ALLARD'S INTERIOR REGULARITY THEOREM: AN INVITATION TO STATIONARY VARIFOLDS

CAMILLO DE LELLIS

0. Introduction

This is a small set of notes, taken from the last lectures of a course given in Spring 2012 at the University of Zürich. The aim is to give a short, reader-friendly but nonetheless detailed introduction to Allard's interior regularity theory for stationary integral varifolds. Allard's results, which are 40 years old (see [2]), form a pillar of the theory of minimal surfaces, which has been used a number of times in the literature, sometimes to reach really spectacular geometric applications. On the other hand I know only one textbook which reports them, Simon's *Lecture notes on geometric measure theory* (see [13]). Because of lack of time, I was not able to cover the material exposed in [13] in my course and I therefore looked for a suitable reduction which would anyway allow me to prove the key results. These lecture notes assume that the reader is familiar with some more advanced measure theory (Hausdorff measures, covering arguments and density theorems, see Chapters 2, 4 and 6 of [9]), has a coarse knowledge of rectifiable sets (definition, area formula and approximate tangents, see Chapter 4 of [5]) and knows a little about harmonic functions (see e.g. [7]).

There is, however, a price to pay. At a first glance an obvious shortcoming is that we restrict the theory to varifolds with bounded generalized mean curvature, whereas a suitable integrability assumption is usually sufficient. This is not a major point, since we cover stationary varifolds in smooth Riemannian manifolds (cf. Exercise 1.6). A second drawback is that hypothesis (H2) in Allard's ε-regularity Theorem 3.2 is redundant. Still, the statement given here suffices to draw the two major conclusions of Allard's theory. A third disadvantage is that a few estimates coming into the proof of Theorem 3.2 are stated in a fairly suboptimal form (the reader might compare, for instance, the crude estimates of Proposition 5.1 to those of Theorem 20.2 of [13]). In spite of these drawbacks, I still hope that these notes will give to the the reader not only a quick access to the most relevant ideas, but also to several important technical points, thus simplifying his task if he eventually explores the deeper results of the literature (see Section 8).

1. Integral and stationary varifolds

Definition 1.1. Let $U \subset \mathbb{R}^N$ be an open set. An integral (or integer-rectifiable) varifold V of dimension k in U is a pair (Γ, f), where $\Gamma \subset U$ is a rectifiable set of dimension k and

$f : \Gamma \to \mathbb{N} \setminus \{0\}$ a Borel map. We can naturally associate to V the following measure:

$$\mu_V(A) = \int_{\Gamma \cap A} f \, d\mathcal{H}^k \qquad \text{for any Borel set } A.$$

The *mass* $\mathbf{M}(V)$ of V is given by $\mu_V(U)$ and in what follows we will assume that it is *finite*.

If M is a closed Riemannian manifold embedded in $\mathbb{R}^N$, an integral varifold V in $U \cap M$ is an integral varifold V in U such that $\mu_V(U \setminus M) = 0$.

Remark 1.2. This definition is fairly different from Allard's original one, which introduces first general varifolds as a certain class of measures in the Grassmanian $G(U)$ (see Definition 3.1 of [2]) and then identifies integral varifolds as an appropriate subset (cf. Section 3.5 of [2]). Here we follow instead the approach of Chapter 4 of [13] ([13] reports also the theory of general varifolds in the subsequent Chapter 8).

We next introduce the concepts of *stationarity* and *generalized mean curvature* (cf. Section 16 of [13]; again the original paper of Allard defines these concepts for general varifolds, cf. Section 4 of [2] and Section 39 of [13]). If $\Phi : U \to W$ is a diffeomorphism and $V = (\Gamma, f)$ an integral varifold in U, we then define the pushforward $\Phi_\sharp V$ as $(\Phi(\Gamma), f \circ \Phi^{-1})$. Obviously $\Phi_\sharp V$ is an integral varifold in W. Given a vector field $X \in C^1_c(U, \mathbb{R}^N)$, the one-parameter family of diffeomorphisms generated by X is $\Phi_t(x) = \Phi(t, x)$ where $\Phi : \mathbb{R} \times U \to U$ is the unique solution of

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

Definition 1.3. If V is a varifold in U and $X \in C^1_c(U, \mathbb{R}^N)$, then the first variation of V along X is defined by

$$\delta V(X) = \left. \frac{d}{dt} \right|_{t=0} \mathbf{M}((\Phi_t)_\sharp V) \, , \tag{1.1}$$

where Φ_t is the one-parameter family generated by X. V has *bounded generalized mean curvature* if there exists a constant $C \geq 0$ such that

$$|\delta V(X)| \leq C \int |X| d\mu_V \qquad \text{for all } X \in C^1_c(U, \mathbb{R}^N). \tag{1.2}$$

If the constant C in (1.2) can be set to 0 the varifold is called *stationary.* If M is a closed Riemannian manifold embedded in $\mathbb{R}^N$ and V an integral varifold in $U \cap M$, V is said to be stationary in $U \cap M$ if $\delta V(X) = 0$ for any vector field $X \in C^1_c(U)$ tangent to M.

By Proposition 1.5 below, the map $X \to \delta V(X)$ is linear. Therefore, the Riesz representation theorem and the Radon-Nikodym Theorem yield the following corollary.

Corollary 1.4. *If V is a varifold in U with bounded generalized mean curvature, then there is a bounded Borel map $H : U \to \mathbb{R}^N$ such that*

$$\delta V(X) = -\int X \cdot H d\mu_V \qquad \textit{for all } X \in C^1_c(U, \mathbb{R}^N). \tag{1.3}$$

H will be called the generalized mean curvature *of the varifold V and is defined up to sets of μ_V-measure zero.*

For these definitions to make sense we need to prove that the derivative in the right hand side of (1.1) exists. This will be shown in the proposition below, where we also give an integral formula for the first variation. In order to do so we introduce the following notation. If $\pi \subset \mathbb{R}^N$ is a k-dimensional plane and $X \in C^1(U, \mathbb{R}^N)$ a vector field, we set

$$\mathrm{div}_\pi X = \sum_{i=1}^k e_i \cdot D_{e_i} X \,,$$

where $e_1, \ldots, e_k$ is any orthonormal base of π. Recall also that, if Γ is a rectifiable k-dimensional subset of $\mathbb{R}^N$, then Γ has an approximate tangent plane $T_x\Gamma$ at $\mathcal{H}^k$–a.e. $x \in \Gamma$.

Proposition 1.5. *Let $V = (\Gamma, f)$ be an integral varifold in $U \subset \mathbb{R}^N$. Then the right hand side of* (1.1) *is well defined and*

$$\delta V(X) = \int \mathrm{div}_{T_x\Gamma} X \, d\mu_V \qquad \textit{for all } X \in C^1_c(U, \mathbb{R}^N). \tag{1.4}$$

Exercise 1.6. Consider a smooth closed Riemannian manifold M isometrically embedded in $\mathbb{R}^N$, a bounded open $U \subset \mathbb{R}^N$ and an integral varifold V *stationary in* $U \cap M$. Show that V has bounded generalized mean curvature H and $|H| \leq C|A_M|$ μ_V–a.e., where A_M denotes the second fundamental form of M and C is a dimensional constant.

Proof of Proposition 1.5. By the definition of rectifiability, there are countably many C^1 embeddings $F_i : \mathbb{R}^k \to \mathbb{R}^N$ and compact sets $K_i \subset \mathbb{R}^k$ such that

(a) $F_i(K_i) \cap F_j(K_j) = \emptyset$ for every $i \neq j$
(b) $F_i(K_i) \subset \Gamma$ for every i;
(c) $\{F_i(K_i)\}$ covers $\mathcal{H}^k$–a.a. Γ

(cf. Lemma 11.1 of [13]). By standard arguments in measure theory we can assume that f is a constant f_i on each $F_i(K_i)$. Since Φ_t is a diffeomorphism of U onto itself, the properties (a)-(b)-(c) hold even if we replace F_i with $\Phi_t \circ F_i$ and Γ with $\Phi_t(\Gamma)$. Therefore, by the area formula (see Proposition 4.3 of [5]) we conclude

$$\mathbf{M}((\Phi_t)_\sharp V) = \sum_i f_i \int_{K_i} |d(\Phi_t \circ F_i)|_y e_1 \wedge \ldots \wedge d(\Phi_t \circ F_i)|_y e_k| \, dy \,, \tag{1.5}$$

where $e_1, \ldots e_k$ is an orthonormal base for $\mathbb{R}^k$ (here $|v_1 \wedge \ldots \wedge v_k|$ denotes the square root of the determinant of the $k \times k$ matrix with coefficients $v_r \cdot v_s$, $r, s \in \{1, \ldots, k\}$). Fix a point $y \in K_i$ and set $x = F_i(y)$ and recall that $dF_i|_y e_1, \ldots, dF_i|_y e_k$ is a base for $T_x\Gamma$ for $\mathcal{H}^k$–a.e. $x \in F_i(K_i)$. If $v_1, \ldots, v_k$ is an orthonormal base of $T_x\Gamma$, by standard multilinear algebra we have

$$|d(\Phi_t \circ F_i)|_y e_1 \wedge \ldots \wedge d(\Phi_t \circ F_i)|_y e_k| = \underbrace{|d\Phi_t|_x v_1 \wedge \ldots \wedge d\Phi_t|_x v_k|}_{=:h_x(t)} |dF_i|_y e_1 \wedge \ldots \wedge dF_i|_y e_k| \,.$$

Using again the area formula and the fact that Φ_0 is the identity we conclude:

$$\frac{1}{t}\left(\mathbf{M}((\Phi_t)_\sharp V) - \mathbf{M}(V)\right) = \int \frac{h_x(t) - h_x(0)}{t}\, d\mu_V(x)\,. \tag{1.6}$$

Next recall that $h_x(t) = \sqrt{\det M_x(t)}$, where $M_x(t)$ is the matrix with entries

$$(M_x(t))_{ij} = \langle d\Phi_t|_x v_i, d\Phi_t|_x v_j\rangle\,.$$

Differentiating in t we then compute

$$(M'_x(t))_{ij} = \langle D_{v_i}(X\circ\Phi_t)(x), d\Phi_t|_x v_j\rangle + \langle d\Phi_t|_x v_i, D_{v_j}(X\circ\Phi_t)(x)\rangle\,.$$

Thus $|M'_x(t)| \le C$, where C is independent of x. So h_x is differentiable and there is $\delta > 0$ and C such that $|h_x(t) - h_x(0)| \le Ct$ for all $t\in[-\delta,\delta]$ and all $x\in U$. Therefore,

$$\lim_{t\downarrow 0}\frac{1}{t}\left(\mathbf{M}((\Phi_t)_\sharp V) - \mathbf{M}(V)\right) = \int h'_x(0)\, d\mu_V(x)\,.$$

Since $(M_x(0))_{ij} = \delta_{ij}$ we compute $h'_x(0) = \frac{1}{2}\mathrm{tr}\, M'_x(0) = \sum_i v_i\cdot D_{v_i}X(x) = \mathrm{div}_{T_x\Gamma}X(x)$. □

2. The Monotonicity Formula and its consequences

The monotonicity formula is a key tool to derive first regularity results for objects which are a weak version of minimal surfaces. The monotonicity formula for general varifolds were derived first by Allard in Section 5 of [2] (cf. also Section 40 of [13]). Here we follow a simpler approach for rectifiable varifolds, taken from Section 17 of [13] with minor technical modifications. Having fixed $\xi\in U\subset\mathbb{R}^N$, we define $r(x) := |x-\xi|$. Next, if $g: U\to\mathbb{R}$ is differentiable and $V = (\Gamma, f)$ an integral varifold in U, for $\mathcal{H}^k$–a.e. $x\in\Gamma$ we denote by $\nabla^\perp g(x)$ the projection of the gradient ∇g on the space orthogonal to $T_x\Gamma$.

Theorem 2.1 (Monotonicity Formula)**.** *Let V be an integral varifold of dimension k in U with bounded generalized mean curvature H. Fix $\xi\in U$. For every $0<\sigma<\rho<\mathrm{dist}\,(\xi,\partial U)$ we have the following* Monotonicity Formula

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

where $m(r) = \max\{r,\sigma\}$. Hence the map $\rho\mapsto e^{\|H\|_\infty\rho}\rho^{-k}\mu_V(B_\rho(\xi))$ is monotone increasing.

Proof. W.l.o.g. we assume $\xi = 0$. Fix a function $\gamma\in C^1_c([0,1[)$ such that $\gamma\equiv 1$ in a neighborhood of 0. For any $s\in]0,\mathrm{dist}(0,\partial U)[$ we let X_s be the vector field $X_s(x) = \gamma(\frac{|x|}{s})x$. Observe that $X_s\in C^1_c(U)$ and hence we can apply (1.3) and (1.4) to conclude

$$\int \mathrm{div}_{T_x\Gamma}X_s\, d\mu_V = -\int H\cdot X_s\, d\mu_V\,. \tag{2.1}$$

Fix now a plane $\pi = T_x\Gamma$. We choose an orthonormal frame $e_1, \dots, e_k$ spanning π and complete it to an orthonormal base of $\mathbb{R}^N$. We are then ready to compute

$$\begin{aligned}\mathrm{div}_\pi X_s &= k\gamma\left(\frac{r}{s}\right) + \sum_{j=1}^k e_j\cdot x\,\gamma'\left(\frac{r}{s}\right)\frac{x\cdot e_j}{|x|s} = k\gamma\left(\frac{r}{s}\right) + \frac{r}{s}\gamma'\left(\frac{r}{s}\right)\sum_{j=1}^k\left(\frac{x\cdot e_j}{|x|}\right)^2\\ &= k\gamma\left(\frac{r}{s}\right) + \frac{r}{s}\gamma'\left(\frac{r}{s}\right)\left(1-\sum_{j=k+1}^N\left(\frac{x\cdot e_j}{|x|}\right)^2\right) = k\gamma\left(\frac{r}{s}\right)+\frac{r}{s}\gamma'\left(\frac{r}{s}\right)\left(1-|\nabla^\perp r|^2\right). \qquad (2.2)\end{aligned}$$

Insert (2.2) into (2.1), divide by s^{k+1} and integrate in s between σ and ρ:

$$\begin{aligned}&\int_\sigma^\rho\int_{\mathbb{R}^N}\frac{k}{s^{k+1}}\gamma\left(\frac{|x|}{s}\right)d\mu_V(x)\,ds + \int_\sigma^\rho\int_{\mathbb{R}^N}\frac{|x|}{s^{k+2}}\gamma'\left(\frac{|x|}{s}\right)(1-|\nabla^\perp r|^2)d\mu_V(x)\,ds\\ &= -\int_\sigma^\rho\int_{\mathbb{R}^N}\frac{H\cdot x}{s^{k+1}}\gamma\left(\frac{|x|}{s}\right)d\mu_V(x)\,ds\,.\end{aligned}$$

We then use Fubini's Theorem and distribute the integrands to obtain:

$$\int_{\mathbb{R}^N}\int_\sigma^\rho\left(\frac{k}{s^{k+1}}\gamma\left(\frac{|x|}{s}\right)+\frac{|x|}{s^{k+2}}\gamma'\left(\frac{|x|}{s}\right)\right)ds\,d\mu_V \qquad (2.3)$$

$$= \int_{\mathbb{R}^N}|\nabla^\perp r|^2\int_\sigma^\rho\frac{|x|}{s^{k+2}}\gamma'\left(\frac{|x|}{s}\right)ds\,d\mu_V(x) - \int_{\mathbb{R}^N}H\cdot x\int_\sigma^\rho\frac{1}{s^{k+1}}\gamma\left(\frac{|x|}{s}\right)ds\,d\mu_V(x)\,. \qquad (2.4)$$

Observe that

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

We can therefore rewrite (2.4) as

$$\begin{aligned}&\rho^{-k}\int\gamma\left(\frac{|x|}{\rho}\right)d\mu_V(x) - \sigma^{-k}\int\gamma\left(\frac{|x|}{\sigma}\right)d\mu_V(x) - \int_{\mathbb{R}^N}H\cdot x\int_\sigma^\rho s^{-k-1}\gamma\left(\frac{|x|}{s}\right)ds\,d\mu_V(x)\\ &= \int_{\mathbb{R}^N}|\nabla^\perp r|^2\left[\rho^{-k}\gamma\left(\frac{|x|}{\rho}\right)-\sigma^{-k}\gamma\left(\frac{|x|}{\sigma}\right)+\int_\sigma^\rho\frac{k}{s^{k+1}}\gamma\left(\frac{|x|}{s}\right)ds\right]d\mu_V(x)\,. \qquad (2.5)\end{aligned}$$

We now test (2.5) with a sequence of nonnegative cut-off functions $\gamma = \gamma_n$ which converge to $\mathbf{1}_{]-1,1[}$ from below. By the Dominated Convergence Theorem, we conclude that we can insert $\gamma = \mathbf{1}_{]0,1[}$ into (2.5). Hence, the Monotonicity Formula follows from

$$\int_\sigma^\rho\frac{k}{s^{k+1}}\mathbf{1}_{]0,1[}\left(\frac{|x|}{s}\right)ds = \mathbf{1}_{B_\rho}(x)\int_{\max\{|x|,\sigma\}}^\rho\frac{k}{s^{k+1}}ds = \left(\frac{1}{\max\{|x|,\sigma\}^k}-\frac{1}{\rho^k}\right)\mathbf{1}_{B_\rho}(x)\,.$$

Finally, define $f(\rho) := \rho^{-k}\mu_V(B_\rho)$. We use the Monotonicity Formula to bound trivially

$$\frac{f(\rho)-f(\sigma)}{\rho-\sigma} \geq -\frac{\|H\|_\infty}{k}\int_{B_\rho}|x|\frac{\max\{|x|,\sigma\}^{-k}-\rho^{-k}}{\rho-\sigma}d\mu_V(x) \geq -\frac{\|H\|_\infty}{k}\mu_V(B_\rho)\rho\frac{\sigma^{-k}-\rho^{-k}}{\rho-\sigma}\,.$$

Since $\rho \mapsto \rho^{-k}$ is convex, setting $\rho = \sigma + \varepsilon$ we conclude

$$\frac{f(\sigma+\varepsilon)-f(\sigma)}{\varepsilon} \geq -\mu_V(B_\rho)\|H\|_\infty(\sigma+\varepsilon)\sigma^{-k-1} = -\|H\|_\infty f(\sigma+\varepsilon)\frac{(\sigma+\varepsilon)^{k+1}}{\sigma^{k+1}}\,. \tag{2.6}$$

If ψ_δ is a standard smooth nonnegative mollifier, we can first take the convolution of both sides of (2.6) (as functions of σ) and then let $\varepsilon \downarrow 0$ to conclude $(f*\psi_\delta)' + \|H\|_\infty (f*\psi_\delta) \geq 0$. Thus the function $g_\delta(\rho) := e^{\|H\|_\infty \rho} f * \psi_\delta(\rho)$ is monotone increasing. Letting $\delta \downarrow 0$ we conclude that $\rho \mapsto e^{\|H\|_\infty \rho}\rho^{-k}\mu_V(B_\rho)$ is also monotone increasing. □

The Monotonicity Formula has several important corollaries: the following proposition can be understood as a first regularity result for varifolds with bounded mean curvature (in what follows $\mathrm{spt}\,(\mu_V)$ denotes the support of the measure μ, i.e. the smallest (relatively) closed set $F \subset U$ such that $\mu_V(U \setminus F) = 0$).

Proposition 2.2. *Let $V = (\Gamma, f)$ be an integral varifold of dimension k in U with bounded mean curvature. Then the limit*

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

exists at every *$x \in U$ and coincides with $f(x)$ for μ_V–a.e. x. Moreover*

(i) *θ_V is upper semicontinuous;*
(ii) *$\theta_V(x) \geq 1$ and $\mu_V(B_\rho(x)) \geq \omega_k e^{-\|H\|_\infty \rho}\rho^k$ $\forall x \in \mathrm{spt}\,(\mu_V)$ and $\forall \rho < \mathrm{dist}(x, \partial U)$;*
(iii) *$\mathcal{H}^k(\mathrm{spt}\,(\mu_V) \setminus \Gamma) = 0$.*

Proof. The existence of the limit is guaranteed by the monotonicity of $e^{\|H\|_\infty \rho}\rho^{-k}\mu_V(B_\rho(x))$. Moreover, $\theta_V = f$ μ_V–a.e. by the standard Density Theorems for rectifiable sets and Radon measures (see Theorem 2.12 and Conclusion (1) of Theorem 16.2 of [9]). Fix next $x \in U$ and $\varepsilon > 0$. Let $0 < 2\rho < \mathrm{dist}(x, \partial U)$ be such that

$$e^{\|H\|_\infty r}\frac{\mu_V(B_r(x))}{\omega_k r^k} \leq \theta_V(x) + \frac{\varepsilon}{2} \qquad \forall r < 2\rho. \tag{2.7}$$

If $\delta < \rho$ and $|x - y| < \delta$, we then conclude

$$\theta_V(y) \leq e^{\|H\|_\infty \rho}\frac{\mu_V(B_\rho(y))}{\omega_k\rho^k} \leq e^{\|H\|_\infty(\rho+\delta)}\frac{\mu_V(B_{\rho+\delta}(x))}{\omega_k\rho^k} \stackrel{(2.7)}{\leq} \left(1+\frac{\delta}{\rho}\right)^k \left(\theta_V(x) + \frac{\varepsilon}{2}\right)\,. \tag{2.8}$$

If δ is sufficiently small we infer $\theta_V(y) \leq \theta_V(x) + \varepsilon$, which proves (i).

Since $\theta_V = f$ μ_V–a.e. and f is integer valued, the set $\{\theta_V \geq 1\}$ has full μ_V measure. Thus $\{\theta_V \geq 1\}$ must be dense in $\mathrm{spt}\,(\mu_V)$ and so for every $x \in \mathrm{spt}\,(\mu_V) \cap U$ the inequality $\theta_V(x) \geq 1$ follows from the upper semicontinuity of θ_V. The remaining assertion in (ii) is then a consequence of Theorem 2.1. Finally, by the classical Density Theorems (see Theorem 6.2 of [9]) $\theta_V = 0$ $\mathcal{H}^k$–a.e. on $U \setminus \Gamma$ and hence (iii) follows from (ii). □

3. Allard's ε-regularity Theorem and its consequences

From now on, given an integral varifold $V = (\Gamma, \theta)$ in U with bounded mean curvature, we assume,w.l.o.g. , that $\overline{\Gamma} \cap U = \Gamma$ and that $\theta(x) = \theta_V(x)$ for every $x \in U$ (this is not completely consistent with Definition 1.1 because θ is not *everywhere* integer valued, but a slight technical adjustment in Definition 1.1 would keep our notation consistent). To state Allard's Theorem we need to introduce the "principal parameter" of its smallness assumption. Given two k-dimensional planes π_1 and π_2 in $\mathbb{R}^N$ and the corresponding orthogonal projections $P_i : \mathbb{R}^N \to \mathbb{R}^N$ onto π_i, we denote by $\|\pi_1 - \pi_2\|$ the Hilbert-Schmidt norm of $P_1 - P_2$.

Definition 3.1 (Excess)**.** Let V be a k-dimensional integral varifold in U, $B_r(x) \subset U$ an open ball and π a k-dimensional plane. The *excess of V in $B_r(x)$ with respect to π* is

$$E(V, \pi, x, r) := r^{-k} \int_{B_r(x)} \|T_y\Gamma - \pi\|^2 \, d\mu_V(y) \,. \tag{3.1}$$

The following is the main result of Allard's regularity theory for stationary integral varifolds. It corresponds essentially to the Regularity Theorem of Section 8 of [2] (cf. also Theorem 23.1 of [13]).

Theorem 3.2 (Allard's ε-regularity Theorem)**.** *For any positive integer $k < N$ there are positive constants α, ε and γ with the following property. Let $V = (\Gamma, \theta)$ be a k-dimensional integral varifold with bounded mean curvature H in $B_r(x_0) \subset \mathbb{R}^N$ such that:*

(H1) $\mu_V(B_r(x_0)) < (\omega_k + \varepsilon) r^k$ *and* $\|H\|_\infty < \varepsilon r^{-1}$*;*

(H2) *There is a plane π such that $E(V, \pi, x_0, r) < \varepsilon$.*

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

We already pointed out the redundancy of (H2) (see Remark 5.4 for a thorough explanation). Nonetheless, the version above is still powerful enough to prove the two main conclusions of Allard's interior regularity theory (cf. Section 8.1 of [2]).

Corollary 3.3. *Let α be as in Theorem 3.2 and $V = (\Gamma, \theta)$ a k-dimensional integral varifold with bounded mean curvature in $U \subset \mathbb{R}^N$. Then there is an open set $W \subset U$ such that $\Gamma \cap W$ is a $C^{1,\alpha}$ submanifold of W without boundary and it is dense in Γ. If in addition $\theta = Q$ μ_V–a.e. for some $Q \in \mathbb{N} \setminus \{0\}$, then $\mu_V(\Gamma \setminus W) = 0$.*

Proof. We first prove the second statement. Observe that, under the assumption that $\theta = Q$ μ_V–a.e., the varifold $(V, Q^{-1}\theta)$ is also integral with bounded generalized mean curvature H. Therefore it suffices to prove the statement when $Q = 1$. This however follows from standard measure theoretic arguments: indeed, for μ_V–a.e. $x_0 \in \Gamma$ we have

$$1 = \theta(x_0) = \lim_{r \downarrow 0} \frac{\mu_V(B_r(x_0))}{\omega_k r^k} \qquad \text{and} \qquad \lim_{r \downarrow 0} r^{-k} \int_{B_r(x_0)} \|T_x\Gamma - T_{x_0}\Gamma\|^2 d\mu_V(x) = 0 \,.$$

Thus, for any such x_0 there is an $r := r_{x_0} > 0$ such that (H1) and (H2) hold with $\pi = T_{x_0}\Gamma$: by Theorem 3.2 we then conclude that $\Gamma \cap B_{\gamma r_{x_0}}(x_0)$ is a $C^{1,\alpha}$ submanifold of $B_{\gamma r_{x_0}}(x_0)$ without boundary. The open set W is then the union of all such balls.

We next come to the first statement. Consider Γ and $\overline{\Gamma}$ as metric spaces, with the metric induced by $\mathbb{R}^N$. $\overline{\Gamma}$ is complete and Γ an open subset of it: therefore Γ is a Baire space. By the semicontinuity of θ, the sets $C_i := \{\theta \geq i\}$ are closed. For each $i \in \mathbb{N} \setminus 0$ we denote by D_i the interior of C_i in the topology of Γ and we define $E_i := D_i \setminus C_{i+1}$ and $E := \bigcup_{i\geq 1} E_i$. Fix $x \in \Gamma \setminus E$ and let $i \in \mathbb{N} \setminus \{0\}$ be such that $i \leq \theta(x) < i + 1$. By upper-semicontinuity, θ takes values in $[1, i+1[$ in a neighborhood of x. If $x \in D_i$, there would then be a neighborhood where θ takes values in $[i, i+1[$ and hence x would belong to E_i. We then conclude that $x \in C_i \setminus D_i$. Therefore $\Gamma \setminus E \subset \bigcup_i C_i \setminus D_i$. Since each $C_i \setminus D_i$ is a closed set with empty interior, it is meager. By the Baire Category Theorem E is a dense open set. For each i let $U_i \subset \mathbb{R}^N$ be an open set such that $U_i \cap \Gamma = E_i$. Consider the varifold $V_i = (\Gamma \cap U_i, \theta|_{U_i})$. V_i is an integral varifold with bounded mean curvature in U_i and $\theta|_{U_i} = i\ \mu_{V_i}$–a.e.. Thus, by the first part of the proof, there is an open set $W_i \subset U_i$ such that $\Gamma \cap W_i$ is a $C^{1,\alpha}$ submanifold of W_i without boundary and $\mu_{V_i}(\Gamma \cap (U_i \setminus W_i)) = 0$. It follows that $\Gamma \cap W_i$ is dense in $\Gamma \cap U_i = E_i$. If we set $W := \bigcup_i W_i$ we then conclude that $\Gamma \cap W$ is a $C^{1,\alpha}$ submanifold of W without boundary and that $\Gamma \cap W$ is dense in E. By the density of E in Γ we conclude the proof. □

The rest of these notes will be dedicated to prove Theorem 3.2. The core of the argument leads to an "excess-decay" Theorem which plays a pivotal role and we state immediately (cf. Theorem 22.5 of [13]).

Theorem 3.4. *Fix any positive integer $k < N$. There are constants $0 < \eta < \frac{1}{2}$ and $\varepsilon_0 > 0$ such that the following holds. If $V = (\Gamma, \theta)$ satisfies the assumptions of Theorem 3.2 with ε_0 in place of ε and $\|H\|_\infty r \leq E(V, \pi, x_0, r)$, then there is a k-dimensional plane $\bar{\pi}$ with*

$$E(V, \bar{\pi}, x_0, \eta r) \leq \frac{1}{2} E(V, \pi, x_0, r)\,. \tag{3.2}$$

The remaining pages are then organized as follows:

- In Section 4 we prove an inequality for the excess which is a direct analogue of the Caccioppoli's inequality for solutions of elliptic partial differential equations.
- In Section 5 we show that, under the assumptions of Theorem 3.2, the set Γ can be well approximated by a Lipschitz graph. The proof given here is relatively short compared to the one in [13], but we pay a high price in the accuracy of the estimates.
- In Section 6 we use the previous sections to prove Theorem 3.4.
- In Section 7 we use Theorem 3.4 and the Lipschitz approximation of Section 5 to conclude the proof of Theorem 3.2.

4. The tilt-excess inequality

The first step towards the proof of Theorem 3.4 is the following inequality (cf. Lemma 8.13 of [2] and Lemma 22.2 of [13]).

Proposition 4.1 (Tilt-excess inequality)**.** *Let $k < N$ be a positive integer. Then there is a constant C such that the following inequality holds for every integral varifold V with* *bounded generalized mean curvature H in $B_r(x_0) \subset \mathbb{R}^N$ and every k-dimensional plane π:*

$$E(V,\pi,x_0,\tfrac{r}{2}) \leq \frac{C}{r^{k+2}} \int_{B_r(x_0)} \mathrm{dist}(y-x_0,\pi)^2\, d\mu_V(y) + \frac{2^{k+1}}{r^{k-2}} \int_{B_r(x_0)} |H|^2\, d\mu_V\,. \tag{4.1}$$

To get an intuition of (4.1), consider $V = (\Gamma, 1)$ and assume Γ is the graph of a function f with small Lipschitz constant. The boundedness of H translates into a suitable elliptic system of partial differential equations for f. The left hand side of (4.1) approximates the Dirichlet energy of f and the first integral in the right hand side compares to the L^2 norm of f. Therefore (4.1) can be interpreted as a Caccioppoli inequality (see for instance [8]). Indeed the proof of (4.1) is achieved by testing (1.3) (the "weak form" of the elliptic system) with a suitable vector field.

Before coming to this we point out some elementary computations which are essential in the proof of (4.1) but will also be used a few more times later. First of all we introduce some notation. Given a k-dimensional plane π, we denote by P_π and $P_\pi^\perp$ respectively the orthogonal projection onto π and the one onto its orthogonal complement. Similarly, for $f \in C^1(\mathbb{R}^N)$, $\nabla_\pi f$ and $\nabla_\pi^\perp f$ will denote, respectively, $P_\pi \circ \nabla f$ and $P_\pi^\perp \circ \nabla f$. Finally, if $\Phi \in C^1(\mathbb{R}^N, \mathbb{R}^k)$, $J_\pi \Phi$ will denote the absolute value of the Jacobian determinant of $\Phi|_\pi$.

Lemma 4.2. *Consider two k-dimensional planes π and T in $\mathbb{R}^N$. Let $X : \mathbb{R}^N \to \mathbb{R}^N$ be the vector field $X(x) = P_\pi^\perp(x)$ and, having fixed an orthonormal base $\nu_{k+1}, \ldots, \nu_N$ of the orthogonal complement of π, consider the functions $f_j(x) = x \cdot \nu_j$. We then have*

$$\frac{1}{2}\|T-\pi\|^2 = \mathrm{div}_T X = \sum_{i=k+1}^N |\nabla_T f_i|^2\,. \tag{4.2}$$

Moreover there is a positive constant C_0, depending only on N and k, such that

$$|J_T P_\pi - 1| \leq C_0 \|T-\pi\|^2\,. \tag{4.3}$$

Proof. Let $\xi_1, \ldots, \xi_k$ be an orthonormal base of T and $e_{k+1}, \ldots, e_N$ an orthonormal base of the orthogonal complement of T. Denote by $\langle\ :\ \rangle$ the standard inner product on $\mathbb{R}^N \otimes \mathbb{R}^N$. Observe that $P_\pi = \mathrm{Id} - \sum_j \nu_j \otimes \nu_j$ and that $P_T = \mathrm{Id} - \sum_i e_i \otimes e_i$. Since $\langle a \otimes b, v \otimes w\rangle = (a \cdot v)(b \cdot w)$, we compute

$$\begin{aligned}
\frac{1}{2}\|\pi - T\|^2 &= \frac{1}{2}\Big\langle \sum_{j=k+1}^N \nu_j \otimes \nu_j - \sum_{i=k+1}^N e_i \otimes e_i : \sum_{j=k+1}^N \nu_j \otimes \nu_j - \sum_{i=k+1}^N e_i \otimes e_i \Big\rangle \\
&= (N-k) - \sum_{j=k+1}^N \sum_{i=k+1}^N (\nu_j \cdot e_i)^2 = \sum_{j=k+1}^N \left(1 - \sum_{i=k+1}^N (\nu_j \cdot e_i)^2\right) \\
&= \sum_{j=k+1}^N \sum_{i=1}^k (\nu_j \cdot \xi_i)^2 = \sum_{j=k+1}^N |\nabla_T f_j|^2
\end{aligned} \tag{4.4}$$

and

$$\sum_{i=k+1}^N |\nabla_T f_i|^2 = \sum_{i=k+1}^N \sum_{j=1}^k (\xi_j\cdot\nu_i)^2 = \sum_{j=1}^k \xi_j \cdot \sum_{i=k+1}^N (\xi_j\cdot\nu_i)\nu_i = \sum_{j=1}^k \sum_{i=k+1}^N \xi_j\cdot \big(D_{\xi_j} f_i \nu_i\big)$$
$$= \sum_{j=1}^k \xi_j \cdot D_{\xi_j} \sum_{i=k+1}^N f_i\nu_i = \sum_{j=1}^k \xi_j\cdot D_{\xi_j} X = \mathrm{div}_T X\,. \tag{4.5}$$

Next, recall that $J_T P_\pi$ is the square root of the determinant of the matrix
$M_{ij} = P_\pi(\xi_j)\cdot P_\pi(\xi_i) = \delta_{ij} - P_\pi^\perp(\xi_j)\cdot\xi_i - \xi_j\cdot P_\pi^\perp(\xi_j) + P_\pi^\perp(\xi_j)\cdot P_\pi^\perp(\xi_i) =: \delta_{ij} + A_{ij} + A_{ji} + B_{ij}$.
Observe that $|P_\pi^\perp(\xi_i)| = |P_\pi^\perp(\xi_i) - P_T^\perp(\xi_i)| \leq C\|T-\pi\|$. Thus $\|A\|\leq C\|T-\pi\|$ and $\|B\|\leq \|T-\pi\|^2$. Moreover,

$$A_{ij} := \xi_i\cdot P_\pi^\perp(\xi_j) = \sum_{i=1}^k \xi_i\cdot \sum_{l=k+1}^N (\xi_j\cdot\nu_l)\nu_l = \xi_i\cdot D_{\xi_j} X\,.$$

Thus, the usual Taylor expansion of the determinant gives

$$\det M = 1 - 2\mathrm{tr}\, A + O(\|\pi - T\|^2) = 1 - 2\mathrm{div}_T X + O(\|\pi-T\|^2)\,.$$

Next, since $J_T P_\pi \geq 0$ we have $|J_T\pi - 1| \leq |J_T\pi - 1|(J_T\pi + 1) = |\det M - 1|$. Combining the inequalities above we then obtain (4.3). □

Proof of Proposition 4.1. By scaling and translating we can assume that $x_0 = 0$ and $r = 1$. Consider a smooth cut-off function $\zeta\in C^\infty_c(B_1)$ such that $\zeta\equiv 1$ on $B_{1/2}$. Let X be the vector field of Lemma 4.2 and test (1.3) with $\zeta^2 X$. We then conclude

$$\int \zeta^2 \mathrm{div}_{T_y\Gamma} X\, d\mu_V(y) \overset{(1.3)}{=} -\int \zeta^2 H\cdot X\, d\mu_V - \int 2\zeta X\cdot \nabla_{T_y\Gamma}\zeta\, d\mu_V\,. \tag{4.6}$$

We next set $T := T_y\Gamma$ and use the notation of the proof of Lemma 4.2 to estimate

$$\zeta|\nabla_T\zeta\cdot X| = \zeta\left|\sum_{j=1}^k (\nabla\zeta\cdot\xi_j)(\xi_j\cdot X)\right| \leq \zeta|\nabla\zeta| \sum_{j=1}^k\sum_{i=k+1}^N |f_i||\xi_j\cdot\nu_i| \leq C\zeta|\nabla\zeta||X| \sum_{i=k+1}^N |\nabla_T f_i|$$
$$\leq C|\nabla\zeta|^2|X|^2 + \frac{1}{4}\zeta^2 \sum_{i=k+1}^N |\nabla_T f_i|^2 \overset{(4.2)}{=} C|\nabla\zeta|^2|X|^2 + \frac{1}{4}\zeta^2\mathrm{div}_T X\,. \tag{4.7}$$

Inserting (4.7) into (4.6), we then conclude

$$\frac{1}{2}\int \zeta^2\mathrm{div}_{T_y\Gamma} X\, d\mu_V(y) \leq \frac{1}{2}\int_{B_1} |H|^2\, d\mu_V + \frac{1}{2}\int \zeta^4 |X|^2\, d\mu_V + C\int_{B_1} |\nabla\zeta|^2|X|^2\,. \tag{4.8}$$

However, by (4.2),

$$E(V,\pi,0,1/2) \leq 2^{k+1}\int \zeta^2\mathrm{div}_{T_y\Gamma} X\, d\mu_V(y)\,. \tag{4.9}$$

Since $|\nabla\zeta| + |\zeta|\leq C$ and $|X(x)| = \mathrm{dist}(x,\pi)$, from (4.8) and (4.9) we easily conclude the desired inequality. □

5. The Lipschitz approximation

This is probably the most technical section of these notes. The aim is to prove the following proposition, which shows that, when the excess and the generalized mean curvatures are sufficiently small, most of the varifold can be covered with a single Lipschitz graph with sufficiently small Lipschitz constant. This type of approximation was pioneered by De Giorgi in [4]. A first proposition of this type for varifolds was proved by Allard in Lemma 8.12 of [2] (cf. Theorem 20.2 of [13]). In what follows, if Ω is a subset of an affine plane $x_0+\pi$ and $f:\Omega\to\pi^\perp$, the graph of f, denoted by Γ_f, is the set $\{y+f(y):y\in\Omega\}$.

Proposition 5.1 (Lipschitz Approximation)**.** *For any positive integer $k<N$ there is a constant C with the following property. For any $\ell,\beta\in]0,1[$ there are $\lambda\in]0,1]$ (depending only on ℓ) and $\varepsilon_L>0$ such that the following holds. If $V=(\Gamma,\theta)$ and π satisfy the assuptions of Theorem 3.2 with ε_L in place of ε, then there is a map $f:(\pi+x_0)\cap B_{r/8}(x_0)\to\pi^\perp$ such that*

- (i) *The Lipschitz constant of f is less than ℓ and the graph of f (from now on denoted by Γ_f) is contained in the βr-neighborhood of $x_0+\pi$;*
- (ii) *$\theta\equiv 1$ $\mathcal{H}^k$–a.e. on $\Gamma\cap B_{r/8}(x_0)$, which is contained in the βr-neighborhood of $x_0+\pi$;*
- (iii) *Γ_f contains the set $G:=\{x\in\Gamma\cap B_{r/8}(x_0):E(V,\pi,x,\rho)\le\lambda\quad\forall\rho\in]0,r/2[\}$;*
- (iv) *The following estimate holds*

$$\mathcal{H}^k\left(\Gamma_f\setminus G\right)+\mathcal{H}^k\left((\Gamma\cap B_{r/8}(x_0))\setminus G\right)\le C\lambda^{-1}E(V,\pi,x_0,r)r^k+C\|H\|_\infty r^{k+1}\,. \tag{5.1}$$

The proof of the Proposition is based on the following Lemma, which in turn is proved using a blow-up argument based on Lemma 5.3.

Lemma 5.2. *Let $k<N$ be a positive integer. For any $\delta\in]0,\frac{1}{2}[$ there is a positive number ε_H such that, if V satisfies the assumptions of Allard's Theorem with ε_H in place of ε, then*

- (i) *$\Gamma\cap B_{r/2}(x_0)$ is contained in the δr-neighborhood of $x_0+\pi$;*
- (ii) *$\mu_V(B_\rho(x))\le(\omega_k+\delta)\rho^k$ for every $x\in B_{r/4}(x_0)$ and any $\rho\le\frac{r}{2}$.*

From now on $\mathcal{H}^k\llcorner A$ will denote the measure assigning to each Borel set B the value $\mathcal{H}^k(A\cap B)$, whereas $\stackrel{*}{\rightharpoonup}$ will denote the usual weak* convergence in the sense of measures.

Lemma 5.3. *Let $V_i=(\Gamma_i,\theta_i)$ be a sequence of k-dimensional integral varifolds on $B\subset\mathbb{R}^N$ satisfying the assumptions of Theorem 3.2 with $\varepsilon=\varepsilon(V_i)\downarrow 0$ for a given plane π. Then $\mu_{V_i}\stackrel{*}{\rightharpoonup}\mathcal{H}^k\llcorner\pi$ in B_1.*

Remark 5.4. Lemma 5.3 can indeed be proved under the only assumption that (H1) in Theorem 3.2 holds. The vanishing of the excess can then be drawn as a conclusion and thus assumption (H2) in Theorem 3.2 is redundant. However this stronger version of Lemma 5.3 requires the theory of general varifolds and in particular the Compactness Theorem for integral varifolds (see Thorem 6.4 of [2] and Theorem 42.7 of [13]).

Remark 5.5. Lemma 5.2 is a "poor-man" version of a more precise height bound (cf. the proof of Theorem 20.2 of [13]) where δ and ε_H are related by $\delta^{2+2k}=\varepsilon_H$. This estimate (first proved in the context of area-minimizing currents by Almgren, see for instance the appendix of [12]) is considerably more difficult to prove and these notes are considerably shorter also because we avoid it. As a result the estimates of Proposition 5.1 are very crude compared to the ones of Theorem 20.2 in [13].

5.1. **Proof of Lemma 5.3.** Fix $\rho \in]0,1[$, denote by H_i the generalized mean curvature of V_i and apply the monotonicity formula to conclude that

$$\int_{B_1\setminus B_\rho} \frac{|\nabla^\perp r|^2}{r^k}\, d\mu_{V_i} \leq \mu_{V_i}(B_1) - \frac{\mu_{V_i}(B_\rho)}{\rho^k} + C\|H_i\|_\infty$$
$$\leq \mu_{V_i}(B_1) - e^{-\|H_i\|_\infty}\omega_k\theta_i(0) + C\|H_i\|_\infty \leq \mu_{V_i}(B_1) - e^{-\|H_i\|_\infty}\omega_k + C\|H_i\|_\infty \quad \stackrel{i\to\infty}{\to} \quad 0.$$

Therefore

$$\int_{B_1} |P_\pi^\perp(y)|^2 d\mu_{V_i} \leq 2\int_{B_1} |P_\pi^\perp(y) - P_{T_y\Gamma}^\perp(y)|^2 d\mu_{V_i}(y) + 2\int_{B_1} |P_{T_y\Gamma}^\perp(y)|^2 d\mu_{V_i}(y)$$
$$\leq C\int_{B_1} \|\pi - T_y\Gamma\|^2 d\mu_{V_i} + C\int_{B_1\setminus B_\rho} \frac{|\nabla^\perp r|^2}{r^k} d\mu_{V_i} + C\rho^2\mu_{V_i}(B_\rho)\,. \tag{5.2}$$

Assume now that a subsequence, not relabeled, of $\{\mu_{V_i}\}$ converges weakly* to some measure μ and fix a nonnegative $\varphi \in C_c(B_1)$. We then conclude

$$\int |P_\pi^\perp(y)|^2\varphi(y)\, d\mu(y) = \lim_{i\to\infty} \int |P_\pi^\perp(y)|^2\varphi(y)\, d\mu_{V_i}(y) \stackrel{(5.2)}{=} 0\,.$$

Thus μ is supported in π. On the other hand, for any $x \in B_1$ and any $\rho < 1 - |x|$ we have, by the monotonicity formula,

$$\frac{\mu(B_\rho(x))}{\rho^k} \leq \liminf_{i\to\infty} \frac{\mu_{V_i}(B_\rho(x))}{\rho^k} \leq \liminf_{i\to\infty} \frac{e^{\|H_i\|_\infty}\mu_{V_i}(B_{1-|x|}(x))}{(1-|x|)^k} \leq \frac{\omega_k}{(1-|x|)^k}\,.$$

Hence the upper k-dimensional density of μ is finite, which in turn implies $\mu = \theta\mathcal{H}^k \llcorner \pi$ for some nonnegative Borel function θ. Fix a vector field $X \in C^1_c(B_1)$ and consider

$$\left|\int_\pi \mathrm{div}_\pi X\theta d\mathcal{H}^k\right| = \lim_{i\to\infty}\left|\int \mathrm{div}_\pi X\, d\mu_{V_i}\right|$$
$$\leq \liminf_{i\to\infty}\left\{\left|\int \mathrm{div}_{T\Gamma_i} X\, d\mu_{V_i}\right| + C\|DX\|_\infty \int \|T\Gamma_i - \pi\|\, d\mu_{V_i}\right\}$$
$$\leq \liminf_{i\to\infty} C\left(\|H_i\|_\infty\|X\|_\infty\mu_{V_i}(B_1) + \|DX\|_\infty E(V_i,\pi,0,1)^{1/2}(\mu_{V_i}(B_1))^{1/2}\right) = 0\,.$$

Let now $z_1,\ldots,z_k,y_1,\ldots,y_{N-k}$ be a system of coordinates such that $\pi = \{y = 0\}$ and denote by $\mathcal{B}_r$ the ball of $\mathbb{R}^k$ with radius r and centered at 0. The last equality implies

$$\int \theta(z)\,\mathrm{div}_z Y(z)\, dz = 0 \qquad \text{for any vector field } Y \in C^1_c(\mathcal{B}_1, \mathbb{R}^k). \tag{5.3}$$

It is well known that (5.3) implies the constancy of θ: take for instance a standard mollifier φ_δ and test (5.3) with vector fields of type $Y * \varphi_\delta$ to conclude that the derivative of $\theta * \varphi_\delta$ vanishes on $\mathcal{B}_{1-\delta}$; letting $\delta \downarrow 0$ we then conclude that θ is a constant θ_0. On the other hand, since $\mu(\partial B_\rho) = 0$, we have $\theta_0 \omega_k \rho^k = \mu(B_\rho) = \lim_{i\to\infty} \mu_{V_i}(B_\rho)$. However, as already observed, by the Monotonicity Formula, $\mu_{V_i}(B_\rho) \to \omega_k \rho^k$. Thus $\theta_0 = 1$.

Summarizing, any convergent subsequence of $\{\mu_{V_i}\}$ converges to $\mathcal{H}^k \llcorner \pi$. By the weak* compactness of bounded closed convex sets in the space of measures, we conclude the proof. □

5.2. **Proof of Lemma 5.2.** By scaling and translating we can assume $x_0 = 0$ and $r = 1$. By possibly rotating, if the proposition were false, then there would be a positive constant $\delta < \frac{1}{2}$, a plane π and a sequence of varifolds $V_i = (\Gamma_i, \theta_i)$ which satisfy the assumptions of Lemma 5.3 and such that, for each i, one of the following two alternatives holds:

(A) there is a point $x_i \in \Gamma_i \cap B_{1/2}$ such that $|P_\pi^\perp(x_i)| \geq \delta$;
(B) there is a point $x_i \in B_{1/4}$ and a radius $\rho_i \in]0, \frac{1}{2}[$ such that $\mu_{V_i}(B_{\rho_i}(x_i)) \geq (\omega_k + \delta)\rho_i^k$.

By Lemma 5.3 we know that $\mu_{V_i} \overset{*}{\rightharpoonup} \mathcal{H}^k \llcorner \pi$ and w.l.o.g. we can assume that one of the two alternatives holds for *every* i.

Case (A). W.l.o.g. we can assume that $x_i \to x$. Then $x \in \overline{B}_{1/2}$ and $|P_\pi^\perp(x)| \geq \delta$. Thus $B_\delta(x) \subset B_1$ and $B_\delta(x) \cap \pi = \emptyset$. On the other hand, for i large enough, $B_{\delta/2}(x_i) \subset B_\delta(x)$. Since $\mathcal{H}^k(\partial B_\delta(x) \cap \pi) = 0$, by the Monotonicity Formula we have

$$0 = \mathcal{H}^k(\pi \cap B_\delta(x)) = \lim_{i\to\infty} \mu_{V_i}(B_\delta(x)) \geq \limsup_{i\to\infty} \mu_{V_i}(B_{\delta/2}(x_i)) \geq \omega_k 2^{-k} \delta^k \,,$$

which obviously is a contradiction.

Case (B). By the Monotonicity formula, $\mu_{V_i}(B_{1/2}(x_i)) \geq e^{-\|H_i\|_\infty/2}(\omega_k + \delta)2^{-k}$. W.l.o.g. we can assume $x_i \to x \in \overline{B}_{1/4}$. Fix any radius $r > \frac{1}{2}$ and observe that the ball of radius r centered at x contains the balls $B_{1/2}(x_i)$ for i large enough. Since $\mathcal{H}^k(\pi \cap \partial B_r(x)) = 0$ we then conclude

$$\mathcal{H}^k(\pi \cap B_r(x)) = \lim_{i\to\infty} \mu_{V_i}(B_r(x)) \geq \lim_{i\to\infty} \mu_{V_i}(B_{1/2}(x_i)) \geq (\omega_k + \delta)2^{-k} \,.$$

Letting $r \downarrow \frac{1}{2}$ we then conclude $\mathcal{H}^k(\pi \cap B_{1/2}(x)) \geq (\omega_k + \delta)2^{-k}$, which is impossible. □

5.3. **Proof of Proposition 5.1.** W.l.o.g. we assume $x_0 = 0$ and $r = 1$. Moreover, to simplify the notation we set $E := E(V, \pi, 0, 1)$.

(C1) We start by choosing λ smaller than the ε_H given by Lemma 5.2 when we set $\delta = \frac{(N-k)^{-\frac{1}{2}}\ell}{6}$.
(C2) We then choose $\varepsilon_L < \lambda$ so that it is also smaller than the ε_H given by Lemma 5.2 when we set $\delta := \min\{\lambda, (N-k)^{-\frac{1}{2}}\beta\}$.

Consider any point $x \in \Gamma \cap B_{1/8}$. By Lemma 5.2 and our choice (C2) of ε_L we have $\mu(B_r(x)) \leq (\omega_k + \lambda)r^k$ for any $r < \frac{1}{2}$. Letting $r \downarrow 0$ we also conclude $\theta(x) \leq (1 + \frac{\lambda}{\omega_k}) < 2$. Since $\theta \in \mathbb{N} \setminus \{0\}$ for $\mathcal{H}^k$–a.e. $x \in \Gamma$, we conclude that $\theta = 1$ $\mathcal{H}^k$–a.e. on $\Gamma \cap B_{1/8}$. Observe also that, again by our choice (C2) of ε_L, $\Gamma \cap B_{1/2}$ is contained in the $(N-k)^{-\frac{1}{2}}\beta$-neighborhood of π. So this shows that conclusion (ii) is satisfied.

Assume next $x \in G$ and pick a second point $y \in G$. Observe that $|y - x| < \frac{1}{4}$. Therefore choose $r > |y - x|$ so that $2r < \min\{\frac{1}{2}, 3|x - y|\}$. Since $2r < \frac{1}{2}$, by the choice of (C2) of ε_L

we have $\mu(B_{2r}(x)) \leq (\omega_k + \lambda)(2r)^k$ and since $x \in G$ we also have $E(V, \pi, x, 2r) < \lambda$. So we can apply Lemma 5.2 in $B_{2r}(x)$ and by our choice (C1) of λ this implies

$$|P_\pi^\perp(x) - P_\pi^\perp(y)| \leq 3^{-1}(N-k)^{-\frac{1}{2}}\ell\, r \leq \tfrac{1}{2}|y-x|\,,$$

because $y \in B_r(x)$. By the triangle inequality, $|P_\pi(x) - P_\pi(y)| \geq \frac{1}{2}|y-x|$. This implies that $P_\pi : G \to \pi$ is an injective map. Thus, if we set $D = P_\pi(G)$, it turns out that G is the graph of a function $f : D \to \pi^\perp$. Observe that $\|f\|_\infty \leq (N-k)^{-\frac{1}{2}}\beta$ and that

$$\begin{aligned}|f(v) - f(w)| &= |P_\pi^\perp(v, f(v)) - P_\pi^\perp(w, f(w))| \leq 2^{-1}(N-k)^{-\frac{1}{2}}\ell|(v, f(v)) - (w, f(w))|\\ &\leq (N-k)^{-\frac{1}{2}}\ell|P_\pi(v, f(v)) - P_\pi(w, f(w))| = (N-k)^{-\frac{1}{2}}\ell|v-w|\,.\end{aligned}$$

Thus $f : D \to \pi^\perp$ has Lipschitz constant $(N-k)^{-\frac{1}{2}}\ell$. Fix a system of orthonormal coordinates on $\pi^\perp$ and let $f_1, \ldots, f_{N-k}$ be the corresponding coordinate functions of f. We can then extend each f_j to $B_{1/8} \cap \pi$ preserving both the Lipschitz constant and the L^∞ norm. The resulting extension of f will then have Lipschitz constant[1] at most ℓ and L^∞ norm at most β. Thus f satisfies conclusion (i).

On the other hand (iii) holds by construction. We therefore come to the estimate (5.1). First of all, for each $x \in F := (\Gamma \setminus G) \cap B_{1/8}$ we choose a radius $\rho_x < \frac{1}{2}$ such that $E(V, \pi, x, \rho_x) \geq \lambda$. By the $5r$-Covering theorem we can find countably many pairwise disjoint balls $B_{\rho_i}(x_i)$ such that $\{B_{5\rho_i}(x_i)\}_i$ covers F and $E(V, \pi, x_i, \rho_i) \geq \lambda$. We can therefore compute:

$$\mathcal{H}^k(F) \leq 5^k\omega_k \sum_i \rho_i^k \leq \frac{5^k\omega_k}{\lambda}\sum_i E(V, \pi, x_i, \rho_i)\rho_i^k \leq C\lambda^{-1}E\,. \tag{5.4}$$

As for estimating $F' := \Gamma_f \setminus G$ observe that, by the Area Formula,

$$\begin{aligned}\mathcal{H}^k(F') &\leq C\left(\omega_k 8^{-k} - \mathcal{H}^k(D)\right) = C\left(\omega_k 8^{-k} - \mathcal{H}^k(P_\pi(G))\right) \leq C\left(\omega_k 8^{-k} - \int_G J_{T\Gamma}P_\pi\, d\mathcal{H}^k\right)\\ &\overset{(4.3)}{\leq} C\left(\omega_k 8^{-k} + CE - \mathcal{H}^k(G)\right) \leq C\left(CE + \omega_k 8^{-k} - \mu_V(B_{1/8}) + \mathcal{H}^k(F)\right)\\ &\overset{(5.4)}{\leq} \frac{C}{\lambda}E + C(\omega_k 8^{-k} - \mu_V(B_{1/8})) \leq \frac{C}{\lambda}E + C\omega_k 8^{-k}\left(1 - e^{-\|H\|_\infty/8}\right) \qquad (5.5)\\ &\leq \frac{C}{\lambda}E + C\|H\|_\infty\,, \qquad (5.6)\end{aligned}$$

where in the last inequality of (5.5) we have used the Monotonicity formula. Finally, (5.6) combined with (5.4) proves (5.1).

[1]We could also apply directly Kirszbraun's extension theorem to f. However the proof of Kirszbraun's theorem is much more difficult than its elementary scalar counterpart, used here on each component of f.

6. Proof of Theorem 3.4

We have almost all the ingredients to prove Theorem 3.4: we only need two additional elementary lemmas about harmonic functions. Before coming to them, let us briefly explain how all these tools will be combined in the proof of Theorem 3.4. We first choose ε_0 sufficiently small so to have a Lipschitz approximation f with small Lipschitz constant, provided by Proposition 5.1. Using Lemma 6.1 we will show that f is L^2-close to a harmonic function u (the heuristic explanation: if the graph of f were minimal, then f would solve a system of PDEs which is a perturbation of the Laplacian, in particular Δf would be small in a suitable sense and the existence of a nearby harmonic function would follow from standard arguments). In a much smaller ball the L^2 norm of u is quite small (cf. Lemma 6.2) and so we conclude that the L^2 norm of f is small. This is then transformed into a bound for the left hand side of (4.1): from (4.1), i.e. the Tilt-excess inequality, we finally conclude (3.2).

Lemma 6.1 (Harmonic approximation)**.** *Let $k \in \mathbb{N}\setminus\{0\}$ and consider the ball $B_r(x) \subset \mathbb{R}^k$. For any $\varrho > 0$ there is $\varepsilon_A > 0$ with the following property. If a function $f \in W^{1,2}(B_r(x))$ with $\int |\nabla f|^2 \leq r^k$ satisfies*

$$\left|\int \nabla\varphi\cdot\nabla f\right| \leq \varepsilon_A r^k \|\nabla\varphi\|_\infty \qquad \forall \varphi \in C^1_c(B_r(x)) \tag{6.1}$$

then there is a harmonic function u on $B_r(x)$ with $\int |\nabla u|^2 \leq r^k$ such that

$$\int (f-u)^2 \leq \varrho r^{2+k}\,. \tag{6.2}$$

Lemma 6.2. *Let $k \in \mathbb{N}\setminus\{0\}$. Then there is a constant $C > 0$ such that, if u is an harmonic function in $B_r(x_0) \subset \mathbb{R}^k$, then*

$$\sup_{x\in B_\rho(x_0)} |u(x) - u(x_0) - \nabla u(x_0)\cdot(x-x_0)| \leq C\rho^2 r^{-\frac{k}{2}-1}\|Du\|_{L^2(B_r(x_0))} \qquad \forall \rho \leq \frac{r}{2}\,. \tag{6.3}$$

Proof of Lemma 6.1. W.l.o.g. we can assume $x_0 = 0$ and $r = 1$. Denote by $\mathscr{H}$ the set of harmonic functions on B_1 with Dirichlet energy at most 1. If the lemma were false, then there would be a $\varrho > 0$ and a sequence of functions $\{f_j\} \subset W^{1,2}(B_1)$ such that

$$\lim_{j\to\infty} \sup_{\varphi\in C^1_c(B_1), \|\nabla\varphi\|_\infty\leq 1} \left|\int \nabla\varphi\cdot\nabla f_j\right| = 0 \tag{6.4}$$

and $\int_{B_1} |\nabla f_j|^2 \leq 1$, but

$$\inf_{u\in\mathscr{H}} \int (u-f_j)^2 \geq \varrho\,. \tag{6.5}$$

Since both statements remain unchanged if we subtract a constant from f_j, we can assume that the average of f_j vanishes. The Poincaré inequality implies that $\|f_j\|_{W^{1,2}}$ is bounded independently of j and by Rellich's theorem we can assume that a subsequence, not relabeled, converges to $u \in W^{1,2}$ strongly in L^2. By semicontinuity of the Dirichlet integral we conclude that $\int |\nabla u|^2 \leq 1$. On the other hand, if we fix a test function $\varphi \in C^1_c(B_1)$, passing into the limit in (6.4) we conclude $\int \nabla\varphi \cdot \nabla u = 0$. Therefore u is harmonic and hence belongs to $\mathscr{H}$. But then the fact that $f_j \to u$ in L^2 contradicts (6.5). □

Proof of Lemma 6.2. W.l.o.g. we assume $r = 1$. Fix $\rho \leq \frac{1}{2}$. By the Taylor expansion:

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

On the other hand, Du is also harmonic and from the mean-value property we conclude the standard estimate $\|D^2u\|_{L^\infty(B_{1/2})} \leq C\|Du\|_{L^2(B_1)}$. □

Proof of Theorem 3.4. W.l.o.g. we assume $x_0 = 0$ and $r = 1$ and in order to simplify the notation we denote $E(V, \pi, 0, 1)$ by $\mathbf{E}$.

Lipschitz approximation. We start by assuming that ε_0 is smaller than the ε_L given by Proposition 5.1 for some choice of positive numbers ℓ and β. ℓ will be specified soon, whereas the value of β will be decided only at the end of the proof. Consider now the Lipschitz approximation $f : B_{1/8} \cap \pi \to \pi^\perp$ and the constant λ given by Proposition 5.1. To simplify the notation, from now on we will denote by $\mathcal{B}_r(x)$ the set $B_r(x) \cap \pi$. We also assume that $(y_1, \ldots, y_k, z_1, \ldots, z_{N-k})$ are orthonormal coordinates on $\mathbb{R}^N$ such that $\pi = \{z = 0\}$. We therefore denote by f_j the corresponding coordinate functions of the map f. Fix a $j \in \{1, \ldots, N-k\}$ and denote by e_j the unit vector $(0, \ldots, 0, 0, \ldots, 1, \ldots 0)$. Let next $\varphi \in C^1_c(\mathcal{B}_{1/16})$ and consider the vector field $X(y,z) = \varphi(y)e_j$. Obviously X is not compactly supported in $B_{1/8}$. However recall that $\Gamma \cap B_{1/8}$ is supported in the β-neighborhood of π. We assume that β is smaller than $\frac{1}{16}$. We can then multiply X by a cut-off function in the variables z to make it compactly supported in $B_{1/8}$ without changing its values on $\Gamma \cap B_{1/16}$. Thus (recalling that by Proposition 5.1 the density θ is 1 μ_V–a.e. on $\Gamma \cap B_{1/8}$) we can use (1.3) and the estimates of Proposition 5.1 to conclude

$$\begin{aligned}
&\left|\int_{\Gamma_f} \nabla_{T_x\Gamma_f}\varphi \cdot e_j \, d\mathcal{H}^k(x)\right| \leq \left|\int_{\Gamma_f} \nabla_{T_x\Gamma_f}\varphi \cdot e_j \, d\mathcal{H}^k(x) - \int_{\Gamma} \nabla_{T_x\Gamma}\varphi \cdot e_j \, d\mathcal{H}^k(x)\right| + |\delta V(X)|\\
\leq &\|\nabla\varphi\|_\infty \left(\mathcal{H}^k\left((\Gamma\setminus\Gamma_f)\cap B_{1/8}\right) + \mathcal{H}^k\left((\Gamma_f\setminus\Gamma)\cap B_{1/8}\right)\right) + \|H\|_\infty \mathcal{H}^k(\Gamma\cap B_{1/8})\|\varphi\|_\infty\\
\leq &\left(C\lambda^{-1}\mathbf{E} + C\|H\|_\infty\right)\|\nabla\varphi\|_\infty \leq C\lambda^{-1}\mathbf{E}\|\nabla\varphi\|_\infty\,. \qquad (6.6)
\end{aligned}$$

(In the last inequality we have used the assumption $\|H\|_\infty \leq \mathbf{E}$). Next, let $\xi_1, \ldots, \xi_k$ be the standard basis vectors of π and denote by g the $k \times k$ matrix with the entries:

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

It follows easily that there is a constant C such that $|g_{ij} - \delta_{ij}| \leq C|Df|^2$. Thus, if ℓ is smaller than some geometric constant, we can conclude the same estimate for the inverse matrix, whose entries will be denoted by g^{ij}: $|g^{ij} - \delta^{ij}| \leq C|Df|^2$.

The projection $P_{T\Gamma_f}$ is then given by the formula

$$P_{T\Gamma_f}(w) = \sum_{i,j} w \cdot v_i\, g^{ij} v_j\,.$$

We easily compute

$$e_j \cdot v_l = \partial_{y_l} f_j \qquad \nabla\varphi \cdot v_m = \partial_{y_m}\varphi\,.$$

Therefore if we fix the point $x = (w, f(w))$,

$$\begin{aligned} P_{T_x\Gamma_f}(\nabla\varphi(w)) \cdot e_j &= \sum_{l,i} \partial_{y_i} f_j(w) g^{il}(w) \partial_{y_l}\varphi(w) \\ &= \sum_{l} \partial_{y_l} f_j(w) \partial_{y_l}\varphi(w) + O(|Df|^3(w)|\nabla\varphi(w)|)\,. \end{aligned} \tag{6.7}$$

Next, we introduce the notation $\bar\nabla\varphi = (\partial_{y_1}\varphi, \ldots, \partial_{y_k}\varphi)$ and we call Jf the Jacobian

$$Jf(w) := \sqrt{1 + |Df(w)|^2 + \sum_{\alpha,\beta} (M^{\alpha,\beta}(w))^2}$$

where the last sum goes over all $n \times n$ minors $M^{\alpha\beta}$ of Df with $n \geq 2$. Recall the area formula:

$$\int_{\Gamma_f} \nabla_{T_x\Gamma_f}\varphi \cdot e_j\, d\mathcal{H}^k(x) = \int_{\mathcal{B}_{1/8}} P_{T_x\Gamma_f}(\nabla\varphi(w)) \cdot e_j Jf(w)\, dw\,. \tag{6.8}$$

On the other hand a simple Taylor expansion gives $|Jf(w) - 1| \leq C|Df(w)|^2$. Combining this last estimate with (6.6), (6.7) and (6.8) and recalling that $\mathrm{spt}\,(\varphi) \subset \mathcal{B}_{1/16}$, we conclude

$$\left| \int_{\mathcal{B}_{1/16}} \bar\nabla\varphi(w) \cdot \bar\nabla f_j(w)\, dw \right| \leq C\lambda^{-1}\mathbf{E}\|\bar\nabla\varphi\|_\infty + C\|\bar\nabla\varphi\|_\infty \int_{\mathcal{B}_{1/16}} |Df|^2\,. \tag{6.9}$$

On the other hand observe that

$$\begin{aligned} \|\pi - T_x\Gamma_f\|^2 \geq |P_\pi(e_j) - P_{T_x\Gamma}(e_j)|^2 = |P_{T_x\Gamma_f}(e_j)|^2 &= \left| \sum_{l,m} \partial_{y_l} f_j(w) g^{lm}(w) v_m(w) \right|^2 \\ &= \left(\sum_{l,m} \partial_{y_l} f_j(w) g^{lm}(w) v_m(w) \right) \cdot \left(\sum_{l',m'} \partial_{y_{l'}} f_j(w) g^{l'm'}(w) v_{m'}(w) \right) \\ &= \sum_{l,m,l',m'} \partial_{y_l} f_j(w) \partial_{y_{l'}} f_j(w) g^{lm}(w) g^{l'm'}(w) g_{m'm}(w) \\ &= \sum_{l,m} \partial_{y_l} f_j(w) \partial_{y_m} f_j(w) g^{lm}(w) \\ &= |\nabla f_j(w)|^2 + \sum_{l,m} \partial_{y_l} f_j(w) \partial_{y_j} f_m(w) (g^{lm}(w) - \delta^{lm}) \\ &\geq |\nabla f_j(w)|^2 (1 - C|Df|^2)\,. \end{aligned} \tag{6.10}$$

Summing over j we conclude that, if the Lipschitz constant of f is smaller than a geometric constant, then

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

We now turn to (6.9): we first use use Proposition 5.1(iv) to conclude

$$\left|\int_{\mathcal{B}_{1/16}} \bar\nabla\varphi(w)\cdot\bar\nabla f_j(w)\,dw\right| \\ \leq C\lambda^{-1}\mathbf{E}\|\bar\nabla\varphi\|_\infty + C\|\bar\nabla\varphi\|_\infty\int_{P_\pi(G)}|Df(w)|^2\,dw\,.$$

Then we observe that $G\subset\Gamma\cap\Gamma_f$ and we thus infer from the previous discussion that $2\|\pi - T_x\Gamma\|^2\geq |Df(w)|^2$ for a.e. $w\in P_\pi(G)$. Hence

$$\left|\int_{\mathcal{B}_{1/16}} \bar\nabla\varphi(w)\cdot\bar\nabla f_j(w)\,dw\right| \\ \leq C\lambda^{-1}\mathbf{E}\|\bar\nabla\varphi\|_\infty + C\|\bar\nabla\varphi\|_\infty\int_G\|\pi - T_x\Gamma\|^2 d\mathcal{H}^k(x)\leq C\lambda^{-1}\mathbf{E}\|\bar\nabla\varphi\|\,.$$

ℓ has now been chosen smaller than a geometric constant which allows to justify the computations above. Therefore (cf. Proposition 5.1) λ is also fixed independently of β (and of all the other parameters which will enter in the rest of the proof). We then record the conclusion

$$\left|\int_{\mathcal{B}_{1/16}} \bar\nabla\varphi(w)\cdot\bar\nabla f_j(w)\,dw\right|\leq C\,\mathbf{E}\|\bar\nabla\varphi\|_\infty \qquad \forall\varphi\in C^1_c(\mathcal{B}_{1/16})\,. \tag{6.11}$$

Moreover, observe that

$$\int_{\mathcal{B}_{1/16}}|\bar\nabla f_j(w)|^2\,dw\leq C\ell^2\lambda^{-1}\mathbf{E} + \int_{P_\pi(G)}|\bar\nabla f_j(w)|^2\,dw\leq C\mathbf{E}\,. \tag{6.12}$$

Harmonic approximation. Fix a positive number ϑ, whose choice will be specified later and consider the ε_A given by Lemma 6.1 when choosing $\varrho=\vartheta$. We next let $j\in\{1,\ldots,N-k\}$ and define $\tilde f_j := c_0\mathbf{E}^{-\frac12}f_j$ where the constant c_0 is chosen so that, according to (6.12), $\int_{\mathcal{B}_{1/16}}|\bar\nabla\tilde f_j|^2\leq 1$. According to (6.11),

$$\left|\int_{\mathcal{B}_{1/16}} \bar\nabla\varphi(y)\cdot\bar\nabla\tilde f_j(y)\,dy\right|\overset{(6.11)}{\leq} C\,\mathbf{E}^{\frac12}\|\bar\nabla\varphi\|_\infty\,. \tag{6.13}$$

Assuming $\varepsilon_0\leq(\varepsilon_A/C)^2$ we can then apply Lemma 6.1 to conclude the existence of an harmonic function $\tilde u_j:\mathcal{B}_{1/16}\to\mathbb{R}$ with $\int|\bar\nabla\tilde u_j|^2\leq 1$ and $\int(\tilde f_j-\tilde u_j)^2\leq\vartheta$. Setting $u_j := c_0^{-1}\mathbf{E}^{\frac12}\tilde u_j$ we then conclude

$$\int_{\mathcal{B}_{1/16}}(f_j-u_j)^2\leq C\vartheta\,\mathbf{E}\,. \tag{6.14}$$

Observe, in particular, that if we define the map $u = (u_1, \ldots, u_{N-k})$ we then have

$$\|\bar{D}u\|_{L^2}^2 \leq C\mathbf{E}\,. \tag{6.15}$$

Height estimate. We denote by $L : \mathbb{R}^k \to \pi^\perp$ the map $L(y) = \sum_j (\bar{\nabla} u_j(0) \cdot y) e_j$, by x_0 the point $(0, u(0))$ and by $\bar{\pi}$ the plane $\{\sum_i y_i \xi_i + L(y) : y \in \mathbb{R}^k\}$. We next claim that

$$\eta^{-k-2} \int_{B_{4\eta}(x_0)} \mathrm{dist}(x - x_0, \bar{\pi})^2 \, d\mu_V(x) \leq C\eta^{-k-2}\vartheta\,\mathbf{E} + C\beta^2\eta^{-k-2}\,\mathbf{E} + C\eta^2\,\mathbf{E}\,. \tag{6.16}$$

We start by observing that, by the mean-value property for the harmonic functions u_j:

$$\mathrm{dist}(x_0, \pi) = |u(0)| \leq C\|u\|_{L^1} \leq C\|u - f\|_{L^2} + C\|f\|_{L^2} \leq C\vartheta^{\frac{1}{2}}\,\mathbf{E}^{\frac{1}{2}} + C\beta\,, \tag{6.17}$$

$$\|P_\pi^\perp - P_{\bar{\pi}}^\perp\| \leq C \sum_j |\bar{\nabla} u_j(0)| \leq C\|u\|_{L^1} \leq C\vartheta^{\frac{1}{2}}\,\mathbf{E}^{\frac{1}{2}} + C\beta\,. \tag{6.18}$$

For $x \in \Gamma \cap B_{1/64}$ we can therefore estimate

$$\mathrm{dist}(x - x_0, \bar{\pi}) = |P_{\bar{\pi}}^\perp(x - x_0)| \overset{(6.18)}{\leq} C\vartheta^{\frac{1}{2}}\,\mathbf{E}^{\frac{1}{2}} + C\beta + |P_\pi^\perp(x - x_0)| \overset{(6.17)}{\leq} C\vartheta^{\frac{1}{2}}\,\mathbf{E}^{\frac{1}{2}} + C\beta\,. \tag{6.19}$$

We thus conclude

$$\begin{aligned}\int_{B_{4\eta}(x_0)\setminus\Gamma_f} \mathrm{dist}(x - x_0, \bar{\pi})^2 \, d\mu_V(x) &= \int_{(\Gamma\setminus\Gamma_f)\cap B_{4\eta}(x_0)} \mathrm{dist}(x - x_0, \bar{\pi})^2 \, d\mathcal{H}^k(x) \\ &\leq C(\vartheta^{\frac{1}{2}}\,\mathbf{E}^{\frac{1}{2}} + \beta)^2\,\mathbf{E}\,. \end{aligned} \tag{6.20}$$

Observe finally that, if $x = (y, f(y)) \in \Gamma_f$, then $\mathrm{dist}(x - x_0, \bar{\pi}) \leq |f(y) - u(0) - L(y)|$. Thus, by (6.14) we conclude

$$\begin{aligned}\int_{\Gamma_f \cap B_{4\eta}(x_0)} \mathrm{dist}(x - x_0, \bar{\pi})^2 \, d\mu_V(x) &\leq \int_{\mathcal{B}_{4\eta}} |f(y) - u(0) - L(y)|^2 \, dy \\ &\overset{(6.14)}{\leq} C\vartheta\,\mathbf{E} + 2\int_{\mathcal{B}_{4\eta}} |u(y) - u(0) - L(y)|^2. \end{aligned}$$

Recalling that $L(y) = \bar{D}u(0) \cdot y$, from (6.3) and (6.15) we infer

$$\sup_{y \in \mathcal{B}_{4\eta}} |u(y) - u(0) - L(y)|^2 \leq C\eta^4 \|\bar{D}u\|_{L^2}^2 \leq C\eta^4\mathbf{E}$$

Summarizing, we have achieved

$$\int_{\Gamma_f \cap B_{4\eta}(x_0)} \mathrm{dist}(x - x_0, \bar{\pi})^2 \, d\mu_V(x) \leq C\vartheta\mathbf{E} + C\eta^{k+4}\mathbf{E}\,. \tag{6.21}$$

(6.20) and (6.21) together imply (6.16).

Tilt-excess inequality and conclusion. Impose that ϑ and β satisfy

$$C\vartheta^{\frac{1}{2}} \leq \frac{\eta}{2} \quad \text{and} \quad C\beta \leq \frac{\eta}{2}\,, \tag{6.22}$$

where C is the constant of (6.17). We then conclude that $B_\eta(0) \subset B_{2\eta}(x_0)$ and, combining(6.16) with Proposition 4.1, we infer

$$E(V,\bar\pi,0,\eta) \le 2^k E(V,\bar\pi,x_0,2\eta) \le C\eta^{-k-2}\vartheta\,\mathbf{E} + C\beta^2\eta^{-k-2}\,\mathbf{E} + C\eta^2\,\mathbf{E} + C\eta^2\mathbf{E}^2\,. \quad (6.23)$$

Recall that the constant C in the last inequality does not depend on the parameters η, β, ϑ and ε_0. We choose η first in such a way that $C\eta^2 = \frac{1}{8}$. We then choose β and ϑ such that $C\eta^{-k-2}\vartheta \le \frac{1}{8}$ and $C\eta^{-k-2}\beta^2 \le \frac{1}{8}$. It is obvious that these last choices are compatible with (6.22). Finally $C\eta\mathbf{E}^2 \le C\eta\mathbf{E} \le \frac{\mathbf{E}}{8}$. Plugging these inequalities in (6.23) we then infer

$$E(V,\bar\pi,0,\eta) \le \frac{1}{2}\mathbf{E} = \frac{1}{2}E(V,\pi,0,1)\,,$$

which is indeed the desired conclusion. □

7. Proof of Theorem 3.2

The rough idea of the proof is as follows. First by a scaling argument we reduce to the case $x_0 = 0$ and $r = 1$. The parameter ε is chosen so small that Theorem 3.4 can be applied to any point $x \in B_{1/2} \cap \Gamma$. The very nature of the latter theorem implies that we can iterate it to conclude a power-law decay for the excess. From Proposition 5.1 we then conclude that $\Gamma \cap B_{1/4}$ is contained in the graph of a Lipschitz function $f : \pi \to \pi^\perp$. Using the Monotonicity Formula we will show that, in a neighborhood of the origin Γ *coincides* with the graph of f. Finally we will show that the decay of the excess translates into a Morrey-type estimate for ∇f, concluding its $C^{1,\alpha}$ regularity.

Proof. W.l.o.g. we assume $x_0 = 0$ and $r = 1$.

Power-law decay of the excess. First let ε_0 be the constant of Theorem 3.4 and choose ε so small that Lemma 5.2 can be applied with $\delta = \varepsilon_0$. We thus know that $\mu(B_r(x)) \le (\omega_k + \varepsilon_0)r^k$ for every $r < \frac{1}{2}$ and any $x \in \Gamma \cap B_{1/4}$. Having fixed such an x we introduce the function

$$F(r) := E(r) + \Lambda\|H\|_\infty r := \min_\tau E(V,\tau,x,r) + \Lambda\|H\|_\infty r$$

where $\Lambda = 4\eta^{-k}$. If $F(r) < \varepsilon_0$, then by Theorem 3.4

- either $\|H\|_\infty r \le E(r)$ and so $F(\eta r) \le \frac{1}{2}E(r) + \Lambda\eta r\|H\|_\infty \le \frac{1}{2}F(r)$;
- or $E(r) \le \|H\|_\infty r$, and so

$$F(\eta r) \le (\eta^{-k}\Lambda^{-1} + \eta)\Lambda\|H\|_\infty r \le \frac{3}{4}\Lambda\|H\|_\infty r \le \frac{3}{4}F(r)\,.$$

Summarizing: $F(r) < \varepsilon_0 \;\Rightarrow\; F(\eta r) \le \frac{3}{4}F(r)$. In particular, $F(\eta r) < \varepsilon_0$ and we can iterate the conclusion with ηr in place of r. Observe next that

$$F(\tfrac{1}{2}) \le 2^{-k}E(V,\pi,0,1) + 2^{-1}\Lambda\|H\|_\infty \le (2^{-k} + 2^{-1}\Lambda)\varepsilon\,.$$

Thus, if ε is sufficiently small, we can start from $r = \frac{1}{2}$ and iterate the argument to infer $F(\eta^n\frac{1}{2}) \le C(\frac{3}{4})^n\varepsilon$ for any $n \in \mathbb{N}$. Given any $r < \frac{1}{2}$, we let $n = \lfloor \log_\eta(2r) \rfloor$ to conclude

$$E(r) \le \eta^{-k}E(\eta^n\tfrac{1}{2}) \le \eta^{-k}F(\eta^n\tfrac{1}{2}) \le C(\tfrac{3}{4})^n\varepsilon \le C(\tfrac{3}{4})^{\log_\eta(2r)-1}\varepsilon \le Cr^{2\alpha}\varepsilon\,, \quad (7.1)$$

where the constants C and $\alpha > 0$ depend only on the dimensions of the varifold and of the ambient euclidean space.

Inclusion in a Lipschitz graph. Fix $x \in B_{1/4}$. Set $\pi_0 = \pi$ and for $n \geq 1$ let π_n be a plane such that $E(V, \pi_n, x, 2^{-n}) = E(2^{-n})$. Recalling that $\mu_V(B_r(x)) \geq C^{-1} r^k$ for any $r < 1 - |x|$ (Monotonicity Formula!), we conclude

$$\begin{aligned} \|\pi_n - \pi_{n+1}\| &\leq \frac{1}{\mu_V(B_{2^{-n-1}}(x))} \int_{B_{2^{-n-1}}(x)} (\|\pi_n - T_y\Gamma\| + \|\pi_{n+1} - T_y\Gamma\|)\, d\mu_V(y) \\ &\leq C(E(V, \pi_n, x, 2^{-n}) + E(V, \pi_{n+1}, x, 2^{-n-1}))^{1/2} \leq C 2^{-n\alpha} \varepsilon^{1/2}\,. \end{aligned} \tag{7.2}$$

Summing this inequality from $n = 0$ to $n = j-1$ we conclude $\|\pi_j - \pi\| \leq C\varepsilon^{1/2}$, where C is a dimensional constant. Thus, $E(V, x, \pi, r) \leq C\varepsilon$ for any $x \in B_{1/4} \cap \Gamma$ and any $r \leq \frac{1}{2}$. Fix next a constant $\ell < \frac{1}{2}$ (whose choice will be specified in the next step) and let λ and ε_L be the corresponding constants given by Proposition 5.1. We assume ε to be smaller than ε_L but also so small that the set G of Proposition 5.1 contains $\Gamma \cap B_{1/4}$. We then conclude that $\Gamma \cap B_{1/4}$ is contained in the graph of a Lipschitz function $f : B_{1/4} \cap \pi \to \pi^\perp$ with Lipschitz constant smaller than ℓ.

Absence of "holes". Consider the ball $\mathcal{B}_1(0) \subset \mathbb{R}^k$, $x \in \partial \mathcal{B}_1(0)$ and let 2ϑ be the value $\mathcal{L}^k(\mathcal{B}_1(0) \setminus \mathcal{B}_1(x))$. Assume now that $D = P_\pi(\Gamma \cap B_{1/4})$ does not contain $\mathcal{B}_{1/16} := B_{1/16} \cap \pi$ and let $w \in \mathcal{B}_{1/16} \setminus D$. Define $r := \inf\{|w - z| : z \in D\}$. $r < \frac{1}{16}$ because $0 \in D$ and thus any infimizing sequence $\{z_n\}$ must be contained in $\mathcal{B}_{1/8}$. Up to extraction of a subsequence we can then assume that $z_n \to z \in \overline{\mathcal{B}}_{1/8}$. Recalling that the origin belongs to Γ we conclude $\|f\|_\infty \leq \ell$. If ℓ is sufficiently small we conclude that $x_n = (z_n, f(z_n)) \in B_{3/16}$ and thus x_n converges $x = (z, f(z)) \in \Gamma \in B_{3/16}$, where $z \in \overline{\mathcal{B}}_{1/8}$. Observe that $\Gamma \cap B_r(x)$ is contained in the graph of f because $r < \frac{1}{16}$. In particular, considering that $\mathcal{B}_r(w) \cap D = \emptyset$, using the area formula we can estimate

$$\mu_V(B_r(x)) \leq \int_{\mathcal{B}_r(z) \setminus \mathcal{B}_r(w)} Jf(u)\, du \leq (\omega_k - 2\vartheta)(1 + C\ell^2) r^k\,. \tag{7.3}$$

We now specify the choice of ℓ so that $(\omega_k - 2\vartheta)(1 + C\ell^2) = \omega_k - \vartheta$. Recall however that, by the Monotonicity Formula, $\mu_V(B_r(x)) \geq \omega_k r^k e^{-\|H\|_\infty r} \geq \omega_k r^k e^{-\varepsilon}$. Thus, choosing ε smaller than a specified dimensional constant, we reach a contradiction.

Morrey estimate for Df. So far we have concluded that Γ coincides with the graph of f on the intersection of the cylinder $\mathcal{B}_{1/8} \times \pi^\perp$ with the ball $B_{1/4}$. For each $z \in \mathcal{B}_{1/16}$ and every $r < \frac{1}{16}$ denote by $\pi_{z,r}$ the k-dimensional plane such that

$$E(V, \pi_{z,r}, (z, f(z)), r) = \min_\tau E(V, \tau, (z, f(z)), r) \leq C\varepsilon r^{2\alpha}\,.$$

Recalling that $E(V, \pi, (z, f(z)), r) \leq C\varepsilon$, we conclude that $\|\pi - \pi_{z,r}\| \leq C\varepsilon^{1/2}$. If ε is sufficiently small, $\pi_{z,r}$ is the graph of a linear map $A_{z,r} : \pi \to \pi^\perp$ with Hilbert-Schmidt norm smaller than 1.

Consider now two linear maps $A, B : \pi \to \pi^\perp$ with $|A|, |B| \leq C\ell$ (where $|\cdot|$ denotes the Hilbert-Schmidt norm), the k-dimensional planes τ_A and τ_B given by the corresponding graphs and P_A and P_B the orthogonal projections onto τ_A and τ_B. Observe that, if ℓ is smaller than a geometric constant, then $|P_A(v)| \le \frac{1}{2}|v|$ for any $v \in \pi^\perp$ (this follows easily because if $\ell \downarrow 0$, then $P_A \to P_\pi$). Fix an orthonormal base $e_1, \ldots, e_k$ of π and consider that

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

We then conclude that $|A(e_i) - B(e_i)| \le C\|\tau_A - \tau_B\|$ and thus, after summing over i, we infer $|A - B| \le C\|\tau_A - \tau_B\|$.

From this discussion we derive, for $r < \frac{1}{16}$,

$$\begin{aligned}\int_{\mathcal{B}_{r/2}(z)} |Df(y) - A_{z,r}|^2\,dy &\le \int_{\mathcal{B}_{r/2}(z)} |Df(y) - A_{z,r}|^2 Jf(y)\,dy \\ &\le Cr^k E(V, \pi_{z,r}, (z, f(z)), 2r) \le Cr^{k+2\alpha}\,. \end{aligned} \tag{7.4}$$

Denoting by $\overline{Df}_{z,r}$ the average of Df on $\mathcal{B}_r(z)$ we then conclude

$$\int_{\mathcal{B}_r(z)} |Df(y) - \overline{Df}_{z,r}|^2\,dy = \min_A \int_{\mathcal{B}_r(z)} |Df(y) - A|^2\,dy \le Cr^{k+2\alpha} \qquad \forall r < \tfrac{1}{2}\,. \tag{7.5}$$

Conclusion. It is well known that (7.5) implies the existence of $g \in C^{0,\alpha}$ which coincides with Df a.e. on $\mathcal{B}_{1/64}$. We briefly sketch here the argument (the reader may consult [8] for more details). First of all, argue as in (7.2) to conclude

$$|\overline{Df}_{x,2^{-k}} - \overline{Df}_{x,2^{-k-1}}| \le C2^{-k\alpha} \qquad \text{for all } k > 5 \text{ and } x \in \mathcal{B}_{1/32}. \tag{7.6}$$

Hence the sequence of continuous functions $x \mapsto \overline{Df}_{x,2^{-k}}$ converges uniformly to a continuous g with $g = Df$ a.e. on $\mathcal{B}_{1/32}$. Next, summing (7.6) over different scales we infer

$$|\overline{Df}_{x,r} - \overline{Df}_{x,\rho}| \le C(\max\{r,\rho\})^\alpha \qquad \text{for all } x \in \mathcal{B}_{1/32} \text{ and all } r, \rho < \tfrac{1}{32}. \tag{7.7}$$

Observe that, if $r = |x - y|$ and $x, y \in \mathcal{B}_{1/64}$, then

$$|\overline{Df}_{x,r} - \overline{Df}_{y,r}|^2 \le Cr^{-k} \int_{\mathcal{B}_r(x)} |Df - \overline{Df}_{x,r}|^2 + Cr^{-k} \int_{\mathcal{B}_r(y)} |Df - \overline{Df}_{y,r}|^2 \le Cr^{2\alpha}\,. \tag{7.8}$$

Combining (7.6) and (7.7) we conclude the existence of a dimensional constant such that

$$|\overline{Df}_{x,2^{-k}} - \overline{Df}_{y,2^{-k}}| \le C\big(\max\{2^{-k}, |x - y|\}\big)^\alpha\,.$$

Thus, fixing x and y and letting $k \uparrow \infty$ we conclude $|g(x) - g(y)| \le C|x - y|^\alpha$.

Finally, mollify f with a standard kernel to get $f * \varphi_\eta$. We have $D(f * \varphi_\eta) = g * \varphi_\eta$ and therefore $\|f * \varphi_\eta\|_{C^{1,\alpha}(\mathcal{B}_{1/64})}$ is bounded independently of η. Letting η go to 0 we then conclude that $f \in C^{1,\alpha}(\mathcal{B}_{1/64})$. $\square$

8. Further readings

To deepen the knowledge on the regularity theory for stationary varifolds the two references [2] and [13] are still the most fundamental sources. An open problem in the area (perhaps the hardest) is to understand whether for stationary varifolds (i.e. when the first variation *vanishes*) the regularity on a dense open set can be improved to an almost everywhere regularity without any assumption on the density. Such result is not possible for varifolds with bounded generalized mean curvature, as it is shown by the Example 8.1(2) of [2]: the question is, therefore, rather subtle.

Decay statements as Theorem 3.4 play a prominent role in the regularity theory of minimal surfaces and much more general results have been proved in the literature, especially in recent times: see for instance [3], [11] and [10] (the latter paper contains an exhaustive list of references). In many of these results one needs quite refined approximation theorems, which use multiple-valued functions to overcome the difficulty posed by points of high density surrounded by points of low density. In this context multiple valued functions were first introduced by Almgren in his Big regularity paper [1]; a more accessible reference for Almgren's theory is [6].

References

[1] F. J. Almgren Jr., Almgrens Big Regularity Paper. Volume 1 of World Scientific Monograph Series in Mathematics, World Scientific, River Edge, NJ, 2000.
[2] W. K. Allard, On the first variation of a varifold. *Ann. of Math (2)* **95** 417–491, 1972.
[3] K. A. Brakke, The motion of a surface by its mean curvature. Mathematical Notes, vol. 20. Princeton University Press, Princeton, N.J., 1978.
[4] E. De Giorgi, Frontiere orientate di misura minima, Seminario di Matematica della Scuola Normale Superiore di Pisa 1960-61. Editrice Tecnico Scientifica, Pisa, 1961.
[5] C. De Lellis, Rectifiable sets, densities and Tangent measures. Zürich lecture notes in advanced mathematics. EMS Publishing Hose, Zürich 2008.
[6] C. De Lellis & E. Spadaro, Almgren's Q-valued functions revisited. *Memoirs of the AMS* **211** no. 991m 2011.
[7] L. C. Evans, Partial Differential equations. Americal Mathematical Society, Providence 2010.
[8] D. Gilbarg & N. Trudinger, Elliptic partial differential equations of second order. *Classics in Mathematics.* Springer-Verlag. Berlin 2001.
[9] P. Mattila, Geometry of sets and measures in Euclidean spaces. Cambridge University Press. Cambridge 1995.
[10] U. Menne, Decay estimates for the quadratic tilt-excess of integral varifolds. *Arch. Rat. Mech. Anal* **204** 1–83 2012.
[11] R. Schätzle, Quadratic tilt-excess decay and strong maximum principle for varifolds. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **3** 171–231, 2004.
[12] R. Schoen and L. Simon, A new proof of the regularity theorem for rectifiable currents which minimize parametric elliptic functionals. *Indiana Univ. Math. J.* **31** 415–434 1982.
[13] L. Simon, Lectures on geometric measure theory. *Proceedings of the Centre for Mathematical Analysis, 3.* Australian National University. Canberra 1983.

Mathematik Institut der Universität Zürich
E-mail address: delellis@math.uzh.ch