HAUSDORFF MEASURES & APPLICATIONS

ABSTRACT. The Hausdorff measure is a generalisation of the Lebesgue measure to an arbitrary metric space. This paper will investigate the s-dimensional Hausdorff measure on $\mathbb{R}^N$ and discuss some of the applications that arise, including the Area and Co-Area Formulae.

1. PRELIMINARIES

Recall the definition of the *Lebesgue (outer) measure* of $A \subseteq \mathbb{R}^N$:

$$m_N(A) := \inf \left\{ \sum_{j=0}^{\infty} \operatorname{vol} R_j : A \subseteq \bigcup_{j \in \mathbb{N}} R_j, R_j \text{ open rectangles} \right\}$$

This generalises the familiar notions of length, area and volume to N dimensions. However, the Lebesgue measure has no way to assign N-dimensional measure to $A \subseteq \mathbb{R}^M$ where $M \neq N$. To see how we might be able to construct a measure which can deal with this problem, we first need to decide what N-dimensional measure in $\mathbb{R}^M$ might mean. Certainly rectangles will not be helpful here, so to start with we need to know the Lebesgue measure of another class of sets.

Lemma 1.1. *Let* $\alpha(s) := \dfrac{\pi^{s/2}}{\Gamma(\frac{s}{2}+1)}$.

For all $N \geq 1$, *if* $x \in \mathbb{R}^N, r > 0$ *then*

$$m_N(B_N(x,r)) = \alpha(N) r^N. \tag{1}$$

Proof. By known properties of the Lebesgue measure, (1) is equivalent to

$$m_N(B_N(0,1)) = \alpha(N) \tag{2}$$

We proceed by induction on N. If $N = 1$, then

$$m_1(B_1(0,1)) = 2 = \frac{2\sqrt{\pi}}{\Gamma(\frac{1}{2})} = \frac{\pi^{1/2}}{\Gamma(\frac{1}{2}+1)} = \alpha(1)$$

so (2) holds for $N = 1$.

Assume now $N > 1$ and (2) holds for all $m < N$. For any hyperplane H through the origin, $H \cap B_N(0,1) = B_{N-1}(0,1)$. Hence by Fubini's Theorem,

$$\begin{aligned}
m_N(B_N(0,1)) &= \int_{-1}^{1} m_{N-1}\left(B_{N-1}(0,\sqrt{1-x^2})\right) \,\mathrm{d}x \\
&= \frac{\pi^{\frac{N-1}{2}}}{\Gamma(\frac{N-1}{2}+1)} \int_{-1}^{1} (1-x^2)^{\frac{N-1}{2}} \,\mathrm{d}x \\
&= \frac{\pi^{\frac{N-1}{2}}}{\Gamma(\frac{N-1}{2}+1)} \int_{0}^{1} u^{\frac{N-1}{2}} (1-u)^{-\frac{1}{2}} \,\mathrm{d}u \\
&= \frac{\pi^{\frac{N-1}{2}}}{\Gamma(\frac{N-1}{2}+1)} \cdot \mathrm{B}\left(\frac{N+1}{2}, \frac{1}{2}\right) \\
&= \frac{\pi^{\frac{N-1}{2}}}{\Gamma(\frac{N-1}{2}+1)} \cdot \frac{\Gamma(\frac{N+1}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{N}{2}+1)} \\
&= \frac{\pi^{\frac{N-1}{2}} \cdot \sqrt{\pi}}{\Gamma(\frac{N}{2}+1)} = \alpha(N)
\end{aligned}$$

and (2) follows by induction. □

Recall, for $A \subseteq \mathbb{R}^N$,

$$\operatorname{diam} A := \sup\{\|x-y\| : x, y \in A\}$$

Then by Lemma 1.1 we have

$$m_N\left(B_N(x,r)\right) = \alpha(N)\left(\frac{\operatorname{diam} B_N(x,r)}{2}\right)^N$$

We could then potentially define s-dimensional volume of a set $A \subseteq \mathbb{R}^N$ by

$$\operatorname{vol}_s A := \alpha(s)\left(\frac{\operatorname{diam} A}{2}\right)^s$$

This motivates the following definition.

Definition 1.2. For $A \subseteq \mathbb{R}^N, s \geq 0, \delta > 0$, define

$$\mathcal{H}^s_\delta(A) := \inf \left\{ \sum_{j=0}^{\infty} \alpha(s) \left(\frac{\operatorname{diam} C_j}{2} \right)^s : A \subseteq \bigcup_{j \in \mathbb{N}} C_j, \operatorname{diam} C_j \leq \delta \right\}$$

The *s-dimensional Hausdorff measure* on $\mathbb{R}^N$ is

$$\mathcal{H}^s(A) := \lim_{\delta \to 0} \mathcal{H}^s_\delta(A) = \sup_{\delta > 0} \mathcal{H}^s_\delta(A).$$

From the definition we immediately get some simple properties, the proofs of which can be found in [1]:

Proposition 1.3 (Properties of the Hausdorff Measure)**.** *Suppose $s < t$, $A \subseteq \mathbb{R}^N$. Then the following are true.*

(i) *The Hausdorff measure is an outer measure on $\mathbb{R}^N$.*
(ii) *$\mathcal{H}^0$ is the counting measure.*
(iii) *If $s > N$ then $\mathcal{H}^s(A) = 0$.*
(iv) *If $\mathcal{H}^s(A) < \infty$ then $\mathcal{H}^t(A) = 0$.*
(v) *If $\mathcal{H}^t(A) > 0$ then $\mathcal{H}^s(A) = 0$.*

2. $\mathcal{H}^N = m_N$

The purpose of this section is to prove that $\mathcal{H}^N = m_N$ on $\mathbb{R}^N$. This is done in two steps: firstly, the Isodiametric Inequality is used to prove $\mathcal{H}^N(A) \leq m_N(A)$, then a corollary to the Vitali Covering Lemma is used to show $m_N(A) \leq \mathcal{H}^N(A)$.

Lemma 2.1 (Brunn-Minkowski Inequality)**.** *Suppose $t \in [0,1]$ and $A, B \subseteq \mathbb{R}^N$. Then*

$$m_N \left(tA + (1-t)B\right)^{1/N} \geq t m_N(A)^{1/N} + (1-t) m_N(B)^{1/N}$$

A proof of the Brunn-Minkowski Inequality can be found in [2].

Theorem 2.2 (Isodiametric Inequality)**.** *If $A \subseteq \mathbb{R}^N$, then*

$$m_N(A) \leq \alpha(N) \left(\frac{\operatorname{diam} A}{2} \right)^N \tag{3}$$

Proof. (3) is obvious if $\operatorname{diam} A = \infty$, so assume $\operatorname{diam} A < \infty$. Set $B = \frac{1}{2}(A - A) = \left\{\frac{1}{2}(a - a') : a, a' \in A\right\}$.
The Brunn-Minkowski Inequality (Lemma 2.1) implies

$$m_N(B) \geq \left(\frac{1}{2}m_N(A)^{1/N} + \frac{1}{2}m_N(-A)^{1/N}\right)^N = m_N(A).$$

If $x \in B$, then $x = \frac{1}{2}(a - a')$ for some $a, a' \in A$. So $-x = \frac{1}{2}(a' - a) \in B$, and hence $\|x\| = \frac{1}{2}\|x - (-x)\| \leq \frac{\operatorname{diam} B}{2}$ for any $x \in B$, so $B \subseteq B_N(0, \frac{\operatorname{diam} B}{2})$.

Since by definition $\operatorname{diam} B \leq \operatorname{diam} A$, we get

$$\begin{aligned} m_N(A) \leq m_N(B) &\leq m_N\left(B\left(0, \frac{\operatorname{diam} B}{2}\right)\right) \\ &= \alpha(N)\left(\frac{\operatorname{diam} B}{2}\right)^N \leq \alpha(N)\left(\frac{\operatorname{diam} A}{2}\right)^N. \end{aligned}$$ □

Corollary 2.3. *For any $A \subseteq \mathbb{R}^N$,*

$$m_N(A) \leq \mathcal{H}^N(A).$$

Proof. Fix $\delta > 0$. Suppose $A \subseteq \bigcup_{j\in\mathbb{N}} C_j$, $\operatorname{diam} C_j \leq \delta$. Then

$$m_N(A) \leq \sum_{j=0}^{\infty} m_N(C_j) \leq \sum_{j-0}^{\infty} \alpha(N)\left(\frac{\operatorname{diam} C_j}{2}\right)^N.$$

Taking infimum, we see $m_N(A) \leq \mathcal{H}^N_\delta(A)$. The result follows by letting $\delta \to 0$. □

Intuitively, the Isodiametric Inequality can be thought of as a consequence of "symmetrisation". If $A \subseteq \mathbb{R}^N$ is any set and H is a hyperplane through the origin such that $H \cap A \neq \emptyset$, then we may apply the *Steiner symmetrisation* of A with respect to H by shifting every perpendicular vector to H in A so that it is symmetric about H. Then the diameter of the symmetrisation is no larger than that of the original set; moreover, Lebesgue measure is invariant under such transformations. Applying this to each coordinate axis fits a set with the same measure as A inside a ball with diameter no larger than A; this is exactly the Isodiametric Inequality.

The reverse inequality unfortunately does not have an analogous intuition, but can be proved using a useful lemma.

Lemma 2.4. *Fix $\delta > 0$. If $U \subseteq \mathbb{R}^N$ is any open set then there exist a countable family of disjoint closed balls $\{B_k\}_{k\in\mathcal{N}}$ in U such that* $\operatorname{diam} B_k \leq \delta$ *and*

$$m_N(U) = m_N\left(\bigcup_{k\in\mathbb{N}} B_k\right)$$

We refer the reader to [1] for the proof of both Lemma 2.4 and the Vitali Covering Theorem on which it is based. We now are almost ready to prove the main result of this section.

Lemma 2.5. *If $m_N(A) = 0$, then $\mathcal{H}^N(A) = 0$.*

Proof. Fix $\delta, \epsilon > 0$. From basic measure theoretic results (see [1]), we may suppose $A \subseteq \bigcup_{j\in\mathbb{N}} Q_j$, Q_j dyadic cubes, with $\operatorname{diam} Q_j \leq \delta$, $\sum_{j=0}^{\infty} \operatorname{vol} Q_j < \frac{2^N \epsilon}{N^{N/2}}$.

Noting that $(\operatorname{diam} Q_j)^N = (\sqrt{N})^N \operatorname{vol} Q_j$, we find

$$\begin{aligned}
\mathcal{H}^N_\delta(A) &\leq \sum_{j=0}^{\infty} \alpha(N)\left(\frac{\operatorname{diam} Q_j}{2}\right)^N \\
&= 2^{-N} N^{N/2} \sum_{j=0}^{\infty} \operatorname{vol} Q_j \\
&< \epsilon.
\end{aligned}$$

Letting ϵ, $\delta \to 0$ concludes the proof. □

Theorem 2.6. *$\mathcal{H}_N = m_N$ on $\mathbb{R}^N$.*

Proof. In light of Corollary 2.3 we need only show that $\mathcal{H}^N(A) \leq m_N(A)$ for all $A \subseteq \mathbb{R}^N$. Fix $\delta > 0, A \subseteq \mathbb{R}^N$ and suppose $A \subseteq \bigcup_{j\in\mathbb{N}} R_j$, R_j open rectangles.

By Lemma 2.4, for each $j \in \mathbb{N}$ there exist disjoint open balls $\{B_{j,k}\}_{k\in\mathbb{N}}$ such that

$$\bigcup_{k\in\mathbb{N}} B_{j,k} \subseteq R_j, \text{ and}$$
$$\sum_{k=0}^{\infty} m_N(B_{j,k}) = m_N\left(\bigcup_{k\in\mathbb{N}} B_{j,k}\right) = m_N(R_j).$$

Since $m_N\left(R_j - \bigcup_{k\in\mathbb{N}} B_{j,k}\right) = 0$, by Lemma 2.5 we have that $\mathcal{H}^N_\delta\left(R_j - \bigcup_{k\in\mathbb{N}} B_{j,k}\right) = 0$, that is, $\mathcal{H}^N_\delta\left(\bigcup_{k\in\mathbb{N}} B_{j,k}\right) = \mathcal{H}^N_\delta(R_j)$.

Hence, by Lemma 1.1,

$$\begin{aligned}\mathcal{H}^N_\delta(A) &\leq \sum_{j=0}^{\infty} \mathcal{H}^N_\delta(R_j)\\ &\leq \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \alpha(N)\left(\frac{\operatorname{diam} B_{j,k}}{2}\right)^N\\ &= \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} m_N(B_{j,k})\\ &= \sum_{j=0}^{\infty} m_N(R_j) = \sum_{j=0}^{\infty} \operatorname{vol} R_j.\end{aligned}$$

Taking infimum and then letting $\delta \to 0$ gives the result. □

3. Rademacher's Theorem

In this next section we investigate Rademacher's Theorem, which asserts that Lipschitz functions are differentiable almost everywhere, that is, except on a set of measure zero. In particular, Lipschitz functions are sufficiently smooth to derive (and apply) results that generalise Fubini's Theorem, transformation formulae and more.

Definition 3.1. Let $f : \mathbb{R}^N \to \mathbb{R}^M$ be a function. We say f is *Lipschitz* if there exists $K > 0$ such that, for every $x, y \in \mathbb{R}^N$,
$$\|f(x) - f(y)\| \leq K\|x - y\|.$$
We define the *Lipschitz constant* of f to be
$$\operatorname{Lip}(f) := \sup\left\{\frac{\|f(x) - f(y)\|}{\|x - y\|} : x, y \in \mathbb{R}^N\right\}.$$

Definition 3.2. With f defined as above, we say f is *differentiable* at $a \in \mathbb{R}^N$ if there exists a linear map $L : \mathbb{R}^N \to \mathbb{R}^M$ such that

$$\lim_{x \to a} \frac{f(x) - f(a) - L(y - x)}{\|x - a\|} = 0.$$

Since clearly such a map must be unique, we define

$$Df(a) := L,$$

the *derivative* of f at a.

We first look at the special case where $N = M = 1$, in which case the derivative coincides with our normal notion of one dimensional derivative. This culminates in Lebesgue's Theorem, an important step in proving Rademacher's Theorem.

Definition 3.3. Define the *total variation* of $f : \mathbb{R} \to \mathbb{R}$ by

$$\|f\|_{TV} := \sup_{x_0 < \ldots < x_n} \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|$$

We say f is of *bounded variation* if $\|f\|_{TV} < \infty$.

Lemma 3.4. *If f is of bounded variation, then f is the sum of two monotone functions.*

Proof. Set

$$f^+(x) := \sup_{x_0 < \ldots < x_n \leq x} \sum_{i=1}^{n} \max\{f(x_i) - f(x_{i-1}), 0\}.$$

Clearly f^+ is increasing. It suffices to show that $f - f^+$ is decreasing, that is, that if $x \leq y$ then

$$f^+(x) + f(y) - f(x) \leq f^+(y). \tag{4}$$

Suppose $x_0 < \ldots < x_n \leq x$. Then

$$\begin{aligned}\sum_{i=1}^{n} \max\{f(x_i) - f(x_{i-1}), 0\} + f(y) - f(x) &\leq \sum_{i=1}^{n+2} \max\{f(x_i) - f(x_{i-1}), 0\} \\ &\leq f^+(y)\end{aligned}$$

where we define $x_{n+1} := x$, $x_{n+2} := y$.
Taking supremum we get (4), completing the proof. ☐

Lemma 3.5. *Any monotone function is differentiable almost everywhere.*

The proof of Lemma 3.5 is quite technical and is omitted. It can be found in [3].

Theorem 3.6 (Lebesgue's Theorem)**.** *If $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz, then f is differentiable almost everywhere.*

Proof. In light of Lemmas 3.4 and 3.5, it suffices to show that any Lipschitz function is of bounded variation. Moreover, since differentiability is a local property, we need only show that f is of bounded variation on any bounded interval $[a, b]$.

Suppose $x_0 < \ldots < x_n$. Then

$$\begin{aligned}\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| &\leq \mathrm{Lip}(f) \sum_{i=1}^{n} |x_i - x_{i-1}| \\ &= \mathrm{Lip}(f)(x_n - x_0) \\ &\leq \mathrm{Lip}(f)(b - a)\end{aligned}$$

Taking supremum we find

$$\|f\|_{TV} \leq \mathrm{Lip}(f)(b - a) < \infty. \qquad \square$$

We now move on to the proof of Rademacher's Theorem, which closely follows that given in [1]. Firstly we need some notation.

Definition 3.7. Suppose $f : \mathbb{R}^N \to \mathbb{R}$ is any function. For $v \in \mathbb{R}^N$ such that $\|v\| = 1$, we call

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

the *directional derivative* of f at x *in the direction of* v, whenever this limit exists.

In the following lemmas, we assume always that $f : \mathbb{R}^N \to \mathbb{R}$ is Lipschitz.

Lemma 3.8. *$D_v f(x)$ exists almost everywhere.*

Proof. Define

$$\begin{aligned}D_v^+ f(x) &:= \limsup_{t \to 0} \frac{f(x + tv) - f(x)}{t} \\ D_v^- f(x) &:= \liminf_{t \to 0} \frac{f(x + tv) - f(x)}{t}.\end{aligned}$$

Clearly $D_v f(x)$ exists if and only if

$$x \notin A_v := \left\{ y \in \mathbb{R}^N : D_v^- f(y) < D_v^+ f(x) \right\}$$

For any $x \in \mathbb{R}^N$, define

$$\varphi_x : \mathbb{R} \to \mathbb{R} : t \mapsto f(x + tv)$$

Then for any $s, t \in \mathbb{R}$,

$$\begin{aligned} |\varphi_x(s) - \varphi_x(t)| &= |f(x + sv) - f(x + tv)| \\ &\leq \operatorname{Lip}(f)\|sv - tv\| = \operatorname{Lip}(f)|s - t| \end{aligned}$$

so φ_x is Lipschitz and hence by Lebesgue's Theorem (Theorem 3.6) is differentiable almost everywhere. But

$$\frac{f(x + tv + hv) - f(x + tv)}{h} = \frac{\varphi_x(t + h) - \varphi_x(t)}{h}$$

which has a limit as $h \to 0$ for almost every $t \in \mathbb{R}$, that is, $D_v f(x + tv)$ exists for almost every $t \in \mathbb{R}$.

In other words, $\mathcal{H}^1 (A_v \cap l(x)) = 0$ for all $x \in \mathbb{R}^N$, where

$$l(x) := \{x + tv : t \in \mathbb{R}\}.$$

As Lebesgue (and Hausdorff) measures are rotation invariant we may assume $v \perp \mathbb{R}^{N-1}$, where we embed $R^{N-1} \subset \mathbb{R}^N$ by the map $x \mapsto (x, 0)$. Hence by Tonelli's Theorem and Theorem 2.6,

$$\begin{aligned} m_N(A_v) &= \int_{\mathbb{R}^N} \mathbf{1}_{A_v}(y)\,\mathrm{d}y \\ &= \int_{\mathbb{R}^{N-1}} \left(\int_{\mathbb{R}} \mathbf{1}_{A_v \cap l(x)}(t)\,\mathrm{d}t \right) \mathrm{d}x \\ &= \int_{\mathbb{R}^{N-1}} \mathcal{H}^1 (A_v \cap l(x))\,\mathrm{d}x = 0. \end{aligned}$$ □

Lemma 3.9. $D_v f(x) = \nabla f(x) \cdot v$ *almost everywhere.*
As usual, $\nabla f(x) := \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_N} \right)$.

Proof. Suppose $\phi \in C_c^\infty(\mathbb{R}^N)$. Then

$$\begin{aligned} \left| \frac{f(x + tv) - f(x)}{t} \phi(x) \right| &\leq \frac{\operatorname{Lip}(f)\|tv\|}{|t|} |\phi(x)| \\ &\leq \operatorname{Lip}(f) \sup\{\phi(x) : x \in \mathbb{R}^N\} < \infty. \end{aligned}$$

Then by two applications of the dominated convergence theorem,

$$\begin{aligned}\int_{\mathbb{R}^N} \lim_{t\to 0}\frac{f(x+tv)-f(x)}{t}\phi(x)\,\mathrm{d}x \\ &= \lim_{t\to 0}\left(\int_{\mathbb{R}^N} \frac{f(x+tv)\phi(x)}{t}\,\mathrm{d}x - \int_{\mathbb{R}^N}\frac{f(x)\phi(x)}{t}\,\mathrm{d}x\right) \\ &= \lim_{t\to 0}\left(\int_{\mathbb{R}^N} \frac{f(x)\phi(x-tv)}{t}\,\mathrm{d}x - \int_{\mathbb{R}^N}\frac{f(x)\phi(x)}{t}\,\mathrm{d}x\right) \\ &= \int_{\mathbb{R}^N} f(x)\lim_{t\to 0}\frac{\phi(x-tv)-\phi(x)}{t}\,\mathrm{d}x.\end{aligned}$$

That is,

$$\int_{\mathbb{R}^N} D_v f(x)\phi(x)\,\mathrm{d}x = -\int_{\mathbb{R}^N} f(x) D_v\phi(x)\,\mathrm{d}x. \tag{5}$$

Now as $\phi \in C_c^\infty(\mathbb{R}^N)$, $D_v\phi(x) = \nabla\phi(x)\cdot v = \displaystyle\sum_{i=1}^N v_i \frac{\partial \phi}{\partial x_i}(x)$.

Applying (5) again to $D_{e_j} = \dfrac{\partial}{\partial x_i}(x)$, $i = 1, \ldots, N$,

$$\begin{aligned}\int_{\mathbb{R}^N} D_v f(x)\phi(x)\,\mathrm{d}x &= -\sum_{i=1}^N v_i \int_{\mathbb{R}^N} f(x)\frac{\partial\phi}{\partial x_i}\,\mathrm{d}x \\ &= -\sum_{i=1}^N v_i\left(-\int_{\mathbb{R}^N}\frac{\partial f}{\partial x_i}(x)\phi(x)\,\mathrm{d}x\right) \\ &= \int_{\mathbb{R}^N}(\nabla f(x)\cdot v)\phi(x)\,\mathrm{d}x\end{aligned}$$

As $\phi \in C_c^\infty(\mathbb{R}^N)$ was arbitrary, by the density of $C_c^\infty(\mathbb{R}^N)$ in $\mathcal{L}^p(\mathbb{R}^N)$ we may conclude the result. $\square$

Lemma 3.10. *Let $\mathcal{F} = \{v_k\}_{k\in\mathbb{N}}$ be a countable dense subset of $\partial B := \left\{x \in \mathbb{R}^N : \|x\| = 1\right\}$.*
Define, for each $k \in \mathbb{N}$,

$$A_k := \left\{x \in \mathbb{R}^N : D_{v_k} f(x), \nabla f(x) \text{ exist and } D_{v_k} f(x) = \nabla f(x)\cdot v_k\right\}.$$

Define $A := \displaystyle\bigcap_{k\in\mathbb{N}} A_k$, and for any $x \in A$ put

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

Then $F_x(v,t) \to 0$ as $t \to 0$ uniformly.

Proof. For any $u, v \in \partial B$, we have

$$\begin{aligned}
|F_x(u,t) - F_x(v,t)| &= \left| \frac{f(x+tu) - f(x+tv)}{t} - \nabla f(x) \cdot (u-v) \right| \\
&\leq \frac{|f(x+tu) - f(x+tv)|}{|t|} + \|\nabla f(x)\| \|u-v\| \\
&\leq \frac{\mathrm{Lip}(f)|tu - tv|}{|t|} + \|\nabla f(x)\| \|u-v\| \\
&= (\mathrm{Lip}(f) + \|\nabla f(x)\|)\|u-v\|
\end{aligned}$$

Now since $\left| \frac{\partial f}{\partial x_j}(x) \right| = \left| \lim_{t \to 0} \frac{f(x+te_j) - f(x)}{t} \right| \leq \mathrm{Lip}(f)$,
we have $\|\nabla f(x)\| = \sqrt{\left| \frac{\partial f}{\partial x_1}(x) \right|^2 + \ldots + \left| \frac{\partial f}{\partial x_N}(x) \right|^2} \leq \sqrt{N}\,\mathrm{Lip}(f)$.
So $|F_x(u,t) - F_x(v,t)| \leq (\sqrt{N}+1)\,\mathrm{Lip}(f)\|u-v\|$.

Now fix $\epsilon > 0$, and put $\alpha := \dfrac{\epsilon}{2(\sqrt{N}+1)\,\mathrm{Lip}(f)}$.
Then $\{B_N(v_k, \alpha)\}$ is a cover for ∂B by density of $\mathcal{F}$, and so by compactness of ∂B there is a finite subcover. That is, there exists $K \in \mathbb{N}$ such that for all $v \in \partial B$,

$$\|v - v_k\| < \alpha \text{ for some } k \in \{1, \ldots, K\}.$$

Now $\lim_{t \to 0} F_x(v_k, t) = 0$ by definition of A. Hence there exists $\delta_k > 0$ such that

$$|F_x(v_k, t)| < \tfrac{\epsilon}{2} \text{ whenever } 0 < |t| < \delta_k.$$

Put $\delta := \min\{\delta_1, \ldots, \delta_K\}$. Then for any $v \in \partial B$, if $0 < |t| < \delta$,

$$\begin{aligned}
|F_x(v,t)| &\leq |F_x(v,t) - F_x(v_k,t)| + |F(v_k,t)| \\
&\leq (\sqrt{N}+1)\,\mathrm{Lip}(f)\|v - v_k\| + |F_x(v_k,t)| \\
&< (\sqrt{N}+1)\,\mathrm{Lip}(f)\frac{\epsilon}{2(\sqrt{N}+1)\,\mathrm{Lip}(f)} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Hence $F(v,t) \to 0$ uniformly as claimed. □

All of this comes together neatly to give us a complete proof of Rademacher's Theorem.

Theorem 3.11 (Rademacher's Theorem)**.** *Let $f : \mathbb{R}^N \to \mathbb{R}^M$ be Lipschitz. Then f is differentiable almost everywhere.*

Proof. First assume $M = 1$. By Lemmas 3.8 and 3.9, $D_v f(x)$ exists and is equal to $\nabla f(x) \cdot v$ almost everywhere. Hence $m_N(\mathbb{R}^N - A) = 0$. Fix $\epsilon > 0$ and $x \in A$. Then by Lemma 3.10, there exists $\delta > 0$ such that, if $0 < \|x - y\| < \delta$ then

$$\begin{aligned}\frac{|f(y) - f(x) - \nabla f(x) \cdot (y - x)|}{\|y - x\|} &= \Bigg| \frac{f(x + \frac{y-x}{\|y-x\|}\|y - x\|) - f(x)}{\|y - x\|} \\ &\quad - \nabla f(x) \cdot \left(\frac{y - x}{\|y - x\|} \right) \Bigg| \\ &= \left| F_x \left(\frac{y - x}{\|y - x\|}, \|y - x\| \right) \right| < \epsilon.\end{aligned}$$

Hence f is differentiable at x, and $Df(x)(v) = \nabla f(x) \cdot v$. Moreover, f is differentiable almost everywhere.

If $M > 1$, then for $j \in \{1, \dots, M\}$ we have

$$|f_j(x) - f_j(y)| \leq \|f(x) - f(y)\| \leq \mathrm{Lip}(f)\|x - y\| \text{ for all } x, y \in \mathbb{R}^N.$$

Hence f_j is also Lipschitz, and so by the $M = 1$ case proved above, f_j is differentiable almost everywhere. Suppose $x \in \mathbb{R}^N$ is such that $f_1, \dots, f_M$ are differentiable at x. Then put $L := \left[\begin{array}{ccc} \nabla f_1(x) & \dots & \nabla f_M(x) \end{array} \right]^T$. Hence

$$\begin{aligned}\frac{f(y) - f(x) - L(y - x)}{\|y - x\|} &= \frac{1}{\|y - x\|} \left[\begin{array}{c} f_1(y) - f_1(x) - \nabla f_1(x) \cdot (y - x) \\ \vdots \\ f_M(y) - f_M(x) - \nabla f_M(x) \cdot (y - x) \end{array} \right] \\ &\to 0 \text{ as } y \to x,\end{aligned}$$

completing the proof. □

4. Applications

Rademacher's Theorem gives us sufficient conditions to use differential calculus on functions that is weaker than the normal condition of functions being C^1. In this section we briefly discuss some of the consequences. Proofs will in general be omitted in this section; refer to [1] for the proofs in full.

Definition 4.1. Let $f : \mathbb{R}^N \to \mathbb{R}^M$ be differentiable at $x \in \mathbb{R}^N$. Let $D'f(x)$ denote the matrix of $Df(x)$ with respect to the standard basis, and define

$$Jf(x) := \sqrt{det\left[(D'f(x))^T \circ (D'f(x))\right]},$$

the *Jacobian* of f at x.

Theorem 4.2 (Area Formula)**.** *Let $f : \mathbb{R}^N \to \mathbb{R}^M$ be Lipschitz and $g : \mathbb{R}^N \to \mathbb{R}$ locally integrable, $N \leq M$. Then*

$$\int_{\mathbb{R}^N} g(x)Jf(x)\,\mathrm{d}x = \int_{\mathbb{R}^M} \sum_{x \in f^{-1}[\{y\}]} g(x)\,\mathrm{d}\mathcal{H}^N(y)$$

Corollary 4.3. *Suppose $U \subseteq \mathbb{R}^N$ is measurable and $f : U \to \mathbb{R}^M$ is a Lipschitz embedding, $N \leq M$. Then $S := f(U)$ is an N-dimensional surface in $\mathbb{R}^M$ and*

$$\mathcal{H}^N(S) = \int_U Jf(x)\,\mathrm{d}x.$$

That is, $\mathcal{H}^N$ is the N-dimensional surface measure on $\mathbb{R}^M$.

Proof. Put $g = \mathbf{1}_U$. Then since f is injective,

$$\begin{aligned}\sum_{x \in f^{-1}[\{y\}]} g(x) &= \sum_{y=f(x)} \mathbf{1}_U(x) \\ &= \mathbf{1}_{f(U)}(x).\end{aligned}$$

Hence by the Area Formula (Theorem 4),

$$\begin{aligned}\mathcal{H}^N(S) &= \int_{\mathbb{R}^M} \mathbf{1}_S(y)\,\mathrm{d}\mathcal{H}^N(y) \\ &= \int_U Jf(x)\,\mathrm{d}x\end{aligned}$$

as claimed. □

Example 4.4 (Arc Length). Putting $U = [a, b] \subset \mathbb{R}$, $\mathcal{C} = f(U)$ we deduce that $\mathcal{H}^1$ coincides with our normal definition of arc length of the curve $\mathcal{C}$:

$$\mathcal{H}^1(\mathcal{C}) = \int_{[a,b]} Jf(t)\,\mathrm{d}t = \int_a^b |f'(t)|\,\mathrm{d}t.$$

Theorem 4.5 (Coarea Formula)**.** *Let $f : \mathbb{R}^N \to \mathbb{R}^M$ be Lipschitz and $g : \mathbb{R}^N \to \mathbb{R}$ locally integrable, $N \geq M$. Then*

$$\int_{\mathbb{R}^N} g(x) Jf(x)\,\mathrm{d}x = \int_{\mathbb{R}^M} \left(\int_{f^{-1}[\{y\}]} g(x)\,\mathrm{d}\mathcal{H}^{N-M}(x) \right) \mathrm{d}y$$

Example 4.6 (Polar Coordinates). Setting $f(x) = \|x - x_0\|$ gives us the regular formula for polar coordinates:

$$\int_{\mathbb{R}^N} g\,\mathrm{d}x = \int_0^\infty \left(\int_{\partial B_N(x_0,r)} g\,\mathrm{d}\mathcal{H}^{N-1} \right) \mathrm{d}r$$

This agrees with our normal polar coordinate formulae for $N = 2, 3$. The Coarea Formula also generalises Fubini's Theorem.

Hausdorff measures can be applied further. We have not even begun to investigate the case of $\mathcal{H}^s$ where s is not an integer. It turns out that for any $s \in [0, N]$, there exists a set $A \subseteq \mathbb{R}^N$ with Hausdorff dimension s. (The Hausdorff dimension is defined as the point where lower values give an infinite Hausdorff measure and higher values give a zero Hausdorff measure; see Proposition 1.3 parts (iii) and (iv).)

Hausdorff measures lead naturally into a study of fractals, as these often do not have integral Hausdorff dimension. For example, the Cantor set and Sierpinski sponge have Hausdorff dimensions of $\frac{log2}{log3}$ and $\frac{log20}{log3}$ respectively. Using Hausdorff measures, we can sensibly assign a useful measure to any subset of $\mathbb{R}^N$.

References

[1] L. C. Evans & R. F. Gariepy: *Measure Theory & Fine Properties of Functions*, CRC PressINC, 1992, ISBN 978-0849371578

[2] R. J. Gardner: *The Brunn-Minkowski Inequality*, Bull. Amer. Math. Soc. (N.S.), 39 (3), 2002, pp. 355-405

[3] J.Yeh: *Real Analysis: Theory of Measure and Integration*, World Scientific, 2, 2006, ISBN 978-9812566546