GMT Notes

Paul Tee

Sp/Sum 23

Follows [Mag12] unless otherwise stated. Occasionally we will switch to [Sim14].

1 Lebesgue equals Hausdorff (3.9.23)

We follow (Thm 2.9, [Sim14]). Recall the definitions of Lebesgue and Hausdorff measures.

$$\mathcal{L}^n(A) = |A| := \inf\{\sum_{j=1}^{\infty} |R_j| : R_j \text{ is a open rectangle}, A \subset \cup_{j=1}^{\infty} R_j\},$$

$$\mathcal{H}^n_\delta(A) := \inf\{\sum_{j=1}^{\infty} \omega_n \left(\frac{\text{diam} C_j}{2}\right)^n : \text{diam}(C_j) < \delta, A \subset \cup_{j=1}^{\infty} C_j\},$$

and $\mathcal{H}^n(A) := \lim_{\delta \searrow 0} \mathcal{H}^n_\delta(A)$.

Theorem 1.1. *For every $\delta > 0$ and $A \subset \mathbb{R}^n$,*

$$\mathcal{H}^n(A) = \mathcal{H}^n_\delta(A) = |A|.$$

Proof. Taking $\delta \searrow 0$, it suffices to show $\mathcal{H}^n_\delta = \mathcal{L}^n$. Since taking closure does not affect diameter, we may assume C_k are closed. Note also $|E| = 0 \implies \mathcal{H}^n_\delta(E) = 0$ since we may inscribe a ball B_j with diameter $< \delta$ in each rectangle R_j.

($\mathcal{H}^n_\delta \leq \mathcal{L}^n$) The idea here is to "uniformly eat up" all the measure A with finitely many pairwise disjoint balls, then iterate this algorithm *ad infinitum.*[1] Fix an arbitrary cover $\{R_j\}$ of A by open rectangles.

Step $k = 1$. For each j, consider a disjoint family of cubes $\{I_k\}$[2] with $\text{diam}(I_k) < \delta$ such that the interiors of I_k are pairwise disjoint and $\cup I_k = R_j$. This is possible since R_j is an open set (Thm 1.4, [SS05]). For each k, inscribe a *closed* ball B_k into I_k such that $\text{diam}(B_k) > \frac{1}{2} s_k$ where s_k is the side length of I_k [Figure 1]. This is obviously not sharp, but it doesn't matter. What does matter is that the most measure we've left off out is

$$|I_k - B_k| \leq s_k^n - \omega_n(\frac{s_k}{4})^n = (1 - \frac{\omega_n}{4^n}) s^k.$$

[1] i.e. Step 1 leaves $\frac{1}{2}$ the measure, step 2 leaves $\frac{1}{4}$ the measure, etc (these are not the correct constants, but the idea is right).

[2] I_k depends on j also, which we suppress from the notation. Note that I_k^j may intersect $I_k^{j'}$ for $j \neq j'$ - the pairwise disjoint condition is with respect to a fixed rectangle.

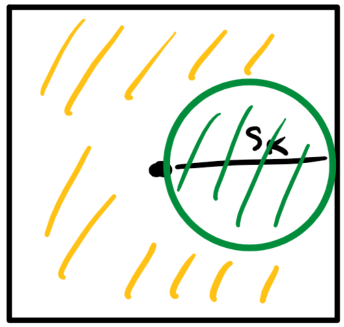

Figure 1: B_k must be bigger than the green ball.

Note the constant in front of s_k is uniform in k and strictly smaller than 1. Therefore,

$$|R_j - \cup_k B_k| = |\cup_k I_k - B_k| \leq (1 - \frac{\omega_n}{4^n}) \sum_k s^k = (1 - \frac{\omega_n}{4^n})|R_j|.$$

Thus, for each j, we may choose finitely many ball $B_j^1, ..., B_j^N$ such that the same inequality holds.

<u>Step $k \geq 2$.</u> For $k = 2$, we repeat the above argument but for $R_j - \cup_{k=1}^N B_j^k$ instead of R_j, which is *again an open set*. Repeat this construction for $k > 2$.

In total, for each j we get a countable disjoint collection of balls $\{B_j^k\}$ of R_j with radius $< \delta$ such that $|R_j - \cup_k B_j^k| = 0$, so $\mathcal{H}_\delta^n(R_j - \cup_k B_j^k) = 0$. Therefore,

$$\mathcal{H}_\delta^n(R_j) = \mathcal{H}_\delta^n(\cup_k B_j^k) \leq \sum_{k=1}^\infty \omega_n \left(\frac{\text{diam} B_j^k}{2} \right)^n = \sum_{k=1}^\infty |B_j^k| = |R_j|.$$

Thus

$$\mathcal{H}_\delta^n(A) \leq \sum_{j=1}^\infty \mathcal{H}_\delta^n(R_j) \leq \sum_{j=1}^\infty |R_j|,$$

and since R_j was an arbitrary cover, the result is shown.

($\mathcal{H}_\delta^n \geq \mathcal{L}^n$) The idea here is to use *Steiner symmetrization* to prove the isodiameteric inequality.

Lemma 1.2. *For any $A \subset \mathbb{R}^n$,*

$$|A| \leq \omega_n \left(\frac{diam A}{2} \right)^n.$$

Assuming the lemma, take C_j an arbitrary collection of sets which cover A, with $\text{diam} C_j < \delta$. Then,

$$|A| \leq |\cup_j C_j| \leq \sum_j |C_j| \leq \sum_j \omega_n \left(\frac{\text{diam} C_j}{2} \right)^n.$$

(of Lemma 1.2). It suffices to prove A compact since taking closure does not increase diameter. We sketch a symmetrization process as follows. For the hyperplane $H_i := \{x^i = 0\}$ for $i = 1, ..., n$, let $\xi_i \in H_i$ parameterize A in the sense that the fibers (under orthogonal projection to H_i) of A are at most 1-dimensional. We record the Lebesgue measure (length) of the fibers, the symmetrically distribute the length across ξ^i. This symmetrizes A across the hyperplane H_i, is diameter non-increasing, and the Lebesgue measure of A is constant. Furthermore, this preserves symmterizations across other hyperplanes. Thus, symmetrizing A across each hyperplane H_i produces a subset of the ball of radius $\frac{\operatorname{diam} A}{2}$, proving the lemma. ■

Both inequalities are proven, and so the result follows. ■

2 Riesz Representation Theorem II (4.6.23)

We state and prove RRT.

Theorem 2.1. *If L is a bounded linear functional on $C_c(\mathbb{R}^n, \mathbb{R}^m)$, then its variation $|L|$ is a (scaler-valued) Radon measure on $\mathbb{R}^n$ and there exists a $|L|$-measurable function $g : \mathbb{R}^n \to \mathbb{R}^m$ such that $g = 1$ $|L|$-a.e., and for all $\phi \in C_c(\mathbb{R}^n, \mathbb{R}^m)$*

$$\langle L, \phi \rangle = \int_{\mathbb{R}^n} (\phi \cdot g) d|L|.$$

Moreover,

$$|L|(A) = \sup\{\int_{\mathbb{R}^n} \phi \cdot g d|L| : \phi \in C_c(A, \mathbb{R}^m), |\phi| \leq 1\}.$$

Proof. Recall the definition of the total variation (measure) of L.

$$|L| :=$$

Zack had previously shown that $|L|$ is Radon, and that a version of RRT holds for L^1. ■

3 Compactness and Regularization (4.27.23)

Maggi proves this by intertwining the functional analysis with the geometric measure theory. I think it's more clear to separate them out.

Let $(V, \|\cdot\|)$ be a normed linear space, and $B \subset V$ the unit ball defined by $\|\cdot\|$. Recall the following basic fact about the strong topology

Proposition 3.1. *A normed linear space V is finite dimensional iff the unit ball is compact.*

The proof is easy but irrelevant. What matters is the moral of the story - if we're out looking for compact sets in infinite dimensional space, the strong topology is not the correct one to look in. Even the most basic candidate for a compact set is not compact, so there are way too many open sets. We need a coarser topology.

Look not in V, but in its dual V^* together with its unit ball B^* defined by the operator norm. We see

$$B^* := \{\mu : V \to \mathbb{R} : |\mu(\phi)| \leq \|\phi\|\} = \{L : V \to \mathbb{R} : |\mu(\phi)| \leq 1 \text{ for } \|\phi\| \leq 1\}.$$

In other words, $\mu : B \to [-1,1]$ and so $\mu \in [-1,1]^B$; in other words, B^* canonically sits inside $[-1,1]^B$, so it is compact in the subspace topology by Tychnoff's. However, the subspace topology agrees does *not* agree with the topology induced by the operator norm. The subspace topology oversees only that the evaluation pairing between V, V^* is continuous, and so it agrees with the weak star topology. That is, a sequence $\{\mu_i\}$ in V^* converges weak-$*$ to μ iff for every $\phi \in V$,

$$\mu_i \overset{*}{\rightharpoonup} \mu \iff \langle \phi, \mu_i \rangle \to \langle \phi, \mu \rangle,$$

where $\langle -, - \rangle$ denotes the evaluation pairing. This is known as *Banach-Alaoglu* and will serve as the backbone functional analysis tool for the compactness result for the space of Radon measures (which by RRT is dual to $V = C_c$ with a mildly strange topology). Truthfully, it's important that we have the sequential version of Banach-Alaoglu - this is true if V is *separable* (counterexamples exist otherwise) which is true in our case.

Theorem 3.2. *Given a sequence $\{\mu_i\}$ of Radon measures which are locally uniformly bounded, i.e. for $K \subset \mathbb{R}^n$ compact*

$$\sup_i \mu_i(K) < \infty,$$

then there exists a subsequence, upon reindexing, $\{\mu_k\}$ which weak-$*$ converge to a Radon measure. That is, by RRT,

$$\int_K \phi d\mu_i \to \int_K \phi d\mu,$$

for each $\phi \in C_c(\mathbb{R}^n)$.

Proof. The idea is basically sequential Banach-Alaoglu and RRT. We first give a construction of $\{\mu_k\}$ via a diagonalization procedure. Consider an exhaustion of $\mathbb{R}^n$ by balls $\{B_j\}$ centered at 0. Define functionals

$$F_{i,j}(\phi) := \int_{B_j} \phi d\mu_i,$$

for $\phi \in C_c(B_j)$. Linearity of $F_{i,j}$ is clear, and boundedness follows from the local uniformity assumption as

$$|F_{i,j}(\phi)| \leq \sup \phi \cdot \mu_i(B_j) \leq C \sup \phi,$$

where $C = C(j)$. By sequential Banach-Alaoglu, there is a functional $F_j : C_c(B_j) \to \mathbb{R}$ and a subsequence $\{F_{i',j}\}$ such that $F_{i',j} \overset{*}{\rightharpoonup} F_j$. By extracting subsequences subsequently, then selecting the diagonal subsequence, we extract the desired $\{\mu_k\}$ upon relabeling. We claim that $F = \lim_k \mu_k$. By construction, this limit is well-defined. Furthermore, by RRT we have

$$F(\phi) := \int_{\mathbb{R}^n} \phi d\mu,$$

for some Radon measure μ. For $\phi \in C_c(\mathbb{R}^n)$, take $\text{spt}\phi \subset B_j$. Linearity follows since $F(\phi) = F_j(\phi)$, and F_j is linear. Boundedness follows from

$$|F(\phi)| \leq \sup \phi \cdot \mu(B_j) \leq C \sup \phi.$$

Here $C = C(j)$; nonetheless the weird topology we put on $C_C(\mathbb{R}^n)$ allows us to consider F as a bounded functional. ■

A similar statement holds for $\mathbb{R}^m$-valued measures by applying the above to the positive and negative part of each component to extract an $\mathbb{R}^m$-valued measure.

We very briefly discuss regularization. The basic theory of regularization of functions extends to measures as one expects. Given a Radon measure μ, define the function $\mu_\epsilon := \rho_\epsilon * \mu$ as

$$\mu_\epsilon(x) := \int_{\mathbb{R}^n} \rho_\epsilon(x-y)d\mu(y),$$

where ρ is a *regularization kernel.* One then makes this a measure by considering it as a density wrt the Lebesgue measure $\mu_\epsilon(x)dx$. The basic proposition that holds here is the following.

Proposition 3.3. *With the above notation, and $B_\epsilon(E)$ the ϵ-ball (neighborhood) of E,*

1. $\mu_\epsilon \overset{*}{\rightharpoonup} \mu$,
2. $|\mu_\epsilon| \overset{*}{\rightharpoonup} |\mu|$,
3. $|\mu_\epsilon|(E) \leq \mu(B_\epsilon(E))$.

Proof. The one word proof of 1 and 3 is Fubini's. 2 is more complicated, requiring an exercise we skipped. ■

4 Area Formula I (6.1.23)

For $f : \mathbb{R}^n \to \mathbb{R}^m$ an *injective, Lipschitz* function (throughout the section, $n \leq m$), set the *Jacobian* of f to be

$$Jf(x) := \begin{cases} \sqrt{\det(\nabla f(x)^* \nabla f(x))} & f \text{ is differentiable} \\ \infty & f \text{ is not differentiable.} \end{cases}$$

By Rademacher's theorem, Jf is integrable, and the area formula (in the case without considering multiplicity) calculates for Lebesgue measurable E the measures of images in terms of Jf as

$$\mathcal{H}^n(f(E)) = \int_E Jf(x)dx. \tag{1}$$

There are two obvious requirements that need to hold for the above formula to be true. First, if $E := \{x \in \mathbb{R}^n : Jf(x) = 0\}$, then $\mathcal{H}^n(f(E)) = 0$ (injectivity is dropped for this statement). This is the content of Proposition 8.7. The second is that under the above assumptions, $f(E)$ had better be $\mathcal{H}^n$ measurable.

Proposition 4.1. *If E is Lebesgue measurable in $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^m$ is an injective Lipschitz function, then $f(E)$ is $\mathcal{H}^n$ measurable in $\mathbb{R}^m$.*

Proof. Assume E is bounded, and exhaust E by a sequence of compact sets $\{K_i\}$ wrt $\mathcal{L}^n$. Since f is Lipschitz, it's continuous and so $f(K_i)$ is compact and $\cup f(K_i)$ is Borel. Then,

$$\mathcal{H}^n(f(E) - \cup f(K_i)) = \mathcal{H}^n(f(E - \cup f(K_i)) \leq \mathrm{Lip} f^n \cdot \mathcal{H}^n(\cup K_i) = \mathrm{Lip} f^n \cdot |\cup K_i| = 0.$$

■

Here, we used that Lipschitz functions play nicely with the Hausdorff measure in the sense that $\mathcal{H}^s(f(E)) \leq \mathrm{Lip} f^s \mathcal{H}^s(E)$, and is equal to the Lebesgue measure. For future reference, recall also that Hausdorff measure plays nicely with scaling in that $\mathcal{H}^s(rE) = r^s \mathcal{H}^s(E)$.

The next statement is to prove that area formula holds on arbitrary (not necessarily injective) linear functions $T : \mathbb{R}^n \to \mathbb{R}^m$.

Proposition 4.2. *If $T : \mathbb{R}^n \to \mathbb{R}^m$ linear, then for every $E \subset \mathbb{R}^n$,*

$$\mathcal{H}^n(T(E)) = JT \cdot |E|.$$

For this, we must first recall some linear algebra. Set $\mathrm{Lin}(n, m) := \{T : \mathbb{R}^n \to \mathbb{R}^m : T \text{ linear}\}$ and $\mathrm{Isom}(n, m) : \{P \in \mathrm{Lin}(n, m) : \langle Px, Py \rangle = \langle x, y \rangle\}$. We think of $\mathrm{Isom}(n, m)$ as the set of linear isometric embeddings - clearly all must be injective.

Lemma 4.3. *For $P \in Isom(\mathbb{R}^n, \mathbb{R}^m)$, we have $P^*P = id_n$.*

Proof. Recall $P^* := (\mathbb{R}^m)^* \to (\mathbb{R}^n)^*$ is precomposition with P. The composition P^*P is understood as the following isomorphism:

$$\mathbb{R}^n \xrightarrow{P} \mathbb{R}^m \to (\mathbb{R}^m)^* \xrightarrow{P^*} (\mathbb{R}^n)^* \to \mathbb{R}^n$$

$$x \mapsto Px \mapsto \langle Px, - \rangle \mapsto \langle Px, P- \rangle \mapsto x$$

where the last map is justified since the covector $y \mapsto \langle Px, Py \rangle = \langle x, y \rangle$ is metrically equivalent to the vector $y \in \mathbb{R}^n$. ■

Since $P \in \mathrm{Isom}(n, m)$ is an isometry, both P and its adjoint P^* (orthogonal projections) have Lipschitz constant 1. Recall also the spectral theorem - every symmetric real-valued matrix diagonalizes with real eigenvalues (and orthogonal eigenspaces) and the polar decomposition - every linear map $T : \mathbb{R}^n \to \mathbb{R}^m$ can be realized as

$$T = PS$$

with $S \in \mathrm{Lin}(n, n)$ symmetric and $P \in \mathrm{Isom}(n, m)$ (both S, T are given explicitly in terms of the diagonalization of T). It turns out linear isometric embeddings do not change the fine structure of the geometry, and in a sense we'll make precise, all the change in the geometry happens with S.

(of 4.2.) Set $\kappa := D_{\mathcal{L}^n} \nu = \frac{\mathcal{H}^n(T(B))}{|B|}$ for $\nu(E) := \mathcal{H}^n(T(E))$ to be the density. The last equality is justified by scaling since for all $r > 0$,

$$\mathcal{H}^n(T(rB)) = r^n \mathcal{H}^n(T(B))$$

$$|rB| = r^n |B|.$$

It suffices to show the following two statements:

1. $\mathcal{H}^n(T(E)) = \kappa|E|$,
2. $JT = \kappa$.

First assume $\kappa = 0$ (T is non-injective). Then, $\mathcal{H}^n(T(B)) = 0 \implies \mathcal{H}^n(T(rB)) = 0$ for every $r > 0$ by linearity of T and so in the limit, $\mathcal{H}^n(\mathbb{R}^n) = 0$. By monotonicity of the measure, $\mathcal{H}^n(T(E)) = 0$ for every E.

Next, assume $\kappa > 0$ (T is injective). Since $E \mapsto \mathcal{H}^n(T(E))$ is Radon, by the decomposition theorem for measures, it suffices to show the equality is true on balls $B(x, r)$ for $r > 0$ and $x \in \mathbb{R}^n$. This follows by translation invariance and scaling, as

$$\begin{aligned}\mathcal{H}^n(T(B(x,r)) &= \mathcal{H}^n(T(x + rB(0,1))) \\ &= r^n\mathcal{H}^n(T(B)) \\ &= r^n\kappa|B| \\ &= \kappa|B(x,r)|,\end{aligned}$$

which shows $\mathcal{H}^n(T(-)) \ll \mathcal{L}^n$ and identifies κ as the density.

Remark 4.4. For linear functions, non-injectivity is very easy to deal with as both sides evaluates to 0. We don't need to consider multiplicity for this case.

To show $JT = \kappa$, we first show for $T = PS$ and $S(Q) = E$ for a cube Q, we have $\frac{\mathcal{H}^n(P(E))}{|E|} = 1$, since

$$|E| = |P^*P(E)| \leq Lip(P^*)^n|P(E)| = |P(E)| = \mathcal{H}^n(P(E)) \leq Lip(P)^n|E| = |E|.$$

Therefore,

$$\kappa = \frac{\mathcal{H}^n(PS(Q))}{|Q|} = \frac{\mathcal{H}^n(P(E))}{|E|}\frac{|S(Q)|}{|Q|} = \frac{|S(Q)|}{|Q|}.$$

Note we can switch to other sets when defining density (Thm 3.22, [Fol13]). By homogeneity, we take Q to be the unit cube, so we must prove $JT = |S(Q)|$. Consider now the spectral decomposition of $S = \sum_i \lambda_i v_i$. Using the fact that $\nabla T = T$ as T is linear, pullback is contravariant, and

$$\langle S^*Sv_i, v_j\rangle = \langle Sv_i, Sv_j\rangle = \begin{cases}\lambda_i^2 & i = j \\ 0 & i \neq j,\end{cases}$$

we compute

$$JT = \sqrt{\det T^*T} = \sqrt{\det(S^*P^*PS)} = \sqrt{\det(S^*S)} = \sqrt{\prod_i |\lambda_i|^2} = |\det(S)| = |S(Q)|.$$

■

The next proof shows that the area formula does not see the *singular set* $\mathcal{S} := \{Jf = 0\}$. For notation, balls without mention to the dimension will be assumed to be dimension n, and without mention to the center will be assumed to be centered at the origin.

Proposition 4.5. *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz, then on the singular set $\mathcal{S}$,*

$$\mathcal{H}^n(f(\mathcal{S})) = 0.$$

Proof. We will deduce $\mathcal{H}^n_\infty(f(\mathcal{S} \cap B_R)) = 0$ for arbitrary $R > 0$ (recall $\mathcal{H}^n \ll \mathcal{H}^n_\infty$). For $x \in \mathcal{S}$ and $1 > \epsilon > 0$,[3] by definition $\nabla f(x)$ is the linear map which satisfies the inequality

$$|f(x+v) - f(x) - \nabla f(x)v| < \epsilon|v|,$$

for $|v| < r(\epsilon, x)$. We can reinterpret the above (think of $f(x) = 0$.) in terms of small balls[4]: for $r = |v| < r(x, \epsilon)$, we have

$$f(B_r(x)) \subset f(x) + B_{\epsilon r}(\nabla f(x)(B_r)).$$

It would serve us well, therefore, to find a bound on the (translation-invariant) size of $B_{\epsilon r}(\nabla f(x)(B_r))$

Lemma 4.6. *For $0 < \epsilon < 1$ and $C = C(n, Lipf)$,*

$$\mathcal{H}^n_\infty(B_{\epsilon r}(\nabla f(x)(B_r))) \leq Cr^n\epsilon.$$

The proof will use that $\mathcal{S}$ is singular by using a dimension drop argument, which produces the ϵ. Assuming this lemma, we get for $x \in \mathcal{S}$,

$$\mathcal{H}^n_\infty(f(B_r(x))) \leq Cr^n\epsilon. \tag{2}$$

We take the family $\mathcal{F}$ of balls with centers in $\mathcal{S} \cap B_R$,

$$\mathcal{F} := \{B_r(x) \subset \mathbb{R}^n : x \in \mathcal{S} \cap B_R, 0 < r < r(\epsilon, x)\}.$$

We cutoff $\mathcal{S}$ by B_R to have bounded centers; we may therefore apply Besicovitch covering theorem to extract subfamilies $\mathcal{F}_1, ..., \mathcal{F}_{\xi(n)}$ such that each $\mathcal{F}_i$ is countable, pairwise disjoint

[3] As far as I can tell, the restriction to $1 > \epsilon$ is for polish.

[4] Compare this to the definition of continuity in terms of balls. f is continuous at x if for $r < r(\epsilon, x)$,

$$f(B_r(x)) \subset f(x) + B_\epsilon.$$

So if you can take a derivative, the information you gain is having a modulus of control over the error tolerance ϵr in terms on your parameter r.

and $\mathcal{S} \cap B_R \subset \cup_{\xi(n)} \mathcal{F}_i$. By inequality 2,

$$\begin{aligned}
\mathcal{H}^n_\infty(f(\mathcal{S} \cap B_R)) &\leq \sum_{i=1}^{\xi(n)} \sum_{B_r(x) \in \mathcal{F}_i} \mathcal{H}^n_\infty(f(B_r(x)) \\
&\leq C\omega_n \epsilon \sum_{i=1}^{\xi(n)} \sum_{B_r(x) \in \mathcal{F}_i} r^n \\
&\leq \frac{C\epsilon}{\omega_n} \sum_{i=1}^{\xi(n)} \sum_{B_r(x) \in \mathcal{F}_i} |B_r(x)| \\
&= \frac{C\epsilon}{\omega_n} \sum_{i=1}^{\xi(n)} \left| \bigcup_{B_r(x) \in \mathcal{F}_i} B_r(x) \right| \\
&\leq \frac{C\epsilon\xi(n)}{\omega_n} |B_1(\mathcal{S} \cap B_R)|.
\end{aligned}$$

Therefore, it suffices to prove the lemma.

(of Lemma 4.6). We first identify that $\nabla f(x)(B_r) \subset B^m_{r\mathrm{Lip}f} \cap \nabla f(x)(\mathbb{R}^n)$, where the right hand side is a disk $D_{r\mathrm{Lip}f}$ of dimension $k < n$ since $x \in \mathcal{S}$. This follows from homogeneity and $\|\nabla f\| \leq \mathrm{Lip}f$, since we may characterize $\mathrm{Lip}f$ as

$$\mathrm{Lip}f = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}.$$

Therefore, we must obtain a bound of the form

$$\mathcal{H}^n_\infty(B_\epsilon(D_s)) \leq C(n, s)\epsilon, \tag{3}$$

since since taking a neighborhood of a disk is linear,

$$\begin{aligned}
\mathcal{H}^n_\infty(B_{\epsilon r}(\nabla f(x)(B_r))) &\leq \mathcal{H}^n_\infty(B_{\epsilon r}(B^m_{r\mathrm{Lip}f} \cap \nabla f(x)(\mathbb{R}^n))) \\
&\leq r^n \mathcal{H}^n_\infty(B_\epsilon(B^m_{\mathrm{Lip}f} \cap \nabla f(x)(\mathbb{R}^n))) \\
&\leq r^n C(n, \mathrm{Lip}f))\epsilon.
\end{aligned}$$

To obtain inequality 3, we produce a covering of $B_\epsilon(D_s) \subset \mathbb{R}^k \times \mathbb{R}^{m-k}$ by translates of

$$B^k_{\epsilon s} \times B^{m-k}_\epsilon \subset \mathbb{R}^k \times \mathbb{R}^{m-k}.$$

Note that we need $C\epsilon^{-k}$ number of translates to cover $B_\epsilon(D_s) \approx B^k_s \times B^{m-k}_\epsilon$ for small ϵ, since area of B^k_s grows like s^k.[5] By Pythagorean theorem,

$$\mathrm{diam}(B^k_{\epsilon s} \times B^{m-k}_\epsilon)^2 = \mathrm{diam}(B^k_{\epsilon s})^2 + \mathrm{diam}(B^{m-k}_\epsilon)^2 = 4\epsilon^2(1 + s^2).$$

[5] e.g. For 2-dimensional disks, need (up to a dimensional constant) one-hundred small disks $D_{1/10}$ to cover one big disk D_1.

Since $k < n$, we have

$$\begin{aligned}
\mathcal{H}^n_\infty(B_\epsilon(D_s)) &\leq \omega_n \sum_{\#\text{ translates}} \left(\frac{\operatorname{diam}(B^k_{\epsilon s} \times B^{m-k}_\epsilon)}{2} \right)^n \\
&= \omega_n C \epsilon^{-k} (2\epsilon^2(1+s^2))^{\frac{n}{2}} \\
&= C(n,s)\epsilon^{n-k} \\
&\leq C\epsilon.
\end{aligned}$$

$\blacksquare$

Therefore, the image of the singular set is too small to be detected by $\mathcal{H}^n$. $\blacksquare$

References

[Fol13] G.B. Folland. *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2013. ISBN: 9781118626399.

[Mag12] Francesco Maggi. *Sets of Finite Perimeter and Geometric Variational Problems: An Introduction to Geometric Measure Theory*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2012. DOI: 10.1017/CBO9781139108133.

[Sim14] L. Simon. *Introduction to Geometric Measure Theory*. https://web.stanford.edu/class/math285/ts-gmt.pdf. 2014.

[SS05] Elias M. Stein and Rami Shakarchi. *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*. Princeton University Press, 2005. ISBN: 9780691113869.