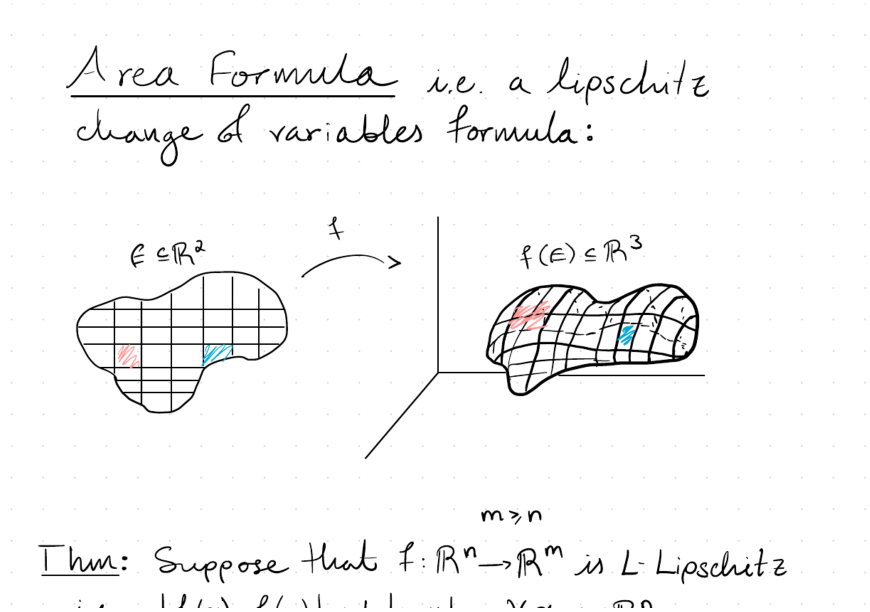
The Area Formula for Lipschitz Maps via Local Linearization

Connor Marrs

May 14th, 2026



Area Formula i.e. a lipschitz change of variables formula:

Thm: Suppose that $f : \mathbb{R}^n \xrightarrow{} \mathbb{R}^m$ ($m \geq n$) is L-Lipschitz

Guiding idea. The area formula is a Lipschitz change-of-variables theorem. A Lipschitz map is not smooth enough for the inverse function theorem, but it is differentiable almost everywhere. The proof isolates the points where the a.e. derivative behaves like an immersion, cuts those points into countably many almost-linear sheets, and then handles the remaining points by showing that they have no n-dimensional image measure. [1]

---

[1] This is a set of notes that was typeset by ChatGPT, but was handwritten by myself based on [1] and [2].

1 Statement and proof strategy

Throughout,

$$f : \mathbb{R}^n \to \mathbb{R}^m, \qquad n \le m,$$

is Lipschitz. At points where $Df(x)$ exists, set

$$Jf(x) = \sqrt{\det\big(Df(x)^T Df(x)\big)}.$$

On the set where Df does not exist, set $Jf(x) = 0$.

For a set $A \subset \mathbb{R}^n$, define the multiplicity function

$$N(f, A, y) = \mathcal{H}^0(A \cap f^{-1}\{y\}) = \#\{x \in A : f(x) = y\},$$

with value $+\infty$ allowed.

Theorem 1 (Area formula)**.** For every Lebesgue measurable $A \subset \mathbb{R}^n$,

$$\int_A Jf(x)\, d\mathcal{L}^n(x) = \int_{\mathbb{R}^m} N(f, A, y)\, d\mathcal{H}^n(y).$$

If f is one-to-one on A, this reduces to

$$\mathcal{H}^n(f(A)) = \int_A Jf(x)\, dx.$$

We use the normalization $\mathcal{H}^n = \mathcal{L}^n$ on $\mathbb{R}^n$. The proof below is written for Borel sets. The Lebesgue measurable case follows by replacing A by a Borel representative up to a Lebesgue-null set and using the Lipschitz null-image lemma below.

The three pieces of the proof

Given a Borel set A, split it as

$$A = A_{\text{good}} \sqcup A_{\text{crit}} \sqcup A_{\text{bad}},$$

where

$$A_{\text{good}} = A \cap \{Df \text{ exists and } Jf > 0\},$$
$$A_{\text{crit}} = A \cap \{Df \text{ exists and } Jf = 0\},$$

and

$$A_{\text{bad}} = A \cap \{Df \text{ does not exist}\}.$$

The proof uses a different idea on each part.

Piece	What happens geometrically	Tool
A_{good}	$Df(x)$ has full rank, so f is infinitesimally an n-dimensional immersion.	Local linear approximation: cover by almost-linear bi-Lipschitz sheets.
A_{crit}	$Df(x)$ has rank $< n$, so infinitesimal balls collapse toward lower-dimensional planes.	Critical set lemma: $\mathcal{H}^n(f(A_{\text{crit}})) = 0$.
A_{bad}	No derivative is available.	Rademacher plus Lipschitz null-image estimate.

The point is that the area formula is proved only on A_{good} in a serious way. The other two pieces contribute zero to both sides.

Indeed, Rademacher gives $\mathcal{L}^n(A_{\text{bad}}) = 0$, and Lipschitz maps send $\mathcal{L}^n$-null sets to $\mathcal{H}^n$-null sets, so

$$\int_{A_{\text{bad}}} Jf = 0, \qquad \mathcal{H}^n(f(A_{\text{bad}})) = 0.$$

The critical set lemma gives

$$\int_{A_{\text{crit}}} Jf = 0, \qquad \mathcal{H}^n(f(A_{\text{crit}})) = 0.$$

The multiplicity function of a set S is supported on $f(S)$. Hence, if $\mathcal{H}^n(f(S)) = 0$, then

$$\int_{\mathbb{R}^m} N(f, S, y)\, d\mathcal{H}^n(y) = 0.$$

Thus only A_{good} remains.

2 Preliminary facts

Theorem 2 (Rademacher)**.** If $f : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz, then $Df(x)$ exists for $\mathcal{L}^n$-a.e. $x \in \mathbb{R}^n$.

Lemma 1 (Lipschitz null-image lemma)**.** If $f : \mathbb{R}^n \to \mathbb{R}^m$ is L-Lipschitz and $N \subset \mathbb{R}^n$ satisfies $\mathcal{L}^n(N) = 0$, then

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

Proof. Cover N by balls $B_i = B(x_i, r_i)$ with $\sum_i r_i^n$ arbitrarily small. Since f is L-Lipschitz, $f(B_i)$ has diameter at most $2Lr_i$. Therefore the n-dimensional Hausdorff content of $f(N)$ is bounded by a dimensional constant times

$$L^n \sum_i r_i^n.$$

Let the covering content tend to zero. $\square$

3 The local linear approximation lemma

This is the central construction. It replaces smooth coordinate charts by countably many measurable sheets on which f is quantitatively close to a fixed linear metric normalization.

3.1 What should one ask for?

At a point $x \in A_{\mathrm{good}}$, the derivative

$$Df(x) : \mathbb{R}^n \to \mathbb{R}^m$$

has full rank. Its polar factorization is

$$Df(x) = O_x S_x, \qquad S_x = (Df(x)^T Df(x))^{1/2},$$

where $O_x : \mathbb{R}^n \to \mathbb{R}^m$ is an isometric embedding and $S_x : \mathbb{R}^n \to \mathbb{R}^n$ is symmetric positive definite. Then

$$|Df(x)v| = |S_x v|, \qquad Jf(x) = \det S_x.$$

The isometry O_x records how the image ellipsoid sits inside $\mathbb{R}^m$, but it does not change n-dimensional area. The metric part S_x records the shape and volume of the infinitesimal ellipsoid.

So the natural question is not primarily:

Which full map $L : \mathbb{R}^n \to \mathbb{R}^m$ is close to $Df(x)$ near x?

Instead, for the area formula, the better question is:

For which points x is the metric part of $Df(x)$ close to one fixed T_k?

Here the T_k are chosen from a countable dense subset of the symmetric positive definite automorphisms of $\mathbb{R}^n$.

Thus, for a fixed tolerance $t > 1$, one wants to group together the points x satisfying

$$t^{-1}|T_k v| \lesssim |Df(x)v| \lesssim t|T_k v| \qquad \forall v \in \mathbb{R}^n.$$

This gives the desired comparison

$$Jf(x) \approx \det T_k.$$

That is the volume information.

But derivative comparison at individual points is not enough. To estimate $\mathcal{H}^n(f(F))$ for an arbitrary subset F of a piece, we need control of the secants

$$f(a) - f(b), \qquad a, b \in F.$$

This is why the pieces also impose a uniform Taylor estimate:

$$f(a) - f(b) = Df(b)(a-b) + \text{small error}$$

whenever a and b lie in the same small piece. Combining the derivative comparison with the Taylor estimate gives

$$t^{-1}|T_k(a-b)| \le |f(a) - f(b)| \le t|T_k(a-b)|.$$

So $T_k x \mapsto f(x)$ is bi-Lipschitz on the piece.

3.2 Why the two conclusions are both needed

On one piece E_k, the lemma gives two kinds of control:

$$\frac{1}{t}|T_k(a-b)| \le |f(a) - f(b)| \le t|T_k(a-b)|, \qquad a, b \in E_k, \tag{secant control}$$

and

$$t^{-n} \det T_k \le Jf(x) \le t^n \det T_k, \qquad x \in E_k. \tag{Jacobian control}$$

They are used for different sides of the area formula. Secant control estimates the image measure $\mathcal{H}^n(f(F))$. Jacobian control estimates the integral $\int_F Jf$.

Secant control does not, by itself, imply the Jacobian control. The reason is that the piece E_k may be very irregular, and secants with endpoints in E_k may not see every ambient direction. For example, in $\mathbb{R}^2$ let

$$f(x,y) = (x, 0), \qquad E = \{(x,0) : x \in \mathbb{R}\}, \qquad T = \operatorname{Id}.$$

For all $a, b \in E$,

$$|f(a) - f(b)| = |a - b|,$$

so the secant estimate holds perfectly on E. But

$$Df = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad Jf = 0,$$

whereas $\det T = 1$. The secants on E only see the horizontal direction; the full Jacobian sees the collapsed vertical direction. This is why the local lemma builds both conclusions from the stronger pointwise derivative comparison.

3.3 The factorization on one piece

The reason T_k is the right normalization is that on E_k the map f factors as

$$E_k \xrightarrow{T_k} T_k(E_k) \xrightarrow{\Phi_k} f(E_k),$$

where

$$\Phi_k(T_k x) = f(x).$$

The first map is **linear**, so its effect on n-volume is exact:

$$\mathcal{H}^n(T_k(F)) = (\det T_k)\mathcal{L}^n(F).$$

The second map is close to an **non-linear but bi-Lipschitz**. In particularly, it is quantitatively close to the identity in the sense needed to control Hausdorff measure:

$$\mathrm{Lip}(\Phi_k) \le t, \qquad \mathrm{Lip}(\Phi_k^{-1}) \le t.$$

Therefore

$$t^{-n}\mathcal{H}^n(T_k(F)) \le \mathcal{H}^n(f(F)) \le t^n\mathcal{H}^n(T_k(F)).$$

> **Heuristic.** The linear map T_k gives the exact volume scale $\det T_k$, and the nonlinear residual $\Phi_k = f \circ T_k^{-1}$ distorts n-measure by at most t^n.

Now we can make this technique precise in the following lemma.

> **Lemma 2** (Local linear approximation lemma)**.** Fix $t > 1$ and put
>
> $$B = \{x \in \mathbb{R}^n : Df(x) \text{ exists and } Jf(x) > 0\}.$$
>
> Then B can be covered by countably many disjoint, Borel sets E_k such that for each k there is a symmetric positive definite automorphism $T_k : \mathbb{R}^n \to \mathbb{R}^n$ such that
>
> 1. $f|_{E_k}$ is one-to-one;
> 2. (**bi-Lipschitz control**)
>
> $$\mathrm{Lip}\left((f|_{E_k}) \circ T_k^{-1}\right) \le t; \quad \mathrm{Lip}\left(T_k \circ (f|_{E_k})^{-1}\right) \le t;$$
>
> 3. (**Volumetric control**) for every $x \in E_k$,
>
> $$t^{-n}\det T_k \le Jf(x) \le t^n \det T_k.$$

Proof. Fix $t > 1$. Choose $\eta > 0$ so small that

$$t^{-1} + \eta < 1 < t - \eta.$$

Let $\mathcal{C} \subset \mathbb{R}^n$ be a countable dense set, and let $\mathcal{T}$ be a countable dense subset of the symmetric positive definite matrices.

For $c \in \mathcal{C}$, $T \in \mathcal{T}$, and $i \in \mathbb{N}$, define $E(c, T, i)$ to be the set of all $b \in B \cap B(c, 1/i)$ such that

$$(t^{-1} + \eta)|Tv| \leq |Df(b)v| \leq (t - \eta)|Tv| \qquad \forall v \in \mathbb{R}^n, \tag{3.1}$$

and

$$|f(a) - f(b) - Df(b)(a - b)| \leq \eta |T(a - b)| \qquad \forall a \in B(b, 2/i). \tag{3.2}$$

This definition asks exactly: at which points b is the metric distortion of $Df(b)$ comparable to the fixed dense matrix T, and at which scale is the Taylor approximation at b good enough to compare b with every other point in the piece? Notice that we do not require Df to be continuous on an open neighborhood, nor do we choose one full map $L : \mathbb{R}^n \to \mathbb{R}^m$ that approximates Df there. We only ask each selected point to pass the same metric test against the same T.

The sets $E(c, T, i)$ are Borel. Indeed, (3.1) can be checked on rational vectors $v \in \mathbb{Q}^n$, and (3.2) can be checked on rational points $a \in \mathbb{Q}^n$ with $|a - b| < 2/i$, then extended by continuity of f. The derivative Df is Borel on the differentiability set of a Lipschitz map.

We claim that the sets $E(c, T, i)$ cover B. Fix $b \in B$. Since $Jf(b) > 0$, the derivative $Df(b)$ has full rank. Write

$$Df(b) = O_b S_b, \qquad S_b = (Df(b)^T Df(b))^{1/2},$$

with S_b symmetric positive definite. Since

$$|Df(b)v| = |S_b v|,$$

and since the inequalities in (3.1) are open conditions on T, density of $\mathcal{T}$ gives a $T \in \mathcal{T}$ such that (3.1) holds at b.

Differentiability of f at b gives

$$|f(a) - f(b) - Df(b)(a - b)| = o(|a - b|) \qquad \text{as } a \to b.$$

Because T is invertible, $|Th| \geq \lambda_{\min}(T)|h|$, so for sufficiently small $r > 0$,

$$|f(a) - f(b) - Df(b)(a - b)| \leq \eta |T(a - b)| \qquad a \in B(b, r).$$

Choose i so large that $2/i < r$, and then choose $c \in \mathcal{C}$ with $|b - c| < 1/i$. Then $b \in E(c, T, i)$. Hence the family covers B.

Now fix one such set $E = E(c, T, i)$. If $a, b \in E$, then both are in $B(c, 1/i)$, so

$$|a - b| < 2/i.$$

Thus (3.2), applied at the base point b, gives

$$f(a) - f(b) = Df(b)(a-b) + R_{a,b}, \qquad |R_{a,b}| \le \eta |T(a-b)|.$$

Using (3.1),

$$|f(a) - f(b)| \le |Df(b)(a-b)| + |R_{a,b}| \le t|T(a-b)|,$$

and

$$|f(a) - f(b)| \ge |Df(b)(a-b)| - |R_{a,b}| \ge t^{-1}|T(a-b)|.$$

Therefore

$$t^{-1}|T(a-b)| \le |f(a) - f(b)| \le t|T(a-b)|. \tag{3.3}$$

In particular, $f|_E$ is one-to-one.

Define

$$\Phi : T(E) \to f(E), \qquad \Phi(Tx) = f(x).$$

The estimate (3.3) says exactly that

$$\operatorname{Lip}(\Phi) \le t, \qquad \operatorname{Lip}(\Phi^{-1}) \le t.$$

Equivalently,

$$\operatorname{Lip}((f|_E) \circ T^{-1}) \le t, \qquad \operatorname{Lip}(T \circ (f|_E)^{-1}) \le t.$$

It remains to get the Jacobian comparison. For $w = Tv$, (3.1) gives

$$(t^{-1} + \eta)|w| \le |Df(b)T^{-1}w| \le (t - \eta)|w|.$$

Hence every singular value of $Df(b)T^{-1}$ lies between $t^{-1} + \eta$ and $t - \eta$. Therefore

$$(t^{-1} + \eta)^n \det T \le Jf(b) \le (t - \eta)^n \det T.$$

In particular,

$$t^{-n} \det T \le Jf(b) \le t^n \det T.$$

Enumerate the countable family $E(c, T, i)$ as $E_1, E_2, \dots$ and replace it by the disjoint refinement

$$\widetilde{E}_1 = E_1, \qquad \widetilde{E}_k = E_k \setminus \bigcup_{j<k} E_j.$$

Each refined piece is a subset of one original piece, so all estimates remain true. $\square$

4 The critical set lemma

The critical set lemma handles the part where the derivative exists but the Jacobian vanishes. This is the Lipschitz version of the Sard-type idea in the notes: critical points have lower-dimensional infinitesimal images, so their image has zero n-dimensional measure.

4.1 Motivation

Let

$$Z = \{x : Df(x) \text{ exists and } Jf(x) = 0\}.$$

If $x \in Z$, then $Df(x)$ has rank $< n$. Thus $Df(x)B(0, r)$ is contained in an at-most-$(n-1)$-dimensional plane. Differentiability says

$$f(x + h) = f(x) + Df(x)h + o(|h|).$$

Consequently, for small r, the image $f(B(x, r))$ lies in a thin tube of radius εr around a lower-dimensional disk of size r.

The key estimate is that such a tube has n-dimensional content of order

$$\varepsilon r^n,$$

not merely r^n. The factor ε can be made arbitrarily small. A Vitali covering argument then covers the critical set by such balls and lets $\varepsilon \downarrow 0$.

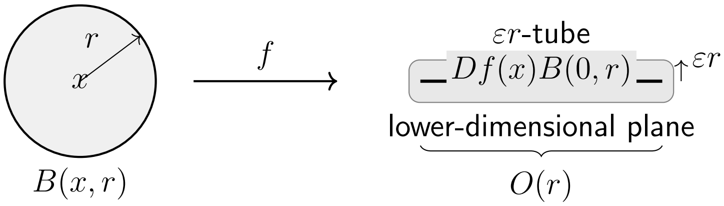


Figure 1: At a point with $Jf = 0$, a small ball maps into a thin tube around a lower-dimensional image. The tube can be covered by about $\varepsilon^{-(n-1)}$ balls of radius εr, so its n-content is $\lesssim \varepsilon r^n$.

Lemma 3 (Critical set lemma). Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz and let

$$Z = \{x \in \mathbb{R}^n : Df(x) \text{ exists and } Jf(x) = 0\}.$$

Then

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

Proof. It suffices to prove $\mathcal{H}^n(f(Z \cap B_R)) = 0$ for every $R > 0$.

Fix $\varepsilon \in (0, 1)$ and $x \in Z$. Since $Jf(x) = 0$, the linear map

$$L_x := Df(x) : \mathbb{R}^n \to \mathbb{R}^m$$

has rank at most $n - 1$. Hence $L_x(B(0, r))$ lies in an at-most-$(n - 1)$-dimensional plane and has diameter at most $2(\operatorname{Lip} f)r$.

By differentiability at x, there is $\rho_x > 0$ such that whenever $0 < r < \rho_x$,

$$f(B(x,r)) \subset f(x) + L_x(B(0,r)) + B(0,\varepsilon r). \tag{4.1}$$

The set on the right is an εr-neighborhood of an at-most-$(n-1)$-dimensional set of diameter $O(r)$. It can be covered by at most

$$C\varepsilon^{-(n-1)}$$

balls of radius $C\varepsilon r$. Therefore

$$\mathcal{H}^n_\infty(f(B(x,r))) \le C\varepsilon^{-(n-1)}(C\varepsilon r)^n \le C\varepsilon r^n, \tag{4.2}$$

where C depends only on n, m, and $\operatorname{Lip} f$.

Now cover $Z \cap B_R$ by balls $B(x, r_x)$ with $x \in Z \cap B_R$, $r_x < \rho_x/5$, and $r_x < \delta$. By the Vitali covering theorem, choose a disjoint subcollection $B_i = B(x_i, r_i)$ such that

$$Z \cap B_R \subset \bigcup_i B(x_i, 5r_i).$$

Applying (4.2) to the balls $B(x_i, 5r_i)$ gives

$$\mathcal{H}^n_{C\delta}(f(Z \cap B_R)) \le \sum_i \mathcal{H}^n_\infty(f(B(x_i, 5r_i))) \le C\varepsilon \sum_i r_i^n.$$

The balls B_i are disjoint and, for δ small, lie inside a fixed ball depending only on R. Hence

$$\sum_i r_i^n \le C_R.$$

Thus

$$\mathcal{H}^n_{C\delta}(f(Z \cap B_R)) \le C_R \varepsilon.$$

Letting $\delta \downarrow 0$ and then $\varepsilon \downarrow 0$ yields

$$\mathcal{H}^n(f(Z \cap B_R)) = 0.$$

Finally, $Z = \bigcup_{R=1}^\infty (Z \cap B_R)$, so $\mathcal{H}^n(f(Z)) = 0$. □

5 Finishing the area formula

We now assemble the proof using the three-piece strategy.

Let $A \subset \mathbb{R}^n$ be Borel. Split

$$A = A_{\text{good}} \sqcup A_{\text{crit}} \sqcup A_{\text{bad}}$$

as above. Rademacher and the Lipschitz null-image lemma handle A_{bad}, and the critical set lemma handles A_{crit}. Therefore it remains to prove

$$\int_{A_{\text{good}}} Jf(x)\,dx = \int_{\mathbb{R}^m} N(f, A_{\text{good}}, y)\,d\mathcal{H}^n(y).$$

Fix $t > 1$. Apply the local linear approximation lemma to $B = \{Df \text{ exists}, Jf > 0\}$ and refine the resulting cover to a disjoint family. Intersecting with A_{good}, write

$$A_{\text{good}} = \bigsqcup_{k=1}^{\infty} F_k,$$

where each F_k lies in one linearization piece with associated symmetric positive definite map T_k.

One-piece estimate through the factorization

Fix k and let $F \subset F_k$ be Borel. Define

$$\Phi_k : T_k(F_k) \to f(F_k), \qquad \Phi_k(T_k x) = f(x).$$

The local lemma says

$$\operatorname{Lip}(\Phi_k) \le t, \qquad \operatorname{Lip}(\Phi_k^{-1}) \le t.$$

Using the factorization

$$F \xrightarrow{T_k} T_k(F) \xrightarrow{\Phi_k} f(F),$$

we estimate the image measure in two stages. First,

$$\mathcal{H}^n(T_k(F)) = (\det T_k)\mathcal{L}^n(F).$$

Second, Lipschitz maps increase $\mathcal{H}^n$ by at most the n-th power of their Lipschitz constant, so

$$\mathcal{H}^n(f(F)) \le t^n \mathcal{H}^n(T_k(F)) = t^n (\det T_k)\mathcal{L}^n(F), \tag{5.1}$$

and the inverse Lipschitz bound gives

$$(\det T_k)\mathcal{L}^n(F) = \mathcal{H}^n(T_k(F)) \le t^n \mathcal{H}^n(f(F)). \tag{5.2}$$

On the other hand, the Jacobian comparison gives

$$t^{-n}(\det T_k)\mathcal{L}^n(F) \le \int_F Jf(x)\,dx \le t^n (\det T_k)\mathcal{L}^n(F). \tag{5.3}$$

Combining (5.1)–(5.3), we get

$$t^{-2n}\mathcal{H}^n(f(F)) \le \int_F Jf(x)\,dx \le t^{2n}\mathcal{H}^n(f(F)). \tag{5.4}$$

This is the whole quantitative heart of the proof.

Summing the sheets

Apply (5.4) with $F = F_k$ and sum over k. Since the F_k are disjoint,

$$\sum_k \int_{F_k} Jf = \int_{A_{\text{good}}} Jf.$$

Since $f|_{F_k}$ is one-to-one,

$$N(f, A_{\text{good}}, y) = \sum_{k=1}^{\infty} \mathbf{1}_{f(F_k)}(y)$$

as an extended-valued function. The images $f(F_k)$ are measurable; for instance, this follows from the Lusin-Souslin theorem because f is continuous and one-to-one on the Borel set F_k. Therefore monotone convergence gives

$$\int_{\mathbb{R}^m} N(f, A_{\text{good}}, y)\, d\mathcal{H}^n(y) = \sum_{k=1}^{\infty} \mathcal{H}^n(f(F_k)).$$

Summing (5.4) gives

$$t^{-2n} \int_{\mathbb{R}^m} N(f, A_{\text{good}}, y)\, d\mathcal{H}^n(y) \le \int_{A_{\text{good}}} Jf(x)\, dx \le t^{2n} \int_{\mathbb{R}^m} N(f, A_{\text{good}}, y)\, d\mathcal{H}^n(y).$$

This is true for every $t > 1$. Letting $t \downarrow 1$ gives

$$\int_{A_{\text{good}}} Jf(x)\, dx = \int_{\mathbb{R}^m} N(f, A_{\text{good}}, y)\, d\mathcal{H}^n(y).$$

Adding back A_{crit} and A_{bad}, both of which contribute zero to both sides, proves the area formula on A.

> **Remark (Injective case).** If f is one-to-one on A, then the sets $f(F_k)$ are disjoint up to null sets, and the multiplicity function is simply $\mathbf{1}_{f(A)}$ on the image. The argument above reduces to
>
> $$t^{-2n}\mathcal{H}^n(f(A_{\text{good}})) \le \int_{A_{\text{good}}} Jf \le t^{2n}\mathcal{H}^n(f(A_{\text{good}})),$$
>
> followed by $t \downarrow 1$ and then adding back the null-image pieces.

6 Weighted version

The weighted area formula follows from the set version. For every nonnegative Borel g,

$$\int_A g(x) Jf(x)\, dx = \int_{\mathbb{R}^m} \sum_{x \in A \cap f^{-1}\{y\}} g(x)\, d\mathcal{H}^n(y).$$

First take $g = \mathbf{1}_C$, then finite simple functions, and then pass to nonnegative Borel g by monotone convergence.

A Applications of the Lipschitz Area Formula

These examples are meant to illustrate the necessity of a change of variables formula for maps with Lipschitz, but not necessarily C^1 regularity. The Area Formula holds for larger classes of functions, but the Lipschitz case often suffices and has a simpler proof.

A.1 Nonlinear elasticity: no interpenetration and cavitation

Let $\Omega \subset \mathbb{R}^n$ be a reference body and let

$$u : \Omega \to \mathbb{R}^n$$

be a deformation. The point $x \in \Omega$ labels a material particle, and $u(x)$ is its physical position. Energies in nonlinear elasticity often have the form

$$E(u) = \int_\Omega W(Du(x))\, dx,$$

where W penalizes stretching, compression, and orientation reversal.

If u is smooth, orientation-preserving, and injective, then

$$\mathcal{L}^n(u(\Omega)) = \int_\Omega \det Du(x)\, dx.$$

But variational deformations are often not smooth. If u is Lipschitz, the area formula gives

$$\int_\Omega |\det Du(x)|\, dx = \int_{\mathbb{R}^n} N(u, \Omega, y)\, dy.$$

When $\det Du \geq 0$ a.e., this says

$$\int_\Omega \det Du(x)\, dx = \int_{u(\Omega)} N(u, \Omega, y)\, dy.$$

The multiplicity $N(u, \Omega, y)$ is the number of material particles occupying the same physical point y. Thus the right-hand side is physical volume counted with overlap.

This is the analytic mechanism behind the Ciarlet-Necas type condition

$$\int_\Omega \det Du(x)\, dx \leq \mathcal{L}^n(u(\Omega)).$$

Since $N(u, \Omega, y) \geq 1$ for a.e. $y \in u(\Omega)$, the area formula shows that this inequality forces

$$N(u, \Omega, y) = 1 \qquad \text{for a.e. } y \in u(\Omega).$$

So the condition prevents bulk interpenetration of matter.

Cavitation is different: the deformation may open a hole in the image without making two pieces of matter overlap. The area formula separates these phenomena. Overlap is detected by multiplicity $N > 1$, while a cavity appears as a region missing from $u(\Omega)$.

A.2 Rectifiable currents, varifolds, and minimal surfaces

A nonsmooth m-dimensional surface in $\mathbb{R}^N$ is often represented by countably many Lipschitz charts

$$f_i : E_i \subset \mathbb{R}^m \to \mathbb{R}^N.$$

On an injective chart, the area formula gives

$$\mathcal{H}^m(f_i(E_i)) = \int_{E_i} J_m f_i(x)\, dx.$$

Without injectivity,

$$\int_{E_i} J_m f_i(x)\, dx = \int_{\mathbb{R}^N} N(f_i, E_i, y)\, d\mathcal{H}^m(y).$$

The parametrized area counts sheets with multiplicity.

This is essential for integer rectifiable currents. Roughly, if a current is represented by oriented Lipschitz sheets with integer multiplicities, its mass is computed by summing chartwise Jacobian integrals. Schematically,

$$\mathbf{M}(T) = \sum_i \int_{E_i} \theta_i(x) J_m f_i(x)\, dx,$$

after arranging the charts so that they represent the current without unwanted double-counting. The area formula is the bridge between the parametrized set $E_i \subset \mathbb{R}^m$ and the geometric object in $\mathbb{R}^N$.

This matters for minimal surfaces because minimizers may not be globally smooth. In high dimensions, area-minimizing hypersurfaces can have singularities. GMT works with rectifiable sets, currents, or varifolds instead of smooth manifolds. The area formula justifies computing their mass through Lipschitz parametrizations and approximate tangent data.

A.3 Optimal transport and nonsmooth change of variables

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ transport a density ρ_0 to a density ρ_1:

$$T_\#(\rho_0\, dx) = \rho_1\, dy.$$

If T were a smooth diffeomorphism, then

$$\rho_0(x) = \rho_1(T(x)) |\det DT(x)|.$$

In optimal transport, a Brenier map has the form $T = \nabla\varphi$ for a convex potential φ, and differentiability is naturally an a.e. statement unless stronger regularity is proved.

When T is Lipschitz, the weighted area formula gives the nonsmooth replacement. For nonnegative Borel h,

$$\int h(x) |\det DT(x)|\, dx = \int \sum_{x \in T^{-1}\{y\}} h(x)\, dy.$$

Choosing

$$h(x) = \frac{\rho_0(x)}{|\det DT(x)|}$$

on the set where $|\det DT| > 0$, one obtains formally

$$\rho_1(y) = \sum_{x \in T^{-1}\{y\}} \frac{\rho_0(x)}{|\det DT(x)|}.$$

If T is one-to-one a.e., the sum has one term and the usual change-of-variables relation follows.

The same formula explains folding in Lagrangian maps for transport equations. If several labels arrive at the same Eulerian point, the density contributions add; that addition is precisely the multiplicity in the area formula.

References

[1] Lawrence C Evans and Ronald F Gariepy. *Measure Theory and Fine Properties of Functions, Revised Edition*. CRC Press, Boca Raton, FL, 2015.

[2] Francesco Maggi. *Sets of Finite Perimeter and Geometric Variational Problems: an Introduction to Geometric Measure Theory*. Cambridge University Press, 2012.