(x) - h(a) = \int_a^x h(y) - 1\, dm(y).$$

Due to Theorem 1.12, $f'(x) = h(x) - 1$ a.e. on I.

$\underline{(c) \implies (a)}$ Since $f' \in L^1$, for every $\varepsilon > 0$, there is a $\delta > 0$ such that $\left|\int_E f\, dm\right| < \varepsilon$ whenever $m(E) < \delta$. The conclusion is immediate now. ■

Definition 2.3. A function $f : I = [a, b] \to \mathbb{R}$ is said to be of *bounded variation* if the *total variation*, defined as

$$\sup \sum_{i=1}^{N} |f(t_i) - f(t_{i-1})|$$

where the supremum is taken over all partitions

$$a = t_0 < t_1 < \cdots < t_N = x,$$

is finite.

Theorem 2.4. Let $f : I = [a, b] \to \mathbb{R}$ be AC. For $a \leqslant x \leqslant b$, let $F(x)$ denote the total variation of f on $[a, x]$. Then the functions $F, F + f, F - f$ are AC and increasing on I.

Proof. The increasing assertion is immediate from the inequality

$$F(y) \geqslant F(x) + |f(y) - f(x)|$$

for all $a \leqslant x \leqslant y \leqslant b$.

As for the assertion about absolute continuity, it suffices to show that F is AC. Let $\varepsilon > 0$, then there is a corresponding δ according to the definition of absolute continuity.

Let $(\alpha_1, \beta_1), \ldots, (\alpha_n, \beta_n)$ be disjoint intervals with $\sum_{i=1}^{n} (\beta_i - \alpha_i) < \delta$. Then,

$$\sum_{i=1}^{n} F(\beta_i) - F(\alpha_i) = \sup \sum_{i,j} |f(t_j^i) - f(t_{j-1}^i)|,$$

where the supremum is taken over partitions of the intervals

$$\alpha_i = t_0^i < \cdots < t_{n_i}^i = \beta_i$$

for $1 \leqslant i \leqslant n$. But since

$$\sum_{i,j} t_j^i - t_{j-1}^i < \delta,$$

we have that $\sum_{i=1}^{n} F(\beta_i) - F(\alpha_i) \leqslant \varepsilon$. Thus, F is absolutely continuous on I. ■

THEOREM 2.5 (FUNDAMENTAL THEOREM OF CALCULUS). If f is a complex-valued function that is AC on $I = [a, b]$, then f is differentiable almost everywhere on I, $f' \in L^1$, and

$$f(x) - f(a) = \int_a^x f'(t)\, dm(t)$$

for all $a \leqslant x \leqslant b$.

Proof. It suffices to prove this for real-valued f. Let F denote its "total variation function". Define

$$g = \frac{F + f}{2} \quad \text{and} \quad h = \frac{F - f}{2}.$$

Due to the preceding result, both g and h are AC and increasing on I. Applying Theorem 2.2, and noting that $f = g - h$, we have the desired conclusion. ■