Departments of Mathematics
Montana State University
Fall 2015
Prof. Kevin Wildrick

# An introduction to non-smooth analysis and geometry

Lecture 6: Rademacher's Theorem

## 1. Statement and idea of proof

To this point, we've been dealing only with functions defined on the line. Essentially, this made life much simpler because of the "order" on the line; we can talk about things like monotonicity in a simple way. Similarly, our arguments in the BV and absolutely continuous setting heavily used the order, although this can be overcome with some trickery that we will not go into here.

However, the concept of a Lipschitz funciton makes perfect sense in any metric space, in particular in higher dimensional Euclidean space. It also makes sense to ask about the differentiability of Lipschitz functions in higher dimensional Euclidean space.

**Definition 1.1.** Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz function. Then $f$ is differentiable at $m_n$-almost every point.

We will first give the most standard classical proof essentially due to Rademacher, and by the end of the semester we will hopefully have seen a second, totally different proof due to Cheeger.

The classical proof goes essentially as follows: using the one-dimensional case, one can easily show that at $m_n$-almost every point, the directional derivatives of $f$ exist for a dense set of directions. The uniformity provided by the Lipschitz condition then allows us to "extend" these derivatives to the remaining directions and prove differentiability.

Here is an explanation of Cheeger's idea, which I heard from Juha Heinonen.

Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be an $L$-Lipschitz function, and let $a \in \mathbb{R}^n$. For $t > 0$, define the mapping $f_t\colon \overline{B}_{\mathbb{R}^n}(0,1) \to \mathbb{R}$ by

$$f_t(v) := \frac{f(a+tv) - f(a)}{t}.$$

Then

$$|f_t(v) - f_t(v')| = \frac{|f(a+tv) - f(a+tv')|}{t} \leq L|v - v'|,$$

and so $f_t$ is again an $L$-Lipschitz function, and so it maps into $\overline{B}_{\mathbb{R}}(0,L)$.

Recall the Arzéla-Ascoli theorem: if $\mathcal{F}$ is an equicontinuous family of mappings from a separable space $X$ to a metric space $Y$ such that for each $x \in X$, the set

$$\overline{\{f(x) : f \in \mathcal{F}\}}$$

is compact, then every infinite sequence in $\mathcal{F}$ has an infinite subsequence that converges uniformly on compact subsets of $X$.

Applying this theorem to the family $\{f_t\}_{t>0}$, we find that for any sequence $t_n \searrow 0$, there is a limit function $f_0$. Cheeger's idea is the following: if this limit function is unique (independent of the choice of sequence) and linear, then $f$ is differentiable at $a$. He shows that this occurs almost everywhere.

As we will see, this sort of idea is substantially different from the classical proof.

## 2. The classical proof of Rademacher's theorem

*Step 1:* Given a direction $v \in \mathbb{S}^n$, the directional derivative

$$\partial_v f(x) := \lim_{t\to 0} \frac{f(x+tv) - f(x)}{t}$$

exists at $m_n$ almost every point $x \in \mathbb{R}^n$.

To see this, we apply the fact that Lipschitz functions on the real line are differentiable almost everywhere, and Fubini's theorem in polar coordinates.

*Step 2:* By Step 1, for each $v \in \mathbb{S}^n$, there is a measure 0 set $E_v$ such that $\partial_v f(x)$ exists for each $x \in \mathbb{R}^n \backslash E_v$. Hence, given any countable set $S \subseteq \mathbb{S}^{n-1}$, there is a measure 0 set $E_S$ such that all the directional derivatives $\{\partial_v f(x)\}_{v\in S}$ exist when $x \in \mathbb{R}^n \backslash E_v$. In particular, the "gradient vector"

$$\nabla f(x) := (\partial_{e_1} f(x), \ldots, \partial_{e_n} f(x))$$

exists $m_n$-almost everywhere.

Let $S \subseteq \mathbb{S}^{n-1}$ be a countable dense set. We claim that there is a measure 0 set $E$ such that all the directional derivatives $\{\partial_v f(x)\}_{v\in S}$ exist, the gradient $\nabla f(x)$ exists, and

$$\partial_v f(x) = \nabla f(x) \cdot v \tag{2.1}$$

for every $x \in \mathbb{R}^n \backslash E$.

To prove this claim, it suffices to show that for every smooth function $\phi\colon \mathbb{R}^n \to \mathbb{R}$ with compact support,

$$\int_{\mathbb{R}^n} \partial_v f(x)\phi(x) dm_n(x) = \int_{\mathbb{R}^n} (\nabla f(x)\cdot v)\phi(x) dm_n(x).$$

To this end, note that by the Lipschitz condition, for every $t \neq 0$ and $x \in \mathbb{R}^n$,

$$\left| \frac{f(x+vt) - f(x)}{t}\phi(x) \right| \leq L\phi(x),$$

the latter being an integrable function (this is the first step at which BV would not suffice). Hence by the Lebesgue dominated convergence theorem and a change of variables

$$\begin{aligned}
\int_{\mathbb{R}^n} \partial_v f(x)\phi(x)\ dm_n(x) &= \lim_{t\to 0} \int_{\mathbb{R}^n} \frac{f(x+vt) - f(x)}{t}\phi(x)\ dm_n(x) \\
&= \lim_{t\to 0} \int_{\mathbb{R}^n} \frac{\phi(x) - \phi(x-vt)}{t}(-f(x))\ dm_n(x) \\
&= -\int_{\mathbb{R}^n} f(x)\partial_v \phi(x)\ dm_n(x) \\
&= -\int_{\mathbb{R}^n} (\nabla\phi(x)\cdot v) f(x)\ dm_n(x) \\
&= -\sum_{i=1}^n v_i \int_{\mathbb{R}^n} f(x)\partial_{e_i}\phi(x)\ dm_n(x).
\end{aligned}$$

Consider the decomposition of $\mathbb{R}^n$ into $V_i = \operatorname{span}(e_i)$ and $V_i^\perp$. For each $z \in V_i^\perp$, the restriction $f|_{z+V_i}$ is still Lipschitz, and so it is absolutely continuous and hence obeys the fundamental theorem of calculus (as does $\phi|_l$). Hence, the familiar integration by parts formula holds:

$$-\int_{z+V_i} f(s)(\phi|_l)'(s)\ dm_1(s) = \int_{z+V_i} (f|_l)'(s)\phi(s)\ dm_1(s).$$

In other words, denoting by

$$-\int_{z+V_i} f(s)\partial_{e_i}\phi(s)\ dm_1(s) = \int_{z+V_i} \partial_{e_i} f(s)\phi(s)\ dm_1(s).$$

Hence, by Fubini's theorem,

$$\begin{aligned}
-\int_{\mathbb{R}^n} f(x)\partial_{e_i}\phi(x)\ dm_n(x) &= \int_{V_i^\perp}\int_{z+V_i} f(s)\partial_{e_i}\phi(s)\ dm_1(s)\ dm_{n-1}(z)\\
&= \int_{V_i^\perp}\int_{z+V_i} \partial_{e_i} f(s)\phi(s)\ dm_1(s)\ dm_{n-1}(z)\\
&= \int_{\mathbb{R}^n} \partial_{e_i} f(x)\phi(x)\ dm_n(x).
\end{aligned}$$

Hence, we may continue the previous sequence of equations:

$$\begin{aligned}
\int_{\mathbb{R}^n} \partial_v f(x)\phi(x)\ dm_n(x) &= -\sum_{i=1}^n v_i \int_{\mathbb{R}^n} f(x)\partial_{e_i}\phi(x)\ dm_n(x)\\
&= \sum_{i=1}^n v_i \int_{\mathbb{R}^n} \partial_{e_i} f(x)\phi(x)\ dm_n(x)\\
&= \int_{\mathbb{R}^n} (\nabla f(x)\cdot v)\phi(x)\ dm_n(x),
\end{aligned}$$

as required, proving (2.1).

*Step 3:* We now complete the proof by showing that $f$ is differentiable at each point $x \in \mathbb{R}^n \backslash E$.

By definition, it suffices to show that

$$\lim_{w\to 0} \frac{||f(x+w) - f(x) - (\nabla f(x)\cdot w)||_{\mathbb{R}^n}}{||w||_{\mathbb{R}^n}} = 0.$$

Writing $w = tv$ for some $v \in \mathbb{S}^{n-1}$, we see that the above equation is equivalent to the following claim:

$$\lim_{t\to 0} \frac{f(x+tv) - f(x)}{t} - (\nabla f(x)\cdot v) = 0$$

*uniformly* in $v$. For ease of notation, set

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

Fix $\epsilon > 0$. Since $S$ is a dense set in $\mathbb{S}^{n-1}$, we may find a finite collection $\{v_j\} \subseteq S$ such that each $v \in \mathbb{S}^{n-1}$ is within distance $\epsilon/4L$ of some $v_j$. Since the directional derivative of $f$ at $x$ in each direction $v_j$ exists, we may find some $t_\epsilon$ such that for each $t < t_\epsilon$,

$$|Q(v_j,t)| < \frac{\epsilon}{2}.$$

Now, fix an arbitary $v \in \mathbb{S}^{n-1}$ and a direction $v_j$ with $||v - v_j|| < \epsilon/4L$. Then for $t < t_\epsilon$,

$$\begin{aligned}
|Q(v,t)| &\le |Q(v_j,t)| + |Q(v,t) - Q(v_j,t)|\\
&\le \frac{\epsilon}{2} + \frac{|f(x+tv) - f(x+tv_j)|}{t} + |\nabla f\cdot(v - v_j)|\\
&\le \frac{\epsilon}{2} + 2L||v - v_j||_{\mathbb{R}^n} < \epsilon.
\end{aligned}$$

Since $t_\epsilon$ is independent of $v$, this shows the desired uniform convergence and completes the proof of Rademacher's theorem.