AN ELEMENTARY PROOF OF RADEMACHER'S THEOREM

James T. Murphy III

December 4, 2015

1. Introduction

This text was written to be given as an hour long talk. The proof is standard and can be found in many sources, including [3, 4, 6, 8, 9, 5, 1]. Some of the proofs require Sobolev spaces, but we give an elementary one here. By "elementary proof" we mean we will only use material found in a first semester graduate course in real analysis or measure theory. We recommend [7, 2] to refresh on these topics. The author thanks Giovanni Leoni for giving the last step of this proof as a sophomore homework problem; we now understand its importance.

1.1. Review

A Lipschitz function $f : [a, b] \to \mathbb{R}$ is absolutely continuous and hence differentiable almost everywhere. The goal of this paper is to prove an analogous result in $\mathbb{R}^n$. Let us recall the definition of differentiability in higher dimensions.

Definition 1.1. *A function $f : U \to \mathbb{R}^m$ where $U \subseteq \mathbb{R}^n$ is open is **differentiable** at x_0 if there is a linear function $T : \mathbb{R}^n \to \mathbb{R}^m$ such that*

$$\lim_{x \to x_0} \frac{\|f(x) - f(x_0) - T(x - x_0)\|}{\|x - x_0\|} = 0.$$

2. Rademacher's Theorem

Theorem 2.1 (Rademacher)**.** *Let $U \subseteq \mathbb{R}^n$ be open and $f : U \to \mathbb{R}^m$ Lipschitz continuous. Then f is differentiable at almost every $x \in U$.*

Proof. Since f is Lipschitz (resp. differentiable) iff every component of f is Lipschitz (resp. differentiable) we may assume without loss that $m = 1$. Moreover, since U can be covered by countably many balls, we may assume U is a ball. The case $n = 1$ follows immediately from the fact that Lipschitz functions are absolutely continuous on compact intervals, so we assume $n \geqslant 2$. Let M be a Lipschitz constant for f. We proceed by showing that f satisfies many necessary (but not sufficient) conditions of differentiability, and then use these properties to show differentiability.

Claim: For all directions $v \in S^{n-1}$ the directional derivative $\partial_v f(x)$ exists at almost every $x \in U$.

Consider restricting f to lines in the direction of v. Then the derivative of f restricted to an oriented line is the same as the directional derivative of f in that direction. Specifically, for any $w \perp v$ let

$$U_w := \{t \in \mathbb{R} : tv + w \in U\}$$
$$f_w(t) := f(tv + w), \quad t \in U_w.$$

Since U is a ball, each U_w is a (possibly empty) open interval of $\mathbb{R}$. Then we have

$$f'_w(t) = \partial_v f(tv + w)$$

in the sense that $t \in U_w$ iff $tv + w \in U$ and if either side exists then both exist and they are equal. We note that f_w is Lipschitz for each $w \perp v$, so f_w' exists at almost every $t \in U_w$ by the $n = 1$ case of this theorem. Thus $\partial_v f(tv + w)$ exists for almost every $t \in U_w$. Let S be the set of $x \in U$ for which $\partial_v f(x)$ does not exist, and let $S_w := U_w \cap \{t : tv + w \in S\}$. We note that S is measurable because f is continuous, so limits can be taken over a countable dense set. Then by Fubini's theorem

$$\begin{aligned} \lambda^n(S) &= \int_U 1_S(x)\,dx \\ &= \int_{v^\perp}\int_{U_w} 1_S(tv + w)\,dt\,dw \\ &= \int_{v^\perp} \lambda^1(S_w)\,dw \\ &= 0. \end{aligned}$$

Thus $\partial_v f$ exists almost everywhere in U, proving the claim.

Claim: For all $v \in S^{n-1}$, we have $\partial_v f(x) = v \cdot \nabla f(x)$ at almost every $x \in U$.

By the previous claim both sides exist at almost every $x \in U$ and since f is Lipschitz they are in $L^\infty(U)$. Thus it suffices to show

$$\int_U (\partial_v f(x) - v \cdot \nabla f(x))g(x)\,dx = 0$$

for all $g \in C_c^\infty(U)$. As before

$$\begin{aligned} \int_U \partial_v f(x) \cdot g(x)\,dx &= \int_{v^\perp}\int_{U_w} \partial_v f(tv + w) \cdot g(tv + w)\,dt\,dw \\ &= \int_{v^\perp}\int_{U_w} f_w'(t) g_w(t)\,dt\,dw \\ &= -\int_{v^\perp}\int_{U_w} f_w(t) g_w'(t)\,dt\,dw \qquad (f_w, g_w \text{ AC, } \operatorname{supp} g \subset\subset U) \\ &= -\int_{v^\perp}\int_{U_w} f(tv + w) \cdot \partial_v g(tv + w)\,dt\,dw \\ &= -\int_U f(x) \cdot \partial_v g(x)\,dx \end{aligned}$$

where the crucial idea here is that we can integrate absolutely continuous function by parts. Similarly

$$\begin{aligned} \int_U (v \cdot \nabla f(x)) \cdot g(x)\,dx &= \sum_{i=1}^n v_i \int_U \partial_{x_i} f(x) \cdot g(x)\,dx \\ &= -\sum_{i=1}^n v_i \int_U f(x) \cdot \partial_{x_i} g(x)\,dx \\ &= -\int_U f(x) \cdot (v \cdot \nabla g(x))\,dx \\ &= -\int_U f(x) \cdot \partial_v g(x)\,dx \end{aligned}$$

where here we used that since g is smooth $\partial_v g = v \cdot \nabla g$. The claim now follows by subtraction.

Claim: f is differentiable at almost every $x \in U$.

By compactness of S^{n-1}, for each k choose a finite cover $\{B(v_{i,k}, 1/k)\}_{k=1}^{n_k}$. Let $V := \{v_{i,k} : 1 \leqslant i \leqslant n_k, k \in \mathbb{N}\}$ be all centers of such balls. Since V is countable, by the previous claim we can choose $\Omega \subseteq U$ with $\lambda^n(U \setminus \Omega) = 0$ on which $\partial_v f(x) = v \cdot \nabla f(x)$ for all $v \in V, x \in \Omega$. Fix $x_0 \in \Omega$, we will show f is differentiable at x_0. Let $\epsilon > 0$ be given. We will find $\delta > 0$ such that $0 < \|x - x_0\| < \delta$ implies

$$\frac{|f(x) - f(x_0) - (x - x_0) \cdot \nabla f(x_0)|}{\|x - x_0\|} \leqslant \epsilon.$$

For any $v \in S^{n-1}, x \in U$ with $x \neq x_0$

$$\begin{aligned} x &= x - x_0 + x_0 \\ &= v_x r_x + x_0 && (v_x := (x - x_0)/\|x - x_0\|, r_x := \|x - x_0\|) \\ &= (v_x - v)r_x + v r_x + x_0 \end{aligned}$$

so

$$\begin{aligned} f(x) - f(x_0) &= f((v_x - v)r_x + v r_x + x_0) - f(v r_x + x_0) + f(v r_x + x_0) - f(x_0) \\ (x - x_0) \cdot \nabla f(x_0) &= (v_x - v)r_x \cdot \nabla f(x_0) + v r_x \cdot \nabla f(x_0). \end{aligned}$$

Subtracting, taking absolute values, dividing, and using triangle inequality gives

$$\begin{aligned} \frac{|f(x) - f(x_0) - (x - x_0) \cdot \nabla f(x_0)|}{\|x - x_0\|} &\leqslant \left| \frac{f(v r_x + x_0) - f(x_0)}{r_x} - v \cdot \nabla f(x_0) \right| \\ &\quad + \left| \frac{f((v_x - v)r_x + v r_x + x_0) - f(v r_x + x_0)}{r_x} \right| \\ &\quad + |(v_x - v) \cdot \nabla f(x_0)|. \end{aligned}$$

We will show each term can be controlled. By construction of V, for any x, k we can find a $v \in V$ with $\|v_x - v\| \leqslant 1/k$. Let $v(x, k)$ denote such a choice. Then

$$|(v_x - v(x,k)) \cdot \nabla f(x_0)| \leqslant \|v_x - v(x,k)\| \cdot \|\nabla f(x_0)\| \leqslant \frac{1}{k} \cdot \sqrt{n} M,$$

and

$$\left| \frac{f((v_x - v(x,k))r_x + v(x,k)r_x + x_0) - f(v(x,k)r_x + x_0)}{r_x} \right| \leqslant M \|v_x - v(x,k)\| \leqslant M \frac{1}{k}$$

where in these steps we used that f is Lipschitz. Lastly, we use that $x_0 \in \Omega$ so $v \cdot \nabla f(x_0) = \partial_v f(x_0)$ for each $v \in V$, and hence

$$\left| \frac{f(v r_x + x_0) - f(x_0)}{r_x} - v \cdot \nabla f(x_0) \right| \to 0$$

for each fixed $v \in V$ as $x \to x_0$. Since there are only finitely many $\{v_{i,k}\}_{k=1}^{n_k}$, the above convergence is uniform over $v \in V$ from a fixed generation k. Take k large enough that

$$\frac{1}{k}(\sqrt{n} M + 1) < \epsilon/2.$$

Then we may choose δ small enough that

$$\left| \frac{f(v_{i,k} r_x + x_0) - f(x_0)}{r_x} - v_{i,k} \cdot \nabla f(x_0) \right| < \epsilon/2$$

for all $1 \leqslant k \leqslant n_k$ and all $0 < \|x - x_0\| < \delta$. In particular,

$$\left| \frac{f(v(x,k) r_x + x_0) - f(x_0)}{r_x} - v(x,k) \cdot \nabla f(x_0) \right| < \epsilon/2$$

for all x with $0 < \|x - x_0\| < \delta$ which then gives

$$\frac{|f(x) - f(x_0) - (x - x_0) \cdot \nabla f(x_0)|}{\|x - x_0\|} \leqslant \epsilon/2 + \epsilon/2 \leqslant \epsilon$$

for all x with $0 < \|x - x_0\| < \delta$, completing the proof. □

3. References

[1] B. Aslan. The rademacher theorem on ae differentiation of lipschitz functions. 2015.

[2] D. L. Cohn. *Measure theory*, volume 1993. Springer, 1980.

[3] H. Federer. *Geometric measure theory*. Springer, 2014.

[4] J. Heinonen. *Lectures on Lipschitz analysis*. Univ., 2005.

[5] M. Muños. Rademacher's theorem. 2008.

[6] A. Nekvinda and L. Zajíček. A simple proof of the rademacher theorem. *Časopis pro pěstování matematiky*, 113(4):337–341, 1988.

[7] W. Rudin. *Real and complex analysis*. Tata McGraw-Hill Education, 1987.

[8] S. Sternberg. Rademacher's theorem. 2005.

[9] T. Zamojski. Rademacher's theorem in rn. 2008.