Rademacher's Theorem

Caleb Castleberry

November 9, 2015

1 Intro

1.1 Statement

Theorem (Rademacher's Theorem)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a locally Lipschitz function. Then f is differentiable $\mathcal{L}^n$ a.e.*

1.2 Relevant Defintions

Definition 1.2.1 (Absolutely Continuous)**.** We say that a function $f : \mathbb{R} \to \mathbb{R}$ is absolutely continuous (shortened to A.C.) on an interval I if for every $\epsilon > 0$ there exists a δ such that if for any finite disjoint collection of intervals $(a_i, b_i) \subset I$ such that

$$\sum_i |b_i - a_i| < \delta$$

then

$$\sum_i |f(b_i) - f(a_i)| < \epsilon.$$

Definition 1.2.2.

2 One Dimension

In one dimension we have that Lipshitz functions are absolutely continuous. We will then show for continuous, non-decreasing functions we have that absolutely continuity implies differentiable a.e. We then show that any absolutely continuous function can be written in terms of such functions and so must also be differentiable a.e. Note that throughout this section m will denote 1-dimensional Lebesgue measure.

Lemma 2.0.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be non-decreasing and absolutely continuous on an interval I. Then f maps sets of measure zero to sets of measure zero.*

Proof. Let f be non-decreasing and absolutely continuous on some interval I. Choose a measurable set $E \subset I$ with $m(E) = 0$. Let $\epsilon > 0$ be given and choose a $\delta > 0$ specified by ϵ, f, and I in the definition of absolutely continuous. Then, since E is measure zero, there exists a disjoint collection of open intervals $V = \cup_i(\alpha_i, \beta_i)$ such that $E \subset V \subset I$ and $\sum_i |\beta_i - \alpha_i| < \delta$. Then we have that

$$m(f(E)) \leq m(f(V)) \leq \sum_i |f(\beta_i) - f(\alpha_i)| \leq \epsilon$$

It follows that $m(f(E)) = 0$ which is what we wanted. □

Lemma 2.0.2. *Let $f : \mathbb{R} \to \mathbb{R}$ be non-decreasing and continuous. If f sends sets of measure zero to sets of measure zero then f is differentiable a.e.*

Proof. Let f be non-decreasing and continuous on an interval I. Define $g(x) = x + f(x)$. Then g is strictly increasing, continuous and sends sets of measure zero to sets of measure zero.

Now suppose that $E \subset I$ is measurable. Since m is Borel regular, we can write $E = E_1 \cup E_0$ where E_1 is an F_σ set and $m(E_0) = 0$. We know that $m(g(E_0) = 0$. We also know that $g(E_1)$ is also an F_σ set since g is continuous. Thus $g(e) = g(E_0) \cup g(E_1)$ is measurable.

This means that we can define a measure μ on I by $\mu(U) = m(g(U))$. Since g sends sets of measure zero to sets of measure zero we have that μ is absolutely continuous as a measure with respect to m. Thus the Radon-Nikodym Theorem says that $d\mu = h\ dm$ where h is an L^1 on I. It follows that for $[a, x] \subset I$

$$g(x) - g(a) = m(g([a, x])) = \mu([a, x]) = \int_{[a,x]} h\ dm = \int_a^x h(t)\ dt$$

which then gives us

$$f(x) - f(a) = \int_a^x (h(t) - 1)\ dt$$

We then have that f is differentiable and $f'(x) = h(x) - 1$ at every Lebesgue point of h. Then since h is L^1 this is almost everywhere. □

Lemma 2.0.3. *Suppose $f : I \to \mathbb{R}$ where $I = [a, b]$, be absolutely continuous. Define*

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

where $x \in [a, b]$ and the supremum is taken over all N and over all choices of $\{t_i\}$ where

$$a = t_1 < t_1 < \cdots < t_N = x.$$

Then the functions F, $F + f$ and $F - f$ are all A.C. and non-decreasing on I.

Proof. Let I, f, and F be as given in the statement of the lemma. Let $a \leq x < y \leq b$. Then

$$F(y) \geq |f(y) - f(x)| + \sup \sum_{i=1}^{N} |f(t_i) - f(t_{i-1})|$$

and so we have that

$$F(y) \geq |f(y) - f(x)| + F(x).$$

It follows with a little algebra that F, $F - f$, and $F + f$ are non-decreasing.

Now the only thing left to prove is that F is A.C. If $(a, b) \subset I$ then

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

with the sup being taken over all partitions of (a, b). Now choose $\epsilon > 0$, and let $\delta > 0$ be the associated value for f and ϵ in the A.C. definition. Now choose disjoint segments $(a_j, b_j) \subset I$ such that

$$\sum_{j=1}^{M} |b_j - a_j| < \delta$$

From the above equality we have that

$$\sum_{j=1}^{M} |F(b_j) - F(a_j)| = \sum_{j=1}^{M} \sup \sum_{i=1}^{N_j} |f(t_{j_i}) - f(t_{j_i - 1})| \leq \epsilon.$$

Thus we have that F is A.C. on I. □

Proof of Theorem in One Dimension. Let $f : \mathbb{R} \to \mathbb{R}$ be a Lipschitz continuous function on an interval $I = [a, b]$. Then f is A.C. on I. We then define

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

as in the previous lemma. Then from above we have that $F - f$ and $F + f$ are absolutely continuous and non-decreasing on I. Thus both of these functions are differentiable almost everywhere from the first two lemmas. We then can write

$$f(x) = \frac{1}{2}(F + f) - \frac{1}{2}(F - f)$$

and so f is differentiable almost everywhere. □

3 Higher Dimensions

Here we are going to generalize this to any dimension. The one dimensional case will allow us to say that directional derivatives exists a.e. and what is left is to show that this implies the general derivatives exists.

Proof of Rademacher's Theorem. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a locally Lipschitz function. For convenience we may assume that $m = 1$ and and f is Lipschitz.

Now fix $v \in \mathbb{R}^n$ with $|v| = 1$ and define

$$D_v f(x) \equiv \lim_{t \to 0} \frac{f(x+t) - f(x)}{t}$$

when this limit exists.

Our first step will be to show that $D_v f$ exists $\mathcal{L}^n$ a.e.

To do this we first note that

$$\overline{D}_v f(x) = \limsup_{t \to 0} \frac{f(x+t) - f(x)}{t}$$

is Borel measurable as is the similarly defined $\underline{D}_v f(x)$. Then the set

$$\begin{aligned} A_v &= \{x \in \mathbb{R}^n : D_v f(x) \text{does not exist}\} \\ &= \{x \in \mathbb{R}^n : \overline{D}_v f(x) > \underline{D}_v f(x)\} \end{aligned}$$

is Borel measurable. Now for each $x \in \mathbb{R}^n$ and $v \in \mathbb{R}^n$ with $|v| = 1$ we define

$$\varphi(t) = f(x + tv).$$

Then φ is Lipschitz and so from the earlier section is differentiable $\mathcal{L}^1$ a.e. It follows that $\mathcal{H}^1(A_v \cap L) = 0$ for every line which is parallel to v. Then using Fubini's Theorem we have that

$$\begin{aligned} \mathcal{L}^n(A_v) &= \int_{\{\langle x,v\rangle = 0\}} \mathcal{H}^1(A_v \cap L_x)\, dx \\ &= 0 \end{aligned}$$

Where L_x is the line parallel to v passing through x.

Thus we have that $D_v f(x)$ exists a.e. It follows that

$$\operatorname{grad} f(x) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)$$

exists a.e.

We will now show that $D_v f(x) = \operatorname{grad} f(x) \cdot v$ a.e.

To do this choose $\zeta \in C_0^\infty(\mathbb{R}^n)$ and note that

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

Now let $t = \frac{1}{k}$ for $k \in \mathbb{N}$ in the above inequality and note that

$$\left| \frac{f(x + \frac{1}{k}v) - f(x)}{\frac{1}{k}} \right| \le \operatorname{Lip}(f)|v| = \operatorname{Lip}(f)$$

Then we can use the Dominated Convergence Theorem to get that

$$\int_{\mathbb{R}^n} D_v f(x) \zeta(x)\, dx = -\int_{\mathbb{R}^n} f(x) D_v \zeta(x)\, dx$$

$$\begin{aligned}
&= -\sum_{i=1}^{n} \int_{\mathbb{R}^n} f(x) \frac{\partial \zeta}{\partial x_1}(x)\, dx \\
&= \sum_{i=1}^{n} \int_{\mathbb{R}^n} \zeta(x) \frac{\partial f}{\partial x_1}(x)\, dx \\
&= \int_{\mathbb{R}^n} (\mathrm{grad} f(x) \cdot v) \zeta(x)\, dx.
\end{aligned}$$

Since the above equality holds for any choice of $\zeta \in C_0^\infty(\mathbb{R}^n)$, we have that $D_v f(x) = \mathrm{grad} f(x) \cdot v$ $\mathcal{L}^n$ a.e.

Now we choose a countable dense subset $\{v_k\}_{k=1}^\infty$ of the unit sphere and set

$$A_k = \{x \in \mathbb{R}^n : D_v f(x), \mathrm{grad} f(x) \text{both exists and } \mathrm{grad} f(x) \cdot v = D_v f(x)\}$$

for $k = 1, 2, \ldots$. Then define

$$A \equiv \bigcap_{k=1}^{\infty} A_k.$$

Then since $\mathcal{L}^n(\mathbb{R}^n - A_k) = 0$ for each k and there is a countable number of k's we have

$$\mathcal{L}^n(\mathbb{R}^n - A) = 0$$

Our last goal is to now show that f is differentiable at every point $x \in A$.

Now fix $x \in A$. Choose $v \in \partial B(0,1)$, $t \in \mathbb{R}$, $t \neq 0$, and write

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

Then if $v' \in \partial B(0,1)$, we have that

$$\begin{aligned}
|Q(x, v, t) - Q(x, v', t)| &\leq \left| \frac{f(x + tv) - f(x + tv')}{t} \right| + |(v - v') \cdot \mathrm{grad} f(x)| \\
&\leq \mathrm{Lip}(f)|v - v'| + |\mathrm{grad} f(x)||v - v'| \\
&\leq (\sqrt{n} + 1)\mathrm{Lip}(f)|v - v'|.
\end{aligned}$$

Now fix $\epsilon > 0$, and choose N large enough so that if $v \in \partial B(01)$, then

$$|v - v_k| \leq \frac{\epsilon}{2(\sqrt{n} + 1)\mathrm{Lip}(f)}$$

for some $k \in \{1, 2, \ldots N\}$. We also have that

$$\lim_{t \to 0} Q(x, v_k, t) = 0$$

for $k = 1, 2, \ldots, N$ and so there is a $\delta > 0$ such that

$$|Q(x, v_k, t)| < \frac{\epsilon}{2}$$

for $k = 1, 2, \ldots, N$ and $0 < |t| < \delta$. It follows that for every $v \in \partial B(0,1)$, there is a $k \in \{1, 2, \ldots, N\}$ such that

$$|Q(x,v,t)| \leq |Q(x,v_k,t)| + |Q(x,v,t) - Q(x,v_k,t)| < \epsilon$$

if $0 < |t| < \delta$. Now Choose any $y \in \mathbb{R}^n$, $y \neq x$. Write $v = \frac{y-x}{|y-x|}$, so that $y = x + tv$ and $t = |x - y|$. Then

$$\begin{aligned}
\frac{f(y) - f(x) - \mathrm{grad} f(x) \cdot (y - x)}{|y - x|} &= \frac{f(x + tv) - f(x)}{t} - v \cdot \mathrm{grad} f(x) \\
&= Q(x, v, t) \\
&\to 0
\end{aligned}$$

as $t = |y - x| \to 0$. It follows that f is differentiable at x with $Df(x) = \mathrm{grad} f(x)$.
$\square$