PROOF OF RADEMACHER'S THEOREM

Theorem. *Let $f(x)$ be a Lipschitz function in $\mathbb{R}^n$. Then $f(x)$ is differentiable almost everywhere.*

Proof. I will break the proof of the theorem into several steps.
Step 1. In the case $n = 1$, the theorem follows from the fact that a Lipschitz function have bounded variation on any finite interval.
Step 2. Let v be a non-zero vector in $\mathbb{R}^n$. The directional derivative

$$f_v(x) = \lim_{\tau \to 0} \frac{f(x + \tau v) - f(x)}{\tau}$$

exists a.e.. In fact, let D_v be the set of all points where $f_v(x)$ exists. Then the one-dimensional measure of the intersection of $\mathbb{R}^n \setminus M_v$ with any line that is parallel to v equals 0 (step 1.) It is an easy exercise to show that the set M_v is measurable. Then Fubini's theorem implies $m(\mathbb{R}^n \setminus M_v) = 0$.
Step 3. By $f_j(x)$ I will denote partial derivatives $\partial f / \partial x_j$. Let $v = (v_1, \dots, v_n)$ be a non-zero vector, and let

$$S_v = \{x : f_v(x), f_1(x), \dots, f_n(x) \text{ exist, and } f_v(x) = v_1 f_1(x) + \dots + v_n f_n(x)\}.$$

Then $m(\mathbb{R}^n \setminus S_v) = 0$. It follows from step 2 that the derivatives $f_v(x)$ and $f_j(x)$, $j = 1, \dots, n$, exist almost everywhere. Take a function $\phi(x) \in C_0^\infty(\mathbb{R}^n)$. By the Dominated Convergence Theorem,

$$\begin{aligned}
\int f_v(x)\phi(x)dx &= \lim_{\tau \to 0} \int \frac{f(x + \tau v) - f(x)}{\tau} \phi(x) dx \\
&= \lim_{\tau \to 0} \int f(x) \frac{\phi(x - \tau v) - \phi(x)}{\tau} dx \\
&= -\int f(x) \sum_{j=1}^{n} v_j \frac{\partial \phi(x)}{\partial x_j} dx = \int \Big(\sum_{j=1}^{n} v_j f_j(x) \Big) \phi(x) dx.
\end{aligned}$$

The last equality is valid for an arbitrary function $\phi(x)$; therefore $f_v = \sum v_j f_j$ a.e..
Step 4. Let Ω be a countable, dense set on the unit sphere in $\mathbb{R}^n$. Take a point $x \in S = \cap_{\omega \in \Omega} S_\omega$. I will show that the function $f(x)$ is differentiable at the point x. For an arbitrary unit vector ω and $\tau > 0$, I define

$$r(\omega, \tau) = \frac{f(x + \tau\omega) - f(x) - \tau \sum \omega_j f_j(x)}{\tau}.$$

One has

$$|r(\tilde{\omega}, \tau) - r(\omega, \tau)| \le C|\tilde{\omega} - \omega| \tag{1}$$

because the function $f(x)$ is Lipschitz. The constant C in (1) depends on the Lipcshitz constant of f only. It follows from step 3 that for every finite set of unit vectors Ω' and for every $\epsilon > 0$ there exists a number $\tau_0(\Omega', \epsilon)$ such that

(2) $$|r(\omega', \tau)| < \frac{\epsilon}{2} \text{ when } \tau < \tau_0(\Omega', \epsilon).$$

The set Ω is dense in the unit sphere, so one can find its finite subset Ω' such that $\operatorname{dist}(\omega, \Omega') < \epsilon/2C$ for every unit vector ω. Here C is the constant from (1). Then, for every unit vector ω there exists a vector $\omega' \in \Omega'$ such that $|\omega - \omega'| < \epsilon/2C$, and, for $\tau < \tau_0(\Omega', \epsilon)$, one has

$$|r(\omega, \tau)| \leq |r(\omega', tau)| + |r(\omega, \tau) - r(\omega', \tau)| < \epsilon.$$

The last inequality means that for every $\epsilon > 0$

$$\left| \frac{f(x+v) - f(x) - \sum v_j f_j(x)}{|v|} \right| < \epsilon$$

when $|v|$ is small enough. This implies differentiability of the function f at the point x.