This part was presented by Alessandro Pigati on March 19 and March 26, 2020, with a small leftover on April 2. The notes are written by Alessandro Pigati. If you have any question or find any mistake, please feel free to contact the author.

## 4. First variation of area and monotonicity. Regularity of almost flat stationary sets: tilt-excess inequality

The presentation, which is essentially the same as in Maggi's book [5] or in Allard's paper [1], follows quite closely the one in the notes [3] by Camillo De Lellis. This choice is motivated by the fact that one of the next speakers will present the adaptations for general stationary varifolds, which is the setting of [3]. The presentation given here contains a few simplifications, due to the simpler setting where we are (reduced boundaries of finite perimeter sets), and the extra addition of Corollary 4.10, which is not known in the general situation of [3].

### 4.1. First variation of area and monotonicity.

**Definition 4.1.** Let $\Omega \subseteq \mathbb{R}^{n+1}$ be an open set and $E \subseteq \Omega$ a set of finite perimeter in $\Omega$. We say that $E$ is a *minimizer* for the perimeter in $\Omega$ if its perimeter satisfies

$$\mathrm{Per}(E, \Omega) \leq \mathrm{Per}(E', \Omega)$$

whenever $E'$ is a set of finite perimeter (in $\Omega$) with $E \Delta E' \subset\subset \Omega$.

**Remark 4.2.** The same definition applies if $E$ is originally defined on a superset $\Omega' \supseteq \Omega$, in which case we ask that $E \cap \Omega$ is a minimizer in $\Omega$.

**Definition 4.3.** Let $E$ be a set of finite perimeter (in $\Omega$ or in a bigger set). We say that $E$ is *stationary* for the perimeter if

$$\frac{d}{dt} \mathrm{Per}(\Phi_t(E), \Omega)|_{t=0} = 0$$

for all "perturbations of the identity" $(\Phi_t)_{t \in (-\varepsilon, \varepsilon)}$ ($\varepsilon$ arbitrary), namely maps $\Phi_t : \Omega \to \Omega$ such that

- $(t, x) \mapsto \Phi_t(x)$ is smooth on $(-\varepsilon, \varepsilon) \times \Omega$,
- $\Phi_t$ is a diffeomorphism for all $t \in (-\varepsilon, \varepsilon)$,
- $\Phi_0 = \mathrm{id}$,
- $\Phi_t(x) = x$ for all $x \in \Omega \setminus K$ and all $t$, for a suitable compact set $K \subset \Omega$.

We will see in a couple of talks that the above derivative always exists and equals

$$\frac{d}{dt} \mathrm{Per}(\Phi_t(E), \Omega)\Big|_{t=0} = \int_{\partial^* E} \mathrm{div}_{T_y \partial^* E} S \, d\mathcal{H}^n(y),$$

where the vector field $S$ is the speed of the deformation at $t = 0$, i.e., $S(x) := \frac{d}{dt} \Phi_t(x)\Big|_{t=0}$, and $\mathrm{div}_{T_y \partial^* E} S$ is its divergence along the approximate tangent $n$-plane $T_y \partial^* E$, namely

$$\mathrm{div}_{T_y \partial^* E} S = \sum_{j=1}^{n} \langle DS(y)[\xi_j], \xi_j \rangle$$

for an arbitrary orthonormal basis $\{\xi_j\}_1^n$ of $T_y\partial^*E$ (the result is independent of the choice).

Given a vector field $X \in C_c^\infty(\Omega, \mathbb{R}^{n+1})$, we can take $\Phi_t$ to be the flow of $X$ at time $t$, in which case $S = X$. We deduce that stationarity is equivalent to ask

$$\int_{\partial^* E} \operatorname{div}_{T_y\partial^* E} X \, d\mathcal{H}^n(y) = 0$$

for all vector fields $X \in C_c^\infty(\Omega, \mathbb{R}^{n+1})$.

**Remark 4.4.** A minimizer is stationary. Indeed, we can choose $E' := \Phi_t(E)$ in the definition of minimality and deduce $\operatorname{Per}(E, \Omega) \le \operatorname{Per}(\Phi_t(E), \Omega)$. Since equality holds for $t = 0$, the function $t \mapsto \operatorname{Per}(\Phi_t(E), \Omega)$ has a minimum for $t = 0$ and hence the derivative vanishes there.

In the sequel, we will denote $\mu := \mathcal{H}^n \llcorner \partial^* E$ for convenience. Note that, in the previous talk, $\mu$ was used instead to denote the vector-valued Gauss–Green measure of $E$. In the present notation, the latter equals $\nu\,\mu$.

In the same forthcoming talk, we will also see that a clever choice of $X$ will imply the following formula, called *monotonicity formula*: for all balls $B_r(x) \subseteq B_s(x) \subseteq \Omega$

$$\int_{B_s(x)\setminus B_r(x)} \frac{|T_y^\perp \partial^* E(y)|^2}{|y|^{n+2}} \, d\mu(y) \le \frac{\mu(B_s(x))}{s^n} - \frac{\mu(B_r(x))}{r^n} \tag{1}$$

where $T_y^\perp \partial^* E$ denotes the line orthogonal to the approximate tangent $n$-plane $T_y\partial^* E$, identified with the orthogonal projection map $\mathbb{R}^{n+1} \to T_y^\perp \partial^* E$ (so that $|T_y^\perp \partial^* E(y)|^2 = \langle \nu(y), y\rangle^2$). The name comes from the fact that it readily implies that the function

$$r \mapsto \frac{\mu(B_r(x))}{r^n}, \quad 0 < r < \operatorname{dist}(x, \partial\Omega),$$

is increasing. We deduce that the *density*

$$\theta(x) := \lim_{r\to 0} \frac{\mu(B_r(x))}{\omega_n r^n}$$

exists and is finite for all $x \in \Omega$. The factor $\omega_n$ normalizes this limit to be 1 when $E$ is a half-space and $x$ belongs to its boundary, which is an $n$-plane. Note that, by the blow-up analysis of the previous talk, we have $\theta(x) = 1$ for all $x \in \partial^* E$.

It is also easy to see that $\theta$ is upper-semicontinuous, i.e., $\theta(x) \ge \limsup_{j\to\infty} \theta(x_j)$ whenever $x_j \to x$ (hint: use the inclusion $B_r(x) \supseteq B_{r-|x_j - x|}(x_j)$ for any fixed $r > 0$). It follows that

$$\theta = 1 \text{ on } \partial^* E, \quad \theta \ge 1 \text{ on } \overline{\partial^* E}, \quad \theta = 0 \text{ on } \Omega \setminus \overline{\partial^* E}, \tag{2}$$

where $\overline{\partial^* E}$ is the relative closure of $\partial^* E$ in $\Omega$. The last assertion is immediate from the fact that for $x \notin \overline{\partial^* E}$ we have $\partial^* E \cap B_r(x) = \emptyset$, for all radii $r$ small enough, and thus $\mu(B_r(x)) = 0$.

**Remark 4.5.** For a minimizer, the fact that $\frac{\mu(B_r(x))}{r^n}$ increases can be derived informally as follows: first, note that

$$\frac{d}{dr} \operatorname{Per}(E, B_r(x)) \ge \operatorname{Per}(E \cap \partial B_r(x)),$$

where the right-hand side is the ($(n-1)$-dimensional) perimeter of $E \cap \partial B_r(x)$ as a subset of the sphere $\partial B_r(x)$, with equality achieved if $\partial E$ meets $\partial B_r(x)$ orthogonally.

Defining $E'$ to equal $E$ outside $B_r(x)$ and to equal the cone of $E \cap \partial B_r(x)$ (with vertex $x$) inside, we get

$$\operatorname{Per}(E', \Omega) = \operatorname{Per}(E, \Omega) - \operatorname{Per}(E, B_r(x)) + \frac{r}{n} \operatorname{Per}(E \cap \partial B_r(x)).$$

Using this $E'$ in the definition of minimality, we deduce

$$\operatorname{Per}(E, B_r(x)) \leq \frac{r}{n} \operatorname{Per}(E \cap \partial B_r(x)) \leq \frac{r}{n} \frac{d}{dr} \operatorname{Per}(E, B_r(x)),$$

which implies that $r^{-n} \operatorname{Per}(E, B_r(x))$ has nonnegative derivative.

4.2. **Regularity of almost flat stationary sets: strategy and statement.** The goal of these two talks is to show that if $E \cap B_r(x)$ is approximately a half-space (meaning that it is approximately $\{y \in B_r(x) : \langle \nu, y - x \rangle > 0\}$ for some unit vector $\nu$) then $E$ is smooth on a smaller concentric ball.

This is similar to many results for nonlinear elliptic PDEs, where one assumes some kind of "smallness" or "flatness" of the solution, in order to make the equation resemble a linear one (the "first order expansion" of the original PDE) plus small higher order terms, thus "taming" the nonlinearity, and one then proves that the solution is smooth on a smaller ball, as it would happen for the linear (elliptic) PDE.

The strategy will be to show that the "flatness" of $E$ on a ball $B_s(y)$, i.e., its closeness to a half-space, improves on a smaller ball in a quantitative way, namely it halves on a ball $B_{\eta s}(y)$. This is the bulk of the work and the proof consists of three steps:

- showing that most of $\partial^* E \cap B_{s/100}(y)$ is contained in the graph of a Lipschitz map $f : \mathbb{R}^n \to \mathbb{R}$ with small Lipschitz constant, up to rotations;
- showing that $f$ is close to being harmonic (on a ball) in a certain weak sense and deducing that it is actually close to a harmonic function $u$ in $L^2$;
- exploiting the fact that harmonic functions do enjoy the desired "improvement of flatness"; the technical difficulty is to transfer this from $u$ to the original boundary $\partial^* E$.

The fact that $f$ is approximately harmonic is expected, since its graph should be essentially a critical point for the area and, if one expands the area functional for graphs, the first term in the expansion is the Dirichlet energy $\int |\nabla f|^2$, whose critical points are harmonic functions.

One can measure flatness on a ball $B_r(x)$ with the $L^1$-distance of $1_{\partial^* E \cap B_r(x)}$ from a half-space, or by averaging the distance of the approximate tangent plane $T_y \partial^* E$ from a reference $n$-plane $\pi$. The second choice turns out to be much more convenient (although one can show a posteriori that they both work).

**Remark 4.6.** There are simple examples showing that one cannot hope for regularity, nor for improvement of flatness, just assuming flatness (in either version): see the picture in the video. In these examples, the culprit is the fact that $\partial^* E \cap B_r(x)$ has a much bigger area than a hyperplane through $x$ (in $B_r(x)$), producing several layers which are not individually boundaries of stationary sets.

In view of this remark, we need two assumptions:

- closeness to a half-space, or more conveniently closeness of the approximate tangent planes to a fixed $n$-plane in average,
- area $\mu(B_r(x)) = \mathcal{H}^n(\partial^* E \cap B_r(x))$ close to $\omega_n r^n$, the one we would have for a hyperplane.

Before giving precise statements, let us define a key quantity, which will make the first assumption precise.

**Definition 4.7.** Given an $n$-plane $\pi \subset \mathbb{R}^{n+1}$ through the origin, we define the *excess* of $E$ on $B_r(x)$, with respect to $\pi$, to be

$$\mathrm{Exc}(E, \pi, x, r) := r^{-n} \int_{B_r(x)} \|T_y \partial^* E - \pi\|^2 \, d\mu(y),$$

where we identify an $n$-plane with the orthogonal projection onto it, which is a linear map $\mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$, and we use the Hilbert–Schmidt norm for linear maps.

We will often write $\mathrm{Exc}(\pi, x, r)$ for simplicity. Note that this quantity is dimensionless, i.e., invariant under dilations, in view of the factor $r^{-n}$. Since we expect $\mu(B_r(x))$ to behave like $r^n$, at least when $x \in \partial^* E$, putting the factor $r^{-n}$ is informally like averaging in the measure $\mu$.

The main result is the following.

**Theorem 4.8.** *Assume $E$ is stationary in $B_r(x)$, with $\theta(x) \geq 1$. There exists a dimensional constant $\varepsilon(n)$ such that if*

$$\mathrm{Exc}(\pi, x, r) < \varepsilon, \quad \mu(B_r(x)) < (\omega_n + \varepsilon) r^n$$

*then $\partial^* E$ is a $C^{1,\gamma}$ embedded submanifold in $B_{r/1000}(x)$, for some $\gamma = \gamma(n) \in (0, 1)$.*

As we said, the bulk of the work lies in showing the improvement of flatness, namely the following.

**Theorem 4.9.** *Under the same assumption, for a possibly different $\varepsilon(n)$, we have*

$$\mathrm{Exc}(\tilde{\pi}, x, \eta r) \leq \frac{1}{2} \mathrm{Exc}(\pi, x, r)$$

*for a new suitable plane $\tilde{\pi}$ (depending of course on all data) and a factor $0 < \eta < 1$ depending only on $n$.*

Note that the nature of Theorem 4.9 allows to iterate it infinitely many times, provided we can apply it once: indeed, by monotonicity

$$\frac{\mu(B_{\eta r}(x))}{(\eta r)^n} \leq \frac{\mu(B_r(x))}{r^n} < \omega_n + \varepsilon$$

and $\mathrm{Exc}(\pi, x, \eta r) < \varepsilon/2 < \varepsilon$ by the conclusion of the theorem, so we can apply Theorem 4.9 again on $B_{\eta r}(x)$, and so on, finding planes relative to which the excess decays exponentially on smaller and smaller balls.

In proving Theorem 4.8 we will apply Theorem 4.9 at all points in a small ball, rather than just at $x$.

We record the following corollary of Theorem 4.8.

**Corollary 4.10.** *There exists a relatively closed set $S \subset \Omega$, with $\mathcal{H}^n(S) = 0$, such that $\partial^* E$ is a smooth embedded $n$-submanifold on $\Omega \setminus S$. In particular, $\mathcal{H}^n$-a.e. point of $\partial^* E$ lies in this smooth manifold.*

*Proof.* First of all, in view of (2), we have $\theta \geq 1$ on $\overline{\partial^* E}$ and hence

$$\mathcal{H}^n(\overline{\partial^* E} \setminus \partial^* E) = 0.$$

This consequence can be proved using standard covering arguments: see, e.g., [2, Theorem 2.56], which gives $\mu \geq \mathcal{H}^n \llcorner \overline{\partial^* E}$ and hence $0 = \mu(\overline{\partial^* E} \setminus \partial^* E) \geq \mathcal{H}^n(\overline{\partial^* E} \setminus \partial^* E)$.

Now, given $x \in \partial^* E$, the assumptions of Theorem 4.8 apply for a suitably small $r > 0$ depending on $x$: indeed, the blow-up analysis of the previous talk (see Theorem 3.10 in the notes) shows that $\frac{\mu(B_r(x))}{r^n} \to \omega_n$ as $r \to 0$, while, letting $\pi = \nu^\perp$ for some unit vector $\nu$ and calling $\nu(y)$ the measure-theoretic outer unit normal,

$$\begin{aligned}
\lim_{r \to 0} \operatorname{Exc}(\pi, x, r) &= \lim_{r \to 0} \frac{\omega_n}{\mu(B_r(x))} \int_{B_r(x)} \|T_y \partial^* E - \pi\|^2 \, d\mu(y) \\
&= \lim_{r \to 0} \frac{\omega_n}{\mu(B_r(x))} \int_{B_r(x)} 2(1 - \langle \nu(y), \nu \rangle^2) \, d\mu(y)
\end{aligned}$$

vanishes if we choose $\nu := \nu(x)$ as in the definition of reduced boundary, since the last integral is $\leq 4 \int_{B_r(x)} (1 - \langle \nu(y), \nu \rangle) \, d\mu(y) = \Big(4 - 4 \frac{\mu_E(B_r(x))}{|\mu_E|(B_r(x))} \cdot \nu(x)\Big) \mu(B_r(x))$ and this is infinitesimal with respect to $\mu(B_r(x))$.[1] The last equality above follows from the identity

$$\|T_y \partial^* E - \pi\| = \|(\mathrm{id} - \pi) - (\mathrm{id} - T_y \partial^* E)\| = \|\pi^\perp - T_y^\perp \partial^* E\|$$

and the fact that, writing $\pi^\perp = \nu \otimes \nu$, $T_y^\perp \partial^* E = \nu(y) \otimes \nu(y)$, we have

$$\|\nu \otimes \nu - \nu(y) \otimes \nu(y)\|^2 = \operatorname{tr}(\nu \otimes \nu - \nu(y) \otimes \nu(y)) = 2 - 2\langle \nu, \nu(y) \rangle^2.$$

So Theorem 4.8 gives that $\partial^* E$ is a $C^{1,\gamma}$ submanifold near $x$. Calling $S$ the complement of the biggest open subset $U \subseteq \Omega$ where $\partial^* E$ is a $C^{1,\gamma}$ submanifold, we get $\partial^* E \subseteq U$ and, trivially, $\Omega \setminus \overline{\partial^* E} \subseteq U$. So

$$S \subseteq \overline{\partial^* E} \setminus \partial^* E$$

and we deduce $\mathcal{H}^n(S) = 0$. To conclude, $C^{1,\gamma}$ can be upgraded to $C^\infty$ using standard elliptic regularity, writing $\partial^* E$ locally in $U$ as a graph of a function: stationarity for the area gives a certain nonlinear elliptic PDE which can be seen as linear with Hölder coefficients, and Schauder theory gives that our function is $C^\infty$. We omit the details. ☐

4.3. **Tilt-excess inequality.** We now show an inequality which is analogous to Cacciopoli's inequality for second order linear elliptic PDEs, both essentially giving a $W^{1,2}$-control on a ball in terms of an $L^2$-control on a bigger ball. Also the proof is substantially the same. It will be needed only at the end of the proof of Theorem 4.9, but we present it here in order to get acquainted with the stationarity condition.

[1] Here $\mu_E$ is the Gauss–Green measure of $E$, in the notation of the previous talk. Here and in the sequel we keep denoting instead $\mu = \mathcal{H}^n \llcorner \partial^* E$, so that $\mu = |\mu_E|$.

**Proposition 4.11** (tilt-excess inequality). *If $E$ is stationary on $B_r(x)$, then*

$$\operatorname{Exc}(\pi, x, r/2) \le C(n) r^{-n-2} \int_{B_r(x)} \operatorname{dist}(y-x,\pi)^2 \, d\mu(y).$$

We call the quantity on the right-hand side the *tilt* of $E$ on $B_r(x)$ with respect to $\pi$. Note that both the excess and the tilt have a normalization which makes them scale invariant (they are "dimensionless").

*Proof.* By translation and dilation, we can assume $x = 0$, $r = 1$. Rotating the space, we can also assume $\pi = \operatorname{span}\{e_1, \dots, e_n\}$.

Let $\varphi$ be a cut-off function, namely we ask $\varphi \in C_c^\infty(B_1)$ and $\varphi = 1$ on $B_{1/2}$. Define the vector fields $Y(y) := y_{n+1} e_{n+1}$ and $X(y) := \varphi^2(y) Y(y)$.

Fixing $y$ where $T_y\partial^* E$ exists and calling $\{\xi_j\}_1^n$ an orthonormal basis of it, we compute

$$\operatorname{div}_{T_y\partial^*E} Y = \sum_j (\xi_j)_{n+1} \langle e_{n+1}, \xi_j\rangle = \sum_j \langle e_{n+1}, \xi_j\rangle^2. \tag{3}$$

Now

$$\operatorname{div}_{T_y\partial^*E} X = \varphi^2(y) \operatorname{div}_{T_y\partial^*E} Y + \sum_j \langle \nabla(\varphi^2), \xi_j\rangle y_{n+1} \langle e_{n+1}, \xi_j\rangle.$$

Since $\int_{B_1} \operatorname{div}_{T_y\partial^*E} X \, d\mu(y) = 0$ by stationarity, we deduce

$$\int_{B_1} \varphi^2(y) \operatorname{div}_{T_y\partial^*E} Y = -2 \int_{B_1} \sum_j \varphi(y) \langle e_{n+1}, \xi_j\rangle y_{n+1} \langle \nabla\varphi, \xi_j\rangle.$$

Applying Young's inequality $2ab \le \frac{a^2}{2} + 2b^2$ and recalling (3), we deduce

$$\int_{B_1} \varphi^2(y) \operatorname{div}_{T_y\partial^*E} Y \le \frac{1}{2} \int_{B_1} \varphi^2(y) \operatorname{div}_{T_y\partial^*E} Y + C(n) \int_{B_1} y_{n+1}^2 \, d\mu(y),$$

which can be rewritten as

$$\int_{B_1} \varphi^2(y) \operatorname{div}_{T_y\partial^*E} Y \le C(n) \int_{B_1} \operatorname{dist}(y-x,\pi)^2 \, d\mu(y).$$

Finally, since $\pi = \operatorname{id} - \pi^\perp = \operatorname{id} - e_{n+1} \otimes e_{n+1}$ (and similarly for $T_y\partial^*E$ with $e_{n+1}$ replaced by $\nu(y)$), we compute

$$\begin{aligned}
\|\pi - T_y\partial^*E\|^2 &= \|e_{n+1} \otimes e_{n+1} - \nu(y) \otimes \nu(y)\|^2 \\
&= 2(1 - \langle e_{n+1}, \nu(y)\rangle^2) \\
&= 2 \sum_j \langle e_{n+1}, \xi_j\rangle^2 \\
&= 2 \operatorname{div}_{T_y\partial^*E} Y
\end{aligned}$$

where we used the identity $\|A\|^2 = \operatorname{tr}(A^2)$ for symmetric matrices. Using the fact that $\varphi = 1$ on $B_{1/2}$, we deduce

$$\operatorname{Exc}(\pi, 0, 1/2) \le 2^n \int_{B_1} \varphi^2(y) \|\pi - T_y\partial^*E\|^2 \, d\mu(y) \le C(n) \int_{B_1} \operatorname{dist}(y-x,\pi)^2 \, d\mu(y).$$

For future use, we also record the inequality

$$J_{T_y\partial^* E}\pi \geq 1 - C(n)\|\pi - T_y\partial^* E\|^2, \tag{4}$$

where $J_{T_y\partial^* E}\pi$ is the Jacobian of the linear map $\pi|_{T_y\partial^* E}$. Indeed, this quantity equals the square root of the determinant of $A$, where $A_{ij} = \langle \pi(\xi_i), \pi(\xi_j)\rangle$. Writing $\pi = \mathrm{id} - e_{n+1} \otimes e_{n+1}$, we find

$$A_{ij} = \delta_{ij} - \langle e_{n+1}, \xi_i\rangle\langle e_{n+1}, \xi_j\rangle$$

and the second term is bounded by $\sum_j \langle e_{n+1}, \xi_j\rangle^2 = \frac{1}{2}\|\pi - T_y\partial^* E\|^2$. Estimate (4) follows. □

## 5. Regularity of almost flat stationary sets: Lipschitz and harmonic approximation

5.1. **Lipschitz approximation.** In this section we implement the first step in our program. The goal is the following proposition, whose statement is a bit technical.

**Proposition 5.1** (Lipschitz approximation). *Assume for simplicity $x = 0$ and fix parameters $\ell, \beta \geq 1$. There exists a constant $\varepsilon_L(\ell, \beta)$ such that, if $E$ satisfies the assumptions of Theorem 4.8 with $\varepsilon_L$ in place of $\varepsilon$, then there exists an $\ell$-Lipschitz map $f : \pi \to \pi^\perp$ with the following properties:*

- *$G := \{y \in \partial^* E \cap B_{r/100}$ s.t. $\mathrm{Exc}(\pi, y, s) < \lambda\ \forall s < \frac{r}{10}\} \subseteq \Gamma_f$, where $\lambda = \lambda(\ell)$ is a suitable constant,*
- *$\partial^* E \cap B_{r/100}$ and $\Gamma_f$ are included in the $(\beta r)$-neighborhood of $\pi$,*
- *$\mathcal{H}^n((\partial^* E \setminus G) \cap B_{r/100}) + \mathcal{H}^n((\Gamma_f \setminus G) \cap B_{r/100}) \leq \frac{C(n)}{\lambda}\mathrm{Exc}(\pi, x, r) r^n$.*

Here $\Gamma_f := \{z + f(z) \mid z \in \pi\}$ is the graph of $f$ (viewing $\pi, \pi^\perp$ as subsets of $\mathbb{R}^{n+1}$). Recall that the $\eta$-neighborhood of a set $S$ is $\bigcup_{p\in S} B_\eta(p)$, or equivalently the set of points whose distance from $S$ is $< \eta$.

The result will follow with little work from the following lemma, whose statement is somewhat cleaner.

**Lemma 5.2** ("height" lemma). *Given $\delta > 0$ there exists $\varepsilon_H(\delta)$ with the following property: if $E$ satisfies the same assumptions as Theorem 4.8 with $\varepsilon_H$ in place of $\varepsilon$, then*

- *$\partial^* E \cap B_{r/2}(x) \subseteq (\delta r)$-neighborhood of $x + \pi$,*
- *$\mu(B_s(y)) < (\omega_n + \delta)s^n$ for all $y \in B_{r/2}(x)$ and all $s \leq \frac{r}{4}$.*

The word "height" refers to the first conclusion, which is the most important one and says that the "height" of $\partial^* E$ over $x + \pi$ is $\leq \delta r$ (on the ball $B_{r/2}(x)$).

*Proof.* Wlog $x = 0$, $r = 1$. We argue by contradiction. Note that we want the constant $\varepsilon_H$ to be independent of $E$, hence if the statement is not true we can find stationary sets $E_j$, satisfying

$$\theta_j(0) := \lim_{s\to 0}\frac{\mu_j(B_s(0))}{\omega_n s^n} \geq 1, \quad \mathrm{Exc}(E_j, \pi, 0, 1) < \varepsilon_j, \quad \mu_j(B_1) < (\omega_n + \varepsilon_j)$$

for a sequence $\varepsilon_j \to 0$ (here $\mu_j := \mathcal{H}^n \llcorner \partial^* E_j$), but such that one of the two conclusions is false.

Up to subsequences, we assume $\mu_j \rightharpoonup \mu_\infty$ as Radon measures. We claim that

$$\mu_\infty = \mathcal{H}^n \llcorner (\pi \cap B_1). \tag{5}$$

Once this claim is proved, it is easy to reach a contradiction. Indeed, assuming that the first conclusion fails for infinitely many $j$'s, up to passing to the corresponding subsequence there exist points $y_j \in B_{1/2}$ with $\operatorname{dist}(y_j, \pi) \geq \delta$ and $y_j \in \partial^* E_j$. Hence, by monotonicity,

$$\mu_j(B_{\delta/2}(y_j)) \geq \omega_n (\delta/2)^n \theta_j(y_j) = \omega_n(\delta/2)^n \tag{6}$$

since $\theta_j = 1$ on $\partial^* E_j$ (wlog $\delta < 1$). Up to further subsequences, we can assume $y_j \to y_\infty$.

Note that $y_\infty \in \overline{B}_{1/2}$ and $\operatorname{dist}(y_\infty, \pi) \geq \delta$. Now (6) implies

$$\mu_\infty(\overline{B}_{\delta/2}(y_\infty)) \geq \omega_n(\delta/2)^n,$$

while (5) trivially gives $\mu_\infty(\overline{B}_{\delta/2}(y_\infty)) = 0$ since $\overline{B}_{\delta/2}(y_\infty) \cap \pi = \emptyset$.

Similarly we reach a contradiction if the second conclusion fails for infinitely many $j$'s, since in the same way we can assume (up to subsequences) that there are sets $E_j$, points $y_j \in B_{1/2}$ and radii $s_j \leq \frac{1}{4}$ such that

$$\mu_j(B_{s_j}(y_j)) \geq (\omega_n + \delta) s_j^n,$$

so that by monotonicity

$$\mu_j(B_{1/4}(y_j)) \geq (\omega_n + \delta)(1/4)^n.$$

Assuming as before that $y_j \to y_\infty$, we deduce $\mu_\infty(\overline{B}_{1/4}(y_\infty)) \geq (\omega_n + \delta)(1/4)^n$, but it is easy to see that

$$\mu_\infty(\overline{B}_{1/4}(y_\infty)) = \mathcal{H}^n(\pi \cap \overline{B}_{1/4}(y_\infty)) \leq \omega_n(1/4)^n.$$

We are left to prove the claim. Recall (1), which gives

$$\int_{B_1 \setminus B_\sigma} \frac{|T_y^\perp \partial^* E_j(y)|^2}{|y|^{n+2}} \, d\mu_j(y) \leq \frac{\mu_j(B_1)}{1^n} - \frac{\mu_j(B_\sigma)}{\sigma^n}$$

for any fixed $0 < \sigma < 1$. By monotonicity again, the second ratio is at least $\omega_n \theta_j(0) \geq \omega_n$, hence the right-hand side is bounded by $(\omega_n + \varepsilon_j) - \omega_n$. We deduce that the left-hand side is infinitesimal as $j \to \infty$. On the other hand,

$$\begin{aligned}\int_{B_1 \setminus B_\sigma} |\pi^\perp y|^2 \, d\mu_j(y) &\leq 2 \int_{B_1 \setminus B_\sigma} \|\pi^\perp - T_y^\perp \partial^* E_j\|^2 \, d\mu_j(y) + 2 \int_{B_1 \setminus B_\sigma} |T_y^\perp \partial^* E_j(y)|^2 \, d\mu_j(y) \\ &\leq 2 \operatorname{Exc}(E_j, \pi, 0, 1) + o(1) \to 0\end{aligned}$$

since $\|\pi^\perp - T_y^\perp \partial^* E_j\| = \|\pi - T_y \partial^* E_j\|$ (being $\pi^\perp = \operatorname{id} - \pi$ and similarly for $T_y \partial^* E_j$). In the limit we deduce

$$\int_{B_1 \setminus \overline{B}_\sigma} |\pi^\perp y|^2 \, d\mu_\infty(y) \leq \liminf_{j \to \infty} \int_{B_1 \setminus \overline{B}_\sigma} |\pi^\perp y|^2 \, d\mu_j(y) = 0,$$

hence $|\pi^\perp y| = 0$ for $\mu_\infty$-a.e. $y \in B_1 \setminus \overline{B}_\sigma$. Since $\sigma$ was arbitrary, we get $\operatorname{supp}(\mu_\infty) \subseteq \pi$. Finally, note that by stationarity

$$\begin{aligned}&\Big|\int_{B_1} \operatorname{div}_\pi X \, d\mu_j(y)\Big| \\ &= \Big|\int_{B_1} (\operatorname{div}_\pi X - \operatorname{div}_{T_y\partial^* E_j} X)\, d\mu_j(y)\Big| \\ &\le C(n)\|DX\|_{L^\infty} \operatorname{Exc}(E_j, \pi, 0, 1)^{1/2},\end{aligned}$$

for all $X \in C_c^\infty(B_1, \mathbb{R}^{n+1})$ (the last inequality is left as an exercise). In the limit we get $\int_{B_1} \operatorname{div}_\pi X \, d\mu_\infty = 0$. Assuming wlog $\pi = \operatorname{span}\{e_1, \dots, e_n\}$, for any $\varphi \in C_c^\infty(B_1^n)$ we have

$$\int_{B_1} \operatorname{div}_\pi(\varphi(y')e_k)\, d\mu_\infty(y) = 0 \tag{7}$$

for all $k = 1, \dots, n$, where we denote $y = (y', y_{n+1})$. Indeed, although this vector field is not supported in $B_1$, we have $\operatorname{supp}(\varphi) \subseteq B^n_{1-\eta}$ for some $\eta > 0$ and we can find a cut-off function $\psi \in C_c^\infty(B^1_\eta)$ such that $\psi = 1$ on $B^1_{\eta/2}$, say. Then the vector field $\varphi(y')\psi(y_{n+1})e_k$ is admissible in the definition of stationarity, namely it is supported in $B_1$, and (7) follows since inserting $\psi(y_{n+1})$ or not makes no difference in the integral (we already know $\operatorname{supp}(\mu_\infty) \subseteq \pi$).

Viewing $\mu_\infty$ as a measure in $\mathbb{R}^n$, we deduce

$$\int_{B_1^n} \frac{\partial \varphi}{\partial x_k}\, d\mu_\infty = 0.$$

Hence, $\mu_\infty$ is a constant multiple of $\mathcal{L}^n \llcorner B_1^n$ (exercise; hint: regularize $\mu_\infty$ by convolution), say $\mu_\infty = \alpha \mathcal{L}^n \llcorner B_1^n$. But

$$\mu_\infty(B_1) \le \liminf_{j\to\infty} \mu_j(B_1) \le \omega_n,$$

hence $\alpha \le 1$, and

$$\mu_\infty(\overline{B}_{1/2}) \ge \limsup_{j\to\infty} \mu_j(\overline{B}_{1/2}) \ge \limsup_{j\to\infty} \omega_n (1/2)^n \theta_j(0) \ge \omega_n (1/2)^n,$$

thus $\alpha \ge 1$. Our claim (5) follows. □

*Proof of Proposition 5.1.* Wlog $r = 1$. We choose:

- $\lambda := \varepsilon_H(\ell/10)$, i.e., we apply Lemma 5.2 with $\delta := \ell/10$ and we let $\lambda$ be the corresponding $\varepsilon_H$ making the lemma work with this $\delta$,
- $\varepsilon_L := \varepsilon_H(\min\{\beta, \lambda\})$.

Note that we immediately have that $\partial^* E \cap B_{1/2}$ is included in a $\beta$-neighborhood of $\pi$, thanks to the second choice and the first conclusion of Lemma 5.2.

If now $y, z \in G$ are distinct, then we fix $s := 3|y - z| < \frac{1}{10}$ and we look at the ball $B_s(y)$. Since $s \le \frac{1}{4}$, by our second choice and the second conclusion of Lemma 5.2 we get

$$\mu(B_s(y)) < (\omega_n + \lambda)s^n.$$

More importantly, we have $\mathrm{Exc}(\pi, y, s) < \lambda$ by the very definition of $G$. Hence, Lemma 5.2 applies on $B_s(y)$: recalling that $\lambda = \varepsilon_H(\ell/10)$, its first conclusion gives

$$\mathrm{dist}(z-y, \pi) < \frac{\ell}{10}s,$$

which can be written as $|\pi^\perp(z-y)| < \frac{3}{10}\ell|z-y|$. Using the triangle inequality (and assuming wlog $\ell < 1$), it is easy to deduce

$$|\pi^\perp(z-y)| \le \ell|\pi(z-y)|.$$

It follows that the projection $\pi : G \to \pi$ is injective, since if $\pi(y) = \pi(z)$ we deduce that also $\pi^\perp(y) = \pi^\perp(z)$, and the inverse can be written as $\mathrm{id} + f$, with $f : \pi \to \pi^\perp$ satisfying $|f(z') - f(y')| \le \ell|z' - y'|$.

We already know $|f| \le \beta$, since $G \subseteq \partial^* E \cap B_{1/2}$, hence as we saw in the first class we can extend $f$ to an $\ell$-Lipschitz map $\pi \to \pi^\perp$, still denoted $f$, with $|f| \le \beta$. Thus we are left to show the third statement.

For all $y \in (\partial^* E \setminus G) \cap B_{1/100}$, by definition of excess there exists a ball $B_{s_y}(y)$ with $\int_{B_{s_y}(y)} \|\pi - T_z\partial^* E\|^2 \, d\mu(z) \ge \lambda s_y^n$ and $s_y < \frac{1}{10}$. Vitali's covering lemma gives a disjoint subcollection $\{B_{s_j}(y_j)\}$ such that the dilated balls $B_{5s_j}(y_j)$ cover $(\partial^* E \setminus G) \cap B_{1/100}$. It may be tempting to use the definition of Hausdorff measure to conclude, but the only upper bound we have on the radii $s_j$ is $\frac{1}{10}$. We use our measure $\mu$ instead:

$$\begin{aligned}
\mu((\partial^* E \setminus G) \cap B_{1/100}) &\le \sum_j \mu(B_{5s_j}(y_j)) \\
&\le C(n) \sum_j s_j^n \\
&\le \frac{C(n)}{\lambda} \int_{\bigcup B_{s_j}(y_j)} \|\pi - T_z\partial^* E\|^2 \, d\mu(z) \\
&\le \frac{C(n)}{\lambda} \mathrm{Exc}(\pi, 0, 1)
\end{aligned}$$

where the third inequality comes from disjointness, while the second follows form monotonicity, as

$$\frac{\mu(B_{5s_j}(y_j))}{(5s_j)^n} \le \frac{\mu(B_{1/2}(y_j))}{(1/2)^n} \le 2^n(\omega_n + \varepsilon_L)$$

(and wlog $\varepsilon_L < 1$, say). In order to estimate $(\Gamma_f \setminus G) \cap B_{1/100}$, we note that this is a subset of $\Gamma_f = (\mathrm{id} \times f)(\mathbb{R}^n)$, where now wlog $\pi = \mathbb{R}^n \times \{0\}$ and we view $f : \mathbb{R}^n \to \mathbb{R}$. Hence, by the area formula,

$$\mathcal{H}^n((\Gamma_f \setminus G) \cap B_{1/100}) \le \sqrt{2}\mathcal{L}^n(\pi((\Gamma_f \setminus G) \cap B_{1/100}))$$

where we used $J(\mathrm{id} \times f) = \sqrt{1 + |\nabla f|^2} \le \sqrt{1 + \ell^2}$ and wlog $\ell < 1$. Since $\pi$ is injective on $\Gamma_f$, the last measure is bounded by

$$\mathcal{L}^n(B^n_{1/100}) - \mathcal{L}^n(\pi(G)).$$

Finally, note that $\mathcal{H}^n(G) \geq \mu(B_{1/100}) - \frac{C}{\lambda}\operatorname{Exc}(\pi,0,1)$ by what we already proved, and $\mu(B_{1/100}) \geq \omega_n(1/100)^n\theta(0) \geq \omega_n(1/100)^n$ by monotonicity, while by the area formula (for the map $\pi$)

$$\mathcal{L}^n(\pi(G)) = \int_G J_{T_yG}\pi \, d\mathcal{H}^n(y)$$

and $J_{T_yG}\pi \geq 1 - C(n)\|\pi - T_y\partial^*E\|^2$ for a.e. $y \in G$ (by (4) and the fact that $G \subseteq \partial^*E$). Hence, $\mathcal{L}^n(\pi(G)) \geq \mathcal{H}^n(G) - C(n)\operatorname{Exc}(\pi,0,1) \geq \mathcal{L}^n(B^n_{1/100}) - \frac{C(n)}{\lambda}\operatorname{Exc}(\pi,0,1)$. Combining these inequalities, we arrive at

$$\mathcal{H}^n((\Gamma_f \setminus G) \cap B_{1/100}) \leq \sqrt{2}(\mathcal{L}^n(B^n_{1/100}) - \mathcal{L}^n(\pi(G))) \leq \frac{C(n)}{\lambda}\operatorname{Exc}(\pi,0,1),$$

as desired. □

5.2. **Harmonic approximation.** As already said, we want to show that $\partial^*E$ is essentially the graph of a harmonic function and exploit the fact that we have the desired improvement of flatness of such graphs. The latter fact is the content of the next very simple proposition.

**Proposition 5.3** ("improvement of flatness" for harmonic graphs)**.** *If* $u : B^n_1 \to \mathbb{R}$ *is harmonic, then for all* $0 < \eta \leq \frac{1}{2}$ *we have*

$$\sup_{x \in B^n_\eta} |u(x) - u(0) - \langle \nabla u(0), x\rangle| \leq C(n)\eta^2 \|\nabla u\|_{L^2(B^n_1)}.$$

The left-hand side bounds the distance of the point $(x, u(x)) \in \Gamma_u$ from the affine $n$-plane $\{u(0) + \langle \nabla u(0), z\rangle \mid z \in \mathbb{R}^n\}$. Hence, with respect to this plane, the graph $\Gamma_u$ has a small tilt on the ball $B_\eta((0, u(0)))$, much smaller than $\|\nabla u\|^2_{L^2}$ (which should be thought informally as the excess) if $\eta$ is chosen small. In the next section we will transfer this information to $f$, using the technical proposition below, and then show that also the excess on a ball of size $\sim \eta$ (with respect to a different plane) is much smaller than the original excess thanks to the tilt-exces inequality.

*Proof.* Let $\varphi \in C^\infty_c(B^n_{1/2})$ be radial and such that $\int_{\mathbb{R}^n} \varphi = 1$. Then the mean-value property enjoyed by $u$ gives $u = u * \varphi$ on $B^2_{1/2}$. By Taylor expansion we have

$$\sup_{x \in B^n_\eta} |u(x) - u(0) - \langle \nabla u(0), x\rangle| \leq \eta^2 \sup_{B^n_\eta} \|\nabla^2 u\|.$$

But $\frac{\partial^2}{\partial x_i \partial x_j}(u * \varphi) = \frac{\partial u}{\partial x_i} * \frac{\partial \varphi}{\partial x_j}$ on $B^n_\eta \subseteq B^n_{1/2}$, hence $\|\nabla^2 u\| \leq |\nabla u| * |\nabla \varphi|$ here (using the Hilbert–Schmidt norm for matrices) and this implies

$$\|\nabla^2 u\|(x) \leq \int_{\mathbb{R}^n} |\nabla u|(x-y)|\nabla\varphi|(y)\, dy \leq \|\nabla u\|_{L^2}\|\nabla\varphi\|_{L^2}$$

by Cauchy–Schwarz, for all $x \in B^n_\eta$. The claim follows since $\|\nabla\varphi\|_{L^2}$ depends only on $n$. □

The next proposition says that if a function $f : B_1^n \to \mathbb{R}$ almost verifies the weak definition of harmonicity, which would be

$$\int_{B_1^n} \langle \nabla f, \nabla \varphi \rangle = 0 \quad \text{for all } \varphi \in C_c^\infty(B_1^n),$$

then it is close in $L^2(B_1^n)$ to a harmonic function.

**Proposition 5.4** (harmonic approximation)**.** *Assume $f \in W^{1,2}(B_1^n)$ has $\int_{B_1^n} |\nabla f|^2 \le 1$. For all $\rho > 0$ there exists $\varepsilon_A(\rho)$ such that if*

$$\Big| \int_{B_1^n} \langle \nabla f, \nabla \varphi \rangle \Big| \le \varepsilon_A \|\nabla \varphi\|_{L^\infty} \quad \textit{for all } \varphi \in C_c^\infty(B_1^n)$$

*then there exists $u : B_1^n \to \mathbb{R}$ harmonic with*

$$\int_{B_1^n} |f-u|^2 \le \rho \quad \textit{and} \quad \int_{B_1^n} |\nabla u|^2 \le 1.$$

*Proof.* By contradiction, assume there exists a sequence of functions $f_j \in W^{1,2}$ with $\|\nabla f_j\|_{L^2} \le 1$ and

$$\Big| \int_{B_1^n} \langle \nabla f_j, \nabla \varphi \rangle \Big| \le \varepsilon_j \|\nabla \varphi\|_{L^\infty} \quad \text{for all } \varphi \in C_c^\infty(B_1^n),$$

for a sequence $\varepsilon_j \to 0$, but such that $\|f_j - u\|_{l^2} \ge \sqrt{\rho}$ for all $u : B_1^n \to \mathbb{R}$ harmonic with $\|\nabla u\|_{L^2} \le 1$.

Subtracting from each $f_j$ its average does not change any of these properties, so we can assume that $\int_{B_!^n} f_j = 0$. Poincaré's inequality then implies that $(f_j)$ is a bounded sequence in $W^{1,2}$, which is a Hilbert space. Hence, up to subsequences we can assume that $u_j \rightharpoonup u_\infty$ (weak convergence in $W^{1,2}$), which gives $u_j \to u_\infty$ strongly in $L^2$ by Rellich's compact embedding theorem.

Due to this strong convergence, we still have $\|f_\infty - u\|_{L^2} \ge \sqrt{\rho}$ for all $u$ as above, but now the weak convergence $\nabla f_j \rightharpoonup \nabla f_\infty$ in $L^2$ gives

$$\int_{B_1^n} \langle \nabla f_\infty, \nabla \varphi \rangle = \lim_{j\to\infty} \int_{B_1^n} \langle \nabla f_j, \nabla \varphi \rangle = 0$$

for all $\varphi \in C_c^\infty(B_1^n)$. This is the weak characterization of harmonicity, so $f_\infty$ is smooth and harmonic, Moreover, by lower semicontinuity of the norm under weak convergence,

$$\|\nabla f_\infty\|_{L^2} \le \liminf_{j\to\infty} \|\nabla f_j\|_{L^2} \le 1.$$

Hence, we can choose $u := f_\infty$ and obtain the contradiction $0 = \|f_\infty - u\|_{L^2} \ge \sqrt{\rho}$. □

**Remark 5.5.** We could have showed a real "improvement of flatness" for $u$, like

$$\sup_{x \in B_\eta^n} |\nabla u(x) - \nabla u(0)| \le C(n)\eta \|\nabla u\|_{L^2},$$

with the same proof. However, Proposition 5.4 forces us to use the tilt in order to transfer the improvement of flatness to $f$, since $f$ is close to $u$ in $L^2$ rather than $W^{1,2}$. The need of using the tilt comes also from the fact that we cannot bound the excess produced by $\partial^* E \setminus \Gamma_f$, while we will be able to say that its tilt is very small since its distance from $\pi$ is $\le \beta$.

We now show that a suitable normalization of the function $f$ given by Lipschitz approximation satisfies the previous proposition (where of course we can change the radius from 1 to 1/200).

**Proposition 5.6.** *Assume again $x = 0$, $r = 1$, $\pi = \mathbb{R}^n \times \{0\}$. If $\ell$ and $E := \mathrm{Exc}(\pi, 0, 1)$ are small enough, then $\int_{B^n_{1/200}} |\nabla \tilde f|^2 \le 1$ and*

$$\sup_{\varphi \in C^\infty_c(B^n_{1/200})\setminus\{0\}} \frac{\left| \int_{B^n_{1/200}} \langle \nabla \tilde f, \nabla \varphi \rangle \right|}{\|\nabla \varphi\|_{L^\infty}}$$

*can be made arbitrarily small, where $\tilde f := (C_0 E)^{-1/2} f$ for some dimensional constant $C_0(n)$ and $f : \mathbb{R}^n \to \mathbb{R}$ is the map given by Proposition 5.1. In more precise words, for all $\tilde\varepsilon$ there exists $\tilde\delta$ such that, if $\ell, E < \tilde\delta$, then $\int_{B^n_{1/200}} |\nabla \tilde f|^2 \le 1$ and*

$$\left| \int_{B^n_{1/200}} \langle \nabla \tilde f, \nabla \varphi \rangle \right| \le \tilde\varepsilon \|\nabla \varphi\|_{L^\infty}$$

*for all $\varphi \in C^\infty_c(B^n_{1/200})$.*

*Proof.* Step 1. Fix $\varphi \in C^\infty_c(B^n_{1/200})$. By stationarity we have

$$0 = \int_{B_{1/100}} \mathrm{div}_{T_y \partial^* E} (\varphi(y') e_{n+1})\, d\mu(y), \tag{8}$$

where we use the notation $y = (y', y_{n+1})$. Note that the vector field $X(y) := \varphi(y') e_{n+1}$ does not have compact support; however $\partial^* E \cap B_{1/100}$ is contained in a $\beta$-neighborhood of $\pi$ (wlog $\beta < 1/200$) and hence we can multiply our vector field by a cut-off function $\psi(y_{n+1})$, with $\psi(t) = 1$ for $|t| \le \beta$ and $\psi(t) = 0$ for $|t| \ge 1/200$, making it supported in $B_{1/100}$, without changing the value of the above integral, so that (8) is justified.

Also, by Proposition 5.1, noting that $T_y \partial^* E = T_y \Gamma_f$ for $\mathcal{H}^n$-a.e. $y \in G \subseteq \partial^* E \cap \Gamma_f$ as seen in the first class, we have

$$\begin{aligned} &\left| \int_{\mathbb{R}^{n+1}} \mathrm{div}_{T_y \Gamma_f} X \, d\mathcal{H}^n \llcorner \Gamma_f(y) - \int_{B_{1/100}} \mathrm{div}_{T_y \partial^* E} X \, d\mu(y) \right| \\ &\le C(n) \|\nabla \varphi\|_{L^\infty} \Big( \mathcal{H}^n((\Gamma_f \setminus G) \cap B_{1/100}) + \mathcal{H}^n((\partial^* E \setminus G) \cap B_{1/100}) \Big) \\ &\le \frac{C(n)}{\lambda} E \|\nabla \varphi\|_{L^\infty}. \end{aligned} \tag{9}$$

Now $T_y \Gamma_f = \mathrm{span}\{v_j(y')\}_1^n$ for $\mathcal{H}^n$-a.e. $y \in \Gamma_f$, where $v_j(y') := e_j + \frac{\partial f}{\partial x_j}(y') e_{n+1}$. Hence the projection onto $T_y \Gamma_f$, still denoted $T_y \Gamma_f$, is given by (exercise)

$$T_y \Gamma_f(z) = \sum_{i,j=1}^n v_i g^{ij} \langle v_j, z \rangle \tag{10}$$

for all $z \in \mathbb{R}^{n+1}$, where $g^{ij}$ is the inverse matrix of $g_{km} := \langle v_k, v_m \rangle$. Note that, since $|g_{km} - \delta_{km}| \le |\nabla f|^2 \le \ell^2$, if $\ell$ is small enough then the inverse $g^{ij}$ really exists and satisfies $|g^{ij} - \delta^{ij}| \le C(n) |\nabla f|^2$ pointwise.

The graph $\Gamma_f$ is the image of the injective map $\mathrm{id}\times f : \mathbb{R}^n \to \mathbb{R}^{n+1}$, so the area formula gives

(11)
$$\Big|\int_{\mathbb{R}^n} \mathrm{div}_{T_y\Gamma_f} X J(\mathrm{id}\times f)(y')\,dy'\Big| = \Big|\int_{\mathbb{R}^{n+1}} \mathrm{div}_{T_y\Gamma_f} X\, d\mathcal{H}^n \llcorner \Gamma_f(y)\Big| \le \frac{C(n)}{\lambda} E\|\nabla\varphi\|_{L^\infty},$$

where the inequality comes from (9) and (8).

Note that $J(\mathrm{id}\times f) = \sqrt{1+|\nabla f|^2} = 1 + O(|\nabla f|^2)$, while formula (10) gives

$$\begin{aligned}\mathrm{div}_{T_y\Gamma_f} X &= \langle T_y\Gamma_f \nabla\varphi(y'), T_y\Gamma_f e_{n+1}\rangle\\ &= \langle \nabla\varphi(y'), T_y\Gamma_f e_{n+1}\rangle\\ &= \sum_{i,j=1}^n \frac{\partial\varphi}{\partial x_i}(y') g^{ij}(y') \frac{\partial f}{\partial x_j}(y')\end{aligned}$$

(viewing $\nabla\varphi(y')$ as a vector in $\mathbb{R}^{n+1}$; we used that $\nabla\varphi(y') \perp e_{n+1}$, hence $\langle\nabla\varphi(y'), v_i\rangle = \frac{\partial\varphi}{\partial x_i}(y')$). Using also the fact that $g^{ij} = \delta^{ij} + O(|\nabla f|^2)$, we deduce that

$$\int_{\mathbb{R}^n} \mathrm{div}_{T_y\Gamma_f} X J(\mathrm{id}\times f)(y')\,dy' = \int_{\mathbb{R}^n} \langle\nabla f, \nabla\varphi\rangle + O\left(\|\nabla\varphi\|_{L^\infty} \int_{B^n_{1/200}} |\nabla f|^2\right). \tag{12}$$

<u>Step 2.</u> We claim that $\int_{B^n_{1/200}} |\nabla f|^2$ is bounded by $E$, provided that $\ell$ is small enough. Indeed,

$$\int_{B_{1/100}} \|\pi - T_y\Gamma_f\|^2\, d\mathcal{H}^n \llcorner \Gamma_f(y) \le E + C(n)\mathcal{H}^n((\Gamma_f \setminus G)\cap B_{1/100}) \le E + \frac{C(n)}{\lambda}E,$$

but, using again the area formula, the left-hand side is bounded below by

$$\int_{B^n_{1/200}} \|\pi - T_y\Gamma_f\|^2 J(\mathrm{id}\times f)\,dy',$$

since $(y', f(y')) \in B_{1/100}$ for all $y' \in B^n_{1/200}$ (recall that we assume $\beta < 1/200$ and $|f| \le \beta$). To conclude, note that

$$\begin{aligned}\|\pi - T_y\Gamma_f\|^2 &\ge |\pi(\pi - T_y\Gamma_f)e_{n+1}|^2\\ &= \Big|\sum_{i,j=1}^n \pi(v_i) g^{ij}\langle v_j, e_{n+1}\rangle\Big|^2\\ &\ge \Big|\sum_{i,j=1}^n e_i g^{ij}\frac{\partial f}{\partial x_j}\Big|^2\\ &\ge \frac{1}{2}\Big|\sum_{i,j=1}^n e_i \delta^{ij}\frac{\partial f}{\partial x_j}\Big|^2 - C(n)|\nabla f|^4,\end{aligned}$$

where we used $\Big|\sum_{i,j=1}^n e_i\delta^{ij}\frac{\partial f}{\partial x_j}\Big|^2 \le 2\Big|\sum_{i,j=1}^n e_i g^{ij}\frac{\partial f}{\partial x_j}\Big|^2 + 2\Big|\sum_{i,j=1}^n e_i(\delta^{ij} - g^{ij})\frac{\partial f}{\partial x_j}\Big|^2$. Hence, using also $J(\mathrm{id}\times f) = 1 + O(|\nabla f|^2)$, we reach the estimate

$$\int_{B^n_{1/200}} \|\pi - T_y\Gamma_f\|^2 J(\mathrm{id}\times f)\,dy' \ge \frac{1}{2}\int_{B^n_{1/200}} |\nabla f|^2(1 - C(n)|\nabla f|^2)\,dy'$$

and choosing $\ell$ so small that $1 - C(n)|\nabla f|^2 \geq 1 - C(n)\ell^2 \geq \frac{1}{2}$ we get

$$\int_{B^n_{1/200}} |\nabla f|^2 \leq 2\int_{B_{1/100}} \|\pi - T_y\Gamma_f\|^2 \, d\mathcal{H}^n \llcorner \Gamma_f(y) \leq C_0(n)E$$

for some big constant $C_0(n)$, since the parameter $\lambda$ depended only on $\ell$, which is now chosen in a precise way depending only on $n$ (so that now also $\lambda = \lambda(n)$). This proves our claim.

Step 3. We deduce that $\tilde f = (C_0E)^{-1/2} f$ satisfies $\int_{B^n_{1/200}} |\nabla \tilde f|^2 \leq 1$ and, combining (11) with (12),

$$\Big| \int_{B^n_{1/200}} \langle \nabla \tilde f, \nabla \varphi \rangle \Big| \leq C(n)\sqrt{E}\|\nabla\varphi\|_{L^\infty}.$$

The statement follows. □

5.3. **Proof of Theorem 4.9.** As usual, wlog we assume $x = 0$, $r = 1$, $\pi = \mathbb{R}^n \times \{0\}$. Let $f : \mathbb{R}^n \to \mathbb{R}$ given by the Lipschitz approximation, namely Proposition 5.1. Thanks to Proposition 5.6, if $E := \mathrm{Exc}(\pi, 0, 1)$ is small enough, we can apply (the "radius 1/200"-version of) Proposition 5.4 and find $\tilde u : B^n_{1/200} \to \mathbb{R}$ harmonic with

$$\int_{B^n_{1/200}} |\tilde f - \tilde u|^2 \leq \rho, \qquad \int_{B^n_{1/200}} |\nabla \tilde u|^2 \leq 1.$$

Setting $u := \sqrt{C_0E}\,\tilde u$ and recalling that $f = \sqrt{C_0E}\,\tilde f$, we find

$$\int_{B^n_{1/200}} |f - u|^2 \leq C(n)\rho E, \qquad \int_{B^n_{1/200}} |\nabla u|^2 \leq C(n)E. \tag{13}$$

Now let $\tilde x := (0, u(0))$ and $\tilde\pi := \mathrm{span}\{e_j + \frac{\partial u}{\partial x_j}(0)e_{n+1}\}_1^n$. Note that

$$\mathrm{dist}(\tilde x, \tilde \pi) \leq |\tilde x| = |u(0)| \leq C(n)\|u\|_{L^1} \leq C(n)\|u - f\|_{L^1} + C(n)\|f\|_{L^1} \leq C(n)(\sqrt{\rho E} + \beta),$$

where the bound on $|u(0)|$ comes from the mean-value property. We also have $|\nabla u(0)| \leq C(n)\|u\|_{L^1}$, because writing $u = u * \varphi$ near 0 as in the proof of Proposition 5.3 we get $|\nabla u(0)| = |(u * \nabla\varphi)(0)| \leq \|u\|_{L^1}\|\nabla\varphi\|_{L^\infty}$. This gives (exercise)

$$\|\pi^\perp - \tilde\pi^\perp\| = \|\pi - \tilde\pi\| \leq C(n)|\nabla u(0)| \leq C(n)(\sqrt{\rho E} + \beta),$$

as well. Now fix $\eta > 0$ to be chosen later on. We want to estimate the tilt

$$(4\eta)^{-n-2}\int_{B_{4\eta}(\tilde x)} \mathrm{dist}(y - \tilde x, \tilde\pi)^2 \, d\mu(y). \tag{14}$$

We first note that

$$\int_{B_{1/100}\setminus\Gamma_f} \mathrm{dist}(y - \tilde x, \tilde \pi)^2 \, d\mu(y) \leq C(n)E\,\mathrm{dist}(\tilde x, \tilde\pi)^2 + 2\int_{B_{1/100}\setminus\Gamma_f} \mathrm{dist}(y, \tilde\pi)^2 \, d\mu(y)$$

and, since $\mathrm{dist}(y, \tilde\pi) = |\tilde\pi^\perp y|$, the last integral is bounded by

$$2\int_{B_{1/100}\setminus\Gamma_f} \|\tilde\pi - \pi\|^2 \, d\mu(y) + 2\int_{B_{1/100}\setminus\Gamma_f} |\pi^\perp y|^2 \, d\mu(y) \leq C(n)(\rho E + \beta^2)E,$$

being $|\pi^\perp y| = \mathrm{dist}(y,\pi) \le \beta$ for $y \in \partial^* E \cap B_{1/100}$ and $\mathcal{H}^n((\partial^* E \setminus \Gamma_f) \cap B_{1/100}) \le C(n)E$ (recall that $\lambda$ was fixed during the proof of Proposition 5.6 and depends now only on $n$). We deduce that

$$\int_{B_{1/100}\setminus\Gamma_f} \mathrm{dist}(y-\tilde{x},\tilde{\pi})^2\, d\mu(y) \le C(n)(\rho E + \beta^2)E.$$

More importantly, using the area formula and bounding $J(\mathrm{id}\times f) \le 2$, we have

$$\int_{\Gamma_f \cap B_{4\eta}(\tilde{x})} \mathrm{dist}(y-\tilde{x},\tilde{\pi})^2\, d\mu(y) \le 2\int_{B^n_{4\eta}(0)} |f(y') - u(0) - \langle \nabla u(0), y'\rangle|^2\, dy',$$

since $y - \tilde{x} = (y', f(y') - u(0))$ and $(y', \langle \nabla u(0), y'\rangle) \in \tilde{\pi}$. The last integral is bounded by

$$C(n)\int_{B^n_{4\eta}} |f-u|^2 + C(n)\eta^n \sup_{y'\in B^n_{4\eta}} |u(y') - u(0) - \langle \nabla u(0), y'\rangle|^2 \le C(n)(\rho E + \eta^{n+4}E),$$

thanks to Proposition 5.3 and (13). Assuming our parameters $\eta,\beta,\rho$ are small enough, we have $B_{4\eta}(\tilde{x}) \subseteq B_{1/100}(0)$, in view of our bound for $|\tilde{x}|$. Hence, we can bound the quantity (14) with the estimates that we found for $B_{1/100}\setminus\Gamma_f$ and $\Gamma_f \cap B_{4\eta}(\tilde{x})$, arriving at

$$(4\eta)^{-n-2}\int_{B_{4\eta}(\tilde{x})} \mathrm{dist}(y-\tilde{x},\tilde{\pi})^2\, d\mu(y) \le C(n)\eta^{-n-2}(\rho E + \beta^2 + \rho + \eta^{n+4})E.$$

If $\beta$ and $\rho$ are small enough with respect to $\eta$ (assuming wlog $E \le 1$), then we have $|\tilde{x}| \le C(n)(\sqrt{\rho E} + \beta) \le \eta$ and hence the inclusion $B_{2\eta}(\tilde{x}) \supseteq B_\eta(0)$. This, together with the tilt-excess inequality (Proposition 4.11) on the ball $B_{4\eta}(\tilde{x})$, finally gives

$$\mathrm{Exc}(\tilde{\pi},0,\eta) \le 2^n\,\mathrm{Exc}(\tilde{\pi},\tilde{x},2\eta) \le C(n)\eta^{-n-2}(\rho E + \beta^2 + \rho)E + C(n)\eta^2 E.$$

It is now a trivial matter to realize that this is $\le E/2$ if we impose $\eta \ll 1$, $\beta^2 \ll \eta^{n+2}$ and $\rho \ll \eta^{n+2}$. Note that imposing small values for $\beta$ and $\rho$ requires having $E$ small enough, in turn, according to Proposition 5.1 and Proposition 5.6.

5.4. **Proof of Theorem 4.8.** As already said, Theorem 4.8 follows with really little work from Theorem 4.9. Assume again $x = 0$, $r = 1$.

In the video these theorems were called Theorem 1 and 2. For the sake of clarity, let us call $\varepsilon_{\mathrm{Thm}\ 2}$ the constant $\varepsilon$ making Theorem 4.9 ("Theorem 2") work. We will show that Theorem 4.8 ("Theorem 1") holds with a certain constant $\varepsilon = \varepsilon_{\mathrm{Thm}\ 1}$ much smaller than $\varepsilon_{\mathrm{Thm}\ 2}$.

First of all, using the "height lemma" (Lemma 5.2), we can take $\varepsilon_{\mathrm{Thm}\ 1}$ so small that

$$\mu(B_s(\hat{x})) < (\omega_n + \varepsilon_{\mathrm{Thm}\ 2})s^n \quad \text{for all } \hat{x} \in B_{1/2},\ s \le \frac{1}{4}. \tag{15}$$

So, fixing any $\hat{x} \in \partial^* E \cap B_{1/100}$ and asking also $4^n\varepsilon_{\mathrm{Thm}\ 1} < \varepsilon_{\mathrm{Thm}\ 2}$, we observe that Theorem 4.9 applies on $B_{1/4}(\hat{x})$ since

$$\mathrm{Exc}(\pi,\hat{x},1/4) \le 4^n\,\mathrm{Exc}(\pi,0,1) < 4^n\varepsilon_{\mathrm{Thm}\ 1} < \varepsilon_{\mathrm{Thm}\ 2}.$$

Hence, for a new $n$-plane $\pi_1$ we get

$$\mathrm{Exc}(\pi_1,\hat{x},\eta/4) \le \frac{1}{2}\,\mathrm{Exc}(\pi,\hat{x},1/4). \tag{16}$$

By the very nature of Theorem 4.9, we can iterate it infinitely many times. Indeed, its conclusion (16) and (15) give $\mu(B_{\eta/4}(\hat{x})) < (\omega_n + \varepsilon_{\text{Thm 2}})(\eta/4)^n$ and $\operatorname{Exc}(\pi_1, \hat{x}, \eta/2) < \varepsilon_{\text{Thm 2}}$. So we can apply Theorem 4.9 on $B_{\eta/4}(\hat{x})$ (with $\pi_1$ in place of $\pi$), and so on. We thus get planes $\pi_0 := \pi, \pi_1, \pi_2, \ldots$ such that

$$\operatorname{Exc}(\pi_{j+1}, \hat{x}, \eta^{j+1}/4) \leq \frac{1}{2} \operatorname{Exc}(\pi_j, \hat{x}, \eta^j/4),$$

hence $\operatorname{Exc}(\pi_j, \hat{x}, \eta^j/4) \leq 2^{-j} 4^n \operatorname{Exc}(\pi, 0, 1)$. We now claim that we can actually replace $\pi_j$ with $\pi$ and still have small excess. This follows comparing $\pi_j, \pi_{j+1}$ with the approximate tangent plane $T_y \partial^* E$:

$$\begin{aligned}\|\pi_j - \pi_{j+1}\|^2 &\leq \frac{2}{\mu(B_{\eta^{j+1}/4}(\hat{x}))} \int_{B_{\eta^{j+1}/4}(\hat{x})} \|\pi_j - T_y \partial^* E\|^2 \, d\mu(y) \\ &\quad + \frac{2}{\mu(B_{\eta^{j+1}/4}(\hat{x}))} \int_{B_{\eta^{j+1}/4}(\hat{x})} \|\pi_{j+1} - T_y \partial^* E\|^2 \, d\mu(y) \\ &\leq C(n) \operatorname{Exc}(\pi_j, \hat{x}, \eta^j/2) + C(n) \operatorname{Exc}(\pi_{j+1}, \hat{x}, \eta^{j+1}/2),\end{aligned}$$

where we used the inequalities $\mu(B_{\eta^{j+1}/4}(\hat{x})) \geq c(n)(\eta^{j+1}/4)^n, c(n)(\eta^j/4)^n$ coming from monotonicity (recall that $\theta(\hat{x}) = 1$ since $\hat{x} \in \partial^* E$). It follows that $\|\pi_j - \pi_{j+1}\| \leq C(n)\sqrt{2^{-j}\varepsilon_{\text{Thm 1}}}$, so that

$$\|\pi - \pi_j\| \leq \sum_{k=0}^{j-1} \|\pi_k - \pi_{k+1}\| \leq C(n) \sum_{k=0}^{j-1} 2^{-k/2} \sqrt{\varepsilon_{\text{Thm 1}}} \leq C(n)\sqrt{\varepsilon_{\text{Thm 1}}}$$

and, using $\frac{\mu(B_{\eta^j/4}(\hat{x}))}{(\eta^j/4)^n} \leq \frac{\mu(B_{1/4}(\hat{x}))}{(1/4)^n} \leq 4^n(\omega_n + \varepsilon_{\text{Thm 1}})$, we deduce

$$\operatorname{Exc}(\pi, \hat{x}, \eta^j/4) \leq 2 \operatorname{Exc}(\pi_j, \hat{x}, \eta^j/4) + 2\frac{\mu(B_{\eta^j/4}(\hat{x}))}{(\eta^j/4)^n} \|\pi - \pi_j\|^2 \leq C(n)\varepsilon_{\text{Thm 1}}.$$

It follows easily that, for a possibly different $C(n)$, we have $\operatorname{Exc}(\pi, \hat{x}, s) \leq C(n)\varepsilon_{\text{Thm 1}}$ for all $s \leq \frac{1}{4}$ (exercise). Now we apply again the Lipschitz approximation, namely Proposition 5.1, on $B_1$. If $\varepsilon_{\text{Thm 1}}$ is so small that the last quantity $C(n)\varepsilon_{\text{Thm 1}}$ is smaller than $\lambda$ (recall that $\lambda$ in the end depends only on $n$), then $\hat{x} \in G$. Hence, with all these assumptions on $\varepsilon_{\text{Thm 1}}$, we deduce that

$$\partial^* E \cap B_{1/100} \subseteq \Gamma_f$$

for some $\ell$-Lipschitz map $f : \pi \to \pi^\perp$. This is already a big achievement, given that general sets of finite perimeter can behave very wildly!

Next, assuming $\pi = \mathbb{R}^n \times \{0\}$, we define the relatively closed subset $G' \subseteq B^n_{1/200}$

$$G' := \{y' \in B^n_{1/200} : (y', f(y')) \in \overline{\partial^* E}\}.$$

Note that $(y', f(y')) \in B_{1/100}$, since wlog $\ell < 1$. We want to show that $G' = B^n_{1/200}$, which implies that in fact $\overline{\partial^* E} \cap B_{1/200} = \Gamma_f \cap B_{1/200}$. If $G'$ is a proper subset of $B^n_{1/200}$ then we can find a ball $B^n_{r_0}(y'_0) \subseteq B^n_{1/200}$ touching $G'$ from outside, meaning that $B^n_{r_0}(y'_0) \cap G' = \emptyset$ and $z'_0 \in G'$ for some $z'_0$ on the sphere $\partial B^n_{r_0}(y'_0)$. For instance, pick a point $y' \in \partial G$ inside the ball $B^n_{1/200}$ and then a point $y'_0 \in (G')^c$ very close to $y'$, and choose the biggest ball $B^n_{r_0}(y'_0)$ disjoint from $G'$: this ball must touch $G'$ by maximality.

Monotonicity gives $\mu(B_s(z_0)) \geq \omega_n s^n$, since $z_0 := (z_0', f(z_0'))$ belongs to $\overline{\partial^* E}$ and hence has $\theta(z_0) \geq 1$. On the other hand, the area formula gives

$$\mu(B_s(z_0)) \leq \int_{B_s^n(z_0') \cap G'} J(\mathrm{id} \times f),$$

since $\mu = \mathcal{H}^n \llcorner \partial^* E$ and $\partial^* E$ is parametrized by $\mathrm{id} \times f$ inside $B_s(z_0) \subseteq B_{1/100}$ (for $s$ small). But $J(\mathrm{id} \times f) = \sqrt{1 + |\nabla f|^2}$ and $B_s^n(z_0') \cap G' \subseteq B_s^n(z_0') \setminus B_{r_0}^n(y_0')$, which is asymptotically like a half-ball as $s \to 0$. Hence,

$$\omega_n s^n \leq \mu(B_s(z_0)) \leq (\omega_n/2 + o(1)) s^n \sqrt{1 + \ell^2}$$

which is a contradiction for $s \to 0$ (wlog $\ell < 1$). Thus we have "no holes" in the closure of the reduced boundary, i.e. $\overline{\partial^* E} \cap B_{1/200} = \Gamma_f \cap B_{1/200}$.

Finally, we turn to the $C^{1,\gamma}$ regularity of $\partial^* E$. Applying (and iterating) Theorem 4.9 as before at any $y \in \overline{\partial^* E} \cap B_{1/400}$, we get

$$\mathrm{Exc}(y, \pi_j, \eta^j/400) \leq C(n) 2^{-j} \varepsilon_{\mathrm{Thm\ 1}}.$$

As already observed, $\pi_j$ is very close to $\pi$. Hence, it equals $\mathrm{span}\{e_k + a_k^j e_{n+1} \mid k = 1, \ldots, n\}$ for some $a^j \in \mathbb{R}^n$ small (exercise). Also, observe that $T_z \partial^* E = \mathrm{span}\{e_j + \frac{\partial f}{\partial x_j}(z') e_{n+1}\}_1^n$ for $\mu$-a.e. $z$. This implies

$$|\nabla f(z') - a^j|^2 \leq C(n) \|T_z \partial^* E - \pi_j\|^2$$

(exercise; hint: reason as in the proof of Proposition 5.6). The graph of $f|_{B^n_{\eta^j/800}(y')}$ is included in $B_{\eta^j/400}(y)$, hence in $\partial^* E \cap B_{\eta^j/400}(y)$ up to a $\mathcal{H}^n$-negligible set (thanks to $\Gamma_f \cap B_{1/200} = \overline{\partial^* E} \cap B_{1/200}$, the inclusion $B_{\eta^j/400}(y) \subseteq B_{1/200}$ and the fact that $\partial^* E \setminus \partial E$ is $\mathcal{H}^n$-negligible). We deduce with the area formula that

$$\int_{B^n_{\eta^j/800}(y')} |\nabla f - a^j|^2 \leq C(n)\, \mathrm{Exc}(\pi_j, y, \eta^j/400) \leq C(n) 2^{-j}.$$

From this we deduce (exercise) that for all $y' \in B^n_{1/400}$ and all radii $s < \frac{1}{400}$ there exists $a = a(y', s) \in \mathbb{R}^n$ with

$$\int_{B_s(y')} |\nabla f - a|^2 \leq C(n) s^\alpha$$

for some dimensional constant $\alpha$. Since the left-hand side is minimized when $a$ equals the average of $\nabla f$ on the ball $B_s^n(y')$, denoted $(\nabla f)_{y',s}$, we arrive at

$$\int_{B_s^n(y')} |\nabla f - (\nabla f)_{y',s}|^2 \leq C(n) s^\alpha$$

for all $y' \in B^n_{1/400}$ and all $s < \frac{1}{400}$. This is the integral characterization of $C^{0,\alpha/2}$-functions, namely it gives $\nabla f \in C^{0,\alpha/2}$ on the ball $B^n_{1/400}$: see for instance [4, Theorem III.1.2]. Theorem 4.8 follows with $\gamma := \frac{\alpha}{2}$.

## References


[1] W. K. Allard. On the first variation of a varifold. *Ann. of Math. (2)* **95** (1972), pp. 417–491.

[2] L. Ambrosio, N. Fusco and D. Pallara. *Functions of bounded variation and free discontinuity problems*. In *Oxford Mathematical Monographs*. Oxford University Press, 2000.

[3] C. De Lellis. *Allard's interior regularity theorem: an invitation to stationary varifolds*. Available online.

[4] M. Giaquinta. *Multiple integrals in the calculus of variations and nonlinear elliptic systems*. Vol. 105 in *Annals of Mathematics Studies*. Princeton University Press, 1983.

[5] F. Maggi. *Sets of finite perimeter and geometric variational problems*. Vol. 135 in *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2012.