The Co-Area Formula and its Applications

Alex Havrilla

1st May 2020

1 Introduction

The co-area formula, along with the area formula as its counterpart, serve as useful tools for computing integrals, allowing us to restructure computations by integrating over level sets of a lipschitz function. In particular for lipschitz $f : \mathbb{R}^m \to \mathbb{R}^n$ with $m \geq n$ and $A \subseteq \mathbb{R}^m$ lebesgue measurable

$$\int_A J_n f(x) d\mathcal{L}^m x = \int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap f^{-1}(y)) d\mathcal{L}^n y$$

where $\mathcal{L}^m$ is the m dimensional lebesgue measure and $\mathcal{H}^{m-n}$ is the $m-n$ dimensional hausdorff measure. $J_n f$ is the n dimensional jacobian of f. Recall this is defined as

$$J_n f(a) = \sqrt{det(Df(a) \cdot Df(a)^t)}$$

and can be thought of as some kind of as a scaling quantity describing how much the differential Df dilates a parallelpiped when pushed forward. Note when $m = n$ this corresponds to the better known definition, ie. the determinant of Df.

To provide some motivation, we will first consider the case for an orthogonal projection $f : \mathbb{R}^m \to \mathbb{R}^n$ with $n < m$ and $m = n + k$ for some $k \in \mathbb{N}$. We know that $f^{-1}(0)$ and more generally $f^{-1}(y), y \in \mathbb{R}^n$ will be a k dimensional subspace of $\mathbb{R}^m$. Since we are othogonally projecting we can interpret thse $f^{-1}(y)$ as translations of $f^{-1}(0)$. We can write

$$\mathbb{R}^m = \bigcup_{y \in \mathbb{R}^n} f^{-1}(y)$$

Then by fubini's theorem we get for any lebesgue mesaurable $A \subseteq \mathbb{R}^m$

$$\int_{\mathbb{R}^n} \mathcal{H}^k(p^{-1}(y) \cap A) d\mathcal{L}^n(y) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^k} 1_{p^{-1}(y) \cap A} d\mathcal{H}^k d\mathcal{L}^n(y) = \int_{\mathbb{R}^m} 1_A d\mathcal{H}^m = \mathcal{H}^m(A)$$

ie. we measure A in $\mathbb{R}^m$ by "summing across" intersections with k dimensional hyperplanes, level sets of the projection. This gives the co-area formula for the simple case of orthogonal projections, but does a good job of motivating the more general idea. Note that $J_n f(x) = 1$ for the orthogonal projection.

This also shows the co-area formula a direct generalization of fubini's theorem since we have the equality

$$\mathcal{L}^m(A) = \int_{\mathbb{R}^m} 1_A d\mathcal{L}^m = \int_{\mathbb{R}^n} \int_{\mathbb{R}^k} 1_A d\mathcal{L}^k d\mathcal{L}^n$$

This result implies fubini via approximation of an integrable f by approximation of simple functions.

2 Some Auxillary Results

Before we can begin proving the co-area formula we need to present some useful facts that will be relied upon.

Fundamentally related to the co-area formula is the area formula, which deals with maps into higher dimensions:

Theorem 1. *If $f : \mathbb{R}^m \to \mathbb{R}^n$ is a lipschitz function and $m \leq n$ then*

$$\int_A J_m f(x) d\mathcal{L}^m x = \int_{\mathbb{R}^n} card(A \cap f^{-1}(y)) d\mathcal{H}^m y$$

The structures of the proofs of both the area and co-area formula closely parallel each other, first examining the setup for linear maps, then extending via approximation and finally to lipschitz functions. We will leverage the area formula to help prove key estimates for the co-area formula.

Theorem 2. *A function $f : \mathbb{R}^m \to \mathbb{R}$ is approximately continuous if for almost every $x \in \mathbb{R}^m$, and every $\epsilon > 0$ we have*

$$0 = \lim_{r \to 0^+} \frac{\mathcal{L}^m(\{y : |f(y) - f(x)| > \epsilon\} \cap B(x,r))}{\mathcal{L}^m(B(x,r))}$$

ie. the set has 0 density. Then if a function is lebesgue measurable then it is approximately continuous.

Theorem 3. ***Polar Decomposition***

1. *If $m \leq n$ and $T : \mathbb{R}^m \to \mathbb{R}^n$ linear then there exists symmetric linear map $S : \mathbb{R}^m \to \mathbb{R}^n$ and orthogonal linear map $U : \mathbb{R}^m \to \mathbb{R}^n$ s.t. $T = U \circ S$*

2. *If $m \geq n$ with T as above then we have $S : \mathbb{R}^n \to \mathbb{R}^n$ symmetric and orthogonal linear $U : \mathbb{R}^n \to \mathbb{R}^m$ s.t. $T = S \circ U^t$*

We will need that linear maps can be decomposed into compositions of symmetric and orthogonal linear transformations. This will be useful in calculating measures of level sets. It is useful to note in the case $U \circ S = T = Df(a)$ that the determinant of S gives the jacobian of f:

$$det(S) = \sqrt{det(S^t U^t U S)} = \sqrt{det(Df(a)^t Dfa)} \implies$$

$$J_m f(a) = det(S)$$

Recall Luisin's and Egoroff's theorems:

Theorem 4. ***Luisin's Theorem***

Let m a Radon measure on locally compact Hausdorff space X. If $f : X \to \mathbb{R}$ measurable with A a compact set with finite measure and $\epsilon > 0$, then $\exists$ compact C s.t. $m(A \backslash C) < \epsilon$ and f restricted to C continuous.

Theorem 5. ***Egoroff's Theorem***

For m a finite measure on X, if measurable functions $f_i : X \to \mathbb{R}$, $i \geq 1$ converge almost everywhere to $g : X \to \mathbb{R}$, for $\epsilon > 0$ $\exists$ an m-measurable set B s.t. $m(X \backslash B) < \epsilon$ and the f_i converge to g uniformly on B.

The pair will be useful as we are able to assume a lipschitz f is has without loss of generality continuous and uniformly converging differential on a measurable set A.

3 Proving the Co-area Formula

The following draws from [3] with notes from [7].

Theorem 6. ***The Co-area Formula***

Suppose $f : \mathbb{R}^m \to \mathbb{R}^n$ is lipschitz and $m \geq n$. Then we have for every lebesgue measurable $A \subseteq \mathbb{R}^m$

$$\int_A J_n f(x) d\mathcal{L}^m x = \int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap f^{-1}(y)) d\mathcal{L}^n y$$

Via approxmation arguments by simple functions this generalizes for arbitrary integrable $g : \mathbb{R}^m \to \mathbb{R}^n$

$$\int_{\mathbb{R}^m} g J_n f(x) d\mathcal{L}^m x = \int_{\mathbb{R}^n} g \mathcal{H}(f^{-1}(y)) d\mathcal{L}^n y$$

The proof is broken down into several parts. First we prove the almost everywhere differentiability of lipschitz mappings and state two theorems grounding the integrability of the right hand side. Of particular interest in this section is the proof that lipschitz functions are almost everywhere differentiable. Then we prove a key estimates in the cases when the differential of f is closely approximated by linear functions. We finishing by reducing arbitrary lipschitz functions to this case, handling the possiblity $Jf = 0$ everywhere on our set A separately.

3.1 Lipschitz Mappings

These results verify integrability of the right hand side of the coarea formula and differentiability of lipschitz functions. Recall a lipschitz mapping is s.t. $|f(x) - f(y)| \leq C_f |x - y|$ for some universal lipschitz constant C_f.

Theorem 7. ***Rademacher's Theorem***

$f : \mathbb{R}^m \to \mathbb{R}^n$ lipschitz is differentiable $\mathcal{L}^m$ almost everywhere and the differential is measurable.

Proof. This will follow from the almost everywhere differentiability of 1 dimensional lipschitz $f : \mathbb{R} \to \mathbb{R}$. Without loss of generality suppose target $n = 1$(we may argue coodinate wise). Induct on m.

The case $m = 1$ is a classical result. We sketch the proof. For a full result see [7]. The bulk of the work comes from proving monotone $f : \mathbb{R} \to \mathbb{R}$ is differentiale almost everywhere. To do so we introduce the notion of *Dini Derivatives*:

1. $\overline{D^+} f(x) := \limsup_{h \to 0^+} \frac{f(x+h)-f(x)}{h}$
2. $\underline{D^+} f(x) := \liminf_{h \to 0^+} \frac{f(x+h)-f(x)}{h}$
3. $\overline{D^-} f(x) := \limsup_{h \to 0^-} \frac{f(x+h)-f(x)}{h}$
4. $\underline{D^-} f(x) := \liminf_{h \to 0^-} \frac{f(x+h)-f(x)}{h}$

Our function will be differentiable at $x \in \mathbb{R}$ when the Dini Derivatives are equal and finite. To show this we prove the One-sided Hardy-Littlewood inequality for f: For $f : [a, b] \to \mathbb{R}$ contonuous monotone non-decreasing and $\lambda > 0$ we have

$$\mathcal{L}(\{x \in [a, b] : \overline{D^+} f(x) \geq \lambda\}) \leq \frac{f(b) - f(a)}{\lambda}$$

and similarly for the other dini derivatives. Once done sending $\lambda \to \infty$ guarnantees finiteness almost everywhere. Conclude by showing $\overline{D_+} f \leq \underline{D_-} f$. This is done by showing for every pair $0 < r < t$ of numbers our set

$$E_{r,t} = \{x \in \mathbb{R} : \overline{D_+} f(x) > t > r > D_- f(x)\}$$

is null. Then letting $r, t \in \mathbb{Q}$ and using density while taking countale unions shows over all of $\mathbb{R}$ this property null.

Note the the above strategy only works in the continuos case, more arguments(but not too many more) are needed to extended to the discontinuous. Once this is accomplished the result follows smoothly, as we may write a function of bounded variation as the difference of two bounded monotone functions. In turn we note lipschitz functions are of locally bounded variation, ie. have bounded variation on every compact interval. Hence almost everywhere differentiability extends from monotone functions to bounded variation type to lipschitz type. This concludes our discussion of the one dimensional case.

Now suppose $m \geq 2$ and assume the induction hypothesis. Since f absolutely continuous we know all m partial derivatives exist almost everywhere and are measurable. We must extract the differential. Let $(y_0, z_0) = x_0 \in \mathbb{R}^{m-1} \times \mathbb{R}$. We know as a function of y_0 f is differentiable. Without loss of generality we suppose $x_0 = 0$ and partial derivatives at x_0 are 0. Compute for each $\epsilon > 0$ an $r > 0$ s.t.

$$|f(y, z)| \leq \epsilon r$$

when $|(y, z)| < r$. Knowing this we can conclude differentiability by applying the inequality in our limit definition.

Set $\epsilon > 0$. For $z = 0$ we can choose $r_0 > 0$ s.t. $|y| < r_0 \implies |f(y,0)| \leq \epsilon|y|$ via lipschivity. We also want the set of the points in the cube $[-r,r]^m$ for which the *mth* partial is large to be small. In particular for we may decrease $r < r_0$ so that this mth partial is greater than ϵ no more than on a set of $\epsilon^m r^m$ measure. Note this comes from the *approximate continuity*(see (2)) of the partials, which is guaranteed by lebesgue measurability.

So inside an $m-1$ dimensional cube $\prod_{i=1}^{m-1}[y_i - \epsilon r, y_i]$ we know the mth partial bounded by ϵ except at a vertical line passing through some y' in the m-1 cube(representing the mth axis). Then via FTC

$$|f(y',z)| \leq |f(y',0)| + \int_0^z |\frac{\partial f}{\partial z}(y',\xi)d\mathcal{L}\xi| \leq \epsilon|y'| + \epsilon|z| + m\epsilon r$$

by the above estimates. Note we bound $\int_0^z |\frac{\partial f}{\partial z}(y',\xi)d\mathcal{L}\xi|$ by $\epsilon|z| + m\epsilon r$.

So

$$|f(y,z)| \leq |f(y',z)| + m\sqrt{m-1}\epsilon r \leq (2 + (1 + \sqrt{m-1}m))\epsilon r$$

where the first bound comes from the regularity of the well-behavedness of m-1st many derivatives. This concludes the proof.

□

Lemma 8. *Suppose $0 \leq n \leq m < \infty$. Then we have some $C_{m,n}$ s.t. if $f : \mathbb{R}^m \to \mathbb{R}^n$ lipschitz and $A \subseteq \mathbb{R}^m$ is measurable then*

$$\int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap f^{-1}(y))d\mathcal{L}^n y \leq C_{m,n} C_f^n \mathcal{L}^m(A)$$

Proof. For proof see [3]. □

This will be used to upper bound the right hand side of the co-area formula to produce an estimate making it arbitrarily small should the jacobian vanish everywhere.

Another lemma showing measurability of this evaluation:

Lemma 9. *Let $f : \mathbb{R}^m \to \mathbb{R}^n$ lipschitz. Then*

$$y \mapsto \mathcal{H}^n(A \cap f^{-1}(y))$$

is $\mathcal{H}^m$ measurable

Proof. For proof see [3]. □

This ensures the integral on the right hand side of the co-area formula is well-defined.

We concludes our look at the measurability/differentiability of lipschitz mappings and associated quantities. Having these tools, in particular Rademacher's theorem, we may prove the co-area formula.

3.2 Key Estimates

We begin by first considering functions whose jacobian is close approximated by orthogonal functions. This will prove the co-area formula for functions whose jacobians are closely approximated by orthogonal transformations. This is described by the following

Lemma 10. *$f : \mathbb{R}^m \to \mathbb{R}^n$. $m > n$ with $U : \mathbb{R}^n \to \mathbb{R}^m$ orthogonal and $0 < \epsilon < 1/2$. If $A \subseteq \mathbb{R}^m$ lebesgue measurable s.t.*

1. *$Df(a)$ exists for $a \in A$*
2. *$||Df(a) - U^t|| < \epsilon$*
3. *$|f(y) - f(a) - \langle Df(a), y - a\rangle| < \epsilon|y - a|$ with $y, a \in A$*

then we have the main estimate for the co-area formula:

$$
\begin{aligned}
(1 - 2\epsilon)^m \int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap f^{-1}(y))d\mathcal{L}^n y &\leq \int_A J_n f(a)d\mathcal{L}^m a \\
&\leq \int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap f^{-1}(y))d\mathcal{L}^n y
\end{aligned}
$$

Proof. Set $V : \mathbb{R}^{m-n} \to \mathbb{R}^m$ orthogonal s.t. $ker(U^t)$,$ker(V^t)$ are orthogonal complements. Set $F : \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^{m-n}$ via $F(x) = (f(x), V^t(x))$. $\Pi : \mathbb{R}^n \times \mathbb{R}^{m-n} \to \mathbb{R}^n$ projects onto the $\mathbb{R}^{m-n}$ factor. Then $J_m F = J_n f$ since V^t orthogonal, causing its contribution to the determinant to be the identity.

We show $F|_A$ is injective. Once this is done we know it is full rank and can apply the area formula since $F : \mathbb{R}^m \to \mathbb{R}^m$. Let $a, y \in A \cap f^{-1}(z)$. Then $F(a) = (f(a), V^t(a)) = (z, V^t(a))$ and $F(y) = (f(y), V^t(y)) = (z, V^t(y))$. Then compute

$$|F(a) - F(y)| = |V^t(a) - V^t(y)|$$

We know V^t shrinks the norm yielding

$$|F(a) - F(y)| \leq |a - y|$$

and also

$$
\begin{aligned}
|U^t(y - a\rangle)| \leq& |Df(a)(y - a\rangle)| + ||Df(a) - U^t|||y - a| \\
& |f(y) - f(a) - Df(a)(y - a)| + ||Df(a) - U^t|||y - a| \\
& < 2\epsilon|y - a|
\end{aligned}
$$

via our second and third assumptions. Also via orthogonal decomposition

$$|y - a|^2 = |V^t(a) - V^t(y)|^2 + |U^t(y - a)|^2$$

giving $|V^t(a) - V^t(y)|^2 \geq |y - a|^2(1 - 4\epsilon^2)$ when combined with the above. We conclude

$$\sqrt{1-4\epsilon^2}|y-a| \leq |F(y)-F(a)| \leq |y-a|$$

This shows injectivity and provides bounds. Since $F|_A$ injective we can apply the area formula to get

$$\mathcal{L}^m(F(A)) = \int_A J_m F d\mathcal{L}^m = \int_A J_n f d\mathcal{L}^m$$

Applying fubini

$$\begin{aligned}\int_A J_n f d\mathcal{L}^m &= \int_{\mathbb{R}^n} \mathcal{H}^{m-n}[F(A) \cap \Pi^{-1}(z)] d\mathcal{L}^n z \\ &= \int_{\mathbb{R}^n} \mathcal{H}^{m-n}[F(A \cap f^{-1}(z))] d\mathcal{L}^n z\end{aligned}$$

where we get the last equality via the set equality $F(A) \cap \Pi^{-1}(z) = F(A \cap f^{-1}(z))$.

We get to the desired inequality by applying the estimates derived above to $F(A \cap f^{-1}(z))$ pointwise in the integral.

□

Having proved the result when the differential is close to an orthogonal map, we now prove the case when the differential is close to an arbitrary linear map. Once this is done we will be ready to tackle the case of arbitrary lipschitz functions directly.

Lemma 11. *$f : \mathbb{R}^m \to \mathbb{R}^n$ and suppose $m > n$ with linear $T : \mathbb{R}^m \to \mathbb{R}^n$ full rank with $0 < \epsilon < 1/2$. If $A \subseteq \mathbb{R}^m$ measurable s.t.*

1. *$Df(a)$ exists for $a \in A$*
2. *$||Df(a) - T|| < \epsilon$ for $a \in A$*
3. *$|f(y) - f(a) - Df(a)(y-a)| < \epsilon|y-a|$ for $y, a \in A$*

then

$$\begin{aligned}(1-2\epsilon)^m \int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap f^{-1}(y)) d\mathcal{L}^n y &\leq \int_A J_n f(a) d\mathcal{L}^m as \\ &\leq \int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap f^{-1}(y)) d\mathcal{L}^n y\end{aligned}$$

Proof. We can use polar decomposition from 3 to write $T = S \circ U^t$ where S symmetric from $\mathbb{R}^n \to \mathbb{R}^n$ and $U : \mathbb{R}^n \to \mathbb{R}^m$ orthogonal. Set $g = S^{-1} \circ f$. Then we can see $||Dg(a) - U^t|| < \epsilon$ as $Dg(a) - U^t = SDg(a) - f$.

Using the above lemma we get

$$\begin{aligned}(1-2\epsilon)^m \int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap g^{-1}(z)) d\mathcal{L}^n z \leq \int_A J_n g(a) d\mathcal{L}^m a \leq \\ \leq \int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap g^{-1}(z)) d\mathcal{L}^n z\end{aligned}$$

For $y = S(z)$, $A \cap g^{-1}(z) = A \cap f^{-1}(y)$ by construction and we can use change of variables to get

$$\int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap g^{-1}(z)) J_n S d\mathcal{L}^n z = \int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap f^{-1}(y)) d\mathcal{L}^n y$$

via chain rule we know $J_n S J_m g = J_n f$ so

$$\int_A J_n g J_m g(a) d\mathcal{L}^m a = \int_A J_n f(a) d\mathcal{L}^m a$$

Thus multiplying

$$(1-2\epsilon)^m \int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap g^{-1}(z)) d\mathcal{L}^n z \leq \int_A J_m g(a) d\mathcal{L}^m a \leq$$
$$\leq \int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap g^{-1}(z)) d\mathcal{L}^n z$$

through by $J_n S$ we obtain the desired inequality.

□

Note that these inequalities sandwhich $\int_A J_n f$ which shows the equality as we let find better approximations. Thus we have completed a proof for functions closely approximated by linear functions.

3.3 Bringing it Together

We now prove the general co-area formula for lipschitz mappings. We start by arguing we may assume f C^1 via approximation arguments. It is possible $Df(a) = 0$ everywhere, so we address this case too.

Proof. If f lipschitz let C_f be the function's lipschitz constant.

It suffices to assume $Df(a)$ exists at every point $a \in A$ by 7. We also assume $J_n f(a) > 0$ wihout loss of generality(if not simply break the integral apart into the positive and negative sets). We may assume $Df(a)$ is the restriction to A of a continuous function Luisin 4. By Egoroff 5 we suppose

$$\frac{|f(y) - f(a) - \langle Df(a), y - a \rangle|}{|y - a|}$$

converges uniformly to 0 as $y \in A$ approaches $a \in A$. For small enough A we know conditions are satisfied for lemma 11 under these assumptions via lipschivity of f. This is sufficient via linearity of the integral(we prove the statement on small subsets). To see the approximation can be made to work find C^1 f_ϵ and measurable with $f_\epsilon = f, Df = Df_\epsilon$ on G_ϵ. Via lipschitvity we may find a $C > 0$ s.t. $0 \leq J_n f(x) \leq CC_f$ almost everywhere(C will depend on the dimension). Then

$$|\int_A J_n f d\mathcal{L}^m - \int_{G_\epsilon} J_n f d\mathcal{L}^m| = \int_{A \setminus G_\epsilon} J_n f d\mathcal{L}^m \leq \epsilon C C_f$$

Similarly

$$
\begin{aligned}
&|\int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap f^{-1}(y)) d\mathcal{L}^n - \int_{\mathbb{R}^n} \mathcal{H}^{m-n}(G_\epsilon \cap f^{-1}(y)) d\mathcal{L}^n| \\
&= \int_{\mathbb{R}^n} \mathcal{H}^{m-n}((A \backslash G_\epsilon) \cap f^{-1}(y)) d\mathcal{L}^n \le D\epsilon
\end{aligned}
$$

for some constant D via 8.

Hoever it may be the that case $J_n f = 0$ on all of A. We must show RHS is 0. We argue by perturbing f by a small amount ϵ. Let $f_\epsilon : \mathbb{R}^{m+n} \to \mathbb{R}^n$ via $(x, y) \mapsto f(x) + \epsilon x$. Consider the set

$$
S = A \times [-1, 1]^n \subseteq \mathbb{R}^m \times \mathbb{R}^n
$$

We know $\mathcal{L}^{m+n}(S) = 2^n \mathcal{L}^m(A)$. Also $J_n f_\epsilon \le \epsilon[\epsilon + C_f]^{n-1}$ as we can use the lipschitz constant again to bound the jacobian of f.

Also

$$
\int_{A \times [-1,1]^n} J_n f_\epsilon d\mathcal{L}^{m+n} = \int_{\mathbb{R}^n} \mathcal{H}^m[(A \times [-1, 1]^n) \cap f_\epsilon^{-1}(z)] d\mathcal{L}^n z
$$

via our previous analysis.

We also have via 8

$$
\begin{aligned}
&C_{m,n} \mathcal{H}^m[(A \times [-1, 1]^n) \cap f_\epsilon^{-1}(z)] \\
&\quad \ge \int_{\mathbb{R}^n} \mathcal{H}^{m-n}[(A \times [-1, 1]^n) \cap f_\epsilon^{-1}(z) \cap \Pi^{-1}(y)] d\mathcal{L}^n y \\
&\quad \ge \int_{[-1,1]} \mathcal{H}^{m-n}[A \cap f^{-1}(z - \epsilon y)] d\mathcal{L}^n y
\end{aligned}
$$

Therefore

$$
\begin{aligned}
&2^n \mathcal{L}^m(A) \epsilon[\epsilon + C_f]^{n-1} \\
&\quad \ge \frac{1}{C_{m,n}} \int_{\mathbb{R}^n} \int_{[-1,1]^n} \mathcal{H}^{m-n}[A \cap f^{-1}(z - \epsilon y)] d\mathcal{L}^n y d\mathcal{L}^n z \\
&\quad = \frac{1}{C_{m,n}} \int_{[-1,1]^n} \int_{\mathbb{R}^n} \mathcal{H}^{m-n}[A \cap f^{-1}(z - \epsilon y)] d\mathcal{L}^n y d\mathcal{L}^n z
\end{aligned}
$$

which via translation invariance simplifies to

$$
\frac{2^n}{C_{m,n}} \int_{\mathbb{R}^n} \mathcal{H}^{m-n}[A \cap f^{-1}(z)] d\mathcal{L}^n z
$$

Letting $\epsilon \to 0$ finishes the proof. $\square$

4 Applications of the Co-area Formula

4.1 Common Level Sets to Integrate Over

We already mentioned the co-area formula genearlizes fubini's theorem, where we are reordering via the level sets of a projetion. Note the co-are formula includes spherical integration as a special case as well:

Proposition 1. *Let $g : \mathbb{R}^m \to \mathbb{R}$ be $\mathcal{L}^m$ integrable. Then*

$$\int_{\mathbb{R}^m} gdx = \int_0^\infty \int_{\partial B(0,r)} gd\mathcal{H}^{m-1}dr$$

Proof. Set $f(x) = |x|$ in the coarea formula. We compute

$$Df(x) = \frac{x}{|x|}$$

with

$$Jf(x) = 1$$

since $\frac{x}{|x|} \cdot \frac{x}{|x|} = 1$, yielding the desired result. □

Proposition 2. *Let $f : \mathbb{R}^m \to \mathbb{R}$ lipschitz with $essinf Jf > 0$. Suppose $g : \mathbb{R}^m \to \mathbb{R}$ $\mathcal{L}^m$ integrable. Then*

$$\int_{f>t} gdx = \int_t^\infty \int_{f=s} \frac{g}{Jf} d\mathcal{H}^{m-1}ds$$

Proof. Compute

$$\begin{aligned}\int_{f>t} gdx =& \int_{\mathbb{R}^m} 1_{f>t} \frac{g}{Jf} Jfdx \\ =& \int_{-\infty}^\infty (\int_{\partial\{f>t\}} \frac{g}{Jf} 1_{f>t} d\mathcal{H}^{m-1})ds \\ & \int_t^\infty (\int_{f=t} \frac{g}{Jf} d\mathcal{H}^{m-1})ds\end{aligned}$$

□

Here we see if our lipschitz function is real valued, we can integrate g over the level sets by simply integrating over each level set and then integrating to infinity.

4.2 Equivalence of the Isoperimetric and Sobolev Inequalities

Informally the isoperimetric inequality in $\mathbb{R}^n$ states for fixed surface area, the sphere maximizes volume. Formally we have for any set $S \subseteq \mathbb{R}^n$ with closure having finite lebesuge measure

$$n\omega_n^{\frac{1}{n}} \mathcal{L}^n(\overline{S})^{\frac{n-1}{n}} \leq M_*^{n-1}(\partial S)$$

where M_*^{n-1} is n-1 dimensional minkowski content defined b by the agreeance of

$$M^{*\mathcal{L}}(A) = \limsup_{r \to 0^+} \frac{\mathcal{L}(\{x : d(x, A) < r\})}{\omega_{n-m} r^{n-m}}$$
$$M_*^{\mathcal{L}}(A) = \liminf_{r \to 0^+} \frac{\mathcal{L}(\{x : d(x, A) < r\})}{\omega_{n-m} r^{n-m}}$$

This can be thought of as a type of volume(although it is not in general a measure).

For simplicity suppose we are on two dimensions $D \subseteq \mathbb{R}^2$. We will show equivalence of the isoperimetric inequality to the sobolev type inequality

$$\int_D |\nabla f| \geq 2\sqrt{\pi} \int_D f^2$$

for compactly supported f.

Note in two dimensions the isoperimetric inequality reduces to

$$L^2 \geq 4\pi A$$

for all simple closed curves where A area and L perimeter. Our argument follows [5].

Proof. First suppose the sobolev inequality. Suppose D has a smooth boundary. Let $\epsilon > 0$and set

$$f_\epsilon(p) = \begin{cases} 1 & d(p, \partial D) \geq \epsilon \\ d(p, \partial D)/\epsilon & d(p, \partial D) < \epsilon \end{cases}$$

We see the sobolev inequality holds for f_ϵ via approximation by smooth and compactly support functions. Letting $\epsilon \to 0$ we see f_ϵ converges pointwise to the 1_D. Further via montone convergence

$$\int_D f_\epsilon^2 \to \int_D 1 = A, \text{the area of D}$$

Setting $C_\epsilon = \{p : d(p, C) < \epsilon\}$ then

$$|\nabla f_\epsilon| = \begin{cases} 1/\epsilon & D \cap C \\ 0 & D \backslash C_\epsilon \end{cases}$$

So

$$\int_D |\nabla f_\epsilon| = \frac{Area(D \cap C_\epsilon)}{\epsilon}$$

by definition. Note that this is the Minkowski content of D in two dimensions which gives the perimeter. Hence using the sobolev inequality tells us

$$L^2 = (\int_D |\nabla f|)^2 \geq 4\pi \int_D f^2 = 4\pi A$$

when we take the limit.

Now we assume the isoperimetric inequality and show sobolev. This will be done with the co-area formula by integrating along level curves of f. Define $D(t) = \{(x, y) \in D : |f(x, y)| > t\}$ with $A(t) = Area(D(t))$ and $C(t) = \{(x, y) \in D : |f(x, y)| = t\}$ with $L(t) = length(C(t))$.

For arbitrary integrable $h(x, y)$ we know

$$\int\int_D h(x, y)|\nabla f|dxdy = \int_0^\infty \int_{C(t)} hdsdt$$

via a direct application of the co-area formula. Setting $h = 1$ we get

$$\int\int_D |\nabla f|dxdy = \int_0^\infty \int_{C(t)} dsdt = \int_0^\infty L(t)dt$$

Applying isoperimetric inequality for fixed t yields

$$\int\int_D |\nabla f|dxdy \geq 2\sqrt{\pi}\int_0^\infty \sqrt{A(t)}dt$$

The next step is to write the right hand side of our sobolev inequality similarly. Compute

$$\int\int_D f^2 dxdy = \int\int_D \int_0^{|f(x,y)|} 2tdtdxdy$$

which is equivalent to the following via fubini

$$\int_0^\infty 2t \int\int_{t \leq |f(x,y)|} 1dxdydt = \int_0^\infty 2tA(t)dt$$

Now we are almost done. We know since $A(t)$ decreasing $t\sqrt{A(t)} \leq \int_0^t \sqrt{A(t)}dt$ and thus $tA(t) \leq \frac{1}{2}\frac{d}{dt}(\int_0^t \sqrt{A(z)}dz)^2$ hence

$$\int_0^\infty 2tA(t)dt \leq (\int_0^\infty \sqrt{A(t)}dt)^2$$

Thus we conclude the our sobolev inequality:

$$(\int\int_D |\nabla f|dxdy)^2 \geq 4\pi(\int_0^\infty \sqrt{A(t)})^2 \geq 4\pi\int_0^\infty 2tA(t)dt \geq 4\pi\int\int_D f^2 dxdy$$

□

Note that via an approximation by bounded sets this generalizes to prove a sobolev this generalizes to show an equivalence over all of $\mathbb{R}^2$(and more generally $\mathbb{R}^n$). In fact the co-area formula can be generalized to $W^{1,p}(\mathbb{R}^n)$ functions(see [8]), leading to a swath of applications of the flavor shown above. Also see [1], [4].

References

[1] Bobkov, S. Houdre, C. Some Connections between Sobolev-type Inequalities and Isoperimetry.

[2] Evans, L., Gariepy R. Measure Theory and Fine Properties of Functions. CRC Press, 1992.

[3] Krantz, S., Parks, H., The Geometry of Domains in Space. Springer, 1999.

[4] Maly, J. Swanson, D. Ziemer, P. The Co-area Formula for Sobolev Mappings

[5] Osserman, R. The Ispoperimetric Inequality. https://projecteuclid.org/download/pdf_1/euclid.bams/1183541466

[6] Simon, L. Introduction to Geometry Measure Theory. https:/./web.stanford.edu/class/math285/ts-gmt.pdf

[7] Tao, T. Differentiation Theorems. https://terrytao.wordpress.com/2010/10/16/245a-notes-5-differentiation-theorems/#more-4290

[8] Walker, R. The Coarea formula for $W^{1,p}$ functions. http://www.ms.uky.edu/ rwalker/research/coarea.pdf