REAL ANALYSIS

LECTURE NOTES

Abstract. The Notes indicate what we do in the lectures, but are not a complete replacement of the book and lectures. The text is from two books of *Real Analysis*:

[1] Xingwei Zhou & Wenchang Sun: Real Variable Analysis, the third edition, Science Press, 2014.

[2] E. Stein & R. Shakarchi: Real Analysis, Princeton University Press, 2005.

1. Differentiability of functions (continued)

We aim to prove the following result.

Theorem 1.1. *Suppose $f \in AC([a,b])$. Then f' exists almost everywhere and is integrable. Moreover,*

$$f(x) - f(a) = \int_a^x f'(t)dt, \quad \text{for all } a \le x \le b.$$

By selecting $x = b$ we get $f(b) - f(a) = \int_a^b f'(t)dt$.

Conversely, if $f \in L^1([a,b])$, then there exists $F \in AC([a,b])$ such that $F'(x) = f(x)$ almost everywhere, and in fact, we may take $F(x) = \int_a^x f(t)dt$.

1.1. Absolutely continuous functions: Proof of Theorem 1.1.

We first show that the total variation of a AC function is also AC.

Lemma 1.1. *Let $f \in AC([a,b])$. Then its total variation $\mathcal{V}_f(a,x) \in AC([a,b])$.*

Proof. Denote in this proof $g(x) = \mathcal{V}_f(a,x)$ for convenience. For disjoint intervals (a_k, b_k), $k = 1, \ldots, N$,

$$\sum_{k=1}^N |g(b_k) - g(a_k)| = \sum_{k=1}^N \mathcal{V}_f(a_k, b_k). \tag{1.1}$$

Given $\varepsilon > 0$, choose $\delta > 0$ small such that for disjoint intervals (a_k, b_k), $k = 1, \ldots, N$,

$$\sum_{k=1}^N |f(b_k) - f(a_k)| < \varepsilon/2 \quad \text{whenever} \quad \sum_{k=1}^N b_k - a_k < \delta. \tag{1.2}$$

For any partition P_k of (a_k, b_k), say $a_k = x_{0,k} < x_{1,k} < \cdots < x_{\ell_k,k} = b_k$, applying (1.2) to disjoint intervals $(x_{j-1,k}, x_{j,k})$, $k = 1, \ldots, N$ and $j = 1, \ldots, \ell_k$, gives

$$\sum_{k=1}^{N} \mathcal{V}_{f|_{[a_k,b_k]}}(P_k) = \sum_{k=1}^{N} \sum_{j=1}^{\ell_k} |f(x_{j,k}) - f(x_{j-1,k})| < \varepsilon/2.$$

Since P_k are arbitrary, it then follows by (1.1) that

$$\sum_{k=1}^{N} |g(b_k) - g(a_k)| < \varepsilon, \quad \text{provided } \sum_{k=1}^{N} b_k - a_k < \delta.$$

$\square$

Proposition 1.1. *Let $f \in AC([a,b])$. Then f is the difference of two increasing absolutely continuous functions.*

Proof. We write $f(x) = [f(x)+\mathcal{V}_f(a,x)]-\mathcal{V}_f(a,x)$. By Lemma 1.1, $\mathcal{V}_f(a,x) \in AC([a,b])$. Recall that $f(x) + \mathcal{V}_f(a,x)$ is increasing. The proposition is thus proved.

$\square$

Suppose $f \in L^1([a,b])$. By Lebesgue differentiation theorem, the indefinite integral

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

is absolutely continuous on $[a,b]$.

Our main goal is to show Theorem 1.1. A covering lemma is needed as a technical tool, which we shall discuss in the following.

1.1.1. *Vitali Covering.*

A collection $\mathcal{B}$ of balls $\{B\}$ is said to be a Vitali covering of a set E if for every $x \in E$ and any $\eta > 0$ there is a ball $B \in \mathcal{B}$, such that $x \in B$ and $m(B) < \eta$. Thus every point is covered by balls of arbitrarily small measure.

Lemma 1.2. *Suppose E is a set of finite measure and $\mathcal{B}$ is a Vitali covering of E. For any $\delta > 0$ we can find finitely many balls $B_1, \ldots, B_N$ in $\mathcal{B}$ that are disjoint and*

$$\sum_{i=1}^{N} m(B_i) \geq m(E) - \delta.$$

Proof of Lemma 1.2. We suppose $m(E) > \delta$ otherwise it is obvious.

Step 1. Let E' be a compact subset of E such that $m(E') \geq \delta$. By the compactness, we can cover E' by finitely many balls in $\mathcal{B}$, and then Lemma **??** allows us to select a disjoint sub-collection of balls $B_1, \dots, B_{N_1}$ such that

$$\sum_{i=1}^{N_1} m(B_i) \geq \frac{1}{A_n} m(E') \geq \frac{\delta}{A_n}.$$

where $A_n = 3^n$, n is the dimension of the background euclidean space.

Step 2. Suppose $\sum_{i=1}^{N_1} m(B_i) < m(E) - \delta$, otherwise we are done. Consider

$$E_2 = E \setminus \Big(\bigcup_{i=1}^{N_1} \overline{B}_i\Big).$$

Then $m(E_2) > \delta$. Repeat the previous argument: choose a compact subset E_2' of E_2; by noting that balls in $\mathcal{B}$ that are disjoint from $\cup_{i=1}^{N_1} \overline{B}_i$ forms a Vitali covering of E_2, we hence can choose a finite disjoint collection of these balls B_i, $N_1 < i \leq N_2$ such that

$$\sum_{N_1 < i \leq N_2} m(B_i) \geq \frac{1}{A_n} m(E_2') \geq \frac{\delta}{A_n}.$$

Therefore balls B_i, $1 \leq i \leq N_2$, are disjoint, and

$$\sum_{i=1}^{N_2} m(B_i) \geq \frac{2\delta}{A_n}.$$

Step 3. We go on such selection if $\sum_{i=1}^{N_2} m(B_i) < m(E) - \delta$. For the k-th stage (if not stopped before then),

$$\sum_{i=1}^{N_k} m(B_i) \geq \frac{k\delta}{A_n}.$$

If we reach $k\delta/A_n \geq m(E) - \delta$, i,e., $k \geq A_n(m(E) - \delta)/\delta$, then $\sum_{i=1}^{N_k} m(B_i) \geq m(E) - \delta$. This proves the lemma.

$\square$

A simple consequence is the following.

Corollary 1.1. *We can arrange the choice of the balls so that*

$$m\Big(E \setminus \bigcup_{i=1}^{N} B_i\Big) < 2\delta.$$

Proof. Take open set $\mathcal{O} \supset E$ with $m(\mathcal{O}) < m(E) + \delta$.

We can restrict all our choices in Lemma 1.2 to balls contained in $\mathcal{O}$. Then

$$[E \setminus \bigcup_{i=1}^{N} B_i] \cup \bigcup_{i=1}^{N} B_i \subset \mathcal{O}.$$

Hence

$$m\Big(E \setminus \bigcup_{i=1}^{N} B_i\Big) \leq m(\mathcal{O}) - m(\bigcup_{i=1}^{N} B_i) \leq m(\mathcal{O}) - m(E) + \delta < 2\delta.$$

□

1.1.2. *Proof of Theorem 1.1.*

The key ingredient of Theorem 1.1 is to show the following.

Theorem 1.2. *If f is absolutely continuous on $[a, b]$, then $f'(x)$ exists almost everywhere. Moreover, if $f'(x) = 0$ for a.e. x, then f is constant.*

Proof. The existence of f' is from $AC([a, b]) \subset BV([a, b])$ and Theorem **??**. We prove the rest of the conclusion.

Step 1. Let

$$E = \{x \in (a, b) : f'(x) \text{ exists and is zero}\}.$$

By our assumption $m(E) = b - a$. Fix $\varepsilon > 0$. Then for each $x \in E$,

$$\lim_{h \to 0} \Big| \frac{f(x + h) - f(x)}{h} \Big| = 0.$$

Then for each $\eta > 0$ we have an open interval $(a_x^\eta, b_x^\eta) \subset [a, b]$ containing x, with

$$|f(b_x^\eta) - f(a_x^\eta)| \leq \varepsilon(b_x^\eta - a_x^\eta) \text{ and } b_x^\eta - a_x^\eta < \eta. \tag{1.3}$$

The collection of such intervals (a_x^η, b_x^η) forms a Vitali covering of E.

Step 2. By Lemma 1.2, for $\delta > 0$, we can select finitely many I_i, $1 \leq i \leq N$, $I_i = (a_i, b_i)$, which are disjoint and such that

$$\sum_{i=1}^{N} m(I_i) \geq m(E) - \delta = b - a - \delta.$$

Since $|f(b_i) - f(a_i)| \le \varepsilon(b_i - a_i)$, we have

$$\sum_{i=1}^{N} |f(b_i) - f(a_i)| \le \varepsilon \sum_{i=1}^{N} (b_i - a_i) \le \varepsilon(b - a). \tag{1.4}$$

Step 3. Write the complement of $\cup_{i=1}^{N} I_i$ in $[a, b]$ as $J = \cup_{j=1}^{M} [\alpha_j, \beta_j]$. By (1.3),

$$m(J) = (b - a) - \sum_{i=1}^{N} m(I_i) \le \delta.$$

Taking δ sufficiently small, we then have by the absolute continuity of f that

$$\sum_{j=1}^{M} |f(\beta_j) - f(\alpha_j)| \le \varepsilon. \tag{1.5}$$

Step 4. Combining (1.4) and (1.5), we conclude

$$|f(b) - f(a)| \le \sum_{i=1}^{N} |f(b_i) - f(a_i)| + \sum_{j=1}^{M} |f(\beta_j) - f(\alpha_j)| \le \varepsilon(b - a) + \varepsilon.$$

Since ε is positive but arbitrary, we infer that $f(a) = f(b)$.

Applying the above argument to $E_{a',b'} = \{x \in (a', b') : f'(x) \text{ exists and is zero}\}$, where $a', b' \in [a, b]$ and $b' > a'$, we find that $f(a') = f(b')$ and so f is constant.

$\square$

It is now the position to finish our main task.

Proof of Theorem 1.1. Let $f \in AC([a, b])$. Then $f = f_1 - f_2$ where $f_1, f_2 \in AC([a, b]) \subset BV([a, b])$ and are both increasing (See Proposition 1.1). The a.e. differentiability of f, f_1, f_2 follows.

Let $\widetilde{f}(x) = \int_a^x f'(t)dt$. Then $\widetilde{f}(x) \in AC([a, b])$; so is the difference

$$g(x) = \widetilde{f}(x) - f(x).$$

By the Lebesgue differentiation theorem, we know that

$$g'(x) = 0 \quad a.e\ x.$$

It follows by Theorem 1.2 that g is a constant. Evaluating this expression at $x = a$ yields

$$f(x) - f(a) = \widetilde{f}(x) = \int_a^x f'(t)dt, \quad \forall\ x \in [a, b].$$

The converse is because $\int_a^x f(t)dt$ is absolutely continuous, and the Lebesgue differentiation theorem gives $g'(x) = f(x)$ almost everywhere.

□

A function of bounded variation is said to be singular provided its derivative vanishes almost everywhere. TheCantor-Lebesgue function is a non-constant singular function. We infer from Theorem 1.1 that an absolutely continuous function is singular if and only if it is constant. Let $f \in BV([a, b])$. Then $f' \in L^1([a, b])$. Define

$$g(x) = \int_a^x f'(t)dt, \quad x \in [a, b].$$

Then we have

$$f(x) = g(x) + h(x) \quad \text{on } [a, b], \tag{1.6}$$

where

$$h(x) = f(x) - \int_a^x f'(t)dt, \quad x \in [a, b]. \tag{1.7}$$

Clearly $g \in AC([a, b])$ and $h \in BV([a, b])$. Moreover, differentiating (1.7) gives $h'(x) = 0$ for a.e. x. Namely h is singular. The decomposition 1.6 of a function of bounded variation f as the sum $g+h$ of two functions of bounded variation, where g is absolutely continuous and h is singular, is called a Lebesgue decomposition of f.

Theorem 1.1 yields the following consequences.

Theorem 1.3 (Integration by parts)**.** *Suppose $f, g \in AC([a, b])$. Then*

$$\int_a^b f'(x)g(x)dx + \int_a^b f(x)g'(x)dx = f(b)g(b) - f(a)g(a).$$

Proof. It is readily seen that $fg \in AC([a, b])$. Then

$$f(b)g(b) - f(a)g(a) = \int_a^b [f(x)g(x)]'dx = \int_a^b [f'(x)g(x) + f(x)g'(x)]dx.$$

□

Theorem 1.4. *Suppose $f \in AC([a, b])$. Then*

$$\mathcal{V}_f(a, b) = \int_a^b |f'(t)|dt. \tag{1.8}$$

Remark 1.1. *The above conclusion holds for complex-valued functions.*

Proof. Given a partition $a = t_0 < t_1 < \cdots < t_N = b$ of $[a, b]$, we have by Theorem 1.1

$$\sum_{j=1}^{N} |f(t_j) - f(t_{j-1})| = \sum_{j=1}^{N} \Big| \int_{t_{j-1}}^{t_j} f'(t)dt \Big| \leq \int_a^b |f'(t)|dt.$$

So this proves

$$\mathcal{V}_f(a, b) \leq \int_a^b |f'(t)|dt. \tag{1.9}$$

For the reverse inequality, fix $\varepsilon > 0$, and find a step function ψ on $[a, b]$ such that $f' = \psi + h$ with [1]

$$\int_a^b |h(t)|dt < \varepsilon.$$

Set

$$\Psi(x) = \int_a^x \psi(t)dt \quad \text{and} \quad H(x) = \int_a^x h(t)dt.$$

Then $f(x) = \Psi(x) + H(x) + f(a)$ [2], and as is easily seen

$$\mathcal{V}_f(a, b) \geq \mathcal{V}_\Psi(a, b) - \mathcal{V}_H(a, b). \tag{1.10}$$

Applying (1.9) to H yields $\mathcal{V}_H(a, b) < \varepsilon$, so that

$$\mathcal{V}_f(a, b) \geq \mathcal{V}_\Psi(a, b) - \varepsilon.$$

Now partition the interval $[a, b]$, as $a = t_0 < \cdots < t_N = b$, so that the step function ψ is constant on each of the intervals (t_{j-1}, t_j), $j = 1, 2, \ldots, N$. Then

$$\mathcal{V}_\Psi(a, b) \geq \sum_{j=1}^{N} \Big| \int_{t_{j-1}}^{t_j} \psi(t) \Big| = \int_a^b |\psi(t)|dt.$$

Consequently, we deduce

$$\mathcal{V}_f(a, b) \geq \int_a^b |\psi(t)|dt - \varepsilon \geq \int_a^b |f'(t)|dt - \int_a^b |h(t)|dt - \varepsilon > \int_a^b |f'(t)|dt - 2\varepsilon.$$

This proves the theorem. $\square$

A geometric application of the above theorem is the length of rectifiable curves. We refer the interested readers to Stein-Shakarchi's book for some details.

[1] Recall that step functions are dense in L^1 space.

[2] As the total variations of f and $f + C$ are the same, one may suppose $f(a) = 0$ directly.

1.2. **Applications: change of variables formula.**

Theorem 1.5 (Variable change formula)**.** *Suppose $g(x)$ is differentiable a.e. on $[a, b]$, $f \in L^1([c, d])$, and $g([a, b]) \subset [c, d]$. Let*

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

Then the following are equivalent:

(i) $F \circ g \in AC([a, b])$;
(ii) $f(g(t))g'(t) \in L^1([a, b])$, *and for* $\alpha, \beta \in [a, b]$

$$\int_{g(\alpha)}^{g(\beta)} f(x)dx = \int_\alpha^\beta f(g(t))g'(t)dt, \tag{1.11}$$

For the sake of Theorem 1.5, we need a couple of lemmas.

Lemma 1.3. *Suppose $f \in AC([a, b])$ and E is measurable subset of $[a, b]$. Then*

(i) $f(E)$ *is measurable.*
(ii) $m(f(E)) = 0$, *provided* $m(E) = 0$.

Proof. This is an exercise.

□

Lemma 1.4. *Let f be a differentiable function on $[a, b]$. Suppose E is a subset of $[a, b]$.*

(i) *If $|f'(x)| \leq M$ for all $x \in [a, b]$, then $m_*(f(E)) \leq Mm_*(E)$,*
(ii) *If E is measurable then $f(E)$ is measurable.*

Proof. This is an exercise.

□

Lemma 1.5. *Let f be a function on $[a, b]$ and E is a subset of $[a, b]$. Suppose f is differentiable on E. Then $m(f(E)) = 0$ if and only if $f'(x) = 0$ for a.e. $x \in E$.*

Proof. "If" part. Fix $N \in \mathbb{N}$. Let

$$\begin{aligned} E_{0,N} &= \left\{x \in E : 0 \leq f(x) < \frac{1}{N}\right\}, \\ E_{1,N} &= \left\{x \in E : \frac{1}{N} \leq f(x) < 1\right\}, \end{aligned}$$

and $E_k = \{x \in E : k \le f(x) < k+1\}$, $k = 1, 2, \cdots$. Then $E = E_{0,N} \cup E_{1,N} \bigcup_{k=1}^{\infty} E_k$ and such union is disjoint. By Lemma 1.4,

$$\begin{aligned} m_*(f(E)) &\le m_*(f(E_{0,N})) + m_*(f(E_{1,N})) + \sum_{k=1}^{\infty} m_*(f(E_k)) \\ &\le \frac{1}{N} m_*(E_{0,N}) + m_*(E_{1,N}) + \sum_{k=1}^{\infty} k m_*(E_k) \\ &\le (b-a)/N. \end{aligned}$$

"Only if" part. Let

$$E_k = \{x \in E : |f(y) - f(x)| > |y - x|/k, \text{ whenever } |x - y| < 1/k\}.$$

Since f is assumed to be differentiable on E, it is not hard to verify that

$$\{x \in E : |f'| > 0\} = \cup_{k \ge 1} E_k.$$

Fix any interval I with $I < 1/k$. We next show that $A = E_k \cap I$ is of measure zero. Once this were proved, we find that $m(E_k) = 0$ by the arbitrariness of I, and thus $\{x \in E : |f'| > 0\}$ is of measure zero.

Since $m(f(A)) = 0$, for each $\varepsilon > 0$, there are disjoint intervals I_j such that

$$f(A) \subset \bigcup_{j=1}^{\infty} I_j, \quad \text{and} \quad \sum_{j=1}^{\infty} |I_j| < \varepsilon.$$

Let $A_j = A \cap f^{-1}(I_j)$. Then $A = \bigcup_{j=1}^{\infty} A_j$. We have

$$m_*(A) \le \sum_{j=1}^{\infty} m_*(A_j) \le \sum_{j=1}^{\infty} k|I_j| < k\varepsilon.$$

Sending $\varepsilon \to 0$, the conclusion is proved.

$\square$

It is now at the position to prove our main result.

Proof of Theorem 1.5.

(ii)$\Longrightarrow$(i). This is by the absolute continuity of integral.

(i)$\Longrightarrow$(ii). Now $F \circ g \in AC([a, b])$ and so is a.e. differentiable. By Theorem 1.1,

$$\int_{g(\alpha)}^{g(\beta)} f = [F \circ g](\beta) - [F \circ g](\alpha) = \int_{\alpha}^{\beta} [F \circ g]'(x) dx \tag{1.12}$$

Let $Z = \{x \in [c,d] : F'(x) \text{ does not exist}\}$ and $X = g^{-1}(Z)$. Note that $m(g(X)) = 0$. Since $F \in AC([c,d])$, we find that $m\big((F \circ g)(X)\big) = 0$ by using Lemma 1.3. Applying Lemma 1.5 to g on X and to $F \circ g$ on X respectively, we deduce that

$$\begin{aligned} g'(x) &= 0 \ \text{ for a.e. } x \in X, \\ [F \circ g]'(x) &= 0 \ \text{ for a.e. } x \in X. \end{aligned}$$

Therefore

$$[F \circ g]'(x) = f(g(x))g'(x) \ \text{ for a.e. } x \in [\alpha, \beta].$$

Hence $f(g(x))g'(x)$ is integrable. Plugging this in (1.12) gives (1.11), as desired.

□

Corollary 1.2. *Suppose $g : [a,b] \to [c,d]$ is absolutely continuous, $f \in L^1([c,d])$. Then* (1.11) *holds if any one of the following holds*

(i) *g is monotone on $[a,b]$;*
(ii) *f is bounded on $[c,d]$;*
(iii) *$f(g(x))g'(x)$ is integrable on $[a,b]$.*

Proof. It is not hard to verify that $F \circ g \in AC([a,b])$, where

$$F(x) = \int_c^x f \in AC([c,d]).$$

The conclusion then follows by Theorem 1.5. □

1.3. **Rademacher differentiable Theorem.**

Definition 1.1. *Let $\Omega \subset \mathbb{R}^n$. A map (function) $f : \Omega \to \mathbb{R}^m$ is called Lipschitz continuous if there is a positive constant C so that*

$$|f(x) - f(y)| \le C|x - y|, \quad \forall \ x, y \in \Omega.$$

The smallest such C is called the Lipschitz constant of f. Namely

$$Lip(f) = \sup\Big\{\frac{|f(x) - f(y)|}{|x - y|} : x, y \in \Omega, \ x \neq y\Big\}.$$

A map (function) $f : \Omega \to \mathbb{R}^m$ is called locally Lipschitz continuous if for each compact subset K of Ω, there is a $C_K > 0$ such that

$$|f(x) - f(y)| \le C_K|x - y|, \quad \forall \ x, y \in K.$$

Definition 1.2. *A map (function) $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $x \in \mathbb{R}^n$ if there is a linear mapping*

$$\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^m$$

such that

$$f(y) = f(x) + \mathcal{L}(y - x) + o(|y - x|) \quad \text{as } y \to x.$$

If such a linear map $\mathcal{L}$ exists, it is clearly unique, and we write $Df(x)$ for $\mathcal{L}$, which is called the derivative of f at x.

The presentation for the celebrated Rademacher differentiable Theorem below is from the book *Measure theory and fine properties of functions* by Evans-Gariepy.

Theorem 1.6 (Rademacher)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz continuous. Then $f(x)$ is differentiable for a.e. x.*

Proof. It suffices to prove the theorem for $m = 1$. Since differentiability is a local property, we suppose f is Lipschitz continuous. Let us denote by $m_{\mathbb{R}^k}(\cdot)$ the Lebesgue measure of $\mathbb{R}^k$ when we need to emphasise the dimension in the following.

Step 1. For each $x \in \mathbb{R}^n$ and $v \in \mathbb{R}^n$ with $|v| = 1$, we define

$$D_v f(x) := \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t}, \quad x \in \mathbb{R}^n.$$

We show that $D_v f(x)$ exists for a.e. x.

Since f is continuous,

$$\overline{D}_v f(x) := \limsup_{t \to 0} \frac{f(x + tv) - f(x)}{t} = \lim_{k \to \infty} \sup_{|t| \in \mathbb{Q} \cap (0, \frac{1}{k})} \frac{f(x + tv) - f(x)}{t}$$

$$\underline{D}_v f(x) := \liminf_{t \to 0} \frac{f(x + tv) - f(x)}{t} = \lim_{k \to \infty} \inf_{|t| \in \mathbb{Q} \cap (0, \frac{1}{k})} \frac{f(x + tv) - f(x)}{t}$$

are measurable functions. As a consequence,

$$\mathcal{S}_v = \{x \in \mathbb{R}^n : D_v f(x) \text{ does not exsit}\} = \{x \in \mathbb{R}^n : \underline{D}_v f(x) < \overline{D}_v f(x)\}$$

is measurable.

Define $\phi : \mathbb{R} \to \mathbb{R}$ by $\phi(t) := f(x + tv)$, $t \in \mathbb{R}$. Observe that ϕ is Lipschitz continuous and so is absolutely continuous, and thus differentiable almost everywhere (Theorem 1.1). Hence

$$m_{\mathbb{R}^1}(\mathcal{S}_v \cap \{tv : t \in \mathbb{R}\}) = 0$$

By Fubini's theorem, we obtain, for each fixed $v \in \mathbb{R}^n$ with $|v| = 1$,

$$m_{\mathbb{R}^n}(\mathcal{S}_v) = 0.$$

Step 2. As as consequence of Step 1, we see that

$$\mathrm{grad} f(x) := (\partial_1 f(x), \ldots, \partial_n f(x)) \ \text{ exists for a.e. } x.$$

We next show that, for each fixed $v \in \mathbb{R}^n$ with $|v| = 1$,

$$D_v f(x) = v \cdot \mathrm{grad} f(x) \ \text{ exists for a.e. } x. \tag{1.13}$$

Write $v = (v_1, \cdots, v_n)$. Let $\zeta \in C_c^\infty(\mathbb{R}^n)$. Then

$$\int_{\mathbb{R}^n} \Big[\frac{f(x+tv) - f(x)}{t}\Big]\zeta(x)dx = -\int_{\mathbb{R}^n} \Big[\frac{\zeta(x) - \zeta(x-tv)}{t}\Big] f(x)dx.$$

Setting $t = 1/k$ above and using the dominated convergence theorem, we find

$$\int_{\mathbb{R}^n} D_v f(x)\zeta(x)dx = -\int_{\mathbb{R}^n} f(x) D_v\zeta(x)dx = -\sum_{i=1}^n v_i \int_{\mathbb{R}^n} f(x)\partial_i\zeta(x)dx.$$

Employing integration by parts (Theorem 1.3) and Fubini's Theorem **??**,

$$\begin{aligned}
\int_{\mathbb{R}^n} f(x)\partial_i\zeta(x)dx &= -\int_{\mathbb{R}^{n-1}} \Big(\int_{\mathbb{R}} f(x)\partial_i\zeta(x)dx_i\Big)dx_1\cdots\widehat{dx_i}\cdots dx_n \\
&= \int_{\mathbb{R}^{n-1}} \Big(\int_{\mathbb{R}} \partial_i f(x)\zeta(x)dx_i\Big)dx_1\cdots\widehat{dx_i}\cdots dx_n = \int_{\mathbb{R}^n} \partial_i f(x)\zeta(x)dx.
\end{aligned}$$

We hence conclude

$$\int_{\mathbb{R}^n} D_v f(x)\zeta(x)dx = \int_{\mathbb{R}^n} v \cdot \mathrm{grad} f(x)\zeta(x)dx.$$

Recall that $C_c^\infty(\mathbb{R}^n)$ is dense in $L^1(\mathbb{R}^n)$. The above equality holding for all $\zeta \in C_c^\infty(\mathbb{R}^n)$ implies (1.13).

Step 3. Choose $\mathcal{U} = \{v_k\}_{k=1}^\infty$ to be a countable, dense subset of unit ball $\mathbb{S}^{n-1}$ in $\mathbb{R}^n$. We take

$$G_k = \{x \in \mathbb{R}^n : D_{v_k} f(x),\ \mathrm{grad} f(x) \text{ exist},\ D_{v_k} f(x) = v_k \cdot \mathrm{grad} f(x)\},$$

and define

$$G := \bigcap_{k=1}^\infty G_k.$$

Then

$$m_{\mathbb{R}^n}(\mathbb{R}^n \setminus G) = 0.$$

Step 4. We finish the proof of the theorem: f is differentiable at each $x \in G$.

For any $x \in G$, $v \in \mathbb{S}^{n-1}$, and $t \in \mathbb{R} \setminus \{0\}$, we define

$$Q(x, v, t) := \frac{f(x + tv)0 - f(x)}{t} - v \cdot \mathrm{grad} f(x).$$

Then if $v' \in S^{n-1}$, we have

$$\begin{aligned} |Q(x, v, t) - Q(x, v', t)| &\leq \left| \frac{f(x + tv) - f(x + tv')}{t} \right| + |v - v'||\mathrm{grad} f(x)| \\ &\leq C_n \mathrm{Lip}(f)|v - v'|. \end{aligned}$$

This together with

$$\lim_{t \to 0} Q(x, v_k, t) = 0, \quad \mathrm{for} v_k \in \mathcal{U}$$

shows that

$$|Q(x, v, t)| \leq |Q(x, v_k, t)| + |Q(x, v, t) - Q(x, v_k, t)| \to 0.$$

For any $y \neq x$, write $v = (y - x)/|y - x|$. Then $y = x + tv$ with $t = |x - y|$. Hence

$$\begin{aligned} f(y) - f(x) - \mathrm{grad} f(x) \cdot (y - x) &= f(x + tv) - f(x) - tv \cdot \mathrm{grad} f(x) \\ &= o(t) = o(|x - y|) \quad \text{as } y \to x. \end{aligned}$$

Hence f is differentiable at x, with

$$Df(x) = \mathrm{grad} f(x).$$

□