Hausdorff measure and surface measure

Definition 1 (Projections and graphs). *Let $d \geq 2$.*

- *For every set $E \subset \mathbb{R}^d$ we denote by $\pi(E)$ the projection of E on $\mathbb{R}^{d-1}$, that is*
$$\pi(E) := \Big\{x' \in \mathbb{R}^{d-1} \ : \ (x', x_d) \in E \text{ for some } x_d \in \mathbb{R}\Big\}.$$
- *Given a set $A \subset \mathbb{R}^{d-1}$ and a function $f : A \to \mathbb{R}$, we denote by $\Gamma(f; A)$ the graph of f over A, that is,*
$$\Gamma(f; A) := \Big\{(x, f(x)) \ : \ x \in A\Big\}.$$

Projections and Hausdorff measure

Lemma 2. *Let $d \geq 2$. Then, for every set $U \subset \mathbb{R}^d$, we have*
$$diam(\pi(U)) \leq diam(U).$$

Proof. Suppose that $x', y' \in \pi(U)$. Let $x_d, y_d \in \mathbb{R}$ be such that
$$x := (x', x_d) \in U \quad \text{and} \quad y := (y', y_d) \in U.$$
Then,
$$|x' - y'| \leq \sqrt{|x' - y'|^2 + (x_d - y_d)^2} = |x - y| \leq \text{diam}(U).$$
Taking the supremum over all $x', y' \in \pi(U)$, we get the claim. □

Proposition 3. *Let $d \geq 2$. Then, for every set $E \subset \mathbb{R}^d$, we have*
$$\mathcal{H}^{d-1}(\pi(E)) \leq \mathcal{H}^{d-1}(E).$$

Proof. Suppose that
$$\mathcal{H}^{d-1}(E) = \sup_{\delta>0} \mathcal{H}^{d-1}_\delta(E) < +\infty.$$
Fix $\varepsilon > 0$.

By the definition of $\mathcal{H}^{d-1}$, for every $\delta > 0$, we can find a covering $\{U_i\}_{i\geq 1}$ of E with sets $U_i \subset \mathbb{R}^{d-1}$ such that
$$\text{diam}(U_i) < \delta \quad \text{for all} \quad i \geq 1,$$
and
$$\omega_{d-1} \sum_{i\geq 1} \left(\frac{\text{diam}(U_i)}{2}\right)^{d-1} \leq \mathcal{H}^{d-1}(E) + \varepsilon.$$
The family of projections $\{\pi(U_i)\}_{i\geq 1}$ is a covering of $\pi(E)$ and is such that
$$\text{diam}(\pi(U_i)) < \delta \quad \text{for all} \quad \geq 1.$$
Then, by the definition of $\mathcal{H}^{d-1}_\delta$, we have
$$\begin{aligned}\mathcal{H}^{d-1}_\delta(\pi(E)) &\leq \omega_{d-1} \sum_{i\geq 1} \left(\frac{\text{diam}(\pi(U_i))}{2}\right)^{d-1} \\ &\leq \omega_{d-1} \sum_{i\geq 1} \left(\frac{\text{diam}(U_i)}{2}\right)^{d-1} \leq \mathcal{H}^{d-1}(E) + \varepsilon.\end{aligned}$$
Taking the supremum over $\delta > 0$ we have
$$\mathcal{H}^{d-1}(\pi(E)) = \sup_{\delta>0} \mathcal{H}^{d-1}_\delta(\pi(E)) \leq \mathcal{H}^{d-1}(E) + \varepsilon.$$
Since $\varepsilon > 0$ is arbitrary, we get the claim. □

More in general, we have the following

Proposition 4. *Let $d \geq 1$, $K \subset \mathbb{R}^d$ be a closed convex set, and $\pi_K : \mathbb{R}^d \to K$ be the projection on K. Then:*
$$diam(\pi_K(U)) \leq diam(U) \quad \textit{for every set} \quad U \subset \mathbb{R}^d;$$
$$\mathcal{H}^s(\pi_K(E)) \leq \mathcal{H}^s(E) \quad \textit{for every set} \quad E \subset \mathbb{R}^d \quad \textit{and every} \quad s \in [0, d].$$

HAUSDORFF MEASURE AND LIPSCHITZ MAPS

Lemma 5. *Let $d \geq 2$. Let $A \subset \mathbb{R}^d$ and let*
$$\Phi : A \to \mathbb{R}^d$$
be an L-Lipschitz map:
$$|\Phi(x) - \Phi(y)| \leq L|x - y| \quad \textit{for all} \quad x, y \in A.$$
Then, for all $s > 0$ we have
$$\mathcal{H}^s(\Phi(A)) \leq L^s \mathcal{H}^s(A).$$

Proof. Suppose that $\mathcal{H}^s(A) < +\infty$. Since Fix $\delta > 0$ and $\varepsilon > 0$. By the definition of $\mathcal{H}^s_\delta(A)$ we can find a covering of A with sets $\{U_i\}_{i\geq 1}$ such that:
$$\mathrm{diam}(U_i) \leq \delta \quad \text{for all} \quad i \geq 1,$$
and
$$\mathcal{H}^s_\delta(A) \leq \sum_{i\geq 1} \omega_s \left(\frac{\mathrm{diam}(U_i)}{2}\right)^s \leq \mathcal{H}^s_\delta(A) + \varepsilon.$$
Consider the family of sets $\widetilde{U}_i := \Phi(U_i)$, for $i \geq 1$. We have that:

- $\{\widetilde{U}_i\}_{i\geq 1}$ is a covering of $\Phi(A)$;
- $\mathrm{diam}(\widetilde{U}_i) \leq L\delta$ for all $i \geq 1$.

Then, we can use the family $\{\widetilde{U}_i\}_{i\geq 1}$ to estimate $\mathcal{H}^s_{L\delta}(\Phi(A))$. Thus,
$$\mathcal{H}^s_{L\delta}(\Phi(A)) \leq \sum_{i\geq 1} \omega_s \left(\frac{\mathrm{diam}(\widetilde{U}_i)}{2}\right)^s \leq L^s \sum_{i\geq 1} \omega_s \left(\frac{\mathrm{diam}(U_i)}{2}\right)^s \leq L^s \Big(\mathcal{H}^s_\delta(A) + \varepsilon\Big).$$
Since
$$\mathcal{H}^s(A) = \sup_{\delta>0} \mathcal{H}^s_\delta(A) = \lim_{\delta\to 0} \mathcal{H}^s_\delta(A),$$
we have that
$$\mathcal{H}^s_\delta(A) \leq \mathcal{H}^s(A),$$
so we get
$$\mathcal{H}^s_{L\delta}(\Phi(A)) \leq L^s \Big(\mathcal{H}^s(A) + \varepsilon\Big).$$
Taking the limit as $\delta \to 0$, we get
$$\mathcal{H}^s(\Phi(A)) \leq L^s \Big(\mathcal{H}^s(A) + \varepsilon\Big).$$
Finally, since $\varepsilon > 0$ is arbitrary, we get
$$\mathcal{H}^s(\Phi(A)) \leq L^s \mathcal{H}^s(A).$$
□

LIPSCHITZ GRAPHS AND HAUSDORFF MEASURE.
PART I. A ROUGH ESTIMATE ON THE HAUSDORFF MEASURE OF A LIPSCHITZ GRAPHS.

Lemma 6. *Let $d \geq 2$ and let $f : \mathbb{R}^{d-1} \to \mathbb{R}$ be an L-Lipschitz function, for some $L > 0$. Then:*
$$\mathit{diam}(\Gamma(f;U)) \leq \sqrt{1 + L^2}\ \mathit{diam}(U) \quad \textit{for all sets} \quad U \subset \mathbb{R}^{d-1}.$$

Proof. Let $x = (x', f(x')) \in \Gamma(f;U)$ and $y = (y', f(y')) \in \Gamma(f;U)$. Then, since
$$|f(x') - f(y')| \leq L|x' - y'|,$$
we get that
$$\begin{aligned} |x - y| &= \sqrt{|x' - y'|^2 + (f(x') - f(y'))^2} \\ &\leq \sqrt{1 + L^2}\ |x' - y'| \\ &\leq \sqrt{1 + L^2}\ \mathrm{diam}(U). \end{aligned}$$
Taking the supremum over all $x, y \in \Gamma(f;U)$ we get the claim. □

Proposition 7. *Let $d \geq 2$ and let $f : \mathbb{R}^{d-1} \to \mathbb{R}$ be an L-Lipschitz function, for some $L > 0$. Then, for every set $A \subset \mathbb{R}^{d-1}$ we have*

$$\mathcal{H}^{d-1}(\Gamma(f;A)) \leq \big(1+L^2\big)^{\frac{d-1}{2}}\ \mathcal{H}^{d-1}(A),$$

where $\Gamma(f;A)$ is the graph of f over A. In particular, since $\mathcal{H}^{d-1} = \mathcal{L}^{d-1}$ on $\mathbb{R}^{d-1}$, for every Lebesgue measurable set $A \subset \mathbb{R}^{d-1}$, we have

$$\mathcal{H}^{d-1}(\Gamma(f;A)) \leq \big(1+L^2\big)^{\frac{d-1}{2}}\ \mathcal{L}^{d-1}(A).$$

Proof. Suppose that

$$\mathcal{H}^{d-1}(A) = \sup_{\delta>0} \mathcal{H}^{d-1}_\delta(A) < +\infty.$$

Then, for every $\delta > 0$, we can find a covering $\{U_i\}_{i\geq 1}$ of A with sets $U_i \subset \mathbb{R}^{d-1}$ such that

$$\operatorname{diam}(U_i) < \frac{\delta}{\sqrt{1+L^2}} \quad \text{for all} \quad i \geq 1,$$

and

$$\omega_{d-1} \sum_{i\geq 1} \left(\frac{\operatorname{diam}(U_i)}{2}\right)^{d-1} \leq \mathcal{H}^{d-1}(A) + \varepsilon.$$

We notice that $\Gamma(f;U_i)$, $i \geq 1$, is a covering of $\Gamma(f;A)$ and is such that:

$$\operatorname{diam}(\Gamma(f;U_i)) \leq \sqrt{1+L^2}\ \operatorname{diam}(U_i) < \sqrt{1+L^2}\ \frac{\delta}{\sqrt{1+L^2}} = \delta.$$

By the definition of $\mathcal{H}^{d-1}_\delta(\Gamma(f;A))$, we have

$$\begin{aligned}
\mathcal{H}^{d-1}_\delta(\Gamma(f;A)) &\leq \omega_{d-1} \sum_{i\geq 1} \left(\frac{\operatorname{diam}(\Gamma(f;U_i))}{2}\right)^{d-1} \\
&\leq \omega_{d-1}\big(1+L^2\big)^{\frac{d-1}{2}} \sum_{i\geq 1} \left(\frac{\operatorname{diam}(U_i)}{2}\right)^{d-1} \\
&\leq \big(1+L^2\big)^{\frac{d-1}{2}} \Big(\mathcal{H}^{d-1}(A) + \varepsilon\Big).
\end{aligned}$$

Taking the supremum in $\delta > 0$, we get

$$\mathcal{H}^{d-1}(\Gamma(f;A)) \leq \big(1+L^2\big)^{\frac{d-1}{2}} \Big(\mathcal{H}^{d-1}(A) + \varepsilon\Big).$$

Since $\varepsilon > 0$ is arbitrary, we get the claim. $\square$

Corollary 8. *Let $d \geq 2$ and let $f : \mathbb{R}^{d-1} \to \mathbb{R}$ be an L-Lipschitz function, for some $L > 0$. Then, for every Lebesgue measurable set $A \subset \mathbb{R}^{d-1}$, we have*

$$\mathcal{L}^{d-1}(A) = 0 \quad \Leftrightarrow \quad \mathcal{H}^{d-1}(\Gamma(f;A)) = 0.$$

Lipschitz graphs and Hausdorff measure.
Part II. The Hausdorff measure of the graph of a linear function.

Lemma 9. *Suppose that $d \geq 2$ and $s \geq 0$. Then, for every $A' \subset \mathbb{R}^{d-1}$, we have*

$$\mathcal{H}^s(A') = \mathcal{H}^s(A' \times \{0\}),$$

where:

- *$\mathcal{H}^s(A')$ is computed with the distance in $\mathbb{R}^{d-1}$ via coverings with sets in $\mathbb{R}^{d-1}$;*
- *$\mathcal{H}^s(A' \times \{0\})$, is calculated with the distance in $\mathbb{R}^d$ via coverings with subsets of $\mathbb{R}^d$.*

Proof. We proceed in two steps.

Step 1. We first prove the inequality $\mathcal{H}^s(A' \times \{0\}) \leq \mathcal{H}^s(A')$.
Let $\delta > 0$ and let $\{U_i'\}_{i\geq 1}$ be a family of subsets of $\mathbb{R}^{d-1}$ such that:

- $\operatorname{diam}(U_i') \leq \delta$ for every $i \geq 1$.
- A' is contained in the union of U_i', $\imath \geq 1$.

Then, the family of sets $U_i := U_i' \times \{0\}$ is a family of subsets of $\mathbb{R}^d$ such that:

- $\operatorname{diam}(U_i) \le \delta$ for every $i \ge 1$.
- $A' \times \{0\}$ is contained in the union of $U_i = U_i' \times \{0\}$, $\imath \ge 1$.

In particular, we have

$$\mathcal{H}^s_\delta(A' \times \{0\}) \le \sum_{i=1}^{+\infty} \omega_s \left(\frac{\operatorname{diam}(U_i' \times \{0\})}{2} \right)^s \le \sum_{i=1}^{+\infty} \omega_s \left(\frac{\operatorname{diam}(U_i')}{2} \right)^s .$$

Taking the infimum on the right, over all coverings $\{U_i'\}_{i\ge 1}$ of A' with sets of diameter at most δ, we get

$$\mathcal{H}^s_\delta(A' \times \{0\}) \le \mathcal{H}^s_\delta(A').$$

Passing to the limit as $\delta \to 0$, we get

$$\mathcal{H}^s(A' \times \{0\}) \le \mathcal{H}^s(A'),$$

which concludes the proof of the inequality in the first step.

Step 2. We next prove the inequality $\mathcal{H}^s(A') \le \mathcal{H}^s(A' \times \{0\})$.
Fix $\delta > 0$. Let $\{U_i\}_{i\ge 1}$ be a family of subsets of $\mathbb{R}^d$ such that:

- $\operatorname{diam}(U_i) \le \delta$ for every $i \ge 1$.
- A is contained in the union of U_i, $\imath \ge 1$.

Consider the family of projections $U_i' := \pi(U_i) \subset \mathbb{R}^{d-1}$, of the sets U_i on $\mathbb{R}^{d-1} = \mathbb{R}^{d-1} \times \{0\}$. Then:

- $\operatorname{diam}(U_i') \le \operatorname{diam}(U_i) \le \delta$ for every $i \ge 1$.
- A' is contained in the union of U_i', $\imath \ge 1$.

In particular, we have

$$\mathcal{H}^s_\delta(A') \le \sum_{i=1}^{+\infty} \omega_s \left(\frac{\operatorname{diam}(U_i')}{2} \right)^s \le \sum_{i=1}^{+\infty} \omega_s \left(\frac{\operatorname{diam}(U_i)}{2} \right)^s .$$

Taking the infimum on the right, over all coverings $\{U_i\}_{i\ge 1}$ of A with sets of diameter at most δ, we get

$$\mathcal{H}^s_\delta(A') \le \mathcal{H}^s_\delta(A' \times \{0\}).$$

Passing to the limit as $\delta \to 0$, we get

$$\mathcal{H}^s(A') \le \mathcal{H}^s(A' \times \{0\}),$$

which concludes the proof. □

Proposition 10. *Let $d \ge 2$ and let $f : \mathbb{R}^{d-1} \to \mathbb{R}$ be a function of the form*

$$f(x) = c + x \cdot V,$$

for some $a \in \mathbb{R}$ and some fixed vector $V \in \mathbb{R}^{d-1} \setminus \{0\}$. Then, for every measurable set $A \subset \mathbb{R}^{d-1}$ we have

$$\mathcal{H}^{d-1}\big(\Gamma(f; A)\big) = \sqrt{1 + |V|^2}\ \mathcal{L}^{d-1}(A).$$

Proof. Since the $(d-1)$-Hausdorff measure in $\mathbb{R}^d$ and the Lebesgue measure in $\mathcal{L}^d$ are both invariant with respect to translations and rotations, we can suppose that

$$c = 0 \quad \text{and} \quad V = ae_{d-1},$$

for some real constant $a > 0$.

Then, the set $\Gamma(f; A)$ is contained in the plane $\Gamma(f; \mathbb{R}^{d-1})$ and is given by:

$$\Gamma(f; A) := \Big\{ (x'', x, ax) \in \mathbb{R}^{d-2} \times \mathbb{R} \times \mathbb{R} \ : \ (x'', x) \in A \Big\}.$$

Let now $\mathcal{R}$ be a rotation sending the plane $\Gamma(f; \mathbb{R}^{d-1})$ in $\mathbb{R}^{d-1} \times \{0\}$. Then,

$$\mathcal{H}^{d-1}(\Gamma(f; A)) = \mathcal{H}^{d-1}(\mathcal{R}(\Gamma(f; A))).$$

On the other hand, for all $(x'', x, y) \in \Gamma(f; A)$, we have that

$$\mathcal{R}(x'', x, y) = (x'', \sqrt{x^2 + y^2}, 0).$$

Thus,

$$\mathcal{R}(\Gamma(f; A)) = \Big\{ (x'', \sqrt{1 + a^2}\, x, 0) \ : \ (x'', x) \in A \Big\}.$$

Since $\mathcal{L}^{d-1} = \mathcal{H}^{d-1}$ on $\mathbb{R}^{d-1}$, we get the claim by the Fubini's theorem. □

Lipschitz graphs and Hausdorff measure.
Part III. The Hausdorff measure of the graph of a Lipschitz function.

In this section we prove the following theorem.

Theorem 11 (Area formula for C^1 graphs)**.** *Let $d \geq 2$ and let $f : \mathbb{R}^{d-1} \to \mathbb{R}$ be a C^1 function. Then, for every measurable set $A \subset \mathbb{R}^{d-1}$ we have*

$$\mathcal{H}^{d-1}(\Gamma(f;A)) = \int_A \sqrt{1 + |\nabla f|^2}\, dx, \tag{1}$$

where $\Gamma(f;A)$ is the graph of f over A.

Theorem 12 (A covering theorem)**.** *Let $A \subset \mathbb{R}^d$ be a set of finite Lebesgue measure. For every $x \in A$, let $\mathcal{F}_x$ be a family of balls of the form*

$$\mathcal{F}_x = \Big\{\overline{B}_{r_{x,n}}(x) \ :\ n \in \mathbb{N}\Big\} \quad \text{with} \quad \lim_{n\to+\infty} r_{x,n} = 0.$$

Then, there is a countable set of centers

$$\{x_k\}_{k\geq 1} \subset A,$$

and radii

$$r_k := r_{x_k, n(x_k)},\ \ k \geq 1,$$

such that:

- $\overline{B}_{r_k}(x_k) \in \mathcal{F}_{x_k}$ *for every* $k \geq 1$;
- *the balls* $\overline{B}_{r_k}(x_k)$ *are disjoint;*
- $\mathcal{L}^d\bigg(A \setminus \Big(\bigcup_{k\geq 1} \overline{B}_{r_k}(x_k)\Big)\bigg) = 0.$

Proof. The proof follows by iterating the following lemma. □

Lemma 13. *Let $A \subset \mathbb{R}^d$ be a set of finite Lebesgue measure. Suppose that, for every $x \in A$, there is a family of balls*

$$\mathcal{F}_x = \Big\{\overline{B}_{r(x,n)}(x) \ :\ n \in \mathbb{N} \ \text{ such that } \lim_{n\to+\infty} r(x,n) = 0\Big\}.$$

Then, for every open set $\Omega \subset \mathbb{R}^d$ containing A, there is a finite set of centers

$$x_k \in A,\ \ k = 1, \dots, m,$$

and radii

$$r_k := r(x_k, n(x_k)),\ \ k = 1, \dots, m,$$

such that:

- $\overline{B}_{r_k}(x_k) \in \mathcal{F}_{x_k}$ *for every* $k = 1, \dots, m$;
- *the balls* $\overline{B}_{r_k}(x_k)$ *are disjoint and contained in* Ω;
- *there is a dimensional constant* $\varepsilon_d \in (0,1)$ *such that*

$$\mathcal{L}^d\bigg(A \setminus \Big(\bigcup_{k=1}^{m} \overline{B}_{r_k}(x_k)\Big)\bigg) \leq (1 - \varepsilon_d)\mathcal{L}^d(A).$$

Proof. We fix a dimensional constant $\kappa = \kappa_d \in (0,1)$, which will be chosen at the end of the proof. Let $U \subset \mathbb{R}^d$ be an open set such that

$$A \subset U \subset \Omega \quad \text{and} \quad \mathcal{L}^d(U \setminus A) \leq \kappa \mathcal{L}^d(A).$$

For every $x \in A$ pick a ball

$$\overline{B}_{r(x,n(x))}(x) \in \mathcal{F}_x \quad \text{such that} \quad \overline{B}_{r(x,n(x))}(x) \subset U.$$

Let $\mathcal{F}$ be the family of closed balls

$$\mathcal{F} = \Big\{\overline{B}_{r(x,n(x))}(x) \ :\ x \in A\Big\}.$$

By the Besicovitch's covering theorem, there are N families of balls

$$\mathcal{F}_1, \dots, \mathcal{F}_N,$$

such that:

- $N \leq N_d$, where N_d is a dimensional constant;
- each family of balls $\mathcal{F}_j$, $j = 1, \dots, N$, is a subset (subfamily) of $\mathcal{F}$;
- the union of the balls in $\mathcal{F}_1 \cup \dots \cup \mathcal{F}_N$ covers A;
- every subfamily $\mathcal{F}_j$ consists of (at most) countably many disjoint closed balls $\{B_{j,k}\}_{k\geq 1}$.

By construction we have

$$A \subset \bigcup_{j=1}^{N} \left(\bigcup_{k\geq 1} B_{j,k} \right) \subset U.$$

Then, we can find j such that

$$\frac{1}{N}\mathcal{L}^d(A) \leq \mathcal{L}^d \left(\bigcup_{k\geq 1} B_{j,k} \right) \leq \mathcal{L}^d(U).$$

Since the balls $B_{j,k}$, $k \geq 1$, are all disjoint, we have

$$\mathcal{L}^d \left(\bigcup_{B_{j,k}\in\mathcal{F}_j} B_{j,k} \right) = \sum_{k=1}^{+\infty} \mathcal{L}^d \left(B_{j,k}\right).$$

Then, we can find a (large) finite number $m > 0$, such that

$$\frac{1}{2N_d}\mathcal{L}^d(A) \leq \frac{1}{2N}\mathcal{L}^d(A) \leq \mathcal{L}^d \left(\bigcup_{k=1}^{m} B_{j,k} \right) = \sum_{k=1}^{m} \mathcal{L}^d \left(B_{j,k}\right) \leq \mathcal{L}^d(U).$$

Now, using the identity

$$\mathcal{L}^d \left(\bigcup_{k=1}^{m} B_{j,k} \right) + \mathcal{L}^d \left(A \setminus \left(\bigcup_{k=1}^{m} B_{j,k} \right) \right) = \mathcal{L}^d \left(A \cup \left(\bigcup_{k=1}^{m} B_{j,k} \right) \right) = \mathcal{L}^d \left(A\right) + \mathcal{L}^d \left(\left(\bigcup_{k=1}^{m} B_{j,k} \right) \setminus A \right),$$

we get that

$$\begin{aligned}
\mathcal{L}^d \left(A \setminus \left(\bigcup_{k=1}^{m} B_{j,k} \right) \right) &= \mathcal{L}^d \left(A\right) + \mathcal{L}^d \left(\left(\bigcup_{k=1}^{m} B_{j,k} \right) \setminus A \right) - \mathcal{L}^d \left(\bigcup_{k=1}^{m} B_{j,k} \right) \\
&\leq \mathcal{L}^d \left(A\right) + \mathcal{L}^d \left(U \setminus A\right) - \mathcal{L}^d \left(\bigcup_{k=1}^{m} B_{j,k} \right) \\
&\leq \mathcal{L}^d \left(A\right) + \mathcal{L}^d \left(U \setminus A\right) - \frac{1}{2N_d}\mathcal{L}^d \left(A\right) \\
&\leq \mathcal{L}^d \left(A\right) + \kappa\mathcal{L}^d \left(A\right) - \frac{1}{2N_d}\mathcal{L}^d \left(A\right).
\end{aligned}$$

Choosing

$$\kappa = \frac{1}{4N_d} = \varepsilon_d,$$

we get that

$$\mathcal{L}^d \left(A \setminus \left(\bigcup_{k=1}^{m} B_{j,k} \right) \right) \leq \mathcal{L}^d \left(A\right) + \frac{1}{4N_d}\mathcal{L}^d \left(A\right) - \frac{1}{2N_d}\mathcal{L}^d \left(A\right) = (1-\varepsilon_d)\mathcal{L}^d \left(A\right),$$

which concludes the proof. $\square$

Lemma 14 (Control on the projections on tangent planes). *Let $L:\mathbb{R}^{d-1}\to\mathbb{R}$ be a function of the form*

$$L(x')=c+x'\cdot V,$$

for some $c\in\mathbb{R}$ and some $V\in\mathbb{R}^{d-1}$. Let P_L be the plane $P_L:=\Gamma(L;\mathbb{R}^{d-1})$ and $\pi_L:\mathbb{R}^d\to\mathbb{R}^d$ be the projection on P_L. Let $B_r'(x_0')\subset\mathbb{R}^d$ and let $f:B_r'(x_0')\to\mathbb{R}$ be a continuous function such that

$$\|f-L\|_{L^\infty(B_r'(x_0'))}\le\varepsilon r.$$

Then,

$$\Gamma\big(L;B'_{(1-\varepsilon)r}(x_0')\big)\subset\pi_L\Big(\Gamma\big(f;B'_r(x_0')\big)\Big)\subset\Gamma\big(L;B'_{(1+\varepsilon)r}(x_0')\big).$$

Proof. Up to translations (in $\mathbb{R}^d$) and rotations (in $\mathbb{R}^{d-1}$), we can assume that:

$$x_0=0\ ,\quad c=0\quad\text{and}\quad V=ae_{d-1},$$

for some $a\ge0$. Then, for every $x'=(x'',x_{d-1})\in\mathbb{R}^{d-2}\times\mathbb{R}$, we have

$$L(x'',x_{d-1})=ax_{d-1},$$

and

$$P_L:=\Big\{(x'',x_{d-1},ax_{d-1})\ :\ x''\in\mathbb{R}^{d-2},\ x_{d-1}\in\mathbb{R}\Big\}.$$

For every $(x'',s,t)\in\mathbb{R}^{d-2}\times\mathbb{R}\times\mathbb{R}$, the map π_L is given by

$$\pi_L(x'',s,t)=\Big(x'',\frac{s+ta}{1+a^2},\frac{a(s+ta)}{1+a^2}\Big).$$

Thus, for every $(x'',x_{d-1})\in B_r'$ we have

$$\pi_L\left(x'',x_{d-1},f(x'',x_{d-1})\right)=\left(x'',\frac{x_{d-1}+af(x'',x_{d-1})}{1+a^2},\frac{a(x_{d-1}+af(x'',x_{d-1}))}{1+a^2}\right),$$

and so

$$\begin{aligned}
&\big|\pi_L\big(x'',x_{d-1},f(x'',x_{d-1})\big)-\big(x'',x_{d-1},ax_{d-1}\big)\big|\\
&\quad=\sqrt{\left(\frac{x_{d-1}+af(x'',x_{d-1})}{1+a^2}-x_{d-1}\right)^2+\left(\frac{a(x_{d-1}+af(x'',x_{d-1}))}{1+a^2}-ax_{d-1}\right)^2}\\
&\quad=\frac{1}{\sqrt{1+a^2}}\left|\big(x_{d-1}+af(x'',x_{d-1})\big)-(1+a^2)x_{d-1}\right|\\
&\quad=\frac{a}{\sqrt{1+a^2}}\left|f(x'',x_{d-1})-ax_{d-1}\right|\\
&\quad\le|f(x'',x_{d-1})-ax_{d-1}|=|f(x'',x_{d-1})-L(x'',x_{d-1})|\le\varepsilon r.
\end{aligned}$$

Let π be the projection on $\mathbb{R}^{d-1}$. It is sufficient to show that

$$B'_{(1-\varepsilon)r}\subset\pi\left(\pi_L\big(x'',x_{d-1},f(x'',x_{d-1})\big)\right)\subset B'_{(1+\varepsilon)r}.\tag{2}$$

We first notice that, thanks to the inequality

$$\big|\pi_L\big(x'',x_{d-1},f(x'',x_{d-1})\big)-\big(x'',x_{d-1},ax_{d-1}\big)\big|\le\varepsilon r,$$

we have

$$\big|\pi\left(\pi_L\big(x'',x_{d-1},f(x'',x_{d-1})\big)\right)-\pi\big(x'',x_{d-1},ax_{d-1}\big)\big|\le\varepsilon r.$$

Now, since

$$\pi\big(x'',x_{d-1},ax_{d-1}\big)=(x'',x_{d-1})\in B_r',$$

we get that

$$\pi\left(\pi_L\big(x'',x_{d-1},f(x'',x_{d-1})\big)\right)\subset B'_{(1+\varepsilon)r},$$

which proves the second inclusion in (2). In order to prove the other inclusion, we fix some $x''\in\mathbb{R}^{d-2}$ with $|x''|<r$ and we consider the segment

$$I_{x''}:=\Big(-\sqrt{r^2-|x''|^2}\ ,\sqrt{r^2-|x''|^2}\Big).$$

Suppose next that

$$|x''|\le r\sqrt{1-\varepsilon^2}.$$

Then,

$$\sqrt{r^2-|x''|^2}\ge \varepsilon r,$$

so we have the following inclusion:

$$(-\varepsilon r,\varepsilon r)\subset I_{x''}.$$

Now, by the continuity of f we know that, for each fixed $x''\in(-r,r)$, the set

$$\Big\{\pi\left(\pi_L\big(x'',x_{d-1},f(x'',x_{d-1})\big)\right)\ :\ x_{d-1}\in I_{x''}\Big\}$$

contains the set

$$\{x''\}\times J_{x''},$$

where $J_{x''}=\big[\alpha(x''),\beta(x'')\big]$ is the interval with endpoints

$$\alpha(x''):=\frac{-\sqrt{r^2-|x''|^2}+af\big(x'',-\sqrt{r^2-|x''|^2}\big)}{1+a^2}\qquad\text{and}\qquad \beta(x''):=\frac{\sqrt{r^2-|x''|^2}+af\big(x'',\sqrt{r^2-|x''|^2}\big)}{1+a^2}\ .$$

Now, since

$$\left|\frac{x_{d-1}+af(x'',x_{d-1})}{1+a^2}-x_{d-1}\right|=\frac{a}{1+a^2}\,|f(x'',x_{d-1})-ax_{d-1}|\le\frac{\varepsilon r}{2},$$

we get that $J_{x''}$ contains the interval

$$K_{x''}:=\Big(-\sqrt{r^2-|x''|^2}+\frac{\varepsilon r}{2}\ ,\ \sqrt{r^2-|x''|^2}-\frac{\varepsilon r}{2}\Big),$$

which is non-empty when $|x''|\le r\sqrt{1-\varepsilon^2}$. Moreover, thanks to the inequality

$$\sqrt{r^2-|x''|^2}-\frac{\varepsilon r}{2}\ge\sqrt{r^2(1-\varepsilon)^2-|x''|^2},$$

which easily follows from the following

$$\begin{aligned}\sqrt{r^2-|x''|^2}-\sqrt{r^2(1-\varepsilon)^2-|x''|^2}&=\frac{(1-(1-\varepsilon)^2)r^2}{\sqrt{r^2-|x''|^2}+\sqrt{r^2(1-\varepsilon)^2-|x''|^2}}\\&\ge\frac{2\varepsilon r^2}{\sqrt{r^2-|x''|^2}+\sqrt{r^2(1-\varepsilon)^2-|x''|^2}}\ge\frac{\varepsilon r^2}{\sqrt{r^2-|x''|^2}}\ge\varepsilon r,\end{aligned}$$

we get that

$$\Big(-\sqrt{r^2(1-\varepsilon)^2-|x''|^2}\ ,\ \sqrt{r^2(1-\varepsilon)^2-|x''|^2}\Big)\subset K_{x''}\subset J_{x''},$$

for all $|x''|\le r(1-\varepsilon)$. Taking the union of all these segments, for all $x''\in B''_{(1-\varepsilon)r}\subset\mathbb{R}^{d-2}$, we get the inclusion

$$B'_{r(1-\varepsilon)}\subset\pi\left(\pi_L\big(x'',x_{d-1},f(x'',x_{d-1})\big)\right),$$

which concludes the proof of (2). □

Proof of the area formula for C^1 graphs. Let A be a Lebesgue measurable set in $\mathbb{R}^{d-1}$. We first notice that if the formula holds for all bounded sets A, then it holds also for all measurable sets A. Indeed, we have

$$\mathcal{H}^{d-1}\big(\Gamma(f;A)\big)=\lim_{R\to+\infty}\mathcal{H}^{d-1}\big(\Gamma(f;A\cap B_R)\big),$$

and

$$\int_A\sqrt{1+|\nabla f|^2}\,d\mathcal{L}^{d-1}=\lim_{R\to+\infty}\int_{A\cap B_R}\sqrt{1+|\nabla f|^2}\,d\mathcal{L}^{d-1}.$$

In what follows, we will assume A to be bounded: $A\subset B_R$. We will also set $L:=\|\nabla f\|_{L^\infty(B_{R+1})}$.

We proceed in two steps.

Step 1. $\displaystyle\int_A\sqrt{1+|\nabla f|^2}\,dx\le\mathcal{H}^{d-1}(\Gamma(f;A))$.

Fix $\varepsilon>0$. Fix an open set Ω containing A, and contained in B_{R+1}, and such that $\mathcal{L}^{d-1}(\Omega\setminus A)\le\varepsilon\mathcal{L}^{d-1}(A)$.

For every $x\in A$ we select a family of closed balls $\mathcal{F}_x$ with center x such that:

- $\overline{B}_r(x)\subset\Omega$ for all $\overline{B}_r(x)\in\mathcal{F}_x$;
- $|f(y)-f(x)-(y-x)\cdot\nabla f(x)|\le\varepsilon|y-x|\le\varepsilon r$ for all $y\in\overline{B}_r(x)$ and all balls $\overline{B}_r(x)\in\mathcal{F}_x$;
- for every ball $\overline{B}_r(x)$, we have

$$\left|\sqrt{1+|\nabla f(x)|^2}\ -\frac{1}{\mathcal{L}^{d-1}(B_r)}\int_{A\cap B_r(x)}\sqrt{1+|\nabla f|^2}\,d\mathcal{L}^{d-1}\right|\le\varepsilon.$$

Thanks to Theorem 12, we can find a countable family of disjoint closed balls $\{\overline{B}_{r_i}(x_i)\}_{i\geq 1}$ such that the following holds

- $x_i \in A$ and $\overline{B}_{r_i}(x_i) \subset \Omega$ for all $i \geq 1$;
- setting
$$U := \bigcup_{i\geq 1} \overline{B}_{r_i}(x_i),$$
we have that
$$\mathcal{L}^{d-1}(A \setminus U) = 0 \qquad \text{and} \qquad \mathcal{L}^{d-1}(U \setminus A) \leq \varepsilon \mathcal{L}^{d-1}(A);$$
- setting $L_i(y) := f(x_i) + (y - x_i) \cdot \nabla f(x_i)$, we have
$$\|f - L_i\|_{L^\infty(B_{r_i}(x_i))} \leq \varepsilon r_i \quad \text{for all} \quad i \geq 1;$$
- for all $i \geq 1$, we have
$$\left| \sqrt{1 + |\nabla f(x_i)|^2} - \frac{1}{\mathcal{L}^{d-1}(B_{r_i})} \int_{A\cap B_{r_i}(x_i)} \sqrt{1 + |\nabla f|^2}\, d\mathcal{L}^{d-1} \right| \leq \varepsilon.$$

Then, thanks to Lemma 14, we have

$$\mathcal{H}^{d-1}\left(\Gamma\big(f; B_{r_i}(x_i)\big)\right) \geq \mathcal{H}^{d-1}\left(\pi_i\Big(\Gamma\big(f; B_{r_i}(x_i)\big)\Big)\right) \geq \mathcal{H}^{d-1}\left(\Gamma\big(L_i; B_{r_i(1-\varepsilon)}(x_i)\big)\right).$$

Now, thanks to Proposition 10, we have that

$$\mathcal{H}^{d-1}\left(\Gamma\big(L_i; B_{r_i(1-\varepsilon)}(x_i)\big)\right) = \sqrt{1 + |\nabla L_i|^2}\, \mathcal{L}^{d-1}\left(B_{r_i(1-\varepsilon)}(x_i)\right) = (1-\varepsilon)^{d-1}\sqrt{1 + |\nabla L_i|^2}\, \mathcal{L}^{d-1}\left(B_{r_i}\right).$$

By the choice of L_i, we have that $\nabla L_i = \nabla f(x_i)$, so we have

$$\mathcal{H}^{d-1}\left(\Gamma\big(L_i; B_{r_i(1-\varepsilon)}(x_i)\big)\right) = (1-\varepsilon)^{d-1}\sqrt{1 + |\nabla f(x_i)|^2}\, \mathcal{L}^{d-1}\left(B_{r_i}\right),$$

which finally gives the inequality

$$\mathcal{H}^{d-1}\left(\Gamma\big(f; B_{r_i}(x_i)\big)\right) \geq (1-\varepsilon)^{d-1}\sqrt{1 + |\nabla f(x_i)|^2}\, \mathcal{L}^{d-1}\left(B_{r_i}\right).$$

Now, since x_i is a Lebesgue point for both A and $\sqrt{1 + |\nabla f|^2}$, we have

$$\left| \sqrt{1 + |\nabla f(x_i)|^2} - \frac{1}{\mathcal{L}^{d-1}(B_{r_i})} \int_{A\cap B_{r_i}(x_i)} \sqrt{1 + |\nabla f|^2}\, d\mathcal{L}^{d-1} \right| \leq \varepsilon,$$

so

$$\frac{1}{\mathcal{L}^{d-1}(B_{r_i})} \int_{A\cap B_{r_i}(x_i)} \sqrt{1 + |\nabla f|^2}\, d\mathcal{L}^{d-1} \leq \varepsilon + \sqrt{1 + |\nabla f(x_i)|^2} \leq (1+\varepsilon)\sqrt{1 + |\nabla f(x_i)|^2},$$

which finally gives

$$\mathcal{H}^{d-1}\left(\Gamma\big(f; B_{r_i}(x_i)\big)\right) \geq (1-\varepsilon)^{d-2} \int_{A\cap B_{r_i}(x_i)} \sqrt{1 + |\nabla f|^2}\, d\mathcal{L}^{d-1}.$$

Now, summing up over all $i \geq 1$, we obtain

$$\begin{aligned}
\mathcal{H}^{d-1}\left(\Gamma\big(f; U\big)\right) &= \sum_{i\geq 1} \mathcal{H}^{d-1}\left(\Gamma\big(f; \overline{B}_{r_i}(x_i)\big)\right) \\
&\geq (1-\varepsilon)^{d-2} \sum_{i\geq 1} \int_{A\cap \overline{B}_{r_i}(x_i)} \sqrt{1 + |\nabla f|^2}\, d\mathcal{L}^{d-1} \\
&= (1-\varepsilon)^{d-2} \int_{A\cap U} \sqrt{1 + |\nabla f|^2}\, d\mathcal{L}^{d-1} \\
&= (1-\varepsilon)^{d-2} \int_{A} \sqrt{1 + |\nabla f|^2}\, d\mathcal{L}^{d-1}.
\end{aligned}$$

Now, we only need to observe that, thanks to Proposition 7, we have

$$\begin{aligned}\mathcal{H}^{d-1}\left(\Gamma(f;U)\right) &= \mathcal{H}^{d-1}\left(\Gamma(f;A\cap U)\right)+\mathcal{H}^{d-1}\left(\Gamma(f;U\setminus A)\right)\\ &= \mathcal{H}^{d-1}\left(\Gamma(f;A)\right)+\mathcal{H}^{d-1}\left(\Gamma(f;U\setminus A)\right)\\ &\le \mathcal{H}^{d-1}\left(\Gamma(f;A)\right)+\left(1+L^2\right)^{\frac{d-1}{2}}\,\mathcal{L}^{d-1}(U\setminus A)\\ &\le \mathcal{H}^{d-1}\left(\Gamma(f;A)\right)+\varepsilon\left(1+L^2\right)^{\frac{d-1}{2}}\,\mathcal{L}^{d-1}(A)\\ &\le \mathcal{H}^{d-1}\left(\Gamma(f;A)\right)+\varepsilon\left(1+L^2\right)^{\frac{d-1}{2}}\int_A\sqrt{1+|\nabla f|^2}\,d\mathcal{L}^{d-1}.\end{aligned}$$

Putting together all these estimates, we get

$$\mathcal{H}^{d-1}\left(\Gamma(f;A)\right)\ge\left((1-\varepsilon)^{d-2}-\varepsilon\left(1+L^2\right)^{\frac{d-1}{2}}\right)\int_A\sqrt{1+|\nabla f|^2}\,d\mathcal{L}^{d-1}.$$

Since $\varepsilon>0$ is arbitrary, the proof of Step 1 is concluded.

Step 2. $\mathcal{H}^{d-1}(\Gamma(f;A))\le\displaystyle\int_A\sqrt{1+|\nabla f|^2}\,dx$.

The construction is similar, yet different with respect to the one from the previous step. Fix $\varepsilon>0$. Fix an open set Ω containing A and such that

$$\int_{\Omega\setminus A}\sqrt{1+|\nabla f|^2}\,dx\le\varepsilon\int_A\sqrt{1+|\nabla f|^2}\,dx.$$

For every $x\in A$ we select a family of closed balls $\mathcal{F}_x$ with center x such that:

- $\overline{B}_r(x)\subset\Omega$ for all $\overline{B}_r(x)\in\mathcal{F}_x$;
- $|\nabla f(y)-\nabla f(x)|\le\varepsilon$ for all $y\in\overline{B}_r(x)$ and all balls $\overline{B}_r(x)\in\mathcal{F}_x$.

By Theorem 12, there is a countable family of disjoint closed balls $\{\overline{B}_{r_i}(x_i)\}_{i\ge1}$ such that the following holds:

- $x_i\in A$ and $\overline{B}_{r_i}(x_i)\subset\Omega$ for all $i\ge1$;
- setting $U:=\displaystyle\bigcup_{i\ge1}\overline{B}_{r_i}(x_i)$, we have that $\mathcal{L}^{d-1}(A\setminus U)=0$;
- setting $L_i(y):=f(x_i)+(y-x_i)\cdot\nabla f(x_i)$, we have

$$\|\nabla f-\nabla L_i\|_{L^\infty(B_{r_i}(x_i))}\le\varepsilon\quad\text{for all}\quad i\ge1.$$

For each $i\ge1$ we define the map

$$\Phi_i:\Gamma\Big(L_i;\overline{B}_{r_i}(x_i)\Big)\to\Gamma\Big(f;\overline{B}_{r_i}(x_i)\Big),$$

as follows:

$$\Phi_i\Big(\big(x',L_i(x')\big)\Big)=\big(x',f(x')\big)\quad\text{for all}\quad x'\in\overline{B}_{r_i}(x_i).$$

We will next prove that Φ_i is Lipschitz. Indeed, for all $x',y'\in\overline{B}_{r_i}(x_i)$, we have

$$\begin{aligned}\Big|\big(x',f(x')\big)-\big(y',f(y')\big)\Big| &= \sqrt{|x'-y'|^2+|f(x')-f(y')|^2}\\ &= \sqrt{|x'-y'|^2+|(x'-y')\cdot\nabla f(tx'+(1-t)y')|^2},\end{aligned}$$

for some $t=t_{x',y'}\in(0,1)$. Now, since $tx'+(1-t)y'\in B_{r_i}(x_i)$, we have that

$$|\nabla f(x_i)-\nabla f(tx'+(1-t)y')|\le\varepsilon,$$

so we get

$$
\begin{aligned}
|x'-y'|^2 + |(x'-y')\cdot \nabla f(tx'+(1-t)y')|^2 \\
&\le |x'-y'|^2 + |(x'-y')\cdot\nabla L_i|^2 \\
&\qquad + 2\Big((x'-y')\cdot\nabla L_i\Big)\Big((x'-y')\cdot\Big(\nabla L_i - \nabla f(tx'+(1-t)y')\Big)\Big) \\
&\qquad\quad + \Big|(x'-y')\cdot\Big(\nabla L_i - \nabla f(tx'+(1-t)y')\Big)\Big|^2 \\
&\le |x'-y'|^2 + |(x'-y')\cdot\nabla L_i|^2 \\
&\qquad + 2|(x'-y')\cdot\nabla L_i|\,|x'-y'|\varepsilon \\
&\qquad\quad + |x'-y'|^2\varepsilon^2 \\
&\le |x'-y'|^2 + |(x'-y')\cdot\nabla L_i|^2 \\
&\qquad + \varepsilon\Big(|(x'-y')\cdot\nabla L_i|^2 + |x'-y'|^2\Big) \\
&\qquad\quad + |x'-y'|^2\varepsilon^2 \\
&\le (1+\varepsilon)^2\Big(|x'-y'|^2 + |(x'-y')\cdot\nabla L_i|^2\Big),
\end{aligned}
$$

so that

$$
\begin{aligned}
\Big|\big(x', f(x')\big) - \big(y', f(y')\big)\Big| &\le (1+\varepsilon)\sqrt{|x'-y'|^2 + |(x'-y')\cdot\nabla L_i|^2} \\
&= \Big|\big(x', L_i(x')\big) - \big(y', L_i(y')\big)\Big|,
\end{aligned}
$$

which proves that Φ_i is $(1+\varepsilon)$-Lipschitz continuous. Thus,

$$
\begin{aligned}
\mathcal{H}^{d-1}\Big(\Gamma\big(f;\overline{B}_{r_i}(x_i)\big)\Big) &\le \mathcal{H}^{d-1}\Big(\Phi_i\Big(\Gamma\big(L_i;\overline{B}_{r_i}(x_i)\big)\Big)\Big) \\
&\le (1+\varepsilon)^{d-1}\mathcal{H}^{d-1}\Big(\Gamma\big(L_i;\overline{B}_{r_i}(x_i)\big)\Big) \\
&= (1+\varepsilon)^{d-1}\int_{B_{r_i}(x_i)} \sqrt{1+|\nabla L_i|^2}\,dx.
\end{aligned}
$$

Now, since

$$
\Big|\sqrt{1+|V|^2} - \sqrt{1+|W|^2}\Big| = \Big|\frac{|V|^2-|W|^2}{\sqrt{1+|V|^2}+\sqrt{1+|W|^2}}\Big| \le |V-W| \quad \text{for all} \quad V, W \in \mathbb{R}^{d-1},
$$

we have that

$$
\int_{B_{r_i}(x_i)} \sqrt{1+|\nabla L_i|^2}\,dx \le (1+\varepsilon)\int_{B_{r_i}(x_i)} \sqrt{1+|\nabla f|^2}\,dx,
$$

so we finally obtain

$$
\mathcal{H}^{d-1}\Big(\Gamma\big(f;\overline{B}_{r_i}(x_i)\big)\Big) \le (1+\varepsilon)^d \int_{B_{r_i}(x_i)} \sqrt{1+|\nabla f|^2}\,dx.
$$

Now, summing over all $i \ge 1$, we get

$$
\begin{aligned}
\mathcal{H}^{d-1}\Big(\Gamma\big(f;A\big)\Big) &\le \sum_{i=1}^{+\infty} \mathcal{H}^{d-1}\Big(\Gamma\big(f;\overline{B}_{r_i}(x_i)\big)\Big) \\
&\le \sum_{i=1}^{+\infty}(1+\varepsilon)^d \int_{B_{r_i}(x_i)} \sqrt{1+|\nabla f|^2}\,dx \\
&= (1+\varepsilon)^d \int_U \sqrt{1+|\nabla f|^2}\,dx \\
&= (1+\varepsilon)^d \int_A \sqrt{1+|\nabla f|^2}\,dx + (1+\varepsilon)^d \int_{\Omega\setminus A} \sqrt{1+|\nabla f|^2}\,dx \\
&= (1+\varepsilon)^{d+1} \int_A \sqrt{1+|\nabla f|^2}\,dx,
\end{aligned}
$$

which concludes the proof since $\varepsilon > 0$ is arbitrary. □

LIPSCHITZ GRAPHS AND HAUSDORFF MEASURE.
PART IV. THE HAUSDORFF MEASURE OF THE GRAPH OF A LIPSCHITZ FUNCTION.

In this section we prove the following theorem.

Theorem 15 (Area formula for Lipschitz graphs). *Let $d \geq 2$ and let $f : \mathbb{R}^{d-1} \to \mathbb{R}$ be an L-Lipschitz function, for some $L > 0$. Then, for every measurable set $A \subset \mathbb{R}^{d-1}$ we have*

$$\mathcal{H}^{d-1}(\Gamma(f;A)) = \int_A \sqrt{1 + |\nabla f|^2}\, dx, \tag{3}$$

where $\Gamma(f;A)$ is the graph of f over A.

The key for the proof of the area formula is the following lemma.

Lemma 16. *Let $f : \mathbb{R}^{d-1} \to \mathbb{R}$ be a Lipschitz function and let $A \subset \mathbb{R}^{d-1}$ be the set of points at which f is differentiable. Then, for every $\varepsilon \in (0, 1/2)$, there is a countable collection of disjoint measurable sets $\{E_k\}_{k\geq1}$ in $\mathbb{R}^{d-1}$ and vectors $\{V_k\}_{k\geq1}$ in $\mathbb{R}^{d-1}$ such that:*

(i) *$diam(E_k) \leq 2$ for all $k \geq 1$;*

(ii) *$E_k \subset A$ for all $k \geq 1$;*

(iii) *$A = \bigcup_{k\geq1} E_k$;*

(iv) *$|V_k - \nabla f(x')| \leq \varepsilon$ for all $x' \in E_k$;*

(v) *for all $x', y' \in E_k$, it holds*

$$\begin{aligned}(1 - 2\varepsilon)&\sqrt{|x' - y'|^2 + (V_k \cdot (x' - y'))^2}\\ &\leq \sqrt{|x' - y'|^2 + (f(x') - f(y'))^2}\\ &\leq (1 + 2\varepsilon)\sqrt{|x' - y'|^2 + (V_k \cdot (x' - y'))^2}.\end{aligned} \tag{4}$$

Proof. Let $\mathcal{C}$ be a dense subset of A. For each $(p', V, n) \in \mathcal{C} \times \mathbb{Q}^{d-1} \times \mathbb{N}$, define the set

$$E(p', V, n) \subset A,$$

as follows. We say that $x' \in E(p', V, n)$, if:

- $x' \in B'_{1/n}(p')$;
- $|V - \nabla f(x')| \leq \varepsilon$;
- for all $y' \in B'_{2/n}(x')$ it holds:

$$|f(y') - f(x') - \nabla f(x') \cdot (y' - x')| \leq \varepsilon |y' - x'|.$$

Step 1. We claim that, if $(p', V, n) \in \mathcal{C} \times \mathbb{Q}^{d-1} \times \mathbb{N}$, then

$$\text{for all} \quad x' \in E(p', V, n) \quad \text{and all} \quad y' \in B'_{2/n}(x'),$$

we have

$$\sqrt{|x' - y'|^2 + (f(x') - f(y'))^2} \leq (1 + 2\varepsilon)\sqrt{|x' - y'|^2 + (V \cdot (x' - y'))^2}. \tag{5}$$

Indeed,

$$\begin{aligned}\sqrt{|x' - y'|^2 + (f(x') - f(y'))^2} &\leq \sqrt{|x' - y'|^2 + \Big((x' - y') \cdot \nabla f(x') + \varepsilon |x' - y'|\Big)^2}\\ &\leq \sqrt{|x' - y'|^2 + \Big((x' - y') \cdot V + (x' - y') \cdot \big(\nabla f(x') - V\big) + \varepsilon |x' - y'|\Big)^2}.\end{aligned}$$

Now, since

$$\begin{aligned}\Big((x'-y')\cdot\nabla f(x')+\varepsilon|x'-y'|\Big)^2 &\le \Big((x'-y')\cdot V+(x'-y')\cdot\big(\nabla f(x')-V\big)+\varepsilon|x'-y'|\Big)^2\\ &\le \Big((x'-y')\cdot V\Big)^2\\ &\quad+2\Big((x'-y')\cdot V\Big)\Big((x'-y')\cdot\big(\nabla f(x')-V\big)+\varepsilon|x'-y'|\Big)\\ &\qquad+\Big((x'-y')\cdot\big(\nabla f(x')-V\big)+\varepsilon|x'-y'|\Big)^2\\ &\le \big|(x'-y')\cdot V\big|^2\\ &\quad+4\varepsilon\big|(x'-y')\cdot V\big|\,|x'-y'|\\ &\qquad+4\varepsilon^2|x'-y'|^2\\ &\le \big|(x'-y')\cdot V\big|^2\\ &\quad+2\varepsilon\Big(\big|(x'-y')\cdot V\big|^2+|x'-y'|^2\Big)\\ &\qquad+4\varepsilon^2|x'-y'|^2,\end{aligned}$$

we get precisely (5).

Step 2. We claim that, if $(p',V,n)\in\mathcal{C}\times\mathbb{Q}^{d-1}\times\mathbb{N}$, then

$$\text{for all}\quad x'\in E(p',V,n)\quad\text{and all}\quad y'\in B'_{2/n}(x'),$$

we have

$$\sqrt{|x'-y'|^2+(f(x')-f(y'))^2}\ge(1-2\varepsilon)\sqrt{|x'-y'|^2+(V\cdot(x'-y'))^2}. \tag{6}$$

Indeed, since

$$\begin{aligned}|f(x')-f(y')| &\ge |(y'-x')\cdot\nabla f(x')|-|f(x')-f(y')-(y'-x')\cdot\nabla f(x')|\\ &\ge |(y'-x')\cdot\nabla f(x')|-\varepsilon|y'-x'|\\ &\ge |(y'-x')\cdot V|-|(y'-x')\cdot(V-\nabla f(x'))|-\varepsilon|y'-x'|\\ &\ge |(y'-x')\cdot V|-2\varepsilon|y'-x'|,\end{aligned}$$

we get

$$\begin{aligned}&\sqrt{|x'-y'|^2+(f(x')-f(y'))^2}\\ &\quad\ge\sqrt{|x'-y'|^2+\Big(|(y'-x')\cdot V|-2\varepsilon|y'-x'|\Big)^2}\\ &\quad=\sqrt{|x'-y'|^2+|(y'-x')\cdot V|^2-4\varepsilon|(y'-x')\cdot V||x-y|+4\varepsilon^2|x'-y'|^2}\\ &\quad\ge\sqrt{|x'-y'|^2+|(y'-x')\cdot V|^2-2\varepsilon\Big(|x'-y'|^2+|(y'-x')\cdot V|^2\Big)+4\varepsilon^2|x'-y'|^2}\\ &\quad\ge\sqrt{|x'-y'|^2+|(y'-x')\cdot V|^2-4\varepsilon\Big(|x'-y'|^2+|(y'-x')\cdot V|^2\Big)+4\varepsilon^2\Big(|x'-y'|^2+|(y'-x')\cdot V|^2\Big)}\\ &\quad=(1-2\varepsilon)\sqrt{|x'-y'|^2+|(y'-x')\cdot V|^2},\end{aligned}$$

where in the last inequality we used that

$$-2\varepsilon\Big(|x'-y'|^2+|(y'-x')\cdot V|^2\Big)+4\varepsilon^2|(y'-x')\cdot V|^2\le 0,$$

which is true since $\varepsilon\le 1/2$. This concludes the proof of (6).

Step 3. We claim that, if $(p',V,n)\in\mathcal{C}\times\mathbb{Q}^{d-1}\times\mathbb{N}$, then for all $x',y'\in E(p',V,n)$ it holds

$$\sqrt{|x'-y'|^2+(f(x')-f(y'))^2}\le(1+2\varepsilon)\sqrt{|x'-y'|^2+(V\cdot(x'-y'))^2}, \tag{7}$$

and xxx

$$\sqrt{|x'-y'|^2+(f(x')-f(y'))^2}\ge(1-2\varepsilon)\sqrt{|x'-y'|^2+(V\cdot(x'-y'))^2}. \tag{8}$$

This follows from the previous steps. Indeed, since $x',y'\in E(p',V,n)$, we have that $x',y'\in B'_{1/n}(p')$. But then, $y'\in B'_{2/n}(x')$, so (7) follows from Step 1, while (8) is a consequence of Step 2.

Step 4. We claim that every $x' \in A$ at which f is differentiable lies in some set $E(p', V, n)$. Indeed, since

$$f(y') - f(x') - (y' - x') \cdot \nabla f(x') = o(|x' - y'|),$$

we can find some $n \in \mathbb{N}$ large enough such that

$$\Big| f(y') - f(x') - (y' - x') \cdot \nabla f(x') \Big| \le \varepsilon |x' - y'| \quad \text{for all} \quad y' \in B'_{2/n}(x').$$

Now, simply pick some $p' \in C \cap B'_{1/n}(x')$ and some $V \in \mathbb{Q}^{d-1}$ sufficiently close to $\nabla f(x')$.

Step 5. Since the family of sets

$$\{E(p', V, n)\}_{(p', V, n) \in \mathcal{C} \times \mathbb{Q}^{d-1} \times \mathbb{N}},$$

is countable we can order it in a sequence $\{\widetilde{E}_k\}_{k \ge 1}$, with

$$\widetilde{E}_k = E(p'_k, V_k, n_k),$$

for some $p'_k \in \mathcal{C}$, $V_k \in \mathbb{Q}^{d-1}$, $n_k \in \mathbb{N}$.

We define the family $\{E_k\}_{k \ge 1}$ as follows. We set $E_1 = \widetilde{E}_1$ and

$$E_k := \widetilde{E}_k \setminus \Big(\bigcup_{j=1}^{k-1} E_j \Big) \quad \text{for all} \quad k \ge 2.$$

By construction, we know that each of the sets E_k is contained in some ball $B'_{1/n_k}(p'_k)$, so clearly $\operatorname{diam}(E_k) \le 2$. The estimate in (iv) follows from the definition of the sets $E(p'_k, V_k, n_k)$, while the last estimate (4) follows from (5) and (6). This concludes the proof of Lemma 16. $\square$

Proof of the area formula for Lipschitz graphs (Theorem 15).

Let A be a Lebesgue measurable set in $\mathbb{R}^{d-1}$.

If $\mathcal{L}^{d-1}(A) = +\infty$, then also $\mathcal{H}^{d-1}(A) = +\infty$. Since A is the projection of $\Gamma(f; A)$ on $\mathbb{R}^{d-1}$,

$$\pi\Big(\Gamma(f; A)\Big) = A,$$

we get that

$$\mathcal{H}^{d-1}(A) \le \mathcal{H}^{d-1}(\Gamma(f; A)).$$

Thus, when $\mathcal{L}^{d-1}(A) = +\infty$, we have:

$$\mathcal{H}^{d-1}(\Gamma(f; A)) \ge \mathcal{H}^{d-1}(A) = \mathcal{L}^{d-1}(A) = +\infty,$$

and at the same time

$$\int_A \sqrt{1 + |\nabla f|^2}\, dx \ge \mathcal{L}^{d-1}(A) = +\infty,$$

so (3) holds trivially as both sides are $+\infty$.

$$\text{In what follows, we will assume that } \mathcal{L}^{d-1}(A) < +\infty.$$

Let $\widetilde{A} \subset A$ be a Borel set in $\mathbb{R}^{d-1}$ such that $\mathcal{L}^{d-1}(A \setminus \widetilde{A}) = 0$ and

$$f \text{ is differentiable at all points } x \in \widetilde{A}.$$

Fix $\varepsilon > 0$. Thanks to Lemma 16, there are countable families of Lebesgue measurable sets $E_k \subset \widetilde{A}$ and linear functions

$$L_k : \mathbb{R}^{d-1} \to \mathbb{R}\ , \qquad L_k(x') = V_k \cdot x' \quad \text{with} \quad V_k \in \mathbb{R}^{d-1},$$

such that:

- $\operatorname{diam}(E_k) \le 2$ for all $k \ge 1$;
- E_k are disjoint and $\widetilde{A} = \bigcup_{k \ge 1} E_k$;
- $|\nabla L_k - \nabla f(x')| \le \varepsilon$ for all $x' \in E_k$;

- for all $x', y' \in E_k$, it holds

$$\begin{aligned}(1-2\varepsilon)&\sqrt{|x'-y'|^2+(L_k(x'-y'))^2}\\ &\le \sqrt{|x'-y'|^2+(f(x')-f(y'))^2}\\ &\qquad \le (1+2\varepsilon)\sqrt{|x'-y'|^2+(L_k(x'-y'))^2}.\end{aligned}$$

Then, since $\mathcal{L}^{d-1}(A\setminus\widetilde{A})=0$, we have

$$\mathcal{H}^{d-1}(\Gamma(f;A))=\sum_{k\ge1}\mathcal{H}^{d-1}(\Gamma(f;E_k))\qquad\text{and}\qquad \int_A\sqrt{1+|\nabla f|^2}\,dx=\sum_{k\ge1}\int_{E_k}\sqrt{1+|\nabla f|^2}\,dx.$$

Now, thanks to Lemma 5, we have that for each k

$$(1-2\varepsilon)^{d-1}\mathcal{H}^{d-1}(\Gamma(L_k;E_k))\le\mathcal{H}^{d-1}(\Gamma(f;E_k))\le(1+2\varepsilon)^{d-1}\mathcal{H}^{d-1}(\Gamma(L_k;E_k)).$$

On the other hand, Proposition 10 gives that

$$\mathcal{H}^{d-1}(\Gamma(L_k;E_k))=\int_{E_k}\sqrt{1+|\nabla L_k|^2}\,dx,$$

while since

$$\sqrt{1+|\nabla L_k|^2}-\sqrt{1+|\nabla f|^2}=\frac{\nabla(f+L_k)\cdot\nabla(f-L_k)}{\sqrt{1+|\nabla L_k|^2}+\sqrt{1+|\nabla f|^2}},$$

we get that

$$\left|\sqrt{1+|\nabla L_k|^2}-\sqrt{1+|\nabla f|^2}\right|\le|\nabla f-\nabla L_k|\le\varepsilon\quad\text{on}\quad E_k,$$

so that

$$\begin{aligned}\mathcal{H}^{d-1}(\Gamma(f;A))&=\sum_{k\ge1}\mathcal{H}^{d-1}(\Gamma(f;E_k))\\ &\le(1+2\varepsilon)^{d-1}\sum_{k\ge1}\mathcal{H}^{d-1}(\Gamma(L_k;E_k))\\ &=(1+2\varepsilon)^{d-1}\sum_{k\ge1}\int_{E_k}\sqrt{1+|\nabla L_k|^2}\,dx\\ &\le(1+2\varepsilon)^{d-1}\sum_{k\ge1}(1+\varepsilon)\int_{E_k}\sqrt{1+|\nabla f|^2}\,dx\le(1+2\varepsilon)^d\int_A\sqrt{1+|\nabla f|^2}\,dx,\end{aligned}$$

and conversely

$$\begin{aligned}\mathcal{H}^{d-1}(\Gamma(f;A))&=\sum_{k\ge1}\mathcal{H}^{d-1}(\Gamma(f;E_k))\\ &\ge(1-2\varepsilon)^{d-1}\sum_{k\ge1}\mathcal{H}^{d-1}(\Gamma(L_k;E_k))\\ &=(1-2\varepsilon)^{d-1}\sum_{k\ge1}\int_{E_k}\sqrt{1+|\nabla L_k|^2}\,dx\\ &\ge(1-2\varepsilon)^{d-1}\sum_{k\ge1}(1-\varepsilon)\int_{E_k}\sqrt{1+|\nabla f|^2}\,dx\ge(1-2\varepsilon)^d\int_A\sqrt{1+|\nabla f|^2}\,dx.\end{aligned}$$

Sending $\varepsilon\to0$, we get the claim. $\square$