Allard Type Regularity Theorems for Rectifiable Varifolds

Hunter Stufflebeam

The University of Texas at Austin

Abstract

In this expository note we embark on a short tour the regularity theory of rectifiable varifolds. We primarily give a detailed accounting of Camillo De Lellis' proof of the Allard ε-regularity theorem simplified to the case of integer rectifiable varifolds with an L^∞ bound on mean curvature. In the process, cosmetic adjustments are made and the exposition is greatly expanded. We also briefly introduce the general theorem as it applies to rectifiable varifolds, as well as a recent generalization of due to Theodora Bourni and Alexander Volkmann in which the L^p bounds on mean curvature are replaced by a regularity condition on the generalized normal of the varifold.

I would like to very warmly thank two people in particular for their guidance over the last year. The first is Dr. Francesco Maggi, who introduced me to the fantastic world of geometric measure theory, supervised this work, and most importantly showed me how to start thinking more like a mathematician. I especially appreciate the warmth and kindness he showed to me throughout the whole process.

The second person I would like to thank is Dr. Salvatore Stuvard, who has become a second adviser and friend throughout my last year at UT Austin. Without his generous help and guidance, none of this would have been possible.

Contents

1 Introduction and Preliminaries

Before diving into the regularity theory of rectifiable varifolds, we first collect together some basic results in measure theory, real analysis, functional analysis, and linear algebra. Unless a proof is particular exciting or seemingly not documented elsewhere, we opt for brevity in forgoing proofs, and refer the reader to the excellent texts of Evans [Eva10], Evans and Gariepy [EG15], Maggi [Mag12], Simon [Sim14], et cetera, for the details.

1.1 Introductory Measure Theory

Borel Measures

We begin by discussing briefly the notion of measures in general. Let X be a set, and $\mathcal{P}(X)$ its power set. An **outer measure** μ on X is a map $\mu\colon \mathcal{P}(X) \to [0,\infty]$ such that the following properties hold:

- $\mu(\emptyset) = 0$;
- For every $E \subset X$ and $\{E_i\}_{i=1}^\infty \subset \mathcal{P}(X)$ with $E \subset \bigcup_{i=1}^\infty E_i$, we have $\mu(E) \leqslant \sum_{i=1}^\infty \mu(E_i)$.

If X is a set, $\Sigma \subset \mathcal{P}(X)$ is a σ-algebra, and $\mu\colon \mathcal{P}(X) \to [0,\infty]$ is an outer measure satisfying the *countable additivity* property

- For every disjoint countable collection $\{E_i\} \subset \Sigma$, $\mu\left(\bigcup_{i=1}^\infty E_i\right) = \sum_{i=1}^\infty \mu(E_i)$,

then we say that the triple (X, Σ, μ) is a measure space, and that μ is a **measure**. In any case, a set $E \subset X$ is said to be **measurable** with respect to an (outer) measure μ if

$$\mu(A) = \mu(A \cap E) + \mu(A \setminus E)$$

for every $A \subset X$. It is an elementary fact that measurable sets form a family with a good set-algebraic structure, as exhibited by the following result:

Theorem 1.1 (Caratheodory)**.** *Let μ be an outer measure on a space X, and let $\mathfrak{M}(\mu)$ be the collection of all μ-measurable subsets of X. Then $\mathfrak{M}(\mu)$ is a complete*[1] *σ-algebra, on which μ is a measure*[2]*.*

We can ask that an outer measure μ on X be determined by its action on $\mathfrak{M}(\mu)$, in the sense that if $A \subset X$ then a μ-measurable set $E \supset A$ exists such that $\mu(A) = \mu(E)$. Such outer measures are called **regular**. The following result, which allows us to pass limits through measures in certain situations, illustrates how regularity can significantly improve the behavior of a measure:

Proposition 1.1. *Let (X, μ, Σ) be a measure space, and $\{E_j\} \subset \Sigma$ a sequence of μ-measurable sets. If $E_j \subset E_{j+1}$, then $\lim_{j\to\infty} \mu(E_j) = \mu(\bigcup_{j=1}^\infty E_j)$. In case μ is regular, then in fact this holds even if the E_j are not measurable. If $E_j \supset E_{j+1}$ and $\mu(E_j) < \infty$ for some j, then $\lim_{j\to\infty} \mu(E_j) = \mu(\bigcap_{j=1}^\infty E_j)$.*

If now X is a topological space, then it is often useful to consider outer measures which play nicely with the topology of X. The family of open sets in X generates a σ-algebra $\mathcal{B}(X)$ called the **Borel σ-algebra**, and if an outer measure μ on X has the property that $\mathcal{B}(X) \subset \mathfrak{M}(\mu)$, then we call μ a **Borel** outer measure. A **Borel regular** outer measure μ on X is a Borel outer measure such that if $A \subset X$, then there exists an $E \subset \mathcal{B}(X)$ such that $\mu(A) = \mu(E)$[3].

If we add a bit more topological structure to X, then we discover that Borel regular measures have good approximation properties in terms of open, closed, and compact sets. In particular, we have the following:

[1] If Σ contains all subsets of X on which μ is trivial, then we say that Σ is **complete**.

[2] We often refer to outer measures and measures interchangeably with the understanding that the object in question will be a measure in earnest when restricted to its distinguished sigma algebra.

[3] Note that this is slightly at odds with the marginally more general definition of a regular outer measure, in that for a Borel regular outer measure we require $E \in \mathcal{B}(X)$ rather than just $E \in \mathfrak{M}(\mu)$.

Theorem 1.2. *Let X be a topological space in which every closed set is a countable intersection of open sets. Let μ be a Borel regular measure on X which is open σ-finite*[4]*. Then if $A \subset X$,*

$$\mu(A) = \inf_{\substack{U \supset A \\ U \, open}} \mu(U)$$

and if $A \in \mathfrak{M}(\mu)$,

$$\mu(A) = \sup_{\substack{C \subset A \\ C \, closed}} \mu(C)$$

If we require X to be Hausdorff and σ-compact[5], then we can replace *closed* sets with *compact* sets the second conclusion.

Before moving onto the construction of the Hausdorff and Lebesgue measures, we introduce a useful construction allowing us to localize measures. Letting X be any space, μ an outer measure on X, and $Y \subset X$ any subset, we can define the **restriction** of μ to Y, denoted by $\mu \llcorner Y$, by

$$(\mu \llcorner Y)(A) := \mu(A \cap Y). \qquad \text{for all } A \subset X$$

It so happens that, regardless of what Y was, $\mathfrak{M}(\mu) \subset \mathfrak{M}(\mu \llcorner Y)$. Moreover, the property of Borel regularity is retained under restriction, provided that the set on which we are localizing is a measurable set with finite measure. In a similar vein, when a measure only "sees" the the subsets of a fixed set, we say that the measure is supported on that set. Precisely, we say that a measure μ on a space X is **concentrated**, or **supported**, on a set $F \subset X$ provided that $\mu(X \setminus F) = 0$. In case X is a separable topological space and μ is Borel, we call the intersection of all closed sets on which μ is concentrated the **support** of μ, denoted by sptμ. Thus, sptμ is the smallest closed set on which μ is concentrated. An especially useful way to view the support of a measure is as

$$\text{spt}\mu = X \setminus \bigcup_{\substack{\mu(U)=0 \\ U \text{ open}}} U,$$

or in the case of a separable metric space X,

$$\text{spt}\mu = \{x : \mu(B_r(x)) > 0 \text{ for all } r > 0\}.$$

The Hausdorff and Lebesgue Measures

We now set to work on producing the measures which underlie all that is to follow. For the full details, see any book on measure theory such as [EG15], [Mag12], or [Sim14]. Let (X, d) be a metric space, and define for each $s \geqslant 0$ the constants

$$\omega_s := \frac{\pi^{s/2}}{\Gamma(1 + s/2)}$$

where Γ is the Gamma function defined for all $s > 0$ by

$$\Gamma(s) := \int_0^\infty t^{s-1} e^{-t} \mathrm{d}t.$$

This seems a bit unmotivated at first, but is done so that ω_s agrees with the volume of the s-dimensional unit ball in $\mathbb{R}^s$ when s is an integer[6].

Fix an $s \geqslant 0$, let $\delta \in (0, \infty]$, and define the **size δ-approximation to the s-dimensional Hausdorff measure** by

$$\mathcal{H}^s_\delta(A) := \inf_{\mathcal{F}} \sum_{F \subset \mathcal{F}} \omega_s \left(\frac{\text{diam}(F)}{2} \right)^s \qquad \text{for any } A \subset X$$

[4] A measure μ on a topological space X is said to be **open σ-finite** if X can be realized as a countable union of open sets, each with finite μ measure.

[5] A topological space is said to be **σ-compact** if it can be realized as a countable union of compact sets.

[6] By the volume of the k dimensional ball in $\mathbb{R}^k$, we mean the result of integrating the constant function 1 over the unit ball by elementary Riemann integration (not requiring a measure)–which agrees with the volume long ago determined by an argument of Archimedes, and apocryphally inscribed on his tombstone.

where the infimum is taken over all countable coverings $\mathcal{F} \subset \mathcal{P}(X)$ of the set A with $\operatorname{diam}(F) < \delta$. In case no such covering exists, we set $\mathcal{H}^s_\delta(A) = \infty$. Notice that, since $\operatorname{diam}(F) = \operatorname{diam}(\bar{F})$, we can require that all the covering sets be closed. Similarly, we can impose that the covering sets be convex, be subsets of A, et cetera. We can also require the covering sets to be open, by taking small open neighborhoods of the sets in any covering. Now, define the **s-dimensional Hausdorff measure** on X by

$$\mathcal{H}^s(A) := \sup_{\delta \in (0,\infty]} \mathcal{H}^s_\delta(A) = \lim_{\delta \searrow 0} \mathcal{H}^s_\delta(A) \qquad \text{for all } A \subset X.$$

Notice that this is well defined since $\mathcal{H}^s_\delta(A)$ is monotone decreasing in δ. Therefore, the Hausdorff measures can be thought of in the following way. We cover our set A with a countable number of sets which do not get too big, and determine the "total volume" of the covering if we had instead used "s-dimensional balls" with diameters equal to those of the covering sets. These coverings should intuitively be thought of as approximations to our set A, with smaller volumes representing increasingly better approximations. We then take the Hausdorff measure of A as the limit of these volumes under the restriction to finer and finer coverings.

The Hausdorff measures have several properties which make them well-suited for use in geometric measure theory. For instance, they appear naturally in the area and co-area formulas which we will soon see in an upcoming section. Amongst the most fundamental reasons, however, is that the Hausdorff measures are Borel measures. This can be proven starting from the definition of the Hausdorff measure as the limit of its δ-approximations, and applying the following result:

Theorem 1.3 (Caratheodory's Criterion)**.** *Let X be a metric space, μ an outer measure on X, and suppose that for all subsets $A, B \subset X$ with $\operatorname{dist}(A, B) > 0$, we have $\mu(A \cup B) = \mu(A) + \mu(B)$. Then μ is a Borel measure.*

To wit, for any two sets A and B with $\operatorname{dist}(A, B) > 0$, as soon as $\delta < \frac{1}{2}\operatorname{dist}(A, B)$ we have that $\mathcal{H}^s_\delta(A \cup B) = \mathcal{H}^s_\delta(A) + \mathcal{H}^s_\delta(B)$ (since we can throw away any sets in the covering not intersecting either A or B). Thus taking $\delta \searrow 0$ yields $\mathcal{H}^s(A \cup B) = \mathcal{H}^s(A) + \mathcal{H}^s(B)$, as desired.

What is more is that Hausdorff measures on topological spaces are Borel regular. Indeed, for any $A \subset X$ we can take a sequence of Borel sets $\{E_k\}$ defined inductively as follows. Fix $\delta > 0$. Assuming $E_1 \supset \cdots \supset E_{k-1} \supset A$ have been constructed, let $\mathcal{F}_k$ be a countable family of closed sets with diameters less than δ covering A, such that

$$\sum_{F \subset \mathcal{F}_k} \omega_s \left(\frac{\operatorname{diam}(F)}{2} \right)^s - \mathcal{H}^s_\delta(A) \leqslant \frac{1}{k}.$$

Take E_k to be the union of all the sets in $\mathcal{F}_k$ intersected with E_{k-1}. Then clearly $\mathcal{H}^s_\delta(E_k) - \mathcal{H}^s_\delta(A) \leqslant 1/k$, and we can take $\delta \searrow 0$ and then $k \to \infty$ to conclude with the Borel set $E := \bigcap_i E_i$.

For a thorough discussion of the properties of the Hausdorff measures, see Chapter 3 of [Mag12] or Chapter 2 of [Sim14]. We wish to isolate one important fact for the sake of culture, however; let $X = \mathbb{R}^k$, on which we consider the k-dimensional Hausdorff measure $\mathcal{H}^k$. This Borel regular measure is in fact just the **k-dimensional Lebesgue measure**, often denoted by $\mathcal{L}^k$, and typically defined as follows: an *interval* I in $\mathbb{R}^k$ is any set of the form $\prod_{i=1}^k (a_i, b_i)$ where $-\infty < a_i < b_i < \infty$, and we *define* the volume of any such set to be the product $|I| := \prod_{i=1}^k (b_i - a_i)$. If now $A \subset \mathbb{R}^k$ is any set, then we define the **Lebesgue outer measure** of A by

$$|A| = \mathcal{L}^k(A) := \inf_{\mathcal{F}} \sum_{I \in \mathcal{F}} |I|,$$

where the infimum is taken over all countable collections $\mathcal{F}$ of intervals covering A. It is more-or-less straightforward to show that the resulting set function is indeed an outer measure, and by applying Caratheodory's Criterion we find that $\mathcal{L}^k$ is Borel. The happy "coincidence" that $\mathcal{H}^k$ coincides with $\mathcal{L}^k$ on $\mathbb{R}^k$ is proven through an interesting argument involving Vitali's property of the Lebesgue measure (see the next section), and the following geometric inequality proved through the process of *Steiner Symmetrization* :

Theorem 1.4 (Isodiametric Inequality)**.** *If $E \subset \mathbb{R}^k$, then $|E| \leqslant \omega_k \left(\frac{\operatorname{diam}(E)}{2} \right)^k$. Thus, amongst all Euclidean sets of a fixed diameter, balls have maximal volume.*

For a proof, see Chapter 3 of [Mag12].

Radon Measures

While Borel measures enjoy the property that Borel sets can be approximated to arbitrarily small error in measure from the inside by compact sets, we can consider a narrower family of measures which have such approximation properties built into their very definition. One might worry that such an imposition would be too restrictive, but such measures, remarkably, are abundant in number. Let us now work in a Hausdorff space X. A **Radon measure** on X is an outer measure μ on X such that the following three conditions hold:

- μ is Borel regular and locally finite[7];
- $\mu(A) = \inf_{\substack{U \supset A \\ U \text{open}}} \mu(U)$ for every subset $A \subset X$;
- $\mu(U) = \sup_{\substack{K \subset U \\ K \text{compact}}} \mu(K)$ for every open subset $U \subset X$.

It turns out that if X is a locally compact Hausdorff space where open sets are σ-compact, then *every* locally finite Borel regular measure on X is Radon (so the first condition alone is sufficient). In particular, the Euclidean spaces are of this form. We also note that the last condition can be greatly improved (in the fully general case), in that U can be taken to be μ-measurable instead of just open. The definition is stated in the way it is to highlight the interplay of μ with the topology of the space. For the detailed arguments of these assertions, see Chapter 1 of [Sim14].

We remark that the Lebesgue measure on $\mathbb{R}^k$ is obviously locally finite and Borel regular, and is thus Radon. Unfortunately, the Hausdorff measures are typically *not* locally finite since, for example, the lower dimensional Hausdorff measure of a higher (Hausdorff) dimensional set is infinite. Nonetheless, it turns out that the operation of restriction can help us out, since localizing a Borel regular measure to a set on which the measure is locally finite results in a locally finite Borel regular measure. Thus, in $\mathbb{R}^k$ we can restrict Hausdorff measures to sets on which they are locally finite to obtain Radon measures, a technique which we use extensively in our study of rectifiable varifolds.

An additional operation on measures that will be useful is the **pushforward** by a map. For our uses, we'll consider the case of an outer measure μ on $\mathbb{R}^n$ and a map $f\colon \mathbb{R}^n \to \mathbb{R}^m$, which come together to define the outer measure $f_\sharp\mu$ on $\mathbb{R}^m$ by

$$f_\sharp\mu(A) := \mu(f^{-1}(A)) \qquad \text{for all } A \subset \mathbb{R}^m.$$

The following appears as Proposition 2.14 of [Mag12]:

Proposition 1.2. *Let μ be a Radon measure on $\mathbb{R}^n$ and $f\colon \mathbb{R}^n \to \mathbb{R}^m$ be a continuous and proper map. Then the pushforward $f_\sharp\mu$ is a Radon measure on $\mathbb{R}^m$, $\operatorname{spt} f_\sharp = f(\operatorname{spt}\mu)$, and for every Borel $g\colon \mathbb{R}^m \to [0,\infty]$,*

$$\int_{\mathbb{R}^m} g \, \mathrm{d}f_\sharp\mu = \int_{\mathbb{R}^n} g \circ f \, \mathrm{d}\mu.$$

Before moving onto the next section, we remark that we can now develop the typical theory of Lebesgue integration based off of these measures, including the typical notions of measurable functions, L^p spaces, Sobolev spaces, et cetera. One critically important result is the following concerning the density of test functions:

Theorem 1.5. *Let X be a locally compact Hausdorff space, μ a Radon measure on X, and $1 \leqslant p < \infty$. Then $C_c(X)$ is dense in $L^p(X)$.*

The Riesz Representation Theorem

Besides having useful approximation properties, Radon measures appear in the following cornerstone result of measure theory and functional analysis. We will use this theorem to help prove a central result concerning the evolution of a varifold (thought of as a generalized C^1 surface) along the flow of a vector field, which introduces a generalized notion of mean curvature for such objects.

[7] An outer measure μ on a topological space X is said to be **locally finite** if $\mu(K) < \infty$ for all compact sets $K \subset X$.

Theorem 1.6 (Riesz Representation Theorem)**.** *Let X be a locally compact Hausdorff space, $(H, \langle -, - \rangle)$ a finite dimensional real Hilbert space, and $L\colon C_c(X; H) \to \mathbb{R}$ a continuous linear functional in the sense that*

$$\sup\{L(f) : f \in C_c(X;H), |f| \leqslant 1, \mathrm{spt} f \subset K\} < \infty$$

whenever $K \subset X$ is compact. Then there is a Radon measure μ on X and a μ-measurable function $g\colon X \to H$ with $|g| = 1$ μ-a.e. on X such that

$$L(f) = \int_X \langle f, g \rangle \mathrm{d}\mu$$

for every $f \in C_c(X; H)$. Moreover, we can characterize μ as the total variation measure $|L|\colon \mathcal{P}(X) \to [0, \infty]$ *of L, which itself is defined as follows: for open sets $A \subset X$ we define*

$$|L|(A) := \sup\{L(f) : f \in C_c(A; H), |f| \leqslant 1\}$$

and for arbitrary sets $E \subset X$

$$|L|(E) := \inf\{|L|(A) : E \subset A \ \ \text{for } A \text{ open}\}.$$

The rough idea of the proof is to extend the standard Riesz Representation Theorem for Hilbert Spaces to $L^1(X)$, associate to L a non-negative functional on $C_c(X; [0, \infty))$, which through some analysis yields the Radon measure μ, and then show that L defines k bounded linear operators on $C_c(X)$ which are just L "restricted" to the different coordinate directions. Applying Riesz' Theorem on $L^1(X)$ to each of these operators yields the component functions of g. For a proof, see either Chapter 4 of [Mag12] or Chapter 1 of [Sim14].

Weak-$*$ Convergence and Compactness for Radon Measures

While there exist several different notions of convergence for sequences of measures, we will be concerned with the notion of **weak-$*$ convergence** for defined as follows. If $\{\mu_j\}$ is a sequence of Radon measures on $\mathbb{R}^n$, then we say that μ_j weak-$*$ converges to a measure μ, written $\mu_j \overset{*}{\rightharpoonup} \mu$, iff

$$\int_{\mathbb{R}^n} f \mathrm{d}\mu = \lim_{j\to\infty} \int_{\mathbb{R}^n} f \mathrm{d}\mu_j \qquad \text{for all } f \in C_c(\mathbb{R}^n).$$

The following frequently used result appears as Proposition 4.26 of [Mag12].

Proposition 1.3. *Let $\{\mu_j\}$ be a sequence of Radon measures on $\mathbb{R}^n$. Then the following are equivalent:*

- $\mu_j \overset{*}{\rightharpoonup} \mu$
- *If K is compact and A is open, then*
  - $\mu(K) \geqslant \limsup_{j\to\infty} \mu_j(K)$
  - $\mu(A) \leqslant \liminf_{j\to\infty} \mu_j(A)$
- *If E is a bounded Borel set with $\mu(\partial E) = 0$, then $\mu(E) = \lim_{j\to\infty} \mu_j(E)$.*

At one point in the proof of Allard's theorem, we will be faced with a special sequence of varifolds and will want to show that they somehow converge to another varifold. By applying the following result to their associated Radon measures we achieve this goal.

Theorem 1.7 (Compactness for Radon Measures)**.** *Let $\{\mu_j\}$ be a sequence of Radon measures on a $\mathbb{R}^n$. Suppose that for every compact set $K \subset \mathbb{R}^n$ we have*

$$\sup_j \mu_j(K) < \infty.$$

Then there exists a Radon measure μ on $\mathbb{R}^n$ such that, up to a subsequence, we have $\mu_j \overset{*}{\rightharpoonup} \mu$.

For a detailed proof, see Theorem 4.16 in [Sim14] or Theorem 4.33 in [Mag12]. Briefly, by considering the sequences $\{\mu_j \llcorner B_h\}$ for a fixed $h \geqslant 1$ and applying a diagonal argument it suffices to consider the case where $\sup_j \mu_j(\mathbb{R}^n) < \infty$. Take a countable dense subset of $C_c(\mathbb{R}^n)$, and for each element f in the subset notice that $\{\int_{\mathbb{R}^n} f \mathrm{d}\mu_j\}$ is a bounded set of real numbers. Thus a subsequence converges to some $\alpha(f) \in \mathbb{R}$. By a diagonal argument we can extract a subsequence of the measures so that $\int_{\mathbb{R}^n} f \mathrm{d}\mu_j \to \alpha(f)$ for every f. By approximating an arbitrary element f of $C_c(\mathbb{R}^n)$ (in the Fréchet topology) by elements of the dense space, we can define a monotone linear functional $L\colon C_c(\mathbb{R}^n) \to \mathbb{R}$ by taking $L(u)$ as a limit of the $\alpha(f_j)$ where the f_j converge to f. Applying Riesz' Theorem then provides the limit Radon measure as the Radon measure representing L.

1.2 Covering Theorems

We collect here a few useful covering lemmas, and a good resource for such results is Section 1.5 of [EG15]. The first covering result will be used in making estimates for the proof of the Lipschitz Approximation Theorem.

Theorem 1.8 (5r Covering Theorem)**.** *Let $\mathcal{F}$ be any family of closed balls in a metric space X with the property that*

$$\sup_{B \in \mathcal{F}} \operatorname{diam}(B) < \infty.$$

Then there is a pairwise disjoint subfamily $\mathcal{F}' \subset \mathcal{F}$ such that

$$\bigcup_{B \in \mathcal{F}} B \subset \bigcup_{B \in \mathcal{F}'} \hat{B},$$

where $\hat{B}$ denotes the ball B dilated by a factor of 5 (i.e. if $B = B_r(x)$, then $\hat{B} = B_{5r}(x)$).

Theorem 1.9 (Besicovitch Covering Theorem)**.** *For each $n \geqslant 1$, there is a dimensional constant $k(n)$ such that the following holds. Let $\mathcal{F}$ be a collection of* closed *balls in $\mathbb{R}^n$, C the set of all centers of the balls in $\mathcal{F}$, and suppose that C is either bounded or that*

$$\sup_{B \in F} \operatorname{diam}(B) < \infty.$$

Then there exist (possibly empty) subfamilies $\mathcal{F}_1, \ldots, \mathcal{F}_{k(n)}$ of $\mathcal{F}$ such that

- *Each family $\mathcal{F}_i$ is pairwise disjoint and countable;*
- *$C \subset \bigcup_{i=1}^{k(n)} \bigcup_{B \in \mathcal{F}_i} B$.*

Corollary 1.1. *If μ is an outer measure on $\mathbb{R}^n$ and $\mathcal{F}$ and C are as in Besicovitch's Covering Theorem, then there is a countable, pairwise disjoint family $\mathcal{F}' \subset \mathcal{F}$ such that*

$$\mu(C) \leqslant k(n) \sum_{B \in \mathcal{F}'} \mu(C \cap B).$$

If μ is Borel, and C is μ-measurable, the we have in addition that

$$\mu(C) \leqslant k(n) \mu \left(C \cap \bigcup_{B \in \mathcal{F}'} B \right).$$

An outer measure μ on a metric space X is said to have the **Symmetric Vitali Property** if, given any $A \subset X$ of finite measure and any collection $\mathcal{F}$ of closed balls with centers in A which forms a **fine cover**[8] of A, then there is a countable, pairwise disjoint subcollection $\mathcal{F}' \subset \mathcal{F}$ with

$$\mu(A \setminus \bigcup_{B \in \mathcal{F}'} B) = 0.$$

Theorem 1.10 (Vitali's Property)**.** *If μ is a Radon measure on $\mathbb{R}^n$, then μ has the Symmetric Vitali Property.*

[8] By a **fine cover** of A, we simply mean that for each $x \in A$ we have $\inf\{r : B_r(x) \in \mathcal{F}\} = 0$.

1.3 The Lebesgue-Besicovitch Differentiation Theorem and Density Results

While a more general formulation of the following result exists for Borel regular measures on metric spaces, we instead present a version for Radon measures on $\mathbb{R}^n$, where we will need the theorem. For the details of the proof of this version, see Chapter 5 of [Mag12], and for the more general version, see Chapter 1 of [Sim14].

First, given two Radon measures μ and ν on $\mathbb{R}^n$ we define the **upper μ density** and **lower μ density** of ν at a point $x \in \text{spt}\mu$ by the respective formulas

$$D^+_\mu \nu(x) := \limsup_{r \searrow 0} \frac{\nu(\overline{B}_r(x))}{\mu(\overline{B}_r(x))}, \qquad D^-_\mu \nu(x) := \liminf_{r \searrow 0} \frac{\nu(\overline{B}_r(x))}{\mu(\overline{B}_r(x))}.$$

We have thus defined two (Borel) functions on $\text{spt}\mu$ taking values in $[0, \infty]$. Of particular interest are those points where both limits exist and are equal. At such points we denote the shared limit by $D_\mu \nu$ and call it the μ density of ν. We remark that by approximation of closed balls by open ones and vice-versa, we can actually use either closed or open balls in the definitions of these densities.

Theorem 1.11 (Lebesgue-Besicovitch Differentiation and a Density Theorem)**.** *Let μ and ν be Radon measures defined on $\mathbb{R}^n$. Then the density $D_\mu \nu$ is defined μ-a.e. on $\mathbb{R}^n$, is Borel measurable, and has $D_\mu \nu \in L^1_{loc}(\mathbb{R}^n, \mu)$. Moreover, there is a Radon measure $\nu^s_\mu \perp \mu$ concentrated on the Borel set*

$$Z = \mathbb{R}^n \setminus \{x \in \text{spt}\mu : D^+_\mu \nu(x) < \infty\} = (\mathbb{R}^n \setminus \text{spt}\mu) \cup \{x \in \text{spt}\mu : D^+_\mu \nu(x) = \infty\}$$

such that we have the unique decomposition

$$\nu = (D_\mu \nu)\mu + \nu^s_\mu$$

which is valid on $\mathfrak{M}(\mu)$.

One particular type of density that we will frequently study is the density with respect to a Hausdorff measure. We define the **upper** and **lower s-dimensional Hausdorff densities** of a Radon measure μ on $\mathbb{R}^n$, at a point $x \in \mathbb{R}^n$, by the respective formulas

$$\Theta^+_s \mu(x) := \limsup_{r \searrow 0} \frac{\mu(\overline{B}_r(x))}{\omega_s r^s}, \qquad \Theta^-_s \mu(x) := \liminf_{r \searrow 0} \frac{\mu(\overline{B}_r(x))}{\omega_s r^s}.$$

In the event of equality, we denote the shared limit by $\Theta_s \mu(x)$ and call it the **s-dimensional Hausdorff density** of μ at x. In the proof of our primary compactness result, we'll use these densities to show that a certain Radon measure is rectifiable. While there are deep theorems concerning the subject of rectifiability for measures (eg. the theorems of Marstrand and Preiss), we can make do with the following much simpler result, which can be found as Proposition 4.17 of [Mai17].

Proposition 1.4. *Let μ be a locally-finite measure on a metric space X, and suppose that there is an $s > 0$ such that for μ almost every $x \in X$,*

$$0 \leqslant \Theta^+_s \mu(x) < \infty.$$

Then $\mu = f\mathcal{H}^s$ for some $f \in L^1_{loc}(X, \mathcal{H}^s)$.

Lastly, we recount the following result (Theorem 6.2 in [Mat95]), which will be used in proving a corollary to the Monotonicity Formula.

Theorem 1.12. *Let $0 \leqslant s < \infty$, $A \subset \mathbb{R}^n$, and $\mathcal{H}^s(A) < \infty$. Then*

- $2^{-s} \leqslant \Theta^+_s(\mathcal{H}^s \llcorner A)(x) \leqslant 1$ *for $\mathcal{H}^s$-a.e. $x \in A$;*
- *If A is $\mathcal{H}^s$ measurable, then $\Theta^+_s(\mathcal{H}^s \llcorner A)(x) = 0$ for $\mathcal{H}^s$-a.e. $x \in \mathbb{R}^n \setminus A$.*

1.4 The Area Formulas

The following four results relate the Hausdorff measures to the notion of area of a graph as obtained by integration. These tools will find frequent use in what is to come. For proofs, see any of the books on measure theory in the references, for instance [Mag12] Chapter 8.

Theorem 1.13 (Area Formula (Take 1)). *Let $1 \leqslant n \leqslant m$, $E \subset \mathbb{R}^n$ Lebesgue measurable, and $f\colon E \to \mathbb{R}^m$ injective and Lipschitz. Then $f(E)$ is $\mathcal{H}^n$ measurable, $\mathcal{H}^n \llcorner f(E)$ is a Radon measure, and*

$$\mathcal{H}^n(f(E)) = \int_E \mathrm{J}f(x)\mathrm{d}x.$$

Theorem 1.14 (Change of Coordinates (Take 1)). *Let $1 \leqslant n \leqslant m$, $E \subset \mathbb{R}^n$ Lebesgue measurable, and suppose that $f\colon E \to \mathbb{R}^m$ is injective and Lipschitz. Let $g\colon \mathbb{R}^m \to [-\infty, \infty]$ be Borel measurable, with either $g \geqslant 0$ or $g \in L^1(\mathbb{R}^m, \mathcal{H}^n \llcorner f(E))$. Then $g \circ f$ is Borel measurable and*

$$\int_{f(E)} g \mathrm{d}\mathcal{H}^n = \int_E g(f(x))\mathrm{J}f(x)\mathrm{d}x$$

In case f is not injective, we have a generalization of the area formula, which works by taking into account the number of times f maps over a point.

Theorem 1.15 (Area Formula (Take 2)). *Let $1 \leqslant n \leqslant m$, $E \subset \mathbb{R}^n$ Lebesgue measurable, and suppose that $f\colon E \to \mathbb{R}^m$ is Lipschitz but not necessarily injective. Then*

$$\int_{\mathbb{R}^m} \mathcal{H}^0(f^{-1}(y) \cap E)\mathrm{d}\mathcal{H}^m(y) = \int_E \mathrm{J}f(x)\mathrm{d}x$$

Theorem 1.16 (Change of Coordinates (Take 2)). *Let $1 \leqslant n \leqslant m$ and suppose that $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz. Let $g\colon \mathbb{R}^n \to [-\infty, \infty]$ be Borel measurable, with either $g \geqslant 0$ or $g \in L^1(\mathbb{R}^n)$. Then*

$$\int_{\mathbb{R}^m} \left(\sum_{x \in f^{-1}\{y\}} g(x) \right) \mathrm{d}\mathcal{H}^n(y) = \int_E g(x)\mathrm{J}f(x)\mathrm{d}x$$

1.5 Campanato's Criterion

In the final step of proving Allard's Theorem, we utilize the following condition for Hölder regularity, known as *Campanato's Criterion*. The criterion states that if a L^p function on a ball has certain integral averages satisfying a uniform decay law, then after modification on/by measure zero sets the function is Hölder regular. Precisely we have the following (taken from Ch. 6 of [Mag12]):

Theorem 1.17 (Campanato's Criterion). *Let $n \geqslant 1$, $p \in [1, \infty)$, and $\alpha \in (0, 1]$. Then there exists a constant $C(n, p, \alpha) > 0$ such that the following holds. Let $u \in L^p(B)$, and define for $x \in B$, $r > 0$*

$$u_{x,r} := \fint_{B \cap B_r(x)} u = \frac{1}{|B \cap B_r(x)|} \int_{B \cap B_r(x)} u$$

Suppose that there exists a $\kappa > 0$ such that the uniform decay condition

$$\left(\frac{1}{r^n} \int_{B \cap B_r(x)} |u - u_{x,r}|^p \right)^{\frac{1}{p}} \leqslant \kappa r^\alpha$$

holds for all $x \in B$, $r > 0$. Then there exists a $\bar{u}\colon B \to \mathbb{R}$ with $\bar{u} = u$ a.e. on B and

$$|\bar{u}(x) - \bar{u}(y)| \leqslant C(n, p, \alpha)\kappa |x - y|^\alpha$$

for all $x, y \in B$.

Proof. First observe that it is possible to find a constant $C(n) > 0$ such that

$$C(n)r^n \leqslant |B \cap B_r(x)| \leqslant \omega_n r^n$$

for all $x \in B$, $r > 0$. Similarly, there is a $\theta(n) > 0$ such that if $x, y \in B$ and $r := |x - y|$, then

$$|B_r(x) \cap B_r(y)| = \theta(n) r^n.$$

Let then $\theta'(n) \leqslant \theta(n)$ be such that

$$\theta'(n) r^n \leqslant |B_r(x) \cap B_r(y) \cap B|.$$

We first claim that if $u \in L^p(B)$, then

$$\lim_{r \downarrow 0} \frac{1}{r^n} \int_{B \cap B_r(x)} |u(y) - u_{x,r}|^p \mathrm{d}y = 0$$

at every Lebesgue point x of u in B (so almost everywhere). Indeed, since $x \in B$, as soon as r is small enough we have

$$\begin{aligned}
\frac{1}{r^n} \int_{B \cap B_r(x)} |u(y) - u_{x,r}|^p \mathrm{d}y &= \frac{\omega_n}{|B_r(x)|} \int_{B_r(x)} |u(y) - u_{x,r}|^p \mathrm{d}y \\
&\leqslant C \left(\frac{1}{|B_r(x)|} \int_{B_r(x)} |u(y) - u(x)|^p \mathrm{d}y + \frac{1}{|B_r(x)|} \int_{B_r(x)} |u(x) - u_{x,r}|^p \mathrm{d}y \right) \\
&\leqslant C \left(\frac{1}{|B_r(x)|} \int_{B_r(x)} |u(y) - u(x)|^p \mathrm{d}y + |u(x) - u_{x,r}|^p \right) \\
&= C \left(\frac{1}{|B_r(x)|} \int_{B_r(x)} |u(y) - u(x)|^p \mathrm{d}y + \left| u(x) - \frac{1}{|B_r(x)|} \int_{B_r(x)} u \right|^p \right)
\end{aligned}$$

Thus, at every Lebesgue point of u we see that $\lim_{r \downarrow 0} \frac{1}{r^n} \int_{B \cap B_r(x)} |u(y) - u_{x,r}|^p \mathrm{d}y = 0$

Now, fix arbitrary radii $0 < r < R$ and $x \in B$. Then

$$\begin{aligned}
C(n) r^n |u_{x,r} - u_{x,R}|^p \leqslant |B \cap B_r(x)| |u_{x,r} - u_{x,R}|^p &= \int_{B \cap B_r(x)} |u_{x,r} - u_{x,R}|^p \mathrm{d}y \\
&\leqslant 2^{p-1} \left(\int_{B \cap B_r(x)} |u_{x,r} - u(y)|^p \mathrm{d}y + \int_{B \cap B_r(x)} |u(y) - u_{x,R}|^p \mathrm{d}y \right)
\end{aligned}$$

which implies that, by our main assumption and the fact that $r < R$,

$$\begin{aligned}
|u_{x,r} - u_{x,R}| &\leqslant C(n,p) \left(\frac{1}{r^n} \int_{B \cap B_r(x)} |u_{x,r} - u(y)|^p \mathrm{d}y + \frac{1}{r^n} \int_{B \cap B_r(x)} |u(y) - u_{x,R}|^p \mathrm{d}y \right)^{1/p} \\
&\leqslant C(n,p) \left(\frac{1}{r^n} \int_{B \cap B_r(x)} |u_{x,r} - u(y)|^p \mathrm{d}y + \left(\frac{R}{r}\right)^n \frac{1}{R^n} \int_{B \cap B_R(x)} |u(y) - u_{x,R}|^p \mathrm{d}y \right)^{1/p} \\
&\leqslant C(n,p) \left(\kappa^p r^{\alpha p} + \left(\frac{R}{r}\right)^n \kappa^p R^{\alpha p} \right)^{1/p} \\
&\leqslant C(n,p) \kappa \left(\frac{R}{r}\right)^{\frac{n}{p}} R^{\alpha}
\end{aligned}$$

We reiterate that this estimate holds for all $x \in B$ and for all radii $0 < r < R$.

Now let $r_k = 2^{-k} r$. Iterating the above argument yields for every $k > h \geqslant 0$ and $x \in B$

$$|u_{x,r_k} - u_{x,r_h}| \leqslant \sum_{j=h}^{k-1} |u_{x,r_{j+1}} - u_{x,r_j}| \leqslant C(n,p)\kappa r^\alpha \sum_{j=h}^{k-1} 2^{-j\alpha}. \tag{✻}$$

Take $h = 0$ and send k to infinity at every Lebesgue point x of u to find that

$$|u(x) - u_{x,r}| \leqslant C(n,p,\alpha)\kappa r^\alpha.$$

Let $x, y \in B$, $r := |x - y|$, and recall that by our choice of $\theta'(n) > 0$ that

$$\theta'(n) r^n |u_{x,r} - u_{y,r}|^p \leqslant 2^{p-1} \left(\int_{B \cap B_r(x)} |u(z) - u_{x,r}|^p \mathrm{d}z + \int_{B \cap B_r(y)} |u(z) - u_{y,r}|^p \mathrm{d}z \right)$$

which implies by our main assumption that

$$\begin{aligned} |u_{x,r} - u_{y,r}| &\leqslant C(n,p) \left(\frac{1}{r^n} \int_{B \cap B_r(x)} |u(z) - u_{x,r}|^p \mathrm{d}z + \frac{1}{r^n} \int_{B \cap B_r(y)} |u(z) - u_{y,r}|^p \mathrm{d}z \right)^{1/p} \\ &\leqslant C(n,p)\kappa r^\alpha \\ &= C(n,p)\kappa |x - y|^\alpha \end{aligned} \tag{✪}$$

Putting our estimates together we find that if $x, y \in B$ are Lebesgue points of u, then

$$|u(x) - u(y)| \leqslant |u(x) - u_{x,r}| + |u_{x,r} - u_{y,r}| + |u_{y,r} - u(y)| \leqslant C(n,p,\alpha)\kappa |x - y|^\alpha \tag{❇}$$

Now, (✻) tells us that the sequence of functions $\{u_{\cdot,r_k}\}_k$ is Cauchy with respect to uniform convergence on B. From (✪), we see that the $u_{\cdot,r_k}$ are all continuous, so by completeness there is a continuous $\bar{u} \colon B \to \mathbb{R}$ such that $u_{\cdot,r_k} \to \bar{u}$ uniformly on B. We also have that $u_{x,r_k} \to u(x)$ at every Lebesgue point x of u, and so at all such points we see that $u(x) = \bar{u}(x)$. From (❇), if x, y are Lebesgue points of u,

$$|\bar{u}(x) - \bar{u}(y)| \leqslant C(n,p,\alpha)\kappa |x - y|^\alpha.$$

By continuity of $\bar{u}$, we conclude the same estimate for all $x, y \in B$. Thus, $\bar{u} \in C^{0,\alpha}(B)$ and $\bar{u} = u$ a.e. on B (at all Lebesgue points of u).

□

1.6 Sobolev Inequalities and Embeddings

During the proof of Allard's Theorem we will construct a harmonic approximation to a Lipschitz graph. To prove the existence of such an approximation, and to obtain an important estimate from it, we will have to utilize two fundamental results from the theory of Sobolev spaces. We collect them here for reference, and refer the reader to any text on functional analysis for the details.

Theorem 1.18 (Rellich-Kondrachov Compactness Theorem). *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain, $1 \leqslant p < \infty$, and $j \geqslant 0$, $m \geqslant 1$ integers. Then $W_0^{j+m,p}(\Omega)$ is compactly embedded in $W^{j,q}(\Omega)$ for every $1 \leqslant q < \frac{dp}{d-mp}$ if $mp \leqslant d$, and in $C^j(\overline{\Omega})$ if $mp > d$.*

If Ω happens to have a bounded extension operator $E \colon W^{m,p}(\Omega) \to W^{m,p}(\tilde{\Omega})$ for some bounded open $\tilde{\Omega} \supset\supset \Omega$ (such as when Ω is Lipschitz), then the same result holds when $W_0^{m,p}(\Omega)$ is replaced with $W^{m,p}(\Omega)$.

Starting from here and the Banach-Alaoglu Theorem, we can readily prove the following by contradiction.

Proposition 1.5 (Poincaré Inequality). *Let $\Omega \subset \mathbb{R}^n$ be a bounded and connected Lipschitz domain. Then there is a constant $C = C(n, \Omega) > 0$ such that*

$$\|u\|_{L^2(\Omega)} \leqslant C \|\nabla u\|_{L^2(\Omega)}$$

for all functions u in the Sobolev space $H^1(\Omega)$ with $\int_\Omega u = 0$.

1.7 Results Concerning Harmonic Functions

As we will see, we can approximate a rectifiable varifold with small curvature by a Lipschitz graph, essentially by using nothing more than a monotonicity formula. This is itself proven by testing the first variation with nice vector fields. Thus, if we were only seeking $C^{0,1}$ regularity, we could get by with only these results (indeed from the Lipschitz Approximation we could jump directly to Step 3 of the main proof and conclude). However, the higher regularity result we seek actually comes from an improved approximation involving harmonic functions, which have decay properties that we can then pull back to the varifold using a "Tilt-Excess Inequality." This is a version of **Caccioppoli's Inequality** for varifolds, the standard version of which we recount below along with several other central results to the general theory of harmonic functions.

Proposition 1.6 (Mean Value Property)**.** *Let $\Omega \subset \mathbb{R}^n$ be open, and suppose that $u \in C^2(\Omega)$ satisfies $\Delta u = 0$ on Ω. Then for any ball $B_r(x_0)$ compactly contained in Ω, we have that*

$$u(x_0) = \fint_{\Omega} u dx = \fint_{\partial\Omega} u \mathrm{d}\sigma.$$

See Theorem 2.1 in [GT01].

Conversely, we have the following result. Thus, we see that the mean value property provides a characterization of harmonic functions.

Proposition 1.7. *Let $\Omega \subset \mathbb{R}^n$ be open, and suppose that $u \in C(\Omega)$ satisfies the mean value property*

$$u(x_0) = \fint_{\partial\Omega} u \mathrm{d}\sigma$$

for every ball $B_r(x_0)$ compactly contained in Ω. Then u is harmonic on Ω.

See Theorem 2.7 in [GT01].

Theorem 1.19 (Harmonic Gradient Estimate)**.** *Let $\Omega \subset \mathbb{R}^n$ be open, and $B_r(x_0) \subset \Omega$. Then given $k \geqslant 0$ there exists a constant $C_k > 0$ such that for each multi-index α of size k,*

$$|D^\alpha u(x_0)| \leqslant \frac{C_k}{r^{n+k}} \|u\|_{L^1(B_r(x_0))}.$$

See Chapter 2, Theorem 7 in [Eva10].

Lemma 1.1 (Weyl)**.** *Let $\Omega \subset \mathbb{R}^n$ be open. A function $u \in L^1_{loc}(\Omega)$ is harmonic iff*

$$\int_\Omega u \Delta \phi = 0 \qquad \forall \phi \in C_c^\infty(\Omega)$$

See Lemma 1.16 in [GM05].

Theorem 1.20 (Caccioppoli Inequality)**.** *Let $\Omega \subset \mathbb{R}^n$ be a domain, and suppose that $u \in H^1(\Omega)$ is a weak solution of $\Delta u = 0$:*

$$\int_\Omega D^\alpha u D^\alpha \phi = 0 \qquad \forall \phi \in H_0^1(\Omega).$$

Then for each $x_0 \in \Omega$ and choice of $0 < \rho < R \leqslant \operatorname{dist}(x_0, \partial\Omega)$ we have that

$$\int_{B_\rho(x_0)} |Du|^2 \leqslant \frac{c}{(R-\rho)^2} \int_{B_R(x_0)\setminus B_\rho(x_0)} |u-\lambda|^2 \qquad \forall \lambda \in \mathbb{R}.$$

See Theorem 4.1 of [GM05].

1.8 Rectifiable Sets

Rectifiable sets are the measure theoretic analogues of differentiable manifolds, and are the worlds upon which much of what is to follow lives. To pose our first definition, let $M \subset \mathbb{R}^n$ be a $\mathcal{H}^k$ measurable set (with $k \leqslant n$). We say that M is **countably $\mathcal{H}^k$ rectifiable** if there exist countably many Lipschitz maps $f_j \colon \mathbb{R}^k \to \mathbb{R}^n$ and a $\mathcal{H}^k$-null set $M_0 \subset \mathbb{R}^n$ such that $M \subset M_0 \cup \bigcup_j f_j(\mathbb{R}^k)$. Notice how the Lipschitz maps play a role analogous to that of parameterizations in smooth differential geometry, and that their coverage of M is allowed to be good up to a $\mathcal{H}^k$ measure 0 set. We say that M is **locally $\mathcal{H}^k$-rectifiable** provided it is countably $\mathcal{H}^k$-rectifiable and $\mathcal{H}^k(K \cap M) < \infty$ for all compact $K \subset \mathbb{R}^n$. Finally, should it occur for such an M that $\mathcal{H}^k(M) < \infty$, then we simply say that M is $\mathcal{H}^k$-rectifiable.

By Whitney's Approximation Theorem (see [EG15] Theorem 6.10) and the C^1 Approximation Theorem to Lipschitz functions (ibid. Theorem 6.11) it follows that a set $M \subset \mathbb{R}^n$ is countably $\mathcal{H}^k$ rectifiable iff there exist countably many k-dimensional embedded C^1 manifolds N_j of $\mathbb{R}^n$ and a $\mathcal{H}^k$-null set N_0 such that $M \subset N_0 \cup \bigcup_j N_j$. Starting from this, it follows that we can realize a countably $\mathcal{H}^k$ rectifiable set as countable disjoint union of $\mathcal{H}^k$-measurable sets, one of which is measure zero, and the others of which are contained in k-dimensional C^1 manifolds embedded in $\mathbb{R}^n$.

One of the most important differential geometric properties which carries over largely intact to rectifiable sets is the existence of tangent spaces. To state what this means (more generally) in a measure theoretic context, we first define the **k-dimensional blowup** of a Radon measure μ on $\mathbb{R}^n$ at the point x and at scale $r > 0$ by

$$\mu_{x,r}(E) \coloneqq \frac{1}{r^k}\mu(x + rE).$$

We can alternatively notate this as the pushforward $\frac{1}{r^k}(\Phi_{x,r})_\sharp\mu$ where $\Phi_{x,r} \colon \mathbb{R}^n \to \mathbb{R}^n$ is defined by $\Phi_{x,r}(y) \coloneqq (y - x)/r$, which has the added benefit of making obvious the fact that $\mu_{x,r}$ is Radon. In the event that there exists a k-dimensional plane π_x and a $\theta(x) \in \mathbb{R}$ such that as $r \searrow 0$ these blow ups weak-$*$ converge to $\theta(x)\mathcal{H}^k \llcorner \pi_x$, then we write $T_x\mu \coloneqq \pi_x$ and call it the **approximate tangent space** to μ at x with multiplicity $\theta(x)$. We'll be primarily concerned with Radon measures occurring as the restriction of Hausdorff measures to locally $\mathcal{H}^k$ rectifiable sets, so let us now discuss such measures. The following appears as Theorems 10.2 and 10.8 in [Mag12]:

Theorem 1.21. *Let $M \subset \mathbb{R}^n$ be a locally $\mathcal{H}^k$-rectifiable set. Then for $\mathcal{H}^k$-a.e. $x \in M$ there exists a unique k-dimensional plane π_x such that as $r \searrow 0$,*

$$(\mathcal{H}^k \llcorner M)_{x,r} = \mathcal{H}^k \llcorner \left(\frac{M - x}{r}\right) \overset{*}{\rightharpoonup} \mathcal{H}^k \llcorner \pi_x.$$

Additionally, the k-dimensional Hausdorff density $\Theta_k(\mathcal{H}^k \llcorner M) = 1$ $\mathcal{H}^k$-a.e. on M.

Conversely, if μ is a Radon measure on $\mathbb{R}^n$, concentrated on a Borel set $M \subset \mathbb{R}^n$, and for every $x \in M$ there exists a k-dimensional plane π_x so that as $r \searrow 0$ the blowups $\mu_{x,r}$ weak-$*$ converge to $\mathcal{H}^k \llcorner \pi_x$, then $\mu \equiv \mathcal{H}^k \llcorner M$ and M is locally $\mathcal{H}^k$-rectifiable.

The more suggestive notation $T_xM \coloneqq \pi_x$ is used for the approximate tangent spaces in this context, and heuristically the forwards direction of the theorem can be shown by observing that if $M = f(E)$ is a Lipschitz image of $f \colon \mathbb{R}^k \to \mathbb{R}^n$, then for $\mathcal{L}^k$-a.e. $x \in E$ we have that $T_{f(x)}M = Df_x(\mathbb{R}^k)$. Here we use the crucially important result (known as **Rademacher's Theorem**) that Lipschitz functions on $\mathbb{R}^n$ are classically differentiable almost everywhere. Once we have this, it is only a matter of figuring out how to modify our rectifiable set by sets of measure zero to obtain a nicer object on which we can apply this result. The reverse direction is more interesting, and involves the following auxiliary result for rectifiability (see Proposition 10.9 of [Mag12]). Here the set $K(\pi, t) \subset \mathbb{R}^n$, for a k-dimensional plane π and $t > 0$ denotes the cone $\{y \in \mathbb{R}^n : |P_\pi^\perp y| < t|P_\pi y|\} = \{y \in \mathbb{R}^n : |y| < \sqrt{1+t^2}|P_\pi y|\}$.

Proposition 1.8. *If $M \subset \mathbb{R}^n$ is compact, π is a k-dimensional plane, and there exist $\delta, t > 0$ such that for all $x \in M$*

$$M \cap B_\delta(x) \subset x + K(\pi, t),$$

then M is $\mathcal{H}^k$-rectifiable.

So, intuitively, a set is locally $\mathcal{H}^k$-rectifiable provided that there is a cone so that by sliding the point of the cone around the set, some uniform amount of the set is always contained in the translated cone.

Therefore, we see that in some sense $\mathcal{H}^k$-rectifiability is characterized by the almost everywhere existence of approximate tangent spaces. By a Hausdorff density argument (see Proposition 10.5 of [Mag12]), we can also show that if $M_1, M_2 \subset \mathbb{R}^n$ are two locally $\mathcal{H}^k$ rectifiable sets, then for $\mathcal{H}^k$-a.e. $x \in M_1 \cap M_2$ we have $T_x M_1 = T_x M_2$. This observation will be frequently useful in relating the tangent planes to our varfiold to the tangent planes of an approximation.

Before moving onto an earnest definition of *varifold*, we recount another fact which will be helpful in effectively wielding the monotonicity formula. The following density result for $\mathcal{H}^k$-measurable sets with multiplicity is adapted from Remark 1.8(2) in [Sim14]:

Proposition 1.9. *Let M be a $\mathcal{H}^k$-measurable subset of $\mathbb{R}^n$, and $\theta\colon M \to \mathbb{R}$ a positive locally $\mathcal{H}^k \llcorner M$-integrable function. Then*

$$\lim_{r \searrow 0} \frac{1}{\omega_k r^k} \int_{M \cap B_r(x)} \theta \mathrm{d}\mathcal{H}^k = \theta(x)$$

wherever $\mathcal{H}^k \llcorner M$ has an approximate tangent space with multiplicity $\theta(x)$.

Notice that this really does follow immediately from the definition of the approximate tangent space with multiplicity. Indeed, writing out the weak-$*$ definition for the approximate tangent space $T_x M$ to M at x with positive, locally $\mathcal{H}^k \llcorner M$ integrable multiplicity θ, we have for all $\phi \in C_c(\mathbb{R}^n)$ that

$$\lim_{r \searrow 0} \int_{\Phi_{x,r}(M)} f(y)\theta(x + \lambda y) \mathrm{d}\mathcal{H}^k(y) = \theta(x) \int_{T_x M} f \mathrm{d}\mathcal{H}^k.$$

Testing with $f \in C_c(\mathbb{R}^n)$ with $\chi_{B_1(0)} \leqslant f \leqslant \chi_{B_{1+\varepsilon}(0)}$ for $\varepsilon > 0$ leads us directly to the proposition.

1.9 Rectifiable Varifolds

In full generality, a **k-dimensional varifold** in an open set $U \subset \mathbb{R}^n$ is any Radon measure on the space $U \times G_k(\mathbb{R}^n)$, where $G_k(\mathbb{R}^n)$ is the k-Grassmanian over $\mathbb{R}^n$ with its standard smooth manifold structure and topology. For our purposes, it suffices to consider a more easily visualizable realization. Our central objects of study are the **rectifiable varifolds**, which intuitively are generalized C^1-manifolds together with a notion of multiplicity. Precisely, let $U \subset \mathbb{R}^n$ be an open set. A k-dimensional rectifiable varifold in U is a pair $V = (\Gamma, \theta)$, where Γ is a countably $\mathcal{H}^k$-rectifiable set in U, and θ is a positive, Borel measurable, locally $\mathcal{H}^k$ integrable function supported on Γ, called the multiplicity of the varifold. The magic of these objects comes from the following construction, which associates to V a natural measure $\mu_V := \theta \mathcal{H}^k \llcorner \Gamma$ called the **mass measure** of V. Its action on $\mathcal{H}^k$ measurable sets $E \subset \mathbb{R}^n$ is given explicitly by

$$\mu_V(A) := \int_{\Gamma \cap A} \theta \mathrm{d}\mathcal{H}^k.$$

Notice that as defined, μ_V is always a Borel measure, and that it is Radon if Γ is locally $\mathcal{H}^k$-rectifiable. We also denote by $\mathbb{M}(V) := \mu_V(U)$ the **mass of V** and from here onward *we assume that this mass is finite.*

At this point it is worth pausing for a moment to consider what these new constructions do for us on an intuitive level. Suppose that we are in a situation requiring us to take a "limit" of surfaces. For instance, we may be trying to minimize a functional acting on surfaces, and have in our hands an infimizing sequence from which we hope to extract a converging subsequence. If we start with classical C^1 surfaces we quickly realize that any reasonable class of limiting objects must be somewhat more general than just C^1 surfaces. Smooth objects might converge to objects with cusps or crinkles, and they might collapse in certain ways. An illustrative example comes from rescalings of the catenoid in $\mathbb{R}^3$. As we "zoom out" from the origin, we both see the upper and lower branches of the catenoid flattening onto the xy-plane, while the neck of the catenoid pinches down to the origin. Thus, the limit of these rescaled catenoids ought to converge to something that not only has a point missing from it, but also has multiplicity two! Indeed, the class of rectifiable varifolds captures these sorts of phenomena. Associating to each catenoid its mass measure, it can be easily shown that these measures weak-$*$ converge to $2\mathcal{H}^2 \llcorner \pi$, where π is the xy-plane. Thus, by expanding our admissible notions of "surface" to rectifiable varifolds, we can concoct notions of convergence and limit for which such problems have a hope of solution.

Yet another motivating reason for the consideration of rectifiable varfiolds for use in geometric variational problems is garnering the continuity of important functionals. A perennial favorite result is the following: Consider a sequence of circles in $\mathbb{R}^2$ of unit radius, and whose centers are the points $(0, \frac{1}{k})$. Now take the union of each such circle with the unit circle centered at the origin. This sequence of two dimensional rectifiable varifolds converges to the unit circle centered at the origin as $k \to \infty$, but consider what happens with the length of these varfiolds in the limit. Each individual varifold has length 4π, but the limit has length 2π! Thus, unless we consider the limiting unit circle with multiplicity 2 (which is of course completely consistent with intuition), we have "lost mass."

For a brief introduction to the idea of general varifolds, see [Men17].

1.10 Miscellaneous Results in Linear Algebra and Real Analysis

Lastly, we collect together a small potpourri of various results which will help us along the way.

Proposition 1.10 (Hadamard's Inequality)**.** *Let $v_1, \ldots, v_n$ be vectors in $\mathbb{R}^n$ forming the columns of a matrix $A := (v_1|\cdots|v_n)$. Then*

$$|\det(A)| \leqslant \prod_{i=1}^{n} |v_i|.$$

We'll use this result in the proof of the Excess Decay Theorem to make an estimate on the size of a Jacobian. The next result, whose proof can be found in Chapter 17 of [Mag12].

Proposition 1.11 (Taylor Expansion of the Determinant)**.** *Let $A \in \mathbb{R}^n \otimes \mathbb{R}^n$ and $I = Id_{\mathbb{R}^n}$. Then for sufficiently small $t > 0$,*

- $(I + tA)^{-1} = I - tA + t^2A^2 + \mathcal{O}(t^3)$
- $\det(I + tA) = 1 + t \text{ trace}(A) + \frac{t^2}{2}\left(\text{trace}(A)^2 - \text{trace}(A^2)\right) + \mathcal{O}(t^3)$

Recall that for a square matrix A we define its **adjugate**, $\text{Adj}(A)$, by the relation

$$A\text{Adj}(A) = \det(A)I.$$

In particular, $\text{Adj}(A)$ is the transpose of the cofactor matrix of A, and the formula is satisfied by the Laplace Expansion for the determinant. Moreover, if A is invertible then we can immediately write A^{-1} in terms of its adjugate as $A^{-1} = \frac{1}{\det A}\text{Adj}(A)$. Part of the importance of the adjugate comes from its appearance in the following formula, which we will utilize in studying the first variation of a varifold:

Proposition 1.12 (Jacobi's Formula)**.** *Let $A\colon \mathbb{R} \to \mathbb{R}^n \otimes \mathbb{R}^n$ be differentiable (here we identify $\mathbb{R}^n \otimes \mathbb{R}^n$ with $n \times n$ real matrices). Then $\det A$ is differentiable, and*

$$\frac{\mathrm{d}}{\mathrm{d}t}\det(A(t)) = \text{trace}\left(\text{Adj}(A(t))\frac{\mathrm{d}}{\mathrm{d}t}A(t)\right).$$

Lastly, we have the following technical lemma, which will be used in the proof of the Excess Decay Theorem to establish the so-called "Height Estimate." In particular, it will allow us to relate the deviation of the tangent plane to a graph relative to the plane it sits over to the derivatives of the graph at the base point of the tangent plane.

Lemma 1.2 (Tilting Subspace Lemma)**.** *Let A be a $k-$dimensional subspace of $\mathbb{R}^N$ with ON basis $\{\xi_i\}$, completed to an ON basis $\{\xi_1, \ldots, \xi_k, e_1, \ldots, e_{N-k}\}$ of $\mathbb{R}^N$. Let $v_1, \ldots, v_{N-k}$ be vectors in $\mathbb{R}^N$, and define the tilted plane*

$$B := \{y + L(y) : y \in A\}$$

where $L\colon A \to \mathbb{R}^N$ is defined by $L(y) = \sum_{j=1}^{N-k}(v_j \cdot y)e_j$. Then for a dimensional constant C we have

$$\|P_A - P_B\| \leqslant C\sum_{j=1}^{N-k}|v_j|.$$

Proof. Let $T := P_A - P_B$ and $K = \sum |v_j|$. We must show that $|Tx| \leqslant CK|x|$ for all $x \in \mathbb{R}^N$. It suffices to show this boundedness on A and $A^\perp$ separately, since every $x \in \mathbb{R}^N$ decomposes uniquely in the form $x = x^\top + x^\perp$ for $x^\top \in A$ and $x^\perp \in A^\perp$, allowing us to estimate

$$|Tx| \leqslant |Tx^\top| + |Tx^\perp| \leqslant CK(|x^\top| + |x^\perp|) \leqslant CK|x|$$

where the last inequality is due to the equivalence of the ℓ_1 and ℓ_2 norms on $\mathbb{R}^2$.

So, first let $x \in A$. Then $|Tx| = |x - P_Bx| \leqslant |x - y| = |x - z - L(z)|$ for every $y \in B$, where $y = z + L(z)$ for some $z \in A$, by the properties of orthogonal projections. Take $z = x$ to find that by Cauchy Schwarz

$$|Tx| \leqslant |L(x)| \leqslant CK|x|.$$

On the other hand let $x \in A^\perp \setminus \{0\}$. Then we have for some $y \in A \setminus \{0\}$

$$\begin{aligned}|Tx|^2 = |P_Bx|^2 = P_Bx \cdot P_Bx = \frac{(P_Bx \cdot P_Bx)^2}{|P_Bx|^2} = \frac{(x \cdot P_Bx)^2}{|P_Bx|^2} &= \frac{(x \cdot (y + L(y)))^2}{|y + L(y)|^2} = \frac{(x \cdot L(y))^2}{|y|^2 + |L(y)|^2} \\ &\leqslant |x|^2 \frac{|L(y)|^2}{|y|^2 + |L(y)|^2} \\ &= |x|^2 \frac{|L(\hat{y})|^2}{1 + |L(\hat{y})|^2} \\ &\leqslant |x|^2 |L(\hat{y})|^2 \\ &\leqslant (CK|x|)^2\end{aligned}$$

where we have written $\hat{y} = y/|y|$ and used the previous result. □

Finally, we define some notation and conventions which will be used throughout the paper: We will frequently write integrals of a function over all of a set while the function may only be defined almost everywhere on that set. For the sake of maintaining some semblance of manageable notation we tacitly assume that whenever this occurs we have accordingly made modifications on measure zero sets so that everything is well-defined. We also freely move between the notations (x, y) and $x + y$ for vectors in the direct sum of two subspaces of $\mathbb{R}^n$, based on whichever is most natural for the purposes of notation. $| - |$ will denote the standard Euclidean norm on $\mathbb{R}^n$ coming from the standard inner product denoted with a $\cdot$, and $\| - \|$ will denote some other operator norm. Since we will be working over finite dimensional spaces, the exact norm will seldom be important, but we'll work with particular ones when it is more convenient to do so. In particular, we will frequently use the norm induced by the **Frobenius inner product** $\langle - : - \rangle$ on real $n \times n$ matrices [9] by $\langle A : B \rangle := \operatorname{trace}(A^*B)$ where $-^*$ denotes the adjoint. We remark that for simple tensors $u \otimes v \in \mathbb{R}^n \otimes \mathbb{R}^n$ acting by $(u \otimes v)x = (v \cdot x)u$ for $x \in \mathbb{R}^n$, we have $\operatorname{trace}(u \otimes v) = u \cdot v$. Thus if $a, b, u, v \in \mathbb{R}^n$ then $\langle a \otimes b : u \otimes v \rangle = (a \cdot c)(b \cdot d)$. Lastly, we state once and for all that all varifolds having $\Gamma \subset U$ are (without loss of generality—see the technical remark at the end of Section 2) assumed to have $\overline{\Gamma} \cap U = \Gamma$.

2 Allard-Type Regularity Theorems for Rectifiable Varifolds

While this note will focus on the proof of one particular version of the regularity theorem, we wish to give not just an intuition for why the result should be true, but also an idea of what can actually be said regarding the subject, and where such results can take us. As we will see in the first section of the next part, a rectifiable varifold has a notion of generalized mean curvature. For the purposes of studying problems such as the famous Plateau Problem, we would be concerned with varifolds whose generalized mean curvature vanishes, analogously to what happens with classical minimal surfaces. The version of the theorem that we will explore is strong enough to handle such problems, since it assumes boundedness of the mean curvature. Nonetheless, better results are still obtainable. We can replace the L^∞ control on the mean curvature with

[9] Note that for operators defined on infinite dimensional spaces this inner product is more commonly known as the **Hilbert-Schmidt inner product**.

a suitable amount of integrability, and we can also remove a technical assumption relating to the oscillation of tangent planes to the varifold. A more recent result does even better, and asks not for control on the mean curvature but on a certain level of regularity of a generalized normal to the varifold. Moreover, the main theorem here will be stated for the class of *integer rectifiable varifolds*, but the later two results apply more generally to the class of *rectifiable varifolds*.

Here comes the main theorem that we will explore. Proper definitions of all the various quantities will be given soon, but for now it suffices to think of the generalized mean curvature $\mathbf{H}$ as a second order notion of curvature to the varifold, and the excess $\mathbb{E}(V, \pi, x_0, r)$ as an L^2 measure of the oscillation of tangent planes to the varifold around a given point and at a certain scale. Thus, a control on these quantities is somehow a control on how much the varifold is allowed to wiggle around and be irregular.

Theorem 2.1 (Allard (Take 1))**.** *Let $k < N$ be a positive integer. Then there are positive constants $\varepsilon, \alpha, \gamma$ such that the following holds. Let $V = (\Gamma, \theta)$ be a k-dimensional integer rectifiable varifold with bounded generalized mean curvature $\mathbf{H}$ supported in the ball $B_r(x_0)$, $x_0 \in \text{spt}\mu_V$, such that*

(A1) $\mu_V(B_r(x_0)) < (\omega_k + \varepsilon)r^k$ and $\|\mathbf{H}\|_\infty < \varepsilon r^{-1}$.

(A2) There is a k-dimensional plane π such that $\mathbb{E}(V, \pi, x_0, r) < \varepsilon$.

Then $B_{\gamma r}(x_0) \cap \Gamma$ is a $C^{1,\alpha}$ submanifold of $B_{\gamma r}(x_0)$ without boundary. Moreover, $\theta \equiv 1$ on $B_{\gamma r}(x_0) \cap \Gamma$.

Indeed the conditions $(A1)$ and $(A2)$ loosely tell us that, in comparison with a flat plane, the varfiold does not fold up too much, is not too wavy, and does not bend too tightly. If these sound somewhat redundant, it is because they are–we can remove the control on the excess by interpolating its smallness from the control on the mass of the varifold and the size of its mean curvature. In any case, if these quantities are all small it seems reasonable that some regularity might be lurking around. In fact, all of the arguments that we will make rely on the intuition that because these quantities are small, there should be nicer surfaces living nearby to the varifold which approximate it well. We begin by proving the Monotonicity Formula 3.1, which not only paves the way for many estimates concerning the growth of the area of the varifold and also provides regularity for the multiplicity function, but also tells us that in the regime of vanishing mean curvature, the area growth of an integral varifold is consistent with the bending experienced by classical minimal surfaces.

Next, we prove the Tilt-Excess Inequality 3.4, which can be thought of as a Caccioppoli inequality for varifolds. Loosely, it allows us to control the excess of the varifold after zooming into a smaller scale through the L^2-norm of a suitable Lipschitz approximation to the varifold. Up next and at the core of the proof are two results, the first of which is the Lipschitz Approximation Theorem 3.2. The Theorem says that when Allard's conditions are in force for a varifold with sufficiently small parameter ε, then a sizable portion of the varifold is covered by the graph of a Lipschitz function. Because the varifold has small mean curvature, the Lipschitz approximation ought to be "close" to being a minimal graph solving the Minimal Surface Equation. Since this equation is loosely a perturbation of the Laplace Equation, we are led to look for a graphical approximation to this Lipschitz approximation whose components are harmonic. This is the content of the Harmonic Approximation Lemma 3.5.

Using the growth properties of harmonic functions, we can then prove the Excess Decay Theorem 3.3, which says that if we start with a varifold satisfying Allard's conditions, then by zooming in a little and tilting our view the excess (which, again, is a measure of the oscillation of tangent planes to the varifold) decreases by at least a factor of 2. From here, we begin the proof of the main theorem, which happens in four parts. We use the Excess Decay Theorem to prove that when we "blow up" the varifold at a point by zooming into it, the excess must decrease to zero at a given rate. This allows us to then utilize the Lipschitz Approximation Theorem, as well as prove that a sizable portion of the varifold is actually covered by the Lipschitz graph. The third part of the proof consists of showing that on a smaller scale, the varifold actually coincides with the Lipshitz graph, and the final part uses the power law decay to prove that the conditions of Campanato's Criterion 1.17 are satisfied, completing the proof.

As we mentioned earlier, one of the assumptions in the statement above is redundant. Indeed, assumption $(A2)$ can be proven as a consequence of the bounds in assumption $(A1)$. The intuition for why this should be true is that the area bound is a zeroth order notion of curvature, and the mean curvature a second order notion of curvature. Since the excess deals with the oscillation of tangent planes, this should be a first order notion of curvature and we might hope to interpolate its smallness from the smallness of these lower and higher order quantities. Notice also that we also can relax the L^∞ control of the mean curvature to having integrability in L^p for $p > k$ with k the dimension of the varifold, and that we have an explicit (and in fact sharper) value of α in the $C^{1,\alpha}$ regularity conclusion. The following formulation of Allard's ε-Regularity Theorem for such varifolds appears as Theorem 5.2 in the text of [Sim14].

Theorem 2.2 (Allard (Take 2))**.** *Let $k < N$ be a positive integer, and $p > k$. Then there are positive constants ε, γ such that the following holds. Let $V = (\Gamma, \theta)$ be a k-dimensional rectifiable varifold with generalized mean curvature $\mathbf{H}$ supported in the ball $B_r(x_0)$, $x_0 \in \operatorname{spt}\mu_V$, such that*

(A'1) $\theta \geqslant 1$ for μ_V-a.e. $x \in \Gamma$.

(A'2) $\mu_V(B_r(x_0)) < (\omega_k + \varepsilon)r^k$ and $\left(\frac{1}{r^{k-p}} \int_{B_r(x_0)} |\mathbf{H}|^p \mathrm{d}\mu_V\right)^{1/p} \leqslant \varepsilon$.

Then there is a k-dimensional subspace π and a map $f \in C^{1,1-k/p}(B_{\gamma r}(x_0) \cap \pi; \pi^\perp)$ with $Df(x_0) = 0$, $\Gamma \cap B_{\gamma r}(x_0) = \Gamma_f \cap B_{\gamma r}(x_0)$, and moreover

$$\frac{1}{r} \sup|f| + \sup|Df| + r^{1-k/p} \sup_{\substack{x,y \in \mathcal{B}_{\gamma r}(x_0) \\ x \neq y}} \frac{|Df(x) - Df(y)|}{|x-y|^{1-k/p}} \leqslant C\varepsilon^{1/(2k+2)}$$

where $C = C(k, N, p)$.

In 2013 Bourni and Volkmann proved (in [BV16]) the following generalization of Allard's theorem, which dispenses with the assumed estimates on the generalized mean curvature, instead asking only for Hölder regularity of the generalized normal vector to the varifold. Before stating their theorem, we define this notion of Hölder regularity. In the definition, the object δV is the *first variation* of the varifold, a construct that we will meet in the next section. Heuristically, the first variation is a linear map of C^1_c vector fields, and is defined to be the initial rate of change of area that the varifold experiences when being moved by the flow of such a vector field.

Definition 2.1 (Generalized $C^{0,\alpha}$ Normal)**.** *Let $U \subset \mathbb{R}^n$ be open, and $V = (\Gamma, \theta)$ a k-dimensional rectifiable varifold in U. Then V is said to have generalized normal of class $C^{0,\alpha}$ in U if there is a constant $K \geqslant 0$ such that for all $B_r(x) \subset U$ and all $X \in C^1_c(B_r(x); \mathbb{R}^N)$,*

$$\delta V(X) \leqslant K r^\alpha \int_U \|\mathrm{d}^{T_x\Gamma} X\| \mathrm{d}\mu_V(x),$$

where $\mathrm{d}^\pi X \coloneqq DX \circ P_\pi$ for P_π the orthogonal projection onto a k-dimensional subspace π.

Theorem 2.3 (Bourni and Volkmann)**.** *Let $k < N$ be a positive integer. Then there are positive constants ε, γ such that the following holds. Let $V = (\Gamma, \theta)$ be a k-dimensional rectifiable varifold supported in the ball $B_r(x_0)$, $x_0 \in \operatorname{spt}\mu_V$, such that*

(A"1) $\theta \geqslant 1$ for μ_V-a.e. $x \in \Gamma$.

(A"2) $\mu_V(B_r(x_0)) < (\omega_k + \varepsilon)r^k$ and V has generalized $C^{0,\alpha}$ normal with $Kr^\alpha \leqslant \varepsilon$.

Then there is a k-dimensional subspace π and a map $f \in C^{1,1-n/p}(B_{\gamma r}(x_0) \cap \pi; \pi^\perp)$ with $\Gamma \cap B_{\gamma r}(x_0) = \Gamma_f \cap B_{\gamma r}(x_0)$.

The proof idea here is similar in spirit to the proof we'll give for the first version of the theorem. A new version of the Monotonicity Formula is proven, and from this corresponding versions of the Lipschitz Approximation Theorem and Excess Decay Theorem are established. Then, the proof follows just as in the case we consider. That the $C^{0,\alpha}$ normal condition really does generalize the mean curvature condition is shown along with the proof of the main theorem in Bourni and Volkmann's Paper [BV16].

Once we have such a result, what can be done with it? One important direction is in building the regularity theory for mass-minimizing integral currents in the effort of studying solutions the Plateau Problem. For a treatment of this topic, see [Sim14] Chapter 7, or [Lel16]. A more immediately accessible result is the following one (appearing as Corollary 3.3 in [Lel18]). Here we use results to be proven in the next section.

Theorem 2.4. *Let α be as in Allard's Theorem (Take 1), and let $V = (\Gamma, \theta)$ be a k-dimensional integer rectifiable varifold in the open set $U \subset \mathbb{R}^N$. Then there is an open set $W \subset U$ such that $W \cap \Gamma$ is a $C^{1,\alpha}$ submanifold of W without boundary. Moreover, $W \cap \Gamma$ is dense in Γ. If in addition θ is constant μ_V-a.e. then $\mu_V(\Gamma \setminus W) = 0$.*

Proof. The second conclusion is relatively easy to get. Without loss of generality, we can assume that $\theta \equiv 1$ μ_V-a.e. on Γ. Then for μ_V-a.e $x_0 \in \Gamma$, we have by Corollary 3.1 that

$$1 = \theta(x_0) = \lim_{r \searrow 0} \frac{\mu_V(B_r(x_0))}{\omega_k r^k}.$$

Moreover, the map $x \mapsto \|T_x\Gamma - T_{x_0}\Gamma\|^2$ is an L^1_{loc} function on Γ by virtue of V being of finite mass, and its value at $x = 0$ is 0. By the Lebesgue Points Theorem, we thus have at μ_V-a.e $x_0 \in \Gamma$ that

$$\lim_{r \searrow 0} \frac{1}{r^k} \int_{B_r(x_0)} \|T_x\Gamma - T_{x_0}\Gamma\|^2 \mathrm{d}\mu_V(x) = 0.$$

Thus, at every such x_0 there is a corresponding $r_{x_0} > 0$ such that $(A1), (A2)$ hold on $B_{r_{x_0}}(x_0)$ and with $\pi = T_{x_0}\Gamma$. Thus by Allard's Theorem we conclude that $B_{\gamma r_{x_0}}(x_0) \cap \Gamma$ is a $C^{1,\alpha}$ submanifold of $B_{\gamma r_{x_0}}(x_0)$ without boundary. We take W to be the union of all such open balls for each such x_0. Notice that if Γ has no isolated points, then it is guaranteed that $\Gamma \subset W$.

The proof of the first conclusion is topological. Recall the standing assumption that $\overline{\Gamma} \cap U = \Gamma$, and consider Γ and its closure $\bar{\Gamma}$ as metric subspaces of $\mathbb{R}^N$ with its standard metric. Then Γ is in fact an open subset of $\overline{\Gamma}$, which is itself a complete metric space. Thus, Γ is a Baire space, which is the crux of the proof to follow. As we will see in the next section, the multiplicity θ is upper semicontinuous, and so the sets $C_k := \{\theta \geqslant k\}$ are closed in $\overline{\Gamma}$ (and relatively closed in Γ) for each $k \geqslant 1$. For each such k, let D_k be the interior of C_k in the topology of Γ and set $E_k := D_k \setminus C_{k+1}$, $E := \cup_{k \geqslant 1} E_k$.

Fix now any $x \in \Gamma \setminus E$, and let $k \geqslant 1$ be such that $\theta(x) \in [k, k+1)$. Note this means that $x \in C_k$. By the upper semi-continuity of θ in some neighborhood of x we have $1 \leqslant \theta < k + 1$. If $x \in D_k = \mathrm{int} C_k$, then we reach the contradiction that $x \in E_k = D_k \setminus C_{k+1}$, since there would exist a neighborhood where $k \leqslant \theta < k + 1$. Thus, $x \in C_k \setminus D_k$, and altogether we have proven that $\Gamma \setminus E \subset \cup_k C_k \setminus D_k$. But $C_k \setminus D_k$ is nowhere dense, so by the Baire Category Theorem we discover that E is a *dense* open set.

Let for each $k \geqslant 1$ $U_k \subset \mathbb{R}^N$ be an open set such that $U_k \cap \Gamma = E_k$, and focus attention on each of the integral varifolds $V_k := (\Gamma \cap U_k, \theta \restriction_{U_k})$. Since these varifolds all have bounded mean curvature in their respective U_k, and have multiplicities $\theta \restriction_{U_k} = k\ \mu_{V_i}$-a.e. in U_k, by the first conclusion of the theorem there is an open set $W_k \subset U_k$ so that $\Gamma \cap W_k$ is a $C^{1,\alpha}$ submanifold of W_k without boundary, with $\mu_{V_k}(\Gamma \cap (U_k \setminus W_k)) = 0$. But then $\Gamma \cap W_k$ is dense in $\Gamma \cap U_k = E_k$. Set $W = \cup_k W_k$ and conclude that $\Gamma \cap W$ is a $C^{1,\alpha}$ submanifold of W without boundary with $\Gamma \cap W$ dense in E, which is itself dense in Γ.

□

We make a technical remark on the hypotheses of the theorem above. The standing assumption that $\overline{\Gamma} \cap U = \Gamma$ is not exactly without cost in generality. Indeed, we can impose this condition on a varifold by modifying Γ, but this runs the risk of introducing non-integral values of the multiplicity via a density argument in Corollary 3.1 (e.g., consider the perils of adding a boundary point to an open half plane). In particular, under the assumption of bounded mean curvature, by this same corollary we have that the multiplicity of the varifold and the density agree almost everywhere, so that we might as well work with the density as it has more mathematical structure behind it. This is, of course, at the cost of not being everywhere integer valued. To circumvent this technicality, we really ought to add the hypothesis that *the density is everywhere integer valued* to the above theorem.

3 Allard's Regularity Theorem for Integer Rectifiable Varifolds with L^∞ Mean Curvature

3.1 Integer Rectifiable Varifolds and the First Variation

In the introduction we briefly considered the idea of rectifiable varifolds. In this section, we specialize to a subclass called the *integer rectifiable varifolds*, which are simply rectifiable varifolds whose multiplicities are positive integer valued (almost everywhere). We study here a construction which brings functional analytic tools to the study of these varifolds, and allows us to view our integral varifolds as objects acting on test vector fields. We thereby uncover a connection between these new objects and classical smooth manifolds through the notion of *generalized mean curvature*.

To begin, let $\Phi\colon U \to W$ be a smooth diffeomorphism of open sets U and W in $\mathbb{R}^N$, and let $V = (\Gamma, \theta)$ be a k-dimensional integral varifold in U. We define the pushforward of V to be the k-dimensional integral varifold $\Phi_\#(V)$ in W given by

$$\Phi_\#(V) = (\Phi(\Gamma), f \circ \Phi^{-1})$$

The pushforward is still k-dimensional as Γ is covered by a countable union of k-dimensional C^1 submanifolds of $\mathbb{R}^N$, which are themselves pushed forward to other C^1 submanifolds of $\mathbb{R}^N$, and since the Borel measurability of $f \circ \Phi^{-1}$ is obvious. Now, for a vector field $X \in C^1_c(U;\mathbb{R}^N)$ we obtain a global flow $\Phi\colon \mathbb{R} \times U \to U$ which is the unique solution of the system

$$\begin{cases} \frac{\partial}{\partial t}\Phi(t,x) = X(\Phi(t,x)) \\ \Phi(0,x) = Id_U(x) = x \end{cases}$$

From the global flow Φ we obtain the one parameter family of flow maps $\Phi_t\colon U \to U$ sending $x \mapsto \Phi(t,x)$ for each $x \in U$.

Definition 3.1 (The First Variation). *If $V = (\Gamma, \theta)$ is an integral varifold in the open set $U \subset \mathbb{R}^N$, and $X \in C^1_c(U;\mathbb{R}^N)$ generates the flow maps $\Phi_t\colon U \to U$, then the **first variation** of V on X is defined by*

$$\delta V(X) := \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \mathbb{M}((\Phi_t)_\#(V))$$

Unfortunately, as defined the first variation is not very easy to work with (in fact we don't even know that it is well-defined to begin with!). Fortunately, we have the following characterization in terms of the *tangential divergence* of X along tangent planes. We define this notion of divergence as follows: let π be a k-dimensional subspace of $\mathbb{R}^N$ with any choice of ON basis $\{e_1, \ldots, e_k\}$, and set

$$\mathrm{div}_\pi X := \sum_{j=1}^{k} e_j \cdot D_{e_j} X \qquad \forall X \in C^1(U;\mathbb{R}^N)$$

where D_{e_j} is the directional derivative of X in the direction of e_j.

Lemma 3.1. *With π a k-dimensional subspace of $\mathbb{R}^N$ and $X \in C^1(U;\mathbb{R}^N)$, we have*

$$\mathrm{div}_\pi X = \mathrm{div}(P_\pi X).$$

Proof. Let $\{\xi_i\}_{i=1}^k$ be an ON base of π, completed to an ON base of $\mathbb{R}^N$. If $X = X^i \xi_i$, then $P_\pi X = \sum_{i=1}^k X^i \xi_i$ and

$$\mathrm{div}_\pi X = \sum_{i=1}^{k} \xi_i \cdot D_{\xi_i} X = \sum_{i=1}^{k} (\xi_i \cdot \xi_j) D_{\xi_i} X^j = \sum_{j=1}^{k} D_{\xi_i} X^i = \mathrm{div}(P_\pi X).$$

□

From these characterizations the following proposition follows immediately:

Proposition 3.1. *Let $U \subset \mathbb{R}^N$ be open and π a k-dimensional subspace of $\mathbb{R}^N$. Then the tangential divergence operator $\mathrm{div}_\pi\colon C^1(U;\mathbb{R}^N) \to C(U)$ is linear and satisfies the following Leibniz rule for all $f \in C^1(U)$ and $X \in C^1(U;\mathbb{R}^N)$:*

$$\mathrm{div}_\pi(fX) = \nabla f \cdot P_\pi X + f \mathrm{div}_\pi X.$$

With this operator in hand we can prove the following versatile characterization of the first variation, which also settles the issue of its well-definedness.

Proposition 3.2. *Let $V = (\Gamma, \theta)$ be an integral varifold in an open subset $U \subset \mathbb{R}^N$. Then the first variation of V can be characterized as*

$$\delta V(X) = \int_U \operatorname{div}_{T_x\Gamma} X \mathrm{d}\mu_V(x) \qquad \forall X \in C_c^1(U; \mathbb{R}^N)$$

Proof. Since Γ is rectifiable, we can find countably many C^1 embeddings $F_i \colon K_i \to \mathbb{R}^N$ where each $K_i \subset \mathbb{R}^k$ is compact, such that for all $i \neq j$

(i) $F_i(K_i) \cap F_j(K_j) = \emptyset$

(ii) $F_i(K_i) \subset \Gamma$

(iii) $\{F_i(K_i)\}$ cover $\mathcal{H}^k$-a.e. of Γ

By looking at the preimages of each positive integer under θ and intersecting with each $F_i(K_i)$, we can further refine our countable collection of K_i so that θ is constant θ_i on each $F_i(K_i)$. Since $\Phi_t \colon U \to U$ is a diffeomorphism for each t, these properties all continue to hold when we replace Γ with $\Phi_t(\Gamma)$, F_i with $\Phi_t \circ F_i$, and θ with $\theta \circ \Phi_t^{-1}$. Let $\{e_i\}$ be an ON basis of $\mathbb{R}^k$, and recall that if $v_1, \ldots, v_k$ are vectors in $\mathbb{R}^k$, then we have $|v_1 \wedge \ldots \wedge v_k|^2 = \det(g_{ij})$ where g_{ij} denotes the $k \times k$ matrix with components $v_i \cdot v_j$. By the area formula, and the fact that $\Phi_t \circ F_i$ is C^1 and injective,

$$\begin{aligned}
\mathbb{M}((\Phi_t)_\#(V)) = \int_{\Phi_t(\Gamma)} \theta \circ \Phi_t^{-1} \mathrm{d}\mathcal{H}^k &= \sum_i \theta_i \int_{\Phi_t(F_i(K_i))} \mathrm{d}\mathcal{H}^k = \sum_i \theta_i \int_{K_i} \mathrm{J}(\Phi_t \circ F_i) \mathrm{d}x \\
&= \sum_i \theta_i \int_{K_i} |\mathrm{d}(\Phi_t \circ F_i)_y e_1 \wedge \cdots \wedge \mathrm{d}(\Phi_t \circ F_i)_y e_k| \mathrm{d}y
\end{aligned}$$

Now let $y \in K_i$ and set $x = F_i(y)$. $\{\mathrm{d}(F_i)_y(e_i)\}$ is a basis for $T_x\Gamma$ (wherever the tangent space exists) since each F_i is an embedding. If $v_1, \ldots, v_k$ denotes any ON basis of $T_x\Gamma$, then by the chain rule we have that

$$|\mathrm{d}(\Phi_t \circ F_i)_y e_1 \wedge \cdots \wedge \mathrm{d}(\Phi_t \circ F_i)_y e_k| = |\mathrm{d}(\Phi_t)_x(v_1) \wedge \cdots \wedge \mathrm{d}(\Phi_t)_x(v_k)| \cdot |\mathrm{d}(F_i)_y(e_1) \wedge \cdots \wedge \mathrm{d}(F_i)_y(e_k)|$$

For ease of notation, let us call $h_x(t) := |\mathrm{d}(\Phi_t)_x(v_1) \wedge \cdots \wedge \mathrm{d}(\Phi_t)_x(v_k)| = \sqrt{\det g_x(t)}$, where $g_x(t)$ is the matrix with entries $g_x(t)_{ij} = \mathrm{d}(\Phi_t)_x(v_i) \cdot \mathrm{d}(\Phi_t)_x(v_j)$. Using the Area Formula Corollary, noting $\Phi_0 = \mathrm{Id}_U$ so that $h_x(0) = 1$, we can form the difference quotient for the mass of V under the flow of X:

$$\begin{aligned}
\frac{\mathbb{M}((\Phi_t)_\#(V)) - \mathbb{M}(V)}{t} &= \sum_i \frac{\theta_i}{t} \int_{K_i} (h_x(t) - 1) \, |\mathrm{d}(F_i)_y(e_1) \wedge \cdots \wedge \mathrm{d}(F_i)_y(e_k)| \mathrm{d}y \\
&= \sum_i \int_{K_i} \frac{h_x(t) - h_x(0)}{t} \theta_i |\mathrm{d}(F_i)_y(e_1) \wedge \cdots \wedge \mathrm{d}(F_i)_y(e_k)| \mathrm{d}y \\
&= \sum_i \int_{K_i} \frac{h_x(t) - h_x(0)}{t} \theta_i \mathrm{J}F_i(y) \mathrm{d}y \\
&= \sum_i \int_{F_i(K_i)} \frac{h_x(t) - h_x(0)}{t} \theta_i \mathrm{d}\mathcal{H}^k(x) \\
&= \int_U \frac{h_x(t) - h_x(0)}{t} \mathrm{d}\mu_V(x)
\end{aligned}$$

We thus need to do two things—we first need to show that $h_x(t)$ is differentiable at 0, and then we want to pass the limit as $t \to 0$ through the integral. First off, since Φ_t is smooth and $\mathrm{d}(\Phi_t)_x(v_i) = D_{v_i}\Phi_t$, we note that

$$\frac{\partial}{\partial t} \left(\mathrm{d}(\Phi_t)_x(v_i) \right) = \frac{\partial}{\partial t} \left(D_{v_i} \Phi_t \right) = \left(D_{v_i} \frac{\partial}{\partial t} \Phi_t \right) = D_{v_i}(X \circ \Phi_t),$$

hence

$$(g_x'(t))_{ij} = D_{v_i}(X \circ \Phi_t)(x) \cdot \mathrm{d}(\Phi_t)_x(v_j) + D_{v_j}(X \circ \Phi_t)(x) \cdot \mathrm{d}(\Phi_t)_x(v_i).$$

Now, notice that $h_x(t)$ is indeed differentiable at $t = 0$ since $\det g_x(t)$ is by the Jacobi Formula and the above direct computation. By said formula we obtain

$$h'_x(t) = \frac{1}{2\sqrt{\det(g_x(t))}} \det(g_x(t))' = \frac{1}{2\sqrt{\det(g_x(t))}} \det(g_x(t))\mathrm{trace}\left(g_x^{-1}(t)g'_x(t)\right)$$

and if we evaluate at $t = 0$ we have, since $g_x(0) = I$,

$$h'_x(0) = \frac{1}{2}\mathrm{trace}(g'_x(0)) = \sum_{i=1}^{k} v_i \cdot D_{v_i}X(x) = \mathrm{div}_{T_x\Gamma}X(x).$$

Moreover, since X is compactly supported, Φ_t reduces to the identity map outside of the compact set $\mathrm{spt}X$. Thus, in your favorite operator norm we have $\|\mathrm{d}(\Phi_t)_x\| \leqslant C$ for a constant C independent of $x \in U$ and $t \in [-1, 1]$. This tells us that $\|g'_x(t)\|_{L^\infty(\Gamma\times[-1,1])} < \infty$. Therefore, h_x is not only differentiable, but there is also a $\delta \in (0, 1)$ and another $C > 0$ such that

$$|h_x(t) - h_x(0)| \leqslant C|t|$$

for all $|t| < \delta$ *and all* $x \in \Gamma$. Hence by applying the Dominated Convergence Theorem (and acknowledging our assumption that V has finite mass) we can safely pass the limit through the integral and conclude that

$$\delta V(X) = \lim_{t\to 0} \frac{\mathbb{M}((\Phi_t)_\#(V)) - \mathbb{M}(V)}{t} = \int_\Gamma h'_x(0)\mathrm{d}\mu_V(x) = \int_U \mathrm{div}_{T_x\Gamma}X\mathrm{d}\mu_V(x).$$

□

As an added bonus of the above analysis, we also see that the first variation is a *linear* functional on $C^1_c(U;\mathbb{R}^N)$. With no small amount of foreshadowing, we say that V has **bounded generalized mean curvature** if δV is a bounded linear functional on $C^1_c(U;\mathbb{R}^N)$ in the sense that for a universal C not depending upon $X \in C^1_c(U;\mathbb{R}^N)$,

$$|\delta V(X)| \leqslant C\|X\|_{L^1(U,\mu_V)} = C\int_U |X|\mathrm{d}\mu_V$$

As a bit of terminology, we say that V is a **stationary** integral varifold if we can take $C = 0$, which just says that the first variation of V vanishes along all vector fields $X \in C^1_c(U;\mathbb{R}^N)$.

Now, if V has bounded generalized mean curvature, we can apply the Riesz-Markov-Kakutani Representation Theorem and the Radon-Nikodym Theorem to establish the following additional characterization of the first variation. This result brings us back to the theory of classical C^2 surfaces having a mean curvature vector.

Proposition 3.3. *Let $V = (\Gamma, \theta)$ be an integral varifold in $U \subset \mathbb{R}^N$. Then V has bounded generalized mean curvature iff there exists a bounded Borel map $\mathbf{H}\colon U \to \mathbb{R}^N$ such that for all $X \in C^1_c(U;\mathbb{R}^N)$,*

$$\delta V(X) = -\int X \cdot \mathbf{H}\mathrm{d}\mu_V.$$

Proof. First of all, we can extend δ from $C^1_c(U;\mathbb{R}^N)$ to $C_c(U;\mathbb{R}^N)$ using boundedness and density. Thus we have in hand a continuous linear functional on $C_c(U;\mathbb{R}^N)$, which by the Riesz-Markov-Kakutani Representation Theorem can be realized as

$$\delta V(X) = \int_U X \cdot g\mathrm{d}|\delta V|$$

for all $X \in C_c(U;\mathbb{R}^N)$, where $|g| = 1$ μ_V-a.e. in U and where $|\delta V|$ is the total variation measure of δV defined as

$$|\delta V|(A) := \sup\{\delta V(X) : X \in C_c(A;\mathbb{R}^N), |X| \leqslant 1\}$$

for $A \subset U$ open and

$$|\delta V|(E) := \inf\{|\delta V|(A) : E \subset A,\ A \text{ open}\}$$

for arbitrary $E \subset U$. We claim that on the level of vector valued Radon measures, $g|\delta V| = -\mathbf{H}\mu_V$ for a bounded Borel map $\mathbf{H}\colon U \to \mathbb{R}^N$.

We must show that $|\delta V| << \mu_V$. If $A \subset U$ is open with $\mu_V(A) = \int_{A\cap\Gamma} \theta \mathrm{d}\mathcal{H}^k = 0$, then for any $X \in C_c(A;\mathbb{R}^N)$ with $|X| \leqslant 1$

$$|\delta V|(A) \leqslant C \int_A |X| \mathrm{d}\mu_V \leqslant C\mu_V(A) = 0 \tag{✻}$$

Thus, $|\delta V|(E) = 0$ for all $E \subset U$ with $\mu_V(E) = 0$, proving the absolute continuity of $|\delta V|$ with respect to μ_V. By the Radon-Nikodym Theorem, the derivative

$$D_{\mu_V}|\delta V|(x) = \lim_{r\to 0} \frac{|\delta V|(B(x,r))}{\mu_V(B(x,r))}$$

exists at μ_V-a.e. $x \in \mathrm{spt}\mu_V$, and $|\delta V| = D_{\mu_V}|\delta V|\mu_V$. Notice that by the inequalities in (✻), $D_{\mu_V}|\delta V|$ is bounded. Define at last $\mathbf{H}\colon U \to \mathbb{R}^N$ by $\mathbf{H} := -gD_{\mu_V}|\delta V|$. Then $\mathbf{H}$ is bounded, Borel, has $|\mathbf{H}| = D_{\mu_V}|\delta V|$, is defined μ_V-a.e. in U, and satisfies for all $X \in C_c(U;\mathbb{R}^N)$

$$\delta V(X) = \int_U X \cdot g\mathrm{d}|\delta V| = -\int_U X \cdot \mathbf{H}\mathrm{d}\mu_V.$$

In the event that $X \in C^1_c(U;\mathbb{R}^N)$, then we have in addition

$$\delta V(X) = \int_U \mathrm{div}_{T_x\Gamma} X \mathrm{d}\mu_V(x) = -\int_U X \cdot \mathbf{H}\mathrm{d}\mu_V.$$

We call $\mathbf{H}$ the **generalized mean curvature vector** of V. □

Before moving onward to our discussion of the Monotonicity Formula, we remark that the theory presented here for integral varifolds in open subsets of $\mathbb{R}^N$ extends to the theory of varifolds in Riemannian manifolds. In particular, let M be a closed Riemannian manifold embedded in $\mathbb{R}^N$. If $U \subset \mathbb{R}^N$ is open, then an integer rectifiable varifold V in $U \cap M$ is just an integral varifold V in U with the added property that $\mu_V(U\setminus M) = 0$. We say that V is stationary in $U \cap M$ iff $\delta V(X) = 0$ for all $X \in C^1_c(U;\mathbb{R}^N)$ which are *tangent* to M.

3.2 Monotonicity Formula

This section is dedicated to the Monotonicity Formula, which is a beautiful result finding use in many places elsewhere in the theory of minimal surfaces. In fact, the Formula tells us that integer rectifiable varifolds with vanishing generalized mean curvature behave in a way similar to classical minimal surfaces, in that they are forced to bend in a way consistent with the area growth that minimal surfaces exhibit due to their vanishing mean curvature. We will also prove here a corollary relating the density of points on the varifold to the values of the multiplicity function θ, enabling us to also garner regularity data for θ.

We first fix some notation for the section. If $g\colon U \to \mathbb{R}$ is differentiable and $V = (\Gamma, \theta)$ is an integral varifold in U, then we define $\nabla^\perp g(x) := P_{N_x\Gamma}(\nabla g(x))$ as the projection of ∇g onto the normal space $N_x\Gamma$ orthogonal to $T_x\Gamma$, which exists $\mathcal{H}^k-$a.e. on Γ.

Theorem 3.1 (Monotonicity Formula). *Let $V = (\Gamma, \theta)$ be a k-dimensional integral varifold in an open subset $U \subset \mathbb{R}^N$ with bounded generalized mean curvature $\mathbf{H}$. Fix any $\xi \in U$ and define the map $r(x) := |x - \xi|$. For every $0 < \sigma < \rho < \mathrm{dist}(\xi, \partial U)$, the following estimate holds true*

$$\frac{\mu_V(B_\rho(\xi))}{\rho^k} - \frac{\mu_V(B_\sigma(\xi))}{\sigma^k} = \int_{B_\rho(\xi)} \frac{1}{k}\mathbf{H}\cdot(x-\xi)\left(\frac{1}{m(r)^k} - \frac{1}{\rho^k}\right)\mathrm{d}\mu_V(x) + \int_{B_\rho(\xi)\setminus B_\sigma(\xi)} \frac{|\nabla^\perp r|^2}{r^k}\mathrm{d}\mu_V$$

with $m(r) := \max\{r, \sigma\}$.

Moreover, the map $\rho \mapsto e^{\|\mathbf{H}\|_\infty \rho}\rho^{-k}\mu_V(B_\rho(\xi))$ is monotone increasing.

Notice here how $\nabla^\perp r(y)$ captures another notion of curvature for the varifold. Indeed, $\nabla^\perp r = P^\perp_{T_y\Gamma}(\nabla|y|) = P^\perp_{T_y\Gamma}(\hat{y})$, so that $\nabla^\perp r(y)$ is the part of the unit vector $\hat{y}$ (in the direction of y) normal to $T_y\Gamma$. Thus, if we stand at the origin and look at a point on Γ, this quantity tells us how our line of sight differs from being tangent to the varifold at that point. In other words, by integrating this (after appropriately rescaling to ensure dimensional consistency) quantity over a region, we can get a sense of how curvy our varifold is compared to a cone.

Proof. By translating we can assume that $\xi = 0$. Fix any $\gamma \in C^1_c([0,1))$ with $\gamma \equiv 1$ in a neighborhood of 0. For any $s \in (0, \operatorname{dist}(0, \partial U))$, define the vector field $X_s \in C^1_c(U)$ by $X_s(x) := \gamma\left(\frac{|x|}{s}\right) x$. Since $X_s \in C^1_c(U)$, we can test

$$\delta V(X_s) = \int_U \operatorname{div}_{T_x\Gamma} X_s \mathrm{d}\mu_V(x) = -\int_U X \cdot \mathbf{H} \mathrm{d}\mu_V$$

Fix any $x \in \Gamma$ at which $T_x\Gamma$ exists, and choose for it an ON basis $\{1, \ldots, e_k\}$, completed to an ON basis $\{e_1, \ldots, e_N\}$ of $\mathbb{R}^N$. Using these coordinates we compute

$$\begin{aligned}
\operatorname{div}_{T_x\Gamma} X_s = \sum_{i=1}^k e_i \cdot D_{e_i} X_s &= \sum_{i=1}^k e_i \cdot D_{e_i} \left\{\gamma\left(\frac{|x|}{s}\right) x\right\} \\
&= \sum_{i=1}^k e_i \cdot \left\{\gamma\left(\frac{|x|}{s}\right) e_i\right\} + \sum_{i=1}^k e_i \cdot \left\{\gamma'\left(\frac{|x|}{s}\right) \frac{x_i}{s|x|} x\right\} \\
&= k\gamma\left(\frac{|x|}{s}\right) + \gamma'\left(\frac{|x|}{s}\right) \frac{1}{s} \sum_{i=1}^k (e_i \cdot x)\left(e_i \cdot \frac{x}{|x|}\right) \\
&= k\gamma\left(\frac{|x|}{s}\right) + \gamma'\left(\frac{|x|}{s}\right) \frac{r}{s} \sum_{i=1}^k \left(e_i \cdot \frac{x}{|x|}\right)^2 \\
&= k\gamma\left(\frac{|x|}{s}\right) + \gamma'\left(\frac{|x|}{s}\right) \frac{r}{s} \left(1 - \sum_{i=k+1}^N \left(e_i \cdot \frac{x}{|x|}\right)^2\right) \\
&= k\gamma\left(\frac{|x|}{s}\right) + \gamma'\left(\frac{|x|}{s}\right) \frac{r}{s} \left(1 - |\nabla^\perp r|^2\right)
\end{aligned}$$

since $\nabla r(x) = \nabla |x| = \frac{x}{|x|} = \sum_{k=1}^N (e_i \cdot \frac{x}{|x|}) e_i$. Therefore,

$$\int_U \operatorname{div}_{T_x\Gamma} X_s \mathrm{d}\mu_V(x) = -\int_U X \cdot \mathbf{H} \mathrm{d}\mu_V$$

yields

$$-\int_U X \cdot \mathbf{H} \mathrm{d}\mu_V = \int_U k\gamma\left(\frac{|x|}{s}\right) \mathrm{d}\mu_V(x) + \int_U \gamma'\left(\frac{|x|}{s}\right) \frac{r}{s} \left(1 - |\nabla^\perp r|^2\right) \mathrm{d}\mu_V(x).$$

We divide by s^{k+1} and integrate from σ to ρ, obtaining

$$\begin{aligned}
-\int_\sigma^\rho \int_U \frac{1}{s^{k+1}} (x \cdot \mathbf{H}) \gamma\left(\frac{|x|}{s}\right) \mathrm{d}\mu_V(x) \mathrm{d}s = &\int_\sigma^\rho \int_U \frac{k}{s^{k+1}} \gamma\left(\frac{|x|}{s}\right) \mathrm{d}\mu_V(x) \mathrm{d}s \\
&+ \int_\sigma^\rho \int_U \frac{|x|}{s^{k+2}} \gamma'\left(\frac{|x|}{s}\right) \left(1 - |\nabla^\perp r|^2\right) \mathrm{d}\mu_V(x) \mathrm{d}s
\end{aligned}$$

Everything in sight has compact support and is bounded, so we can safely apply Fubini's theorem and rearrange to find

$$\begin{aligned}
\int_U \int_\sigma^\rho \left(\frac{k}{s^{k+1}} \gamma\left(\frac{|x|}{s}\right) + \frac{|x|}{s^{k+2}} \gamma'\left(\frac{|x|}{s}\right)\right) \mathrm{d}s \mathrm{d}\mu_V(x) = &\int_U |\nabla^\perp r|^2 \int_\sigma^\rho \frac{|x|}{s^{k+2}} \gamma'\left(\frac{|x|}{s}\right) \mathrm{d}s \mathrm{d}\mu_V(x) \\
&- \int_U x \cdot \mathbf{H} \int_\sigma^\rho \frac{1}{s^{k+1}} \gamma\left(\frac{|x|}{s}\right) \mathrm{d}s \mathrm{d}\mu_V(x)
\end{aligned} \tag{$\circledast$}$$

Next, we integrate by parts to compute

$$-\int_\sigma^\rho \frac{k}{s^{k+1}} \gamma\left(\frac{|x|}{s}\right) \mathrm{d}s = \frac{1}{s^k} \gamma\left(\frac{|x|}{s}\right)\Big|_\sigma^\rho + \int_\sigma^\rho \frac{|x|}{s^{k+2}} \gamma'\left(\frac{|x|}{s}\right) \mathrm{d}s$$

so that

$$-\int_\sigma^\rho \left(\frac{k}{s^{k+1}}\gamma\left(\frac{|x|}{s}\right)+\frac{|x|}{s^{k+2}}\gamma'\left(\frac{|x|}{s}\right)\right)\mathrm{d}s=\frac{1}{\rho^k}\gamma\left(\frac{|x|}{\rho}\right)-\frac{1}{\sigma^k}\gamma\left(\frac{|x|}{\sigma}\right)$$

Consequently, it follows that (✻) becomes, after rearrangement,

$$\begin{aligned}\frac{1}{\rho^k}\int_U\gamma\left(\frac{|x|}{\rho}\right)\mathrm{d}\mu_V(x)-\frac{1}{\sigma^k}\int_U\gamma\left(\frac{|x|}{\sigma}\right)\mathrm{d}\mu_V(x)&=\int_U x\cdot\mathbf{H}\int_\sigma^\rho\frac{1}{s^{k+1}}\gamma\left(\frac{|x|}{s}\right)\mathrm{d}s\mathrm{d}\mu_V(x)\\&+\int_U|\nabla^\perp r|^2\left[\frac{1}{\rho^k}\gamma\left(\frac{|x|}{\rho}\right)-\frac{1}{\sigma^k}\gamma\left(\frac{|x|}{\sigma}\right)+\int_\sigma^\rho\frac{k}{s^{k+1}}\gamma\left(\frac{|x|}{s}\right)\mathrm{d}s\right]\mathrm{d}\mu_V(x)\end{aligned}\tag{✻✻}$$

Recall that this statement holds for every $\gamma\in C_c^1([0,1))$, and in particular it holds on a sequence $\gamma_k\nearrow\mathcal{X}_{[0,1)}$, $0\leqslant\gamma_k\leqslant\mathcal{X}_{[0,1)}$. By the Monotone Convergence Theorem, we discover that (✻✻) is in fact true with $\gamma=\mathcal{X}_{[0,1)}$. We have

$$\begin{aligned}\frac{\mu_V(B_\rho)}{\rho^k}-\frac{\mu_V(B_\sigma)}{\sigma^k}&=\int_U x\cdot\mathbf{H}\int_\sigma^\rho\frac{1}{s^{k+1}}\mathcal{X}_{B_s}\mathrm{d}s\mathrm{d}\mu_V(x)\\&+\int_U|\nabla^\perp r|^2\left[\frac{1}{\rho^k}\mathcal{X}_{B_\rho}-\frac{1}{\sigma^k}\mathcal{X}_{B_\sigma}+\int_\sigma^\rho\frac{k}{s^{k+1}}\mathcal{X}_{B_s}\mathrm{d}s\right]\mathrm{d}\mu_V(x)\end{aligned}\tag{✻✻✻}$$

Consider now the integral terms in s. We compute

$$\int_\sigma^\rho\frac{k}{s^{k+1}}\mathcal{X}_{B_s}(x)\mathrm{d}s=\mathcal{X}_{B_\rho}\int_{\max\{|x|,\sigma\}}^\rho\frac{k}{s^{k+1}}\mathrm{d}s=\mathcal{X}_{B_\rho}\cdot\left(\frac{1}{\max\{|x|,\sigma\}^k}-\frac{1}{\rho^k}\right)=\mathcal{X}_{B_\rho}\cdot\left(\frac{1}{m(r)^k}-\frac{1}{\rho^k}\right)$$

since $\mathcal{X}_{B_s}(x)=0$ if $|x|\geqslant s$. This takes care of the mean curvature term, and all we have left to simplify is the final integral:

$$\frac{1}{\rho^k}\mathcal{X}_{B_\rho}-\frac{1}{\sigma^k}\mathcal{X}_{B_\sigma}+\mathcal{X}_{B_\rho}\cdot\left(\frac{1}{m(r)^k}-\frac{1}{\rho^k}\right)=\frac{1}{m(r)^k}\mathcal{X}_{B_\rho}-\frac{1}{\sigma^k}\mathcal{X}_{B_\sigma}=\frac{1}{r^k}\mathcal{X}_{B_\rho\setminus B_\sigma}=\frac{1}{|x|^k}\mathcal{X}_{B_\rho\setminus B_\sigma}.$$

Putting everything together in (✻✻✻) yields the Monotonicity Formula:

$$\frac{\mu_V(B_\rho)}{\rho^k}-\frac{\mu_V(B_\sigma)}{\sigma^k}=\int_{B_\rho}\frac{1}{k}(x\cdot\mathbf{H})\left(\frac{1}{m(r)^k}-\frac{1}{\rho^k}\right)\mathrm{d}\mu_V(x)+\int_{B_\rho\setminus B_\sigma}\frac{|\nabla^\perp r|^2}{|x|^k}\mu_V(x)$$

Next we need to show that the map $\rho\mapsto e^{\|\mathbf{H}\|_\infty\rho}\rho^{-k}\mu_V(B_\rho(\xi))$ is monotone increasing. To this end, we define $f(\rho):=\rho^{-k}\mu_V(B_\rho)$. Then by the Monotonicity Formula and Cauchy-Schwarz, we discover that

$$f(\rho)-f(\sigma)\geqslant\frac{1}{k}\int_{B_\rho}x\cdot\mathbf{H}\left(m(r)^{-k}-\rho^{-k}\right)\mathrm{d}\mu_V(x)\geqslant-\|\mathbf{H}\|_\infty\int_{B_\rho}|x|\left(m(r)^{-k}-\rho^{-k}\right)\mathrm{d}\mu_V(x)$$

Dividing both sides by $\rho-\sigma$ and recalling that $m(r)=\max\{|x|,\sigma\}\geqslant\sigma$, we find

$$\frac{f(\rho)-f(\sigma)}{\rho-\sigma}\geqslant-\|\mathbf{H}\|_\infty\rho\mu_V(B_\rho)\cdot\frac{\sigma^{-k}-\rho^{-k}}{\rho-\sigma}$$

Recall that a real valued function g is convex iff the quantity $R(x,y):=\frac{g(x)-g(y)}{x-y}$ is monotone increasing in one variable when the other is fixed. Since $\rho\mapsto\rho^{-k}$ is convex, we see that

$$-\sigma^{-(k+1)}\leqslant\sigma^{-(k+1)}=R(0,\sigma)\leqslant R(\rho,\sigma)=-\frac{\sigma^{-k}-\rho^{-k}}{\rho-\sigma}.$$

Writing $\rho=\sigma+\varepsilon$, we obtain

$$\frac{f(\sigma+\varepsilon)-f(\sigma)}{\varepsilon}\geqslant-\|\mathbf{H}\|_\infty(\sigma+\varepsilon)\mu_V(B_\rho)\sigma^{-(k+1)}=-\|\mathbf{H}\|_\infty f(\sigma+\varepsilon)\frac{(\sigma+\varepsilon)^{k+1}}{\sigma^{k+1}}\tag{✭}$$

Let Ψ_δ be a standard mollifier and convolve both sides of the above inequality as functions of σ (we first harmlessly extend f to all of $\mathbb{R}$ by setting it equal to 0 on the negative reals). We obtain by a change of variables

$$\left(\frac{f(\cdot+\varepsilon)-f(\cdot)}{\varepsilon}*\Psi_\delta\right)(\sigma)=\int_{\mathbb{R}}\frac{f(t+\varepsilon)-f(t)}{\varepsilon}\Psi_\delta(\sigma-t)\mathrm{d}t=\int_{\mathbb{R}}\frac{\Psi_\delta(\sigma-t+\varepsilon)-\Psi_\delta(\sigma-t)}{\varepsilon}f(t)\mathrm{d}t$$

and we estimate that (if, say, $\varepsilon<1$)

$$\begin{aligned}\left|\left(\frac{f(\cdot+\varepsilon)-f(\cdot)}{\varepsilon}*\Psi_\delta\right)(\sigma)-(f*\Psi_\delta')(\sigma)\right| &\leqslant \int_{\mathbb{R}}\left|\frac{\Psi_\delta(\sigma-t+\varepsilon)-\Psi_\delta(\sigma-t)}{\varepsilon}f(t)-\Psi_\delta'(\sigma-t)f(t)\right|\mathrm{d}t\\ &\leqslant \|f\|_\infty\int_K\left|\frac{\Psi_\delta(\sigma-t+\varepsilon)-\Psi_\delta(\sigma-t)}{\varepsilon}-\Psi_\delta'(\sigma-t)\right|\mathrm{d}t\end{aligned}$$

where K is compact, since $\mathrm{spt}\Psi_\delta=\overline{B}_\delta$. Note that $\|f\|_\infty\leqslant\omega_k$, since $f(\rho)=\rho^{-k}\mu_V(B_\rho)\leqslant\rho^{-k}\mathcal{H}^k(B_\rho)=\omega_k$ for every $\rho>0$. Since Ψ_δ is smooth, the integrand is uniformly bounded in $\varepsilon<1$ and converges to 0 as ε goes to 0, so we can apply the Dominated Convergence Theorem to conclude that

$$\lim_{\varepsilon\downarrow 0}\left(\frac{f(\cdot+\varepsilon)-f(\cdot)}{\varepsilon}*\Psi_\delta\right)(\sigma)=(f*\Psi_\delta')(\sigma)=(f*\Psi_\delta)'(\sigma)$$

for every $\sigma>0$. Similarly, we find that

$$\lim_{\varepsilon\downarrow 0}\left(f(\cdot+\varepsilon)\frac{(\cdot+\varepsilon)^{k+1}}{(\cdot)^{k+1}}*\Psi_\delta\right)(\sigma)=(f*\Psi_\delta)(\sigma)$$

for every $\sigma>0$. We therefore establish that after convolving both sides of (✯) with Ψ_δ and sending ε to 0, we are left with

$$(f*\Psi_\delta)'(\sigma)+\|\mathbf{H}\|_\infty(f*\Psi_\delta)(\sigma)\geqslant 0.$$

But this tells us that the function $g_\delta(\sigma):=e^{\|\mathbf{H}\|_\infty\sigma}(f*\Psi_\delta)(\sigma)$ has for all $\sigma>0$

$$g_\delta'(\sigma)=e^{\|\mathbf{H}\|_\infty\sigma}\left\{(f*\Psi_\delta)'(\sigma)+\|\mathbf{H}\|_\infty(f*\Psi_\delta)(\sigma)\right\}\geqslant 0.$$

Therefore, g_δ is monotone increasing for any $\delta>0$. But then

$$\lim_{\delta\downarrow 0}g_\delta(\sigma)=e^{\|\mathbf{H}\|_\infty\sigma}f(\sigma)=e^{\|\mathbf{H}\|_\infty\sigma}\sigma^{-k}\mu_V(B_\sigma)$$

is also monotone increasing, as desired. □

Corollary 3.1. *Let $V=(\Gamma,\theta)$ be a $k-$dimensional rectifiable varifold in an open set $U\subset\mathbb{R}^N$ with bounded generalized mean curvature. Then*

$$\theta_V(x):=\lim_{\rho\searrow 0}\frac{\mu_V(B_\rho(x))}{\omega_k\rho^k}$$

*exists at **every** $x\in U$, and $\theta_V=\theta$ for μ_V-a.e. $x\in U$. Moreover, we have that*

(i) θ_V is upper semicontinuous.

(ii) $\theta_V(x)\geqslant 1$ at every $x\in\mathrm{spt}\mu_V$.

(iii) $\mu_V(B_\rho(x))\geqslant\omega_k e^{-\|\mathbf{H}\|_\infty\rho}\rho^k$ for every $x\in\mathrm{spt}\mu_V$ and $\rho<\mathrm{dist}(x,\partial U)$.

(iv) $\mathcal{H}^k(\mathrm{spt}\mu_V\setminus\Gamma)=0$.

Proof. We first settle the existence of the limit defining θ_V by recalling that the map $\rho\mapsto e^{\|\mathbf{H}\|_\infty\rho}\rho^{-k}\mu_V(B_\rho(x))$ is monotone increasing and bounded below. Thus $\lim_{\rho\searrow 0}e^{\|\mathbf{H}\|_\infty\rho}\rho^{-k}\mu_V(B_\rho(x))$ exists, and since $e^{\|\mathbf{H}\|_\infty\rho}\to 1$ as $\rho\to 0$ it follows that

$$\omega_k\theta_V(x)=\lim_{\rho\searrow 0}\frac{\mu_V(B_\rho(x))}{\rho^k}$$

must exist as well. The fact that $\theta_V = \theta$ for $\mathcal{H}^k$-a.e. $x \in \Gamma$ is due to Proposition 1.9. In particular, the two functions agree wherever Γ has a tangent space.

To prove the upper semicontinuity property (i), we fix an $x \in U$ and $\varepsilon > 0$, letting $\rho \in (0, \frac{1}{2}\mathrm{dist}(x, \partial U))$ be such that

$$e^{\|\mathbf{H}\|_\infty \sigma} \frac{\mu_V(B_\sigma(x))}{\omega_k \sigma^k} \leqslant \theta_V(x) + \frac{\varepsilon}{2} \qquad \forall \sigma < 2\rho$$

If we now select $0 < \delta < \rho$ and any $y \in U$ with $|x - y| < \delta$, then

$$\theta_V(y) \leqslant e^{\|\mathbf{H}\|_\infty \rho} \frac{\mu_V(B_\rho(y))}{\omega_k \rho^k} \leqslant e^{\|\mathbf{H}\|_\infty(\rho+\delta)} \frac{\mu_V(B_{\rho+\delta}(x))}{\omega_k(\rho+\delta)^k} \frac{(\rho+\delta)^k}{\rho^k} \leqslant \left(1 + \frac{\delta}{\rho}\right)^k \left(\theta_V(x) + \frac{\varepsilon}{2}\right)$$

If we take δ sufficiently small relative to ρ then it will follow that $\theta_V(y) \leqslant \theta_V(x) + \varepsilon$ whenever $y \in B_\delta(x)$, which says that $\limsup_{y \to x} \theta_V(y) \leqslant \theta_V(x)$.

Next, since $\theta_V = \theta$ $\mathcal{H}^k$-a.e. on Γ, and since $\theta \colon \Gamma \to \mathbb{Z}^{\geqslant 1}$, it follows that $\{x : \theta_V(x) \in \mathbb{Z}^{\geqslant 1}\}$ has full μ_V measure. This tells us that $\{x : \theta_V(x) \in \mathbb{Z}^{\geqslant 1}\}$ is dense in $\mathrm{spt}\mu_V$, and so if $x \in \mathrm{spt}\mu_V \cap U$, the claim that $\theta(x) \geqslant 1$ follows from upper semicontinuity. If $x \in \mathrm{spt}\mu_V$ and $\rho < \mathrm{dist}(x, \partial U)$, then $e^{\|\mathbf{H}\|_\infty \rho} \omega_k^{-1} \rho^{-k} \mu_V(B_\rho(x)) \geqslant \omega_k^{-1} \rho^{-k} \mu_V(B_\rho(x))$ implies that

$$\lim_{\rho \searrow 0} e^{\|\mathbf{H}\|_\infty \rho} \omega_k^{-1} \rho^{-k} \mu_V(B_\rho(x)) \geqslant 1.$$

Since $e^{\|\mathbf{H}\|_\infty \rho} \omega_k^{-1} \rho^{-k} \mu_V(B_\rho(x))$ is monotonic, it follows that $\mu_V(B_\rho(x)) \geqslant \omega_k e^{-\|\mathbf{H}\|_\infty \rho} \rho^k$.

Lastly, (iv) follows from Theorem 1.12 (showing that $\theta_V(x) = 0$ for $\mathcal{H}^k$-a.e. $x \in U \setminus \Gamma$) coupled with conclusion (ii). □

3.3 The Tilt-Excess Inequality

In this section we introduce and prove a counterpart to Caccioppoli's Inequality for use in the theory of varifolds. We first prove the following Lemma which is also used later in the proof of Allard's Theorem.

Lemma 3.2. *Let π and T be k-dimensional planes in $\mathbb{R}^N$, and let $X \in C(\mathbb{R}^N; \mathbb{R}^N)$ be the vector field $X(x) := P_\pi^\perp(x)$. If $\{v_{k+1}, \ldots, v_N\}$ is an ON basis of $\pi^\perp$, and $f_j(x) := x \cdot v_j$, then*

$$\tfrac{1}{2}\|T - \pi\|^2 = \mathrm{div}_T X = \sum_{j=k+1}^{N} |\nabla_T f_j|^2.$$

Moreover, there is a positive dimensional constant $C_0 = C_0(N, k)$ such that

$$|\mathrm{J}_T P_\pi - 1| \leqslant C_0 \|T - \pi\|^2$$

Proof. Let $\xi_1, \ldots \xi_k$ be an ON basis of T and $e_{k+1}, \ldots, e_N$ an ON basis of $T^\perp$. In coordinates we thus have

$$P_\pi = I - \sum_{j=k+1}^{N} v_j \otimes v_j \qquad \text{and} \qquad P_T = I - \sum_{j=k+1}^{N} e_j \otimes e_j.$$

We therefore compute

$$\begin{aligned}
\frac{1}{2}\|\pi - T\|^2 = \frac{1}{2}\|P_\pi - P_T\|^2 &= \frac{1}{2}\left\|\sum_{i=k+1}^{N} v_i\otimes v_i - \sum_{j=k+1}^{N} e_j\otimes e_j\right\|^2 \\
&= \frac{1}{2}\left\langle \sum_{i=k+1}^{N} v_i\otimes v_i - \sum_{j=k+1}^{N} e_j\otimes e_j : \sum_{i=k+1}^{N} v_i\otimes v_i - \sum_{j=k+1}^{N} e_j\otimes e_j \right\rangle \\
&= \frac{1}{2}\sum_{i,j=k+1}^{N}\langle v_i\otimes v_i : v_j\otimes v_j\rangle + \frac{1}{2}\sum_{i,j=k+1}^{N}\langle e_i\otimes e_i : e_j\otimes e_j\rangle \\
&\qquad - \sum_{i,j=k+1}^{N}\langle v_i\otimes v_i : e_j\otimes e_j\rangle \\
&= \frac{1}{2}\sum_{i,j=k+1}^{N}(v_i\cdot v_j)^2 + \frac{1}{2}\sum_{i,j=k+1}^{N}(e_i\cdot e_j)^2 - \sum_{i,j=k+1}^{N}(v_i\cdot e_j)^2 \\
&= (N-k) - \sum_{i,j=k+1}^{N}(v_i\cdot e_j)^2 = \sum_{i=k+1}^{N}\left(1 - \sum_{j=k+1}^{N}(v_i\cdot e_j)^2\right) \\
&= \sum_{i=k+1}^{N}\sum_{j=1}^{k}(v_i\cdot\xi_j)^2 = \sum_{i=k+1}^{N}|\nabla_T f_i|^2
\end{aligned}$$

which proves half of the first claim. We can also compute, starting at the last line,

$$\begin{aligned}
\sum_{i=k+1}^{N}|\nabla_T f_i|^2 = \sum_{i=k+1}^{N}\sum_{j=1}^{k}(v_i\cdot\xi_j)^2 = \sum_{j=1}^{k}\xi_j\cdot\left(\sum_{i=k+1}^{N}(v_i\cdot\xi_j)v_i\right) &= \sum_{j=1}^{k}\sum_{i=k+1}^{N}\xi_j\cdot(D_{\xi_j}f_i)v_i \\
= \sum_{j=1}^{k}\xi_j\cdot\left(D_{\xi_j}\sum_{i=k+1}^{N}f_i v_i\right) &= \sum_{j=1}^{k}\xi_j\cdot D_{\xi_j}X = \mathrm{div}_T X
\end{aligned}$$

Next we prove the estimate $|\mathrm{J}_T P_\pi - 1| \leqslant C_0\|T-\pi\|^2$, and begin by recalling that $\mathrm{J}_T P_\pi := \sqrt{\det(M^t M)}$ where M is a matrix representation of P_π restricted to T. Thus

$$\begin{aligned}
(M^t M)_{ij} = P_\pi(\xi_i)\cdot P_\pi(\xi_j) &= \delta_{ij} - P_\pi^\perp(\xi_i)\cdot\xi_j - P_\pi^\perp(\xi_j)\cdot\xi_i + P_\pi^\perp(\xi_i)\cdot P_\pi^\perp(\xi_j) \\
&= \delta_{ij} - A_{ij} - A_{ji} + B_{ij}
\end{aligned}$$

We have that $|P_\pi^\perp(\xi_i)| = |P_\pi^\perp(\xi_i) - P_T^\perp(\xi_i)| \leqslant C\|P_\pi^\perp - P_T^\perp\| = C\|P_\pi - P_T\| = C\|T-\pi\|$, so if A and B are the matrices with respective entries A_{ij}, B_{ij}, then $\|A\| \leqslant C\|T-\pi\|$ and $\|B\| \leqslant C\|T-\pi\|^2$. Moreover, we have that, by the above calculations,

$$\mathrm{trace} A = \sum_{i=1}^{k} A_{ii} = \sum_{i=1}^{k} P_\pi^\perp(\xi_i)\cdot\xi_i = \sum_{i=1}^{k}\sum_{j=k+1}^{N}((\xi_i\cdot v_j)v_j)\cdot\xi_i = \sum_{i=1}^{k}\sum_{j=k+1}^{N}(\xi_i\cdot v_j)^2 = \mathrm{div}_T X$$

By employing the Taylor expansion of the determinant we determine that

$$\det(M^t M) = 1 - 2\mathrm{trace}(A) + \mathcal{O}(\|T-\pi\|^2) = 1 - 2\mathrm{div}_T X + \mathcal{O}(\|T-\pi\|^2)$$

which together with the first part yields, since $\mathrm{J}_T P_\pi \geqslant 0$,

$$|\mathrm{J}_T P_\pi - 1| \leqslant |\mathrm{J}_T P_\pi - 1|(\mathrm{J}_T P_\pi + 1) = |\det(M^t M) - 1| \leqslant 2|\mathrm{div}_T X| + \mathcal{O}(\|T-\pi\|^2) \leqslant C_0\|T-\pi\|^2$$

□

With this result in hand, we can now prove the main assertion of the current section.

Proposition 3.4 (Tilt-Excess Inequality)**.** *Let $k < N$ be a positive integer. There exists a dimensional constant C such that if V is an integer rectifiable varifold with bounded generalized mean curvature $\mathbf{H}$ in the ball $B_r(x_0) \subset \mathbb{R}^N$, and π is a k-dimensional plane, then*

$$\mathbb{E}(V,\pi,x_0,r/2) \leqslant \frac{C}{r^{k+2}}\int_{B_r(x_0)} \operatorname{dist}(y-x_0,\pi)^2 \mathrm{d}\mu_V(y) + \frac{2^{k+1}}{r^{k-2}}\int_{B_r(x_0)} |\mathbf{H}|^2 \mathrm{d}\mu_V$$

Proof. By scaling and translating, we may assume that $x_0 = 0$ and $r = 1$. We let X be the vector field $X(x) := P_\pi^\perp(x)$ of the previous lemma, and choose a smooth cutoff function $\zeta \in C_c^\infty(B_1;[0,\infty))$ with $\zeta \equiv 1$ on $B_{1/2}$. We test the first variation

$$\delta(X) = \int_{B_1} \operatorname{div}_{T_x\Gamma} X \mathrm{d}\mu_V(x) = -\int_{B_1} X\cdot\mathbf{H}\mathrm{d}\mu_V$$

with the vector field $\zeta^2 X$ yielding, since $\operatorname{div}_{T_x\Gamma}(\zeta^2 X) = 2\zeta(\nabla_{T_x\Gamma}\zeta)\cdot X + \zeta^2 \operatorname{div}_{T_x\Gamma} X$,

$$\int_{B_1} \zeta^2 \operatorname{div}_{T_x\Gamma} X \mathrm{d}\mu_V(x) = -\int_{B_1} \zeta^2 X\cdot\mathbf{H}\mathrm{d}\mu_V - 2\int_{B_1} \zeta(\nabla_{T_x\Gamma}\zeta)\cdot X\mathrm{d}\mu_V(x)$$

If for a moment we set $T = T_x\Gamma$, and take an ON basis $\{\xi_1,\ldots,\xi_k\}$ of T, $\{v_{k+1},\ldots,v_N\}$ an ON basis of $\pi^\perp$, and $f_j(x) := x\cdot v_j$ as in the previous lemma, then we can estimate the last integrand pointwise at $x \in B_1$ using the observations

- $|\xi_i \cdot X(x)| = |\xi_i \cdot \sum_{j=k+1}^N (x\cdot v_j)v_j| \leqslant \sum_{j=k+1}^N |(x\cdot v_j)(\xi_i\cdot v_j)| = \sum_{j=k+1}^N |f_j(x)||(\xi_i\cdot v_j)|$
- $|f_j(x)| \leqslant \left(\sum_{i=k+1}^N |f_i(x)|^2\right)^{\frac12} = \left(\sum_{i=k+1}^N |x\cdot v_i|^2\right)^{\frac12} = |X(x)|$
- $\nabla_T f_j(x) = P_T(\nabla f_j(x)) = \sum_{i=1}^k (\xi_i\cdot\nabla f_j(x))\xi_i = \sum_{i=1}^k (\xi_i\cdot v_j)\xi_i \implies \sum_{i=1}^k |\xi_i\cdot v_j| \leqslant C|\nabla_T f_j(x)|$

as follows:

$$\zeta\,|\nabla_T\zeta\cdot X| = \zeta\left|\sum_{i=1}^k (\nabla\zeta\cdot\xi_i)(\xi_i\cdot X)\right| \leqslant \sum_{i=1}^k\sum_{j=k+1}^N \zeta|\nabla\zeta||f_j||\xi_i\cdot v_j| \leqslant C\sum_{j=k+1}^N \zeta|\nabla\zeta||X||\nabla_T f_j|$$

applying Young's inequality with $p = q = 2$ to each term $(C|\nabla\zeta||X|)\left(\frac12\zeta|\nabla_T f_j|\right)$ in the sum yields

$$\zeta\,|\nabla_T\zeta\cdot X| \leqslant C|\nabla\zeta|^2|X|^2 + \tfrac14\zeta^2\sum_{j=k+1}^N |\nabla_T f_j|^2 = C|\nabla\zeta|^2|X|^2 + \tfrac14\zeta^2 \operatorname{div}_T X$$

where in the last equality we applied Lemma 3.2. Applying this estimate to our original quantity (using Young's inquality once again in the first integral)

$$\begin{aligned}\int_{B_1} \zeta^2 \operatorname{div}_{T_x\Gamma} X\mathrm{d}\mu_V(x) &= -\int_{B_1}\zeta^2 X\cdot\mathbf{H}\mathrm{d}\mu_V - 2\int_{B_1}\zeta(\nabla_{T_x\Gamma}\zeta)\cdot X\mathrm{d}\mu_V(x)\\ &\leqslant \int_{B_1}|\zeta|^2|X||\mathbf{H}|\mathrm{d}\mu_V + 2\int_{B_1}\zeta|\nabla_{T_x\Gamma}\zeta\cdot X|\mathrm{d}\mu_V(x)\\ &\leqslant \frac12\int_{B_1}|\zeta|^4|\mathbf{H}|^2\mathrm{d}\mu_V + \frac12\int_{B_1}|X|^2\mathrm{d}\mu_V\\ &\qquad + C\int_{B_1}|\nabla\zeta|^2|X|^2\mathrm{d}\mu_V + \frac12\int_{B_1}\zeta^2\operatorname{div}_{T_x\Gamma}X\mathrm{d}\mu_V(x)\end{aligned}$$

which implies that

$$\frac12\int_{B_1}\zeta^2\operatorname{div}_{T_x\Gamma}X\mathrm{d}\mu_V(x) \leqslant \frac12\int_{B_1}|\zeta|^4|\mathbf{H}|^2\mathrm{d}\mu_V + \frac12\int_{B_1}|X|^2\mathrm{d}\mu_V + C\int_{B_1}|\nabla\zeta|^2|X|^2\mathrm{d}\mu_V$$

On the other hand, applying Lemma 3.2 again tells us that

$$\mathbb{E}(V,\pi,0,\tfrac{1}{2}) = 2^k\int_{B_{1/2}} \|T_x\Gamma-\pi\|^2 \mathrm{d}\mu_V(x) = 2^{k+1}\int_{B_{1/2}} \mathrm{div}_{T_x\Gamma}X\mathrm{d}\mu_V(x) \leqslant 2^{k+2}\cdot\frac{1}{2}\int_{B_1}\zeta^2\mathrm{div}_{T_x\Gamma}X\mathrm{d}\mu_V(x)$$

since $\zeta\equiv 1$ on $B_{1/2}$. But recall that $\zeta\in C_c^\infty$, so that $\|\nabla\zeta\|_\infty\leqslant C$. Therefore, since $|X(x)| = \mathrm{dist}(x,\pi)$,

$$\begin{aligned}\mathbb{E}(V,\pi,0,\tfrac{1}{2}) &\leqslant 2^{k+2}\cdot\frac{1}{2}\int_{B_1}\zeta^2\mathrm{div}_{T_x\Gamma}X\mathrm{d}\mu_V(x)\\ &\leqslant 2^{k+1}\int_{B_1}|\zeta|^4|\mathbf{H}|^2\mathrm{d}\mu_V + 2^{k+1}\int_{B_1}|X|^2\mathrm{d}\mu_V + C\int_{B_1}|\nabla\zeta|^2|X|^2\mathrm{d}\mu_V\\ &\leqslant 2^{k+1}\int_{B_1}|\mathbf{H}|^2\mathrm{d}\mu_V + C\int_{B_1}\mathrm{dist}(x,\pi)^2\mathrm{d}\mu_V(x)\end{aligned}$$

just as advertised. □

3.4 The Lipschitz Approximation

If the generalized mean curvature and the excess of our varifold are both small, then intuitively the varifold does not "wiggle too much" when compared with a plane. In this situation, it might thus be expected that the varifold (or at least a decent part of it) be realizable as some nice object, say as the image of a map with some nice regularity. Indeed, this is what happens–a sufficiently well behaved varifold can be realized as a Lipschitz image over a plane, at least on a slightly smaller scale. This approximation is applied several times throughout the proofs of the Excess Decay Theorem and Allard's Theorem, but is of course of interest in its own right.

Theorem 3.2 (The Lipschitz Approximation)**.** *Let $k<N$ be a positive integer. Then there exists a $C>0$ satisfying the following. For any fixed $l,\beta\in(0,1)$ there are $\lambda=\lambda(l)\in(0,1]$ and $\varepsilon_L>0$ such that if $V=(\Gamma,\theta)$ satisfies Allard's conditions with $\varepsilon=\varepsilon_L$, then there is a Lipschitz map $f\colon (x_0+\pi)\cap B_{r/4}(x_0)\to\pi^\perp$ such that*

(i) $Lip(f)<l$ and $\Gamma_f\subset I_{\beta r}(x_0+\pi)$

(ii) $\theta\equiv 1$ $\mathcal{H}^k$-a.e. on $\Gamma\cap B_{r/4}(x_0)\subset I_{\beta r}(x_0+\pi)$

(iii) $G\subset\Gamma_f$, where $G:=\{x\in\Gamma\cap B_{r/4}(x_0) : \mathbb{E}(V,\pi,x,\rho)\leqslant\lambda \quad \forall\rho\in(0,r/2]\}$

(iv) $\mathcal{H}^k(\Gamma_f\setminus G)+\mathcal{H}^k((\Gamma\cap B_{r/4}(x_0))\setminus G)\leqslant C\lambda^{-1}\mathbb{E}(V,\pi,x_0,r)r^k + C\|\mathbf{H}\|_\infty r^{k+1}$

In the effort of proving this result, we prove two preliminary lemmas. The first lemma is of particular interest, as leverages the versatile compactness properties of Radon measures in extracting convergence from particular sequences of varifolds. This is then used to prove a technical lemma which provides a "height bound" for the varifold with respect to a plane, allowing us to transfer certain data to smaller scales.

From here on the notation $\mathcal{B}_r(x):=B_r(x)\cap\pi$ is assumed to be in force, whenever it is clear what π is.

Lemma 3.3. *Let $V_i=(\Gamma_i,\theta_i)$ be a sequence of k-dimensional integral varifolds in $B_1\subset\mathbb{R}^n$ satisfying Allard's conditions with $\varepsilon=\varepsilon(V_i)\downarrow 0$ for a given fixed plane π. Then $\mu_{V_i}\overset{*}{\rightharpoonup}\mathcal{H}^k\llcorner\pi$ in B_1.*

Proof. Fix $\rho\in(0,1)$ and let $\mathbf{H}_i$ be the generalized mean curvature vector of V_i. Then by assumption $V_i=(\Gamma_i,\theta_i)$ satisfies $\mu_{V_i}(B_1)\leqslant(\omega_k+\varepsilon_i)$ and $\|\mathbf{H}_i\|_\infty<\varepsilon_i$.

First of all, observe that up to a subsequence we have $\mu_{V_i}\overset{*}{\rightharpoonup}\mu$ for some Radon measure μ on B_1, since

$$\sup_i \mu_{V_i}(B_1)\leqslant\sup_i(\omega_k+\varepsilon_i)<\infty$$

. We will show that $\mu=\mathcal{H}^k\llcorner\mathcal{B}_1$ where $\mathcal{B}_1=B_1\cap\pi$, and as a first step in this direction we prove that $\mathrm{spt}\mu\subset\mathcal{B}_1$. In particular, fix an arbitrary $\phi\in C_c(B_1)$ and consider the integral

$$\int_{B_1}|P_\pi^\perp(y)|^2\phi(y)\mathrm{d}\mu(y).$$

If this integral vanishes for all $\phi \in C_c(B_1)$, then μ must to be supported in $\mathcal{B}_1$. For if there is a $x \in B_1 \setminus \pi$ which is in sptμ, then for a small enough $r > 0$ we have both $B_r(x) \subset B_1 \setminus \pi$ and $\mu(B_r(x)) > 0$, and testing the integral with a nontrivial $\phi \in C_c(B_r(x)) \subset C_c(B_1)$ yields a contradiction.

Notice that $|P_\pi^\perp(y)|^2\phi \in C_c(B_1)$, and so by the weak-$*$ convergence of the μ_{V_i},

$$\int_{B_1} |P_\pi^\perp(y)|^2\phi(y)\mathrm{d}\mu(y) = \lim_{i\to\infty}\int_{B_1} |P_\pi^\perp(y)|^2\phi(y)\mathrm{d}\mu_{V_i}(y) \leqslant \|\phi\|_\infty \lim_{i\to\infty}\int_{B_1} |P_\pi^\perp(y)|^2\mathrm{d}\mu_{V_i}(y).$$

We show that this limit is 0. If $y \in B_1 \setminus B_\rho$, $\hat{y} = y/|y|$, then we observe that

$$|P_{T_y\Gamma_i}^\perp(y)|^2 \leqslant |P_{T_y\Gamma_i}^\perp(\hat{y})|^2 = |P_{T_y\Gamma_i}^\perp(\nabla|y|)|^2 = |\nabla_{T_y\Gamma_i}^\perp|y||^2 \leqslant \frac{|\nabla_{T_y\Gamma_i}^\perp|y||^2}{|y|^k}.$$

Therefore, we can estimate

$$\begin{aligned}
\int_{B_1} |P_\pi^\perp(y)|^2\mathrm{d}\mu_{V_i}(y) &\leqslant 2\int_{B_1} |P_\pi^\perp(y) - P_{T_y\Gamma_i}^\perp(y)|^2\mathrm{d}\mu_{V_i}(y) + 2\int_{B_1} |P_{T_y\Gamma_i}^\perp(y)|^2\mathrm{d}\mu_{V_i}(y)\\
&\leqslant C\int_{B_1} \|T_y\Gamma_i - \pi\|^2\mathrm{d}\mu_{V_i}(y) + 2\int_{B_1\setminus B_\rho} |P_{T_y\Gamma_i}^\perp(y)|^2\mathrm{d}\mu_{V_i}(y) + 2\int_{B_\rho} |P_{T_y\Gamma_i}^\perp(y)|^2\mathrm{d}\mu_{V_i}(y)\\
&\leqslant C\mathbb{E}(V_i,\pi,0,1) + C\int_{B_1\setminus B_\rho} \frac{|\nabla_{T_y\Gamma_i}^\perp|y||^2}{|y|^k}\mathrm{d}\mu_{V_i}(y) + C\rho^2\mu_{V_i}(B_\rho)\\
&\leqslant C\varepsilon_i + C\int_{B_1\setminus B_\rho} \frac{|\nabla_{T_y\Gamma_i}^\perp|y||^2}{|y|^k}\mathrm{d}\mu_{V_i}(y) + C\rho^2(\omega_k + \varepsilon_i)
\end{aligned}$$

By the Monotonicity Formula and its Corollary we establish that the second term obeys the estimate

$$\begin{aligned}
\int_{B_1\setminus B_\rho} \frac{|\nabla_{T_y\Gamma_i}^\perp r|^2}{r^k}\mathrm{d}\mu_{V_i}(y) &= \mu_{V_i}(B_1) - \frac{\mu_{V_i}(B_\rho)}{\rho^k} - \int_{B_1} \frac{1}{k} y\cdot\mathbf{H}_i\left(\frac{1}{m(r)^k} - 1\right)\mathrm{d}\mu_{V_i}(y)\\
&\leqslant \mu_{V_i}(B_1) - \omega_k e^{-\|\mathbf{H}_i\|_\infty\rho} + \int_{B_1}\frac{1}{k}|y|\|\mathbf{H}_i\|_\infty\left(\frac{1}{\rho^k} - 1\right)\mathrm{d}\mu_{V_i}(y)\\
&\leqslant \mu_{V_i}(B_1) - \omega_k e^{-\|\mathbf{H}_i\|_\infty} + C\|\mathbf{H}_i\|_\infty\\
&\leqslant (\omega_k + \varepsilon_i) - \omega_k e^{-\varepsilon_i} + C\varepsilon_i
\end{aligned}$$

so that

$$\int_{B_1} |P_\pi^\perp(y)|^2\mathrm{d}\mu_{V_i}(y) \leqslant C\varepsilon_i + \{(\omega_k + \varepsilon_i) - \omega_k e^{-\varepsilon_i} + C\varepsilon_i\} + C\rho^2(\omega_k + \varepsilon_i)$$

Taking $i \to \infty$ therefore yields for any $\rho \in (0,1)$

$$\lim_{i\to\infty}\int_{B_1} |P_\pi^\perp(y)|^2\mathrm{d}\mu_{V_i}(y) \leqslant C\rho^2.$$

Since $\rho \in (0,1)$ was arbitrary, we conclude that indeed

$$\int_{B_1} |P_\pi^\perp(y)|^2\phi(y)\mathrm{d}\mu(y) = \lim_{i\to\infty}\int_{B_1} |P_\pi^\perp(y)|^2\phi(y)\mathrm{d}\mu_{V_i}(y) \leqslant \|\phi\|_\infty \lim_{i\to\infty}\int_{B_1} |P_\pi^\perp(y)|^2\mathrm{d}\mu_{V_i}(y) = 0$$

for any $\phi \in C_c(B_1)$, thereby proving that μ is supported in $\mathcal{B}_1$.

We now apply Proposition 1.4 to conclude that $\mu = \theta\mathcal{H}^k \llcorner \mathcal{B}_1$ for some Borel map $\theta\colon \mathcal{B}_1 \to \mathbb{R}^{\geqslant 0}$. Indeed, we can show that the k-dimension upper Hausdorff density $\theta^{*k}(\mu)$ of μ is finite everywhere in $\mathcal{B}_1$. Let $x \in \mathcal{B}_1$ and fix a $\rho \in (0, 1 - |x|)$. By the Monotonicity Formula, and the fact that by weak-$*$ convergence $\mu(U) \leqslant \liminf_{i\to\infty}\mu_{V_i}(U)$ for every open set U, we see that

$$\frac{\mu(B_\rho(x))}{\omega_k\rho^k} \leqslant \liminf_{i\to\infty}\frac{\mu_{V_i}(B_\rho(x))}{\omega_k\rho^k} \leqslant \liminf_{i\to\infty}\frac{e^{\|\mathbf{H}_i\|_\infty}\mu_{V_i}(B_{1-|x|}(x))}{\omega_k(1-|x|)^k} \leqslant \liminf_{i\to\infty}\frac{e^{\varepsilon_i}(\omega_k+\varepsilon_i)}{\omega_k(1-|x|)^k} = \frac{1}{(1-|x|)^k}.$$

Therefore

$$0 \leqslant \theta^{*k}(\mu)(x) = \limsup_{\rho \searrow 0} \frac{\mu(B_\rho(x))}{\omega_k \rho^k} \leqslant \frac{1}{(1-|x|)^k} < \infty$$

for every $x \in \mathcal{B}_1$. By Proposition 1.4, we conclude that $\mu = \theta \mathcal{H}^k$ for a locally integrable Borel map $\theta \colon \mathbb{R}^N \to \mathbb{R}^{\geqslant 0}$ (hence integrable on $\mathcal{B}_1$), and since $\operatorname{spt}\mu \subset \mathcal{B}_1$ we have that $\mu = \theta \mathcal{H}^k \llcorner \mathcal{B}_1$.

We now proceed onward to show that $\theta \equiv 1$ on $\mathcal{B}_1$, first showing that for any $X \in C^1_c(\mathcal{B}_1; \mathbb{R}^N)$ we have

$$\int_{\mathcal{B}_1} \theta(y) \operatorname{div}_\pi X(y) \mathrm{d}\mathcal{H}^k(y) = 0$$

where for convenience we have chosen an ON basis $\{e_j\}$ of $\mathbb{R}^N$ with corresponding coordinate system $\{y_1, \ldots, y_k, z_{N-k}, \ldots, z_N\}$ such that $\pi = \{z = 0\}$. We first observe that at a point $x \in \mathcal{B}_1 \times \mathbb{R}^{N-k} \cap \Gamma_i$,

$$\operatorname{div}_\pi X = (\operatorname{div}_\pi X - \operatorname{div}_{T_x\Gamma_i} X) + \operatorname{div}_{T_x\Gamma_i} X = (\operatorname{div}(P_\pi X) - \operatorname{div}(P_{T_x\Gamma_i} X)) + \operatorname{div}_{T_x\Gamma_i} X.$$

If we let (π_{ij}) and (T_{ij}) denote the matrix representations of the projections P_π and $P_{T_x\Gamma_i}$ in the standard basis, then

$$\begin{aligned}\operatorname{div}(P_\pi X) - \operatorname{div}(P_{T_x\Gamma_i} X) &= \sum_{k=1}^N \partial_k (P_\pi X)^k - \partial_k (P_{T_x\Gamma_i} X)^k = \sum_{k=1}^N \partial_k(\pi_{kj} X^j) - \partial_k(T_{kj} X^j) \\ &= \sum_{k=1}^N (\pi_{kj} - T_{kj}) \partial_k X^j\end{aligned}$$

so that

$$|\operatorname{div}(P_\pi X) - \operatorname{div}(P_{T_x\Gamma_i} X)| \leqslant C \|DX\|_\infty \|T_x\Gamma_i - \pi\|.$$

We can thus estimate

$$\begin{aligned}\left| \int_{\mathbb{R}^N} \operatorname{div}_\pi X \mathrm{d}\mu \right| &= \left| \int_{\mathcal{B}_1} \theta \operatorname{div}_\pi X \mathrm{d}\mathcal{H}^k \right| = \lim_{i\to\infty} \left| \int_{B_1} \operatorname{div}_\pi X \mathrm{d}\mu_{V_i} \right| \\ &\leqslant \liminf_{i\to\infty} \left\{ C\|DX\|_\infty \int_{B_1} \|T_x\Gamma_i - \pi\| \mathrm{d}\mu_{V_i}(x) + \left| \int_{B_1} \operatorname{div}_{T_x\Gamma_i} X \mathrm{d}\mu_{V_i}(x) \right| \right\} \\ &\leqslant \liminf_{i\to\infty} \left\{ C\|DX\|_\infty \int_{B_1} \|T_x\Gamma_i - \pi\| \mathrm{d}\mu_{V_i}(x) + \left| \int_{B_1} X \cdot \mathbf{H}_i \mathrm{d}\mu_{V_i} \right| \right\} \\ &\leqslant \liminf_{i\to\infty} \left\{ C\|DX\|_\infty \left(\int_{B_1} \|T_x\Gamma_i - \pi\|^2 \mathrm{d}\mu_{V_i}(x) \right)^{\frac12} (\mu_{V_i}(B_1))^{\frac12} + \|X\|_\infty \|\mathbf{H}_i\|_\infty \mu_{V_i}(B_1) \right\} \\ &= \liminf_{i\to\infty} \left\{ C\|DX\|_\infty \mathbb{E}(V_i, \pi, 0, 1)^{\frac12} (\mu_{V_i}(B_1))^{\frac12} + \|X\|_\infty \|\mathbf{H}_i\|_\infty \mu_{V_i}(B_1) \right\} \\ &\leqslant \liminf_{i\to\infty} \left\{ C\|DX\|_\infty \varepsilon_i^{\frac12} (\omega_k + \varepsilon_i)^{\frac12} + \|X\|_\infty \varepsilon_i (\omega_k + \varepsilon_i) \right\} = 0\end{aligned}$$

Now let $Y \in C^1_c(\mathcal{B}_1; \pi \times \{0\}^{N-k})$ (recalling our coordinate choice making $\pi = \{z = 0\}$), extend it to $\mathcal{B}_1 \times \mathbb{R}^{N-k}$ by making it constant on the sets $\{x\} \times \mathbb{R}^{N-k}$ for $x \in \mathcal{B}_1$, and multiply by a smooth cutoff function in the variables z to obtain a vector field $X \in C^1_c(\mathbb{R}^N; \mathbb{R}^N)$ which is supported in $\mathcal{B}_1 \times \mathbb{R}^{N-k}$ while agreeing with Y on $\mathcal{B}_1$. By the above observations, we conclude that

$$\int_{\mathcal{B}_1} \theta(y) \operatorname{div}_\pi Y(y) \mathrm{d}\mathcal{H}^k(y) = \int_{\mathbb{R}^N} \theta(x) \operatorname{div}_\pi Y(x) \mathrm{d}(\mathcal{H}^k \llcorner \mathcal{B}_1)(x) = \int_{\mathbb{R}^N} \operatorname{div}_\pi X(x) \mathrm{d}\mu(x) = 0.$$

By the arbitrariness of $Y \in C^1_c(\mathcal{B}_1; \pi \times \{0\}^{N-k})$, we conclude that θ is a constant on $\mathcal{B}_1$. Indeed, take a standard radial mollifier Ψ_δ on $\mathcal{B}_\delta$, and test with vector fields $Y \in C^1_c(\mathcal{B}_{1-\delta}; \pi \times \{0\}^{N-k})$ to establish that $(\theta * \Psi_\delta)' \equiv 0$ on the set $\mathcal{B}_{1-\delta}$. We then send δ to 0 to conclude that θ is constant on $\mathcal{B}_1$. To wit, we have that in coordinates

$$\operatorname{div}_\pi(Y * \Psi_\delta) = \sum_{j=1}^k e_j \cdot D_{e_j}(Y * \Psi_\delta) = \sum_{j=1}^k \left(\left(Y * \frac{\partial \Psi_\delta}{\partial y_j} \right)(y) \right)_j = \sum_{j=1}^k \int_{\mathcal{B}_{1-\delta}} Y_j(t) \frac{\partial \Psi_\delta}{\partial y_j}(y - t) \mathrm{d}\mathcal{H}^k(t)$$

which by Fubini's Theorem (as Ψ_δ is compactly supported) yields

$$
\begin{aligned}
0 = \int_{\mathcal{B}_1} \theta(y)\mathrm{div}_\pi(Y * \Psi_\delta)\mathrm{d}\mathcal{H}^k(y) &= \sum_{j=1}^k \int_{\mathcal{B}_1}\int_{\mathcal{B}_{1-\delta}} Y_j(t)\theta(y)\frac{\partial \Psi_\delta}{\partial y_j}(y-t)\mathrm{d}\mathcal{H}^k(t)\mathrm{d}\mathcal{H}^k(y) \\
&= \sum_{j=1}^k \int_{\mathcal{B}_{1-\delta}} Y_j(t)\int_{\mathcal{B}_1} \theta(y)\frac{\partial \Psi_\delta}{\partial y_j}(y-t)\mathrm{d}\mathcal{H}^k(y)\mathrm{d}\mathcal{H}^k(t) \\
&= \sum_{j=1}^k \int_{\mathcal{B}_{1-\delta}} Y_j(t)\int_{\mathcal{B}_1} \theta(y)\frac{\partial \Psi_\delta}{\partial y_j}(-(t-y))\mathrm{d}\mathcal{H}^k(y)\mathrm{d}\mathcal{H}^k(t) \\
&= \sum_{j=1}^k \int_{\mathcal{B}_{1-\delta}} Y_j(t)\frac{\partial}{\partial y_j}(\theta * \Psi_\delta)(-t)\mathrm{d}\mathcal{H}^k(t)
\end{aligned}
$$

By the arbitrariness of $Y \in C_c^1(\mathcal{B}_{1-\delta}; \pi \times \{0\}^{N-k})$, it follows that $\frac{\partial}{\partial y_j}(\theta * \Psi_\delta)(t) = 0$ for all j and all $t \in \mathcal{B}_{1-\delta}$, so that $\theta * \Psi_\delta$ is a constant on $\mathcal{B}_{1-\delta}$. Since θ is integrable on $\mathcal{B}_1$, and since Ψ_δ can be chosen to have sufficiently fast decay, $\theta * \Psi_\delta \to \theta$ pointwise $\mathcal{H}^k$-a.e. on $\mathcal{B}_1$, and since θ is constant on each $\mathcal{B}_{1-\delta}$, sending δ to 0 tells us that indeed θ is constant on $\mathcal{B}_1$. To see that $\theta = 1$, we simply note that $\mu(\partial B_\rho) = 0$ for any $\rho \in (0,1)$, so that by weak-$*$ convergence $\theta\omega_k\rho^k = \theta\mu(B_\rho) = \lim_{i\to\infty}\mu_{V_i}(B_\rho)$. By part (iii) of the Corollary to the Monotonicity Formula

$$
\omega_k\rho^k = \lim_{i\to\infty}\omega_k e^{-\varepsilon_i\rho}\rho^k \leqslant \lim_{i\to\infty}\omega_k e^{-\|\mathbf{H}_i\|_\infty\rho}\rho^k \overset{(M.F.)}{\leqslant} \lim_{i\to\infty}\mu_{V_i}(B_\rho) \leqslant \lim_{i\to\infty}(\omega_k+\varepsilon_i)\rho^k = \omega_k\rho^k
$$

from which we conclude that $\theta \equiv 1$.

Lastly, we remark that this convergence is in fact enjoyed by the entire sequence we began with. Indeed, every subsequence of $\{\mu_{V_i}\}$ has, by the compactness theorem, a sub-subsequence which weak-$*$ converges to $\mu = \mathcal{H}^k \llcorner \mathcal{B}_1$. Therefore, $\mu_{V_i} \overset{*}{\rightharpoonup} \mu$.

□

We now prove the second preliminary lemma:

Lemma 3.4. *Let $k < N$ be a positive integer, and fix positive constants $\delta, \eta, \sigma < r$. Then there is an $\varepsilon_H > 0$ such that if $V = (\Gamma, \theta)$ satisfies Allard's conditions with $\varepsilon = \varepsilon_H$, then*

(i) $\Gamma \cap B_\eta(x_0) \subset I_{\delta r}(x_0 + \pi)$.

(ii) $\mu_V(B_\rho(x)) \leqslant (\omega_k + \delta)\rho^k$ *for every* $x \in B_{r-\sigma}(x_0)$ *and every* $\rho \in (0, \sigma]$.

Proof. After scaling and translating we can assume that $x_0 = 0$ and $r = 1$. Suppose to the contrary that the lemma were false, so that there exist some positive constants $\delta, \eta, \sigma < 1$ such that the following holds. For every $\varepsilon_H > 0$ there exists a varifold $V = (\Gamma, \theta)$ supported in B and satisfying Allard's conditions with respect to a plane π and $\varepsilon = \varepsilon_H$ such that at least one of the following alternatives holds:

1. There exists a point $x \in B_\eta \cap \Gamma$ such that $|P_\pi^\perp x| \geqslant \delta$.
2. There exists a point $x \in B_{r-\sigma}$ and a radius $\rho \leqslant \sigma$ such that $\mu_V(B_\rho(x)) > (\omega_k + \delta)\rho^k$

Take a sequence of such varifolds $V_i = (\Gamma_i, \theta_i)$ satisfying Allard's conditions with respect to the same fixed plane π (by rotating if necessary) with $\varepsilon_{H,i} \downarrow 0$. Let $\{x_i\}$ and $\{\rho_i\}$ be the corresponding sequences of points and radii satisfying, for each i, at least one of conditions 1, 2. By passing to a subsequence, we can assume that at least one of the two conditions holds for every x_i. We can thus work by cases. Note that, by Lemma 3.3, $\mu_{V_i} \overset{*}{\rightharpoonup} \mathcal{H}^k \llcorner \pi$.

Case 1 By compactness of $\overline{B}_\eta$, we can pass to a convergence subsequence, not relabeled. Suppose $x_i \to x$. Then $x \in \overline{B}_\eta$ and by continuity $|P_\pi^\perp(x)| \geqslant \delta$. Let $0 < t < \min\{\delta, 1-|x|\}$. Then $B_t(x) \subset B$ and $B_t(x) \cap \pi = \emptyset$. Now, for all i sufficiently large, we have $B_{t/2}(x_i) \subset B_t(x)$. By Corollary 3.1 to the Monotonicity Formula, part (ii), it follows that $\mu_{V_i}(B_{t/2}(x_i)) \geqslant C > 0$. Moreover, $\mathcal{H}^k(\partial B_t(x) \cap \pi) = 0$, so by Proposition 1.3 we see that by weak star convergence $\mu_{V_i}(B_t(x)) \to \mathcal{H}^k(B_t(x) \cap \pi)$. Putting everything together, we reach a contradiction:

$$0 = \mathcal{H}^k(B_t(x) \cap \pi) = \lim_{i\to\infty} \mu_{V_i}(B_t(x)) \geqslant \limsup_{i\to\infty} \mu_{V_i}(B_{t/2}(x_i)) \geqslant C > 0.$$

Case 2 By compactness of $\overline{B}_{1-\sigma}$, we can pass to a convergence subsequence, not relabeled. Suppose $x_i \to x$. Then $x \in \overline{B}_{1-\sigma}$. Now, since $0 < \rho_i \leqslant \sigma < 1 - |x_i| = \operatorname{dist}(x_i, \partial B)$ we can apply the Monotonicity Formula and property 2 of the sequence to conclude that

$$(\omega_k + \delta) \leqslant \frac{\mu_{V_i}(B_{\rho_i}(x_i))}{\rho_i^k} \leqslant \frac{\mu_{V_i}(B_\sigma(x_i))}{\sigma^k}.$$

Fix now any $t > \sigma$. Then for all i large $B_\sigma(x_i) \subset B_t(x)$. By comparing Hausdorff dimension, we see that $\mathcal{H}^k(\partial B_t(x) \cap \pi) = 0$. Thus, from Proposition 1.3 we have that

$$\mathcal{H}^k(B_t(x) \cap \pi) = \lim_{i\to\infty} \mu_{V_i}(B_t(x)) \geqslant \limsup_{i\to\infty} \mu_{V_i}(B_\sigma(x_i)) \geqslant (\omega_k + \delta)\sigma^k.$$

Let $t \downarrow \sigma$ to conclude the contradiction

$$\mathcal{H}^k(B_\sigma(x) \cap \pi) \geqslant (\omega_k + \delta)\sigma^k > \omega_k \sigma^k.$$

□

Remark 3.1. *For our subsequent uses, it suffices to just work with* $\delta < \frac{1}{2}$, $\eta = \frac{r}{2}$, *and* $\sigma = \frac{3r}{4}$.

We are now ready to prove the Lipschitz Approximation Theorem. The general idea is that if some part of the varifold does not wiggle too much with respect to a plane (in a uniform sense), then we should be able to flatten out that section of the varifold without causing any creasing or doubling up. In more precise terms, we take the set of points in the varifold that have uniformly small excess at all small scales, and then project this region down to the plane. We can prove that on this set the projection is injective, so that we can invert and obtain a Lipschitz parametrization of the "good" part of the varifold. We then can extend the map to obtain our approximation.

Proof of the Lipschitz Approximation. As always, we can assume by translating and scaling that $x_0 = 0$ and $r = 1$. We also set $\mathbb{E}(V, \pi, 0, 1) := \mathbb{E}$.

We need to produce a $\lambda \in (0, 1]$ and an $\varepsilon_L > 0$. To this end we make two choices:

(C1) We choose $\lambda < \min\{\omega_k, \varepsilon_H(\delta_1)\}$, where $\varepsilon_H(\delta_1)$ is the $\varepsilon_H > 0$ given by Lemma 3.4 when we take $\delta = \delta_1 = \frac{(N-k)^{-\frac{1}{2}} l}{6}$, $\eta = \frac{1}{2}$, and $\sigma = \frac{3}{4}$.

(C2) We next choose $\varepsilon_L < \min\{\lambda, \varepsilon_H(\delta_2)\}$ where $\varepsilon_H(\delta_2)$ is the $\varepsilon_H > 0$ given by Lemma 3.4 when we take $\delta = \delta_2 < \min\{\lambda, (N-k)^{-1/2}\beta\}$, $\eta = \frac{1}{2}$, and $\sigma = \frac{3}{4}$.

We first prove (ii). Let $V = (\Gamma, \theta)$ satisfy Allard's conditions with $\varepsilon \leqslant \varepsilon_L$, and fix any point $x \in \Gamma \cap B_{1/4}$. Since $\varepsilon_L < \varepsilon_H(\delta_2)$, by Lemma 3.4 (ii) we find that for all $\rho \in (0, 1/2)$

$$\mu_V(B_\rho(x)) \leqslant (\omega_k + \delta_2)\rho^k < (\omega_k + \lambda)\rho^k.$$

Dividing by $\omega_k \rho^k$ we find that

$$\frac{\mu_V(B_\rho(x))}{\omega_k \rho^k} < 1 + \frac{\lambda}{\omega_k} < 2,$$

and sending $\rho \to 0$ yields, by Corollary 3.1 to the Monotonicity Formula, $\theta_V(x) = \theta(x) < 2$. Since $\theta\colon \Gamma \to \mathbb{Z}^{\geqslant 1}$, we thus find that $\theta \equiv 1$ $\mathcal{H}^k$-a.e. on $\Gamma \cap B_{1/4}$. Moreover, by Lemma 3.4 (i), we see that

$$\Gamma \cap B_{1/4} \subset \Gamma \cap B_{1/2} \subset I_{\delta_2}(\pi) \subset I_{(N-k)^{-\frac{1}{2}}\beta}(\pi) \subset I_\beta(\pi).$$

Next we prove (i) and (iii). Let $x, y \in G$, and notice that $|x-y| < \frac{1}{2}$ since $x, y \in B_{1/4}$. Choose $r > |x-y|$ such that $r < \{\frac{1}{2}, \frac{3}{2}|x-y|\}$. Since $r < \frac{3}{4}$, and since $x \in B_{1/4}$, by Lemma 3.4 (ii)

$$\mu_V(B_r(x)) \leqslant (\omega_k + \delta_2)r^k < (\omega_k + \lambda)r^k < (\omega_k + \varepsilon_H(\delta_1))r^k.$$

By assumption, $x \in G$ implies that $\mathbb{E}(V, \pi, x, r) \leqslant \lambda < \varepsilon_H(\delta_1)$. Additionally, $\|\mathbf{H}\|_\infty \leqslant \varepsilon_L < \frac{\varepsilon_H(\delta_1)}{r}$, since $r < 1$. Thus, V satisfies Allard's conditions on the smaller ball $B_r(x)$ with respect to $\varepsilon \leqslant \varepsilon_H(\delta_1)$, so by Lemma 3.4 (i) we conclude that $\Gamma \cap B_{1/2}(x) \subset I_{\delta_1 r}(x+\pi)$. But then $y \in I_{\delta_1 r}(x+\pi)$, so $y - x \in I_{\delta_1 r}(\pi)$. Thus,

$$|P_\pi^\perp(y-x)| < 2\delta_1 r = \frac{1}{3}(N-k)^{-\frac{1}{2}} l r < \frac{1}{3}r < \frac{1}{2}|x-y|.$$

By the triangle inequality and the fact that $P_\pi^\perp$ is an orthogonal projection, we find that for any $x, y \in G$,

$$|P_\pi x - P_\pi y| \geqslant \frac{\sqrt{3}}{2}|x-y| \geqslant \frac{1}{2}|x-y|.$$

Thus, the map $P_\pi\colon G \to \pi$ is injective, and is invertible on its image $D \coloneqq P_\pi(G)$. Let $f\colon D \to \mathbb{R}^N$ denote the inverse map. By viewing $\mathbb{R}^N = \pi \oplus \pi^\perp$, we can interpret $f\colon D \to \pi^\perp$ and G as the graph $G = \Gamma_f = \{v + f(v) : v \in D\}$ of f over D. Notice also that $\|f\|_\infty \leqslant (N-k)^{-\frac{1}{2}}\beta$, since $G \subset I_{\delta_2}(\pi)$ and $\delta_2 < (N-k)^{-\frac{1}{2}}\beta$.

Moreover, we can show that f is Lipschitz, by observing that for any pair $v, w \in D$,

$$\begin{aligned}
|f(v) - f(w)| &= |P_\pi^\perp(v + f(v)) - P_\pi^\perp(w + f(w))| \\
&= |P_\pi^\perp(v + f(v) - w - f(w))| \\
&< \frac{1}{2}(N-k)^{-\frac{1}{2}} l |v + f(v) - w - f(w)| \\
&\leqslant (N-k)^{-\frac{1}{2}} l |P_\pi(v + f(v)) - P_\pi(w + f(w))| \\
&= (N-k)^{-\frac{1}{2}} l |v - w|
\end{aligned}$$

which shows that $\mathrm{Lip}(f, D) < (N-k)^{-\frac{1}{2}} l$. We can now extend f to all of $B_{1/4} \cap \pi$ in one of two ways. We could either apply Kirzbraun's Theorem directly to f, or we could use the more elementary (and constructive) McShane's Lemma to extend each coordinate function of f individually. To illustrate the lower-tech route, fix a system of orthonormal coordinates on $\pi^\perp$ such that $f = (f_1, \dots, f_{N-k})$. Then each $f_j\colon D \to \mathbb{R}$ has $\mathrm{Lip}(f_j, D) < (N-k)^{-\frac{1}{2}} l$, and can be extended to all of $\mathbb{R}^n$ while preserving this Lipschitz bound. In particular it can be extended to $B_{1/4} \cap \pi$ while also preserving $\|f_j\|_\infty \leqslant (N-k)^{-\frac{1}{2}}\beta$. The resulting extended function $f\colon B_{1/4} \cap \pi \to \pi^\perp$ then satisfies

$$\mathrm{Lip} f \leqslant (N-k)^{\frac{1}{2}}(N-k)^{-\frac{1}{2}} l = l$$

and

$$\|f\|_\infty \leqslant (N-k)^{\frac{1}{2}}(N-k)^{-\frac{1}{2}}\beta = \beta$$

just as desired, since $\|f\|_\infty \leqslant \beta$ implies that $\Gamma_f \subset I_\beta(\pi)$. Notice that as an immediate consequence, part (iii) follows as well, since $G \subset \Gamma_f$ by definition of f!

Lastly, we need to prove the estimate in (iv). For each $x \in F \coloneqq (\Gamma \setminus G) \cap B_{1/4}$ we can choose a $\rho_x < \frac{1}{2}$ such that $\mathbb{E}(V, \pi, x, \rho_x) > \lambda$. The collection $\{B_{\rho_x}(x)\}_{x \in F}$ covers F, and $\sup_x \rho_x \leqslant \frac{1}{2}$. By the $5r$ covering theorem, there exists a countable, disjoint subcollection $\{B_{\rho_i}(x_i)\}$ such that $F \subset \bigcup_i B_{5\rho_i}(x_i)$. Thus,

$$\mathcal{H}^k(F) \leqslant \sum_i \omega_k (5\rho_i)^k < \omega_k 5^k \lambda^{-1} \sum_i \mathbb{E}(V, \pi, x_i, \rho_i)\rho_i^k \leqslant C\lambda^{-1}\mathbb{E},$$

where we have used the observation that, by disjointness of the collection,

$$\mathbb{E} = \int_B \|T_y\Gamma - \pi\|^2 \mathrm{d}\mu_V(y) \geqslant \sum_i \int_{B_{\rho_i}(x_i)} \|T_y\Gamma - \pi\|^2 \mathrm{d}\mu_V(y) = \sum_i \mathbb{E}(V, \pi, x_i, \rho_i)\rho_i^k$$

Next, set $K = \Gamma_f \setminus G$. By the area formula, for any $y \in \Gamma$ we have

$$\mathcal{H}^k(P_\pi(G \cap T_y\Gamma)) = \int_G J_{T_y\Gamma} P_\pi \mathrm{d}\mathcal{H}^k$$

since $J_{T_y\Gamma}P_\pi$ is the Jacobian of P_π after being restricted to the k-plane $T_y\Gamma$, and $P_\pi(G\cap T_y\Gamma)$ is an (injective Lipshitz) image of this restriction. In particular, this formula is verifiable after action by a suitable element of $\mathcal{O}(n)$ bringing the plane $T_y\Gamma$ to $\mathbb{R}^k \hookrightarrow \mathbb{R}^N$. But then

$$\mathcal{H}^k(P_\pi(G)) \geqslant \mathcal{H}^k(P_\pi(G) \cap T_y\Gamma) = \int_G J_{T_y\Gamma} P_\pi \mathrm{d}\mathcal{H}^k$$

so we can compute

$$\mathcal{H}^k(K) = \mathcal{H}^k(\Gamma_f \setminus G) = \mathcal{H}^k(f(\pi \cap B_{1/4}) \setminus f(D)) \leqslant \mathcal{H}^k(f(B_{1/4} \setminus D)) \leqslant (\mathrm{Lip} f)^k \mathcal{H}^k(B_{1/4} \setminus D)$$

and thus

$$\mathcal{H}^k(K) \leqslant C(\omega_k 4^{-k} - \mathcal{H}^k(D)) = C(\omega_k 4^{-k} - \mathcal{H}^k(P_\pi(G))) \leqslant C\left(\omega_k 4^{-k} - \int_G J_{T_y\Gamma}P_\pi \mathrm{d}\mathcal{H}^k\right).$$

By Lemma 3.2 we have that $|J_{T_y\Gamma}P_\pi - 1| \leqslant C_0\|T_y\Gamma - \pi\|^2$ for some dimensional constant $C_0 > 0$. Observe also that by the area assumption (A1) of Allard's Theorem, $\mu_V(B_1) \leqslant C$ for a dimensional C. Therefore,

$$\begin{aligned}
\mathcal{H}^k(K) &\leqslant C\left(\omega_k 4^{-k} - \int_G \left(J_{T_y\Gamma}P_\pi - 1\right) \mathrm{d}\mathcal{H}^k - \mathcal{H}^k(G)\right) \\
&\leqslant C\left(\omega_k 4^{-k} + \int_G \left|J_{T_y\Gamma}P_\pi - 1\right| \mathrm{d}\mathcal{H}^k - \mathcal{H}^k(G)\right) \\
&\leqslant C\left(\omega_k 4^{-k} + C_0 \int_G \|T_y\Gamma - \pi\|^2 \mathrm{d}\mathcal{H}^k - \mathcal{H}^k(G)\right) \\
&= C\left(\omega_k 4^{-k} + C\|T_y\Gamma - \pi\|^2 - \mathcal{H}^k(G)\right)
\end{aligned}$$

Integrating this inequality over all $y \in B_1$ with respect to μ_V then yields

$$\mathcal{H}^k(K) \leqslant C(\omega_k 4^{-k} + C\mathbb{E} - \mathcal{H}^k(G)).$$

Next, since $\theta \equiv 1$ on $B_{1/4}$ by part (ii), we see that $\mathcal{H}^k(F) = \mathcal{H}^k(B_{1/4} \cap \Gamma) - \mathcal{H}^k(G) = \mu_V(B_{1/4}) - \mathcal{H}^k(G)$, so

$$\begin{aligned}
\mathcal{H}^k(K) &\leqslant C(\omega_k 4^{-k} + C\mathbb{E} + \mathcal{H}^k(F) - \mu_V(B_{1/4})) \\
&\leqslant C\lambda^{-1}\mathbb{E} + C(\omega_k 4^{-k} - \mu_V(B_{1/4})) \\
&= C\lambda^{-1}\mathbb{E} + C\omega_k 4^{-k}\left(1 - \frac{\mu_V(B_{1/4})}{\omega_k 4^{-k}}\right)
\end{aligned}$$

where we applied the estimate $\mathcal{H}^k(F) \leqslant C\lambda^{-1}\mathbb{E}$ from above. Recalling conclusion (ii) of Corollary 3.1, we conclude the estimate for $\mathcal{H}^k(K)$:

$$\begin{aligned}
\mathcal{H}^k(K) &\leqslant C\lambda^{-1}\mathbb{E} + C\omega_k 4^{-k}\left(1 - \frac{\mu_V(B_{1/4})}{\omega_k 4^{-k}}\right) \\
&\leqslant C\lambda^{-1}\mathbb{E} + C\left(1 - e^{-\|H\|_\infty 4^{-1}}\right) \\
&= C\lambda^{-1}\mathbb{E} + \mathcal{O}(\|H\|_\infty) \\
&\leqslant C\lambda^{-1}\mathbb{E} + C\|H\|_\infty
\end{aligned}$$

Putting the estimates for $\mathcal{H}^k(F)$ and $\mathcal{H}^k(K)$ together yields conclusion (iv), completing the proof of the Lipschitz Approximation Theorem.

□

3.5 Harmonic Approximations

By our last result, an integer rectifiable varifold satisfying Allard's conditions is well approximated by the graph of a Lipschitz function. As we mentioned earlier, if we now jumped to Step 3 of the proof of Allard's Theorem, we could "immediately" conclude $C^{0,1}$ regularity of the varifold on a smaller scale. The next two lemmas are what allow us to do better, and garner $C^{1,\alpha}$ regularity. Intuitively, because the varifold has small bounded mean curvature, the Lipschitz graph approximating it must somehow be close to solving the minimal surface equation, which itself is a sort of perturbation of the Laplace equation. Thus, we might hope that nearby to f is a graph whose components are harmonic. To make this intuition work for us we prove a weakened form of Weyl's Lemma 1.1, which we will apply to the components of the Lipschitz approximation in the next part.

Lemma 3.5 (Harmonic Approximation)**.** *Let $k \geqslant 1$ and consider the ball $B_r(x) \subset \mathbb{R}^k$. For any $\rho > 0$, there exists an $\varepsilon_A > 0$ such that if $f \in H^1(B_r(x))$ with $\int |\nabla f|^2 \leqslant r^k$ satisfies*

$$\left| \int \nabla\phi \cdot \nabla f \right| \leqslant \varepsilon_A r^k \|\nabla\phi\|_\infty, \qquad \forall \phi \in C^1_c(B_r(x))$$

then there exists a harmonic function u on $B_r(x)$ with $\int |\nabla u|^2 \leqslant r^k$ and

$$\int (f-u)^2 \leqslant \rho r^{2+k}$$

Proof. Without loss of generality, by translating and scaling we can take $x_0 = 0$ and $r = 1$. Define the class

$$\mathcal{H} := \left\{ u\colon B_1 \to \mathbb{R} : \Delta u = 0 \text{ on } B_1 \text{ and } \int_{B_1} |\nabla u|^2 \mathrm{d}x \leqslant 1 \right\}.$$

Supposing that the lemma were false, there would exist a $\rho > 0$ and a sequence of $f_j \in H^1(B_1)$ such that

$$\lim_{j\to\infty} \sup_{\substack{\phi \in C^1_c(B_1) \\ \|\nabla\phi\|_\infty \leqslant 1}} \left| \int_{B_1} \nabla\phi \cdot \nabla f_j \mathrm{d}x \right| = 0$$

and $\int_{B_1} |\nabla f_j|^2 \mathrm{d}x \leqslant 1$, but

$$\inf_{u\in\mathcal{H}} \int_{B_1} (u - f_j)^2 \mathrm{d}x > \rho.$$

By subtracting the average of f_j over B_1 from f_j, we assume that the f_j have average 0 while retaining the properties above. By the Poincaré Inequality, we calculate that

$$\|f_j\|_{H^1(B_1)} \leqslant C\|f_j\|_{L^2(B_1)} + C \sum_{|\alpha|=1} \|D^\alpha f_j\|_{L^2(B_1)} \leqslant C\|f_j\|_{L^2(B_1)} + C\|\nabla f_j\|_{L^2(B_1)} \leqslant C\|\nabla f_j\|_{L^2(B_1)} \leqslant C.$$

Thus the sequence $\{f_j\}$ is bounded in $H^1(B_1)$, so by Banach-Alaoglu we have up to a subsequence $f_j \rightharpoonup u$ for some $u \in H^1(B_1)$. Moreover, $H^1(B_1) \subset\subset L^2(B_1)$ and we may assume up to a sub-subsequence that $f_j \xrightarrow{L^2} v \in L^2(B_1)$, by the Rellich-Kondrachov Compactness Theorem. By uniqueness of weak limits, $u = v$, so by semicontinuity of the Dirichlet energy (proof: $\nabla f_j \overset{*}{\rightharpoonup} \nabla u \implies |\nabla u| \leqslant \liminf_{j\to\infty} |\nabla f_j|$, and Fatou's Lemma) we obtain

$$\int_{B_1} |\nabla u|^2 \mathrm{d}x \leqslant \liminf_{j\to\infty} \int_{B_1} |\nabla f_j|^2 \mathrm{d}x \leqslant 1.$$

On the other hand, fix a $\phi \in C^1_c(B_1)$. Then weak convergence of $f_j \rightharpoonup u$ in the Hilbert space $H^1(B_1)$ implies

$$\left| \int_{B_1} \nabla\phi \cdot \nabla u \mathrm{d}x \right| = \lim_{j\to\infty} \left| \int_{B_1} \nabla\phi \cdot \nabla f_j \mathrm{d}x \right| = 0.$$

By applying Green's First Identity and Weyl's Lemma 1.1 we conclude at once that u is harmonic, and so $u \in \mathcal{H}$. But $f_j \to u$ strongly in $L^2(B_1)$, contradicting the assumption that $\int_{B_1} (f_j - u)^2 \mathrm{d}x > \rho$. ☐

Lemma 3.6. *Let $k \geqslant 1$. Then there exists a $C > 0$ such that if u is harmonic in $B_r(x) \subset \mathbb{R}^k$, then*

$$\sup_{x \in B_\rho(x_0)} |u(x) - u(x_0) - \nabla u(x_0) \cdot (x - x_0)| \leqslant C\rho^2 r^{-\frac{k}{2}-1} \|Du\|_{L^2(B_r(x_0))}$$

Proof. Without loss of generality we take $r = 1$. Fix a $\rho \leqslant \frac{1}{2}$. By a Taylor expansion we have

$$\sup_{x \in B_\rho} |u(x) - u(0) - \nabla u(0) \cdot x| \leqslant \frac{\rho^2}{2} \|D^2 u\|_{L^\infty(B_{1/2})}.$$

Since u is harmonic, $D^2 u$ is also harmonic, so we can apply the Mean Value Property 1.6, Hölder's Inequality, and Caccioppoli's Inequality 1.20 to the components of $D^2 u$, finding that for any $x \in B_{1/2}$

$$|(D^2 u)_{ij}| = C \left| \int_{B_{1/2}} (D^2 u)_{ij} \mathrm{d}x \right| \leqslant C \int_{B_{1/2}} |(D^2 u)_{ij}| \mathrm{d}x \leqslant C \|(D^2 u)_{ij}\|_{L^2(B_{1/2})} \leqslant C \|(Du)_{ij}\|_{L^2(B_1)}$$

which yields the result. □

3.6 The Excess Decay Theorem

We now enter the meatiest part of the enterprise–proving the Excess Decay Theorem. The theorem says, loosely, that if a varifold satisfies Allard's conditions on a ball, then on a smaller ball the excess of the varifold decreases by at least a factor of two. This theorem allows us to prove a power law decay result for the excess, which ultimately allows us apply our Lipschitz Approximation to a sizable portion of our varifold.

Theorem 3.3 (Excess Decay Theorem)**.** *Let $k < N$ be a positive integer. There exist constants $\eta \in (0, \frac{1}{2})$ and $\varepsilon_0 > 0$ such that if $V = (\Gamma, \theta)$ satisfies Allard's conditions on the ball $B_r(x_0)$ with respect to $\varepsilon = \varepsilon_0$, and $\|H\|_\infty r \leqslant \mathbb{E}(V, \pi, x_0, r)$, then there exists a k-dimensional plane $\bar{\pi}$ such that*

$$\mathbb{E}(V, \bar{\pi}, x_0, \eta r) \leqslant \frac{1}{2} \mathbb{E}(V, \pi, x_0, r).$$

Proof. Even though the proof of this result is somewhat sizable, the idea is reasonably simple and goes roughly as follows. Because the excess of our varifold is small, it is possible to approximate the varifold with a Lipschitz function. We then produce a harmonic approximation to this Lipschitz graph, apply the estimates for harmonic functions proven earlier, and use these estimates to produce bounds for the Tilt-Excess Inequality, thereby yielding the result. The overarching theme, then, is to use successively better-behaved approximations to the varifold, and then transfer estimates for these nicer objects *back* to the original varifold we started with.

Estimates for the Lipschitz Approximation By translating and scaling, we can assume $x_0 = 0$ and $r = 1$. We also abbreviate $\mathbb{E} := \mathbb{E}(V, \pi, 0, 1)$. We need to produce an $\varepsilon_0 > 0$ and an $\eta \in (0, \frac{1}{2})$. To start, we restrict our choice of ε_0 to $\varepsilon_0 < \min\{1, \varepsilon_L\}$ for the constant of the Lipschitz approximation corresponding to some choice of l, β which will be specified over the course of the proof.

Let $f\colon \mathcal{B}_{1/4} \to \pi^\perp$ and $\lambda = \lambda(l) \in (0, 1]$ be the Lipschitz approximation and constant corresponding to l, β in the Lipschitz Approximation Theorem. Choose coordinates $(y, z) = (y_1, \ldots, y_k, z_1, \ldots, z_{N-k})$ on $\mathbb{R}^N$ such that $\pi = \{z = 0\}$. Let $f_1, \ldots, f_{N-k}$ be the corresponding coordinate functions of f. For each $j \in \{1, \ldots, N-k\}$, we set

$$e_j := (\overbrace{0, \ldots, 0}^{k}, 0, \ldots, \overset{j}{1}, \ldots, 0).$$

Fix such a j, and choose $\phi \in C_c^1(\mathcal{B}_{1/16})$. Define the vector field $X \in C_c^1(\mathbb{R}^N; \mathbb{R}^N)$ by $X(y, z) := \phi(y)\psi(z)e_j$ where ψ is a cutoff function in the z variables chosen as follows. By (ii) of the Lipschitz Approximation Theorem 3.2 we have that $B_{1/8} \cap \Gamma \subset B_{1/4} \cap \Gamma \subset I_\beta(\pi)$. At this point we stipulate that $\beta < 1/16$. Then $\psi(z) \in C^1(\mathbb{R}^N)$ is chosen with $\operatorname{spt}\psi \subset I_{1/8}(\pi)$, $\psi \equiv 1$ on $I_{1/16}(\pi)$, and $\|\psi\|_\infty \leqslant 1$. Then $X(y, z) = \phi(y)\psi(z)e_j = \phi(y)e_j$ on $I_{1/16}(\pi)$, while $\operatorname{spt}X \subset B_{1/4}$.

Recall that by $(i), (ii)$ of the Lipschitz Approximation Theorem 3.2, $\theta \equiv 1$ $\mathcal{H}^k$-a.e. on $B_{1/4} \cap \Gamma \subset I_{1/16}(\pi)$ and $\Gamma_f \subset I_{1/16}(\pi)$, where $\psi \equiv 1$. Thus if $\xi_1, \dots, \xi_k$ are an ON basis of $T_x\Gamma$ with $x \in B_{1/4} \cap \Gamma$ or $x \in \Gamma_f$, then

$$\begin{aligned}\mathrm{div}_{T_x\Gamma} X &= \sum_{i=1}^k \xi_i \cdot D_{\xi_i} X \\ &= \sum_{i=1}^k (D_{\xi_i}\phi)\xi_i \cdot e_j \\ &= (\nabla_{T_x\Gamma}\phi) \cdot e_j\end{aligned}$$

Recalling Proposition 3.2, we use this to compute

$$\begin{aligned}\left|\int_{\Gamma_f} (\nabla_{T_x\Gamma_f}\phi) \cdot e_j \mathrm{d}\mathcal{H}^k(x)\right| &\leqslant \left|\int_{\Gamma_f} (\nabla_{T_x\Gamma_f}\phi) \cdot e_j \mathrm{d}\mathcal{H}^k(x) - \delta V(X)\right| + |\delta V(X)| \\ &= \left|\int_{\Gamma_f} (\nabla_{T_x\Gamma_f}\phi) \cdot e_j \mathrm{d}\mathcal{H}^k(x) - \int_\Gamma \mathrm{div}_{T_x\Gamma} X \mathrm{d}\mathcal{H}^k(x)\right| + |\delta V(X)| \\ &= \left|\int_{\Gamma_f} (\nabla_{T_x\Gamma_f}\phi) \cdot e_j \mathrm{d}\mathcal{H}^k(x) - \int_\Gamma (\nabla_{T_x\Gamma}\phi) \cdot e_j \mathrm{d}\mathcal{H}^k(x)\right| + \left|\int_\Gamma X \cdot \mathbf{H} \mathrm{d}\mathcal{H}^k\right|\end{aligned}$$

Recall that if A, B are rectifiable sets, then for $\mathcal{H}^k$-a.e. $x \in A \cap B$, we have $T_xA = T_xB$. Recall also that $\mathrm{spt}\phi \subset \mathcal{B}_{1/16} \subset B_{1/4}$. This allows us to estimate

$$\begin{aligned}\left|\int_{\Gamma_f} (\nabla_{T_x\Gamma_f}\phi) \cdot e_j \mathrm{d}\mathcal{H}^k(x)\right| &\leqslant \left|\int_{\Gamma_f} (\nabla_{T_x\Gamma_f}\phi) \cdot e_j \mathrm{d}\mathcal{H}^k(x) - \int_\Gamma (\nabla_{T_x\Gamma}\phi) \cdot e_j \mathrm{d}\mathcal{H}^k(x)\right| + \left|\int_\Gamma X \cdot \mathbf{H} \mathrm{d}\mathcal{H}^k\right| \\ &= \left|\int_{\Gamma_f \setminus \Gamma} (\nabla_{T_x\Gamma_f}\phi) \cdot e_j \mathrm{d}\mathcal{H}^k(x) - \int_{\Gamma \setminus \Gamma_f} (\nabla_{T_x\Gamma}\phi) \cdot e_j \mathrm{d}\mathcal{H}^k(x)\right| + \left|\int_\Gamma X \cdot \mathbf{H} \mathrm{d}\mathcal{H}^k\right| \\ &\leqslant \left|\int_{(\Gamma_f \setminus \Gamma) \cap B_{1/4}} (\nabla_{T_x\Gamma_f}\phi) \cdot e_j \mathrm{d}\mathcal{H}^k(x)\right| + \left|\int_{(\Gamma \setminus \Gamma_f) \cap B_{1/4}} (\nabla_{T_x\Gamma}\phi) \cdot e_j \mathrm{d}\mathcal{H}^k(x)\right| \\ &\qquad + \left|\int_{\Gamma \cap B_{1/4}} X \cdot \mathbf{H} \mathrm{d}\mathcal{H}^k\right| \\ &\leqslant \|\nabla\phi\|_\infty \left\{\mathcal{H}^k((\Gamma_f \setminus \Gamma) \cap B_{1/4}) + \mathcal{H}^k((\Gamma \setminus \Gamma_f) \cap B_{1/4})\right\} + \|\phi\|_\infty \|\mathbf{H}\|_\infty \mathcal{H}^k(B_{1/4} \cap \Gamma)\end{aligned}$$

Now recall that the set G of the Lipschitz Approximation Theorem has $G \subset B_{1/4} \cap \Gamma$ and $G \subset \Gamma_f$, meaning that

$$(\Gamma_f \setminus \Gamma) \cap B_{1/4} \subset \Gamma_f \setminus (B_{1/4} \cap \Gamma) \subset \Gamma_f \setminus G$$

and

$$(\Gamma \setminus \Gamma_f) \cap B_{1/4} = (B_{1/4} \cap \Gamma) \setminus \Gamma_f \subset (B_{1/4} \cap \Gamma) \setminus G$$

so that by (iv) of the Lipschitz Approximation Theorem 3.2, the assumption $\|\mathbf{H}\|_\infty \leqslant \mathbb{E}$, and $1 \leqslant \lambda^{-1}$,

$$\begin{aligned}\mathcal{H}^k((\Gamma_f \setminus \Gamma) \cap B_{1/4}) + \mathcal{H}^k((\Gamma \setminus \Gamma_f) \cap B_{1/4}) &\leqslant \mathcal{H}^k(\Gamma_f \setminus G) + \mathcal{H}^k(B_{1/4} \cap \Gamma \setminus G) \\ &\leqslant C\lambda^{-1}\mathbb{E} + C\|\mathbf{H}\|_\infty \\ &\leqslant C\lambda^{-1}\mathbb{E}\end{aligned}$$

Moreover, $\|\phi\|_\infty\|\mathbf{H}\|_\infty\mathcal{H}^k(B_{1/4}\cap\Gamma)\leqslant C\mathbb{E}\|\nabla\phi\|_\infty$ because (recalling that $\phi(y)\to 0$ as $y\to\partial\mathcal{B}_{1/16}$ in π)

$$\|\phi\|_\infty\leqslant\sup_{a,b\in\mathcal{B}_{1/16}}|\phi(a)-\phi(b)|=\sup\left|\int_0^1\frac{\mathrm{d}}{\mathrm{d}t}(\phi\circ\gamma(t))\mathrm{d}t\right|=\sup\left|\int_0^1\nabla(\phi\circ\gamma(t))\cdot\gamma'(t)\mathrm{d}t\right|\leqslant\|\nabla\phi\|_\infty\mathrm{diam}\mathcal{B}_{1/16}$$

where the latter suprema are taken over all straight paths γ between $a,b\in\mathcal{B}_{1/16}$. Therefore, we conclude that

$$\left|\int_{\Gamma_f}(\nabla_{T_x\Gamma_f}\phi)\cdot e_j\mathrm{d}\mathcal{H}^k(x)\right|\leqslant C\lambda^{-1}\mathbb{E}\|\nabla\phi\|_\infty \tag{✤}$$

Onward! Let $\xi_1,\ldots,\xi_k$ be an ON basis of π corresponding to the coordinates (y,z) on $\mathbb{R}^N$ from above, in which f has the coordinate representation $f=(f_1,\ldots,f_{N-k})$. The metric tensor of the graph Γ_f is given by the $k\times k$ matrix

$$g_{ij}=\left(\xi_i+\sum_{l=1}^{N-k}(\partial_{y_i}f_l)e_l\right)\cdot\left(\xi_j+\sum_{l=1}^{N-k}(\partial_{y_j}f_l)e_l\right):=v_i\cdot v_j$$

Here we are writing $v_i|_y=\mathrm{d}F_y(\xi_i)$, where $F\colon\mathcal{B}_{1/4}\to\mathbb{R}^N$ is defined by $F(y)=(y,f(y))$. Whenever the base point y is clear, or immaterial, we ommit it from the notation. Because f is Lipschitz, this construction exists almost everywhere on $\mathcal{B}_{1/4}$. Notice that because $\{\xi_i\}$ are ON and $\xi\perp e_j$, we have

$$\begin{aligned}
|g_{ij}-\delta_{ij}|&=|v_i\cdot v_j-\delta_{ij}|\\
&=\Bigg|\xi_i\cdot\xi_j+\xi_i\cdot\left(\sum_l(\partial_{y_j}f_l)e_l\right)+\xi_j\cdot\left(\sum_l(\partial_{y_i}f_l)e_l)\right)\\
&\qquad\qquad+\left(\sum_l(\partial_{y_i}f_l)e_l\right)\cdot\left(\sum_l(\partial_{y_j}f_l)e_l\right)-\delta_{ij}\Bigg|\\
&=\left|\sum_{l,h}(\partial_{y_i}f_l)(\partial_{y_j}f_h)e_l\cdot e_h\right|\\
&=\left|\sum_l(\partial_{y_i}f_l)(\partial_{y_j}f_l)\right|\\
&\leqslant\sum_l|(\partial_{y_i}f_l)|\,|(\partial_{y_j}f_l)|\\
&\leqslant\frac{1}{2}\sum_l|\partial_{y_i}f_l|^2+\frac{1}{2}\sum_l|\partial_{y_j}f_l|^2 &&(2ab\leqslant a^2+b^2)\\
&\leqslant C\|Df\|^2
\end{aligned}$$

Because f is Lipschitz, $\|Df\|\leqslant C\mathrm{Lip}f$ for a dimensional constant C, so if l (the bounding constant for $\mathrm{Lip}f$) is smaller than a dimensional constant, then the same estimate holds for the inverse metric:

$$|g^{ij}-\delta^{ij}|\leqslant C\|Df\|^2.$$

Indeed, recall from operator theory that if X is a Banach space and $V\in B(X,X)$ with $\|V\|<1$, then $(I-V)^{-1}\in B(X,X)$, $(I-V)^{-1}=\sum_{k=0}^\infty V^k$, and thus $\|(I-V)^{-1}\|=(1-\|V\|)^{-1}$. If we choose l small enough so that $\|I-g\|\leqslant C\|Df\|^2\leqslant 3/4$, say, then we have that $g=I-(I-g)$ is boundedly invertible with $\|g^{-1}\|\leqslant(1-\|I-g\|)^{-1}\leqslant 4$. Thus,

$$|g^{ij}-\delta^{ij}|=|g^{ij}||g_{ij}-\delta_{ij}|\leqslant C\|g^{-1}\|\|Df\|^2\leqslant C\|Df\|^2$$

It will also be useful to compute the projection $P_{T_x\Gamma_f}$ for a point $x\in\Gamma_f$ (where the tangent plane exists of course). Observe that if $w\in\mathbb{R}^N$, then because $\{v_i|_x\}$ is a basis for $T_x\Gamma$ and $P_{T_x\Gamma_f}(w)\in T_x\Gamma_f$, we have (suppressing the base point x for readability) $P_{T_x\Gamma_f}(w)\cdot v_j = w\cdot v_j$. On the other hand, since $P_{T_x\Gamma_f}(w)\in T_x\Gamma_f$, there are λ^j such that $P_{T_x\Gamma_f}(w)=\lambda^j v_j$. Therefore, $\lambda^j g_{ij}=\lambda^j v_i\cdot v_j = v_i\cdot P_{T_x\Gamma_f}(w)=v_i\cdot w$, and thus $\lambda^j = g^{ij}v_i\cdot w$, and consequently,

$$P_{T_x\Gamma_f}(w)=g^{ij}(w\cdot v_i)v_j$$

for every $w\in\mathbb{R}^N$.

Two easy facts that we will need are the following:

$$e_j\cdot v_l = e_j\cdot\left(\xi_l+\sum_{i=1}^{N-k}(\partial_{y_l}f_i)e_i\right)=\partial_{y_l}f_j$$

and

$$\begin{aligned}\nabla\phi\cdot v_m &= \left(\sum_{i=1}^{k}(\partial_{y_i}\phi)\xi_i+\sum_{i=1}^{N-k}(\partial_{z_i}\phi)e_i\right)\cdot\left(\xi_m+\sum_{l=1}^{N-k}(\partial_{y_m}f_l)e_l\right)\\ &= \left(\sum_{i=1}^{k}(\partial_{y_i}\phi)\xi_i\right)\cdot\left(\xi_m+\sum_{l=1}^{N-k}(\partial_{y_m}f_l)e_l\right)=\partial_{y_m}\phi\end{aligned}$$

Now fix $x=(w,f(w))$ for $w\in\mathcal{B}_{1/4}$. Then using the above facts

$$P_{T_x\Gamma_f}(\nabla\phi(w))\cdot e_j = (\nabla\phi(w)\cdot v_i)g^{ik}(v_k\cdot e_j)=\partial_{y_i}\phi(w)g^{ik}(w)\partial_{y_k}f_j(w)$$

and applying the estimate $|g^{ik}-\delta^{ik}|\leqslant C\|Df\|^2$, which tells us that $g^{ik}=\delta^{ik}+\mathcal{O}(\|Df\|^2)$,

$$\begin{aligned}P_{T_x\Gamma_f}(\nabla\phi(w))\cdot e_j &= \partial_{y_i}\phi(w)g^{ik}(w)\partial_{y_k}f_j(w)\\ &= \sum_i\partial_{y_i}\phi(w)\partial_{y_i}f_j(w)+\mathcal{O}(\|Df(w)\|^2)\sum_{i,k}\partial_{y_i}\phi(w)\partial_{y_k}f_j(w)\\ &\leqslant \sum_i\partial_{y_i}\phi(w)\partial_{y_i}f_j(w)+\mathcal{O}(\|Df(w)\|^2)\left(\sum_i|\partial_{y_i}\phi(w)|^2\right)^{\frac12}\left(\sum_k|\partial_{y_k}f_j(w)|^2\right)^{\frac12}\\ &\leqslant \sum_i\partial_{y_i}\phi(w)\partial_{y_i}f_j(w)+\mathcal{O}(\|Df(w)\|^2)|\nabla\phi(w)|\|Df(w)\|\\ &= \sum_i\partial_{y_i}\phi(w)\partial_{y_i}f_j(w)+\mathcal{O}(\|Df(w)\|^3)|\nabla\phi(w)|.\end{aligned}\tag{✥}$$

Now set the notation $\bar{\nabla}\phi \coloneqq (\partial_{y_1}\phi,\ldots,\partial_{y_k}\phi)$, and recall the definition of the Jacobian of f as

$$\mathrm{J}f(w)\coloneqq\sqrt{1+\|Df(w)\|^2+\sum_{h,\alpha,\beta}(M^{h,\alpha,\beta}(w))^2}$$

where the sum is taken over all the size $h\times h$ minors (with coordinates (α,β)) of Df with $h\geqslant 2$. We thus have the estimate

$$\mathrm{J}f(w)=1+\frac12\left(\|Df(w)\|^2+\sum_{h,\alpha,\beta}(M^{h,\alpha,\beta}(w))^2\right)+\mathcal{O}\left(\left(\|Df(w)\|^2+\sum_{h,\alpha,\beta}(M^{h,\alpha,\beta}(w))^2\right)^2\right).$$

Recall Hadamard's Inequality (Proposition 1.10), which tells us that

$$|M^{h,\alpha,\beta}(w)|\leqslant\prod_{i=1}^{h}\|Df_i^{h,\alpha,\beta}(w)\|$$

where $Df_i^{h,\alpha,\beta}(w)$ is the i^{th} row (or column) of the (α,β) minor matrix of size h of $Df(w)$. It follows that the sum of the minors is controlled by $\|Df(w)\|$, as

$$|M^{h,\alpha,\beta}(w)| \leqslant \prod_{i=1}^{h} \|Df_i^{h,\alpha,\beta}(w)\| \leqslant C\|Df(w)\|^h$$

and thus recalling $h \geqslant 2$, and the fact that $\|Df\|$ is uniformly bounded,

$$\sum_{h,\alpha,\beta} (M^{h,\alpha,\beta}(w))^2 = \mathcal{O}\left(\|Df(w)\|^2\right).$$

We have thereby established the estimate

$$\mathrm{J}f(w) = 1 + \mathcal{O}(\|Df(w)\|^2)$$

so that

$$|\mathrm{J}f(w) - 1| \leqslant C\|Df(w)\|^2 \tag{✦}$$

for all $w \in \mathcal{B}_{1/4}$.

Next we apply the area formula to deduce

$$\int_{\Gamma_f} (\nabla_{T_x\Gamma_f}\phi)\cdot e_j \mathrm{d}\mathcal{H}^k(x) = \int_{\mathcal{B}_{1/4}} P_{T_x\Gamma_f}(\nabla\phi(w))\cdot e_j \mathrm{J}f(w)\mathrm{d}w, \tag{⊛}$$

which is just a change of variables and follows because $\nabla_{T_x\Gamma_f}\phi \coloneqq P_{T_x\Gamma_f}(\nabla\phi)$.

We now collect the results (✤), (✥), (✦), and (⊛) together to find that

$$\begin{aligned}
\int_{\mathcal{B}_{1/16}} \bar{\nabla}\phi(w)\cdot\bar{\nabla}f_j(w)\mathrm{d}w &= \int_{\mathcal{B}_{1/16}} \sum_{i=1}^{k} (\partial_{y_i}\phi(w))(\partial_{y_i}f_j(w))\mathrm{d}w \\
&\overset{(✥)}{=} \int_{\mathcal{B}_{1/16}} (P_{T_x\Gamma_f}(\bar{\nabla}\phi(w))\cdot e_j - \mathcal{O}(\|Df(w)\|^3)\|\bar{\nabla}\phi(w)\|)\mathrm{d}w \\
&= \int_{\mathcal{B}_{1/16}} P_{T_x\Gamma_f}(\bar{\nabla}\phi(w))\cdot e_j\mathrm{J}f(w)\mathrm{d}w + \int_{\mathcal{B}_{1/16}} (1-\mathrm{J}f(w))P_{T_x\Gamma_f}(\bar{\nabla}\phi(w))\cdot e_j\mathrm{d}w \\
&\qquad - \int_{\mathcal{B}_{1/16}} \mathcal{O}(\|Df(w)\|^3)\|\bar{\nabla}\phi(w)\|\mathrm{d}w \\
&\overset{(⊛)}{=} \int_{\Gamma_f} (\nabla_{T_x\Gamma_f}\phi)\cdot e_j\mathrm{d}\mathcal{H}^k(x) + \int_{\mathcal{B}_{1/16}} (1-\mathrm{J}f(w))P_{T_x\Gamma_f}(\bar{\nabla}\phi(w))\cdot e_j\mathrm{d}w \\
&\qquad - \int_{\mathcal{B}_{1/16}} \mathcal{O}(\|Df(w)\|^3)\|\bar{\nabla}\phi(w)\|\mathrm{d}w
\end{aligned}$$

which implies that

$$\begin{aligned}
\left|\int_{\mathcal{B}_{1/16}} \bar{\nabla}\phi(w)\cdot\bar{\nabla}f_j(w)\mathrm{d}w\right| &\overset{(✤)}{\leqslant} C\lambda^{-1}\mathbb{E}\|\bar{\nabla}\phi\|_\infty + \int_{\mathcal{B}_{1/16}} |1-\mathrm{J}f(w)||P_{T_x\Gamma_f}(\bar{\nabla}\phi(w))|\mathrm{d}w \\
&\qquad + C\|\bar{\nabla}\phi\|_\infty \int_{\mathcal{B}_{1/16}} \|Df(w)\|^2\mathrm{d}w \\
&\overset{(✤)}{\leqslant} C\lambda^{-1}\mathbb{E}\|\bar{\nabla}\phi\|_\infty + C\int_{\mathcal{B}_{1/16}} \|Df(w)\|^2|\bar{\nabla}\phi(w)|\mathrm{d}w \\
&\qquad + C\|\bar{\nabla}\phi\|_\infty \int_{\mathcal{B}_{1/16}} \|Df(w)\|^2\mathrm{d}w \\
&\leqslant C\lambda^{-1}\mathbb{E}\|\bar{\nabla}\phi\|_\infty + C\|\bar{\nabla}\phi\|_\infty \int_{\mathcal{B}_{1/16}} \|Df(w)\|^2\mathrm{d}w
\end{aligned}$$

We turn now to estimating the second term above, and seek the relation $\|Df(w)\|^2 \leqslant 2\|T_x\Gamma_f - \pi\|^2$ for $w \in \mathcal{B}_{1/4}$. This is just another computation as follows (suppressing dependence on w):

$$
\begin{aligned}
\|T_x\Gamma_f - \pi\|^2 \geqslant |P_\pi(e_j) - P_{T_x\Gamma_f}(e_j)|^2 &= |P_{T_x\Gamma_f}(e_j)|^2 = \left|\partial_{y_l} f_j g^{lm} v_m\right|^2 = \left(\partial_{y_l} f_j g^{lm} v_m\right) \cdot \left(\partial_{y_{l'}} f_j g^{l'm'} v_{m'}\right) \\
&= (\partial_{y_l} f_j)(\partial_{y_{l'}} f_j) g^{lm} g^{l'm'} (v_m \cdot v_{m'}) \\
&= (\partial_{y_l} f_j)(\partial_{y_{l'}} f_j) g^{lm} g^{l'm'} g_{mm'} \\
&= (\partial_{y_l} f_j)(\partial_{y_m} f_j) g^{lm} \\
&= \sum_{l,m=1}^{k} \left((\partial_{y_l} f_j)^2 + (\partial_{y_l} f_j)(\partial_{y_m} f_j)(g^{lm} - \delta^{lm})\right) \\
&= |\bar{\nabla} f_j|^2 + \sum_{l,m=1}^{k} (\partial_{y_l} f_j)(\partial_{y_m} f_j)(g^{lm} - \delta^{lm}) \\
&\geqslant |\bar{\nabla} f_j|^2 - C\|Df\|^2 \sum_{l,m=1}^{k} (\partial_{y_l} f_j)(\partial_{y_m} f_j) \\
&\geqslant |\bar{\nabla} f_j|^2 - C\|Df\|^3
\end{aligned}
$$

Seeing as though $\|Df\|^2 = \sum_{j=1}^{N-k} |\bar{\nabla} f_j|^2$, we sum the above over all such j and apply the bound $\|Df\| \leqslant Cl$ to determine

$$\|T_x\Gamma_f - \pi\|^2 \geqslant \|Df\|^2 - C\|Df\|^3 \geqslant \|Df\|^2(1 - Cl).$$

By taking l small enough, depending only on the dimensional constant C, we have our desired estimate

$$2\|T_x\Gamma_f - \pi\| \geqslant \|Df(w)\|^2.$$

Continuing, we observe that

$$
\begin{aligned}
\mathcal{B}_{1/16} = (P_\pi(G) \cap B_{1/16}) \cup (\mathcal{B}_{1/16} \setminus P_\pi(G)) \subset P_\pi(G) \cup (\mathcal{B}_{1/4} \setminus P_\pi(G)) &= P_\pi(G) \cup (P_\pi(\Gamma_f) \setminus P_\pi(G)) \\
&= P_\pi(G) \cup P_\pi(\Gamma_f \setminus G).
\end{aligned}
$$

thus allowing us to write

$$
\begin{aligned}
\left|\int_{\mathcal{B}_{1/16}} \bar{\nabla}\phi(w) \cdot \bar{\nabla} f_j(w) \mathrm{d}w\right| &\leqslant C\lambda^{-1}\mathbb{E}\|\bar{\nabla}\phi\|_\infty + C\|\bar{\nabla}\phi\|_\infty \int_{\mathcal{B}_{1/16}} \|Df(w)\|^2 \mathrm{d}w \\
&\leqslant C\lambda^{-1}\mathbb{E}\|\bar{\nabla}\phi\|_\infty + C\|\bar{\nabla}\phi\|_\infty \int_{P_\pi(G)} \|Df(w)\|^2 \mathrm{d}w \\
&\qquad + C\|\bar{\nabla}\phi\|_\infty \int_{P_\pi(\Gamma_f \setminus G)} \|Df(w)\|^2 \mathrm{d}w \\
&\leqslant C\lambda^{-1}\mathbb{E}\|\bar{\nabla}\phi\|_\infty + C\|\bar{\nabla}\phi\|_\infty \int_{P_\pi(G)} \|Df(w)\|^2 \mathrm{d}w
\end{aligned}
$$

where we applied estimate (iv) of the Lipschitz Approximation Theorem 3.2 (and recalled that projections are Lipschitz) to deduce that

$$\int_{P_\pi(\Gamma_f \setminus G)} \|Df(w)\|^2 \mathrm{d}w \leqslant C\mathcal{H}^k(P_\pi(\Gamma_f \setminus G)) \leqslant C\mathcal{H}^k(\Gamma_f \setminus G) \leqslant C\lambda^{-1}\mathbb{E} + C\|\mathbf{H}\|_\infty = C\lambda^{-1}\mathbb{E}.$$

We thus need to estimate the last integral. To do so, first observe that $G \subset \Gamma \cap \Gamma_f$ since $G \subset \Gamma_f$ and $G \subset B_{1/4} \cap \Gamma \subset \Gamma$. Since Γ and Γ_f are rectifiable, for $\mathcal{H}^k$-a.e. $x \in G$ we have $T_x\Gamma_f = T_x\Gamma$, so

$$2\|T_x\Gamma - \pi\|^2 = 2\|T_x\Gamma_f - \pi\|^2 \geqslant \|Df(w)\|^2$$

for $\mathcal{H}^k$-a.e. $w \in P_\pi(G)$ where as before $x = (w, f(w))$. Therefore,

$$\left|\int_{\mathcal{B}_{1/16}} \bar{\nabla}\phi(w) \cdot \bar{\nabla} f_j(w) \mathrm{d}w\right| \leqslant C\lambda^{-1}\mathbb{E}\|\bar{\nabla}\phi\|_\infty + C\|\bar{\nabla}\phi\|_\infty \int_{P_\pi(G)} \|T_x\Gamma - \pi\|^2 \mathrm{d}w$$

But we observe

$$\begin{aligned}
\int_{P_\pi(G)} \|T_x\Gamma - \pi\|^2 \mathrm{d}w &= \int_{P_\pi(G)} \|T_x\Gamma - \pi\|^2 \mathrm{J}f(w)\mathrm{d}w + \int_{P_\pi(G)} \|T_x\Gamma - \pi\|^2 (1 - \mathrm{J}f(w))\mathrm{d}w \\
&= \int_G \|T_x\Gamma - \pi\|^2 \mathrm{d}\mathcal{H}^k(x) + \int_{P_\pi(G)} \|T_x\Gamma - \pi\|^2 (1 - \mathrm{J}f(w))\mathrm{d}w \\
&\leqslant \mathbb{E} + \int_{P_\pi(G)} \|T_x\Gamma - \pi\|^2 |1 - \mathrm{J}f(w)|\mathrm{d}w \\
&\overset{(\blacklozenge)}{\leqslant} \mathbb{E} + C\int_{P_\pi(G)} \|T_x\Gamma - \pi\|^2 \|Df(w)\|^2 \mathrm{d}w \\
&\leqslant \mathbb{E} + Cl\int_{P_\pi(G)} \|T_x\Gamma - \pi\|^2 \mathrm{d}w
\end{aligned}$$

so that, provided we take l small enough to ensure $Cl < 1$,

$$\int_{P_\pi(G)} \|T_x\Gamma - \pi\|^2 \mathrm{d}w \leqslant \frac{1}{1 - Cl}\mathbb{E}.$$

We now freeze l small in accordance with all of the above arguments. From the Lipschitz Approximation Theorem 3.2 we obtain a fixed $\lambda = \lambda(l) \leqslant 1$, and we conclude the main estimate of this section:

$$\left|\int_{\mathcal{B}_{1/16}} \bar{\nabla}\phi(w) \cdot \bar{\nabla} f_j(w) \mathrm{d}w\right| \leqslant C\mathbb{E}\|\bar{\nabla}\phi\|_\infty \qquad \forall \phi \in C_c^1(\mathcal{B}_{1/16}) \tag{✺}$$

Observe that at every step in determining the required smallness of l, we only had to take dimensional data into account. Thus, l, and also λ, are dimensional constants. Thus is important because it now allows us to bundle these constants with C.

Lastly, we notice that by a similar argument

$$\begin{aligned}
\int_{\mathcal{B}_{1/16}} |\bar{\nabla} f_j(w)|^2 \mathrm{d}w &\leqslant C\int_{\mathcal{B}_{1/16}} \|Df(w)\|^2 \mathrm{d}w \\
&\leqslant C\int_{P_\pi(G)} \|Df(w)\|^2 \mathrm{d}w + C\int_{P_\pi(\Gamma_f \setminus G)} \|Df(w)\|^2 \mathrm{d}w \\
&\leqslant C\mathbb{E}
\end{aligned} \tag{✳}$$

The Harmonic Approximation We now produce a harmonic approximation to the Lipschitz approximation of our varifold. Fix a $\vartheta > 0$, arbitrary for now, and let $\varepsilon_A > 0$ be the constant of Lemma 3.5 corresponding to $\rho = \vartheta$. For each $j \in \{1, \ldots, N - k\}$ set $\tilde{f}_j \coloneqq c_0\mathbb{E}^{-1/2} f_j$, where c_0 is chosen such that

$$\int_{\mathcal{B}_{1/16}} |\bar{\nabla}\tilde{f}_j(w)|^2 \mathrm{d}w \leqslant 16^{-k}.$$

This is possible by (✳), if we take $c_0 = 16^{-k}C^{-1/2}$ in that inequality. By (✺) we see also that

$$\left|\int_{\mathcal{B}_{1/16}} \bar{\nabla}\phi(w) \cdot \bar{\nabla}\tilde{f}_j(w) \mathrm{d}w\right| \leqslant C\mathbb{E}^{1/2}\|\bar{\nabla}\phi\|_\infty \qquad \forall \phi \in C_c^1(\mathcal{B}_{1/16}).$$

Recall that we are in the business of finding an $\varepsilon_0 > 0$. We further restrict our choice to $\varepsilon_0 \leqslant (\frac{\varepsilon_A}{16^k C})^2$ for this dimensional constant C, and freeze it for good (so that this is the ε_0 of the claim). Then if V satisfies Allard's conditions with such an ε_0, we see that

$$\left| \int_{\mathcal{B}_{1/16}} \bar{\nabla}\phi(w) \cdot \bar{\nabla}\tilde{f}_j(w) \mathrm{d}w \right| \leqslant C\mathbb{E}^{1/2} \|\bar{\nabla}\phi\|_\infty \leqslant \varepsilon_A 16^{-k} \|\bar{\nabla}\phi\|_\infty \qquad \forall \phi \in C_c^1(\mathcal{B}_{1/16})$$

which is just what we need to apply Lemma 3.5 and obtain harmonic functions $\tilde{u}_j \colon \mathcal{B}_{1/16} \to \mathbb{R}$ with

$$\int_{\mathcal{B}_{1/16}} |\bar{\nabla}\tilde{u}_j|^2 \mathrm{d}w \leqslant 16^{-k} \leqslant 1 \quad \text{and} \quad \int_{\mathcal{B}_{1/16}} (\tilde{f}_j - \tilde{u}_j)^2 \mathrm{d}w \leqslant 16^{-2-k}\vartheta \leqslant \vartheta.$$

For each j set $u_j := c_0^{-1}\mathbb{E}^{1/2}\tilde{u}_j$ to conclude that

$$\int_{\mathcal{B}_{1/16}} (f_j - u_j)^2 \mathrm{d}w \leqslant C\vartheta\mathbb{E} \tag{❃1}$$

Lastly, set $u := (u_1, \dots, u_{N-k})$. Then

$$\|\bar{D}u\|_2^2 \leqslant C \sum_j \int_{\mathcal{B}_{1/16}} |\bar{\nabla}u_j|^2 \mathrm{d}w \leqslant C\mathbb{E}. \tag{❃2}$$

The Height Estimate The main estimate to be established in this section is the following:

$$\eta^{-2-k} \int_{B_{4\eta(x_0)}} \operatorname{dist}(x - x_0, \bar{\pi})^2 \mathrm{d}\mu_V(x) \leqslant C\eta^{-2-k}\vartheta\mathbb{E} + C\beta^2\eta^{-2-k}\mathbb{E} + C\eta^2\mathbb{E} \tag{✲}$$

To this end let $L \colon \pi \to \pi^\perp$ be the map defined by

$$L(y) := \sum_{j=1}^{N-k} (\bar{\nabla}u_j(0) \cdot y)e_j = \sum_{j=1}^{N-k} \bar{D}_y u_j(0) e_j.$$

Set $x_0 := (0, u(0))$ and let $\bar{\pi}$ be the k-plane defined by

$$\bar{\pi} := \{y + L(y) \ : \ y \in \pi\}.$$

By the mean value property of harmonic functions applied in each coordinate of u, we first observe that

$$\begin{aligned} \operatorname{dist}(x_0, \pi) = |u(0)| = \left| \fint_{\mathcal{B}_{1/16}} u \mathrm{d}w \right| &\leqslant C\|u\|_{L^1(\mathcal{B}_{1/16})} \overset{\text{(Hölder)}}{\leqslant} C\|u\|_{L^2(\mathcal{B}_{1/16})} \\ &\leqslant C\|u - f\|_{L^2(\mathcal{B}_{1/16})} + C\|f\|_{L^2(\mathcal{B}_{1/16})} \\ &\leqslant C\vartheta^{1/2}\mathbb{E}^{1/2} + C\beta \end{aligned} \tag{❃}$$

using the estimates (❃1) and (❃2) of the Harmonic Approximation above, and the bound $\|f\|_\infty \leqslant \beta$ from part (i) of the Lipschitz Approximation Theorem 3.2. Moreover, we have the same bound for the quantity $\|P_\pi^\perp - P_{\bar{\pi}}^\perp\|$. We apply the Tilting Subspace Lemma 1.2 from the preliminaries to secure the estimate

$$\|P_\pi^\perp - P_{\bar{\pi}}^\perp\| = \|P_\pi - P_{\bar{\pi}}\| \leqslant C \sum_{j=1}^{N-k} |\bar{\nabla}u_j(0)|$$

Next, we apply Theorem 1.19 from the preliminaries to conclude:

$$\|P_\pi^\perp - P_{\bar{\pi}}^\perp\| \leqslant C \sum_{j=1}^{N-k} |\bar{\nabla}u_j(0)| \leqslant C\|u\|_{L^1(\mathcal{B}_{1/16})} \leqslant C\vartheta^{1/2}\mathbb{E}^{1/2} + C\beta$$

Now let $x \in B_{1/64} \cap \Gamma$. We will soon specify ϑ, β in such a way that $|x_0| = |u(0)| \leqslant C\vartheta^{1/2}\mathbb{E}^{1/2} + C\beta \leqslant 1/80$ (in addition to several additional conditions). Thus, $|x - x_0| \leqslant |x| + |x_0| < 1$ and we have

$$\begin{aligned}\operatorname{dist}(x - x_0, \bar{\pi}) = |P_{\bar{\pi}}^{\perp}(x - x_0)| &\leqslant |(P_{\bar{\pi}}^{\perp} - P_{\pi}^{\perp})(x - x_0)| + |P_{\pi}^{\perp}(x - x_0)| \\ &\leqslant \|P_{\bar{\pi}}^{\perp} - P_{\pi}^{\perp}\||x - x_0| + |P_{\pi}^{\perp}x| + |P_{\pi}^{\perp}x_0| \\ &\leqslant \|P_{\bar{\pi}}^{\perp} - P_{\pi}^{\perp}\| + \operatorname{dist}(x, \pi) + \operatorname{dist}(x_0, \pi) \\ &\leqslant C\vartheta^{1/2}\mathbb{E}^{1/2} + C\beta\end{aligned}$$

where we have applied the two previous estimates and noted that $\operatorname{dist}(x, \pi) < \beta$ since $B_{1/64} \cap \Gamma \subset I_\beta(\pi)$.

We have thereby paved the way for the following estimates, where $\eta \in (0, \frac{1}{2})$ is to be determined. In particular, we will be choosing $\vartheta, \beta, \varepsilon_0, \eta$ so that $B_{4\eta}(x_0) \subset B_{1/16}$, thus ensuring that $\theta \equiv 1$ $\mathcal{H}^k$-a.e. on $B_{4\eta}(x_0) \cap \Gamma$. Therefore, recalling the estimate $\mathcal{H}^k((B_{1/4} \cap \Gamma) \setminus G) \leqslant C\lambda^{-1}\mathbb{E} = C\mathbb{E}$ of the Lipschitz Approximation Theorem 3.2, we can calculate

$$\begin{aligned}\int_{B_{4\eta}(x_0)\setminus\Gamma_f} \operatorname{dist}(x - x_0, \bar{\pi})^2 \mathrm{d}\mu_V(x) &= \int_{(\Gamma\setminus\Gamma_f)\cap B_{4\eta}(x_0)} \operatorname{dist}(x - x_0, \bar{\pi})^2 \mathrm{d}\mathcal{H}^k(x) \\ &\leqslant (C\vartheta^{1/2}\mathbb{E}^{1/2} + C\beta)^2 \mathcal{H}^k((\Gamma \setminus \Gamma_f) \cap B_{4\eta}(x_0)) \\ &\leqslant C(\vartheta^{1/2}\mathbb{E}^{1/2} + \beta)^2 \mathcal{H}^k((B_{1/4} \cap \Gamma) \setminus \Gamma_f) \\ &\leqslant C(\vartheta^{1/2}\mathbb{E}^{1/2} + \beta)^2 \mathcal{H}^k((B_{1/4} \cap \Gamma) \setminus G) \\ &\leqslant C(\vartheta^{1/2}\mathbb{E}^{1/2} + \beta)^2 \mathbb{E}.\end{aligned}$$

Let $x = (y, f(y)) \in \Gamma_f$, so that $\operatorname{dist}(x - x_0, \bar{\pi}) \leqslant |(y, f(y)) - (0, u(0)) - (y, L(y))| = |f(y) - u(0) - L(y)|$ as $(y, L(y)) \in \bar{\pi}$. From estimate (❃1) of the Harmonic Approximation section,

$$\begin{aligned}\int_{\Gamma_f \cap B_{4\eta}(x_0)} \operatorname{dist}(x - x_0, \bar{\pi})^2 \mathrm{d}\mu_V(x) &\leqslant \int_{\mathcal{B}_{4\eta}(x_0)} \operatorname{dist}((y, f(y)) - x_0, \bar{\pi})^2 \mathrm{d}y \\ &\leqslant \int_{\mathcal{B}_{4\eta}(x_0)} |f(y) - u(0) - L(y)|^2 \mathrm{d}y \\ &\leqslant 2\int_{\mathcal{B}_{4\eta}(x_0)} |f(y) - u(y)|^2 \mathrm{d}y + 2\int_{\mathcal{B}_{4\eta}(x_0)} |u(y) - u(0) - L(y)|^2 \mathrm{d}y \\ &\leqslant C\vartheta\mathbb{E} + 2\int_{\mathcal{B}_{4\eta}(x_0)} |u(y) - u(0) - L(y)|^2 \mathrm{d}y.\end{aligned}$$

By Lemma 3.6 and (❃2) we have that

$$\sup_{y \in \mathcal{B}_{4\eta}(x_0)} |u(y) - u(x_0) - L(y)|^2 \leqslant C\eta^4 \|\bar{D}u\|_{L^2}^2 \leqslant C\eta^4\mathbb{E}$$

because $L(y) = \bar{D}u(0) \cdot y$. Therefore, we conclude that

$$\int_{\Gamma_f \cap B_{4\eta}(x_0)} \operatorname{dist}(x - x_0, \bar{\pi})^2 \mathrm{d}\mu_V(x) \leqslant C\vartheta\mathbb{E} + C\eta^{k+4}\mathbb{E}.$$

At last, we are ready to exhibit estimate (✲):

$$\begin{aligned}\int_{B_{4\eta}(x_0)} \operatorname{dist}(x - x_0, \bar{\pi})^2 \mathrm{d}\mu_V(x) &= \int_{B_{4\eta}(x_0)\setminus\Gamma_f} + \int_{B_{4\eta}(x_0)\cap\Gamma_f} \\ &\leqslant C(\vartheta^{1/2}\mathbb{E}^{1/2} + \beta)^2\mathbb{E} + C\vartheta\mathbb{E} + C\eta^{k+4}\mathbb{E} \\ &= C\vartheta\mathbb{E}^2 + C\beta\vartheta^{1/2}\mathbb{E}^{3/2} + C\beta^2\mathbb{E} + C\vartheta\mathbb{E} + C\eta^{k+4}\mathbb{E} \\ &\leqslant C\vartheta\mathbb{E} + C\beta^2\mathbb{E} + C\eta^{k+4}\mathbb{E}\end{aligned}$$

using that $\vartheta, \beta, \mathbb{E} < 1$ and $\beta\vartheta^{1/2} \leqslant \frac{1}{2}\beta^2 + \frac{1}{2}\vartheta$.

Applying the Tilt-Excess Inequality We now impose the conditions

$$C\vartheta^{1/2} \leqslant \tfrac{\eta}{2} \qquad C\beta \leqslant \tfrac{\eta}{2}$$

on ϑ and β, where η is still to be determined, and where C is the constant in (❄) of the Height Estimate. Then it follows that $B_\eta \subset B_{2\eta}(x_0)$ since $|x_0| = |u(0)| \leqslant C\vartheta^{1/2}\mathbb{E}^{1/2} + C\beta \leqslant \frac{\eta}{2}(\mathbb{E}^{1/2}+1)\eta < \eta$ and so if $|x| < \eta$, then $|x - x_0| \leqslant |x| + |x_0| < 2\eta$. This allows us to compute

$$\begin{aligned}\mathbb{E}(V, \bar{\pi}, 0, \eta) &= \frac{1}{\eta^k}\int_{B_\eta(0)} \|T_y\Gamma - \bar{\pi}\|^2 \mathrm{d}\mu_V(x) \\ &\leqslant \frac{2^k}{(2\eta)^k}\int_{B_{2\eta}(x_0)} \|T_y\Gamma - \bar{\pi}\|^2 \mathrm{d}\mu_V(x)\end{aligned}$$

so that by the Tilt-Excess Inequality

$$\begin{aligned}\mathbb{E}(V, \bar{\pi}, 0, \eta) \leqslant 2^k\mathbb{E}(V, \bar{\pi}, x_0, 2\eta) &\leqslant \frac{C}{\eta^{2+k}}\int_{B_{4\eta}(x_0)} \mathrm{dist}(y - x_0, \bar{\pi})^2 \mathrm{d}\mu_V(y) + \frac{C}{\eta^{k-2}}\int_{B_{4\eta}(x_0)} \|\mathbf{H}\|^2 \mathrm{d}\mu_V \\ &\leqslant C\eta^{-2-k}\vartheta\mathbb{E} + C\eta^{-2-k}\beta^2\mathbb{E} + C\eta^2\mathbb{E} + C\eta^2\mathbb{E}^2 \\ &\leqslant C\eta^{-2-k}\vartheta\mathbb{E} + C\eta^{-2-k}\beta\mathbb{E} + C\eta^2\mathbb{E}.\end{aligned}$$

Because we were so thoughtful in making sure this particular C is purely dimensional and independent of $\eta, \beta, \vartheta, \varepsilon_0$, we can now place our restrictions on the later constants in terms of this C. In any case we can safely assume without loss of generality that $C \geqslant 1600$ by increasing it if necessary. We then fix $\eta \in (0, \frac{1}{2})$ such that

$$C\eta^2 = \frac{1}{4}.$$

Then $\eta \leqslant 1/80$, implying that $|x_0| < \eta \leqslant 1/80$ and $4\eta \leqslant 1/20$ ensuring that as we requested, $B_{4\eta}(x_0) \subset B_{1/16}$. Then, choose β and ϑ (both less than 1 of course) so small that in addition to the previous conditions we also have

$$C\eta^{-k-2}\vartheta \leqslant \tfrac{1}{8} \qquad C\eta^{-k-2}\beta \leqslant \frac{1}{8}.$$

Freezing every constant at this point, we conclude the result:

$$\begin{aligned}\mathbb{E}(V, \bar{\pi}, 0, \eta) &\leqslant C\eta^{-2-k}\vartheta\mathbb{E} + C\eta^{-2-k}\beta\mathbb{E} + C\eta^2\mathbb{E} \\ &\leqslant \tfrac{1}{8}\mathbb{E} + \tfrac{1}{8}\mathbb{E} + \tfrac{1}{4}\mathbb{E} \\ &= \tfrac{1}{2}\mathbb{E}.\end{aligned}$$

□

3.7 The Main Theorem

We at last come to the proof of the main theorem, which we restate for easy reference:

Theorem 3.4 (Allard)**.** *Let $k < N$ be a positive integer. Then there are positive constants $\varepsilon, \alpha, \gamma$ such that the following holds. Let $V = (\Gamma, \theta)$ be a k-dimensional rectifiable varifold with bounded generalized mean curvature $\mathbf{H}$ supported in the ball $B_r(x_0)$, $x_0 \in \mathrm{spt}\mu_V$, such that*

(A1) $\mu_V(B_r(x_0)) < (\omega_k + \varepsilon)r^k$ and $\|\mathbf{H}\|_\infty < \varepsilon r^{-1}$.

(A2) There is a k-dimensional plane π such that $\mathbb{E}(V, \pi, x_0, r) < \varepsilon$.

Then $B_{\gamma r}(x_0) \cap \Gamma$ is a $C^{1,\alpha}$ submanifold of $B_{\gamma r}(x_0)$ without boundary. Moreover, $\theta \equiv 1$ on $B_{\gamma r}(x_0) \cap \Gamma$.

Proof. By translating and scaling, we can assume without loss of generality that $x_0 = 0$ and $r = 1$. We proceed by breaking the proof into four parts. The first establishes a decay result for the excess, and the second utilizes it in covering part of the varifold with a Lipschitz graph. The next step shows that this covering coincides with the varifold in a neighborhood of the origin. Finally, we use the result of part one to show the desired regularity of the Lipschitz image.

Step 1: The Power-Law Decay for the Excess We need to produce positive constants $\varepsilon, \alpha, \gamma$. The Excess Decay Theorem 3.3 gives us an $\eta \in (0, \frac{1}{2})$ and an $\varepsilon_0 > 0$ such that if V satisfies $(A1), (A2)$ with $\varepsilon = \varepsilon_0$ and $\|\mathbf{H}\|_\infty < \mathbb{E}(V, \pi, 0, 1)$, then there exists a k dimensional plane $\bar{\pi}$ such that

$$\mathbb{E}(V, \bar{\pi}, 0, \eta) \leqslant \frac{1}{2}\mathbb{E}(V, \pi, 0, 1).$$

Without loss of generality, we can impose the condition that $\varepsilon_0 < \frac{1}{2}$ as well. This allows us to utilize Lemma 3.4 with $\delta = \varepsilon_0$, yielding an $\varepsilon_H > 0$. We need to find $\varepsilon > 0$, so let's start by restricting our choice to $\varepsilon < \min\{\varepsilon_0, \varepsilon_H\}$.

If $V = (\Gamma, \theta)$ satisfies Allard's conditions with this ε, then we have the following:

(i) V satisfies the hypotheses of the Excess Decay Theorem 3.3, so that there exists a k-dimensional plane $\bar{\pi}$ with $\mathbb{E}(V, \bar{\pi}, 0, \eta) \leqslant \frac{1}{2}\mathbb{E}(V, \pi, 0, 1)$, where $\eta < \frac{1}{2}$ is purely dimensional.

(ii) V satisfies the hypotheses of Lemma 3.4, so that $B_{1/2} \cap \Gamma \subset I_{\varepsilon_0}(\pi)$ and $\mu_V(B_r(x)) < (\omega_k + \varepsilon_0)r^k$ for all $x \in B_{1/4}$ and all $r < \frac{1}{2}$.

Fix then any $x \in B_{1/4} \cap \Gamma$, and define the map

$$F(r) := \mathbb{E}(r) + \Lambda\|\mathbf{H}\|_\infty r$$

where $\Lambda := 4\eta^{-k}$ and $\mathbb{E}(r) := \min_\tau \mathbb{E}(V, \tau, x, r)$, the minimization taking place over all k-dimensional subspaces τ. $\mathbb{E}(r)$ is indeed well defined because for fixed x, r the map $\tau \mapsto \mathbb{E}(V, \tau, x, r)$ is a continuous map from from the compact Grassmanian $G_k(\mathbb{R}^N)$ (in the standard topology induced from its smooth manifold structure) to $\mathbb{R}$. Notice also that $\Lambda^{-1} < 2^{-2-k}$ since $\eta < \frac{1}{2}$.

Now, suppose that $F(r) < \varepsilon_0$. Then $\mathbb{E}(r) = \mathbb{E}(V, \tau_0, x, r) < \varepsilon_0$ for some τ_0, and there are two possibilities we could be dealing with:

Case 1: $r\|\mathbf{H}\|_\infty \leqslant \mathbb{E}(r)$: From point (ii) above, we have that $\mu_V(B_r(x)) < (\omega_k + \varepsilon_0)r^k$ for all $r < \frac{1}{2}$, and moreover $r\|\mathbf{H}\|_\infty \leqslant \mathbb{E}(r) < \varepsilon_0$. Thus, we can apply the Excess Decay Theorem 3.3 to V in the ball $B_r(x) \subset B$ to find, with the same universal η, a k plane $\tilde{\pi}$ such that $\mathbb{E}(V, \tilde{\pi}, x, \eta r) \leqslant \frac{1}{2}\mathbb{E}(V, \tau_0, x, r)$. We can then estimate

$$\begin{aligned}
F(\eta r) &= \mathbb{E}(\eta r) + \Lambda\eta r\|\mathbf{H}\|_\infty \\
&\leqslant \mathbb{E}(V, \tilde{\pi}, x, \eta r) + \Lambda\eta r\|\mathbf{H}\|_\infty \\
&\leqslant \frac{1}{2}\left\{\mathbb{E}(V, \tau_0, x, r) + \Lambda r\|\mathbf{H}\|_\infty\right\} \\
&= \frac{1}{2}\left\{\mathbb{E}(r) + \Lambda r\|\mathbf{H}\|_\infty\right\} \\
&= \frac{1}{2}F(r)
\end{aligned}$$

Case 2: $r\|\mathbf{H}\|_\infty \geqslant \mathbb{E}(r)$**:** Similarly we compute, using $\Lambda^{-1} < 2^{-2-k}$, that

$$\begin{aligned}
F(\eta r) &\leqslant \frac{1}{2}\left\{\mathbb{E}(r) + \Lambda r\|\mathbf{H}\|_\infty\right\} \\
&\leqslant \frac{1}{2}\left\{r\|\mathbf{H}\|_\infty + \Lambda r\|\mathbf{H}\|_\infty\right\} \\
&= \Lambda r\|\mathbf{H}\|_\infty \left\{\frac{1}{2}\Lambda^{-1} + \frac{1}{2}\right\} \\
&= \Lambda r\|\mathbf{H}\|_\infty \left\{\frac{1}{2^{3+k}} + \frac{1}{2}\right\} \\
&\leqslant \frac{3}{4}\Lambda r\|\mathbf{H}\|_\infty \\
&\leqslant \frac{3}{4}F(r)
\end{aligned}$$

Therefore, in any case if $F(r) < \varepsilon_0$, then $F(\eta r) \leqslant \frac{3}{4}F(r)$. Thus, $F(\eta r) < \varepsilon_0$, and we can iterate our result to conclude that $F(\eta^k r) \leqslant \left(\frac{3}{4}\right)^k F(r)$. In particular, notice that with $r = 2^{-1}$, we can compute

$$F(2^{-1}) = \mathbb{E}(2^{-1}) + 2^{-1}\Lambda\|\mathbf{H}\|_\infty \leqslant 2^k\mathbb{E}(V,\pi,0,1) + 2^{-1}\Lambda\|\mathbf{H}\|_\infty \leqslant (2^k + 2^{-1}\Lambda)\varepsilon$$

where we have noticed that, since $B_{1/2}(x) \subset B$,

$$\begin{aligned}
\mathbb{E}\left(2^{-1}\right) \leqslant \mathbb{E}(V,\pi,x,2^{-1}) &= 2^k \int_{B_{1/2}(x)} \|T_y\Gamma - \pi\|^2 \mathrm{d}\mu_V(y) \\
&\leqslant 2^k \int_B \|T_y\Gamma - \pi\|^2 \mathrm{d}\mu_V(y) \\
&= 2^k\mathbb{E}(V,\pi,0,1)
\end{aligned}$$

Thus, if we further add the restriction that $\varepsilon < \frac{\varepsilon_0}{2^k + 2^{-1}\Lambda}$, then $F(2^{-1}) \leqslant (2^k + 2^{-1}\Lambda)\varepsilon < \varepsilon_0$. So, we can start the iteration at $r = 2^{-1}$ and find that for every $n \geqslant 0$

$$F(\eta^n 2^{-1}) \leqslant \left(\tfrac{3}{4}\right)^n F(2^{-1}) \leqslant C\left(\tfrac{3}{4}\right)^n \varepsilon.$$

Note that this holds for any fixed $x \in B_{1/4} \cap \Gamma$.

Now let $r \leqslant 2^{-1}$, and set $n = \lfloor \log_\eta(2r) \rfloor (\geqslant 0)$. Then we can show that

$$\mathbb{E}(r) \leqslant 2^k\eta^{-k}\mathbb{E}(\eta^n 2^{-1}) \leqslant 2^k\eta^{-k}F(\eta^n 2^{-1}) \leqslant C\left(\tfrac{3}{4}\right)^n \varepsilon \leqslant C\left(\tfrac{3}{4}\right)^{\log_\eta(2r)-1}\varepsilon \leqslant Cr^{2\alpha}\varepsilon$$

where C and α are positive constants depending only upon N, k. The first inequality is due to the following observation. Since $r \leqslant 2^{-1}$, and since $\eta < 2^{-1}$, there is an $n \geqslant 0$ such that $\eta^{n+1}2^{-1} < r \leqslant \eta^n 2^{-1}$. This is where the peculiar choice of n comes from, since $n + 1 > \log_\eta 2r \geqslant n$. Then

$$\begin{aligned}
\mathbb{E}(r) \leqslant \mathbb{E}(V,\hat{\pi},x,r) &= r^{-k}\int_{B_r(x)} \|T_y\Gamma - \hat{\pi}\|^2 \mathrm{d}\mu_V(y) \\
&\leqslant 2^k\eta^{-(n+1)k}\int_{B_{\eta^n 2^{-1}}(x)} \|T_y\Gamma - \hat{\pi}\|^2 \mathrm{d}\mu_V(y) \\
&= 2^k\eta^{-k}\mathbb{E}(\eta^n 2^{-1})
\end{aligned}$$

where $\hat{\pi}$ is such that $\mathbb{E}(\eta^n 2^{-1}) = \mathbb{E}(V,\hat{\pi},x,\eta^n 2^{-1})$. The last inequality follows by choosing α small, say $\alpha \leqslant \frac{1}{2}\log_\eta(\frac{3}{4})$, which depends only on η and thus only on N and k. To wit, with such an α we see that

$$(2r)^{2\alpha} = \left(\tfrac{3}{4}\right)^{2\alpha\log_{3/4} 2r} = \left(\tfrac{3}{4}\right)^{\frac{2\alpha}{\log(3/4)}\log(2r)} \geqslant \left(\tfrac{3}{4}\right)^{\frac{\log_\eta(3/4)}{\log(3/4)}\log(2r)} = \left(\tfrac{3}{4}\right)^{\frac{1}{\log\eta}\log(2r)} = \tfrac{3}{4}\left(\tfrac{3}{4}\right)^{\log_\eta(2r)-1}$$

In conclusion, we have shown that for any $r \leqslant 2^{-1}$,

$$\mathbb{E}(r) \leqslant Cr^{2\alpha}\varepsilon.$$

Step 2: Covering by a Lipshitz Graph Next, we show that $B_{1/4} \cap \Gamma$ is contained in the graph of a Lipschitz function $f\colon \mathcal{B}_{1/4} \to \pi^{\perp}$, where we write $\mathcal{B}_r(x) := B_r(x) \cap \pi$.

We begin by fixing any point $x \in B_{1/4}$ and setting $\pi = \pi_0$. For each $n \geqslant 1$, we select a k plane π_n such that $\mathbb{E}(2^{-n}) = \mathbb{E}(V, \pi_n, x, 2^{-n})$. By Corollary 3.1 part (ii), we have that for all $x \in \operatorname{spt}\mu_V = B \cap \Gamma$, and for all $r < \operatorname{dist}(x, \partial B) = 1 - |x|$,

$$\frac{\mu_V(B_r(x))}{r^k} \geqslant \omega_k e^{-\|\mathbf{H}\|_\infty r} \geqslant \omega_k e^{-1},$$

since $\|\mathbf{H}\|_\infty < \varepsilon < 1$. In other words for any $r < 1 - |x|$ we have for some dimensional C the estimate $\mu_V(B_r(x)) \geqslant Cr^k$.

Now, by the triangle inequality we find that $\|\pi_n - \pi_{n+1}\| \leqslant \|\pi_n - T_y\Gamma\| + \|T_y\Gamma - \pi_{n+1}\|$, and moreover $x \in B_{1/4}$ implies that $1 - |x| > 3/4 > 1/2$, so we can compute that for any $n \geqslant 0$

$$\begin{aligned}
\|\pi_n - \pi_{n+1}\| &\leqslant \frac{1}{\mu_V(B_{2^{-(n+1)}}(x))} \int_{B_{2^{-(n+1)}}(x)} (\|\pi_n - T_y\Gamma\| + \|T_y\Gamma - \pi_{n+1}\|)\, \mathrm{d}\mu_V(y) \\
&\leqslant \frac{C}{2^{-(n+1)k}} \int_{B_{2^{-(n+1)}}(x)} \|\pi_n - T_y\Gamma\| \mathrm{d}\mu_V(y) + \frac{C}{2^{-(n+1)k}} \int_{B_{2^{-(n+1)}}(x)} \|\pi_{n+1} - T_y\Gamma\| \mathrm{d}\mu_V(y) \\
&\leqslant \frac{C}{2^{-(n+1)k}} \left(\int_{B_{2^{-(n+1)}}(x)} \|\pi_n - T_y\Gamma\|^2 \mathrm{d}\mu_V(y) \right)^{\frac{1}{2}} (\mu_V(B_{2^{-(n+1)}}(x)))^{\frac{1}{2}} \\
&\qquad + \frac{C}{2^{-(n+1)k}} \left(\int_{B_{2^{-(n+1)}}(x)} \|\pi_{n+1} - T_y\Gamma\|^2 \mathrm{d}\mu_V(y) \right)^{\frac{1}{2}} (\mu_V(B_{2^{-(n+1)}}(x)))^{\frac{1}{2}}. \quad \text{(Hölder)}
\end{aligned}$$

Recall that V satisfies the hypotheses of Lemma 3.4 with $\delta = \varepsilon_0$, so that, as $x \in B_{1/4}$ and $2^{-(n+1)} \leqslant 2^{-1}$, we have

$$\mu_V(B_{2^{-(n+1)}}(x)) < (\omega_k + \varepsilon_0) 2^{-(n+1)k} \leqslant C 2^{-(n+1)k}$$

where C depends only on N, k since ε_0 and ω_k do. Thus, we can continue estimating with

$$\begin{aligned}
\|\pi_n - \pi_{n+1}\| &\leqslant \frac{C}{2^{-(n+1)k}} \left(\int_{B_{2^{-(n+1)}}(x)} \|\pi_n - T_y\Gamma\|^2 \mathrm{d}\mu_V(y) \right)^{\frac{1}{2}} 2^{-\frac{(n+1)k}{2}} \\
&\qquad + \frac{C}{2^{-(n+1)k}} \left(\int_{B_{2^{-(n+1)}}(x)} \|\pi_{n+1} - T_y\Gamma\|^2 \mathrm{d}\mu_V(y) \right)^{\frac{1}{2}} 2^{-\frac{(n+1)k}{2}} \\
&= C \left(2^{(n+1)k} \int_{B_{2^{-(n+1)}}(x)} \|\pi_n - T_y\Gamma\|^2 \mathrm{d}\mu_V(y) \right)^{\frac{1}{2}} \\
&\qquad + C \left(2^{(n+1)k} \int_{B_{2^{-(n+1)}}(x)} \|\pi_{n+1} - T_y\Gamma\|^2 \mathrm{d}\mu_V(y) \right)^{\frac{1}{2}} \\
&= C\mathbb{E}(V, \pi_n, x, 2^{-(n+1)})^{\frac{1}{2}} + C\mathbb{E}(2^{-(n+1)})^{\frac{1}{2}}
\end{aligned}$$

But notice that

$$\begin{aligned}\mathbb{E}(V,\pi_n,x,2^{-(n+1)}) &= 2^{n+1}\int_{B_{2^{-(n+1)}}(x)} \|\pi_n - T_y\Gamma\|^2 \mathrm{d}\mu_V(y)\\ &\leqslant 2\cdot 2^n \int_{B_{2^{-n}}(x)} \|\pi_n - T_y\Gamma\|^2 \mathrm{d}\mu_V(y) = 2\mathbb{E}(2^{-n})\end{aligned}$$

so applying the power law decay proved in Step 1, for any $n \geqslant 1$,

$$\begin{aligned}\|\pi_n - \pi_{n+1}\| &\leqslant C\mathbb{E}(2^{-n})^{\frac{1}{2}} + C\mathbb{E}(2^{-(n+1)})^{\frac{1}{2}}\\ &\leqslant C\left\{((2^{-n})^{2\alpha}\varepsilon)^{\frac{1}{2}} + ((2^{-(n+1)})^{2\alpha}\varepsilon)^{\frac{1}{2}}\right\}\\ &= C\left\{2^{-n\alpha}\varepsilon^{\frac{1}{2}} + 2^{-(n+1)\alpha}\varepsilon^{\frac{1}{2}}\right\}\\ &\leqslant C2^{-n\alpha}\varepsilon^{\frac{1}{2}}\end{aligned}$$

Additionally, by $(A2)$ it also follows that $\|\pi - \pi_1\| \leqslant C\left\{\mathbb{E}^{\frac{1}{2}} + \mathbb{E}(2^{-1})^{\frac{1}{2}}\right\} < C\varepsilon^{\frac{1}{2}}$.

Thus, we conclude that for any $j \geqslant 1$,

$$\|\pi - \pi_j\| \leqslant \sum_{k=0}^{j-1}\|\pi_{k+1} - \pi_k\| \leqslant C\varepsilon^{\frac{1}{2}}\sum_{k=0}^{j-1}(2^{-\alpha})^k \leqslant C\varepsilon^{\frac{1}{2}}.$$

Double checking all the calculations reassures us that C still depends only upon N and k.

Therefore, for any $x \in B_{1/4}$ and $r \leqslant 2^{-1}$, we have that

$$\begin{aligned}\mathbb{E}(V,\pi,x,r) &= r^{-k}\int_{B_r(x)} \|\pi - T_y\Gamma\|^2 \mathrm{d}\mu_V(y)\\ &\leqslant 2r^{-k}\int_{B_r(x)} \|\pi_n - T_y\Gamma\|^2 \mathrm{d}\mu_V(y) + 2r^{-k}\int_{B_r(x)} \|\pi_n - \pi\|^2 \mathrm{d}\mu_V(y)\end{aligned}$$

Choose $n \geqslant 1$ so large that $2^{-(n+1)} < r \leqslant 2^{-n}$. Then

$$\begin{aligned}r^{-k}\int_{B_r(x)} \|\pi_n - T_y\Gamma\|^2 \mathrm{d}\mu_V(y) &\leqslant 2^{(n+1)k}\int_{B_{2^{-n}}(x)} \|\pi_n - T_y\Gamma\|^2 \mathrm{d}\mu_V(y)\\ &= 2^k\mathbb{E}(2^{-n})\\ &\leqslant C(2^{-n})^{2\alpha}\varepsilon = C\varepsilon\end{aligned}$$

Meanwhile, we have by Lemma 3.4 (ii) again that

$$r^{-k}\int_{B_r(x)} \|\pi_n - \pi\|^2 \mathrm{d}\mu_V(y) \leqslant C\varepsilon r^{-k}\mu_V(B_r(x)) \leqslant C\varepsilon$$

In conclusion, we find that $\mathbb{E}(V,\pi,x,r) \leqslant C\varepsilon$ (❇) for every $x \in B_{1/4}\cap\Gamma$ and every $r \leqslant 2^{-1}$.

—

We now show that $B_{1/4}\cap\Gamma$ is covered by a Lipschitz graph. To this end, we will utilize the Lipschitz Approximation, for which we need to specify constants $l, \beta \in (0,1)$. β can be fixed at will, but we will need to be more careful with l. Start by assuming that we have chosen $l < 2^{-1}$, and let λ and ε_L be the corresponding constants of the Lipschitz Approximation Theorem 3.2. We further restrict our choice of ε to $\varepsilon < \min\{\varepsilon_L, C^{-1}\varepsilon_L, C^{-1}\lambda\}$, where C is the constant in (❇). Then for any $x \in B_{1/4}\cap\Gamma$ and any $r \leqslant 2^{-1}$, we find that $\mathbb{E}(V,\pi,x,r) < \lambda$. Thus, there exists a Lipschitz map $f\colon B_{1/4}\cap\pi \to \pi^{\perp}$ as in the Lipschitz Approximation Theorem 3.2 with the property that $B_{1/4}\cap\Gamma$ is covered by Γ_f.

Step 3: The Covering Coincides with the Varifold We just showed that $B_{1/4} \cap \Gamma$ is contained in a Lipschitz graph, but we don't yet know that any neighborhood of $x_0 = 0$ in π actually lies completely under Γ. In other words, Γ might be punctured with holes around the origin. Here we show that this is not the case, and that the set $D := P_\pi(B_{1/4} \cap \Gamma)$ contains $\mathcal{B}_{1/16}$. This implies that over $\mathcal{B}_{1/16}$, Γ is not just covered by the graph of f but actually coincides with it. Before proceeding with the proof, recall the notation $\mathcal{B}_r(x) := B_r(x) \cap \pi$, which is used whenever it is clear what π is.

Let then $\xi \in \partial\mathcal{B}_1(0)$, and let θ be such that $\omega_k - 2\theta := \mathcal{H}^k(\mathcal{B}_1(0) \setminus \mathcal{B}_1(\xi))$. Suppose that $D := P_\pi(B_{1/4} \cap \Gamma)$ does NOT contain $\mathcal{B}_{1/16}$, and let $w \in \mathcal{B}_{1/16} \setminus D$ (so that w represents a hole in Γ over $\mathcal{B}_{1/16}$). Define $r := \inf_{z \in D} |w - z|$, which must be less than $1/16$ since $0 \in D$ and $w \in \mathcal{B}_{1/16}$. If $r = 0$, we extend Γ to include (the measure zero set) w by leveraging the continuity of f while preserving all norms. Thus we can assume that $r > 0$. Any sequence of points $\{z_n\} \subset D$ realizing this infimum must therefore be contained, without loss of generality, in $\mathcal{B}_{1/8}$. By compactness, we can also assume without loss of generality that $z_n \to z \in \overline{\mathcal{B}}_{1/8}$.

Now, since $0 \in \Gamma$, $f(0) = 0$, and $\mathrm{Lip} f \leqslant l$, we conclude that $\|f\|_\infty \leqslant l$ as well since for any $x \in \mathcal{B}_{1/4}$ we have $|f(x)| = |f(x) - f(0)| \leqslant l|x - 0| < l$. By taking $l < 1$ sufficiently small, we can thus ensure that $x_n := (z_n, f(z_n)) \in B_{3/16}$ and also (by continuity of f) $x_n \to x := (z, f(z)) \in B_{3/16}$. Because $r < 1/16$, we thus find that $B_r(x) \cap \Gamma \subset B_{1/4}(x) \cap \Gamma \subset \Gamma_f$. Notice also that by definition of r, $\mathcal{B}_r(w) \cap D = \emptyset$.

Thus by the area formula,

$$\mu_V(B_r(x)) = \mathcal{H}^k(B_r(x) \cap \Gamma) \leqslant \int_{\mathcal{B}_r(z) \setminus \mathcal{B}_r(w)} Jf$$

since no points of Γ lie over the set $B_r(w)$. But recall that

$$Jf(x) \leqslant \sqrt{1 + l^2 + h.o.t.} = 1 + \tfrac{1}{2} l^2 + h.o.t \leqslant 1 + Cl^2$$

since $l < 1$. Additionally, by scaling and translation invariance we find

$$\begin{aligned} \mathcal{H}^k(\mathcal{B}_r(z) \setminus \mathcal{B}_r(w)) &= \mathcal{H}^k(r(\mathcal{B}_1(0) \setminus \mathcal{B}_1(\tfrac{w-z}{r}))) \\ &= r^k \mathcal{H}^k(\mathcal{B}_1(0) \setminus \mathcal{B}_1(\tfrac{w-z}{r})) \\ &\leqslant r^k \mathcal{H}^k(\mathcal{B}_1(0) \setminus \mathcal{B}_1(\xi)) \\ &= r^k(\omega_k - 2\theta), \end{aligned}$$

so we estimate $\mu_V(B_r(x)) \leqslant (\omega_k - 2\theta)(1 + Cl^2) r^k$.

Now, if necessary, make C bigger and take l smaller so that

$$(\omega_k - 2\theta)(1 + Cl^2) = \omega_k - \theta. \qquad (✾)$$

On the other hand, by Corollary 3.1

$$\mu_V(B_r(x)) \geqslant \omega_k r^k e^{-\|\mathbf{H}\|_\infty r} \geqslant \omega_k r^k e^{-\varepsilon}$$

since $r < 1/16$ and $x \in B_{3/16}$. We are therefore faced with

$$\omega_k r^k e^{-\varepsilon} \leqslant (\omega_k - \theta) r^k$$

which yields a contradiction as soon as ε is smaller than a dimensional constant. Thus, $D = P_\pi(B_{1/4} \cap \Gamma)$ must contain $\mathcal{B}_{1/16} = B_{1/16} \cap \pi$.

Step 4: The Regularity of the Lipschitz Covering Lastly, we prove the $C^{1,\alpha}$ regularity of Γ. From Step 3, we know that $\Gamma_f = B_{1/4} \cap (\mathcal{B}_{1/16} \times \pi^\perp) \cap \Gamma$. For each $z \in \mathcal{B}_{1/32}$, and each $r < 1/32$, let $\pi_{z,r}$ be a k-dimensional plane such that

$$\mathbb{E}(V, \pi_{z,r}, (z, f(z)), r) = \min_\tau \mathbb{E}(V, \tau, (z, f(z)), r) \leqslant Cr^{2\alpha}\varepsilon < C\varepsilon$$

where we noticed that $(z, f(z)) \in B_{1/4} \cap \Gamma$ and applied the power law decay of the excess from Step 1. We also know that $\mathbb{E}(V, \pi, (z, f(z)), r) \leqslant C\varepsilon$, so by calculations identical to those of Step 2, we conclude that $\|\pi - \pi_{z,r}\| \leqslant C\varepsilon^{\frac{1}{2}}$ (note that the right hand side is independent of z). By now choosing ε smaller than a dimensional constant, we can thus realize each $\pi_{z,r}$ as the image of a linear map $A_{z,r} \colon \pi \to \pi^{\perp}$ with $\|A_{z,r}\| \leqslant 1$.

Let now $A, B \colon \pi \to \pi^{\perp}$ be any two linear maps with $\|B\| \leqslant Cl$. Let τ_A and τ_B denote their respective graphs over π, and P_A, P_B the corresponding projections onto τ_A and τ_B. If we choose l in (✾) to be smaller than yet another a geometric constant (adjusting the constant C in that formula as necessary), then we can ensure that $|P_B(v)| \leqslant \frac{1}{2}|v|$ for every $v \in \pi^{\perp}$. Indeed, as $l \downarrow 0$, $\|P_B - P_\pi\| \to 0$, and $P_\pi(v) = 0$ for all $v \in \pi^{\perp}$. Therefore, $|P_B(v)| = |(P_B - P_\pi)(v)| \leqslant \|P_B - P_\pi\||v| < \frac{1}{2}|v|$ provided l is sufficiently small.

Now fix an orthonormal basis $e_1, \ldots, e_k$ of π and observe that

$$\begin{aligned}
|A(e_i) - B(e_i)| &= |(e_i + A(e_i)) - (e_i + B(e_i))| \\
&= |P_A(e_i + A(e_i)) - P_B(e_i + B(e_i))| \\
&\leqslant |P_A(e_i) - P_B(e_i)| + |P_A(Ae_i) - P_B(Ae_i)| + |P_B(Ae_i) - P_B(Be_i)| \\
&= |P_A(e_i) - P_B(e_i)| + |P_A(Ae_i) - P_B(Ae_i)| + |P_B(\underbrace{Ae_i - Be_i}_{\in \pi^{\perp}})| \\
&\leqslant C\|P_A - P_B\| + \tfrac{1}{2}|A(e_i) - B(e_i)| \\
&= C\|\tau_A - \tau_B\| + \tfrac{1}{2}|A(e_i) - B(e_i)|
\end{aligned}$$

Thus we conclude that $|A(e_i) - B(e_i)| \leqslant C\|\tau_A - \tau_B\|$, which implies that $\|A - B\| \leqslant C\|\tau_A - \tau_B\|$ since if $v = \lambda^i e_i \in \pi$, then

$$|(A - B)(v)| = |(A - B)(\lambda^i e_i)| \leqslant \lambda^i |(A - B)(e_i)| \leqslant C\|\tau_A - \tau_B\||v|.$$

We finally freeze l for good.

Recall that since f is Lipschitz, $Df(y) \colon B_{1/4} \cap \pi \to \pi^{\perp}$ has (a.e.) as its image $T_{(y,f(y))}\Gamma$ whenever $y \in \mathcal{B}_{1/16}$, since we just proved that Γ coincides with Γ_f over $\mathcal{B}_{1/16}$.

For our $z \in \mathcal{B}_{1/64}$ and $r < 1/64$ we have the inclusion $f(\mathcal{B}_r(z)) \subset B_{2r}(z, f(z))$. Indeed, if $x = (w, f(w))$ for some $w \in \mathcal{B}_{1/64}$, then $|(w, f(w)) - (z, f(z))|^2 = |w - z|^2 + |f(w) - f(z)|^2 \leqslant r^2(1 + l^2) < 4r^2$. Since $Jf \geqslant 1$, we estimate by the Area Formula (Theorem 1.14) and the above calculations that

$$\begin{aligned}
\int_{\mathcal{B}_r(z)} \|Df(y) - A_{z,2r}\|^2 \mathrm{d}y &\leqslant \int_{\mathcal{B}_r(z)} \|Df(y) - A_{z,2r}\|^2 Jf(y) \mathrm{d}y \\
&\leqslant C \int_{\mathcal{B}_r(z)} \|T_{(y,f(y))}\Gamma - \pi_{z,2r}\|^2 Jf(y) \mathrm{d}y \\
&\leqslant C \int_{B_{2r}(z,f(z))} \|T_y\Gamma - \pi_{z,2r}\|^2 \mathrm{d}\mu_V(y) \\
&= Cr^k \mathbb{E}(V, \pi_{z,r}, (z, f(z)), 2r) \\
&\leqslant Cr^{k+2\alpha}
\end{aligned}$$

Again, this holds for any $z \in \mathcal{B}_{1/64}$ and $r < 1/64$. Let $\overline{Df}_{z,r}$ denote the average of Df over $\mathcal{B}_r(z)$. Then we compute for any such z, r

$$\begin{aligned}\int_{\mathcal{B}_r(z)} \|Df(y) - \overline{Df}_{z,r}\|^2 \mathrm{d}y &= \sum_{i,j} \int_{\mathcal{B}_r(z)} |(Df)^i_j(y) - \overline{(Df)^i_j}_{z,r}|^2 \mathrm{d}y \\ &= \sum_{i,j} \min_{a \in \mathbb{R}} \int_{\mathcal{B}_r(z)} |(Df)^i_j(y) - a|^2 \mathrm{d}y \\ &= \min_{A \in M^{k\times(N-k)}} \int_{\mathcal{B}_r(z)} \|Df(y) - A\|^2 \mathrm{d}y \\ &\leqslant \int_{\mathcal{B}_r(z)} \|Df(y) - A_{z,2r}\|^2 \mathrm{d}y \\ &\leqslant Cr^{k+2\alpha}\end{aligned}$$

From this we conclude that there exists a $C^{0,\alpha}$ function g such that $g = Df$ on $\mathcal{B}_{1/64}$. To do so, we apply Campanato's Criterion 1.17. Indeed, notice that for all $z \in \mathcal{B}_{1/64}$ and all $r > 0$ the following uniform decay condition holds for each component function of Df:

$$\left(\frac{1}{r^k} \int_{\mathcal{B}_{1/64} \cap \mathcal{B}_r(z)} |Df^i_j(y) - \overline{(Df^i_j)}_{z,r}|^2 \mathrm{d}y \right)^{\frac{1}{2}} \leqslant Cr^\alpha$$

Thus, for every i, j there is a $g^j_i \in C^{0,\alpha}(\mathcal{B}_{1/64})$ such that $g^j_i = Df^j_i$ a.e. on $\mathcal{B}_{1/64}$. Let $g \in C^{0,\alpha}(\mathcal{B}_{1/64}; M^{k\times(N-k)})$ be defined by $g = (g^j_i)$. Then $g = Df$ a.e. on $\mathcal{B}_{1/64}$, which proves that $f \in C^{1,\alpha}(\mathcal{B}_{1/64})$.

□

4 References

References

[BV16] T. Bourni and A. Volkmann. An allard type regularity theorem for varifolds with a hölder condition on the first variation. *Calculus of Variations and Partial Differential Equations*, 55(3):46, 2016.

[EG15] L. C. Evans and R. Gariepy. *Measure Theory and Fine Properties of Functions*. CRC Press, 2015.

[Eva10] L. C. Evans. *Partial Differential Equations*. AMS, 2010.

[GM05] M. Giaquinta and L. Martinazzi. *An Introduction to the Regularity Theory for Elliptic Systems, Harmonic Maps and Minimal Graphs*. Scuola Normale Superiore Pisa, 2005.

[GT01] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer, 2001.

[LA00] D. Pallara L. Ambrosio, N. Fusco. *Functions of Bounded Variation and Free Discontinuity Problems*. Clarendon Press, Oxford, 2000.

[Lel16] C. De Lellis. The size of the singular set of area minimizing currents. *Surv. Differ. Geom.*, 2016.

[Lel18] C. De Lellis. Allard's interior regularity theorem: An invitation to stationary varifolds. *Proceedings of CMSA Harvard*, 2018.

[Mag12] F. Maggi. *Sets of Finite Perimeter and Geometric Variational Problems*. Cambridge, 2012.

[Mai17] F. P. Maiale. Geometric measure theory: Lecture notes. Course taught at Pisa University by G. Alberti, 2017.

[Mat95] P. Mattila. *Geometry of Sets and Measures in Euclidean Spaces*. Cambridge, 1995.

[Men17] U. Menne. The concept of varifold. *Notices of Amer. Math. Soc.*, 2017.

[Sim14] Leon Simon. *Introduction to Geometric Measure Theory*. 2014.