Lecture Notes

Zihui Zhao

May 28, 2026

Contents

1 Introduction

Suppose $u : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ satisfies that its graph minimizes the area, among competitors which fix the boundary.

Note that for any $v : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ the area functional is given by

$$\operatorname{Area}\operatorname{Graph}(v) = \int_\Omega \sqrt{1 + |\nabla v|^2}\, dx.$$

For any $\varphi \in C_0^\infty(\Omega)$, we have

$$\begin{aligned}\operatorname{Area}\operatorname{Graph}(u + t\varphi) = &\int_\Omega \sqrt{1 + |\nabla u|^2}\, dx + \int_\Omega \frac{\langle \nabla u, \nabla \varphi\rangle}{\sqrt{1 + |\nabla u|^2}}\, dx \cdot t \\ &+ \int_\Omega \frac{|\nabla \varphi|^2 + |\nabla u|^2 |\nabla \varphi|^2 - |\langle \nabla u, \nabla \varphi\rangle|^2}{(1 + |\nabla u|^2)^{3/2}}\, dx \cdot \frac{t^2}{2} + O(t^3).\end{aligned}$$

The minimality assumption implies that

$$\frac{d}{dt}\Big|_{t=0} \operatorname{Area}\operatorname{Graph}(u+t\varphi) = \int_\Omega \frac{\langle \nabla u, \nabla \varphi\rangle}{\sqrt{1+|\nabla u|^2}}\, dx = 0. \tag{1}$$

This is the weak formulation of the divergence form elliptic equation

$$\operatorname{div}\left(\frac{\nabla u}{\sqrt{1+|\nabla u|^2}}\right) = 0. \tag{2}$$

Or equivalently, to phrase it in non-divergence form

$$\Delta u - \frac{\partial_i u \cdot \partial_j u}{1+|\nabla u|^2}\partial_{ij} u = 0. \tag{3}$$

They are referred to as the **minimal surface equation (MSE)**. They resemble the familiar Laplacian equation $\Delta u = \operatorname{div}(\nabla u) = 0$.

Note that the condition (1) is only necessary for minimality: It just says that u is a critical point of the area functional. However, because of the convexity of $\operatorname{Area}\operatorname{Graph}(u+t\varphi)$, minimality actually follows from stationarity. In fact, a calibration argument shows that among all surfaces $\Sigma \subset \Omega\times\mathbb{R}$ with $\partial\Sigma = \partial \operatorname{Graph}(u)$, we still have $\operatorname{Area}\operatorname{Graph}(u) \leq \operatorname{Area}\Sigma$, see [CM, Lemma 1.1].

We remark that not all minimal surfaces are graphical (i.e. given by the graph of a function), for example the catenoid is a two-dimensional minimal surface in $\mathbb{R}^3$ and it is clearly not the graph of a function. We will focus on minimal graphs in this class, where one can already see many interesting features and techniques in the regularity theory of minimal surfaces. In particular, we will focus on the following **Bernstein's problem** for minimal graphs.

Theorem 1.1. *Suppose $u : \mathbb{R}^n \to \mathbb{R}$ is a global solution of the MSE. Then*

i) u must be linear when $n \leq 7$, that is to say $\operatorname{Graph}(u)$ *is just a hyperplane;*

ii) nonlinear solutions exist when $n \geq 8$.

This is in stark contrast with the classical Liouville's theorem for harmonic functions: Any harmonic function on $\mathbb{R}^n$ that is bounded from above or from below is a constant. More generally, if a harmonic function is bounded by a polynomial of degree k near infinity, it must be a polynomial of degree at most k. (e.g. there exists homogeneous harmonic polynomial of any degree)

There are analogous theories for

- variational solutions of $\int_\Omega F(x,u,\nabla u)\, dx$ where F is uniformly convex in the last variable ∇u (e.g. $F(x,z,p) = \frac{1}{2}|p|^2$).

  Exercise. *The Euler-Lagrange equation for the functional $\int_\Omega F(x,u,\nabla u)\, dx$ is*

  $$\operatorname{div}(\nabla_p F(x,u,\nabla u)) - \partial_z F(x,u,\nabla u) = 0;$$

- prescribed minimal surface equation

  $$\frac{1}{\sqrt{1+|\nabla u|^2}}\left(\Delta u - \frac{\partial_i u\cdot\partial_j u}{1+|\nabla u|^2}\partial_{ij} u\right) = nH,$$

  for given H;

- the stationary flow of irrotational compressible fluid

$$\operatorname{div}(\rho\nabla u) = 0,$$

  where u is the velocity and $\rho = \rho(\nabla u)$ is the density of the fluid;

- equation for capillary surfaces in a gravitational field, etc.

On the other hand, even for minimal surface system, i.e. a vector-valued function $u : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ such that its graph minimizes the n-dimensional area in $\mathbb{R}^{n+m}$, the theory is quite different.

2 Regularity ladder of second-order elliptic PDEs

To fix ideas, consider a **linear** elliptic PDE

$$\operatorname{div}(A(x)\nabla u) + b \cdot \nabla u \,(+cu) = 0 \tag{4}$$

or

$$a_{ij}(x)\partial_{ij} u + b_i(x)\partial_i u = 0, \tag{5}$$

sometimes with prescribed boundary data $u = f$ on $\partial\Omega$. Here $A(x) = (a_{ij}(x))$ is assumed to be an $n \times n$ real-symmetric matrix. **Ellipticity** means that there exists a uniform constant $\lambda > 0$ such that

$$a_{ij}\xi_i\xi_j \geq \lambda|\xi|^2 \quad \text{for any } \xi \in \mathbb{R}^n \setminus \{0\}. \tag{6}$$

Here I use the Einstein summation convention. Assuming certain differentiability of the coefficients, these two forms of elliptic PDEs are interchangeable; but they each have their own merits and some techniques are better suited to study one form than the other, as we shall see below.

The equations (4) and (5) are linear in u: If u and v are two solutions of the same PDE, their sum $u + v$ also solves the same PDE. On the other hand the MSE is not linear, and it belongs to a class of **quasilinear** elliptic PDEs

$$a_{ij}(x, u, \nabla u)\partial_{ij} u + b(x, u, \nabla u) = 0, \tag{7}$$

with the property that the coefficient for the highest order, a_{ij}, does not depend on the highest-order derivatives. The other examples mentioned in Section 1 are also quasilinear elliptic PDEs; the p-Laplacian equation $\operatorname{div}(|\nabla u|^{p-2}\nabla u) = 0$ is another example of quasilinear elliptic PDE.

Exercise. *Check if the MSE and the p-Laplacian equation (for $p \neq 2$) are uniformly elliptic.*

Functional analysis (Lax-Milgram theorem) guarantees that for elliptic PDEs of divergence form, weak solutions $u \in W^{1,2}(\Omega)$ exist and satisfy the weak formulation: for any $\varphi \in C_0^\infty(\Omega)$,

$$\int_\Omega \langle A(x)\nabla u, \nabla\varphi\rangle - b \cdot \nabla u \,\varphi \, dx = 0. \tag{8}$$

Moreover, a function $u \in C^2(\Omega)$ that satisfies (4) in the classical sense also satisfies the weak formulation (8). A natural follow-up question is: How regular is the weak solution? Historically this is Hilbert's 19th problem: solutions to problems from calculus of variations with analytic coefficients are analytic. Next we give a brief overview of regularity theories of second-order elliptic PDEs. Proving some of the results below rigorously will take an entire semester and will be a digression from the minimal surface theory, but I still want to summarize known theories, in order to highlight the analytic difficulty in studying the MSE.

1. u is bounded by the weak maximum principle $\sup_\Omega |u| = \sup_{\partial\Omega} |f|$. Alternatively $|u|$ is bounded by the L^2 norm of u in a bigger ball, by Moser's estimate.

2. **Schauder estimates** for non-divergence form: Assuming a_{ij}, b_i have Hölder regularity with norm bounded by Λ, we have

$$\|u\|_{C^{2,\alpha}} \leq C \sup |u|$$

where C depends on λ and Λ. (See [GT, Chapter 6].) More generally, assuming the coefficients have $C^{k,\alpha}$ regularity with $k = 0, 1, \cdots$, we get estimates for the $C^{k+2,\alpha}$ norm of u.

Why is it natural to use the Hölder space $C^{2,\alpha}$ instead of the C^2 space? The proof of Schauder estimates is a perturbative argument by freezing the coefficient, so that

$$a_{ij}(x_0)\partial_{ij}u = (a_{ij}(x_0) - a_{ij}(x))\, \partial_{ij}u.$$

On the left hand side, we have constant coefficient $a_{ij}(x_0)$. A model case for constant-coefficient elliptic PDE is the Poisson equation $\Delta v = f$. When $f \in C^{0,\alpha}$, it follows that $v \in C^{2,\alpha}$; however when f is merely C^0, we can't guarantee that $u \in C^2$.

Check if the minimal surface equation, or a more general quasilinear equation $a_{ij}(x, u, \nabla u)\partial_{ij}u + b(x, u, \nabla u) = 0$, satisfies both the elliptic condition and the regularity condition required to apply the Schauder theory. Note that for the MSE $a_{ij} = \delta_{ij} - \frac{\partial_i u \cdot \partial_j u}{1+|\nabla u|^2}$. It satisfies

$$a_{ij}\xi_i\xi_j = \frac{|\xi|^2 + |\xi|^2|\nabla u|^2 - |\langle \xi, \nabla u\rangle|^2}{1 + |\nabla u|^2} \geq \frac{|\xi|^2}{1 + |\nabla u|^2} \geq \lambda|\xi|^2 > 0,$$

if $|\nabla u|$ is uniformly bounded; and moreover if we also have bounds on the Hölder norm of ∇u, the Schauder theory applies to show $u \in C^{2,\alpha}$.

3. **De Giorgi-Nash-Moser theory** for divergence form[1] equation with bounded measurable coefficients: Suppose u is a weak solution $\mathrm{div}(A(x)\nabla u) + b \cdot \nabla u = 0$ in B_2, with coefficients $A(x)$ and $b(x)$ merely measurable and bounded by Λ. Then there exists $\alpha \in (0, 1)$ depending on n, λ, Λ such that $u \in C^{0,\alpha}$ and

$$\|u\|_{C^{0,\alpha}(B_1)} \leq C \sup_{B_2} |u|.$$

(See [GT, Chapter 8], or [HL, Chapter 4].) Though phrased for linear PDE, this theory is super useful for quasilinear equation where the coefficient $a_{ij}(x, u, \nabla u)$ depends on the function itself and thus no continuity is expected apriori.

[1] Krylov-Safonov developed the theory to prove analogous statement for non-divergence form.

4. **Iteration.** Combined with the Schauder estimate, we now have that $u \in C^{2,\alpha}$. Iterating this argument gives that $u \in C^k$ for any $k \in \mathbb{N}$, and thus $u \in C^\infty$. Analyticity follows by Morrey-Nirenberg.

Applications of De Giorgi-Nash-Moser theory to the MSE. Suppose u is a weak solution to the MSE. We claim that its derivative $v := \partial_k u$ satisfies the following PDE

$$\operatorname{div}(\bar{A}\nabla v) = 0, \qquad \text{where } \bar{A} = \frac{1}{\sqrt{1+|\nabla u|^2}}\left(\mathrm{Id} - \frac{\nabla u \otimes \nabla u}{1+|\nabla u|^2}\right).$$

In the weak formulation of the MSE (1), we plug in test function $\varphi = \partial_k \eta$ in order to deduce the PDE satisfied by the derivative $v := \partial_k u$. Since (again we're using Einstein summation convention)

$$\begin{aligned}
\frac{\partial_j u\, \partial_{jk}\eta}{\sqrt{1+|\nabla u|^2}} &= \partial_k\left(\frac{\partial_j u\, \partial_j \eta}{\sqrt{1+|\nabla u|^2}}\right) - \frac{\partial_{jk} u\, \partial_j \eta}{\sqrt{1+|\nabla u|^2}} - \partial_j u\, \partial_j \eta\, \partial_k \frac{1}{\sqrt{1+|\nabla u|^2}} \\
&= \partial_k\left(\frac{\partial_j u\, \partial_j \eta}{\sqrt{1+|\nabla u|^2}}\right) - \frac{\partial_j v\, \partial_j \eta}{\sqrt{1+|\nabla u|^2}} + \partial_j u\, \partial_j \eta\, \frac{\partial_i u\, \partial_i v}{(\sqrt{1+|\nabla u|^2})^3},
\end{aligned}$$

thus by the integration by parts we have

$$\int_\Omega \left[\frac{\langle \nabla v, \nabla \eta\rangle}{\sqrt{1+|\nabla u|^2}} - \partial_i v\, \frac{\partial_i u\, \partial_j u}{(\sqrt{1+|\nabla u|^2})^3}\, \partial_j \eta\right] dx = 0, \qquad \text{for any } \eta \in C_0^\infty(\Omega).[2]$$

This is the weak formulation for the PDE

$$\partial_i(\bar{a}_{ij}\partial_j v) = 0, \qquad \text{where } \bar{a}_{ij} = \frac{\delta_{ij}}{\sqrt{1+|\nabla u|^2}} - \frac{\partial_i u \cdot \partial_j u}{(\sqrt{1+|\nabla u|^2})^3};$$

or in matrix form

$$\operatorname{div}(\bar{A}\nabla v) = 0, \qquad \text{where } \bar{A} = \frac{1}{\sqrt{1+|\nabla u|^2}}\left(\mathrm{Id} - \frac{\nabla u \otimes \nabla u}{1+|\nabla u|^2}\right).$$

Under the assumption that $|\nabla u|$ is uniformly bounded, the coefficient matrix $\bar{A}$ is elliptic and bounded. Hence we may apply the De Giorgi-Nash-Moser theory to conclude that $v = \partial_k u \in C^{0,\alpha}$ for some $\alpha \in (0,1)$ (depending on the upper bound of $|\nabla u|$). This in turn allows us to apply Schauder theory to the MSE.

To summarize, the regularity (and quantitative estimates of norms) of solutions to the MSE relies heavily on the boundedness of the gradient. In fact, we can prove the following **twist of the Bernstein's problem**, using the De Giorgi-Nash-Moser theory:

Theorem 2.1. *Suppose $u : \mathbb{R}^n \to \mathbb{R}$ is a global solution of the MSE and the dimension n is arbitrary. If $|\nabla u|$ is uniformly bounded in $\mathbb{R}^n$, then u must be linear.*

Proof. By rescaling, we can show the following variant of the De Giorgi-Nash-Moser estimate: Suppose v is a weak solution to $\operatorname{div}(A(x)\nabla v) + b \cdot \nabla v = 0$ in B_{2R}, where the

[2]Note that using the difference quotient, we can show that $u \in W^{2,2}$, which makes the above calculation rigorous.

coefficient matrix $A(\cdot)$ is uniformly elliptic with parameter λ and bounded from above by Λ, and the coefficient $b(\cdot)$ is bounded from above by Λ/R. Then $v \in C^{0,\alpha}$, with estimate

$$R^\alpha \sup_{x,y\in B_R} \frac{|v(x)-v(y)|}{|x-y|^\alpha} \le C \sup_{B_{2R}} |v|, \tag{9}$$

where the constants $\alpha \in (0,1)$ and $C > 0$ only depend on n, λ, Λ.

Next we apply the above estimate to the function $v = \partial_k u$ solving $\operatorname{div}(\bar{A}\nabla v) = 0$. Let $x_0, y_0 \in \mathbb{R}^n$ be fixed, and let R be sufficiently large so that $x_0, y_0 \in B_R$. Then (9) implies that

$$\frac{|v(x_0)-v(y_0)|}{|x_0-y_0|^\alpha} \le CR^{-\alpha} \sup_{B_{2R}} |v| \le CMR^{-\alpha},$$

where M denotes the uniform upper bound of $|\nabla u|$. Passing $R \to +\infty$, we get that $v(x_0) = v(y_0)$, i.e. v is a constant. Therefore u is a linear function. $\square$

Sketch of De Giorgi's approach. Next we give a sketch of the De Giorig-Nash-Moser theory, mainly following De Giorgi's approach. We will only prove Theorem 2.2 rigorously, since the proof argument of choosing good test functions to get desired integral estimates is the starting point of many estimates in the minimal surface theory and calculus of variations. We also emphasize that the proof does not use *linearity* at all, only the ellipticity and boundedness of the coefficient matrix $A(\cdot)$.

From now on, we always assume that $A(\cdot)$ is elliptic with parameter $\lambda > 0$ (see (6)) and bounded from above by Λ.

Theorem 2.2. *Suppose u is a subsolution to the operator $L = -\operatorname{div}(A\nabla)$ in B_2, namely*

$$\int \langle A\nabla u, \nabla\varphi\rangle dx \le 0, \qquad \text{for any } \varphi \in C_c^\infty(B_2) \text{ s.t. } \varphi \ge 0. \tag{10}$$

Then

$$\sup_{B_1} u \le C\|u^+\|_{L^2(B_2)},$$

where $u^+ = \max\{u, 0\}$, and the constant C only depends on n, λ, Λ.

Remark 2.3. The above estimate can be generalized to subsolutions of elliptic operators with lower order terms and right hand side, namely $\operatorname{div}(A\nabla u) + cu \le f$ where $c, f \in L^q$ for some $q > n/2$. See [HL, Theorem 7.3] for details.

Proof. Recall that $C_c^\infty(B_2)$ is dense in the Sobolev space $W_0^{1,2}(B_2)$. So the weak formulation (10) in fact holds for all test functions $\varphi \in W_0^{1,2}(B_2)$ such that $\varphi \ge 0$. In order to estimate the upper bound of u, we consider $v := (u-k)^+$ for $k > 0$ large. Take a smooth test function $\eta \in C_c^\infty(B_2)$ to be chosen later. Plugging the test function $\varphi := v\eta^2 \ge 0$ into (10), we get

$$\begin{aligned}
\lambda \int |\nabla v|^2\eta^2 \le \int \langle A\nabla v, \nabla v\rangle \eta^2 &\le -2\int \eta v \langle A\nabla u, \nabla\eta\rangle \\
&\le 2\Lambda \int \eta v |\nabla u||\nabla\eta| \\
&\le 2\Lambda \left(\int |\nabla v|^2\eta^2\right)^{1/2} \left(\int v^2|\nabla\eta|^2\right)^{1/2}.
\end{aligned}$$

Therefore

$$\int |\nabla v|^2\eta^2 \le \left(\frac{2\Lambda}{\lambda}\right)^2 \int v^2|\nabla\eta|^2. \tag{11}$$

This is usually referred to as a Cacciopolli-type inequality, where we bound the integral norm of ∇v by an integral norm of v, and it can be thought of a reverse Poincaré inequality.

By the Sobolev inequality, we have

$$\left(\int |v\eta|^{\frac{2n}{n-2}}\right)^{\frac{n-2}{n}} \lesssim \int |\nabla(v\eta)|^2 \lesssim \int v^2|\nabla\eta|^2,$$

where we use $\lesssim$ to indicate bounded from above modulo a uniform constant (depending on n,λ,Λ). Moreover, Hölder's inequality implies that

$$\int |v\eta|^2 \le \left(\int |v\eta|^{\frac{2n}{n-2}}\right)^{\frac{n-2}{n}} |\{v\eta\neq 0\}|^{\frac{2}{n}} \lesssim \left(\int v^2|\nabla\eta|^2\right) |\{v\eta\neq 0\}|^{\frac{2}{n}}.$$

Fixing $0<r<R<2$, we choose $\eta\in C_c^\infty(B_2)$ such that $0\le\eta\le 1$, $\eta\equiv 1$ on B_r, and $|\nabla\eta|\le C/(R-r)$ with a uniform constant C. It follows that

$$\int_{B_r} v^2 \lesssim \frac{1}{(R-r)^2}\left(\int_{B_R} v^2\right) |\{x\in B_R : v>0\}|^{\frac{2}{n}}. \tag{12}$$

To write the above estimate in u, denote the super level set

$$A(k,r) := \{x\in B_r : u(x)\ge k\}.$$

Then (12) becomes

$$\int_{A(k,r)} |u-k|^2 \lesssim \frac{1}{(R-r)^2}\left(\int_{A(k,R)} |u-k|^2\right) |A(k,R)|^{\frac{2}{n}}.$$

For any two levels $0<h<k$, note that

$$A(k,R)\subset A(h,R), \qquad \text{and} \qquad |u-h|^2 \ge (k-h)^2 \text{ on } A(k,R).$$

Hence

$$\int_{A(k,r)} |u-k|^2 \lesssim \frac{1}{(R-r)^2}\frac{1}{(k-h)^{\frac{4}{n}}}\left(\int_{A(h,R)} |u-h|^2\right)^{1+\frac{2}{n}},$$

or equivalently by denoting $\varphi(k,r) = \|(u-k)^+\|_{L^2(B_r)}$, we have

$$\varphi(k,r) \lesssim \frac{1}{R-r}\frac{1}{(k-h)^{\frac{2}{n}}}\varphi(h,R)^{1+\frac{2}{n}}. \tag{13}$$

We aim to show that the *superlinear decay* (i.e. the exponent $1+2/n>1$ above) as we move from h up to k implies the existence of some level $k_{\max}$ such that $\varphi(k_{\max},1)=0$. In other words $u\le k_{\max}$ on B_1. We prove this inductively on the levels and radii. Consider an increasing (but finite) sequence $k_\ell = A(1-2^{-\ell})\ge 0$ and a decreasing sequence $r_\ell = 1+2^{-\ell}\in[1,2]$, with index $\ell=0,1,\cdots$, where A is a positive constant to be determined later. Since

$$k_\ell - k_{\ell-1} = \frac{A}{2^\ell}, \quad r_{\ell-1}-r_\ell = \frac{1}{2^\ell},$$

the estimate (13) becomes

$$\varphi(k_\ell,r_\ell) \le \frac{C}{A^{\frac{2}{n}}} 2^{\ell(1+\frac{2}{n})}\varphi(k_{\ell-1},r_{\ell-1})^{1+\frac{2}{n}}.$$

Exercise. *Suppose a positive sequence a_ℓ satisfies the super-linear decay as follows*

$$a_\ell \le M\beta^\ell a_{\ell-1}^{1+\alpha}, \qquad \text{for some } \alpha > 0 \text{ and } \beta > 1.$$

Show that $a_\ell \to 0$ as $\ell \to +\infty$, as long as $Ma_0^\alpha \beta^{1+1/\alpha} \le 1$.

We claim that

$$\varphi(k_\ell, r_\ell) \le \frac{\varphi(k_0, r_0)}{\gamma^\ell}, \qquad \text{for all } \ell \ge 0, \tag{14}$$

with the constant $\gamma = 2^{n/2+1} > 1$, which immediately implies that $\varphi(k_\ell, r_\ell) \to 0$ as $\ell \to +\infty$. The case for $\ell = 0$ holds trivially. Assuming (14) holds for $\ell - 1$, it follows that

$$\begin{aligned} \varphi(k_\ell, r_\ell) \le \frac{C}{A^{\frac{2}{n}}} 2^{\ell(1+\frac{2}{n})} \left(\frac{\varphi(k_0, r_0)}{\gamma^{\ell-1}}\right)^{1+\frac{2}{n}} &= \frac{C\varphi(k_0, r_0)^{\frac{2}{n}} \gamma^{1+\frac{2}{n}}}{A^{\frac{2}{n}}} \left(\frac{2^{1+\frac{2}{n}}}{\gamma^{\frac{2}{n}}}\right)^\ell \frac{\varphi(k_0, r_0)}{\gamma^\ell} \\ &= \frac{C\varphi(k_0, r_0)^{\frac{2}{n}} \gamma^{1+\frac{2}{n}}}{A^{\frac{2}{n}}} \frac{\varphi(k_0, r_0)}{\gamma^\ell}, \end{aligned}$$

by the choice of γ; additionally we choose A such that the constant in front is 1, namely

$$A = \left(C\varphi(k_0, r_0)^{\frac{2}{n}} \gamma^{1+\frac{2}{n}}\right)^{\frac{n}{2}} = C^{\frac{n}{2}} \gamma^{1+\frac{n}{2}} \varphi(0, 2) = C' \|u^+\|_{L^2(B_2)}.$$

This finishes the proof of the claim (14). Since $k_\ell \nearrow A$ and $r_\ell \searrow 1$, it follows that $\varphi(A, 1) = 0$, and thus

$$u \le A = C' \|u^+\|_{L^2(B_2)} \text{ on } B_1.$$

□

Theorem 2.2 can also be used to give a lower bound for positive *supersolutions*:

Theorem 2.4 (Density theorem). *Suppose u is a positive supersolution to the operator $L = -\operatorname{div}(A\nabla)$ in B_2, namely*

$$\int \langle A\nabla u, \nabla\varphi \rangle dx \ge 0, \qquad \text{for any } \varphi \in C_c^\infty(B_2) \text{ s.t. } \varphi \ge 0.$$

Assume that

$$|\{x \in B_2 : u(x) \ge 1\}| \ge \epsilon |B_2|. \tag{15}$$

Then there exists a constant $c > 0$ depending on n, λ, Λ and ϵ such that

$$\inf_{B_1} u \ge c.$$

Proof.

Exercise. *Since u is a positive supersolution, one can check that $v := (\log u)^-$ is a subsolution (since $t \mapsto \log t$ is concave).*

Hence an upper bound on v gives a lower bound on u. On the other hand, the assumption (15) implies that $|\{x \in B_2 : v(x) = 0\}| \ge \epsilon |B_2|$, hence Poincaré inequality applies and we can estimate the L^2-norm of v by the L^2-norm of its gradient. See [HL, Theorem 4.9] for details of the proof. □

Theorem 2.5 (Decay of the oscillation of u)**.** *Suppose u is a weak solution to $-\operatorname{div}(A\nabla u) = 0$ in B_2. There exists $\theta \in (0,1)$ such that*

$$\operatorname{osc}_{B_{1/2}} u \le \theta \operatorname{osc}_{B_1} u. \tag{16}$$

Proof. Applying Theorem 2.2 to $\pm u$, we get that u is bounded from above and below in B_1. Let

$$M = \sup_{B_1} u, \quad m = \sup_{B_1} u.$$

Consider the two solutions

$$\frac{M-u(x)}{M-m} \text{ and } \frac{u(x)-m}{M-m}$$

which take values in $[0,1]$ in B_1 and add up to 1. So either one of them is $\ge 1/2$ on half of the ball B_1. Applying Theorem 2.4 to either one of them gives a uniform lower bound $c \in (0,1)$. Therefore we get the decay of the oscillation of u, with parameter $\theta = 1-c$. ☐

Note that the Hölder regularity of u follows immediately from iterating the decay estimate (16). This finishes the sketch of De Giorgi's proof.

Going back to the MSE, we have shown that the regularity of solutions to the MSE entirely boils down to estimating its gradient. In the 1960s Bombieri, De Giorgi and Giusti proved the following interior gradient estimate, by studying PDEs on the minimal surface Graph(u).

Theorem 2.6. *Let $u \in C^2(B_{2R})$ be a solution to the MSE in B_{2R}. Then for any $x \in B_R$, we have*

$$|\nabla u(x)| \le C_1 \exp\left(C_2 \frac{\sup_{B_{2R}} u - u(x)}{R}\right), \tag{17}$$

where C_1, C_2 are universal constants.

Before setting out to prove the interior gradient estimate, we need to introduce a few preliminaries.

3 Preliminaries from geometric measure theory

3.1 Hausdorff measure

Definition 3.1. Let $E \subset \mathbb{R}^n$ and $s > 0$. For any $\delta > 0$ we denote

$$\mathcal{H}^s_\delta(E) := \inf\left\{\sum_i \omega_s r_i^s : E \subset \cup_i B(x_i, r_i), r_i \le \delta\right\}, \quad \text{where } w_s = \frac{\pi^{\frac{s}{2}}}{\Gamma\left(\frac{s}{2}+1\right)};$$

and the **s-dimensional Hausdorff measure** is defined as

$$\mathcal{H}^s(E) := \lim_{\delta \to 0} \mathcal{H}^s_\delta(E) = \sup_{\delta > 0} \mathcal{H}^s_\delta(E).$$

Remark 3.2. If we do not impose the restriction $r_i \le \delta$ or do not pass to the limit $\delta \to 0$, then for any E bounded

$$\mathcal{H}^s_\infty(E) := \inf \left\{ \sum_i \omega_s r_i^s : E \subset \cup_i B(x_i, r_i) \right\} \lesssim \operatorname{diam}(E)^s < \infty,$$

which does not reflect the true geometry of the set E.[3] e.g. a fractal curve in $\mathbb{R}^3$.

Proposition 3.3. *The Hausdorff measure in $\mathbb{R}^n$ satisfies the following properties:*

i) $\mathcal{H}^0$ *is the counting measure, and* $\mathcal{H}^n(E) = |E|$*, the Lebesgue measure of* E*;*

ii) $\mathcal{H}^s(\lambda E) = \lambda^s \mathcal{H}^s(E)$ *for any* $\lambda > 0$*;*

iii) $\mathcal{H}^s(L(E)) = \mathcal{H}^s(E)$ *for a rigid motion* $L : \mathbb{R}^n \to \mathbb{R}^n$*;*

iv) $\mathcal{H}^s(E) = 0$ *for any* $s > n$*;*

v) let $0 \le s < t < +\infty$*, then*

$$\mathcal{H}^s(E) < \infty \implies \mathcal{H}^t(E) = 0;$$

$$\mathcal{H}^t(E) > 0 \implies \mathcal{H}^s(E) = +\infty;$$

vi) if $\mathcal{H}^s_\infty(E) = 0$*, then* $\mathcal{H}^s(E) = 0$*.*

The above properties make it reasonable to define the **Hausdorff dimension** of sets as

$$\dim_{\mathcal{H}}(E) := \inf \{0 \le s < \infty : \mathcal{H}^s(E) = 0\}$$

Exercise. *Show that the Cantor set in* $[0, 1]$ *has Hausdorff dimension* $\log 2 / \log 3$*.*

3.2 Area formula

Theorem 3.4. *Suppose* $f : \mathbb{R}^n \to \mathbb{R}^m$ *is an* **injective Lipschitz** *map, and* $1 \le n \le m$*. Then for any Lebesgue-measurable set* $E \subset \mathbb{R}^n$*,*

$$\mathcal{H}^n(f(E)) = \int_E Jf(x)\, dx, \tag{18}$$

where $Jf : \mathbb{R}^n \to [0, \infty]$ *is the Jacobian of* f *and defined as*

$$Jf(x) = \begin{cases} \sqrt{\det\left(\nabla f(x)^* \nabla f(x)\right)}, & \text{if } f \text{ is differentiable at } x, \\ +\infty, & \text{if } f \text{ is not differentiable at } x. \end{cases} \tag{19}$$

Remark 3.5. By Rademacher's theorem, every Lipschitz function is differentiable almost everywhere. So the Jacobian of f is finite almost everywhere.

Exercise. *It is not too difficult to prove the following preliminary estimate*

$$\mathcal{H}^n(f(E)) \le \operatorname{Lip}(f)^n |E|.$$

[3] Another downside is that $\mathcal{H}^s_\infty$ is not a Borel regular measure, unlike $\mathcal{H}^s$.

Remark 3.6. When f is not assumed to be injective, we need to take into account the multiplicities in order to write down an analogous area formula:

$$\int_{\mathbb{R}^m} \mathcal{H}^0 \left(E \cap \{f = y\}\right) d\mathcal{H}^n(y) = \int_E Jf(x)\, dx.$$

Proof of the area formula. *Step 1. Area formula for linear maps.* We aim to show that the area formula (18) holds for any linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ with $1 \le n \le m$.

By the polar decomposition theorem in linear algebra, the linear transformation L can be decomposed into a scaling and an orthogonal transformation. Namely, there exist a symmetric map $S : \mathbb{R}^n \to \mathbb{R}^n$ and an orthogonal map $O : \mathbb{R}^n \to \mathbb{R}^m$ such that $L = O \circ S$. Note that the Jacobian of L is a constant

$$JL = \sqrt{\det L^* L} = |\det S|,$$

which exactly measures how much the linear map L distorts the area/volume of cubes Q

$$\mathcal{H}^n(L(Q)) = \mathcal{H}^n(S(Q)) = |\det S| \cdot |Q|.$$

The statement for general measurable sets E

$$\frac{\mathcal{H}^n(L(E))}{|E|} = \frac{\mathcal{H}^n(L(Q))}{|Q|} = |\det S|$$

follows from applying the Radon-Nikodym theorem for Radon measures, see [EG, Theorem 1.30].

Step 2. Approxination of Lipschitz maps by linear maps. Consider the set

$$F := \{x \in \mathbb{R}^n : 0 < Jf(x) < \infty\}. \tag{20}$$

Note that for any $x \in F$, by the polar decomposition we have $\nabla f(x) = O_x \circ S_x$, where $O_x : \mathbb{R}^n \to \mathbb{R}^m$ is an orthogonal map and $S_x : \mathbb{R}^n \to \mathbb{R}^n$ is a symmetric map, such that $|\det S_x| = Jf(x) \in (0, \infty)$. In particular $S_x \in \mathrm{GL}(n)$. Using a dense subset of $\mathrm{GL}(n)$, we aim to prove the following.

Proposition 3.7. *For every $t > 1$, the set F defined in (20) has a partition into Borel sets $\{F_k\}_k$ such that for each $k \in \mathbb{N}$ there is an invertible linear map $S_k \in \mathrm{GL}(n)$ such that $f|_{F_k} \approx S_k$. More precisely, for every $x, y \in F_k$ and $v \in \mathbb{R}^n$,*

$$t^{-1}|S_k x - S_k y| \le |f(x) - f(y)| \le t|S_k x - S_k y|, \tag{21}$$

$$t^{-1}|S_k v| \le |\nabla f(x) v| \le t|S_k v|, \tag{22}$$

$$t^{-n} JS_k \le Jf(x) \le t^n JS_k. \tag{23}$$

Remark 3.8. The first estimate (21) is equivalent to saying

$$\mathrm{Lip}(f|_{F_k} \circ (S_k)^{-1}) \le t, \qquad \mathrm{Lip}(S_k \circ (f|_{F_k})^{-1}) \le t \tag{24}$$

on their respective domains of definitions.

Assuming Proposition 3.7, we can prove the area formula for any set $E \subset F$. On one hand,

$$\begin{aligned}
\mathcal{H}^n(f(E)) = \sum_k \mathcal{H}^n(f(E \cap F_k)) &= \sum_k \mathcal{H}^n\left((f|_{F_k} \circ (S_k)^{-1}) \circ S_k(E \cap F_k)\right) \\
&\leq \sum_k \mathrm{Lip}(f|_{F_k} \circ (S_k)^{-1})^n \cdot |S_k(E \cap F_k)| \\
&\leq t^n \sum_k JS_k \cdot |E \cap F_k| \qquad \text{by (24) and Step 1,} \\
&\leq t^{2n} \sum_k \int_{E \cap F_k} Jf(x)\, dx \qquad \text{by (23),} \\
&= t^{2n} \int_E Jf(x)\, dx;
\end{aligned}$$

on the other hand,

$$\begin{aligned}
\int_E Jf(x)\, dx = \sum_k \int_{E \cap F_k} Jf(x)\, dx &\leq t^n \sum_k JS_k \cdot |E \cap F_k| \\
&= t^n \sum_k |S_k(E \cap F_k)| \\
&= t^n \sum_k \left|(S_k \circ (f|_{F_k})^{-1}) \circ f(E \cap F_k)\right| \\
&\leq t^n \sum_k \mathrm{Lip}(S_k \circ (f|_{F_k})^{-1})^n \cdot \mathcal{H}^n(f(E \cap F_k)) \\
&\leq t^{2n} \sum_k \mathcal{H}^n(f(E \cap F_k)) \\
&= t^{2n} \mathcal{H}^n(f(E)).
\end{aligned}$$

The area formula follows by passing $t \to 1$.

Now we are motivated to return to **the proof of Proposition 3.7**. To give us some wiggle room, let $\epsilon > 0$ be fixed such that $t^{-1} + \epsilon < 1 < t - \epsilon$. The space $\mathrm{GL}(n)$ has a countable dense subset, denoted by $\mathcal{G}$. For each $S \in \mathcal{G}$ and $j \in \mathbb{N}$, let $F(S, j)$ denote the set of points $x \in F$ such that

$$|f(x+v) - f(x) - \nabla f(x)v| \leq \epsilon |Sv|, \quad \text{for any } v \in \mathbb{R}^n \text{ such that } |v| < 1/j, \tag{25}$$

$$(t^{-1} + \epsilon)|Sv| \leq |\nabla f(x)v| \leq (t - \epsilon)|Sv|, \quad \text{for any } v \in \mathbb{R}^n. \tag{26}$$

The set $F(S, j)$ may well be empty, but when $S \in \mathcal{G}$ is fixed, as j increases the set clearly gets larger. In fact we claim that

$$F = \bigcup_{S \in \mathcal{G}, j \in \mathbb{N}} F(S, j). \tag{27}$$

This is because for each $x \in F$, the linear map $\nabla f(x)$ has the polar decomposition as above $\nabla f(x) = O_x \circ S_x$, where $S_x \in \mathrm{GL}(n)$. Let δ be sufficiently small (depending on the value of $|(S_x)^{-1}|$), and let $S \in \mathcal{G}$ satisfy $|S - S_x| < \delta$. We can guarantee that

$$|S_x \circ S^{-1}| = |(S_x - S) \circ S^{-1} + \mathrm{Id}\,| \leq |S_x - S||S^{-1}| + 1 \leq t - \epsilon,$$

and similarly

$$|S \circ (S_x)^{-1}| \le |S - S_x||S_x^{-1}| + 1 \le (t^{-1} + \epsilon)^{-1}.$$

Therefore for any $v \in \mathbb{R}^n$

$$|\nabla f(x)v| = |O_x \circ S_x(v)| = |S_x(v)| < (t - \epsilon)|Sv|,$$

and

$$|Sv| = |(S \circ (S_x)^{-1}) \circ S_x v| \le (t^{-1} + \epsilon)^{-1}|S_x v| = (t^{-1} + \epsilon)^{-1}|\nabla f(x)v|.$$

This proves the estimates in (26). By the definition of $\nabla f(x)$ when f is differentiable, we know

$$|f(x+v) - f(x) - \nabla f(x)v| \le o(1)|v| \le o(1)|S^{-1}| \cdot |Sv|.$$

By choosing $|v|$ sufficiently small (depending on the point x), the right hand side can be chosen to be bounded by $\epsilon|Sv|$.[4] We finish proving the claim (27).

Moreover, by letting $\{z_i\}_i$ be a countable dense subset of $F \subset \mathbb{R}^n$, we also have

$$F = \bigcup_{S \in \mathcal{G}, i, j \in \mathbb{N}} \left(F(S, j) \cap B(z_i, 1/2j)\right).$$

Hence for any $x, y \in F(S, j) \cap B(z_i, 1/2j)$, they satisfy (26) and thus (22) with $S \in \mathrm{GL}(n)$; and by letting $v := y - x$, since $|v| = |y - x| < 1/j$, the estimate (25) applies to show

$$|f(y) - f(x)| \le |\nabla f(x)v| + \epsilon|Sv| \le t|Sv| = t|Sy - Sx|,$$

and

$$|f(y) - f(x)| \ge |\nabla f(x)v| - \epsilon|Sv| \ge t^{-1}|Sv| = t^{-1}|Sy - Sx|.$$

They are the desirable estimates in (21). The last estimates (23) follows from (22). □

It remains to study the singular set $F_0 := \{x \in \mathbb{R}^n : Jf(x) = 0\}$.

Step 3. The nullity of the set $f(F_0)$. Assume without loss of generality that F_0 is bounded (if not we simply replace F_0 by its intersection with large enough balls that eventually exhaust $\mathbb{R}^n$). For any $x \in F_0$, since $|\det S_x| = Jf(x) = 0$, we have that $P_x := \nabla f(x)(\mathbb{R}^n) = O_x(S_x(\mathbb{R}^n))$ is a linear subspace in $\mathbb{R}^m$ of dimension $k \le n-1$, where k is the rank of S_x.

Since f is differentiable at x, for any $\epsilon > 0$ there exists $r_0(\epsilon, x) \in (0, 1]$ such that

$$|f(x+v) - f(x) - \nabla f(x)v| \le \epsilon|v|, \qquad \text{whenever } v \in \mathbb{R}^n \text{ satisfies } |v| < r_0(\epsilon, x). \quad (28)$$

In particular, for any $r \le r_0(\epsilon, x)$ we have

$$f(B^n(x, r)) \subset f(x) + \mathcal{N}_{\epsilon r}\left(\nabla f(x)(B^n(0, r)\right),$$

where $\mathcal{N}_\delta(A)$ denotes the δ-tubular neighborhood of a set A in $\mathbb{R}^m$

$$\mathcal{N}_\delta(A) = \{x \in \mathbb{R}^m : \mathrm{dist}(x, A) \le \delta\}.$$

Besides, by the notation above we also know that

$$\nabla f(x)(B^n(0, r)) \subset P_x \cap B^m(0, \mathrm{Lip}(f)r),$$

and the right hand side is the intersection of a ball and a k-dimensional linear subspace, $k \le n - 1 < m$. To estimate the $\mathcal{H}^n$-measure of $\mathcal{N}_{\epsilon r}\left(P_x \cap B^m(0, \mathrm{Lip}(f)r)\right)$, we observe the following.

[4]Note that by choosing S sufficiently close to S_x, we can always guarantee $|S^{-1}| \le 2|(S_x)^{-1}|$.

Exercise. *Let A be a k-dimensional disk in $\mathbb{R}^m$ of radius s,*

$$A = \{(x, y) \in \mathbb{R}^k \times \mathbb{R}^{m-k} : |x| < s, y = 0\},$$

with $k \leq n - 1 < m$. Then its δs-tubular neighborhood

$$\mathcal{N}_{\delta s}(A) \subset \{(x, y) \in \mathbb{R}^k \times \mathbb{R}^{m-k} : |x| < (1 + \delta)s, |y| < \delta s\}$$

satisfies

$$\mathcal{H}^n_\infty(\mathcal{N}_{\delta s}(A)) \leq C\delta s^n,$$

by counting the number of δs-balls needed to cover $\mathcal{N}_{\delta s}(A)$ and using the assumption $k \leq n - 1$. Here the constant C depends on the dimensions.

The exercise implies that

$$\mathcal{H}^n_\infty(f(B^n(x, r)) \leq \mathcal{H}^n_\infty\left(\mathcal{N}_{\epsilon r}\left(P_x \cap B^m(0, \mathrm{Lip}(f)r)\right)\right) \leq C\epsilon r^n,$$

where the constant C also depends on $\mathrm{Lip}(f)$. Note that the following collection of sets

$$\mathcal{F} := \{B^n(x, r)) : x \in F_0, 0 < r \leq r_0(\epsilon, x)\}$$

form a Vitali covering of F_0, from which we can extract a disjoint family $\{B^n(x_i, r_i)\}_i \subset \mathcal{F}$ such that

$$F_0 \subset \bigsqcup_i B^n(x_i, r_i), \qquad \text{except for a set of } \mathcal{H}^n\text{-measure zero.}$$

(See Vitali covering lemma in [WZ, Theorem 7.17], for example.) This allows us to estimate

$$\mathcal{H}^n_\infty(f(F_0)) \leq \sum_i \mathcal{H}^n_\infty(f(B^n(x_i, r_i)) \leq C\epsilon \sum_i r_i^n = \frac{C\epsilon}{\omega_n} \sum_i |B^n(x_i, r_i)| \lesssim \epsilon |\mathcal{N}_1(F_0)| < \infty.$$

In the last two inequalities we use the assumption that $r_0(\epsilon, x)$ is uniformly bounded by 1, and that F_0 is bounded. Since ϵ can be chosen arbitrarily, it follows that $\mathcal{H}^n_\infty(f(F_0)) = 0$, and thus $\mathcal{H}^n(f(F_0)) = 0$. Combined with Step 2, this finishes the proof of the area formula (18).

In the special case of the graph of a function, we immediately have the following expression of the area formula.

Corollary 3.9. *Let $u : \Omega \to \mathbb{R}$ be a Lipschitz function. Then for any measurable set $E \subset \Omega$,*

$$\mathcal{H}^n(\mathrm{Graph}(u; E)) = \int_E \sqrt{1 + |\nabla u(x)|^2}\, dx. \tag{29}$$

Proof. Consider the injective Lipschitz map $f : \mathbb{R}^n \to \mathbb{R}^{n+1}$ defined as $f(x) = (x, u(x))$. We have $\nabla f(x) = (\mathrm{Id}, \nabla u(x))^T$, and thus

$$\det((\nabla f(x))^* \nabla f(x)) = \det(\mathrm{Id} + \nabla u(x) \otimes \nabla u(x)) = 1 + |\nabla u(x)|^2.$$

Therefore $Jf(x) = \sqrt{1 + |\nabla u(x)|^2}$, and the area formula (18) implies

$$\mathcal{H}^n(\mathrm{Graph}(u; E)) = \mathcal{H}^n(f(E)) = \int_E Jf(x)\, dx = \int_E \sqrt{1 + |\nabla u(x)|^2}\, dx.$$

$\square$

3.3 Second variational formula

Let $\Sigma \subset \mathbb{R}^{n+1}$ be an n-dimensional minimal surface. Consider a variation of Σ by the map $F : (-\epsilon, \epsilon) \times \Sigma \to \mathbb{R}^{n+1}$ such that $F_0 = \mathrm{Id}$ and for each t, F_t is equal to the identity map outside of a compact set. Let X and Z denote the initial velocity and acceleration vectors of F_t, namely

$$\frac{d}{dt}\Big|_{t=0} F_t = X, \qquad \frac{d^2}{dt^2}\Big|_{t=0} F_t = Z.$$

Then for $|t| \ll 1$ we have

$$F_t(x) = x + tX_x + \frac{t^2}{2} Z_x + O(t^3). \tag{30}$$

By the area formula

$$\mathrm{Area}(F_t(\Sigma)) := \mathcal{H}^n(F_t(\Sigma)) = \int_\Sigma JF_t(x)\, d\mathcal{H}^n(x).$$

Next, we compute the Jacobian of F_t. For any tangent vector $\tau \in T_x\Sigma$, by (30) we have

$$dF_t\big|_x(\tau) = \tau + t\nabla_\tau X + \frac{t^2}{2}\nabla_\tau Z + O(t^3).$$

Let $\{\tau_1, \cdots, \tau_n\}$ denote a local orthonormal basis of Σ, and $\{e_1, \cdots, e_{n+1}\}$ denote an orthonormal basis for $\mathbb{R}^{n+1}$. Then the map $dF_t\big|_x$ has matrix representation

$$\left(dF_t\big|_x\right)_{ij} = \tau_j^i + t\nabla_{\tau_j} X^i + \frac{t^2}{2}\nabla_{\tau_j} Z^i + O(t^3).$$

Hence $\left(dF_t\big|_x\right)^* dF_t\big|_x$ has entries

$$\begin{aligned}
b_{ij} &= \sum_{k=1}^{n+1} \left(dF_t\big|_x\right)_{ki} \left(dF_t\big|_x\right)_{kj} \\
&= \delta_{ij} + t\left[\langle \tau_i, \nabla_{\tau_j} X\rangle + \langle \tau_j, \nabla_{\tau_i} X\rangle\right] + t^2\left[\langle \nabla_{\tau_i} X, \nabla_{\tau_j} X\rangle + \frac{1}{2}\langle \tau_i, \nabla_{\tau_j} Z\rangle + \frac{1}{2}\langle \tau_j, \nabla_{\tau_i} Z\rangle\right] + O(t^3),
\end{aligned}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{R}^{n+1}$. Using the formula

$$\det(\mathrm{Id} + tA + t^2 B) = 1 + t\,\mathrm{Tr}\, A + t^2\left(\mathrm{Tr}\, B + \frac{1}{2}(\mathrm{Tr}\, A)^2 - \frac{1}{2}\mathrm{Tr}\, A^2\right) + O(t^3),$$

and $\mathrm{div}_\Sigma X := \sum_i \langle \tau_i, \nabla_{\tau_i} X\rangle$, we have that

$$\begin{aligned}
(JF_t)^2 &= 1 + 2t\,\mathrm{div}_\Sigma X \\
&\quad + t^2\left[\sum_i |\nabla_{\tau_i} X|^2 + \mathrm{div}_\Sigma Z + 2(\mathrm{div}_\Sigma X)^2 - \frac{1}{2}\sum_{i,j}\left(\langle \tau_i, \nabla_{\tau_j} X\rangle + \langle \tau_j, \nabla_{\tau_i} X\rangle\right)^2\right] + O(t^3) \\
&= 1 + 2t\,\mathrm{div}_\Sigma X \\
&\quad + t^2\left[\sum_i |\left(\nabla_{\tau_i} X\right)^\perp|^2 + \mathrm{div}_\Sigma Z + 2(\mathrm{div}_\Sigma X)^2 - \sum_{i,j}\langle \tau_i, \nabla_{\tau_j} X\rangle\langle \tau_j, \nabla_{\tau_i} X\rangle\right] + O(t^3),
\end{aligned}$$

where in the last line, we have used $(\nabla_{\tau_i} X)^\perp$ to denote the normal part of $\nabla_{\tau_i} X$, i.e.

$$(\nabla_{\tau_i} X)^\perp = \nabla_{\tau_i} X - \sum_j \langle \nabla_{\tau_i} X, \tau_j \rangle \tau_j.$$

Finally, we obtain

$$JF_t = 1 + t\,\mathrm{div}_\Sigma X + \frac{t^2}{2}\left(\sum_i |(\nabla_{\tau_i} X)^\perp|^2 + \mathrm{div}_\Sigma Z + (\mathrm{div}_\Sigma X)^2 - \sum_{i,j} \langle \tau_i, \nabla_{\tau_j} X\rangle \langle \tau_j, \nabla_{\tau_i} X\rangle \right) + O(t^3). \tag{31}$$

Plugging (31) into the area formula, we conclude that

$$\frac{d}{dt}\Big|_{t=0} \mathrm{Area}(F_t(\Sigma)) = \int_\Sigma \mathrm{div}_\Sigma X \, d\mathcal{H}^n, \tag{32}$$

$$\frac{d^2}{dt^2}\Big|_{t=0} \mathrm{Area}(F_t(\Sigma)) = \int_\Sigma \left(\sum_i |(\nabla_{\tau_i} X)^\perp|^2 + \mathrm{div}_\Sigma Z + (\mathrm{div}_\Sigma X)^2 - \sum_{i,j} \langle \tau_i, \nabla_{\tau_j} X\rangle \langle \tau_j, \nabla_{\tau_i} X\rangle \right) d\mathcal{H}^n. \tag{33}$$

The stationarity of Σ implies that $\int_\Sigma \mathrm{div}_\Sigma X \, d\mathcal{H}^n = 0$ for every compactly supported vector field X. In particular we also have $\int_\Sigma \mathrm{div}_\Sigma Z \, d\mathcal{H}^n = 0$.

On the other hand, since Σ has codimension 1 in the ambient space $\mathbb{R}^{n+1}$, assuming it is orientable, we can reduce and simplify the above formula of vector fields into one of functions.[5] Let $X = f\nu$ where ν is the unit normal vector to Σ and $f \in C_c^\infty(\Sigma)$. We have

$$\mathrm{div}_\Sigma X = \sum_i \langle \tau_i, \nabla_{\tau_i} X\rangle = f \sum_i \langle \tau_i, \nabla_{\tau_i} \nu \rangle = -f \sum_i A_{ii} = -fH,$$

where $A_{ij} = -\langle \tau_i, \nabla_{\tau_j} \nu\rangle$ denote the entries of the second fundamental form of Σ, and $H = \mathrm{Tr}\, A$ denotes the mean curvature of Σ. Since this holds for all test functions f, in particular we have that $H = 0$. Moreover, since $\nabla_{\tau_i} \nu \perp \nu$, we have

$$(\nabla_{\tau_i} X)^\perp = ((\nabla_{\tau_i} f)\nu + f \nabla_{\tau_i} \nu)^\perp = (\nabla_{\tau_i} f)\nu. \tag{34}$$

Hence (33) is simplified to the following second variational formula

$$\frac{d^2}{dt^2}\Big|_{t=0} \mathrm{Area}(F_t(\Sigma)) = \int_\Sigma \left(|\nabla_\Sigma f|^2 - |A|^2 f^2 \right) d\mathcal{H}^n, \tag{35}$$

where $\nabla_\Sigma f$ denotes the tangential derivative of f on Σ, i.e.

$$\nabla_\Sigma f := P_{T_x\Sigma}(\nabla f) = \nabla f - \langle \nabla f, \nu\rangle \nu = \sum_i \langle \nabla f, \tau_i \rangle \tau_i,$$

and $|A|^2 = \sum_{i,j} A_{ij}^2$. By (35), the stability of the minimal surface Σ is equivalent to

$$(f, Lf) := \int_\Sigma \left(|\nabla_\Sigma f|^2 - |A|^2 f^2 \right) d\mathcal{H}^n \geq 0, \qquad \text{for every } f \in C_c^\infty(\Sigma), \tag{36}$$

[5] If Σ has higher codimensions, the tangential derivative of the normal vector may not lie entirely on the tangent space, so the expression (34) is no longer true, for example.

where the self-adjoint operator $L = -\Delta_\Sigma - |A|^2$ is called the Jacobi operator (or the stability operator).

In particular, applying the constant vector field e_j to Σ clearly does not change the area of Σ. Namely, the normal variation F_t with initial velocity

$$\frac{d}{dt}\Big|_{t=0} F_t = \langle e_j, \nu\rangle \nu$$

satisfies $\mathrm{Area}(F_t(\Sigma)) \equiv \mathrm{Area}(\Sigma)$. So the second variational formula implies[6]

$$(\nu_j, L\nu_j) = 0.$$

In fact for any minimal surface Σ, $\nu_j = \langle e_j, \nu\rangle$ belongs to the kernel of the Jacobi operator (referred to as the Jacobi field in the literature), i.e. $L\nu_j = 0$ for all j. We will give a PDE proof of this fact later when Σ is a minimal graph; but the second variation formula justifies why the elliptic operator $\Delta_\Sigma + |A|^2$ is meaningful on minimal surfaces.

As a result of the second variational formula, we can prove the following result about the Simons' cone.

Theorem 3.10. *The cone*

$$C := C^{p,p} = \{(x,y) \in \mathbb{R}^p \times \mathbb{R}^p : |x| = |y|\} \subset \mathbb{R}^{2p} \tag{37}$$

is stationary for all p, and is stable if and only if $p \geq 4$.

Proof. We first compute the second fundamental form of C. For any point (x,y) on C with norm r,

$$\nu_{(x,y)} = \frac{(x,-y)}{r}$$

is the unit normal vector to C at (x,y). When $p > 1$ the tangent space at (x,y) can be decomposed as follows

$$T_{(x,y)}C = \{(u,0) \in \mathbb{R}^p \times \mathbb{R}^p : u \perp x\} \bigcup \{(0,v) \in \mathbb{R}^p \times \mathbb{R}^p : v \perp y\} \bigcup \{(\lambda x, \lambda y) : \lambda \in \mathbb{R}\}.$$

It is clear that the second fundamental form in the radial direction is trivial, since

$$\nabla_{(\frac{x}{r},\frac{y}{r})}\nu \text{ points to the normal direction.}$$

For a unit tangent vector $(u,0) \in T_{(x,y)}C$, i.e. $u \perp x$ and $|u| = 1$ in $\mathbb{R}^p$, we have

$$\nabla_{(u,0)}\nu = \left(\frac{u}{r}, 0\right).$$

This implies that the principal curvature of C in the $(u,0)$-direction is $1/r$. Similarly, for a unit tangent vector $(0,v) \in T_{(x,y)}C$, we have

$$\nabla_{(0,v)}\nu = \left(0, -\frac{v}{r}\right),$$

and thus the principal curvature in the $(0,v)$-direction is $-1/r$. Therefore the mean curvature H vanishes, and the second fundamental form satisfies

$$|A|^2 = (p-1)\left(\frac{1}{r}\right)^2 + (p-1)\left(-\frac{1}{r}\right)^2 = \frac{2(p-1)}{r^2}. \tag{38}$$

[6] Note that this observation is true regardless of the stability of the minimal surface Σ.

In particular C is minimal/stationary.

We study the second variational formula using the separation of variables. Let $\Sigma := C \cap \mathbb{S}^{2p-1}$ denote the link of the cone. Then

$$|\nabla_C f|^2 = f_r^2 + \frac{1}{r^2}|\nabla_\Sigma f|^2.$$

Suppose $f(r\theta) = \phi(r)u(\theta)$, where $r > 0$ and $\theta \in \Sigma$. By a weighted Hardy's inequality, the second variation formula becomes

$$\begin{aligned}
Q(f,f) &= \int_C |\nabla_C f|^2 - |A_C|^2 f^2 d\mathcal{H}^{2p-1} \\
&= \int_0^\infty \phi_r^2 r^{2p-2} dr \cdot \int_\Sigma u^2 d\mathcal{H}^{2p-2} + \int_0^\infty \phi^2 r^{2p-4} dr \cdot \int_\Sigma |\nabla_\Sigma u|^2 - |A_\Sigma|^2 u^2 d\mathcal{H}^{2p-2} \\
&\geq \frac{(2p-3)^2}{4} \int_0^\infty \phi^2 r^{2p-4} dr \cdot \int_\Sigma u^2 d\mathcal{H}^{2p-2} + \int_0^\infty \phi^2 r^{2p-4} dr \cdot \int_\Sigma |\nabla_\Sigma u|^2 - |A_\Sigma|^2 u^2 d\mathcal{H}^{2p-2} \\
&\geq \int_0^\infty \phi^2 r^{2p-4} dr \cdot \left[\int_\Sigma |\nabla_\Sigma u|^2 + \left(\frac{(2p-3)^2}{4} - 2(p-1) \right) u^2 d\mathcal{H}^{2p-2} \right],
\end{aligned}$$

which is non-negative whenever $p \geq 4$. Hence C is stable whenever $p \geq 4$. Since the above inequality is sharp (i.e. equality in the weighted Hardy's inequality can be approximated by a sequence of test functions ϕ), we also show the instability of C when $p < 4$. $\square$

4 Interior gradient estimate for the MSE

We will prove the interior gradient estimate in Theorem 2.6, which we restate here:

Theorem 4.1. *Let $u \in C^2(B_{2R})$ be a solution to the MSE in B_{2R}. Then for any $x \in B_R$, we have*

$$|\nabla u(x)| \leq C_1 \exp\left(C_2 \frac{\sup_{B_{2R}} u - u(x)}{R} \right), \tag{39}$$

where C_1, C_2 are universal constants.

Step 1. Let Σ denote Graph(u). Since Σ is a minimal surface in $\mathbb{R}^{n+1}$, we **claim** that its unit normal derivative $\boldsymbol{\nu}$ satisfies the equation $\Delta_\Sigma \boldsymbol{\nu} + |A_\Sigma|^2 \boldsymbol{\nu} = 0$. (Recall that $L = \Delta_\Sigma + |A_\Sigma|^2$ is the stability/Jacobi operator introduced in Section 3.3.) In particular when Σ is given by the graph of u, one of its unit normal derivative is given by

$$\boldsymbol{\nu} = \frac{(-\nabla u, 1)}{\sqrt{1 + |\nabla u|^2}}. \tag{40}$$

Thus $\nu_{n+1} = 1/\sqrt{1 + |\nabla u|^2} > 0$ is in the kernel of L.[7] Consider the function

$$w = -\log(\nu_{n+1}) = \log(\sqrt{1 + |\nabla u|^2}). \tag{41}$$

It satisfies

$$\Delta_\Sigma w = -\operatorname{div}_\Sigma \left(\frac{\nabla_\Sigma \nu_{n+1}}{\nu_{n+1}} \right) = -\frac{\Delta_\Sigma \nu_{n+1}}{\nu_{n+1}} + \frac{|\nabla \nu_{n+1}|^2}{\nu_{n+1}^2} = |A_\Sigma|^2 + |\nabla_\Sigma w|^2.$$

[7]Note that $\Delta_\Sigma \boldsymbol{\nu} + |A_\Sigma|^2 \boldsymbol{\nu} = 0$ holds for any minimal surface. However, as long as one of the components of $\boldsymbol{\nu}$ has a sign, locally Σ can be written as a graph pointing in the direction of that component. So already in Step 1, the proof relies on the assumption that Σ is graphical.

In particular

$$\Delta_\Sigma w \geq |\nabla_\Sigma w|^2 \geq 0, \tag{42}$$

i.e. w is a sub-harmonic function on Σ.

Step 2. We **claim** that any non-negative sub-harmonic function w on a minimal surface Σ satisfies the sub-mean value property:

$$w(0) \leq \frac{1}{\omega_n R^n} \int_{\Sigma \cap B_R} w \, d\mathcal{H}^n. \tag{43}$$

Here we state the mean value property centered at $0 \in \Sigma$ for the simplicity of notation[8], and we always assume $0 < 3R < \operatorname{dist}(0, \partial\Sigma)$. In fact, we will show that the ratio on the right hand side of (43) is monotone non-decreasing with respect to R. Since the ratio tends to $w(0)$ as $R \searrow 0$, (43) immediately follows.

Assuming the above two claims are true, we will prove the interior gradient estimate using PDE arguments in the spirit of De Giorgi's proof of the L^∞ estimate in Theorem 2.2. By the definition of w in (41), to get a pointwise estimate of $|\nabla u|$ it suffices to the estimate the integral on the RHS of (43). To fix notation, let $h = \sqrt{1 + |\nabla u|^2}$. By the area formula (29), h measures the change of volume from $\Omega \subset \mathbb{R}^n$ to $\Sigma = \operatorname{Graph}(u)$; besides, $1/h$ is also the coefficient in the divergence-form MSE.

Step 3. Note that

$$\begin{aligned}
\int_{\Sigma \cap B_R} w \, d\mathcal{H}^n \leq \int_{|x|,|u|<R} wh \, dx &= \int_{|x|,|u|<R} \frac{w(1 + |\nabla u|^2)}{h} \, dx \\
&= \int_{|x|,|u|<R} \frac{w}{h} \, dx + \int_{|x|,|u|<R} \frac{w|\nabla u|^2}{h} \, dx \\
&\leq \omega_n R^n + \int_{|x|,|u|<R} \frac{w|\nabla u|^2}{h} \, dx, \qquad \text{since } 0 \leq w = \log h \leq h.
\end{aligned}$$

So we need to estimate $\int_{|x|,|u|<R} \frac{w|\nabla u|^2}{h} \, dx$.

Since it involves $|\nabla u|^2/h$, we use the minimal surface equation for u. To that purpose, we plug in the test function $\varphi = u_R w \eta$, where u_R is defined as

$$u_R = \begin{cases} 2R, & \text{if } u > R, \\ u + R, & \text{if } |u| \leq R, \\ 0, & \text{if } u < -R, \end{cases}$$

and $\eta \in C_c^\infty(B_{2R})$ satisfies $0 \leq \eta \leq 1$, $\eta \equiv 1$ on $B_R \subset \mathbb{R}^n$ and $|\nabla \eta| \leq C/R$. From the weak formulation

$$\int \frac{\langle \nabla u, \nabla \varphi \rangle}{h} \, dx = 0,$$

we deduce a Cacciopolli-type inequality[9]

$$\int_{|x|,|u|<R} \frac{w|\nabla u|^2}{h} \, dx \leq C \int_{|x|<2R, u>-R} w \, dx + CR \int_{|x|<2R, u>-R} \eta |\nabla w| \, dx. \tag{44}$$

[8] In other words, we assume that $0 \in \Omega$ and $u(0) = 0$.

[9] Here we also use the simple observation that for any vector $\boldsymbol{a}$, we have $|\langle \nabla u, \boldsymbol{a} \rangle|/h \leq |\nabla u|/h \cdot |\boldsymbol{a}| \leq |\boldsymbol{a}|$. Hence the terms on the RHS do not contain $|\nabla u|$.

Step 4. Since the second term on the RHS of (44) involves ∇w, we need to use the equation of w to estimate it. Note that (42) is a PDE on Σ, we need to first rewrite it as an integral over Σ and change ∇w into $\nabla_\Sigma w$.

Firstly we observe that for any function $g \in C^1(\mathbb{R}^{n+1})$ in the ambient space,

$$\nabla_i g \equiv 0 \text{ on } \Sigma \implies |\nu_i||\nabla g| \le |\nabla_\Sigma g|.$$

This is because $|\langle \nabla g, \boldsymbol{\nu}\rangle|^2 = |\sum_{j\neq i} \partial_j g\, \nu_j|^2 \le (1-\nu_i^2)|\nabla g|^2$, and thus $|\nabla_\Sigma g|^2 = |\nabla g|^2 - |\langle \nabla g, \boldsymbol{\nu}\rangle|^2 \ge |\nu_i|^2 |\nabla g|^2$. Applying to the function w, which is independent of the x_{n+1}-variable, we have that

$$\int_{|x|<2R, u>-R} \eta|\nabla w|\, dx \le \int_{\Sigma\cap\cdots} \eta|\nabla w| \frac{d\mathcal{H}^n}{\sqrt{1+|\nabla u|^2}} \le \int_\Sigma \phi|\nabla_\Sigma w| d\mathcal{H}^n. \tag{45}$$

Here $\phi(x, x_{n+1}) = \eta(x)\tau(x_{n+1}) \ge 0$ is a test function in the ambient space $\mathbb{R}^{n+1}$, with the function $\tau \in C_c^\infty(-2R, R + \sup_{B_{2R}} u)$ satisfying $0 \le \tau \le 1$, $\tau \equiv 1$ on $[-R, \sup_{B_{2R}} u]$ and $|\tau'| \le C/R$.

To estimate the last term, we plug the non-negative test function ϕ^2 into the weak formulation of (42) and get

$$\int_\Sigma |\nabla_\Sigma w|^2 \phi^2 d\mathcal{H}^n \le -2\int_\Sigma \phi\langle \nabla_\Sigma w, \nabla_\Sigma \phi\rangle.$$

Applying the Cauchy-Schwartz inequality to the right hand side, we obtain

$$\int_\Sigma |\nabla_\Sigma w|^2 \phi^2 d\mathcal{H}^n \le 4\int_\Sigma |\nabla_\Sigma \phi|^2 d\mathcal{H}^n \le CR^{-2}\mathcal{H}^n(\Sigma \cap \operatorname{supp}\phi).$$

Hence

$$\int_\Sigma \phi|\nabla_\Sigma w| d\mathcal{H}^n \le CR^{-1}\mathcal{H}^n(\Sigma \cap \operatorname{supp}\phi). \tag{46}$$

Step 5. It remains to estimate $\mathcal{H}^n(\Sigma \cap \operatorname{supp}\phi) \le \int_{|x|<2R, u>-2R} h\, dx$. We remark that since $w \le h$, the first term on the RHS of (44) is also bounded by it.

Note that

$$\frac{|\nabla u|^2}{h} = \frac{1+|\nabla u|^2}{h} - \frac{1}{h} = h - \frac{1}{h},$$

and $1/h \le 1$. The integral estimate of h is reduced to that of $|\nabla u|^2/h$, which is very similar and in fact simpler than the estimate (44) in Step 3. Plugging the test function $\varphi = (u+2R)_+\eta$ into the MSE (here we need to choose $\eta \equiv 1$ when $|x| < 2R$ and that η is supported in $|x| < 3R$), we get[10]

$$\int_{|x|<2R, u>-2R} \frac{|\nabla u|^2}{h} dx \le \int \frac{|\nabla \eta|(u+2R)_+}{h} dx \le CR^n(1 + \sup_{|x|<3R} u/R).$$

Hence

$$\int_{|x|<2R, u>-2R} h dx \le \int_{|x|<2R, u>-2R} \left(1 + \frac{|\nabla u|^2}{h}\right) dx \le CR^n\left(1 + \frac{\sup_{|x|<3R} u}{R}\right).$$

[10] Alternatively, we could also use the test function $\varphi = u\eta$, in which case the upper bound will be $CR^n \cdot \sup_{B_{3R}} |u|/R$,where we need the boundedness of u both from above and from below.

In other words, we have shown

$$\frac{\mathcal{H}^n(\Sigma \cap \operatorname{supp} \phi)}{\omega_n R^n} \leq C\left(1 + \frac{\sup_{|x|<3R} u}{R}\right). \tag{47}$$

Combining the estimates (45), (46), (47) and plugging into (44), it follows that

$$\int_{|x|,|u|<R} \frac{w|\nabla u|^2}{h} \, dx \leq CR^n\left(1 + \frac{\sup_{|x|<3R} u}{R}\right).$$

Therefore

$$\log(\sqrt{1+|\nabla u(0)|^2}) = w(0) \leq \frac{1}{\omega_n R^n}\int_{\Sigma\cap B_R} w d\mathcal{H}^n \leq C\left(1 + \frac{\sup_{|x|<3R} u}{R}\right),$$

and finally

$$|\nabla u(0)| \leq C_1 \exp\left(C_2 \sup_{|x|<3R} u/R\right)$$

assuming that $u(0) = 0$. This finishes the proof of (39).

To prove the claims in Step 1 and 2, we introduce some notation. Recall that the tangential derivative on Σ is

$$\nabla_\Sigma f = \nabla f - \langle \nabla f, \boldsymbol{\nu}\rangle \boldsymbol{\nu},$$

where ∇f is the gradient of f in the ambient space $\mathbb{R}^{n+1}$. Let δ_i denote the i-th compoment of ∇_Σ, for $i \in \{1, \cdots, n+1\}$, i.e.

$$\delta_i f := \partial_i f - \langle \nabla f, \boldsymbol{\nu}\rangle \nu_i,$$

where $\partial_i f$ is the directional derivative of f in the e_i-direction. Then

$$\Delta_\Sigma f = \operatorname{div}_\Sigma(\nabla_\Sigma f) = \delta_i\delta_i f.$$

Under this notation, the MSE is equivalent to

$$\delta_i \nu_i = 0. \tag{48}$$

In fact, by (40) and $|\boldsymbol{\nu}| = 1$ we have

$$\begin{aligned}
\delta_i\nu_i = \partial_i\nu_i - (\partial_j\nu_i)\nu_j\nu_i &= \sum_{i=1}^n \partial_i\left(\frac{-\partial_i u}{\sqrt{1+|\nabla u|^2}}\right) - \langle \partial_j \boldsymbol{\nu}, \boldsymbol{\nu}\rangle \nu_j \\
&= -\operatorname{div}\left(\frac{\nabla u}{\sqrt{1+|\nabla u|^2}}\right) - \partial_j \frac{|\boldsymbol{\nu}|^2}{2}\nu_j \\
&= 0.
\end{aligned}$$

We also observe the following properties:

i) $\nu_i\delta_i f = \langle \boldsymbol{\nu}, \nabla_\Sigma f\rangle = 0$;

ii) $\delta_i\nu_j = \delta_j\nu_i$ for any pair $i, j \in \{1, \cdots, n+1\}$ (this is the same as saying $\nabla_\Sigma \boldsymbol{\nu}$ is a symmetric matrix);

iii) $\delta_i\delta_j f - \delta_j\delta_i f = (\nu_i\delta_j\nu_\ell - \nu_j\delta_i\nu_\ell)\delta_\ell f$.

Therefore

$$\begin{aligned}\Delta_\Sigma\nu_k = \delta_i\delta_i\nu_k &= \delta_i\delta_k\nu_i && \text{by ii)}\\ &= \delta_k\delta_i\nu_i + (\nu_i\delta_k\nu_\ell - \nu_k\delta_i\nu_\ell)\delta_\ell\nu_i && \text{by iii)}\\ &= 0 + \delta_k\nu_\ell\cdot\nu_i\delta_i\nu_\ell - \nu_k|\delta_i\nu_\ell|^2 && \text{by (48) and ii)}\\ &= -|A_\Sigma|^2\nu_k && \text{by i).}\end{aligned}$$

We remark that for general hypersurfaces Σ, the above equality shows that $\Delta_\Sigma\boldsymbol{\nu}+|A_\Sigma|^2\boldsymbol{\nu}+\nabla_\Sigma\boldsymbol{H}=0$.

To prove the **sub-mean value property** (43) for non-negative sub-harmonic functions, we consider a radial test function $\psi(r)=\psi(|x|)$. We have

$$\delta_i\psi(r) = \left(\frac{x_i}{r} - \nu_i\nu_j\frac{x_j}{r}\right)\psi'(r),$$

and by the MSE (48),

$$\begin{aligned}\Delta_\Sigma\psi(r) = \delta_i\delta_i\psi(r) &= \left(\frac{x_i}{r} - \nu_i\nu_j\frac{x_j}{r}\right)\delta_i\psi'(r) + \left(\delta_i\frac{x_i}{r} - \delta_i\left(\nu_i\nu_j\frac{x_j}{r}\right)\right)\psi'(r)\\ &= \left(1-\frac{\langle x,\boldsymbol{\nu}\rangle^2}{r^2}\right)\psi''(r) + \left(\frac{n-1}{r}+\frac{\langle x,\boldsymbol{\nu}\rangle^2}{r^3}\right)\psi'(r)\\ &= \frac{n\psi'(r)}{r} + \left(\psi''(r)-\frac{\psi'(r)}{r}\right)\left(1-\frac{\langle x,\boldsymbol{\nu}\rangle^2}{r^2}\right)\\ &= \frac{n\psi'(r)}{r} + \left(\psi''(r)-\frac{\psi'(r)}{r}\right)\left|\left(\frac{x}{r}\right)^\Sigma\right|^2,\end{aligned}$$

where for each unit vector u we use the notation $u^\Sigma := u - \langle u,\boldsymbol{\nu}\rangle\boldsymbol{\nu}$ to denote its projection onto Σ. Since w is a non-negative sub-harmonic function on Σ, by the weak formulation we have

$$\int_\Sigma w\Delta_\Sigma\psi(r)\,d\mathcal{H}^n \ge 0, \tag{49}$$

if the test function $\psi\ge 0$.

Let $\chi:\mathbb{R}_+\to\mathbb{R}_+$ be a non-negative, non-increasing test function supported in $[0,1]$, and let $\rho>0$. Letting

$$\psi(r) = \int_r^\infty \tau\chi\left(\frac{\tau}{\rho}\right)d\tau,$$

we have $\psi'(r) = -r\chi(\frac{r}{\rho})$ and $\psi''(r)-\frac{\psi'(r)}{r} = -\frac{r}{\rho}\chi'(\frac{r}{\rho})$. Plugging into (49), we get

$$\begin{aligned}0 &\ge \int_\Sigma\left[n\chi\left(\frac{r}{\rho}\right)+\frac{r}{\rho}\chi'\left(\frac{r}{\rho}\right)\left|\left(\frac{x}{r}\right)^\Sigma\right|^2\right]w\,d\mathcal{H}^n\\ &= \int_\Sigma \rho^{n+1}\frac{d}{d\rho}\left[\frac{1}{\rho^n}\chi\left(\frac{r}{\rho}\right)\right]w - \frac{r}{\rho}\chi'\left(\frac{r}{\rho}\right)\left[1-\left|\left(\frac{x}{r}\right)^\Sigma\right|^2\right]w\,d\mathcal{H}^n.\end{aligned}$$

Hence

$$\rho^{n+1}\frac{d}{d\rho}\left[\frac{1}{\rho^n}\int_\Sigma w\chi\left(\frac{r}{\rho}\right)d\mathcal{H}^n\right] \le \int_\Sigma\frac{r}{\rho}\chi'\left(\frac{r}{\rho}\right)\left[1-\left|\left(\frac{x}{r}\right)^\Sigma\right|^2\right]w\,d\mathcal{H}^n \le 0.$$

In other words, the map

$$\rho \mapsto \frac{1}{\rho^n}\int_\Sigma w\chi\left(\frac{r}{\rho}\right)\, d\mathcal{H}^n$$

is monotone non-increasing. By approximating the characteristic function with χ, it follows that the map

$$\rho \mapsto \frac{1}{\omega_n\rho^n}\int_{\Sigma\cap B_\rho} w\, d\mathcal{H}^n$$

is monotone non-increasing; and it is a constant if and only if

$$1 - \left|\left(\frac{x}{r}\right)^\Sigma\right|^2 \equiv 1 \quad \text{for all } x\in\Sigma,$$

in other words $\langle x, \boldsymbol{\nu}\rangle \equiv 0$ and that Σ is a cone.

In particular, $w\equiv 1$ is clearly a non-negative sub-harmonic function, and thus the density ratio

$$\rho \mapsto \frac{\mathcal{H}^n(\Sigma\cap B_\rho(x_0))}{\omega_n\rho^n} =: \theta_\Sigma(x_0,\rho) \tag{50}$$

is monotone non-increasing; and it is a constant if and only if Σ is a cone with vertex at x_0. This monotonicity formula will be very useful later.

We have finished proving the interior gradient estimate in Theorem 2.6. By the discussions in Section 2, higher regularity immediately follows:

Corollary 4.2. *Suppose Ω is a domain in $\mathbb{R}^n$ and $u\in C^2(\Omega)$ is a solution to the MSE. Then $u\in C^\infty(\Omega)$, and for any $x\in\Omega$ and any multi-index $\boldsymbol{\beta}$ with $|\boldsymbol{\beta}| = k\in\mathbb{N}$, we have*

$$|D^{\boldsymbol{\beta}}u(x)| \le C$$

where $C = C(n,k,\sup|u|, \operatorname{dist}(x,\partial\Omega))$.

Combined with Theorem 2.1, the Bernstein theorem assuming bounded gradients, we obtain the following analogy of the Liouville's theorem for harmonic functions:

Corollary 4.3. *Suppose $u\in C^2(\mathbb{R}^n)$ is a global solution to the MSE. Assume u is bounded on one side by a cone, i.e.*

$$u(x)\le C(1+|x|) \qquad \text{for all } x\in\mathbb{R}^n \tag{51}$$

for some constant $C>0$. Then u must be linear.

Impressionistic sketch of what's to come next. As Corollary 4.2 says, solutions to the MSE are smooth on the interior; and Corollary 4.3 shows that non-trivial global solutions have super-linear growth at infinity. This motivates the following blow-down argument: For any $x_0\in\Sigma$ and $R\ll 1$, we consider the rescaling of Σ as

$$\Sigma_R := \frac{\Sigma - x_0}{R},$$

and study the limit of Σ_R as $R\nearrow +\infty$. In general there's no reason why the limit exists, but assume that the sequential limit exists, that is: for any sequence $R_k\to+\infty$, modulo passing to a subsequence Σ_{R_k} *converges* to some Σ_∞ (which is often referred to as *a tangent of* Σ *at* ∞); and also assume if Σ is a minimal hypersurface (resp. minimizing), the corresponding tangent Σ_∞ is also minimal (resp. minimizing). Then morally speaking, for any $\rho > 0$ we have (assuming a strong enough notion of *convergence*)

$$\theta_{\Sigma_\infty}(\rho) = \lim_{R_k \to +\infty} \theta_{\Sigma_{R_k}}(\rho) = \lim_{R_k \to +\infty} \theta_\Sigma(x_0, R_k\rho) \equiv \theta_\Sigma(x_0, +\infty).$$

By the rigidity of the monotonicity formula (50), it follows that Σ_∞ must be a cone; and Σ_∞ is a hyperplane if and only if Σ is also a hyperplane.

What does this blowdown process mean for the function u? If $\Sigma = \mathrm{Graph}(u)$, then $\Sigma_{R_k} = \mathrm{Graph}(u_{R_k})$ where

$$u_{R_k}(y) = \frac{u(R_k y)}{R_k},$$

and u_{R_k} also solves the MSE. Formally speaking, let

$$v(y) := \lim_{R_k \to +\infty} u_{R_k}(y) = \lim_{R_k \to +\infty} \frac{u(R_k y)}{R_k}.$$

If u were to violate (51) from above or below somewhere, we must have that

$$P := \{y \in \mathbb{R}^n : v(y) = +\infty\} \neq \emptyset, \quad Q := \{y \in \mathbb{R}^n : v(y) = -\infty\} \neq \emptyset.$$

This shows the limitation of restricting to minimal graphs, since this class is not closed after taking the limit and v is not a finite-valued function. Assuming that P and Q account for the entire domain $\mathbb{R}^n$, then Σ_∞, being the boundary of the sub-graph of v, corresponds to $C \times \mathbb{R}$, where $C = \partial P = \partial Q$ in $\mathbb{R}^n$. Moreover, C is not a halfplane; and Σ_∞ is minimizing in $\mathbb{R}^{n+1}$ implies that C is minimizing in $\mathbb{R}^n$. On the other hand, we have discussed that (at least among Simons cones) non-flat cones $C \subset \mathbb{R}^n$ are area-minimizing if and only if $n \geq 8$. This explains the significance of the dimension; and it also implies that non-trivial global solutions must blow down to $\Sigma_\infty = C \times \mathbb{R}$, and knowing what non-flat minimizing cones C there are indicates what the behavior of non-trivial global solutions near ∞.

In the second part of the topic course, we will (largely depart from PDEs and) introduce the necessary tools to make the above discussion rigorous.

5 Compactness among sets of finite perimeter

As we explained in the synopsis, the family of minimal graphs will not be closed (unless perhaps if we include the graph of functions taking values in the extended reals $\mathbb{R} \cup \{\pm\infty\}$). Instead, we will work with the more general class of sets of finite perimeter, which are essentially measurable sets in $\mathbb{R}^n$ whose boundaries can be *measured*.[11]

Definition 5.1. Let E be a measurable set in $\mathbb{R}^n$. We say that E is a **set of (locally) finite perimeter** in $\mathbb{R}^n$ if

$$\sup\left\{\int_E \operatorname{div} X : X \in C_c^1(\mathbb{R}^n; \mathbb{R}^n), |X| \leq 1\right\} < \infty.$$

[11] The materials of this section are mostly based on [Ma, Chapters 12 and 21]. We refer interested readers to [Ma], especially for the illustrative pictures accompanying the proofs.

Proposition 5.2. *Let E be a measurable set in $\mathbb{R}^n$. Then E is a set of (locally) finite perimeter if and only if there exists a $\mathbb{R}^n$-valued Radon measure*[12] *μ_E on $\mathbb{R}^n$ such that*

$$\int_E \operatorname{div} X = \int_{\mathbb{R}^n} X \cdot d\mu_E, \qquad \text{for any } X \in C_c^1(\mathbb{R}^n; \mathbb{R}^n).$$

Moreover, under the above assumption, the total variation of μ_E satisfies $|\mu_E|(\mathbb{R}^n) < \infty$; and there exists an $|\mu_E|$-measurable function $\nu_E : \mathbb{R}^n \to \mathbb{R}^n$ with $|\nu_E| = 1$ a.e. such that

$$\mu_E = \nu_E |\mu_E|.$$

Proof. The proof follows from Riesz representation theorem (for bounded linear functional on the space $C_c^0(\mathbb{R}^n; \mathbb{R}^m)$), see [Ma, Theorem 4.7 and Proposition 12.1]. The total variation of the vector-valued Radon measure μ_E is defined as follows: For any open set $A \subset \mathbb{R}^n$,

$$|\mu_E|(A) := \sup \left\{ \int_E \operatorname{div} X : X \in C_c^1(A; \mathbb{R}^n), |X| \leq 1 \right\}.$$

We also define the (relative) perimeter of E in A as

$$P(E; A) := |\mu_E|(A), \qquad P(E) := |\mu_E|(\mathbb{R}^n).$$

□

The notion of sets of finite perimeter is motivated by the Gauss-Green theorem (divergence theorem): Suppose $E \subset \mathbb{R}^n$ is an open set with C^1 boundary, then for every $X \in C_c^1(\mathbb{R}^n; \mathbb{R}^n)$,

$$\int_E \operatorname{div} X = \int_{\partial E} X \cdot \nu_E \, d\mathcal{H}^{n-1}.$$

Clearly in this case, the map

$$X \in C_c^1(\mathbb{R}^n; \mathbb{R}^n) \mapsto \int_E \operatorname{div} X$$

is a bounded linear functional, which can be extended to $C_c^0(\mathbb{R}^n; \mathbb{R}^n)$ by density.

PDE-inclined people may prefer to think of sets of finite perimeter as a special case of **BV functions**: A function $f \in L^1$ has bounded variation in $\mathbb{R}^n$ if

$$\sup \left\{ \int f \operatorname{div} X : X \in C_c^1(\mathbb{R}^n; \mathbb{R}^n), |X| \leq 1 \right\} < \infty.$$

Roughly speaking, these are functions in L^1 whose derivatives are Radon measures, instead of integrable functions. This notion fixes the issue that $W^{1,1}$ is not compact; in fact it is the *completion* of the $W^{1,1}$ space.

Example Consider a sequence of cuntions on $(-1, 1)$ defined as

$$u_k(x) = \begin{cases} 0, & \text{if } -1 < x < 0 \\ kx, & \text{if } 0 \leq x \leq \frac{1}{k} \\ 1, & \text{if } \frac{1}{k} < x < 1 \end{cases}$$

[12] Radon measures are locally finite, Borel regular measures.

They are uniformly bounded in $L^1(-1,1)$, and their derivatives all have L^1 norm equal to 1. However, they do not (subsequentially) converge in $W^{1,1}(-1,1)$; in fact, their limit function

$$u(x) = \begin{cases} 0, & \text{if } -1 < x \le 0 \\ 1, & \text{if } 0 < x < 1 \end{cases}$$

has derivative δ_0, the Dirac delta function at 0, which is a Radon measure.

It's easy to see that E is a set of (locally) finite perimeter in $\mathbb{R}^n$ if and only if the characteristic function χ_E is a BV function.

To study the compactness, we introduce the following notion of convergence (for sets and Radon measures, respectively).

Definition 5.3. We say a sequence of measurable sets E_k converges to E, and write $E_k \to E$, if their set differences satisfy $\lim_{k\to\infty} |E \Delta E_k| = 0$.

Definition 5.4 (Convergence of Radon measures)**.** Let $\{\mu_k\}, \mu$ be Radon measures on $\mathbb{R}^n$ with values in $\mathbb{R}^m$. We say that $\mu_k \rightharpoonup \mu$ if

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} \varphi \cdot d\mu_k = \int_{\mathbb{R}^n} \varphi \cdot d\mu, \qquad \text{for any } \varphi \in C_c^0(\mathbb{R}^n; \mathbb{R}^m).$$

Exercise. *Let ρ_ϵ be a smooth approximation of the identity. If E is a set of finite perimeter in $\mathbb{R}^n$, then*

$$-\nabla(\chi_E * \rho_\epsilon) \rightharpoonup \mu_E.$$

Exercise. *Suppose $\mu_k \rightharpoonup \mu$. Then*

1. *If $U, K \subset \mathbb{R}^n$ with U open and K compact, then*
$$\mu(U) \le \liminf_{k\to\infty} \mu_k(U), \qquad \mu(K) \ge \limsup_{k\to\infty} \mu_k(K); \tag{52}$$
2. *If A is a bounded Borel set with $\mu(\partial A) = 0$, then $\lim_{k\to\infty} \mu_k(A) = \mu(A)$;*
3. *For every $x \in \operatorname{supp} \mu$, there exists a sequence $\{x_k\} \subset \mathbb{R}^n$ such that*
$$\lim_{k\to\infty} x_k = x, \quad x_k \in \operatorname{supp} \mu_k \text{ for every } k \in \mathbb{N}. \tag{53}$$

A family of sets of finite perimeter is compact, assuming a uniform bound on their perimeters, see [Ma, Theorem 12.26 and Proposition 12.25].

Theorem 5.5 (Compactness among sets of finite perimeter)**.** *Suppose $\{E_k\}$ are sets of finite perimeter in $\mathbb{R}^n$ such that*

$$\sup_k P(E_k) < \infty,$$

and there is a ball $B \subset \mathbb{R}^n$ such that each E_k is contained in B. Then there exist a set of finite perimeter, E, and a subsequence such that

$$E_k \to E, \quad \mu_{E_k} \rightharpoonup \mu_E, \quad E \subset B.$$

Moreover, the perimeter is lower semi-continuous: for every open set $A \subset \mathbb{R}^n$, we have

$$P(E; A) \le \liminf_{k\to\infty} P(E_k; A). \tag{54}$$

Remark 5.6. The assumption that E_k's are contained in the same ball is to pin them down; without it, one can have a sequence of unit balls going off the infinity. Alternatively, we can impose the *uniform perimeter bounds* assumption locally on compact sets, and obtain a similar conclusion: If $\{E_k\}$ are sets of locally finite perimeter in $\mathbb{R}^n$ such that

$$\sup_k P(E_k; B_R) < \infty, \qquad \text{for any } R > 0,$$

then there exist a set of locally finite perimeter E and a subsequence such that

$$E_k \to E \text{ on any compact set}, \qquad \mu_{E_k} \rightharpoonup \mu_E.$$

Proof. The main ingredient of the proof is to show the following. Consider the complete metric space

$$\mathcal{X} := \{E \subset \mathbb{R}^n : E \text{ is Lebesgue-measurable with } |E| < \infty\}$$

endowed with the metric

$$d(E, F) := |E \Delta F|.$$

(Here we identify two sets $E, F \in \mathcal{X}$ if $|E\Delta F| = 0$.) We claim that for any ball $B \subset \mathbb{R}^n$ and any $p > 0$, the set

$$\mathcal{X}_{B,p} := \{E \in \mathcal{X} : E \subset B, P(E) \leq p\}$$

is compact. To prove the compactness, we just need to show it is closed and totally bounded. It is closed, because of the lower semi-continuity of the perimeter. The condition that $\mathcal{X}_{B,p}$ is totally bounded means for any $\epsilon > 0$, there exists finitely many sets $T_1, \cdots, T_K \in \mathcal{X}$ such that for any $E \in \mathcal{X}_{B,p}$,

$$d(E, T_i) \leq \epsilon \qquad \text{for some } i \in \{1, \cdots, K\}.$$

To prove this, we first observe that: Let E be a set of finite perimeter with $|E| < \infty$, then for every $r > 0$, there exists a finite union T of disjoint cubes of side length r such that

$$|E\Delta T| \leq \sqrt{n} r P(E).$$

See the details in the proof of [Ma, Theorem 12.26]. □

Sets of finite perimeter provide the framework to study the minimization problem under the volume constraint.

Theorem 5.7 (Isoperimetric inequality). *If E is a measurable set in $\mathbb{R}^n$ with $|E| < \infty$, then*

$$P(E) \geq n\omega_n^{\frac{1}{n}} |E|^{\frac{n-1}{n}}. \tag{55}$$

Equality holds if and only if $|E\Delta B| = 0$ for some ball $B \subset \mathbb{R}^n$.

It is also natural to consider the *relative isoperimetric problem* in a set $A \subset \mathbb{R}^n$, namely

$$\inf\{P(E; A) : E \subset A, |E| = m\}$$

where the volume constraint $m \in (0, |A|)$. In the case that A is a ball, one can prove:

Theorem 5.8. *Let $t \in (0,1)$. There exists a constant $c(n,t) > 0$ such that for any ball B and any set of locally finite perimeter E such that $|E \cap B| \le t|B|$, we have*

$$P(E;B) \ge c(n,t)|E \cap B|^{\frac{n-1}{n}}. \tag{56}$$

In particular, the inequality with $t = 1/2$ implies that

$$P(E;B) \ge c(n) \min\{|E \cap B|, |B \setminus E|\}^{\frac{n-1}{n}}. \tag{57}$$

Lastly, we have the following **structure theorem for sets of finite perimeter**.

Theorem 5.9. *If E is a set of locally finite perimeter in $\mathbb{R}^n$, then there is a set $\partial^* E \subset \operatorname{supp}\mu_E \subset \partial E$ called the* reduced boundary *of E, such that*

$$\mu_E = \nu_E \mathcal{H}^{n-1}|_{\partial^* E}, \qquad |\mu_E| = \mathcal{H}^{n-1}|_{\partial^* E}.$$

Thus the generalized Gauss-Green formula *holds:*

$$\int_E \operatorname{div} X = \int_{\partial^* E} X \cdot \nu_E \, d\mathcal{H}^{n-1}, \qquad \textit{for any } X \in C_c^1(\mathbb{R}^n;\mathbb{R}^n).$$

Moreover $\partial^* E$ is $(n-1)$-rectifiable, meaning that

$$\partial^* E = N \cup \bigsqcup_{i \in \mathbb{N}} \Sigma_i,$$

where Σ_i is a compact subset of a C^1-hypersurface M_i in $\mathbb{R}^n$, $\mathcal{H}^{n-1}(N) = 0$, and for any $x \in \Sigma_i$, $\nu_E(x)$ is a unit normal vector to the hypersurface M_i at x.

Proof. See [Ma, Section 15.2]. In particular, the reduced boundary consists of points x near which χ_E blows up to the characteristic function of a halfspace

$$H_x := \{z \in \mathbb{R}^n : \langle z, \nu_E(x)\rangle \le 0\}.$$

□

Exercise. *1) For the set of finite perimeter $E = B(0,1) \setminus$ the positive x_1-axis in $\mathbb{R}^2$, we have*

$$\partial^* E = \partial B(0,1) \subsetneq \partial E.$$

In fact, we have that $\mu_E = \mu_F$ where $F = B(0,1)$ and $|E \Delta F| = 0$. (From the viewpoint the Gauss-Green formula, the sets E and F are indistinguishable, even though $\partial E \supsetneq \partial F$.)

2) For the set of finite perimeter $E =$ a unit square in $\mathbb{R}^2$, $\partial^* E = \partial E \setminus$ the four vertices.

3) Let E be the complement in $\mathbb{R}^2$ of the four-corner Cantor set. Somewhat counter-intuitively, E is also a set of finite perimeter, but with $\partial^* E = \emptyset$.

The above examples show that despite the nice characterization of μ_E and the reduced boundary in Theorem 5.9, there are gaps in the inclusion below $\partial^* E \subset \partial E$. For $t \in [0,1]$, let

$$E^{(t)} := \left\{ x \in \mathbb{R}^n : \lim_{r \to 0} \frac{|E \cap B(x,r)|}{\omega_n r^n} \text{ exists and is equal to } t \right\}.$$

By the Lebesgue density theorem, we have $E = E^{(1)}$ and $\mathbb{R}^n \setminus E = E^{(0)}$ modulo sets of Lebesgue measure zero. By the characterization of the reduced boundary in Theorem 5.9 above, we know that $\partial^* E \subset E^{(1/2)}$. In fact, we also have the density characterization for $\mathcal{H}^{n-1}$:

Theorem 5.10 (Federer's Theorem). *If E is a set of (locally) finite perimeter in $\mathbb{R}^n$, then $\partial^* E \subset E^{(1/2)} \subset \mathbb{R}^n \setminus (E^{(0)} \cup E^{(1)})$, with*

$$\mathcal{H}^{n-1}((\mathbb{R}^n \setminus (E^{(0)} \cup E^{(1)})) \setminus \partial^* E) = 0.$$

Proof. By the relative isoperimetric inequality (57) and the trivial bound $|E \cap B(x,r)| \leq \omega_n r^n$, we have

$$\frac{P(E; B(x,r))}{r^{n-1}} \geq c(n) \min \left\{ \frac{|E \cap B(x,r)|}{r^n}, \frac{|B(x,r) \setminus E|}{r^n} \right\}.$$

Hence if $x \notin E^{(0)} \cup E^{(1)}$, the lower density of $|\mu_E| = \mathcal{H}^{n-1}|_{\partial^* E}$ at x is strictly positive, and thus

$$(\mathbb{R}^n \setminus (E^{(0)} \cup E^{(1)})) \setminus \partial^* E \subset \left\{ x \in \mathbb{R}^n \setminus \partial^* E : \theta^*_{n-1}(\mathcal{H}^{n-1}|_{\partial^* E})(x) > 0 \right\}.$$

The analogue of the Lebesgue density theorem for $\mathcal{H}^s$, see [Ma, Corollary 6.5], implies that the latter set is $\mathcal{H}^{n-1}$-negligible. $\square$

Definition 5.11. Let E be a set of (locally) finite perimeter, and A be a bounded open set in $\mathbb{R}^n$. We say that E is a **perimeter minimizer** in A if $\operatorname{supp} \mu_E = \partial E$ and

$$P(E; A) \leq P(F; A), \qquad \text{whenever } E \Delta F \subset\subset A.$$

We say E is a perimeter minimizer in $\mathbb{R}^n$ is it is a perimeter minimizer in any bounded open set $A \subset \mathbb{R}^n$.

Remark 5.12. The assumption $\operatorname{supp} \mu_E = \partial E$ is not too restrictive: In fact, for any set of (locally) finite perimeter E, there exists a Borel set F such that

$$|E \Delta F| = 0, \qquad \operatorname{supp} \mu_F = \partial F.$$

To prove the compactness of perimeter minimizers, there are two issues we need to resolve: firstly, prove a uniform upper bound on the perimeters in any ball, so we can appeal to the compactness result for sets of finite perimeter in Theorem 5.5; secondly, show that the limit also minimizes the perimeter and this is a closed family.

Theorem 5.13 (Density estimates for perimeter minimizers). *There exists a constant $c(n) > 0$ with the following property. If A is an open set in $\mathbb{R}^n$, and E is a perimeter minimizer in A, then for any $x \in \partial E \cap A$ and $r < \operatorname{dist}(x, \partial A)$, we have*

$$\frac{1}{2^n} \leq \frac{|E \cap B(x,r)|}{\omega_n r^n} \leq 1 - \frac{1}{2^n}, \tag{58}$$

$$c(n) \leq \frac{P(E; B(x,r))}{r^{n-1}} \leq n\omega_n. \tag{59}$$

In particular,

$$\mathcal{H}^{n-1}((\partial E \setminus \partial^* E) \cap A) = 0. \tag{60}$$

Proof. Note that

$$\mathcal{H}^{n-1}(\partial^* E \cap \partial B(x,r)) = 0, \qquad \text{for a.e. } r < \operatorname{dist}(x, \partial A). \tag{61}$$

For every such r, let $F := E \setminus B(x,r)$. Since $E \Delta F \subset\subset B(x,s) \subset A$ for any $r < s < \operatorname{dist}(x, \partial A)$, we have

$$\begin{aligned} P(E; B(x,s)) &\leq P(F; B(x,s) = P(E \setminus B(x,r); B(x,s)) \\ &= P(E; B(x,s) \setminus \overline{B(x,r)}) + \mathcal{H}^{n-1}(E^{(1)} \cap \partial B(x,r)). \end{aligned}$$

Letting $s \to r+$, it follows from (61) that

$$P(E; B(x,r)) \leq \mathcal{H}^{n-1}(E^{(1)} \cap \partial B(x,r)) \leq P(B(x,r)) = n\omega_n r^{n-1}.$$

This proves the upper bound in (59).

It also implies

$$P(E \cap B(x,r)) = P(E; B(x,r)) + \mathcal{H}^{n-1}(E^{(1)} \cap \partial B(x,r)) \leq 2\mathcal{H}^{n-1}(E^{(1)} \cap \partial B(x,r)).$$

Thus by the isoperimetric inequality (55),

$$n\omega_n^{\frac{1}{n}} |E \cap B(x,r)|^{\frac{n-1}{n}} \leq P(E \cap B(x,r)) \leq 2\mathcal{H}^{n-1}(E^{(1)} \cap \partial B(x,r)).$$

Let $m(r) := |E^{(1)} \cap B(x,r)| = |E \cap B(x,r)|$. We have

$$n\omega_n^{\frac{1}{n}} m(r)^{\frac{n-1}{n}} \leq 2m'(r), \qquad \text{for a.e. } r.$$

Integrating this inequality, we get $m(r) \geq \omega_n r^n / 2^n$, which is the lower bound in (58). Applying the same argument to $\mathbb{R}^n \setminus E$, which is also a perimeter minimizer, gives the upper bound in (58).

The lower bound in (59) follows from the relative isoperimetric inequality in the ball, see (57). And lastly, (60) follows from Federer's Theorem 5.10. $\square$

Theorem 5.14 (Pre-compactness for perimeter minimizers)**.** *Suppose $\{E_k\}$ is a sequence of perimeter minimizers in the open set $A \subset \mathbb{R}^n$. Then for every open set $A_0 \subset\subset A$ with $P(A_0) < \infty$, there exists a set of finite perimeter $E \subset A_0$ and a subsequence such that*

$$E_k \cap A_0 \to E, \qquad \mu_{E_k \cap A_0} \rightharpoonup \mu_E.$$

Proof. By the proof of the density estimates, for any $x \in A_0$ and $r > 0$ such that $E_k \cap B(x,r) \neq \emptyset$, we have

$$P(E_k \cap B(x,r)) \leq P(E_k; B(x,r)) + P(B(x,r)) \leq n\omega_n (2r)^{n-1} + n\omega_n r^{n-1},$$

uniformly in k. Therefore we may apply Theorem 5.5 to extract a convergent subsequence. $\square$

Lastly, we show that the limit E above also minimizes the perimeter.

Theorem 5.15. *Suppose $\{E_k\}$ is a sequence of perimeter minimizers in the open set $A \subset \mathbb{R}^n$, and $A_0 \subset\subset A$ is an open set with $P(A_0) < \infty$ such that $E_k \cap A_0 \to E$ for a set of finite perimeter E. Then E is a* **perimeter minimizer** *in A_0.*

Moreover,

$$\mu_{E_k \cap A_0} \rightharpoonup \mu_E, \qquad |\mu_{E_k}| \rightharpoonup |\mu_E| \text{ in } A_0, \tag{62}$$

and that

i) if $x_k \in \partial E_k \cap A_0$, $x_k \to x \in A_0$, then $x \in \partial E$;

ii) if $x \in \partial E \cap A_0$, then there exists $\{x_k\} \subset \partial E_k \cap A_0$ such that $x_k \to x$.

Remark 5.16. In fact, near a regular point $x \in \partial^* E \cap A_0$, we can say more about the convergence: for k sufficiently large $x_k \in \partial^* E_k \cap A_0$ and that $\nu_{E_k}(x_k) \to \nu_E(x)$.

Proof. Step 1. Suppose F is a set of finite perimeter such that $E \Delta F \subset\subset A_0$, we aim to show that

$$P(E; A_0) \leq P(F; A_0).$$

To use the minimality of E_k's, we want to modify the set F so that it becomes a competitor of E_k. For any $x \in A_0$, we have that for a.e. $0 < r < r_x := \operatorname{dist}(x, \partial A_0)$

$$\mathcal{H}^{n-1}(\partial^* F \cap \partial B(x,r)) = 0 = \mathcal{H}^{n-1}(\partial^* E_k \cap \partial B(x,r)), \qquad \text{for all } k, \tag{63}$$

$$\liminf_{k\to\infty} \mathcal{H}^{n-1}((E^{(1)} \Delta E_k^{(1)}) \cap \partial B(x,r)) = 0. \tag{64}$$

The last equality follows from:

$$\begin{aligned} 0 = \lim_{k\to\infty} |(E \Delta E_k) \cap B(x, r_x)| &= \lim_{k\to\infty} |(E^{(1)} \Delta E_k^{(1)}) \cap B(x, r_x)| \\ &= \lim_{k\to\infty} \int_0^{r_x} \mathcal{H}^{n-1}((E^{(1)} \Delta E_k^{(1)}) \cap \partial B(x,r))\, dr \\ &\geq \int_0^{r_x} \liminf_{k\to\infty} \mathcal{H}^{n-1}((E^{(1)} \Delta E_k^{(1)}) \cap \partial B(x,r))\, dr. \end{aligned}$$

Since $E \Delta F$ is compactly contained in A_0, it can be covered by finitely many such balls, i.e.

$$E \Delta F \subset\subset G := \bigcup_{i=1}^{N} B(x_i, r_i)$$

with $x_i \in A_0$ and $r_i < \operatorname{dist}(x_i, \partial A_0)$ such that (63) and (64) hold. In particular, it follows that

$$\begin{aligned} &\mathcal{H}^{n-1}(\partial^* F \cap \partial G) = 0 = \mathcal{H}^{n-1}(\partial^* E_k \cap \partial G), \qquad \text{for all } k, \\ &\liminf_{k\to\infty} \mathcal{H}^{n-1}((E^{(1)} \Delta E_k^{(1)}) \cap \partial G) = 0. \end{aligned} \tag{65}$$

Define a set of finite perimeter

$$F_k := (F \cap G) \sqcup (E_k \setminus G).$$

Then $E_k \Delta F_k \subset G \subset\subset A_0$, i.e. F_k is a competitor of E_k in the open set A_0. Hence

$$P(E_k; A_0) \leq P(F_k; A_0) \leq P(F; G) + P(E_k; A_0 \setminus \overline{G}) + \mathcal{H}^{n-1}(\partial G \cap (F^{(1)} \Delta E_k^{(1)})),$$

which implies

$$P(E_k; G) \leq P(F; G) + \mathcal{H}^{n-1}(\partial G \cap (F^{(1)} \Delta E_k^{(1)})).$$

Passing $k \to \infty$ and by the lower semi-continuity of the perimeter (see (54)) and (65), we have

$$P(E; G) \leq \liminf_{k\to\infty} P(E_k; G) \leq P(F; G).$$

Since $E \Delta G \subset\subset G$, this shows that E minimizes the perimeter.

Modulo modifying E on a set of Lebesgue measure zero (which doesn't affect the convergence $E_k \cap A_0 \to E$ and the perimeter minimality of E), we can guarantee that $\partial E = \operatorname{supp} \mu_E$.

Step 2. The convergence of the vector-valued measures

$$\mu_{E_k \cap A_0} \rightharpoonup \mu_E \tag{66}$$

follows from $E_k \cap A_0 \to E$. By the density upper bound, their total variations also converge: there exists a Radon measure μ such that

$$|\mu_{E_k \cap A_0}| \rightharpoonup \mu. \tag{67}$$

In general, there could be cancellations in the vector-valued measures $\mu_{E_k \cap A_0}$ which lead to μ_E to have a smaller support; and we aim to prove that for perimeter minimizers E_k's, this does not happen and that $\mu|_{A_0} = |\mu_E|$ restricted to A_0 as desired.[13]

Firstly, we claim that the convergences of (66) and (67) always imply that $\mu \geq |\mu_E|$. For any bounded open set $V \subset \mathbb{R}^n$, let δ be sufficiently small and let φ be a smooth test function such that

$$\chi_{V_\delta} \leq \varphi \leq \chi_V, \qquad \text{where } V_\delta := \{x \in V : \operatorname{dist}(x, \partial V) > \delta\}.$$

By (67) and the definition of the total variations $|\mu_{E_k \cap A_0}|, |\mu_E|$, we have

$$\mu(V) \geq \int \varphi \, d\mu = \lim_{k\to\infty} \int \varphi \, d|\mu_{E_k \cap A_0}| \geq \limsup_{k\to\infty} |\mu_{E_k \cap A_0}|(V_\delta) \geq |\mu_E|(V_\delta).$$

By taking a sequence $\delta_j \to 0$ such that $|\mu_E|(\partial V_{\delta_j}) = 0$, we obtain that $|\mu_E|(V) \leq \mu(V)$, and thus $|\mu_E| \leq \mu$ on any Borel set.

On the other hand, let $x \in \operatorname{supp} \mu \cap A_0$ and $0 < r < r_x := \operatorname{dist}(x, \partial A_0)$. Arguing as in *Step 1*, we may consider the set

$$\tilde{F}_k := (E \cap B(x,r)) \sqcup (E_k \setminus B(x,r))$$

for a.e. $r < r_x$, which is a competitor for the minimizer E_k in A_0. Hence

$$P(E_k; B(x,r)) \leq P(E; B(x,r)) + \mathcal{H}^{n-1}(\partial B(x,r) \cap (E^{(1)} \Delta E_k^{(1)})).$$

Combined with (52) and the choice of a.e. r, we have

$$\begin{aligned} \mu(B(x,r)) \leq \liminf_{k\to\infty} |\mu_{E_k \cap A_0}|(B(x,r)) &= \liminf_{k\to\infty} P(E_k; B(x,r)) \\ &\leq P(E; B(x,r)) = |\mu_E|(B(x,r)). \end{aligned}$$

Combined with $|\mu_E| \leq \mu$, we conclude that these two measures agree on A_0.

Step 3. The statement ii) always follows from $|\mu_{E_k}| \rightharpoonup |\mu_E|$ and (53). To prove i), let $x_k \in \partial E_k \cap A_0$ such that $x_k \to x \in A$. Let $s > 0$ such that $B(x, 2s) \subset A_0$. Then for k sufficiently large $B(x_k, s) \subset B(x, 2s)$. Again by $|\mu_{E_k}| \rightharpoonup |\mu_E|$ and (52), we have

$$|\mu_E|(\overline{B(x,2s)}) \geq \limsup_{k\to\infty} |\mu_{E_k}|(\overline{B(x,2s)}) \geq \limsup_{k\to\infty} P(E_k; (B(x,s)) \geq c(n) s^{n-1},$$

by the density lower bound for perimeter minimizers (see (59)). In particular, it follows that $x \in \operatorname{supp} \mu_E = \partial E$. ☐

[13] For experts, the possible gap between $\mu := \lim |\mu_{E_k \cap A_0}|$ and $|\mu_E|$ is due to the discrepancy between the convergence of currents and varifold convergence.

In the proof of Theorems 5.13 and 5.5, we have heavily used the *minimality* assumption. The study of analogous compactness results among sets of finite perimeter which are merely stable (or have finite Morse index, or just stationary, respectively) is much more difficult and closely related to the regularity theory of minimal surfaces. We refer interested readers to [SS].

6 Asymptotic behavior of global minimal graphs

Now we're ready to combine all the ingredients to prove the following theorem, see [Si1, Theorem 2].

Theorem 6.1. *Suppose u is a C^2 solution of the MSE on $\mathbb{R}^n \setminus \Omega$, with Ω bounded. Then*

a) either $\nabla u(x)$ is bounded and has a limit as $|x| \to \infty$;

b) or else all tangent cones of $\operatorname{Graph}(u)$ at ∞ are cylinders of the form $C \times \mathbb{R}$, where C is a $(n-1)$-dimensional minimizing cone in $\mathbb{R}^n$ and $0 \in \operatorname{sing} C$.

Remark 6.2. Since there are no non-flat stable minimal cones in $\mathbb{R}^n$ when $n \leq 6$, the second alternative is impossible in low dimensions. This finishes the proof of the Bernstein theorem in low dimensions. In high dimensions, the second alternative also indicates what the behavior of non-trivial solutions should be near ∞.

Assume without loss of generality $u(0) = 0$. Recall by the interior gradient estimate, we have either $|\nabla u|$ is bounded, or else by applying Theorem 2.6 to u and $-u$,

$$\lim_{\rho_j \to \infty} \frac{\sup_{B_{\rho_j}} u}{\rho_j} = +\infty \quad \text{and} \quad \lim_{\rho_j \to \infty} \frac{\inf_{B_{\rho_j}} u}{\rho_j} = -\infty. \tag{68}$$

Blow-down procedure. As we discussed in the end of Section 4, the class of minimal graphs is not closed, since the interesting behavior happens when they tend to $\pm\infty$. Instead, we consider a minimal graph as the (reduced) boundary of a set of locally finite perimeter, which is the subgraph of u: Let

$$E = \operatorname{SGraph}(u) := \{(x, y) : x \in \mathbb{R}^n \setminus \Omega, y < u(x)\} \subset \mathbb{R}^n \times \mathbb{R}$$

be a set of finite perimeter in $\mathbb{R}^{n+1} \setminus \Omega \times \mathbb{R}$ and $\Sigma := \partial E = \operatorname{Graph}(u)$. By the interior gradient estimate E is actually a smooth domain with C^2 boundary, so in particular $\partial^* E = \partial E = \Sigma$. We consider rescalings of E by $R \gg 1$, namely

$$E_R := \frac{E}{R} = \operatorname{SGraph}(u_R), \quad \Sigma_R := \partial E_R = \operatorname{Graph}(u_R),$$

where $u_R(y) = u(Ry)/R$ for $y \in \mathbb{R}^n \setminus \Omega_R$ also solves the MSE, with $\Omega_R := \Omega/R$. Each E_R is a perimeter minimizer in $\mathbb{R}^{n+1}$ and passes through the origin. Then by the compactness proven in Section 5, for every sequence $R_k \to \infty$, modulo passing to a subsequence we have $E_k := E_{R_k} \to$ a perimeter minimizer E_∞, uniformly on compact sets in $\mathbb{R}^{n+1} \setminus \{0\} \times \mathbb{R}$. Moreover, we have

i) the generalized hypersurface in the limit $\Sigma_\infty := \partial E_\infty = \operatorname{supp} \mu_{E_\infty}$ satisfies $\mathcal{H}^n(\Sigma_\infty \setminus \partial^* E_\infty) = 0$, by (60);

ii) the Gauss-Green measures satisfy $\mu_{E_k} \rightharpoonup \mu_{E_\infty}$ and $|\mu_{E_k}| = \mathcal{H}^n|_{\partial E_k} \rightharpoonup |\mu_{E_\infty}| =: \mu$;

iii) $\Sigma_k \to \Sigma_\infty$ in the Hausdorff distance sense, again by Theorem 5.15.

Our goal is to prove that $\Sigma_\infty = C \times \mathbb{R}$ for some cone $C \subset \mathbb{R}^n$. Then the minimality of C follows from the perimeter minimality of E_∞. To that end, we study the properties of Σ_∞.

By the De Giorgi/Allard regularity theorem, for any $p \in \Sigma_\infty$ with density 1, i.e.

$$\lim_{r\to 0} \frac{\mathcal{H}^n(\Sigma_\infty \cap B(p,r))}{\omega_n r^n} = 1,$$

there is a neighborhood around p such that Σ_∞ is a $C^{1,\alpha}$ minimal graph over some n-dimensional hyperplane. As we have seen in Section 1, the regularity of Σ_∞ can be upgraded to C^∞. This indicates that any $p \in \partial^* E_\infty$ is a regular point, denoted as $\operatorname{reg}\Sigma_\infty$. Combined with i), we know that the singular set $\operatorname{sing}\Sigma_\infty := \Sigma_\infty \setminus \operatorname{reg}\Sigma_\infty$ is a closed set with $\mathcal{H}^n$-measure zero.

By the second convergence in ii), for a.e. $\rho > 0$

$$\begin{aligned}\mu(B_\rho) = \lim_{R_k\to\infty} \mathcal{H}^n(\Sigma_k \cap B_\rho) &= \lim_{R_k\to\infty} \frac{\mathcal{H}^n(\Sigma \cap B_{R_k\rho})}{R_k^n} \\ &= \lim_{R_k\to\infty} \frac{\mathcal{H}^n(\Sigma \cap B_{R_k\rho})}{\omega_n(R_k\rho)^n} \cdot \omega_n\rho^n =: \Theta(\Sigma,\infty)\cdot\omega_n\rho^n. \qquad (69)\end{aligned}$$

Hence by the monotonicity formula (for critical point of the area functional among sets of finite perimeter), we conclude that Σ_∞ is a cone, i.e. $\lambda\Sigma_\infty = \Sigma_\infty$ for all $\lambda > 0$. We call Σ_∞ a tangent cone to Σ at ∞. In particular, by the classification of stable minimal cones, as well as the minimality of Σ_∞, we have that Σ_∞ is a hyperplane if $n+1 \le 7$.

By the first convergence in ii), we know that for any $X \in C_c^1(\mathbb{R}^{n+1};\mathbb{R}^{n+1})$,

$$\int_{\Sigma_k} X \cdot \boldsymbol{\nu}_{\boldsymbol{E_k}} d\mathcal{H}^n \to \int_{\operatorname{reg}\Sigma_\infty} X \cdot \boldsymbol{\nu}_{\boldsymbol{E_\infty}} d\mathcal{H}^n.$$

In particular let $X = \varphi e_{n+1}$ with φ being a test function, we get that

$$\int_{\Sigma_k} \varphi\langle \boldsymbol{\nu}_{\boldsymbol{E_k}}, e_{n+1}\rangle \to \int_{\operatorname{reg}\Sigma_\infty} \varphi\langle \boldsymbol{\nu}_{\boldsymbol{E_\infty}}, e_{n+1}\rangle. \qquad (70)$$

Since Σ_k is a graph, we always have $\langle \boldsymbol{\nu}_{\boldsymbol{E_k}}, e_{n+1}\rangle > 0$. Thus (70) implies that $\langle \boldsymbol{\nu}_{\boldsymbol{E_\infty}}, e_{n+1}\rangle \ge 0$ on $\operatorname{reg}\Sigma_\infty$. On the other hand, recalling the discussion about the second variational formula and that $\eta(x) := \langle \boldsymbol{\nu}_{\boldsymbol{E_\infty}}(x), e_{n+1}\rangle$ is a Jacobi field, namely $\Delta_{\Sigma_\infty}\eta + |A_{\Sigma_\infty}|^2\eta = 0$ on $\operatorname{reg}\Sigma_\infty$. In particular $\Delta_{\Sigma_\infty}\eta \le 0$, namely η is a non-negative super-harmonic function. Hence for any B that intersects $\operatorname{reg}\Sigma_\infty$, either $\eta \equiv 0$ on B or else $\inf_B \eta > 0$.

By (68), for any $\sigma > 0$, we have

$$\sup_{B_\sigma} u_{R_k} = \frac{\sup_{B_{\sigma R_k}} u}{R_k} \nearrow +\infty, \quad \inf_{B_\sigma} u_{R_k} = \frac{\inf_{B_{\sigma R_k}} u}{R_k} \searrow -\infty. \qquad (71)$$

Then by the maximum principle for quasilinear elliptic PDE we have

$$\max_{\partial B_\sigma} u_{R_k} \nearrow +\infty, \qquad \min_{\partial B_\sigma} u_{R_k} \searrow -\infty.$$

Hence for any $t > 0$, both $\{u_{R_k} > t\} \cap \partial B_\sigma$ and $\{u_{R_k} < t\} \cap \partial B_\sigma$ are non-empty for k sufficiently large. In particular there exists $z_k \in \partial B_\sigma$ such that $u_{R_k}(z_k) = t$. Modulo passing to a subsequence z_k converges to $z_t \in \partial B_\sigma$. Hence $(z_t, t) = \lim_{k\to\infty}(z_k, u_{R_k}(z_k) = t) \in \Sigma_\infty$ by Theorem 5.15. Since Σ_∞ is a cone, it follows that $(\lambda z_t, \lambda t) \in \Sigma_\infty$ for all $\lambda > 0$. In other words, for any $s > 0$, the point

$$(\frac{s}{t} z_t, s) = (\lambda z_t, \lambda t) \in \Sigma_\infty, \quad \text{where } \left|\frac{z_t}{t}\right| = \frac{\sigma}{t} \to 0 \text{ as } t \to +\infty.$$

This implies that $(0, s) \in \Sigma_\infty$ for all $s > 0$. Analogously we can also show that $(0, s) \in \Sigma_\infty$ for all $s < 0$. Therefore the line $L := \{te_{n+1} : t \in \mathbb{R}\} \subset \Sigma_\infty$.

Let $y \in L$ and B be a ball centered at y. Assume for contradiction that $\inf_B \eta > 0$. By the definition of η this implies that $\operatorname{reg}\Sigma_\infty$ is a Lipschitz weak solution to the MSE in B. This is a contradiction with the vertical line $L \subset \Sigma_\infty = \operatorname{supp}\mu_{E_\infty} = \overline{\operatorname{reg}\Sigma_\infty}$. Therefore $\eta = \langle \boldsymbol{\nu}_{\boldsymbol{E_\infty}}, e_{n+1}\rangle \equiv 0$ everywhere on $\operatorname{reg}\Sigma_\infty$. Since Σ_∞ has no boundary, this implies that it is invariant under translation in the e_{n+1}-direction, and therefore $E_\infty = U \times \mathbb{R}$ where U is a perimeter minimizing cone in $\mathbb{R}^n$, $\Sigma_\infty = C \times \mathbb{R}$ and $C = \partial U$.

It remains to show that C is not a hyperplane. This can be easily ruled out if $\Omega = \emptyset$ and that Σ is complete: If C is a hyperplane, then $\Sigma_\infty = C \times \mathbb{R}$ has density 1, and thus it follows by (69) that $\theta(\Sigma, \infty) = 1$. Since $\theta(\Sigma, x, 0+) = 1$ for any $x \in \operatorname{reg}\Sigma_\infty$, by the rigidity of the monotonicity formula Σ must also be a hyperplane, which falls into the first alternative.

When $\Omega \neq \emptyset$ we argue as follows. Suppose for the sake of contradiction that C is indeed a hyperplane H in $\mathbb{R}^n$. To fix notation we identify

$$\mathbb{R}^n = \{(x', x_n) : x' \in H, x_n \in \mathbb{R}\}.$$

Again by Allard's regularity theorem, if a tangent cone to Σ at ∞, given by $H \times \mathbb{R}_{n+1}$, is flat, then away from a compact set Σ is a $C^{1,\alpha}$-graph over $H \times \mathbb{R}_{n+1}$. Namely, there exist $R > 0$, a function h defined on $(H \times \mathbb{R}) \setminus B_R$, and a compact set in $\mathbb{R}^{n+1}$ such that

$$\Sigma \setminus K = \{(y', h(y', y_n), y_n) : (y', y_n) \in \mathbb{R}^{n+1} \setminus B_R\};$$

moreover h satisfies

$$|h(y)| + |y||\nabla h(y)| \leq C|y|^{1-\alpha} \tag{72}$$

for some $\alpha \in (0, 1)$ in $\mathbb{R}^n \setminus B_R$. In particular, the minimal hypersurface Σ is the graph of h over $H \times \mathbb{R}_{n+1} \setminus B_R$, and thus h satisfies the MSE. Since Σ is also the graph of u over $\mathbb{R}^{n+1} = H \times \mathbb{R}_n$, we have that

$$h(x', u(x', x_n)) = x_n \tag{73}$$

whenever they are defined.

In particular (73) implies that $u(x', \cdot)$ is injective for each x' fixed. Then $u(x', \cdot)$ must be strictly increasing or decreasing for all x', because if not, by the continuity of u we can find some x'_0 at the interface such that $u(x'_0, t_1) = u(x'_0, t_2)$ for some $t_1 < t_2$. Assume without loss of generality that it is strictly increasing, and thus $\partial_n u > 0$ everywhere. Since (73) implies that $\partial_n u(x) \cdot \partial_n h(y) = 1$, it follows that $\partial_n h > 0$ everywhere. Moreover, we also have $u(x', te_n) \nearrow +\infty$ as $t \nearrow +\infty$, otherwise h is not well-defined at a finite point by (73). By (72) $|\nabla h|$ is uniformly bounded near ∞, so by the elliptic theory ∇h has a finite limit as $|y| \to \infty$. Taking derivatives of (73), we obtain that ∇u also has a finite limit as $|x| \to \infty$, and we fall into the first alternative. Therefore C must have singularity in the second alternative.

7 Example of non-trivial global minimal graph

Bombieri-De Giorgi-Giusti constructed a non-trivial global minimal graph modelled on the Simons cone. Later, Simon constructed non-trivial global minimal graphs modelled on any strictly minimizing isoparametric cones with isolated singularity.

Theorem 7.1. *There exists a global solution u to the MSE in $\mathbb{R}^8$ such that the tangent cone to* $\mathrm{Graph}(u)$ *at ∞ is $C \times \mathbb{R}$, where $C = C^{4,4}$ as in* (37).

Proof. We first observe the Simons cone has $O(4)\times O(4)$ symmetry, so we aim to construct a solution that respects such symmetry: For $(x, y) \in \mathbb{R}^4 \times \mathbb{R}^4$, let $s = |x|$ and $t = |y|$, and we look for solutions of the form $u(x, y) = F(s, t)$ for some function F defined on the first quadrant

$$T := \{(s, t) \in \mathbb{R}^2 : s \geq 0, t \geq 0\}.$$

Denote

$$T_1 := \{(s, t) \in \mathbb{R}^2 : 0 \leq t < s\}, \qquad T_2 := \{(s, t) \in \mathbb{R}^2 : 0 \leq s < t\},$$

which corresponds to the orbits of the two connected components of $\mathbb{R}^8 \setminus C^{4,4}$ under the $O(4) \times O(4)$ group action. To build intuition, we remark that we use the convention the component T_1 corresponds to the inside of $C^{4,4}$, denoted as S (which is also the interior of E_∞ in the proof of Theorem 6.1). Under this new variables, the MSE becomes

$$\mathcal{H}(F) := (1 + F_t^2)F_{ss} + (1 + F_s^2)F_{tt} - 2F_sF_tF_{st} + 3\left(\frac{F_s}{s} + \frac{F_t}{t}\right)(1 + F_s^2 + F_t^2) = 0.$$

The idea of the proof is to construct two barriers $F_1(s, t)$ and $F_2(s, t)$, such that

$$0 < F_1 < F_2 \text{ in } T_1, \qquad F_2 < F_1 < 0 \text{ in } T_2, \qquad F_1 = F_2 = 0 \text{ when } s = t, \tag{74}$$

and

$$\mathcal{H}(F_1) > 0 \text{ in } \operatorname{int} T_1, \qquad \mathcal{H}(F_1) < 0 \text{ in } \operatorname{int} T_2, \qquad F_1(s, t) = -F_1(t, s), \tag{75}$$

$$\mathcal{H}(F_2) < 0 \text{ in } \operatorname{int} T_1, \qquad \mathcal{H}(F_2) > 0 \text{ in } \operatorname{int} T_2. \tag{76}$$

Suppose such barriers exist. Then let $u_i(x, y) = F_i(s, t)$ for $i = 1, 2$. For any $R > 0$, let $u^{(R)}$ be the solution to the Dirichlet problem

$$\begin{cases} \operatorname{div}\left(\frac{\nabla u}{\sqrt{1+|\nabla u|^2}}\right) = 0, & \text{in } B_R \\ u = u_1, & \text{on } \partial B_R. \end{cases}$$

The uniqueness of the solution indicates that $u^{(R)}(x, y)$ must agree with some $F^{(R)}(s, t)$; moreover, the symmetry of F_1 in (75) indicates that $u^{(R)} = 0$ on ∂S. By the comparison principle for quasilinear elliptic PDE (see for example [GT, Theorem 10.1]), we have

$$0 \leq u_1 \leq u^{(R)} \leq u_2, \qquad \text{in } B_R \cap S,$$

$$u_2 \leq u^{(R)} \leq u_1 \leq 0, \qquad \text{in } B_R \setminus S.$$

In other words,

$$|u_1| \leq |u^{(R)}| \leq |u_2|, \qquad \text{in } B_R.$$

For any $\rho > 0$ fixed, for R sufficiently large, by the interior gradient estimate we have

$$\sup_{B_\rho} |\nabla u^{(R)}| \le C_1 \exp\left(C_2\left(1 + \sup_{B_{\rho+1}} u^{(R)}\right)\right) \le C_1 \exp\left(C_2\left(1 + \sup_{B_{\rho+1}} u_2\right)\right),$$

where the right hand side is independent of R. Therefore by Arzela-Ascoli theorem we can extract a subsequence from $u^{(R)}$ which converges uniformly on compact sets of $\mathbb{R}^8$ to a function u, which also solves the minimal surface equation. Moreover $|u| \ge |u_1|$ in $\mathbb{R}^8$, so u is not a hyperplane as long as u_1 has superlinear growth.

Bombieri-De Giorgi-Giusti constructed the barriers

$$F_1(s,t) = (s^2 - t^2)(s^2 + t^2)^{\frac{1}{2}},$$

$$F_2(s,t) = H\left\{(s^2 - t^2)\left[1 + (s^2+t^2)^{\frac{1}{2}}\left(1 + A\left|\frac{s^2 - t^2}{s^2 + t^2}\right|^{\lambda - 1}\right)\right]\right\},$$

$$H(z) = \int_0^z \exp\left(B\int_{|y|}^\infty \frac{dx}{x^{2-\lambda} + x^{2\lambda - 1}}\right) dy,$$

where $\frac{21}{16} < \lambda < \frac{4}{3}$ and A, B are sufficiently large, which satisfy the desired properties. By another change of variables

$$r := s^2 + t^2, \quad \rho := \frac{s^2 - t^2}{s^2 + t^2},$$

we can rewrite

$$F_1 = \rho r^{\frac{3}{2}}, \quad F_2 = H(\rho r + g(\rho) r^{\frac{3}{2}}), \qquad \text{where } g(\rho) = \rho + A\rho^\lambda.$$

The exponents 1 and $\frac{3}{2}$ of $r = |(x,y)|^2$ comes from the first eigenfunction of the Jacobi operator of the Simons cone, and the level sets $\{F_1 = c\}$ foliate S. $\square$

We also briefly sketch Simon's argument, which applies in more general settings, but requires more PDE analysis. Let $\Sigma := C \cap \mathbb{S}^{n-1}$ be the link of the cone, and that Σ separates the sphere into two connected components U_+ and U_-. Let $u_k \in C^2(B_1(0) \cup U_+ \cup U_-)$ be the solution of the MSE in $B_1(0)$ with boundary data

$$u_k = \begin{cases} k, & \text{on } U_+, \\ -k, & \text{on } U_-. \end{cases}$$

Clearly u_k respects all the symmetries of the cone, and that $u_k \equiv 0$ on $C \cap B_1(0)$. For any $\rho \in (0,1)$ fixed, we can show $\sup_{B_\rho(0)} |\nabla u_k| \to \infty$ as $k \to \infty$; and $\nabla u_k(0) = 0$ by symmetry. Choose $\rho_k \in (0,1)$ such that

$$\sup_{B_{\rho_k}(0)} |\nabla u_k| = 1, \qquad |\nabla u_k| < 1 \text{ in } B_{\rho_k}(0).$$

Note that we must have $\rho_k \to 0$. We define

$$v_k(x) := \frac{u_k(\rho_k x)}{\rho_k}, \qquad \text{for } x \in B_{\rho_k^{-1}}(0).$$

Each v_k solves the MSE in $B_{\rho_k^{-1}}(0)$ and

$$v_k(0) = 0, \qquad \nabla v_k(0) = 0, \qquad \sup_{B_1(0)} |\nabla v_k| = 1.$$

Modulo passing to a subsequence, v_k converges locally in C^2 to a function $v \in C^2(\Omega)$ that solves the MSE in some domain Ω, with $\Omega \supset \overline{B_1(0)}$ and $\partial\Omega \cap C = \emptyset$. Either $\Omega = \mathbb{R}^n$, then v is a global solution to the MSE and we are done. Or $\partial\Omega$ consists of two connected components $\Sigma_\pm$, where $\Sigma_\pm$ is a hypersurface in $U_\pm$ respectively and that $v(x)$ tends to $\pm\infty$ as x tends to $\Sigma_\pm$. In this case, we can show that $\Sigma_\pm$ are minimizing hypersurfaces, and thus they must be leaves of the Hardt-Simon foliation, so their asymptotics towards C at ∞ are known explicitly: $\Sigma_\pm \setminus B_R = \operatorname{Graph}_C h_\pm$ where $h_\pm(r\theta) = cr^{-\gamma_1} + \text{L.O.T.}$ for some explicit $-\gamma_1 < 0$. This indicates that

$$\operatorname{Graph}(u) \setminus (B_R \times \mathbb{R}) = \operatorname{Graph}_{C\times\mathbb{R}} w,$$

for some w defined on a unbounded subset of $C \times \mathbb{R}$, such that w solves the minimal surface equation on $C \times \mathbb{R}$.

Moreover $\Omega = \mathbb{R}^n$, and that the tangent cone to v at ∞ is $C \times \mathbb{R}$.

8 Precise asymptotics of non-trivial minimal graph modelled on some $C \times \mathbb{R}$

We have seen that a global solution u to the MSE is affine if its gradient is bounded at ∞. In fact, by the work of Caffarelli-Nirenberg-Spruck, Esker-Huisken, and Nitsche, this condition can be relaxed: u is affine if $|\nabla u(x)| = o(|x|^{\frac{1}{2}})$, or $|\nabla u(x)| = o\left(\sqrt{|x|^2 + |u(x)|^2}\right)$, or $|\nabla u(x)| = O(|x|^\mu)$ with $\mu < 1$. These results also give lower bounds for the growth rate of $|\nabla u|$, for non-trivial solutions. In fact, given some information on the tangent cone of $\operatorname{Graph}(u)$ at ∞, the growth rate for $|\nabla u|$ can be further identified. The following theorem is proven in [Si2, Theorem 5].

Theorem 8.1. *Suppose $u \in C^2(\mathbb{R}^n \setminus \Omega)$, Ω bounded, is a solution to the MSE, such that* $\operatorname{Graph}(u)$ *has a tangent cone $C \times \mathbb{R}$ at ∞, with* $\operatorname{sing} C = \{0\}$ *and C strictly minimizing and strictly stable. Then for each $\epsilon > 0$, there exists $c = c(\epsilon, u) > 0$ and $R_0 = R_0(\epsilon, u)$ such that*

$$|\nabla u(x)| \le c|x|^{\gamma_1+\epsilon}, \qquad |u(x)| \le c|x|^{\gamma_1+1+\epsilon}, \qquad \text{for all } |x| \ge R_0,$$

and there is a sequence x_j with $|x_j| \to \infty$ such that

$$|\nabla u(x_j)| \ge c^{-1}|x_j|^{\gamma_1-\epsilon}, \qquad |u(x_j)| \ge c^{-1}|x_j|^{\gamma_1+1-\epsilon}.$$

Here $\gamma_1 = \gamma_1(C) > 1$ is an explicit constant determined by the cone C.

Remark 8.2. The exponent $\gamma_1 = 2$ when $C = C^{4,4}$ is the Simons cone.

We first explain where the exponent $\gamma_1(C)$ comes from. Suppose M is an m-dimensional minimal hypersurface in $\mathbb{R}^{m+1}$. We aim to deduce the linearization of the mean curvature on M. To that end, consider a continuous family of deformation $u_t : M \to \mathbb{R}^{m+1}$ with initial velocity $\frac{d}{dt}\big|_{t=0} u_t = h\nu$, such that each $u_t(M)$ is also a minimal hypersurface. Here ν is the unit normal vector to M and h is a function on M. Then

$$0 = \frac{d}{dt}\Big|_{t=0} H_{u_t(M)} = \Delta_M h + |A_M|^2 h.$$

In other words, any continuous family of minimal hypersurface near M is given by a function h in the kernel of the *Jacobi operator* $\mathcal{L}_M = \Delta_M + |A_M|^2$. In particular, when M is a cone C with smooth link $\Sigma := C \cap \mathbb{S}^m$, the kernel of $\mathcal{L}_C$ can be further analyzed using the separation of variables.

Consider functions of the form $h(x) = f(r)g(\theta)$, where $x \in C$, $r := |x|$ and $\theta := x/|x| \in \mathbb{S}^m$. Then

$$\begin{aligned}
\mathcal{L}_C h &= \Delta_C h + |A_C|^2 h \\
&= \frac{1}{r^{m-1}} \frac{\partial}{\partial r}\left(r^{m-1} \frac{\partial}{\partial r} h \right) + \frac{1}{r^2}\Delta_\Sigma h + \frac{1}{r^2}|A_\Sigma|^2 h \\
&= f''(r)g(\theta) + \frac{m-1}{r} f'(r)g(\theta) + \frac{f(r)}{r^2}\Delta_\Sigma g(\theta) + \frac{m-1}{r^2} f(r)g(\theta) \\
&= \left(f''(r) + \frac{m-1}{r} f'(r) - \lambda \frac{f(r)}{r^2} \right) g(\theta) + \left(\Delta_\Sigma g(\theta) + |A_\Sigma|^2 g(\theta) + \lambda g(\theta) \right) \frac{f(r)}{r^2}.
\end{aligned}$$

This means that if g is an eigenfunction of the operator $\mathcal{L}_\Sigma := \Delta_\Sigma + |A_\Sigma|^2$ with $-\mathcal{L}_\Sigma g = \lambda g$, then $h(r,\theta) = r^\gamma g(\theta)$ is in the kernel of $\mathcal{L}_C$, where $\gamma(\gamma + m - 2) = \lambda$, namely

$$\gamma_\pm = -\frac{m-2}{2} \pm \sqrt{\left(\frac{m-2}{2}\right)^2 + \lambda}.$$

By the standard elliptic theory $\mathcal{L}_\Sigma$ has discrete spectrum

$$\lambda_1 < \lambda_2 \le \lambda_3 \le \cdots \lambda_k \le \cdots \nearrow \infty;$$

and we denote the corresponding exponents $\gamma_j^\pm$.

In particular when $C = C^{p,q}$ is the Simons cone, we have seen in Theorem 3.10 that $|A_\Sigma|^2 = m - 1$. Then the first eigenvalue of $-\Delta_\Sigma - |A_\Sigma|^2$ is $\lambda_1 = -(m-1)$, with the corresponding eigenfunction being the constant function. Whenever

$$\left(\frac{m-2}{2}\right)^2 + \lambda_1 = \left(\frac{m-2}{2}\right)^2 - (m-1) \ge 0 \iff m \ge 7,$$

the above discussion holds and that $cr^{\gamma_1^\pm}$ is the only kernel element in $\mathcal{L}_C$ with a sign. Recall we have seen how the above quantifier is related to the stability of the Simons cone in the proof of Theorem 3.10.

In Theorem 8.1, the cone C is a $m := n-1$ dimensional minimal/minimizing hypersurface in $\mathbb{R}^n$. Let $\Sigma := C \cap \mathbb{S}^{n-1}$ be the smooth link of C, and let $\lambda_1(C)$ denote the first eigenvalue of $-\Delta_\Sigma - |A_\Sigma|^2$. Then the exponent

$$\gamma_1(C) = -\gamma_1^+(C) = \frac{n-3}{2} - \sqrt{\left(\frac{n-3}{2}\right)^2 + \lambda_1(C)}.$$

Whenever C is singular, Σ is a minimal hypersurface in $\mathbb{S}^{n-1}$ that is different from the equator $\mathbb{S}^{n-2}$. Simons showed that $|A_\Sigma|^2 \geq n-2$, with equality if and only if Σ is the link of a Simons cone. Hence $\lambda_1(C) \leq -(n-2)$, and

$$\begin{aligned}
\gamma_1(C) = \frac{n-3}{2} - \sqrt{\left(\frac{n-3}{2}\right)^2 + \lambda_1(C)} &\geq \frac{n-3}{2} - \sqrt{\left(\frac{n-3}{2}\right)^2 - (n-2)} \\
&> \frac{n-3}{2} - \sqrt{\left(\frac{n-3}{2}\right)^2 - (n-4)} \\
&= \frac{n-3}{2} - \frac{n-5}{2} \\
&= 1.
\end{aligned}$$

With a lot more work, the above linear analysis hints at the following asymptotics (77), for the nonlinear problem of studying minimal hypersurfaces lying on one side of a cone.

Theorem 8.3 (Hardt-Simon foliation)**.** *Let C be a minimizing (and strictly stable[14]) hypercone in $\mathbb{R}^{m+1}$, and let U be either one of the connected components $U_\pm$ of $\mathbb{R}^{m+1} \setminus C$. There is an oriented connected embedded complete real-analytic hypersurface $S \subset U$ with S being the boundary of a perimeter minimizer E, $\overline{E} \subset U$, and* $\operatorname{dist}\{S, \{0\}\} = 1$. *Moreover,*

i) for any $\xi \in U$, the ray $\{\lambda\xi : \lambda > 0\}$ intersects S in a unique point, and the intersection is transverse;

ii) there exist $R > 0$ and a C^2 function v on $C \setminus B_R$ such that

$$S \setminus B_R = \operatorname{Graph}_C(v) := \{x + v(x)\nu_C(x) : x \in C \setminus B_R\},$$

and v has the asymptotics

$$\textit{either } v = cr^{\gamma_1^+} g_1 + O(r^{\gamma_1^+ - \alpha}), \qquad \textit{or } v = cr^{\gamma_1^-} g_1 + O(r^{\gamma_1^- - \alpha}), \tag{77}$$

as $r \to +\infty$, with some constant $c > 0$;

iii) if T is the boundary of any other perimeter minimizer F for some $F \subset \overline{U}$, then either $T = C$ or T is some rescaling of S.

If in addition C is strictly minimizing, then v has the slower decay rate $O(r^{\gamma_1^+})$ at ∞. Moreover, for any non-negative function $h \in C^1_c(S)$ and any ϵ sufficiently small, there are positive solutions $w^\epsilon \in C^2(S)$ of $\mathcal{M}_S w^\epsilon = -\epsilon h$, and furthermore $\limsup_{r\to+\infty} w^\epsilon / r^{\gamma_1^-} < +\infty$.

The precise asymptotics in Theorem 8.1 comes from the fact that the level set $S_y := \{x \in \mathbb{R}^n \setminus \Omega : u(x) = y\}$ is very close to the rescalings of one of the Hardt-Simon foliations $S_\pm$ in Theorem 8.3, depending on $y > 0$ or $y < 0$.

[14] If we do not assume *strict* stability, the exponents $\gamma_1^+ = \gamma_1^- = -\frac{m-2}{2}$ and the possible asymptotics in (77) are $O(r^{-\frac{m-2}{2}} \log r)$ or $O(r^{-\frac{m-2}{2}})$.

Proposition 8.4. *Suppose $u \in C^2(\mathbb{R}^n \setminus \Omega)$ is a solution to the MSE, with Ω open and bounded, and* $\mathrm{Graph}(u)$ *has a tangent cylinder $C \times \mathbb{R}$ at ∞ with* $\mathrm{sing}(C) = \{0\}$*. Then $C \times \mathbb{R}$ is the unique tangent cylinder of* $\mathrm{Graph}(u)$ *at ∞, and on the two connected components $U_\pm$ of $\mathbb{R}^n \setminus C$ we have*

$$\lim_{r\to\infty} \frac{u(r\omega)}{r} = \begin{cases} +\infty, & \text{if } \omega \in \mathbb{S}^{n-1} \cap U_+, \\ -\infty, & \text{if } \omega \in \mathbb{S}^{n-1} \cap U_-, \end{cases}$$

where the convergence is uniform on compact subsets of $\mathbb{S}^{n-1} \cap U_\pm$.

Furthermore

$$\lim_{r\to\infty} \inf_{\omega\in\mathbb{S}^{n-1}} |\nabla u(r\omega)| = \infty, \tag{78}$$

and for any given $\epsilon > 0$, there is $y(\epsilon) > 0$ such that if $y > y(\epsilon)$ ($y < -y(\epsilon)$),

$$\frac{S_y}{\mathrm{dist}(S_y, \{0\})} \text{ is within } \epsilon\text{-distance of } S_+ \text{ (resp. } S_-\text{) in the } C^2_* \text{ sense.} \tag{79}$$

Finally, the length of the second fundamental form A of $\mathrm{Graph}(u)$ *satisfies*

$$|A(x, u(x))| \le \frac{C}{|x|}, \qquad \text{for } x \in \mathbb{R}^n \setminus \Omega, \tag{80}$$

and the gradient function $h := \sqrt{1 + |\nabla u(x)|^2}$ satisfies the Harnack inequality

$$\sup_{B_\rho(X_0)\cap\mathrm{Graph}(u)} h \le c \inf_{B_\rho(X_0)\cap\mathrm{Graph}(u)} h, \tag{81}$$

where $X_0 = (x_0, u(x_0))$ and $\rho > 0$ such that $|x_0| > 2\rho + \mathrm{diam}\,\Omega$.

Definition 8.5. Let S_1, S_2 be two complete embedded hypersurfaces. We say S_2 is within ϵ-distance of S_1 in the C^2_* sense if

$$S_2 = \mathrm{Graph}_{S_1}(\eta) := \{x + \eta(x)\nu_{S_1}(x) : x \in S_1\}$$

where ν_{S_1} is a unit normal vector for S_1, and $\eta \in C^2(S_1)$ satisfies

$$\|\eta\|_{C^2_*} := \sup_{x\in S_1\setminus\{0\}} \left(\frac{|\eta(x)|}{|x|} + |\nabla\eta(x)| + |x||\nabla^2\eta(x)| \right) < \epsilon.$$

The definition of the C^2_* norm is to guarantee it is invariant under rescalings.

Proof. We prove (80) and (79). The control on the second fundamental form (80) is a quantitative property related to the assumption that $\mathrm{sing}(C) = \{0\}$, which roughly speaking implies that the curvature should blow up according to the blowup rate of a regular cone. The statement (79) on the level set indicates that the asymptotics of u near ∞ follows from the asymptotics of the Hardt-Simon foliation $S_\pm$, which is given by γ_1^+. The Harnack inequality (81) follows from the fact that $\nu_{n+1} = 1/\sqrt{1 + |\nabla u(x)|^2}$ is a Jacobi field (recall similar argument in the proof of the interior gradient estimate in Theorem 2.6), as well as connectedness.

Assume (80) is false, then there exists a sequence $x_k \in \mathbb{R}^n \setminus \Omega$ with $|x_k| \to \infty$ such that

$$|x_k| \cdot |A(x_k, u(x_k))| \to \infty.$$

Let G_k denote the graph of the function v_k, defined as

$$v_k(x) := \frac{u(\lambda_k x) - u(x_k)}{\lambda_k}, \qquad \text{with } \lambda_k := |x_k| \to \infty \text{ and } x \in \mathbb{R}^n \setminus (\Omega/r_k).$$

In other words, G_k is the image of G under the map $X \mapsto (X - \mu_k e_{n+1})/\lambda_k$, where $\mu_k := u(x_k)$. The translation by $\mu_k e_{n+1}$ is to guarantee that no matter μ_k is bounded or blows up, we do not lose them after passing to the limit. By similar argument as in the proof of Theorem 6.1, we have that modulo passing to a subsequence G_k converges to a minimizing hypersurface H. If μ_k is bounded, since $C \times \mathbb{R}$ is the unique tangent of Graph(u) at ∞, we have $H = C \times \mathbb{R}$. If μ_k blows up, assume without loss of generality that $\mu_k \to +\infty$. We claim that $H \subset \overline{U}_+ \times \mathbb{R}$.

Since Graph$(u) \setminus (B_R \times \mathbb{R})$ is graphical over $C \times \mathbb{R}$ for some R large, $S_0 \setminus B_R$ is a smooth embedded hypersurface that tends to C asymptotically. Hence $\mathbb{R}^n \setminus (S_0 \cup B_R)$ has two unbounded connected components, denoted as $V_\pm$, which tends to $U_\pm$ near ∞ respectively. Moreover, each level set S_y also tends to C near ∞, so it only has one unbounded connected component which lies in V_+ or V_- depending on the sign of $y > 0$ or $y < 0$. It also can not have a bounded connected component by the maximum principle, if $|y| > \sup_{\partial\Omega} |u|$. Therefore S_y is connected and $S_y \subset V_+$ if $y > 0$.

Let $a \in \mathbb{R}$ be fixed. Then for any z such that $\lambda_k z \in S_{u(x_k)+a\lambda_k}$, we have $v_k(z) = a$, namely $(z, a) \in G_k$ and $\{z\} \times (-\infty, a) \subset \mathrm{SGraph}(v_k)$. On the other hand, if $u(x_k)/\lambda_k \to +\infty$, we have $u(x_k)+a\lambda_k \to \infty$. If $u(x_k)/\lambda_k \to M$ for some $M \in \mathbb{R}$, then for any $a > -M$ we have $u(x_k) + a\lambda_k \to \infty$. In both cases, by the discussion above $\lambda_k z \in S_{u(x_k)+a\lambda_k} \subset V_+$. Since $V_+/\lambda_k \to \overline{U}_+$, it follows that $\{v_k = a\} \subset \overline{U}_+$ either for all $a \in \mathbb{R}$ or for all $a > -M$. Therefore the limit H of G_k satisfies $H \subset \overline{U}_+ \times \mathbb{R}$.

Recall that $G_k \to H$ locally in the Hausdorff distance sense in $\mathbb{R}^{n+1} \setminus \{0\} \times \mathbb{R}$. Since $(x_k/|x_k|, 0) \in G_k$ for each k, modulo passing to subsequences we have $x_k/|x_k| \to x_0 \in \mathbb{S}^{n-1}$ such that $(x_0, 0) \in H$. By similar argument as in the proof of Theorem 6.1, either H is a vertical cylinder of the form $H_1 \times \mathbb{R}$, or else H is the graph of a minimal surface equation w over a unbounded domain $\Omega_\infty \subset \mathbb{R}^n$, such that $|w(x)|$ blows up as $x \to \partial\Omega_\infty$. In the first case, we have that H_1 is minimizing and $H_1 \subset \overline{U}_+$. So by the uniqueness in Theorem 8.3, either $H_1 = C$ or H_1 is equal to a rescaling of the foliation S_+. In either of the three cases we always have that $x_0 \in H \cap \mathbb{S}^{n-1}$ is a regular point, i.e. $x_0 \in \operatorname{reg} H$. So in a neighborhood of x_0, the sequence G_k converges to H in the classical sense. This is a contradiction with $G_k \ni (x_k/|x_k|, 0) \to (x_0, 0)$ and $|A_{G_k}(x_k/|x_k|, 0)| = |x_k| \cdot |A(x_k, u(x_k))| \to \infty$. Therefore (80) is proven.

Now let $\theta_\pm$ be points in $\mathbb{S}^{n-1} \cap U_\pm$ respectively, and let $\epsilon > 0$. We claim that there exists $K = K(\epsilon, u, \theta_\pm)$ such that if $|\nabla u(r\theta_\pm)| \geq K$, then $S_{u(r\theta_\pm)}/\operatorname{dist}(S_{u(r\theta_\pm)}, 0)$ is within ϵ-distance of $S_\pm$ in the C^2_* sense, and $|\nabla u| \geq \epsilon^{-1}$ on $S_{u(r\theta_\pm)}$. This, combined with (78), will give (79). Assume the statement is false for some θ_+, namely there exist $\epsilon > 0$ and a sequence r_k with $|\nabla u(r_k\theta_+)| \geq k$, such that one of the conclusions is false. Firstly the assumption $|\nabla u(r_k\theta_+)| \geq k$ implies that $r_k \to \infty$.

We consider the rescaling

$$v_k(x) := \frac{u(r_k x) - u(r_k\theta_+)}{r_k}, \qquad G_k := \operatorname{Graph} v_k.$$

In other words, G_k is the image of G under the map $X \mapsto (X - u(r_k\theta_+)e_{n+1})/r_k$. Then

$$\nabla v_k(x) = \nabla u(r_k x) \implies |\nabla v_k(\theta_+)| \geq k, \qquad \theta_+ \in S_0^k := \{v_k = 0\} = \frac{S_{u(r_k\theta_+)}}{r_k}.$$

By the same argument as in the proof of (80), we can show that G_k converges to $H \subset \overline{U}_+ \times \mathbb{R}$. Since $(\theta_+, 0) \in G_k$ and $|\nabla v_k(\theta_+)| \to \infty$, H can not be a minimal graph or $C \times \mathbb{R}$. So $H = S_+^\lambda \times \mathbb{R}$, where S_+^λ is a rescaling of S_+ and λ is chosen so that $\theta_+ \in S_+^\lambda$. In particular, for k sufficiently large S_0^k is within ϵ-distance of S_+^λ in the C_*^2 sense. By the Harnack inequality (81), we can show that $\inf_{S_0^k} |\nabla v_k| \to \infty$, and thus $\inf_{S_{u(r_k\theta_+)}} |\nabla u_k| \to \infty$, contradicting the second hypothesis. On the other hand,

$$\operatorname{dist}(S_{u(r_k\theta_+)}, 0) = r_k \operatorname{dist}(S_0^k, 0) \approx r_k \operatorname{dist}(S_+^\lambda, 0),$$

so by the scale-invariance of the C_*^2-norm,

$$\frac{S_{u(r_k\theta_+)}}{\operatorname{dist}(S_{u(r_k\theta_+)}, 0)} = \frac{r_k}{\operatorname{dist}(S_{u(r_k\theta_+)}, 0)} S_0^k \approx \frac{S_0^k}{\operatorname{dist}(S_+^\lambda, 0)} \text{ is within } \epsilon\text{-distance of } S_+.$$

□

References

[CM] Tobias H. Colding and William P. Minicozzi II, *A Course in Minimal Surfaces.* Graduate Studies in Mathematics, 2011

[GT] David Gilbarg and Neil S. Trudinger, *Elliptic Partial Differential Equations of Second Order.* Classics in Mathematics, Springer 2001

[EG] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions, Revised Edition.* Textbooks in Mathematics, CRC Press, 2015

[HL] Qing Han and Fanghua Lin, *Elliptic Partial Differential Equations: Second Edition.* Courant Lecture Notes, 2011

[Ma] Francesco Maggi, *Sets of Finite Perimeter and Geometric Variational Problems: An Introduction to Geometric Measure Theory.* Cambridge Studies in Advanced Mathematics, 2012

[SS] R. Schoen and L. Simon, *Regularity of stable minimal hypersurfaces.* Comm. Pure and Appl. Math. **34** (1981), 742-797.

[Si1] L. Simon, *Asymptotic behaviour of minimal graphs over exterior domains.* Ann. Inst. H. Poincare 4 (1987) 231-242.

[Si2] L. Simon, *Entire solutions of the minimal surface equation.* J. Differential Geometry 30 (1989) 643-688.

[WZ] R. L. Wheeden and A. Zygmund, *Measure and Integral: An Introduction to Real Analysis, Second Edition*