Presentation: Rademacher's theorem

Atte Haarala

Geometric measure theory

- Theorem. (Fubini's theorem) Let X and Y be sets. Let μ be a measure on X and ν a measure on Y. If $S \subset X \times Y$ is σ-finite with respect to $\mu \times \nu$, then $S_x := \{ y \mid (x,y) \in S \}$ is ν-measurable for μ-a.e. $x \in X$, in addition the function $x \longmapsto \nu(S_x)$ is μ-integrable and

$$(\mu \times \nu)(S) = \int_X \nu(S_x)\, d\mu(x)$$

(This is the only part of the theorem needed in the proof of Rademacher's theorem.).

Theorem (Rademacher's theorem) Let $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a locally Lipschitz function. Then f is differentiable almost everywhere with respect to Lebesgue measure m_n.

Proof: Assume first that $m = 1$. Later we will prove the general case by simply considering the coordinate functions. Assume also that f is Lipschitz. As differentiability is a local property this can be done without loss of generality. (Consider $f|_{B(0,R)}$.)

Assume $v \in \mathbb{R}^n$ and $|v| = 1$. We will first prove that the directional derivative

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

exists for m_n-a.e. $x \in \mathbb{R}^n$.

Define $\overline{D}_v f: \mathbb{R}^n \longrightarrow \mathbb{R}$ and $\underline{D}_v f: \mathbb{R}^n \longrightarrow \mathbb{R}$ by setting

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

and

$$\underline{D}_v f(x) = \liminf_{t \to 0} \frac{f(x+tv) - f(x)}{t}$$

for every $x \in \mathbb{R}^n$. Because f is continuous we find that

$$\overline{D}_v f(x) = \limsup_{t \to 0} \frac{f(x+tv) - f(x)}{t} = \lim_{k \to \infty} \sup_{0<|t|<\frac{1}{k}} \frac{f(x+tv) - f(x)}{t}$$

$$= \lim_{k \to \infty} \sup_{\substack{t \in \mathbb{Q} \\ 0<|t|<\frac{1}{k}}} \frac{f(x+tv) - f(x)}{t}$$

and

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

for every $x \in \mathbb{R}^n$. As supremums, infimums and limits of measurable function are measurable, we see that $\overline{D}_v f$ and $\underline{D}_v f$ are both measurable. Hence the set

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

is m_n-measurable. Also $D_v f(x) = \overline{D}_v f(x)$ m_n-a.e. $x \in \mathbb{R}^n$ so $D_v f$ is measurable.

Assume then that $x \in \mathbb{R}^n$. Define $\varphi_x : \mathbb{R} \to \mathbb{R}$ by setting $\varphi_x(t) = f(x+tv)$. We see that φ_x is Lipschitz and hence absolutely continuous. Thus φ_x is differentiable for m_1-a.e. $t \in \mathbb{R}$. We notice that

$$D_v f(x+sv) = \lim_{t \to 0} \frac{f(x+sv+tv) - f(x+sv)}{t} = \lim_{t \to 0} \frac{\varphi_x(s+t) - \varphi_x(s)}{t}$$
$$= \varphi_x'(s)$$

for every such $s \in \mathbb{R}$ that $\varphi_x'(s)$ exists. Hence $D_v f(x+sv)$ exists for m_1-a.e. $s \in \mathbb{R}$. Thus $m_1(A_v^x) = 0$ for the set

$$A_v^x := \{ s \in \mathbb{R} : x+sv \in A_v \} = \{ s \in \mathbb{R} : D_v f(x+sv) \text{ does not exist} \}.$$

We can express $\mathbb{R}^n = (\mathbb{R}^n \setminus V) \times V$ where $V = \mathrm{span}(v)$ and $m_n = m_{n-1, \mathbb{R}^n \setminus V} \times m_{1,V}$. We notice that

$$\{ y \in V : (x,y) \in A \} = \{ s \in \mathbb{R} : (x, sv) \in A \}$$
$$= \{ s \in \mathbb{R} : x+sv \in A \}$$
$$= A_v^x$$

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

Because A_v is m_n-measurable and has finite measure, we know by Fubini's theorem that

$$m_n(A_v) = (m_{n-1,\mathbb{R}^n\setminus v} \times m_{1,v})(A_v)$$

$$= \int_{\mathbb{R}^n\setminus v} m_{1,v}(A_v^x)\, dm_{n-1,\mathbb{R}^n\setminus v}(x)$$

$$= \int_{\mathbb{R}^n\setminus v} 0 \, dm_{n-1,\mathbb{R}^n\setminus v}(x) = 0$$

Thus $D_v f(x)$ exists for m_n-a.e. $x \in \mathbb{R}^n$.

So for each $v \in \mathbb{R}^n$, we know that $D_v f(x)$ exists almost everywhere. Thus this holds particularly for the vectors in the standard basis of $\mathbb{R}^n$. As finite union of sets of measure zero has measure zero, we know that $\nabla f(x) = \left(\frac{\partial f}{\partial x^1}(x), \dots, \frac{\partial f}{\partial x^n}(x)\right)$ exists for m_n-a.e. $x \in \mathbb{R}^n$.

Assume $v \in \mathbb{R}^n$. Assume $\varphi \in C_c^\infty(\mathbb{R}^n)$.

We notice that

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

$$= -\int_{\mathbb{R}^n} f(x) \frac{\varphi(x) - \varphi(x-tv)}{t}$$

for every $t \in \mathbb{R}$. We also note that for every $k \in \mathbb{N}$ we have

$$\left| \frac{f(x+\frac{1}{k}v) - f(x)}{\frac{1}{k}} \right| \le \mathrm{Lip}(f)|v| < \infty$$

(Also shows that $|D_v f(x)| \le \mathrm{Lip}(f)|v|$ when $D_v f(x)$ exists)

Hence we may apply dominated convergence theorem to obtain

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

$$= -\sum_{k=1}^n v_k \int f(x) \frac{\partial \varphi}{\partial x^k}(x)\, dx$$

$$= -\sum_{k=1}^n v_k \int \frac{\partial f}{\partial x^k}(x)\varphi(x)\, dx$$

$$= -\int (v \cdot \nabla f(x))\varphi(x)\, dx$$

So $D_v f(x) = v \cdot \nabla f(x)$ for m_n-a.e. $x \in \mathbb{R}^n$

Choose a dense sequence $(v_k)_{k=1}^{\infty}$ of $\partial B(0,1)$. (Take a dense sequence in $\mathbb{Q}^n$ and take directions of those points) Define

$$A_k = \{ x \in \mathbb{R}^n \mid D_{v_k} f(x) \text{ and } \nabla f(x) \text{ exists and } D_{v_k} f(x) = v_k \cdot \nabla f(x) \}$$

for each $k \in \mathbb{N}$. Define also $\hat{A} = \bigcap_k A_k$. As $m_n(A_k^c) = 0$ for each $k \in \mathbb{N}$, we know that $m_n(A^c) = 0$.

We will prove that f is differentiable at each $x \in A$. Assume $x \in A$. Define

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

for all $v \in \partial B(0,1)$ and $t \in \mathbb{R} \setminus \{0\}$. Assume then that $v, v' \in \partial B(0,1)$ and $t \in \mathbb{R} \setminus \{0\}$. We notice that

$$\begin{aligned} |Q(x,v,t) - Q(x,v',t)| &\leq \left| \frac{f(x+tv) - f(x+tv')}{t} \right| + \left| (v - v') \cdot \nabla f(x) \right| \\ &\overset{f \text{ Lipschitz}}{\leq} \operatorname{Lip}(f) |v - v'| + |\nabla f| |v - v'| \\ &\leq (1 + \sqrt{n}) \operatorname{Lip}(f) |v - v'| \end{aligned}$$

Let $\varepsilon > 0$. As $\partial B(0,1)$ is compact, we can find finite number of balls of radius

$$r = \frac{\varepsilon}{4(\sqrt{n}+1) \operatorname{Lip}(f)}$$

so that they cover $\partial B(0,1)$. Because $(v_k)_{k=1}^{\infty}$ is dense in $\partial B(0,1)$, for each of these balls we can choose such $k \in \mathbb{N}$ that v_k is in that ball. Let N be the maximum of these finite number of indices. Then for each $v \in \partial B(0,1)$ we can find such $k \in \{1, \dots, N\}$ that

$$|v - v_k| < 2r = \frac{\varepsilon}{4(\sqrt{n}+1) \operatorname{Lip}(f)}$$

We know that $Q(x, v_k, t) \xrightarrow{t \to 0} 0$ for each $k \in \{1, \dots, N\}$ because $x \in A$ (by the definition of A). Hence we can find such $\delta > 0$ that

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

for each $k \in \{1, \dots, N\}$ and $0 < |t| < \delta$.

Assume now that $v \in \partial B(0,1)$. Choose then such $k \in \mathbb{N}$ that
$$|v - v_k| \leq \frac{\varepsilon}{2(\sqrt{n}+1)\,\mathrm{Lip}(f)}$$
Now combining previous inequalities we find that
$$\begin{aligned} |Q(x,v,t)| &\leq |Q(x,v_k,t)| + |Q(x,v,t) - Q(x,v_k,t)| \\ &< \frac{\varepsilon}{2} + (\sqrt{n}+1)\,\mathrm{Lip}(f) \cdot \frac{\varepsilon}{2(\sqrt{n}+1)\,\mathrm{Lip}(f)} = \varepsilon \end{aligned}$$
for all $0 < |t| < \delta$. So $|Q(x,v,t)| < \varepsilon$ for all $v \in \partial B(0,1)$ when $0 < |t| < \delta$. Thus if $h \in \mathbb{R}^n$ and $|h| < \delta$, and we write $h = |h| v$ where $v \in \partial B(0,1)$, we have
$$\begin{aligned} \frac{|f(x+h) - f(x) - h \cdot \nabla f(x)|}{|h|} &= \left| \frac{f(x+|h|v) - f(x)}{|h|} - v \cdot \nabla f(x) \right| \\ &= |Q(x,v,|h|)| < \varepsilon \end{aligned}$$

We have proved that f is differentiable at x and $Df(x) = \nabla f(x)$. So f is differentiable at each $x \in A$ and hence differentiable m_n-a.e.

If we assume $m > 1$, each coordinate function of f is differentiable m_n-a.e. As there are finite number of coordinate functions, we see that all coordinate functions are differentiable at x for m_n-a.e. $x \in \mathbb{R}^n$. At these points f is differentiable because all the coordinate functions are differentiable.