

# Hardy–Littlewood Maximal Functions and the Lebesgue Differentiation Theorem

Lecture Notes on Differentiation, Bounded Variation, and Absolute Continuity

## Abstract

These notes present a self-contained route from one-dimensional differentiation theory to the Hardy–Littlewood maximal theorem and the Lebesgue differentiation theorem in Euclidean space. The exposition begins with Vitali’s covering lemma on the line, proves almost-everywhere differentiability of monotone functions, develops functions of bounded variation and absolute continuity, and then turns to the maximal function, its weak and strong boundedness properties, and the differentiation of locally integrable functions. The guiding principle is that differentiation theorems are consequences of two complementary mechanisms: covering arguments that control exceptional sets and approximation arguments that reduce measurable functions to continuous or bounded ones.

## Contents

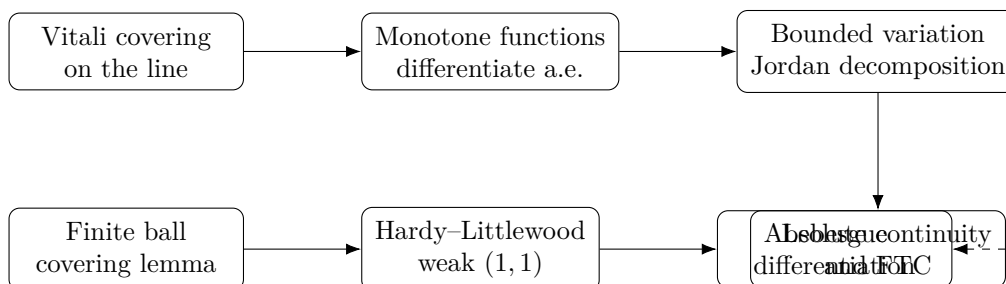
<b>Preface and organization</b>	<b>1</b>
<b>1 Introduction and guiding problems</b>	<b>2</b>
<b>2 A one-dimensional Vitali covering lemma</b>	<b>3</b>
<b>3 Dini derivatives and monotone functions</b>	<b>4</b>
<b>4 Functions of bounded variation</b>	<b>6</b>
<b>5 Differentiation of indefinite integrals on an interval</b>	<b>7</b>
<b>6 Absolute continuity</b>	<b>10</b>
<b>7 From intervals to balls</b>	<b>11</b>
<b>8 The Hardy–Littlewood maximal function</b>	<b>12</b>
<b>9 The Lebesgue differentiation theorem</b>	<b>15</b>
<b>Concluding perspective</b>	<b>16</b>
<b>Exercises</b>	<b>16</b>

## Preface and organization

These notes are written for a first graduate course in real analysis. The intended prerequisites are Lebesgue measure, measurable functions, the Lebesgue integral, basic properties of outer measure, and the standard approximation of integrable functions by continuous compactly

supported functions. All functions are real-valued unless explicitly stated otherwise. Lebesgue measure is denoted by  $m$ , and  $m^*$  denotes outer Lebesgue measure when measurability is not yet available.

The logical structure is as follows.



The dashed arrow records the conceptual connection: the one-dimensional fundamental theorem of calculus says that averages of an integrable function over intervals shrinking to a point recover the function almost everywhere; the Lebesgue differentiation theorem is the corresponding statement for balls in  $\mathbb{R}^n$ .

## 1 Introduction and guiding problems

A central theme of real analysis is the reciprocity between integration and differentiation. On an interval  $[a, b]$ , two natural questions arise.

- (1) If  $f \in L^1([a, b])$  and

$$F(x) = \int_a^x f(t) dt,$$

does one have  $F'(x) = f(x)$  for almost every  $x \in [a, b]$ ?

- (2) Conversely, if  $F : [a, b] \rightarrow \mathbb{R}$  is differentiable almost everywhere, under what additional hypotheses is  $F'$  integrable and

$$\int_a^b F'(x) dx = F(b) - F(a)?$$

The first question has an affirmative answer for Lebesgue integrable functions. The second question is more delicate. An increasing function is differentiable almost everywhere and satisfies

$$\int_a^b F'(x) dx \leq F(b) - F(a),$$

but equality may fail. The Cantor–Lebesgue function is continuous and increasing on  $[0, 1]$ , has  $F(0) = 0$ ,  $F(1) = 1$ , and yet satisfies  $F' = 0$  almost everywhere. The additional condition that restores equality is absolute continuity.

The higher-dimensional analogue asks whether, for  $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ , the averages of  $f$  over balls shrinking to  $x$  converge to  $f(x)$  for almost every  $x$ . This is the Lebesgue differentiation theorem. Its proof is governed by the Hardy–Littlewood maximal function and a geometric covering lemma.

Throughout, intervals are non-degenerate unless the contrary is explicitly stated. In  $\mathbb{R}^n$ ,

$$B(x, r) = \{y \in \mathbb{R}^n : |x - y| < r\}, \quad v_n = m(B(0, 1)),$$

so that  $m(B(x, r)) = v_n r^n$ .

## 2 A one-dimensional Vitali covering lemma

The first indispensable tool is Vitali's covering lemma. We use the following compact-interval form, which is sufficient for the differentiation arguments below.

**Definition 2.1** (Vitali cover). Let  $E \subset \mathbb{R}$ . A family  $\mathcal{I}$  of closed intervals is called a *Vitali cover* of  $E$  if, for every  $x \in E$  and every  $\varepsilon > 0$ , there exists  $I \in \mathcal{I}$  such that

$$x \in I, \quad 0 < |I| < \varepsilon,$$

where  $|I|$  denotes the length of  $I$ .

**Lemma 2.2** (Vitali covering lemma on the line). *Let  $E \subset \mathbb{R}$  with  $m^*(E) < \infty$ , and let  $\mathcal{I}$  be a Vitali cover of  $E$ . Then, for every  $\varepsilon > 0$ , there exist finitely many pairwise disjoint intervals  $I_1, \dots, I_N \in \mathcal{I}$  such that*

$$m^* \left( E \setminus \bigcup_{j=1}^N I_j \right) < \varepsilon.$$

*Proof.* Choose an open set  $O \supset E$  with  $m(O) < \infty$ . Since  $O$  is open and  $\mathcal{I}$  is a Vitali cover of  $E$ , the subfamily of intervals in  $\mathcal{I}$  that are contained in  $O$  is still a Vitali cover of  $E$ . We therefore restrict to intervals contained in  $O$ .

We select intervals greedily. Having chosen pairwise disjoint intervals  $I_1, \dots, I_n$ , put

$$k_n = \sup \left\{ |I| : I \in \mathcal{I}, I \subset O, I \cap \bigcup_{j=1}^n I_j = \emptyset \right\}.$$

If no admissible interval remains, the construction stops. Otherwise choose  $I_{n+1} \in \mathcal{I}$ , disjoint from  $I_1, \dots, I_n$ , such that

$$|I_{n+1}| > \frac{k_n}{2}.$$

The selected intervals are pairwise disjoint and contained in  $O$ ; hence

$$\sum_{j=1}^{\infty} |I_j| \leq m(O) < \infty.$$

In particular  $|I_j| \rightarrow 0$ . Choose  $N$  so large that

$$\sum_{j>N} |I_j| < \frac{\varepsilon}{5}.$$

Let

$$R = E \setminus \bigcup_{j=1}^N I_j.$$

Fix  $x \in R$ . Since  $\bigcup_{j=1}^N I_j$  is closed and does not contain  $x$ , the Vitali property gives an interval  $I \in \mathcal{I}$  such that  $x \in I$ ,  $I \subset O$ , and  $I \cap I_j = \emptyset$  for  $1 \leq j \leq N$ . The interval  $I$  cannot be disjoint from all selected intervals; if it were, then, for all sufficiently large  $n$ , it would be admissible in the definition of  $k_n$ , so  $|I| \leq k_n < 2|I_{n+1}|$ , and the right-hand side tends to 0, a contradiction.

Let  $j > N$  be the first index for which  $I \cap I_j \neq \emptyset$ . By the choice of  $I_j$ , the interval  $I$  was admissible at step  $j-1$ , and therefore

$$|I| \leq k_{j-1} < 2|I_j|.$$

Since  $x \in I$  and  $I \cap I_j \neq \emptyset$ , the point  $x$  belongs to the interval having the same midpoint as  $I_j$  and five times its length. Denote this enlarged interval by  $5I_j$ . We have proved

$$R \subset \bigcup_{j>N} 5I_j.$$

Consequently,

$$m^*(R) \leq \sum_{j>N} |5I_j| = 5 \sum_{j>N} |I_j| < \varepsilon.$$

□

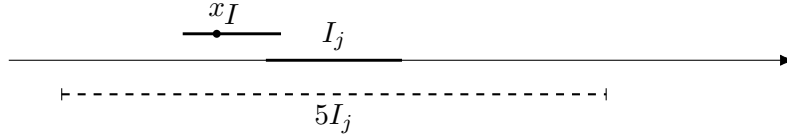


Figure 1: The one-dimensional Vitali argument: if a small interval  $I$  containing  $x$  first intersects a selected interval  $I_j$ , then  $x \in 5I_j$ .

### 3 Dini derivatives and monotone functions

Let  $f : [a, b] \rightarrow \mathbb{R}$ . For  $x \in (a, b)$ , define the four Dini derivatives by

$$\begin{aligned} D^+ f(x) &= \limsup_{h \downarrow 0} \frac{f(x+h) - f(x)}{h}, & D_+ f(x) &= \liminf_{h \downarrow 0} \frac{f(x+h) - f(x)}{h}, \\ D^- f(x) &= \limsup_{h \downarrow 0} \frac{f(x) - f(x-h)}{h}, & D_- f(x) &= \liminf_{h \downarrow 0} \frac{f(x) - f(x-h)}{h}. \end{aligned}$$

Thus  $D_+ f \leq D^+ f$  and  $D_- f \leq D^- f$ . The ordinary derivative exists and equals  $\alpha \in \mathbb{R}$  precisely when all four quantities are equal to  $\alpha$ .

**Lemma 3.1** (Discontinuities of monotone functions). *If  $f : [a, b] \rightarrow \mathbb{R}$  is monotone, then  $f$  has at most countably many discontinuities. In particular,  $f$  is measurable and is continuous almost everywhere.*

*Proof.* Assume, for definiteness, that  $f$  is increasing. For  $x \in (a, b)$ , the one-sided limits

$$f(x-) = \sup_{t < x} f(t), \quad f(x+) = \inf_{t > x} f(t)$$

exist and satisfy  $f(x-) \leq f(x) \leq f(x+)$ . If  $f$  is discontinuous at  $x$ , then  $f(x-) < f(x+)$ , and hence the interval  $(f(x-), f(x+))$  contains a rational number. For distinct discontinuity points these open intervals are disjoint. Assigning one rational number to each discontinuity gives an injection into  $\mathbb{Q}$ . The discontinuities in  $(a, b)$  are therefore countable, and the endpoints add at most two more. □

**Theorem 3.2** (Lebesgue's theorem for monotone functions). *Let  $f : [a, b] \rightarrow \mathbb{R}$  be increasing. Then  $f'(x)$  exists and is finite for almost every  $x \in (a, b)$ . Moreover,  $f'$  is measurable,  $f' \in L^1([a, b])$ , and*

$$\int_a^b f'(x) dx \leq f(b) - f(a).$$

*Proof.* We first prove almost-everywhere existence, allowing temporarily the value  $+\infty$ . It is enough to show that the four Dini derivatives coincide outside a null set.

Fix rational numbers  $0 \leq \alpha < \beta$ , and set

$$E_{\alpha,\beta} = \{x \in (a, b) : D^+f(x) > \beta > \alpha > D_-f(x)\}.$$

We prove  $m^*(E_{\alpha,\beta}) = 0$ . Put  $\delta = m^*(E_{\alpha,\beta})$ , and let  $\eta > 0$ . Choose an open set  $O \supset E_{\alpha,\beta}$  such that  $m(O) < \delta + \eta$ .

For each  $x \in E_{\alpha,\beta}$ , the inequality  $D_-f(x) < \alpha$  gives arbitrarily small  $h > 0$  such that  $[x - h, x] \subset O$  and

$$f(x) - f(x - h) < \alpha h.$$

These intervals form a Vitali cover of  $E_{\alpha,\beta}$ . By Lemma 2.2, choose pairwise disjoint intervals

$$I_j = [x_j - h_j, x_j], \quad 1 \leq j \leq N,$$

such that

$$m^* \left( E_{\alpha,\beta} \setminus \bigcup_{j=1}^N I_j \right) < \eta.$$

Let  $A$  be the part of  $E_{\alpha,\beta}$  covered by the interiors of the intervals  $I_j$ . Only finitely many endpoints are lost, so

$$m^*(A) > \delta - \eta.$$

For every  $y \in A$ , the inequality  $D^+f(y) > \beta$ , together with the fact that  $y$  lies in the interior of some  $I_j$ , gives arbitrarily small  $k > 0$  such that  $[y, y + k]$  is contained in that  $I_j$  and

$$f(y + k) - f(y) > \beta k.$$

A second application of Lemma 2.2 gives pairwise disjoint intervals

$$J_i = [y_i, y_i + k_i], \quad 1 \leq i \leq M,$$

which cover all but a set of outer measure less than  $\eta$  of  $A$ . Hence

$$\sum_{i=1}^M k_i > \delta - 2\eta.$$

Since the intervals  $J_i$  are disjoint and each lies inside one of the intervals  $I_j$ , monotonicity gives, after grouping the  $J_i$ 's according to the containing  $I_j$ ,

$$\begin{aligned} \beta(\delta - 2\eta) &< \sum_{i=1}^M (f(y_i + k_i) - f(y_i)) \\ &\leq \sum_{j=1}^N (f(x_j) - f(x_j - h_j)) < \alpha \sum_{j=1}^N h_j. \end{aligned}$$

Because the  $I_j$ 's are disjoint and contained in  $O$ ,

$$\sum_{j=1}^N h_j \leq m(O) < \delta + \eta.$$

Thus

$$\beta(\delta - 2\eta) < \alpha(\delta + \eta).$$

Letting  $\eta \downarrow 0$ , we obtain  $\beta\delta \leq \alpha\delta$ . Since  $\alpha < \beta$ ,  $\delta = 0$ . Therefore  $D^+f \leq D_-f$  almost everywhere. By the same argument with left and right interchanged,  $D^-f \leq D_+f$  almost everywhere. Outside a null set we have

$$D^+f \leq D_-f \leq D^-f \leq D_+f \leq D^+f,$$

and hence all four Dini derivatives coincide, possibly with value  $+\infty$ .

It remains to exclude an infinite derivative on a set of positive measure and to obtain the integral estimate. Extend  $f$  to  $[a, b+1]$  by setting  $f(x) = f(b)$  for  $x > b$ , and put

$$q_h(x) = \frac{f(x+h) - f(x)}{h}, \quad 0 < h < 1.$$

At every point where  $f'$  exists,  $q_h(x) \rightarrow f'(x)$  as  $h \downarrow 0$ . Since  $q_h \geq 0$ , Fatou's lemma gives

$$\begin{aligned} \int_a^b f'(x) dx &\leq \liminf_{h \downarrow 0} \int_a^b q_h(x) dx \\ &= \liminf_{h \downarrow 0} \frac{1}{h} \left( \int_b^{b+h} f(u) du - \int_a^{a+h} f(u) du \right) \\ &\leq f(b) - f(a). \end{aligned}$$

The last bound is finite; consequently  $f' < \infty$  almost everywhere. Measurability follows by taking the pointwise limit of the measurable difference quotients  $q_{1/n}$  on the set where the derivative exists and defining  $f' = 0$  elsewhere.  $\square$

## 4 Functions of bounded variation

**Definition 4.1** (Positive, negative, and total variation). Let  $f : [a, b] \rightarrow \mathbb{R}$ , and let

$$P = \{a = x_0 < x_1 < \cdots < x_k = b\}$$

be a partition of  $[a, b]$ . Put  $\Delta_i f = f(x_i) - f(x_{i-1})$ , and define

$$\begin{aligned} p(P, f) &= \sum_{i=1}^k (\Delta_i f)^+, & n(P, f) &= \sum_{i=1}^k (\Delta_i f)^-, \\ v(P, f) &= \sum_{i=1}^k |\Delta_i f|, \end{aligned}$$

where  $t^+ = \max\{t, 0\}$  and  $t^- = \max\{-t, 0\}$ . The positive, negative, and total variations of  $f$  on  $[a, b]$  are

$$P_a^b(f) = \sup_P p(P, f), \quad N_a^b(f) = \sup_P n(P, f), \quad V_a^b(f) = \sup_P v(P, f),$$

where the supremum is taken over all partitions of  $[a, b]$ . We say that  $f$  has *bounded variation* on  $[a, b]$ , and write  $f \in \text{BV}([a, b])$ , if  $V_a^b(f) < \infty$ .

**Lemma 4.2** (Variation identities). *If  $f \in \text{BV}([a, b])$ , then*

$$V_a^b(f) = P_a^b(f) + N_a^b(f), \quad f(b) - f(a) = P_a^b(f) - N_a^b(f).$$

*Proof.* For every partition  $P$ ,

$$p(P, f) - n(P, f) = \sum_{i=1}^k \Delta_i f = f(b) - f(a), \quad v(P, f) = p(P, f) + n(P, f).$$

Taking suprema in the first identity gives

$$P_a^b(f) = N_a^b(f) + f(b) - f(a),$$

which is equivalent to the second asserted identity. Combining this relation with  $v(P, f) = 2p(P, f) - [f(b) - f(a)]$ , and taking suprema, gives

$$V_a^b(f) = 2P_a^b(f) - [f(b) - f(a)] = P_a^b(f) + N_a^b(f).$$

□

**Theorem 4.3** (Jordan decomposition for bounded variation). *A function  $f : [a, b] \rightarrow \mathbb{R}$  is of bounded variation if and only if it is the difference of two increasing functions.*

*Proof.* Assume  $f \in \text{BV}([a, b])$ . For  $x \in [a, b]$ , define

$$P(x) = P_a^x(f), \quad N(x) = N_a^x(f).$$

Both  $P$  and  $N$  are increasing. Applying Lemma 4.2 to  $[a, x]$ , we obtain

$$f(x) - f(a) = P(x) - N(x).$$

Thus

$$f(x) = (f(a) + P(x)) - N(x),$$

a difference of two increasing functions.

Conversely, suppose  $f = g - h$ , where  $g$  and  $h$  are increasing. For every partition  $P$ ,

$$\begin{aligned} \sum_{i=1}^k |f(x_i) - f(x_{i-1})| &\leq \sum_{i=1}^k |g(x_i) - g(x_{i-1})| + \sum_{i=1}^k |h(x_i) - h(x_{i-1})| \\ &= g(b) - g(a) + h(b) - h(a). \end{aligned}$$

Taking the supremum over all partitions gives  $V_a^b(f) < \infty$ . □

**Corollary 4.4** (Differentiability and integrability for BV functions). *If  $f \in \text{BV}([a, b])$ , then  $f'$  exists almost everywhere,  $f' \in L^1([a, b])$ , and*

$$\int_a^b |f'(x)| \, dx \leq V_a^b(f).$$

*Proof.* Write  $f = u - v$ , where  $u$  and  $v$  are increasing. By Theorem 3.2,  $u'$  and  $v'$  exist almost everywhere and belong to  $L^1$ . At points where both derivatives exist,  $f' = u' - v'$ , and therefore

$$|f'| \leq u' + v' \quad \text{a.e.}$$

For the particular Jordan decomposition in the proof of Theorem 4.3, the total increase of  $u + v$  over  $[a, b]$  equals  $P_a^b(f) + N_a^b(f) = V_a^b(f)$ . Applying the integral estimate in Theorem 3.2 to  $u$  and  $v$  gives the asserted bound. □

## 5 Differentiation of indefinite integrals on an interval

We now answer the first guiding question. The proof is arranged so that it depends only on the one-dimensional results already established.

**Lemma 5.1.** Let  $f \in L^1([a, b])$ , and define

$$F(x) = \int_a^x f(t) dt.$$

Then  $F$  is continuous and of bounded variation. More precisely,

$$V_a^b(F) \leq \int_a^b |f(t)| dt.$$

*Proof.* Continuity follows from the absolute continuity of the Lebesgue integral: if  $x+h, x \in [a, b]$ , then

$$F(x+h) - F(x) = \int_x^{x+h} f(t) dt \rightarrow 0 \quad (h \rightarrow 0).$$

For a partition  $a = x_0 < \dots < x_k = b$ ,

$$\sum_{i=1}^k |F(x_i) - F(x_{i-1})| = \sum_{i=1}^k \left| \int_{x_{i-1}}^{x_i} f(t) dt \right| \leq \int_a^b |f(t)| dt.$$

Taking the supremum over all partitions proves the variation estimate.  $\square$

**Lemma 5.2.** Let  $g \in L^1([a, b])$ . If

$$\int_a^x g(t) dt = 0 \quad \text{for every } x \in [a, b],$$

then  $g = 0$  almost everywhere.

*Proof.* The hypothesis implies  $\int_s^t g = 0$  for every subinterval  $[s, t] \subset [a, b]$ . Define a finite signed measure  $\mu$  on Borel subsets of  $[a, b]$  by

$$\mu(E) = \int_E g(t) dt.$$

Since  $\mu$  vanishes on all intervals, it vanishes on the algebra of finite disjoint unions of intervals. By the monotone class theorem,  $\mu$  vanishes on all Borel sets. If  $g > 0$  on a set of positive measure, then  $\{g > 1/N\}$  has positive measure for some  $N$ , forcing  $\mu(\{g > 1/N\}) > 0$ , a contradiction. Applying the same argument to  $-g$  excludes a negative part. Hence  $g = 0$  almost everywhere.  $\square$

**Lemma 5.3** (Bounded case). Let  $f \in L^\infty([a, b])$ , and define

$$F(x) = F(a) + \int_a^x f(t) dt.$$

Then

$$F'(x) = f(x) \quad \text{for almost every } x \in [a, b].$$

*Proof.* By Lemma 5.1,  $F$  is of bounded variation; hence  $F'$  exists almost everywhere by Corollary 4.4. Let  $|f| \leq M$  almost everywhere. For  $h > 0$  and  $x, x+h \in [a, b]$ , set

$$Q_h(x) = \frac{F(x+h) - F(x)}{h} = \frac{1}{h} \int_x^{x+h} f(t) dt.$$

Then  $|Q_h(x)| \leq M$ , and  $Q_h(x) \rightarrow F'(x)$  at every point where  $F'$  exists.

Fix  $c \in (a, b)$ , and take  $h > 0$  with  $c + h \leq b$ . Dominated convergence gives

$$\begin{aligned} \int_a^c F'(x) dx &= \lim_{h \downarrow 0} \int_a^c Q_h(x) dx \\ &= \lim_{h \downarrow 0} \frac{1}{h} \int_a^c (F(x+h) - F(x)) dx \\ &= \lim_{h \downarrow 0} \frac{1}{h} \left( \int_c^{c+h} F(u) du - \int_a^{a+h} F(u) du \right) \\ &= F(c) - F(a). \end{aligned}$$

Since  $F(c) - F(a) = \int_a^c f$ , we have

$$\int_a^c (F'(x) - f(x)) dx = 0 \quad (a < c < b).$$

By Lemma 5.2,  $F' = f$  almost everywhere. □

**Theorem 5.4** (Fundamental theorem for Lebesgue integrals). *Let  $f \in L^1([a, b])$ , and define*

$$F(x) = F(a) + \int_a^x f(t) dt.$$

*Then*

$$F'(x) = f(x) \quad \text{for almost every } x \in [a, b].$$

*Proof.* First assume  $f \geq 0$ . Then  $F$  is increasing, so Theorem 3.2 gives that  $F'$  exists almost everywhere and

$$\int_a^b F'(x) dx \leq F(b) - F(a) = \int_a^b f(x) dx.$$

For  $N \in \mathbb{N}$ , let  $f_N = \min\{f, N\}$ , and put

$$F_N(x) = \int_a^x f_N(t) dt.$$

By Lemma 5.3,  $F'_N = f_N$  almost everywhere. Moreover,

$$F(x) - F_N(x) = \int_a^x (f(t) - f_N(t)) dt$$

is increasing; hence its derivative, wherever it exists, is non-negative. Outside a null set depending only on the countable family  $N = 1, 2, \dots$ , we therefore have

$$F'(x) - f_N(x) \geq 0 \quad \text{for every } N.$$

Letting  $N \rightarrow \infty$ , we obtain  $F'(x) \geq f(x)$  almost everywhere. Thus  $F' - f \geq 0$  almost everywhere. Since

$$\int_a^b (F' - f) dx \leq 0,$$

it follows that  $F' = f$  almost everywhere.

For a general  $f \in L^1([a, b])$ , write  $f = f^+ - f^-$ , apply the non-negative case to  $f^+$  and  $f^-$ , and subtract the resulting identities. □

## 6 Absolute continuity

Absolute continuity is precisely the regularity condition that restores the equality

$$\int_a^b F'(x) dx = F(b) - F(a).$$

**Definition 6.1** (Absolute continuity). A function  $F : [a, b] \rightarrow \mathbb{R}$  is *absolutely continuous* on  $[a, b]$ , written  $F \in AC([a, b])$ , if, for every  $\varepsilon > 0$ , there exists  $\delta > 0$  such that

$$\sum_{j=1}^N |F(y_j) - F(x_j)| < \varepsilon$$

whenever  $(x_j, y_j)$ ,  $1 \leq j \leq N$ , are pairwise disjoint subintervals of  $[a, b]$  and

$$\sum_{j=1}^N (y_j - x_j) < \delta.$$

**Lemma 6.2.** *Every absolutely continuous function on  $[a, b]$  is of bounded variation.*

*Proof.* Choose  $\delta > 0$  corresponding to  $\varepsilon = 1$  in the definition of absolute continuity. Let  $N$  be an integer such that  $(b - a)/N < \delta$ , and put

$$q_i = a + i \frac{b - a}{N}, \quad i = 0, 1, \dots, N.$$

For any partition of  $[q_{i-1}, q_i]$ , the total length of its subintervals is  $q_i - q_{i-1} < \delta$ ; hence the corresponding total variation is at most 1. Therefore

$$V_a^b(F) \leq \sum_{i=1}^N V_{q_{i-1}}^{q_i}(F) \leq N < \infty.$$

□

**Lemma 6.3.** *If  $F \in AC([a, b])$  and  $F'(x) = 0$  for almost every  $x \in [a, b]$ , then  $F$  is constant.*

*Proof.* It suffices to show that  $F(c) = F(a)$  for each  $c \in (a, b]$ . Fix such a  $c$ , and let

$$E = \{x \in (a, c) : F'(x) = 0\}.$$

Then  $m(E) = c - a$ . Let  $\varepsilon > 0$ , and choose  $\delta > 0$  from absolute continuity so that every disjoint family of intervals with total length less than  $\delta$  has total  $F$ -oscillation less than  $\varepsilon$ . Fix also  $\eta > 0$ .

For each  $x \in E$ , the intervals  $[x, x + h] \subset (a, c)$  for which

$$|F(x + h) - F(x)| < \eta h$$

form a Vitali cover of  $E$ . By Lemma 2.2, choose pairwise disjoint intervals

$$[x_j, y_j], \quad 1 \leq j \leq N,$$

which cover all of  $E$  except a set of measure less than  $\delta$ . Order these intervals from left to right and include the complementary gaps in  $[a, c]$ , including the two endpoint gaps if present. Since  $E$  has full measure in  $(a, c)$ , the total length of the complementary gaps is less than  $\delta$ . Hence their total  $F$ -oscillation is less than  $\varepsilon$ . On the selected intervals,

$$\sum_{j=1}^N |F(y_j) - F(x_j)| < \eta \sum_{j=1}^N (y_j - x_j) \leq \eta(c - a).$$

Adding the contributions from the selected intervals and the complementary gaps along the resulting partition of  $[a, c]$ , we get

$$|F(c) - F(a)| \leq \varepsilon + \eta(c - a).$$

Since  $\varepsilon > 0$  and  $\eta > 0$  are arbitrary,  $F(c) = F(a)$ . □

**Theorem 6.4** (Characterization of indefinite integrals). *For a function  $F : [a, b] \rightarrow \mathbb{R}$ , the following are equivalent.*

- (i)  $F \in \text{AC}([a, b])$ .
- (ii) *There exists  $f \in L^1([a, b])$  such that*

$$F(x) = F(a) + \int_a^x f(t) dt, \quad x \in [a, b].$$

*In this case  $f = F'$  almost everywhere, and*

$$F(x) = F(a) + \int_a^x F'(t) dt, \quad \int_a^b F'(t) dt = F(b) - F(a).$$

*Proof.* Assume first that  $F(x) = F(a) + \int_a^x f(t) dt$  for some  $f \in L^1([a, b])$ . Given  $\varepsilon > 0$ , choose  $\delta > 0$  such that  $m(E) < \delta$  implies  $\int_E |f| < \varepsilon$ . If  $(x_j, y_j)$  are pairwise disjoint and  $\sum_j (y_j - x_j) < \delta$ , then

$$\sum_j |F(y_j) - F(x_j)| \leq \sum_j \int_{x_j}^{y_j} |f(t)| dt = \int_{\cup_j (x_j, y_j)} |f(t)| dt < \varepsilon.$$

Thus  $F$  is absolutely continuous.

Conversely, suppose  $F \in \text{AC}([a, b])$ . By Lemma 6.2,  $F \in \text{BV}([a, b])$ ; hence  $F'$  exists almost everywhere and belongs to  $L^1([a, b])$  by Corollary 4.4. Define

$$G(x) = \int_a^x F'(t) dt.$$

The first part of the proof shows that  $G \in \text{AC}([a, b])$ . Hence  $F - G \in \text{AC}([a, b])$ , and Theorem 5.4 gives

$$(F - G)' = F' - G' = 0 \quad \text{almost everywhere.}$$

By Lemma 6.3,  $F - G$  is constant. Since  $G(a) = 0$ , the constant is  $F(a)$ , and therefore

$$F(x) = F(a) + \int_a^x F'(t) dt.$$

□

## 7 From intervals to balls

The identity  $F' = f$  for indefinite integrals can be rewritten as a statement about local averages. If

$$F(x) = \int_a^x f(t) dt,$$

then

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h} \int_x^{x+h} f(t) dt.$$

Thus the one-dimensional theorem says that averages over intervals shrinking to  $x$  recover  $f(x)$  for almost every  $x$ .

In  $\mathbb{R}^n$ , the corresponding assertion is that, for  $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ ,

$$\lim_{\substack{m(B) \rightarrow 0 \\ x \in B}} \frac{1}{m(B)} \int_B f(y) dy = f(x) \quad \text{for almost every } x \in \mathbb{R}^n,$$

where  $B$  ranges over Euclidean balls containing  $x$ . If  $f$  is continuous at  $x$ , this is immediate: whenever  $x \in B(c, r)$  and  $r \rightarrow 0$ , every  $y \in B(c, r)$  satisfies  $|y - x| < 2r$ , so

$$\left| \frac{1}{m(B)} \int_B f(y) dy - f(x) \right| \leq \frac{1}{m(B)} \int_B |f(y) - f(x)| dy \rightarrow 0.$$

For general integrable functions, pointwise continuity is replaced by a maximal-function estimate.

## 8 The Hardy–Littlewood maximal function

**Definition 8.1** (Centered and uncentered maximal functions). For  $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ , the *uncentered Hardy–Littlewood maximal function* is

$$Mf(x) = \sup_{B \ni x} \frac{1}{m(B)} \int_B |f(y)| dy,$$

where the supremum is taken over all Euclidean balls  $B \subset \mathbb{R}^n$  containing  $x$ . The centered maximal function is

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

*Remark 8.2.* The two maximal functions are pointwise comparable:

$$M_c f \leq Mf \leq 2^n M_c f.$$

Indeed, if  $x \in B(c, r)$ , then  $B(c, r) \subset B(x, 2r)$ , and therefore

$$\frac{1}{m(B(c, r))} \int_{B(c, r)} |f| \leq 2^n \frac{1}{m(B(x, 2r))} \int_{B(x, 2r)} |f|.$$

Thus all qualitative conclusions below are the same for the centered and uncentered versions, although the constants differ.

**Lemma 8.3** (Finite covering lemma for balls). *Let  $\mathcal{B} = \{B_1, \dots, B_N\}$  be a finite collection of balls in  $\mathbb{R}^n$ . Then there exists a pairwise disjoint subcollection  $B_{i_1}, \dots, B_{i_k}$  such that*

$$m \left( \bigcup_{j=1}^N B_j \right) \leq 3^n \sum_{\ell=1}^k m(B_{i_\ell}).$$

*Proof.* Choose a ball of largest radius from  $\mathcal{B}$ , call it  $B_{i_1}$ , and remove from  $\mathcal{B}$  this ball together with every ball intersecting it. From the remaining family choose a ball of largest radius, call it  $B_{i_2}$ , and again remove it and all balls intersecting it. Since the family is finite, the process terminates and gives pairwise disjoint balls  $B_{i_1}, \dots, B_{i_k}$ .

Every original ball  $B$  either is selected or was removed because it intersects some selected ball  $B_{i_\ell}$  whose radius is at least the radius of  $B$ . If  $B = B(c, r)$  intersects  $B_{i_\ell} = B(c_\ell, r_\ell)$  and  $r \leq r_\ell$ , then for every  $y \in B$ ,

$$|y - c_\ell| \leq |y - c| + |c - c_\ell| < r + (r + r_\ell) \leq 3r_\ell.$$

Thus  $B \subset 3B_{i_\ell}$ , where  $3B_{i_\ell} = B(c_\ell, 3r_\ell)$ . Hence

$$\bigcup_{j=1}^N B_j \subset \bigcup_{\ell=1}^k 3B_{i_\ell}.$$

Taking measures gives

$$m\left(\bigcup_{j=1}^N B_j\right) \leq \sum_{\ell=1}^k m(3B_{i_\ell}) = 3^n \sum_{\ell=1}^k m(B_{i_\ell}).$$

□

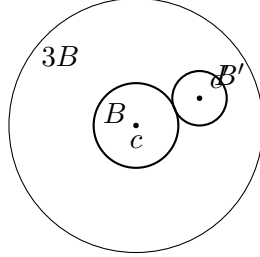


Figure 2: If  $B'$  intersects  $B$  and  $r(B') \leq r(B)$ , then  $B' \subset 3B$ .

**Theorem 8.4** (Weak type (1,1)). *If  $f \in L^1(\mathbb{R}^n)$ , then  $Mf$  is measurable, finite almost everywhere, and for every  $\alpha > 0$ ,*

$$m\{x \in \mathbb{R}^n : Mf(x) > \alpha\} \leq \frac{3^n}{\alpha} \|f\|_{L^1(\mathbb{R}^n)}.$$

*Proof.* For  $\alpha > 0$ , set

$$E_\alpha = \{x \in \mathbb{R}^n : Mf(x) > \alpha\}.$$

If  $x \in E_\alpha$ , then there is a ball  $B \ni x$  such that

$$\frac{1}{m(B)} \int_B |f(y)| dy > \alpha.$$

Every point of this same ball also belongs to  $E_\alpha$ . Hence  $E_\alpha$  is a union of open balls and is therefore open. In particular,  $Mf$  is measurable.

Let  $K \subset E_\alpha$  be compact. For each  $x \in K$ , choose a ball  $B_x \ni x$  such that

$$m(B_x) < \frac{1}{\alpha} \int_{B_x} |f(y)| dy.$$

The balls  $B_x$  cover  $K$ , so a finite subcollection  $B_{x_1}, \dots, B_{x_N}$  covers  $K$ . By Lemma 8.3, choose a pairwise disjoint subcollection  $B_{i_1}, \dots, B_{i_k}$  such that

$$\begin{aligned} m(K) &\leq m\left(\bigcup_{j=1}^N B_{x_j}\right) \leq 3^n \sum_{\ell=1}^k m(B_{i_\ell}) \\ &< \frac{3^n}{\alpha} \sum_{\ell=1}^k \int_{B_{i_\ell}} |f(y)| dy \leq \frac{3^n}{\alpha} \|f\|_1. \end{aligned}$$

Since  $E_\alpha$  is open and Lebesgue measure is inner regular on open sets,

$$m(E_\alpha) \leq \frac{3^n}{\alpha} \|f\|_1.$$

Finally,  $\{Mf = \infty\} \subset E_\alpha$  for every  $\alpha > 0$ . Letting  $\alpha \rightarrow \infty$  in the weak estimate gives  $m\{Mf = \infty\} = 0$ . □

**Theorem 8.5** (Strong type  $(p, p)$ ,  $1 < p \leq \infty$ ). For every  $1 < p \leq \infty$ , the maximal operator  $M$  maps  $L^p(\mathbb{R}^n)$  boundedly into itself. More precisely, if  $1 < p < \infty$ , then

$$\|Mf\|_p \leq \left(\frac{3^n p 2^p}{p-1}\right)^{1/p} \|f\|_p.$$

For  $p = \infty$ ,  $\|Mf\|_\infty \leq \|f\|_\infty$ .

*Proof.* The case  $p = \infty$  is immediate from the definition. Let  $1 < p < \infty$ , and first assume  $f \in L^1 \cap L^p$ . For  $\lambda > 0$ , write

$$f = f_1 + f_2, \quad f_1 = f \mathbf{1}_{\{|f| \leq \lambda/2\}}, \quad f_2 = f \mathbf{1}_{\{|f| > \lambda/2\}}.$$

Then  $Mf_1 \leq \lambda/2$ , and by sublinearity,

$$\{Mf > \lambda\} \subset \{Mf_2 > \lambda/2\}.$$

Using Theorem 8.4,

$$m\{Mf > \lambda\} \leq \frac{2 \cdot 3^n}{\lambda} \int_{\{|f| > \lambda/2\}} |f(x)| dx.$$

The layer-cake formula gives

$$\begin{aligned} \|Mf\|_p^p &= p \int_0^\infty \lambda^{p-1} m\{Mf > \lambda\} d\lambda \\ &\leq 2 \cdot 3^n p \int_0^\infty \lambda^{p-2} \int_{\{|f| > \lambda/2\}} |f(x)| dx d\lambda \\ &= 2 \cdot 3^n p \int_{\mathbb{R}^n} |f(x)| \int_0^{2|f(x)|} \lambda^{p-2} d\lambda dx \\ &= \frac{3^n p 2^p}{p-1} \|f\|_p^p. \end{aligned}$$

For a general  $f \in L^p$ , apply the estimate to  $f_k = f \mathbf{1}_{B(0,k) \cap \{|f| \leq k\}}$ . Then  $|f_k| \uparrow |f|$  pointwise, hence  $Mf_k \uparrow Mf$  pointwise, and Fatou's lemma yields the stated bound.  $\square$

*Remark 8.6* (Why the weak estimate is genuinely weak). The weak type estimate does not say that  $Mf \in L^1$ . In fact, if  $f \in L^1(\mathbb{R}^n)$  is not zero, then  $Mf \notin L^1(\mathbb{R}^n)$ . Choose  $R > 0$  such that

$$c_R = \int_{B(0,R)} |f(y)| dy > 0.$$

For  $|x| > 2R$ , a ball centered at  $x$  with radius comparable to  $|x|$  contains  $B(0, R)$ . Therefore

$$Mf(x) \geq \frac{C}{|x|^n}$$

for all sufficiently large  $|x|$ , where  $C > 0$ . Since  $|x|^{-n}$  is not integrable at infinity in  $\mathbb{R}^n$ ,  $Mf \notin L^1(\mathbb{R}^n)$ .

**Exercise 8.7** (A local non-integrability example). Let

$$f(x) = \frac{1}{x(\log x)^2} \mathbf{1}_{(0,1/2)}(x), \quad x \in \mathbb{R}.$$

Then  $f \in L^1(\mathbb{R})$ , but  $Mf \notin L^1_{\text{loc}}(\mathbb{R})$ . Indeed, for  $0 < x < 1/4$ , the interval  $(0, 2x)$  contains  $x$ , and

$$Mf(x) \geq \frac{1}{2x} \int_0^x \frac{dy}{y(\log y)^2} = \frac{1}{2x |\log x|},$$

which is not integrable near 0.

## 9 The Lebesgue differentiation theorem

**Theorem 9.1** (Lebesgue differentiation theorem). *If  $f \in L^1(\mathbb{R}^n)$ , then*

$$\lim_{\substack{m(B) \rightarrow 0 \\ x \in B}} \frac{1}{m(B)} \int_B f(y) dy = f(x) \quad \text{for almost every } x \in \mathbb{R}^n.$$

*Equivalently, the averages of  $f$  over all balls containing  $x$  and shrinking to  $x$  converge to  $f(x)$  for almost every  $x$ .*

*Proof.* For  $\alpha > 0$ , define the exceptional set

$$N_\alpha = \left\{ x : \limsup_{\substack{m(B) \rightarrow 0 \\ x \in B}} \left| \frac{1}{m(B)} \int_B f(y) dy - f(x) \right| > 2\alpha \right\}.$$

It suffices to prove  $m(N_\alpha) = 0$  for every  $\alpha > 0$ .

Let  $\varepsilon > 0$ . Since  $C_c(\mathbb{R}^n)$  is dense in  $L^1(\mathbb{R}^n)$ , choose  $g \in C_c(\mathbb{R}^n)$  such that

$$\|f - g\|_1 < \varepsilon.$$

For every  $x \in \mathbb{R}^n$ , continuity of  $g$  gives

$$\lim_{\substack{m(B) \rightarrow 0 \\ x \in B}} \frac{1}{m(B)} \int_B g(y) dy = g(x).$$

Moreover,

$$\begin{aligned} \left| \frac{1}{m(B)} \int_B f(y) dy - f(x) \right| &\leq \frac{1}{m(B)} \int_B |f(y) - g(y)| dy \\ &\quad + \left| \frac{1}{m(B)} \int_B g(y) dy - g(x) \right| + |g(x) - f(x)|. \end{aligned}$$

Taking the limsup over balls  $B \ni x$  with  $m(B) \rightarrow 0$ , we get

$$\limsup_{\substack{m(B) \rightarrow 0 \\ x \in B}} \left| \frac{1}{m(B)} \int_B f(y) dy - f(x) \right| \leq M(f - g)(x) + |f(x) - g(x)|.$$

Hence

$$N_\alpha \subset \{x : M(f - g)(x) > \alpha\} \cup \{x : |f(x) - g(x)| > \alpha\}.$$

By Theorem 8.4 and Chebyshev's inequality,

$$m(N_\alpha) \leq \frac{3^n}{\alpha} \|f - g\|_1 + \frac{1}{\alpha} \|f - g\|_1 < \frac{3^n + 1}{\alpha} \varepsilon.$$

Since  $\varepsilon > 0$  is arbitrary,  $m(N_\alpha) = 0$ . Taking the union over rational  $\alpha > 0$  proves the theorem.  $\square$

**Corollary 9.2** (Local form). *If  $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ , then*

$$\lim_{\substack{m(B) \rightarrow 0 \\ x \in B}} \frac{1}{m(B)} \int_B f(y) dy = f(x) \quad \text{for almost every } x \in \mathbb{R}^n.$$

*Proof.* For each  $k \in \mathbb{N}$ , apply Theorem 9.1 to  $f \mathbf{1}_{B(0,k)} \in L^1(\mathbb{R}^n)$ . If  $x \in B(0, k-1)$ , then every ball  $B \ni x$  with sufficiently small radius is contained in  $B(0, k)$ . Hence, for almost every  $x \in B(0, k-1)$ , the averages of  $f \mathbf{1}_{B(0,k)}$  over sufficiently small balls containing  $x$  coincide with the corresponding averages of  $f$ . Taking the countable union over  $k$  proves the assertion.  $\square$

**Definition 9.3** (Lebesgue points). Let  $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ . A point  $x \in \mathbb{R}^n$  is called a *Lebesgue point* of  $f$  if

$$\lim_{r \downarrow 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)| dy = 0.$$

**Corollary 9.4.** *Almost every point of  $\mathbb{R}^n$  is a Lebesgue point of every  $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ .*

*Proof.* It is enough to prove the assertion on each ball  $B(0,k)$ . Approximate  $f\mathbf{1}_{B(0,k+1)}$  in  $L^1$  by a continuous compactly supported function  $g$ , and repeat the proof of Theorem 9.1 with  $|f(y) - f(x)|$  estimated by

$$|f(y) - g(y)| + |g(y) - g(x)| + |g(x) - f(x)|.$$

The middle term has vanishing average at every  $x$  by continuity of  $g$ , and the two remaining terms are controlled by the maximal theorem and Chebyshev's inequality. Letting the approximation error tend to zero gives the result.  $\square$

**Corollary 9.5** (Lebesgue density theorem). *Let  $E \subset \mathbb{R}^n$  be Lebesgue measurable. Then for almost every  $x \in E$ ,*

$$\lim_{\substack{m(B) \rightarrow 0 \\ x \in B}} \frac{m(B \cap E)}{m(B)} = 1.$$

*For almost every  $x \notin E$ , the same limit is 0.*

*Proof.* Apply Corollary 9.2 to  $\mathbf{1}_E \in L^1_{\text{loc}}(\mathbb{R}^n)$ . The displayed quotient is the average of  $\mathbf{1}_E$  over  $B$ , and the limit equals  $\mathbf{1}_E(x)$  for almost every  $x$ .  $\square$

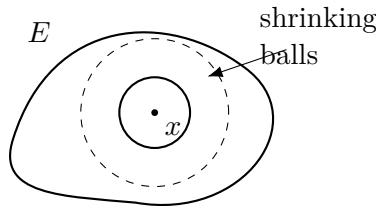


Figure 3: At a density point  $x$  of  $E$ , small balls containing  $x$  are asymptotically filled by  $E$ .

## Concluding perspective

The route followed in these notes may be summarized in one sentence: differentiation theorems are obtained by combining good approximations with quantitative control of the exceptional set. On the line, Vitali's covering lemma controls the oscillations of monotone functions and leads to almost-everywhere differentiability, bounded variation, and absolute continuity. In  $\mathbb{R}^n$ , the finite ball-selection lemma gives the weak  $(1,1)$  estimate for the Hardy–Littlewood maximal operator. This maximal estimate is the analytic mechanism that allows the passage from continuous functions to arbitrary  $L^1$ -functions, and hence proves the Lebesgue differentiation theorem and the density theorem.

## Exercises

**Exercise 9.6.** Show that the exceptional set in Theorem 3.2 can be described using only rational parameters  $\alpha < \beta$ . In particular, justify carefully why proving  $m^*(E_{\alpha,\beta}) = 0$  for all rational  $\alpha < \beta$  is enough.

**Exercise 9.7.** Let  $F \in \text{BV}([a, b])$ . Prove directly from partitions that the variation function  $x \mapsto V_a^x(F)$  is increasing and that both

$$x \mapsto \frac{1}{2}(V_a^x(F) + F(x) - F(a)), \quad x \mapsto \frac{1}{2}(V_a^x(F) - F(x) + F(a))$$

are increasing.

**Exercise 9.8.** Prove the weak type  $(1, 1)$  estimate for the centered maximal function  $M_c$  and compare the resulting constant with the uncentered estimate proved above.

**Exercise 9.9.** Let  $E \subset \mathbb{R}^n$  be measurable. A point  $x \in E$  is called a density point of  $E$  if

$$\lim_{r \downarrow 0} \frac{m(E \cap B(x, r))}{m(B(x, r))} = 1.$$

Derive this centered formulation from Corollary 9.5.

## References

- [1] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*, 2nd ed., Wiley, 1999.
- [2] H. L. Royden and P. M. Fitzpatrick, *Real Analysis*, 4th ed., Pearson, 2010.
- [3] E. M. Stein and R. Shakarchi, *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*, Princeton University Press, 2005.
- [4] R. L. Wheeden and A. Zygmund, *Measure and Integral: An Introduction to Real Analysis*, Marcel Dekker, 1977.