COMMUNICATIONS IN
ANALYSIS AND GEOMETRY
Volume 1, Number 2, 281–326, 1993

EXISTENCE OF SURFACES MINIMIZING THE WILLMORE FUNCTIONAL

LEON SIMON

For compact surfaces Σ embedded in $\mathbb{R}^n$, the Willmore functional is defined by

$$\mathcal{F}(\Sigma) = \frac{1}{4}\int_\Sigma |\mathbf{H}|^2$$

where the integration is with respect to ordinary 2-dimensional area measure, and $\mathbf{H}$ is the mean curvature vector of Σ (in case $n = 3$ we have $|\mathbf{H}| = |\kappa_1 + \kappa_2|$, where κ_1, κ_2 are the principal curvatures of Σ). In particular $\mathcal{F}(\mathbb{S}^2) = 4\pi$.

For surfaces Σ without boundary we have the important fact that $\mathcal{F}(\Sigma)$ is invariant under conformal transformations of $\mathbb{R}^n$; thus if $\tilde{\Sigma} \subset \mathbb{R}^n$ is the image of Σ under an isometry or a scaling ($x \mapsto \lambda x, \lambda > 0$) or an inversion in a sphere with centre not in Σ (e.g. $x \mapsto x/|x|^2$ if $0 \notin \Sigma$) then

$$\mathcal{F}(\Sigma) = \mathcal{F}(\tilde{\Sigma}). \tag{0.1}$$

(See [WJ], [LY], [W] for general discussion.)

For each genus $g = 0, 1, 2, \ldots$ and each $n \geq 3$ we let

$$\beta_g^n = \inf \mathcal{F}(\Sigma),$$

where the inf is taken over all compact genus g surfaces without boundary embedded in $\mathbb{R}^n$. We note some inequalities concerning the numbers β_g^n. Firstly we claim

$$4\pi \leq \beta_g^n < 8\pi \tag{0.2}$$

Research partly supported by NSF grant DMS–9012718 at Stanford University.

with equality on the left if and only if $g = 0$ (indeed $\mathcal{F}(\Sigma) \geq 4\pi$, with equality if and only if Σ is a round sphere—see the simple argument of [W]). The right-hand-side inequality in (0.2) was pointed out to the author by Pinkall [P] and (independently) by Kusner [K], who both noted that, by an area comparison argument, the genus g minimal surfaces $\Sigma_g \subset \mathbb{S}^3$ constructed by Lawson [L] have area $< 4\pi$. It then follows (using the conformal invariance of the Willmore functional between general Riemannian 3-manifolds as discussed in [WJ]) that $\mathcal{F}(\widetilde{\Sigma}_g) < 8\pi$, where $\widetilde{\Sigma}_g$ is the stereographic image of Σ_g in $\mathbb{R}^3$. Another inequality concerning the numbers β_g^n is as follows: if $e_g = \beta_g^n - 4\pi (= \beta_g^n - \beta_0^n)$ then

$$e_g \leq \sum_{j=1}^{q} e_{\ell_j} \tag{0.3}$$

for any integers $q \geq 2$ and $\ell_1, \dots, \ell_q \geq 1$ with $\sum_{j=1}^{q} \ell_j = g$. To see this we take a genus ℓ_j surface $\Sigma_k^{(j)}$ with $\mathcal{F}(\Sigma_k^{(j)}) \leq \beta_{\ell_j}^n + 1/k$ which is C^2 close to $\mathbb{S}^2$ except near some preassigned spherical cap of $\mathbb{S}^2$ (we get sum of a sequence by first taking any sequence $\{\Sigma_k^{(j)}\}$ satisfying the given inequality, and then for each k taking an inversion in a suitable sphere with center y_k very close to $\Sigma_k^{(j)}$); near this spherical cap $\Sigma_k^{(j)}$ looks like a spherical cap with ℓ_j handles. Then by cutting out these spherical caps with handles and sewing them back into a copy of $\mathbb{S}^2$ with q spherical caps removed, we get a genus g surface $\widetilde{\Sigma}_k$ with

$$\mathcal{F}(\widetilde{\Sigma}_k) \leq 4\pi + \sum_{j=1}^{q} e_{\ell_j} + \epsilon_k, \quad \epsilon_k \downarrow 0 \quad \text{as} \quad k \to \infty.$$

It is of course tempting to conjecture that the stereographic image of $\widetilde{\Sigma}_g \subset \mathbb{S}^3$ (as above) actually minimizes $\mathcal{F}$ (so that we would have $\mathcal{F}(\widetilde{\Sigma}_g) = \beta_g^3$). There is some evidence to support this in case $g = 1$ (see [LY], [WJ]).

One of the main results of this paper is that for each $n \geq 3$ there *exists* a compact embedded real analytic torus T in $\mathbb{R}^n$ with $\mathcal{F}(T) = \beta_1^n$. For arbitrary genus $g \geq 2$ the result is almost as clear-cut; we prove that *there is a genus g embedded real analytic surface Σ in $\mathbb{R}^n$ with $\mathcal{F}(\Sigma) = \beta_g^n$ unless equality holds in* (0.3) *for some choice of $q \geq 2, \ell_1, \dots, \ell_q$, $\sum_{j=1}^{q} \ell_j = g$, in which case we can construct, by the cut-and-paste procedure used to establish* (0.3), *a minimizing sequence explicitly in terms of lower genus minimizers for $\mathcal{F}$.* It is not clear at the moment whether or not equality *can* hold in (0.3); certainly since $\beta_g^n < 8\pi$ by (0.2), it is clear that equality *cannot* hold if $\beta_\ell^n \geq 6\pi \; \forall \, \ell = 1, \ldots, g-1$. (At the moment it is not known whether or not $\beta_\ell^n \geq 6\pi$, although this seems extremely likely.)

The proof of the above existence results are given in §§1–4 below.

The present paper gives a detailed exposition (in arbitrary codimension) of arguments which were only briefly sketched for the codimension 1 case in the conference proceedings paper [SL1].

1. Lemmas valid for arbitrary compact $\Sigma \subset \mathbb{R}^n$

In each of the 3 lemmas below, Σ denotes a smooth surface in $\mathbb{R}^n$ and $\partial\Sigma = \overline{\Sigma}\backslash\Sigma$, where $\overline{\Sigma}$ denotes the closure of Σ (as a subset of $\mathbb{R}^n$). C will denote a constant depending only on n (and not on Σ). The first two lemmas give bounds on $\operatorname{diam}\Sigma$, where $\operatorname{diam}\Sigma$ denotes diameter of Σ as a subset of $\mathbb{R}^n$; that is, $\operatorname{diam}\Sigma = \sup_{x \neq y,\, x,y \in \Sigma} |x-y|$.

Lemma 1.1. *If $\partial\Sigma = \emptyset$ and if Σ is compact and connected, then*

$$\sqrt{|\Sigma|/\mathcal{F}(\Sigma)} \leq \operatorname{diam}\Sigma \leq C\sqrt{|\Sigma|\mathcal{F}(\Sigma)}.$$

Here $|\Sigma|$ denotes the area of Σ.

Lemma 1.2. *In the general case when $\partial\Sigma \neq \emptyset$ is allowed, and when Σ is connected and $\overline{\Sigma}$ is compact, we have*

$$\operatorname{diam}\Sigma \leq C\Big(\int_\Sigma |\mathbf{A}| + \sum_j \operatorname{diam}\Gamma_j\Big),$$

where $|\mathbf{A}|$ is the length of the second fundamental form of Σ and Γ_j are the connected components of $\partial\Sigma$.

Remark 1.1. Notice that $\partial\Sigma$ need not be smooth or even rectifiable here.

Lemma 1.2 has the following useful corollary, where the notation and hypotheses on Σ are as in the lemma, and where we use the notation

$$B_\rho(y) = \{x \colon |x-y| < \rho\}, \qquad B_\rho = B_\rho(0).$$

Corollary 1.3. *If $\theta \in (0,1)$, there is $\alpha_0 = \alpha_0(n,\theta) > 0$ such that if $\int_{\Sigma\cap B_\rho} |\mathbf{A}| < \alpha_0\rho$, if $\sum_j \operatorname{diam}\Gamma_j \le \alpha_0\rho$, if and $\Sigma \cap \partial B_\rho \ne \emptyset$ and $\Sigma \cap \partial B_{\theta\rho} \ne \emptyset$, then $|\Sigma \cap B_\rho| \ge C\rho^2$, where $C = C(\theta)$.*

Proof. By applying Lemma 1.2 to $\Sigma \cap B_\sigma$, for $\sigma \in (\frac{1+\theta}{2}\rho, \rho)$, we conclude that $\operatorname{length}(\Sigma \cap \partial B_\sigma(0)) \ge C\rho$ for a fixed constant C, and the corollary follows by virtue of the coarea formula. □

In the third lemma we give a result which can be viewed as a variant of a lemma of Li and Yau (see [LY, Theorem 6]).

Lemma 1.4. *Suppose Σ is a compact surface without boundary, ∂B_ρ intersects Σ transversely, and $\Sigma \cap B_\rho$ contains disjoint subsets Σ_1, Σ_2 with $\Sigma_j \cap B_{\theta\rho} \ne \emptyset$, $\partial\Sigma_j \subset \partial B_\rho$, and $|\partial\Sigma_j| \le \beta\rho$, $j = 1, 2$, where $\theta \in (0, \frac{1}{2})$ and $\beta > 0$. Then*

$$\mathcal{F}(\Sigma) \ge 8\pi - C\beta\theta,$$

(where C does not depend on Σ, β, θ).

In the proofs of Lemmas 1.1, 1.4 we use the first variation identity

$$\int_\Sigma \operatorname{div}_\Sigma \Phi = -\int_\Sigma \Phi \cdot \mathbf{H}, \tag{1.1}$$

for any C^1 vector field $\Phi = (\Phi^1, \ldots, \Phi^n)$ defined in a neighbourhood of Σ. Here we use the notation

$$\operatorname{div}_\Sigma \Phi = \sum_{j=1}^n \nabla_j \Phi^j,$$

where $\nabla_j f = e_j \cdot \nabla^\Sigma f$, ∇^Σ denoting the gradient operator on Σ. In particular if Σ is the restriction to Σ of a C^1 function $\bar{f}$ defined in a neighborhood of Σ, then $\nabla_j f(x) = \sum_{i=1}^n g^{ij}(x) D_i \bar{f}(x)$, $x \in \Sigma$, where g^{ij} is the matrix of the orthogonal projection P_x of $\mathbb{R}^n$ onto the tangent space $T_x\Sigma$. Also notice that the identity (1.1) is valid if X is merely a Lipschitz vector field on Σ. Using these facts, we easily check, for any fixed $y \in \mathbb{R}^n$ and $0 < \sigma < \rho$, that we can substitute $\Phi(x) \equiv (|X|_\sigma^{-2} - \rho^{-2})_+ X$, where $X = x - y$ and $|X|_\sigma = \max(|X|, \sigma)$. Since $\operatorname{div} X \equiv 2$, direct computation in(1.1) then shows that

$$2\sigma^{-2}|\Sigma_\sigma| + 2\int_{\Sigma_{\sigma,\rho}} \frac{|X^\perp|^2}{|X|^4} = 2\rho^{-2}|\Sigma_\rho| - \int_{\Sigma_\rho} (|X|_\sigma^{-2} - \rho^{-2}) X \cdot \mathbf{H},$$

where $\Sigma_\sigma = \Sigma \cap B_\sigma(y)$, $\Sigma_{\sigma,\rho} = \Sigma \cap B_\rho(y) \backslash B_\sigma(y)$, and $X^\perp = X - P_x X$. By using the identity $\frac{|X^\perp|^2}{|X|^4} + \frac{1}{2}\frac{X}{|X|^2} \cdot \mathbf{H} = \left|\frac{1}{4}\mathbf{H} + \frac{X^\perp}{|X|^2}\right|^2 - \frac{1}{16}|\mathbf{H}|^2$ we then conclude

$$(1.2) \quad \sigma^{-2}|\Sigma_\sigma| + \int_{\Sigma_{\sigma,\rho}} \left(\tfrac{1}{4}\mathbf{H}(x) + \tfrac{X^\perp}{|X|^2}\right)^2 = \rho^{-2}|\Sigma_\rho| + \tfrac{1}{4}\mathcal{F}(\Sigma_\rho) + \tfrac{1}{2}\int_{\Sigma_\rho} \rho^{-2} X \cdot \mathbf{H} - \tfrac{1}{2}\int_{\Sigma_\sigma} \sigma^{-2} X \cdot \mathbf{H}.$$

Of course since $\sigma^{-2}|\Sigma_\sigma| \to \pi$ as $\sigma \downarrow 0$ and $|X^\perp| \le \beta|X|^2$ (with β depending on Σ), we can let $\sigma \downarrow 0$ in (1.2), thus giving

$$(1.2') \quad \pi + \int_{\Sigma_\rho} \left(\tfrac{1}{4}\mathbf{H}(x) + \tfrac{X^\perp}{|X|^2}\right)^2 = \rho^{-2}|\Sigma_\rho| + \tfrac{1}{4}\mathcal{F}(\Sigma_\rho) + \tfrac{1}{2}\int_{\Sigma_\rho} \rho^{-2} X \cdot \mathbf{H}.$$

Also by dropping the square terms on the left of (1.2) and using the Cauchy inequality on the right, we have

$$(1.3) \qquad \sigma^{-2}|\Sigma_\sigma| \le C(\rho^{-2}|\Sigma_\rho| + \mathcal{F}(\Sigma_\rho)), \quad 0 < \sigma \le \rho < \infty,$$

where C depends only on n and not on Σ or σ or ρ, and in particular, by letting $\sigma \downarrow 0$,

$$(1.4) \qquad \pi \le C(\rho^{-2}|\Sigma_\rho| + \mathcal{F}(\Sigma_\rho))$$

Notice that minor modifications of the discussion leading to (1.2) can be applied in the case when $\overline{\Sigma}$ is compact with smooth boundary $\partial\Sigma \ne \emptyset$, yielding an identity like (1.2′) with $\rho \uparrow \infty$, but with an extra boundary term on the right; more precisely, we have the identity

$$(1.5) \qquad \pi + \int_\Sigma \left(\tfrac{1}{4}\mathbf{H}(x) + \tfrac{(X)^\perp}{|X|^2}\right)^2 = \tfrac{1}{2}\int_{\partial\Sigma} \eta \cdot \frac{X}{|X|^2} + \tfrac{1}{4}\mathcal{F}(\Sigma),$$

where η is the outward pointing unit conormal for $\partial\Sigma$. This is proved in exactly the same way as (1.2′), starting with

$$(1.1') \qquad \int_\Sigma \operatorname{div}_\Sigma \Phi = \int_{\partial\Sigma} \eta \cdot \Phi - \int_\Sigma \Phi \cdot \mathbf{H}$$

in place of (1.1), and then letting $\rho \uparrow \infty$.

Proof of Lemma 1.1. To prove the inequality on the left in Lemma 1.1, we simply choose $\Phi(x) = x - y$ in (1.1), where y is a fixed element of Σ, and note again that $\operatorname{div}_\Sigma x \equiv 2$ on Σ; then the required inequality follows by using the Hölder inequality on the right side.

The proof of the inequality on the right side of Lemma 1.1 involves the identity (1.2′). Take $y \in \Sigma$ and $d = \max_{x\in\Sigma}|x-y|$, let $\rho \in (0, \frac{d}{2}]$ and let $N =$ integer part of $\rho^{-1}d$, and for each $j = 1, \dots, N-1$ take $y_j \in \partial B_{(j+\frac{1}{2})\rho}(y)$. (Notice that here we use the connectivity of Σ.) Then with $y_0 = y$, the balls $B_{\rho/2}(y_j)$, $j = 0, \dots, N-1$ are pairwise disjoint. Thus by using (1.4) with y_j in place of y and summing over j, we deduce that

$$N\pi \le C\Big(\mathcal{F}(\Sigma) + \frac{|\Sigma|}{\rho^2}\Big). \tag{1.6}$$

Now select $\rho = \frac{1}{4}\sqrt{|\Sigma|/\mathcal{F}(\Sigma)}$. By the first inequality of the lemma (which we already proved above), we deduce $d \ge 2\rho$, so that the condition $\rho \in (0, d/2]$ is satisfied. Since $N \ge \frac{1}{2}\rho^{-1}d$, (1.6) above gives

$$d \le C(\rho\mathcal{F}(\Sigma) + \rho^{-1}|\Sigma|) \le 5C\sqrt{|\Sigma|\mathcal{F}(\Sigma)}. \quad \square$$

Proof of Lemma 1.2. First note that it is enough to prove

$$\operatorname{diam}\Sigma \le C\int_\Sigma |\mathbf{A}|, \tag{1.7}$$

subject to the assumption that

$$\sum_j \operatorname{diam}\Gamma_j \le \theta \operatorname{diam}\Sigma, \tag{1.8}$$

provided $\theta \in (0, \frac{1}{4})$ is a fixed constant depending only on n, because otherwise the required inequality is trivially true with $C = \theta^{-1}$.

So let $\theta \in (0, \frac{1}{4})$ be for the moment arbitrary (we will select θ independent of Σ below), and let $y_1, y_2 \in \Sigma$ be such that $|y_1 - y_2| = \operatorname{diam}\Sigma$; for convenience of notation write $d = |y_1 - y_2|$ and let $y_t = y_1 + t(y_1 - y_2)$, $0 \le t \le 1$. Choose $e \in \mathbb{S}^{n-1}$ such that $|e - d^{-1}(y_1 - y_2)| < \frac{1}{2}$, such that there is an open set $I \subset [0, d]$ with

$$|I| \ge \theta d, \tag{1.9}$$

and such that the hyperplane normal to e passing through y_t meets Σ transversely in a family $\Gamma_t = \bigcup_{j=1}^{N_t}\Gamma_t^{(j)}$ of smooth Jordan curves $\Gamma_t^{(j)}$, and

$$\sup_{x\in\Gamma_t^{(1)},\, \tau\in T_x\Sigma,\, |\tau|=1} |e\cdot\tau| \ge \theta. \tag{1.10}$$

Subject to the assumption (1.8), and with the help of Sard's theorem, it is an easy matter to check that this can be arranged by taking θ small enough (but not depending on Σ).

Now for $t \in I$ fixed, let $\gamma(s)$ be the arc-length parametrization for $\Gamma_t^{(1)}$ with $x = \gamma(0)$ such that the sup in (1.10) is attained. Since $\int_{\Gamma_t^{(1)}} \gamma'(0) \cdot \gamma'(s)\, ds = 0$ (so that $\gamma'(0) \cdot \gamma'(s)$ changes sign on $(0, \operatorname{length} \Gamma_t^{(1)})$), we can select $s_1 \in (0, \operatorname{length} \Gamma_t^{(1)})$ such that $\gamma'(0) \cdot \gamma'(s_1) = 0$. Using the notation $s_0 = 0$, we then have by (1.10) that if $v_j = e^{(j)}/|e^{(j)}|$, where $e^{(j)}$ is the orthogonal projection of e onto $T_{\gamma(s_j)}\Sigma$, then $v_j \wedge \gamma'(s_j) = \boldsymbol{\tau}(\gamma(s_j))$, $j = 0, 1$, where $\boldsymbol{\tau}(x)$ denotes the orienting unit 2-vector for $T_x\Sigma$. Then, since $e, \gamma'(s_0), \gamma'(s_1)$ are orthonormal,

$$\begin{aligned} (1.11) \quad & |\boldsymbol{\tau}(\gamma(s_1)) - \boldsymbol{\tau}(\gamma(s_0))| = |v_1 \wedge \gamma'(s_1) - v_0 \wedge \gamma'(s_0)| \\ & \geq |(v_1 \cdot e) e \wedge \gamma'(s_1) - (v_0 \cdot e) e \wedge \gamma'(s_0)| = \sqrt{(e \cdot v_0)^2 + (e \cdot v_1)^2} \geq \theta, \end{aligned}$$

and hence

$$\left| \int_0^{s_1} \frac{d}{ds} \boldsymbol{\tau}(\gamma(s))\, ds \right| \geq \theta.$$

Therefore, since $|\frac{d}{ds}\boldsymbol{\tau}(\gamma(s))| \leq 2|\mathbf{A}(\gamma(s))|$, we deduce that

$$\theta \leq 2 \int_{\Gamma_t} |\mathbf{A}|,$$

and integrating over $t \in I$ and using (1.9) and the coarea formula we conclude the required inequality (1.7). □

Proof of Lemma 1.4. Here we are going to use the identity (1.5). We actually apply this identity separately to the two components $\widetilde{\Sigma}_1$, $\widetilde{\Sigma}_2$ obtained as the image of Σ_1, Σ_2 (as in the statement of Lemma 1.4) under an inversion in the sphere $B_\rho(0)$. (By a slight perturbation we may assume that $0 \notin \Sigma$.) Take points y_1, y_2 in $\widetilde{\Sigma}_1$, $\widetilde{\Sigma}_2$ respectively with $|y_j| \geq \theta^{-1}\rho$. (Such y_j exist because $\Sigma_j \cap B_{\theta\rho}(0) \neq \emptyset$, $j = 1, 2$.) Since

$$\mathcal{F}(\widetilde{\Sigma}_1) + \mathcal{F}(\widetilde{\Sigma}_2) \leq \mathcal{F}(\widetilde{\Sigma}) = \mathcal{F}(\Sigma),$$

we thus conclude from (1.5) with $\widetilde{\Sigma}_j$ in place of Σ and with y_j in place of y that

$$2\pi \leq \frac{1}{4}\mathcal{F}(\Sigma) + \sum_{j=1}^{2} \int_{\partial \widetilde{\Sigma}_j} |x - y_j|^{-2} \eta_j \cdot (x - y_j), \quad j = 1, 2,$$

where η_j is the unit conormal of $\partial\widetilde{\Sigma}_j$, $j = 1, 2$. Since $|y_j| \geq \theta^{-1}\rho$ and since $|\partial\widetilde{\Sigma}_j| = |\partial\Sigma_j| \leq \beta\rho$ we have

$$\sum_{j=1}^{2}\left|\int_{\partial\widetilde{\Sigma}_j} |x - y_j|^{-2}\eta_j \cdot (x - y_j)\right| \leq 4\theta\beta,$$

provided $\theta \in (0, \frac{1}{2})$. This gives the required inequality. □

2. Approximate Graphical Decomposition and Biharmonic Comparison

Here, as in the previous section, we continue to work with arbitrary smooth compact surfaces Σ in $\mathbb{R}^n$. The following lemma asserts that, in balls where the integral of the square length of the second fundamental form (i.e. $\int|\mathbf{A}|^2$) is small, we can decompose such a surface into a union of discs, each of which is well approximated by a graph of small Lipschitz norm. In this lemma B_ρ continues to denote the open ball of radius $\rho > 0$ (ρ given) in $\mathbb{R}^n$ with centre 0. Also, we adopt the convention that if L is a plane in $\mathbb{R}^n$ then we write $u = (u_1, \dots, u_n) \in C^2(\overline{\Omega}; L^\perp)$, where $\Omega \subset L$, if $u(x) \in L^\perp\, \forall x \in \Omega$. In this case we write

$$\text{graph}\, u = \{x + u(x) \colon x \in \Omega\}. \tag{2.1}$$

We need to include here the possibility that u is k-valued for some integer $k \geq 2$; in this case we write $u \in C^2(\overline{\Omega}; L^\perp)$ if for each $x_0 \in \overline{\Omega}$ there is $\sigma > 0$ such that $u(x) = \{u_1(x), \dots, u_k(x)\}$ for each $x \in \overline{B}_\sigma(x_0) \cap \overline{\Omega}$, where the u_j are C^2 functions on $\overline{B}_\sigma(x_0) \cap \overline{\Omega}$ with values in $L^\perp$. Then we again write (2.1), keeping in mind that now (2.1) says that graph u is locally expressed as the disjoint union of k single-valued graphs.

Lemma 2.1. *For any $\beta > 0$, there is $\epsilon_0 = \epsilon_0(n, \beta) > 0$ (independent of Σ, ρ) such that if $\epsilon \in (0, \epsilon_0]$, if $\partial\Sigma \cap \overline{B}_\rho = \emptyset$, if $0 \in \Sigma$, if $|\Sigma \cap \overline{B}_\rho| \leq \beta\rho^2$, and if $\int_{\Sigma\cap B_\rho} |\mathbf{A}| \leq \epsilon\rho$, then the following holds:*

There are pairwise disjoint closed sets $P_1, \dots, P_N \subset \Sigma$ with

$$\sum_{j=1}^{N} \operatorname{diam} P_j \leq C\epsilon^{1/2}\rho$$

and

$$\Sigma \cap B_{\rho/2}\backslash\Big(\bigcup_{j=1}^{N} P_j\Big) = \Big(\bigcup_{i=1}^{M} \operatorname{graph} u_i\Big) \cap B_{\rho/2},$$

where each $u_i \in C^\infty(\overline{\Omega}_i; L_i^\perp)$ *is a* k_i*-valued function for some* $k_i \geq 1$ $(k_i = 1\,\forall\, i$ *if* $n = 3)$*, with* L_i *a plane in* $\mathbb{R}^n$*,* Ω_i *a smooth bounded connected domain in* L_i *of the form* $\Omega_i = \Omega_i^0\backslash(\bigcup_k d_{i,k})$*, where* Ω_i^0 *is simply connected and* $d_{i,k}$ *are pairwise disjoint closed discs in* L_i *which do not intersect* $\partial\Omega_i^0$*, with* graph u_i *connected, and with*

$$\sum_{i=1}^{M} k_i \leq C\beta, \quad \sup_{\Omega_i} \rho^{-1}|u_i| + \sup_{\Omega_i} |Du_i| \leq C\epsilon^{1/2(2n-3)}.$$

If we also have $\int_{B_\rho} |\mathbf{A}|^2 \leq \epsilon^2$*, then* $k_i = 1\,\forall i$ *and, in addition to the above conclusions, for any* $\sigma \in (\rho/4, \rho/2)$ *such that* ∂B_σ *intersects* Σ *transversely and* $\partial B_\sigma \cap (\bigcup_j P_j) = \emptyset$*, we have*

$$\Sigma \cap \overline{B}_\sigma = \bigcup_{i=1}^{M} D_{\sigma,i},$$

where each $D_{\sigma,i}$ *is topologically a disc with* graph $u_i \cap \overline{B}_\sigma \subset D_{\sigma,i}$ *and*

$$D_{\sigma,i}\backslash \operatorname{graph} u_i$$

is a union of a subcollection of the P_j*, and each* P_j *is topologically a disc.*

(Note in particular this means that if $\widetilde{\Omega}_j \subset \Omega_j$ is the projection of graph $u_j \cap B_\sigma$ onto L_j and if Γ_j is the outermost component of $\partial\widetilde{\Omega}_j$, then $\partial D_{\sigma,j} = \operatorname{graph}(u_j|\Gamma_j) \subset \Sigma \cap \partial B_\sigma$.)

Roughly speaking the lemma says that $\Sigma \cap B_\sigma$ is a union of discs with smooth boundary contained in ∂B_σ, and each of the discs can be expressed as a graph with small gradient, together with some "pimples" P_j, the sum of diameters of which are small.

Proof. For any $\sigma > 0$, let $\Sigma_\sigma = \Sigma \cap \overline{B}_\sigma$. Let $\eta \in (0, \frac{1}{2})$ be for the moment arbitrary (we choose η to be a power of ϵ below), let $\sigma_0 \in (\frac{7}{8}\rho, \rho)$ be chosen such that Σ intersects ∂B_{σ_0} transversely, and let

$$S = \{\tau_1 \wedge \tau_2 \colon \tau_1, \tau_2 \in \mathbb{S}^{n-1}, |\tau_1| = |\tau_2| = 1, \tau_1 \cdot \tau_2 = 0\},$$

so that in particular $\boldsymbol{\tau}(x) \in S$ for each $x \in \Sigma$, and assume S is equipped with its usual metric. Then S is a compact manifold of dimension $2n-4$ and for any $\tau_0 \in S$ we define

$$f(x) = |\boldsymbol{\tau}(x) - \tau_0|, \quad x \in \Sigma_\rho.$$

Notice that f is then a smooth function on the region of Σ_ρ where it is non-zero, and $|\nabla f| \le |\mathbf{A}(x)|$. Thus by Sard's theorem and the coarea formula we have $t \in (\eta/2, \eta)$ such that the set $\Gamma_t = \{x \in \Sigma_{\sigma_0} \colon f(x) = t\}$ is contained in the union of finitely many pairwise disjoint Jordan curves and Jordan arcs, the endpoints of the arcs being in $\partial\Sigma_{\sigma_0}$, and

$$\mathcal{H}^1(\Gamma_t \cap \Sigma_{\sigma_0}) \le \frac{C}{\eta} \int_{\Sigma_{\sigma_0} \cap \{\frac{\eta}{2} < f(x) < \eta\}} |\mathbf{A}| \le \frac{C}{\eta} \int_{\Sigma_\rho} |\mathbf{A}| \le \frac{C}{\eta} \epsilon\rho. \tag{2.2}$$

Now cover all of S by balls $\mathbf{B}_{\eta/2}(\tau_j)$, $j = 1, \ldots, M$, with

$$M \le \frac{C(n)}{\eta^{2n-4}}.$$

Then corresponding to each $j = 1, \ldots, M$, by applying the above argument with τ_j in place of τ_0, there is $t_j \in (\eta/2, \eta)$ such that

$$\Gamma^{(j)} \equiv \{x \in \Sigma_{\sigma_0} \colon |\boldsymbol{\tau}(x) - \tau_j| = t_j\} \tag{2.3}$$

is contained in a finite union of Jordan curves, and

$$\mathcal{H}^1(\Gamma^{(j)}) \le \frac{C}{\eta} \int_{\Sigma_\rho} |\mathbf{A}|, \quad j = 1, \ldots, M.$$

Actually we note that by selecting the t_j successively, applying Sard's theorem at each stage, we can also arrange that each of the curves Γ_i is either disjoint from the remaining Γ_j or intersects them transversely in a finite set of points.

Hence in particular, selecting $\eta = \epsilon^{1/2(2n-3)}$, we have

$$\sum_{j=1}^{M} \mathcal{H}^1(\Gamma^{(j)}) \le C\epsilon^{1/2}\rho, \quad C = C(n). \tag{2.4}$$

Now let $\{Q_i\}_{i=1,\ldots,N}$ be the (finitely many) components of $\Sigma_{\sigma_0} \setminus \bigcup_j \Gamma^{(j)}$; notice that, since the $\mathbf{B}_{\eta/2}(\tau_j)$ cover all of S, we have (in view of the above choice of η)

$$\sup |\boldsymbol{\tau}(x) - \boldsymbol{\tau}(y)| \le C\epsilon^{1/2(2n-3)}, \quad x, y \in Q_i \tag{2.5}$$

for each i. Notice also that then each Q_i is an open subset of Σ_{σ_0} and

$$\partial Q_i \cap \overline{B}_{\sigma_0} \subset \left(\bigcup_{j=1}^{M} \Gamma^{(j)}\right) \cup \partial B_{\sigma_0}. \tag{2.6}$$

For each i pick a plane L_i containing a point $y_i \in Q_i \cap \overline{B}_{\sigma_0}$ with orienting 2-vector $\boldsymbol{\tau}(y_i)$, and select discs $\{d_{i,k}\}_{k=1,\dots,R_i} \subset L_i$ such that

$$\Pi_i\left(\bigcup_j \Gamma_j\right) \subset \bigcup_k d_{i,k}, \tag{2.7}$$

where Π_i is the orthogonal projection of $\mathbb{R}^n$ onto L_i, and such that

$$\sum_{k=1}^{R_i} \operatorname{diam}(d_{i,k}) \le C\epsilon^{1/2}\rho. \tag{2.8}$$

Now without loss of generality these discs can be selected to be pairwise disjoint; here we use the easily checked general fact that if $d_1, \dots, d_N$ are closed disks in the plane $\mathbb{R}^2$ then there is a pairwise-disjoint collection $\widetilde{d}_1, \dots, \widetilde{d}_M$ with $M \le N$, $\bigcup_{j=1}^M \widetilde{d}_j \supset \bigcup_{j=1}^N d_j$ and with $\sum_{j=1}^M \operatorname{diam} \widetilde{d}_j \le \sum_{j=1}^N \operatorname{diam} d_j$. (This is checked by induction on N, starting with $N = 2$.) We therefore assume in the following discussion that the discs $d_{i,k}$, $k = 1, \dots, R_i$ are pairwise disjoint for each given $i = 1, \dots, N$. Also, by (2.5), we know that if there is a point in $x \in L_i \backslash (\bigcup_k d_{i,k})$ such that $\Pi_i^{-1}(x) \cap \overline{B}_{3\rho/4} \cap Q_i$ has k distinct points, then $Q_i \cap \Pi_i^{-1}(y)$ has $\ge k$ distinct points for each $y \in \overline{B}_{\rho/32}(x) \backslash (\bigcup_k d_{i,k})$; then since the area of Σ_ρ is bounded by $\beta\rho^2$, we must have that (i) there is a bound

(2.9)
$$\text{number of points in } \Pi_i^{-1}(x) \cap Q_i \cap \overline{B}_{3\rho/4} \le C(n, \beta), \quad x \in L_i \backslash (\bigcup_k d_{i,k}),$$

and (ii) there is an upper bound

$$M \le C(n, \beta) \tag{2.10}$$

on the number M of distinct Q_i such that $Q_i \cap \overline{B}_{3\rho/4} \backslash \Pi_i^{-1}(\bigcup_k d_{i,k}) \ne \emptyset$.

Now pick $\sigma_1 \in (\frac{5}{8}\rho, \frac{3}{4}\rho)$ such that Σ intersects ∂B_{σ_1} transversely, and such that $\partial B_{\sigma_1} \cap L_i$ intersects none of the discs $d_{i,k}$, $k = 1, \dots, R_i$, $i = 1, \dots, N$. Suppose that the labelling is such that $\{i\colon Q_i \cap \overline{B}_{\sigma_1\rho} \backslash \Pi_i^{-1}(\bigcup_{k=1}^{R_i} d_{i,k}) \ne \emptyset\} =$

$\{1,\dots,M\}$, and for each $i=1,\dots,M$ let $\{Q_i^q\}_{q=1,\dots,K_i}$ denote the components of $Q_i\backslash\Pi_i^{-1}(\bigcup_{k=1}^{R_i} d_{i,k})$. In view of (2.9) and (2.5) we know that

$$Q_i^q\cap\overline{B}_{\sigma_1}=\bigcup_{s=1}^{m_{i,q}}\operatorname{graph}u_i^{s,q}\cap\overline{B}_{\sigma_1}, \tag{2.11}$$

where each $u_i^{s,q}$ is a $k_{i,s,q}$-valued function over $\Omega_i^q\subset L_i$ for some $k_{i,s,q}\ge 1$, and (by (2.5))

$$\sup_{\Omega_i^s}|\nabla u_i^{s,q}|+\sup_{x,y\in\Omega_i^{s,q}}|u_i^{s,q}(x)-u_i^{s,q}(y)|\le C\epsilon^{1/2(2n-3)}\rho. \tag{2.12}$$

Also from (2.9)

$$\sum_{\{(i,s,q):\ \operatorname{graph}u_i^{s,q}\cap\overline{B}_{\sigma_1}\neq\emptyset\}}k_{s,i,q}\le C\beta. \tag{2.13}$$

Notice also that each $\partial d_{i,k}$ which is one of the boundary components of Ω_i^q lifts via $u_i^{s,q}$ to a curve $\gamma_{i,q,k,s}=\operatorname{graph}u_i^{s,q}|\partial d_{i,k}$ on $\Sigma_{3\rho/4}$. By (2.5) and (2.8), for each i we have

$$\sum_{q,k,s}\operatorname{diam}\gamma_{i,q,k,s}\le C\epsilon^{1/2}\rho,\quad C=C(n,\beta). \tag{2.14}$$

Notice also that by construction the entire collection $\{\gamma_{i,q,k,s}\}$ over all possible i,q with $Q_i^q\cap\overline{B}_{\sigma_1}\neq\emptyset$ and all possible k,s is a pairwise disjoint collection. Notice further that if $J_{i,j}$ denotes the components of $Q_i\cap\overline{B}_{\sigma_1}\backslash(\bigcup_{s,q}\operatorname{graph}u_i^{s,q})$, then each $J_{i,j}$ is a smooth surface in $\overline{B}_{\sigma_1}$ with boundary components in $(\bigcup\gamma_{i,q,k,s})\cup\partial B_{\sigma_1}\cup(\bigcup_j\Gamma_j)$.

Next we claim that for each $i=1,\dots,M$ we have

$$\sum_j|J_{i,j}\cap\overline{B}_{9\rho/16}|\le C\epsilon^{1/2}\rho^2. \tag{2.15}$$

To see this, note first that $\sum_j|J_{i,j}|\le|\Sigma_\rho|\le\beta\rho^2$, hence we can pick $\sigma_2\in(\frac{9}{16}\rho,\sigma_1)$ such that $\bigcup_j J_{i,j}$ intersects ∂B_{σ_2} transversely and length$((\bigcup_j J_{i,j})\cap\partial B_{\sigma_2})\le C\beta\rho$. In view of (2.14) we can also select σ_2 to ensure that

$$\bigcup_{k,q,s}\gamma_{i,q,k,s}\cap\partial B_{\sigma_2}=\emptyset. \tag{2.16}$$

Then apply (1.1′) with $J_{i,j}\cap\overline{B}_{\sigma_2}$ in place of Σ and with $\Phi(x)=\Pi_i(x)-\Pi_i(z_{i,j})$, where $z_{i,j}$ is any fixed point of $J_{i,j}$. By connectedness of $J_{i,j}$ we have

$\Pi_i(J_{i,j}) \subset d_{i,\ell}$ for some ℓ (so that $|\Phi(x)| \leq \operatorname{diam} d_{i,\ell}$ on $J_{i,j}$) and by (2.5) $\operatorname{div}_\Sigma \Pi_i(x) \geq 2 - C\epsilon^\alpha$ on $J_{i,j}$, hence (1.1′) gives

$$\sum_{\{j:\, \Pi_i(J_{i,j}) \subset d_{i,\ell}\}} |J_{i,j} \cap \overline{B}_{\sigma_1}| \leq C\beta\rho \operatorname{diam} d_{i,\ell}. \tag{2.17}$$

After summing on ℓ this gives (2.13) as required by virtue of (2.8). Notice that by the corollary to Lemma 1.2 and by (2.13) we have the implication

$$J_{i,j} \cap \overline{B}_{9\rho/16} \neq \emptyset \Rightarrow J_{i,j} \cap \partial B_{\sigma_1} = \emptyset. \tag{2.18}$$

Now let P_j be the components of $\Sigma \cap \overline{B}_{\sigma_1} \backslash (\bigcup_{i,s,q} \operatorname{graph} u_i^{s,q})$, and note that by (2.18) we have

$$P_j \cap \partial B_{9\rho/16} \neq \emptyset \Rightarrow P_j \cap \partial B_{\sigma_1} = \emptyset. \tag{2.19}$$

Thus for any j such that $P_j \cap \overline{B}_{9\rho/16} \neq \emptyset$, we have that P_j is a compact manifold with boundary equal to a finite subcollection of the curves $\gamma_{i,q,k,s}$, and hence by (2.14) and Lemma 1.2 (with P_j in place of Σ, and keeping in mind that at most 2 of the P_j can have a given $\gamma_{i,q,k,s}$ as a boundary component), we deduce

$$\sum_{\{j:\, P_j \cap \overline{B}_{9\rho/16} \neq \emptyset\}} \operatorname{diam} P_j \leq C\epsilon^{1/2}\rho. \tag{2.20}$$

After some relabelling (and possibly a translation of the planes L_i), (2.19) and (2.20) imply that there exist planes $L_1, \ldots, L_M$ (where the L_j are not necessarily distinct) and k_i-valued functions $u_i \in C^2(\overline{\Omega}_i; L_i^\perp)$, where each Ω_i is a smooth connected domain of the form $\Omega_i^0 \backslash (\bigcup_k d_{i,k})$ with Ω_i^0 a simply connected domain and $d_{i,k}$ are pairwise disjoint closed discs in L_i such that $d_{i,k} \cap \partial\Omega_i^0 = \emptyset \ \forall k$, and where u_i satisfies

$$\sum_{i=1}^M k_i \leq C, \quad \sup_{\Omega_i} \rho^{-1}|u_i| + \sup_{\Omega_i} |\nabla u_i| \leq C\epsilon^{1/2(2n-3)}, \tag{2.21}$$

and

$$\left(\Sigma \backslash \bigcup_{i=1}^M \operatorname{graph} u_i\right) \cap \overline{B}_{9\rho/16} = \left(\bigcup_{j=1}^R P_j\right) \cap \overline{B}_{9\rho/16}, \tag{2.22}$$

where $P_1, \dots, P_R$ are pairwise disjoint compact surfaces with boundary and

$$\sum_{j=1}^{R} \operatorname{diam} P_j \leq C\epsilon^{1/2}\rho. \tag{2.23}$$

In view of (2.21), (2.22), and (2.23) the proof of the first part of the lemma is complete; in fact there is evidently a set S of measure $\geq \frac{\rho}{32}$ in $(\frac{\rho}{2}, \frac{9\rho}{16})$ such that

$$\Big(\bigcup_j P_j\Big) \cap \partial B_\sigma = \emptyset, \quad \forall \sigma \in S, \tag{2.24}$$

such that Σ intersects ∂B_σ transversely, and such that

$$\int_{\Sigma \cap \partial B_\sigma} |\mathbf{A}|^2 \leq C\frac{\epsilon^2}{\rho}, \quad \forall \sigma \in S. \tag{2.25}$$

Since by (2.21) we have $M \leq C(n,\beta)$ it of course follows that $\exists\, \widetilde{S} \subset S$ with measure $\geq \frac{\rho}{64}$ such that $\forall\, i = 1, \dots, M$ either

$$L_i \cap \overline{B}_\sigma = \emptyset \quad \forall \sigma \in \widetilde{S}, \tag{2.26}$$

or

$$\operatorname{diam}(L_i \cap \overline{B}_\sigma) \geq C^{-1}\rho, \quad \forall \sigma \in \widetilde{S}, \tag{2.27}$$

with $C = C(n,\beta)$. Now we relabel $L_1, \dots, L_M$ so that (2.27) holds for $i = 1, \dots, M_0$ and (2.26) holds for $i = M_0 + 1, \dots, M$. Select $\rho_0 \in \widetilde{S}$, take $j \in \{1, \dots, M_0\}$, and let z_j be the center of the disc $L_j \cap \overline{B}_{\rho_0}$. Without loss of generality we assume $L_j = z_j + \mathbb{R}^2 \times \{0\}$. Let $\alpha_j\rho_0$ be the radius of $L_j \cap \overline{B}_{\rho_0}$. Then notice that any point in $\operatorname{graph} u_j \cap \partial B_{\rho_0}$ can be uniquely written as

(2.28)
$$z_j + (1 - \sigma_j(\theta))\alpha_j\rho_0 e^{i\theta} + u_j(z_j + (1 - \sigma_j(\theta))\alpha_j\rho_0 e^{i\theta}), \quad 0 \leq \theta \leq 2k_j\pi,$$

where $\sigma_j(\theta) \in (-\frac{1}{8}, \frac{1}{8})$ is to be determined implicitly by the relation

$$|(1 - \sigma_j(\theta))\alpha_j\rho_0 e^{i\theta}|^2 + |u_j(z_j + (1 - \sigma_j(\theta))\alpha_j\rho_0 e^{i\theta})|^2 = \alpha_j^2\rho_0^2;$$

that is, by the relation

$$\sigma_j \in 1 - \sqrt{1 - \alpha_j^{-2}\rho_0^{-2}|u_j(z_j + (1 - \sigma_j)\alpha_j\rho_0 e^{i\theta})|^2}, \tag{2.29}$$

keeping in mind that the right side is k_j-valued for any given $\sigma_j \in (-\frac{1}{8}, \frac{1}{8})$. Notice that the smooth solution of (2.29) (which is unique modulo choosing an initial value for σ_j from the k_j possible choices) satisfies

$$|\sigma_j'(\theta)| \le C|\nabla u_j| \tag{2.30}$$

and

$$|\sigma_j''(\theta)| \le C(|\nabla u_j| + \rho|\nabla^2 u_j|), \tag{2.31}$$

where all functions on the right side are evaluated at the point $p_j(\theta) = z_j + (1 - \sigma_j(\theta))\alpha_j\rho_0 e^{i\theta}$. Now by definition of second fundamental form we have

$$|\nabla^2 u_j(x)| \le C|\mathbf{A}(x + u_j(x))|,$$

and hence (2.25) implies

$$\int_0^{2k_j\pi} |\nabla^2 u_j(p_j(\theta))|^2 \le C\frac{\epsilon^2}{\rho^2}. \tag{2.32}$$

Combining (2.21), (2.30), (2.31) and (2.32) we conclude that

$$\int_0^{2k_j\pi} (|\sigma_j'|^2 + |\sigma_j''|^2) \le C\epsilon^{1/2n-3}. \tag{2.33}$$

It now follows directly from (2.21), (2.28) and (2.33) that if $\boldsymbol{\kappa}_j$ denotes the curvature vector of the curve $\gamma(\theta) \equiv p_j(\theta) + u_j(p_j(\theta))$ (as a space curve in $\mathbb{R}^n$), then

$$\boldsymbol{\kappa}_j = (\alpha_j r_0)^{-1}\eta_j + E_j,$$

where η_j is the inward pointing (tangent to Σ_{σ_0}) unit normal of $\partial\Sigma_{\rho_0}$ and where $\int_{\partial\Sigma_{\sigma_0}} |E_j| \le C\epsilon^{1/2(2n-3)}$. Thus if κ denotes the geodesic curvature of $\partial\Sigma_{\rho_0}$ (so that $\kappa = \eta_j \cdot \boldsymbol{\kappa}_j$ on the boundary component $\operatorname{graph} u_j \cap \partial B_{\rho_0}$ of $\partial\Sigma_{\rho_0}$), then

$$\left|\int_{\partial\Sigma_{\rho_0}} \kappa - 2\pi\sum_{j=1}^{M_0} k_j\right| \le C\epsilon^{1/2(2n-3)}.$$

Then by virtue of the Gauss-Bonnet formula $\int_{\Sigma_{\rho_0}} K = 2\pi(2M_1 - 2g - M_0) - \int_{\partial\Sigma_{\rho_0}} \kappa$, with M_1 equal to the number of components of Σ_{ρ_0}, with $M_0(\ge M_1)$ equal to the number of components of $\partial\Sigma_{\rho_0}$, and with g the total genus of Σ_{ρ_0}, and keeping in mind that $\int_{\Sigma_{\rho_0}} |K| \le \epsilon^2$ by hypothesis, we conclude that (for ϵ small enough)

$$2M_1 - 2g - M_0 - \sum_{j=1}^{M_0} k_j = 0,$$

and hence that $g = 0$, $k_j = 1$ for each j and $M_1 = M_0$. This evidently completes the proof. □

Next we derive an important inequality involving biharmonic functions.

Lemma 2.2. *Let $\Sigma \subset \mathbb{R}^n$ be smooth embedded, $\xi \in \mathbb{R}^n$, L a plane containing ξ, $u \in C^\infty(U)$ for some open (L-)neighborhood U of $L \cap \partial B_\rho(\xi)$, and*

$$\operatorname{graph} u \subset \Sigma, \qquad |Du| \le 1.$$

Also, let $w \in C^\infty(L \cap \overline{B}_\rho(\xi))$ satisfy

$$\begin{cases} \triangle^2 w = 0 & \text{on } L \cap B_\rho(\xi) \\ w = u, \ \ Dw = Du & \text{on } L \cap \partial B_\rho(\xi)\,. \end{cases}$$

Then

$$\int_{L \cap B_\rho(\xi)} |D^2 w|^2 \le C\rho \int_\Gamma |\mathbf{A}|^2 d\mathcal{H}^1,$$

where $\Gamma = \operatorname{graph}(u|L \cap \partial B_\rho(\xi))$, $\mathbf{A}$ is the second fundamental form of Σ, and $\mathcal{H}^1$ is 1-dimensional Hausdorff measure (i.e. arc-length measure) on Γ; C is a fixed constant independent of Σ, ρ.

Remark 2.1. Of course there *exists* a w as above, because u is C^∞, so we can use the existence and regularity theory for the Dirichlet problem; the solution w is also clearly *unique*.

Proof of Lemma 2.2. Let $\Omega = L \cap B_\rho(\xi)$. Recall that the function w minimizes $\int_\Omega |\triangle w|^2$ subject to the given boundary conditions, and hence by the identity

$$\int_\Omega \sum_{i,j} ((D_{ij} w)^2 - (D_{ij} v)^2) = \int_\Omega ((\triangle w)^2 - (\triangle v)^2),$$

valid for any $v, w \in C^2(\overline{\Omega})$ with $Dv = Dw$ on $\partial\Omega$, we see that w also minimizes the integral $\int_\Omega \sum_{i,j} (D_{ij} v)^2$ over all $v \in C^2(\overline{\Omega})$ with $Dv = Dw$ on $\partial\Omega$.

Then, after rescaling so that $\rho = 1$, by the appropriate Sobolev-space trace lemma—see e.g. [TF, 26.5, 26.9 with $m = 2$]—we have, with $\Omega_1 = L \cap B_1(\xi)$ and $\gamma = L \cap \partial B_1(\xi)$,

$$\int_{\Omega_1} |D^2 w|^2 \le C \left(|u|^2_{H^{3/2}(\gamma)} + |Du|^2_{H^{1/2}(\gamma)} \right).$$

Applying the same to $w - \ell$ (ℓ any linear function + constant) we get

$$\int_{\Omega_1} |D^2 w|^2 \le C \int_\gamma \left((u-\ell)^2 + (Du - D\ell)^2 + |D^2 u|^2 \right).$$

By selecting ℓ suitably we can then establish that the first two terms on the right are dominated by a fixed multiple of the third. Thus (in the original scale)

$$\int_\Omega |D^2 w|^2 \le C \int_\gamma |D^2 u|^2.$$

Since $|Du|^2 \le 1$ on γ we also have $|D^2 u|^2(x) \le C\,|\mathbf{A}|^2(X)$, where X is the point $(x, u(x))$ of graph u corresponding to $x \in \gamma$, hence Lemma 2.2 follows. □

3. Regularity of measure-theoretic limits of minimizing sequences

A sequence of compact embedded surfaces $\Sigma_k \subset \mathbb{R}^n$ with $\partial \Sigma_k = 0$ is called a genus g minimizing sequence for $\mathcal{F}$ if genus $\Sigma_k = g$ $\forall k$ and if

$$\mathcal{F}(\Sigma_k) \le \beta_g + \epsilon_k, \quad \epsilon_k \downarrow 0.$$

By translation and scaling we can (and we shall) assume

$$0 \in \Sigma_k, \quad |\Sigma_k| = 1.$$

Notice that then by Lemma 1.1 we have a fixed constant $C > 0$ such that

(*) $$C^{-1} \le \operatorname{diam} \Sigma_k \le C.$$

Our main result here is the following:

Theorem 3.1. *Given any genus g minimizing sequence Σ_k as above, there is a subsequence $\Sigma_{k'}$, and a compact embedded real analytic surface Σ such that $\Sigma_{k'} \to \Sigma$ both in the Hausdorff distance sense and in the measure-theoretic sense that*

$$\int_{\Sigma_{k'}} f \to \int_\Sigma f$$

for each fixed continuous f on $\mathbb{R}^n$. This Σ has genus $g_0 \le g$, and Σ minimizes $\mathcal{F}$ relative to all compact smooth embedded genus g_0 surfaces $\widetilde{\Sigma} \subset \mathbb{R}^n$.

Remark 3.1. It can of course happen that $g_0 = 0$ (and Σ is a round sphere) even if $g \ge 1$. This is a problem in proving existence of the required genus 1 (or higher genus) minima which we show how to overcome in the next section.

Proof. We first prove that u is a $C^{1,\alpha} \cap W^{2,2}$ surface. First note that since $|\Sigma_k| = 1$ we may choose a subsequence $\Sigma_{k'}$ such that the corresponding sequence of measures $\mu_{k'}$, given by $\mu_{k'}(A) = |A \cap \Sigma_{k'}|$ for Borel sets $A \subset \mathbb{R}^n$, converges to a Borel measure μ of compact support. Thus

$$\int_{\Sigma_{k'}} f \to \int_{\mathbb{R}^n} f \, d\mu$$

for each fixed continuous function f in $\mathbb{R}^n$, and by (*) the support of μ is compact.

In spt μ (the support of μ) we say ξ is a *bad* point relative to a preassigned number $\epsilon > 0$ if

$$\lim_{\rho\downarrow 0} \left(\liminf_{k'\to\infty} \int_{\Sigma_{k'} \cap B_\rho(\xi)} |\mathbf{A}_{k'}|^2 \right) > \epsilon^2,$$

where $\mathbf{A}_k$ is the second fundamental form of Σ_k. Evidently, since, by the Gauss-Bonnet theorem,

$$\tfrac{1}{4} \textstyle\int_{\Sigma_k} |\mathbf{A}_k|^2 = \mathcal{F}(\Sigma_k) - \pi(2 - 2g)$$

we have that $\int_{\Sigma_k} |\mathbf{A}_k|^2$ is bounded, and hence there are only finitely many bad points for each $\epsilon > 0$. Indeed if $\xi_1, \dots, \xi_N$ are distinct bad points, let $\rho = \min_{i\neq j} |\xi_i - \xi_j|$, and note that for k' sufficiently large we have $\int_{\Sigma_{k'} \cap B_\rho(\xi)} |\mathbf{A}_{k'}|^2 > \epsilon^2$, so by summing over j and using the fact that $B_\rho(\xi_i) \cap B_\rho(\xi_j) = \emptyset$ for each $i \ne j$, we obtain

$$N\epsilon^2 \le \int_{\Sigma_{k'}} |\mathbf{A}_{k'}|^2 \equiv 4\mathcal{F}(\Sigma_{k'}) - 4\pi(2 - 2g),$$

so we have an upper bound on N in terms of ϵ. Denoting the subsequence simply by Σ_k, we can actually assume

$$\lim_{\rho\downarrow 0} \left(\liminf_{k\to\infty} \int_{\Sigma_k \cap B_\rho(\xi)} |\mathbf{A}_k|^2 \right) > \epsilon^2$$

for the finitely many bad points $\xi = \xi_1, \dots, \xi_P$ $(P = P(\epsilon))$.

On the other hand for any $\xi \in \operatorname{spt}\mu\backslash\{\xi_1, \dots, \xi_P\}$ we can select $\rho(\xi, \epsilon) > 0$ such that for $\rho \le \rho(\xi, \epsilon)$ we have $\int_{\Sigma_k \cap B_\rho(\xi)} |\mathbf{A}_k|^2 \le \epsilon^2$ for infinitely many k, and hence the last part of Lemma 2.1 is applicable to Σ_k in $B_\rho(\xi)$ for infinitely many k. At the same time we have, since $\beta_g < 8\pi$, that we can apply Lemma 1.4 to deduce that for large enough k and for small enough θ (θ fixed, independent of k, ϵ, ξ), only one of the disks $D_j^{(k)}$, say $D_1^{(k)}$, given by applying Lemma 2.1 can intersect the ball $B_{\theta\,\rho}(\xi)$. Thus, in accordance with Lemma 2.1, for ϵ small enough (which we subsequently assume), for infinitely many k there is a plane L_k containing ξ and a $C^\infty(\bar{\Omega}_k)$ function $u_k\colon \overline{\Omega}_k \to L_k^\perp$ ($L_k^\perp$ the subspace of vectors orthogonal to L_k) with

$$\begin{gathered}\rho^{-1}|u_k| + |Du_k| \le C\epsilon^{1/2(2n-3)},\\ (\operatorname{graph} u_k \cup_j P_{k,j}) \cap B_\sigma(\xi) = D_1^{(k)} \cap B_\sigma(\xi),\\ \sum_j \operatorname{diam} P_{k,j} \le C\epsilon^{1/2}\rho,\end{gathered} \tag{3.1}$$

where each $P_{k,j}$ is diffeomorphically a closed disk disjoint from $\operatorname{graph}(u_k|\Omega_k)$, and where $\sigma \in (\theta\rho/2, \theta\rho)$ does not depend on k.

With $C_\sigma(\xi) \equiv \{x + y\colon x \in B_\sigma(\xi) \cap L_k,\ y \in L_k^\perp\}$, the selection principle of Appendix B guarantees that we can then choose a set $T \subset (\theta\rho/2, \theta\rho)$ of measure $\ge \theta\rho/8$ such that for each $\sigma \in T$ we have $\partial C_\sigma(\xi) \cap P_{k,j} = \emptyset$ for infinitely many k, and hence for any such σ for infinitely many k we can apply Lemma 5 to obtain a biharmonic function w_k on $B_\sigma(\xi) \cap L_k$ such that

$$\int_{L_k \cap B_\sigma(\xi)} |D^2 w_k|^2 \le C \int_{\Gamma_k} |\mathbf{A}_k|^2.$$

Letting $\widetilde{\mathbf{A}}_k$ be the second fundamental form of $\operatorname{graph} w_k$, we then in particular have

$$\int_{\operatorname{graph} w_k} |\widetilde{\mathbf{A}}_k|^2 \le C \int_{\Gamma_k} |\mathbf{A}_k|^2.$$

On the other hand Σ_k is a minimizing sequence for the functional $\frac{1}{4}\int_\Sigma |\mathbf{A}|^2$, and hence the $C^{1,1}$ composite surface $\widetilde{\Sigma}_k = (\Sigma_k \backslash D_1^{(k)}) \cup \operatorname{graph} w_k$ satisfies

$$\mathcal{F}(\widetilde{\Sigma}_k) \ge \mathcal{F}(\Sigma_k) - \epsilon_k,\ \ \epsilon_k \downarrow 0,$$

so that

$$\int_{\operatorname{graph} w_k} |\widetilde{\mathbf{A}}_k|^2 \ge \int_{D_1^{(k)}} |\mathbf{A}_k|^2 - \epsilon_k.$$

Thus we conclude that for infinitely many k

$$\int_{\Sigma_k \cap B_\sigma(\xi)} |\mathbf{A}_k|^2 \le C \int_{\partial D_1^{(k)}} |\mathbf{A}_k|^2 + \delta_k,$$

where $\delta_k \downarrow 0$. Since σ was selected arbitrarily from the set T of Lebesgue measure $\ge \theta\,\rho/8$ in the interval $(\theta\rho/2, \theta\rho)$ we can arrange that

$$\int_{\partial D_1^{(k)}} |\mathbf{A}_k|^2 \le 4 \int_{\Sigma_k \cap B_{\theta\rho}(\xi) \setminus B_{\theta\rho/2}(\xi)} |\mathbf{A}_k|^2,$$

for infinitely many k, so that in fact we get, for $\rho \le \theta\rho(\epsilon)$ arbitrary, and for infinitely many k (depending on ρ),

$$\int_{\Sigma_k \cap B_{\rho/2}(\xi)} |\mathbf{A}_k|^2 \le C \int_{\Sigma_k \cap B_\rho(\xi) \setminus B_{\rho k}(\xi)} |\mathbf{A}_k|^2 + \delta_k,$$

where $\delta_k \downarrow 0$. Notice that by adding C times the left side to both sides of this inequality (i.e., by "hole filling") we deduce that

$$\int_{\Sigma_k \cap B_{\rho/2}(\xi)} |\mathbf{A}_k|^2 \le \gamma \int_{\Sigma_k \cap B_\rho(\xi)} |\mathbf{A}_k|^2 + \delta_k$$

for infinitely many k, where $\gamma = C/(1+C)$ is a fixed constant in the interval $(0, 1)$.

We also need to make the remark that $\rho(\xi, \epsilon)$ above merely had to be chosen so that $\int_{\Sigma_k \cap B_\rho(\xi)} |\mathbf{A}_k|^2 \le \epsilon^2$ for infinitely many k. In particular this means that if $\xi_0 \in \operatorname{spt}\mu \setminus \{\xi_1, \dots, \xi_P\}$, then we may take $\rho(\xi, \epsilon) = \rho(\xi_0, \epsilon)/2$ for any $\xi \in \operatorname{spt}\mu \cap B_{\rho(\xi_0,\epsilon)/2}(\xi_0)$. Thus we see that the following is established:

If we let

$$\psi(\xi, \rho) = \liminf_{k\to\infty} \int_{\Sigma_k \cap B_\rho(\xi)} |\mathbf{A}_k|^2,$$

then we have for all $\xi_0 \in \operatorname{spt}\mu \setminus \{\xi_1, \dots, \xi_P\}$ and all $\rho \le \theta\rho(\xi_0)/2$, and all $\xi \in \operatorname{spt}\mu \cap B_{\rho(\xi_0)/2}(\xi_0)$ that

$$\psi(\rho/2, \xi) \le \gamma\psi(\rho, \xi)$$

for some fixed $\gamma \in (0, 1)$ independent of ρ, ξ. Thus

$$\psi(\rho, \xi) \le C(\rho/\rho_0)^\alpha \psi(\rho_0, \xi) \le C(\rho/\rho_0)^\alpha \psi(\rho(\xi_0), \xi_0) \tag{3.2}$$

for some $\alpha \in (0, 1)$ and for all such ρ, ξ, where $\rho_0 = \theta\rho(\xi_0)/2$.

Henceforth $\xi_0 \in \operatorname{spt}\mu\backslash\{\xi_1,\ldots,\xi_P\}$ is fixed and we take $\xi \in \operatorname{spt}\mu\cap B_{\rho(\xi_0)/2}(\xi_0)$ and $\rho \in (0, \rho(\xi_0)/2)$, and let

$$\alpha_k = \alpha_k(\rho,\, \xi) = \int_{\Sigma_k\cap B_\rho(\xi)} |\mathbf{A}_k|^2 \ (< \epsilon^2 \text{ for infinitely many } k\ ),$$

and let L_k, Ω_k, u_k, ρ_k, d_k^i be as in (3.1). Also let $U_k = P_k(\operatorname{graph} u_k \cap B_\sigma(\xi))$ (σ as in (1)), and let $\bar{u}_k$ denote an extension of $u_k|U_k$ to all of L_k such that

$$\rho^{-1}\sup|\bar{u}_k| + |D\bar{u}_k| \le C\epsilon^{1/2(2n-3)}. \tag{3.3}$$

(It is easy to see that such an extension exists—first extend $u_k|\partial d_{i,k}$ to $d_{i,k}$ appropriately to give $\bar{u}_k$ on $U_k \cup (\bigcup_i d_{i,k})$.) Since $\sum_i \operatorname{diam} d_k^i \le C\sqrt{\alpha_k}\,\rho$ (by Lemma 2.1), the variant of Poincaré's inequality in the Appendix A below gives

$$\inf_{\lambda\in\mathbb{R}} \int_{\Omega_k} |f-\lambda|^2 \le C\rho^2 \int_{\Omega_k} |Df|^2 + C\sqrt{\alpha_k}\, \sup|f|^2\rho^2,$$

with C independent of k. Applying this with $f = D_j u_k$, we have a constant vector η_k so that

$$\int_{\Omega_k} |Du_k - \eta_k|^2 \le C\rho^2 \int_{\Omega_k} |D^2u_k|^2 + C\sqrt{\alpha_k}\,\rho^2 \le C\rho^2\sqrt{\alpha_k}.$$

Then since, by Lemma 2.1, $\sum_i |d_k^i| \le C\sqrt{\alpha_k}\,\rho^2$, we have

$$\int_{L_k\cap B_{\rho_k}(\xi)} |D\bar{u}_k - \eta_k|^2 \le C\sqrt{\alpha_k}\,\rho^2,$$

so finally, by (3.2), for suitable $\gamma > 0$

$$\int_{B_{\theta\rho/2}(\xi)} |D\bar{u}_k - \eta_k|^2 \le C\rho^{2+\gamma}. \tag{3.4}$$

Take a subsequence so that the L_k converge to L, $\eta_k \to \eta$, and so that (by the Arzela-Ascoli theorem) $\operatorname{graph}\bar{u}_k$ converges in the Hausdorff distance sense to $\operatorname{graph} u$, with $u \in \operatorname{Lip} L$, $\rho^{-1}\sup|u| + \sup|Du| \le C\epsilon^{1/2(2n-3)}$ and

$$\int_{B_{\theta\rho/2}(\xi)\cap L} |Du - \eta|^2 \le C\rho^{2+\gamma}. \tag{3.5}$$

In measure theoretic terms (provided we take ϵ small enough to begin with) this means we have established that for all $\xi \in \operatorname{spt}\mu \cap B_{\theta\rho(\xi_0)/2}(\xi_0)$ and for all $\rho < \theta\rho(\xi_0)/4$

$$\mathcal{H}^2 \llcorner (\Sigma_k \cap B_\rho(\xi)) = \mathcal{H}^2 \llcorner\ (\operatorname{graph}\bar{u}_k \cap B_\rho(\xi)) + \theta_k,$$

where θ_k is a signed measure with total mass $\le C\rho^{2+\gamma}$ and (taking limits in the measure-theoretic sense)

$$\mu \llcorner B_\rho(\xi) = \mathcal{H}^2 \llcorner (\operatorname{graph} u \cap B_\rho(\xi)) + \theta, \tag{3.6}$$

where total mass of $\theta \le C\,\rho^{2+\gamma}$ and where u satisfies (3.5) (with $\eta = \eta(\rho, \xi)$). Of course all the constants C here are independent of ρ, ξ, provided we continue to assume that $\xi \in \operatorname{spt}\mu \cap B_{\rho(\xi_0)/2}(\xi_0)$ and $\rho \in (0, \rho(\xi_0)/2)$.

In view of the arbitrariness of ρ, ξ it then follows from (3.5) and (3.6) that if ϵ is small enough, firstly

(3.7)
$$\begin{cases} \textit{the measure } \mu \textit{ has a unique multiplicity } 1 \textit{ tangent plane at each point} \\ \xi \in \operatorname{spt}\mu \cap B_{\theta\rho(\xi_0)/4}(\xi_0) \textit{ with normal-space } N(\xi) \textit{ such that} \\ \|N(\xi_1) - N(\xi_2)\| \le C|\xi_1 - \xi_2|^\gamma, \quad \xi_1, \xi_2 \in \operatorname{spt}\mu \cap B_{\theta\rho(\xi_0)/4}(\xi_0), \end{cases}$$

and also that then for any preassigned $\delta > 0$ there is a neighbourhood U of x_0 such that

$$\mu \llcorner U = \mathcal{H}^2 \llcorner (\Sigma \cap U), \tag{3.8}$$

where Σ is an embedded $C^{1,\gamma/2}$ surface expressible as graph w for some $w \in C^{1,\gamma/2}(U \cap L_0)$ with $\sup_{U\cap L_0} |Dw| \le \delta$, where L_0 is the tangent plane of μ at ξ_0.

On the other hand, since $\int_{\Sigma\cap B_\rho(\xi)} \mathbf{H}_k^2 \le C\rho^\gamma$ (by (3.2)), where $\mathbf{H}_k$ denotes the mean-curvature vector of Σ_k, and since Σ (with multiplicity 1) is the varifold limit of Σ_k in $B_{\theta\rho(\xi_0)/8}(\xi_0)$, we deduce that Σ has generalized mean curvature $\mathbf{H}$ satisfying

$$\int_{\Sigma\cap B_\rho(\xi)} \mathbf{H}^2 \le C\rho^\gamma,$$

for $\xi = x + w(x) \in \operatorname{graph} w$ such that $\operatorname{dist}(x, \partial U) > 2\rho$, and that w is a C^1 weak solution of the mean curvature system

$$\sum_{i,j}^{2} D_i\left(\sqrt{g}g^{ij}D_jw\right) = \sqrt{g}\,\mathbf{H},$$

(where $(g^{ij}) = (g_{ij})^{-1}$, $g = \det(g_{ij})$, $g_{ij} = \det(\delta_{ij} + D_iw \cdot D_jw)$). It then follows from a standard difference quotient argument (e.g. by the obvious modifications of the argument used in [GT, Theorem 8.8]) that $w \in W^{2,2}_{loc}(U)$ and that each $w_\ell = D_\ell w$ satisfies a system of the form

$$\sum_{i,j=1}^{2} D_i(g^{ij} D_j w_\ell) = f_\ell + \sum_{j=1}^{2} D_j f_j^\ell,$$

with $|f_\ell| \le C(1 + \delta|D^2 w|)$ and $|f_j^\ell| \le C|\mathbf{H}|$ in a neighbourhood of x_0, where C does not depend on δ. By using the weak form of the equation for $w_\ell - a_\ell$, where a_ℓ is the mean of w_ℓ over $B_\sigma(\xi)$, one then very easily checks that, for suitable $\rho > 0$, w satisfies an inequality of the form

$$\int_{L_0 \cap B_{\sigma/2}(\xi)} |D^2 w|^2 \le C \int_{B_\sigma(\xi) \setminus B_{\sigma/2}(\xi)} |D^2 w|^2 + C\sigma^\gamma,$$

for each $\xi \in L_0 \cap B_\rho(x_0)$ and each $\sigma \in (0, \rho)$, where C depends on ρ but not on σ. By hole-filling (that is, by adding $C \int_{B_{\sigma/2}(\xi)} |D^2 w|^2$ to each side of the inequality and iterating the consequent inequality), we then have for suitable $\alpha > 0$

$$\int_{L_0 \cap B_\sigma(\xi)} |D^2 w|^2 \le C\sigma^{2\alpha}, \quad 0 < \sigma < \rho \tag{3.9}$$

for each $\xi \in L_0 \cap B_\rho(x_0)$, where C does not depend on σ, thus by virtue of Morrey's lemma completing the proof that Σ is a $C^{1,\alpha} \cap W^{2,2}$ surface away from the bad points $\xi_1, \dots, \xi_P$. □

We now show that w is actually $C^{2,\alpha}$ for some $\alpha > 0$. (Higher regularity, and real-analyticity, of w is standard—see e.g. [MCB]—once we get as far as $C^{2,\alpha}$.) To establish $C^{2,\alpha}$ regularity of u we need the following lemma:

Lemma 3.2. *Let β, γ, $L > 0$, $\mathbb{D} = \{x \in \mathbb{R}^2 \colon |x| < 1\}$, and let*

$$u = (u^1, \dots, u^m) \in W^{2,2}(\mathbb{D}; \mathbb{R}^m) \cap C^{1,\gamma}(\mathbb{D}; \mathbb{R}^m)$$

satisfy $|u| + |Du| \le 1$ in $\mathbb{D}$ and

$$\int_{\mathbb{D} \cap \{x \colon |x - \xi| < \rho\}} |D^2 u|^2 \le \beta \rho^{2\gamma} \tag{i}$$

for each $\xi \in \mathbb{D}$ and $\rho < 1$. Suppose further that u is a weak solution of the 4^{th}-order quasilinear system

$$D_j D_s \left(A^{ijrs}_{\alpha\beta}(x, u, Du) D_i D_r u^\beta\right) + D_j B^j_\alpha(x, u, Du, D^2 u) + B^0_\alpha(x, u, Du, D^2 u) = 0,$$

where $A^{ijrs}_{\alpha\beta} = A^{ijrs}_{\alpha\beta}(x,z,p)$ *and* $B^j_\alpha = B^j_\alpha(x,z,p,q)$ *satisfy the following for* $|z|+|p|\le 1$

(ii) $$\sum_{i,j,r,s,\alpha,\beta} A^{ijrs}_{\alpha\beta}\xi^\alpha_{ij}\xi^\beta_{rs} \ge L^{-1}\sum_{i,j,\alpha}|\xi^\alpha_{ij}|^2,$$

(iii) $$|A^{ijrs}_{\alpha\beta}(x,z,p)|\le L, \quad |D_{(x,z,p)}A^{ijrs}_{\alpha\beta}(x,z,p)|\le L,$$

(iv) $$|B^j_\alpha(x,z,p,q)| + |D_{(x,z,p)}B^j_\alpha(x,z,p,q)| \le L(1+|q|^2),$$
$$|D_q B^j_\alpha(x,z,p,q)| \le L(1+|q|),$$

where we use the notation that $D_P F$ *means the tensor of all first order partial derivatives with respect to the variables* P. *Then* $u\in W^{3,2}_{loc}(\mathbb{D})\cap C^{2,\alpha}$ *for some* $\alpha>0$; *in fact there are* $C>0$, $\alpha\in(0,1)$, *depending only on* β,γ,n,L *such that*

$$\int_{\{x:\ |x-\xi|<\rho\}} |D^3u|^2 \le C\rho^{2\alpha}$$

for each $\xi\in\mathbb{D}$ *with* $\operatorname{dist}(\xi,\partial\mathbb{D})>2\rho$.

Proof. The weak form of the equation is

(3.10) $$\int A^{ijrs}_{\alpha\beta}(x,u,Du)D_{ir}u^\beta D_{js}\zeta^\alpha - B^j_\alpha(x,u,Du,D^2u)D_j\zeta^\alpha + B^0_\alpha(x,u,Du,D^2u)\zeta^\alpha = 0,$$

valid for any $\zeta^\alpha\in W^{2,2}_0(\mathbb{D})$, where, here and subsequently, repeated Latin indices are summed from 1 to 2 and repeated Greek indices are summed from 1 to m.

We are going to use the difference quotient operators

(3.11)
$$\delta_h f(x) = h^{-1}(f(x+he)-f(x)), \quad \bar\delta_h f(x) = h^{-1}(f(x)-f(x-he)), \quad h\neq 0,$$

on $\mathbb{D}_{|h|}\equiv\{x\in\mathbb{D}:\ \operatorname{dist}(x,\partial\mathbb{D})>|h|\}$, where $e=(1,0)$ or $(0,1)$. Concerning these, recall the formulae

(3.12) $$\begin{cases} \delta_h(fg)(x) = (\delta_h f(x))g(x+he) + f(x)\delta_h g(x) \\ \int_{\mathbb{D}} g\delta_h f = -\int_{\mathbb{D}} f\bar\delta_h g, \end{cases}$$

the first being valid on $\mathbb{D}_{|h|}$ and the second requiring that the product fg vanishes outside a compact subset of $\mathbb{D}_{|h|}$.

Now take a disk $B_\rho(\xi)$ with $|\xi| < 1-2\rho$, take $0 < |h| < \rho/4$, and replace ζ^α in (3.10) by $\bar{\delta}_h((\delta_h u^\alpha - \ell_h^\alpha)\zeta)$, where $\zeta \in C_c^\infty(\mathbb{R}^2)$ is arbitrary with support in $B_{3\rho/4}(\xi)$, and where $\ell_h^\alpha = a_h + b_h \cdot (x-\xi)$, with a_h the mean value of $\delta_h u^\alpha$ in the annulus $A \equiv B_\rho(\xi)\backslash B_{\rho/2}(\xi)$ and b_h the mean value of Du_h on this annulus. Notice that then we have

(3.13)
$$\int_A |u_h^\alpha - \ell_h^\alpha|^2 \le C\rho^2 \int_A |Du_h^\alpha|^2 \text{ and also } \int_A |u_h^\alpha - \ell_h^\alpha|^2 \le C\rho^4 \int_A |D^2 u_h^\alpha|^2$$

by the Poincaré inequality, where we use the abbreviation

$$u_h^\alpha = \delta_h u^\alpha.$$

Also, in view of the given L^2 bounds on D^2u and the fact that $u_h(x) = \int_0^1 D_1 u(x+she)\,ds$, one readily checks the following inequalities for $|h| < 1$, $0 < \sigma < \rho < 1$ and $|y| < 1$:

(3.14)
$$\int_{\mathbb{D}_{|h|}\cap B_\sigma(y)} |D\ell_h|^2 \le C\sigma^2 \sup_{\mathbb{D}} |D\ell_h|^2 \le C\sigma^2\rho^{2\gamma-2} \equiv C(\sigma/\rho)^{2-2\gamma}\sigma^{2\gamma} \le C\sigma^{2\gamma}$$
$$\int_{\mathbb{D}_{|h|}\cap B_\sigma(y)} |Du_h|^2 \le \int_0^1 \int_{\mathbb{D}\cap B_\sigma(y+she)} |D^2u|^2\,dxds \le C\sigma^{2\gamma},$$

where $\mathbb{D}_{|h|} = \{x\colon |x| < 1-|h|\}$. Now using (3.12) and the above choice of ζ^α in (3.10) we obtain

$$\int_{B_\rho(\xi)} \Big(a_{\alpha\beta}^{ijrs} D_{ir} u_h^\beta + \delta_h(A^{ijrs}(x,u,Du))D_{ir}u^\beta)D_{js}((u_h^\alpha - \ell_h^\alpha)\zeta)$$
$$-\delta_h(B_\alpha^j(x,u,Du,D^2u))D_j((u_h^\alpha - \ell_h^\alpha)\zeta) + \delta_h(B_\alpha^0(x,u,Du,D^2u)\Big)(u_h^\alpha - \ell_h^\alpha)\zeta = 0,$$

where $a_{\alpha\beta}^{ijrs}(x) = A^{ijrs}(x+he, u(x+he), Du(x+he))$. Using the given conditions (iii), (iv) we can then check that this is an identity of the form

$$(3.15)\quad \int_{B_\rho(\xi)} (a^{ijrs} D_{ir} u_h^\beta D_{js} u_h^\alpha + E_\alpha \cdot D^2 u_h^\alpha + F_\alpha)\zeta$$
$$+ \int_{B_\rho(\xi)} (E_\alpha^j \cdot D^2 u_h^\alpha + F_\alpha^j) D_j \zeta^\alpha + (E_\alpha^{jk} \cdot D^2 u_h^\alpha + F_\alpha^{jk}) D_{jk}\zeta^\alpha = 0,$$

where

$$
\begin{aligned}
&|E_\alpha| \le C(1+|Du_h+|D(u_h-\ell_h)|+|u_h-\ell_h|)(1+|D^2u|)\\
&|F_\alpha| \le C(1+|Du_h|^2+|D(u_h-\ell_h)|^2+|u_h-\ell_h|^2)(1+|D^2u|^2)\\
&|E_\alpha^j| \le C(|D(u_h-\ell_h)|+(1+|D^2u|)|u_h-\ell_h|)\\
&|F_\alpha^j| \le C((1+|Du_h|^2+|D(u_h-\ell_h)|^2)|D^2u|\\
&\qquad\qquad +(1+|D^2u|^2)|u_h-\ell_h|(1+|Du_h|))\\
&|E_\alpha^{jk}| \le C|u_h-\ell_h|\\
&|F_\alpha^{jk}| \le C|u_h-\ell_h|(1+|Du_h|),
\end{aligned} \tag{3.16}
$$

where C depends only on L, n.

After replacing ζ by ζ^4, using the ellipticity condition (ii), the Cauchy-Schwarz inequality, and the inequalities (3.16), and keeping in mind that $\sup|u_h|, \sup|\ell_h| \le C$ (the latter being true by (3.14)), we obtain

(3.17)

$$
\begin{aligned}
&\int_{B_\rho(\xi)} |D^2u_h|^2\zeta^4\\
&\le C\int_{B_\rho(\xi)} \zeta^4(1+|Du_h|^2+|D(u_h-\ell_h)|^2)(1+|D^2u|^2)\\
&\quad + C\int_{B_\rho(\xi)} |D\zeta|^2|\zeta|^2(1+|Du_h|^2+|D(u_h-\ell_h)|^2+|D^2u|^2)|u_h-\ell_h|^2\\
&\quad + C\int_{B_\rho(\xi)} (\zeta^2|D^2\zeta|^2+|D\zeta|^4)|u_h-\ell_h|^2.
\end{aligned}
$$

Now we use a result of Morrey [MCB, Lemma 5.4.2], which says that if $q \ge 0$ on $B_1(0) \subset \mathbb{R}^2$ and if there are constants β, $\gamma > 0$ with $\int_{B_\sigma(\xi)\cap B_1(0)} q \le \beta\sigma^\gamma$ for all $\xi \in B_1(0)$ and $\sigma \in (0,1)$, then

$$\int_{B_1(0)} q|v|^2 \le \epsilon\int_{B_1(0)} |Dv|^2 + C\int_{B_1(0)} |v|^2$$

for each $v \in W_0^{1,2}(B_1(0))$ and for each $\epsilon > 0$, where C depends only on the constants β, γ, ϵ. By the scaling $x \to z = \rho^{-1}x$ we see that this implies that if $q \ge 0$ on a ball $B_\rho(\xi)$ in $\mathbb{R}^2$ and if $\int_{B_\sigma(y)\cap B_\rho(\xi)} q \le \beta(\sigma/\rho)^\gamma$ for all $y \in B_\rho(\xi)$ and $\sigma \le \rho$, then

$$\int_{B_\rho(\xi)} q|v|^2 \le \epsilon\int_{B_\rho(\xi)} |Dv|^2 + C\rho^{-2}\int_{B_\rho(\xi)} |v|^2,$$

for any $v \in W_0^{1,2}(B_\rho(\xi))$ and any $\epsilon > 0$, where C depends only on β, γ, ϵ. That is, stated in another way, if $q \geq 0$ on $B_\rho(\xi)$ and $\int_{B_\rho(\xi)\cap B_\sigma(y)} q \leq \beta\sigma^\gamma$ for all $\sigma \in (0, \rho]$ and all $y \in B_\rho(\xi)$, then for each $\epsilon > 0$

(3.18)
$$\int_{B_\rho(\xi)} q|v|^2 \leq \epsilon\rho^\gamma \int_{B_\rho(\xi)} |Dv|^2 + C\rho^{\gamma-2} \int_{B_\rho(\xi)} |v|^2, \quad v \in W_0^{1,2}(B_\rho(\xi)),$$

where C depends only on β, γ, ϵ. (Because the previous version can be applied with $\rho^{-\gamma}q$ in place of q.)

So let $B_\rho(\xi)$ continue to be such that $|\xi| < 1 - \rho$ and take $0 < |h| < \rho/4$, $\rho < 1/4$ in the above. We use (3.18) to estimate some of the terms on the right in (3.17); in fact, according to (i) we can use (3.18) with $q = (1 + |D^2u|^2)$, so in particular, assuming $\zeta \in C_c^\infty(B_\rho(\xi))$,

$$\begin{aligned}\int_{B_\rho(\xi)} &\zeta^4(1 + |Du_h|^2 + |D(u_h - \ell_h)|^2)(1 + |D^2u|^2) \\ &\leq \epsilon\rho^\gamma \int_{B_\rho(\xi)} \Big(|D(\zeta^2(1 + |Du_h|^2)^{1/2})|^2 + |D(\zeta^2(1 + |D(u_h - \ell_h)|^2)^{1/2})|^2\Big) \\ &\quad + C\rho^{\gamma-2} \int_{B_\rho(\xi)} \zeta^4(1 + |Du_h|^2 + |D(u_h - \ell_h)|^2),\end{aligned}$$

which evidently gives

$$\begin{aligned}(3.19) \quad \int_{B_\rho(\xi)} &\zeta^4(1 + |Du_h|^2 + |D(u_h - \ell_h)|^2)(1 + |D^2u|^2) \\ &\leq 16\epsilon\rho^\gamma \int_{B_\rho(\xi)} \zeta^4|D^2u_h|^2 + \\ &\quad C\rho^\gamma \int_{B_\rho(\xi)} (\rho^{-2}\zeta^4 + \zeta^2|D\zeta|^2)(1 + |Du_h|^2 + |D(u_h - \ell_h)|^2).\end{aligned}$$

Also, using (3.18) with the same choice of q, we have

$$\begin{aligned}\int_{B_\rho(\xi)} &\zeta^2|D\zeta|^2(1 + |D^2u|^2)|u_h - \ell_h|^2) \\ &\leq \epsilon\rho^\gamma \int_{B_\rho(\xi)} |D(\zeta D\zeta \otimes (u_h - \ell_h))|^2 + C\rho^{\gamma-2} \int_{B_\rho(\xi)} \zeta^2|D\zeta|^2|u_h - \ell_h|^2,\end{aligned}$$

which evidently gives

$$(3.20) \quad \int_{B_\rho(\xi)} \zeta^2|D\zeta|^2(1 + |D^2u|^2)|u_h - \ell_h|^2)$$

$$\leq 16\epsilon\rho^\gamma \int_{B_\rho(\xi)} \zeta^2 |D\zeta|^2 |Du_h|^2 +$$
$$C\rho^\gamma \int_{B_\rho(\xi)} (|D\zeta|^4 + \zeta^2 |D^2\zeta|^2 + \rho^{-2}\zeta^2 |D\zeta|^2)|u_h - \ell_h|^2.$$

Notice that by virtue of (3.14), assuming (as we subsequently do) that support of ζ is contained in $B_{3\rho/4}(\xi)$ and that $|h| < \rho/4$, we can also use the choice $q = 1 + |Du_h|^2 + |D(u_h - \ell_h)|^2$ (rather than $q = 1 + |D^2 u|^2$) in the above, giving in place of (3.20) the inequality

$$\tag{3.20$'$} \begin{aligned} \int_{B_\rho(\xi)} \zeta^2 |D\zeta|^2 (1 + |Du_h|^2 + |D(u_h - \ell_h)|^2)|u_h - \ell_h|^2 \\ \leq 16\epsilon\rho^\gamma \int_{B_\rho(\xi)} \zeta^2 |D\zeta|^2 |Du_h|^2 \\ + C\rho^\gamma \int_{B_\rho(\xi)} (|D\zeta|^4 + \zeta^2 |D^2\zeta|^2 + \rho^{-2}\zeta^2 |D\zeta|^2)|u_h - \ell_h|^2. \end{aligned}$$

Using (3.19), (3.20) and (3.20$'$) in (3.17) we then conclude that

$$\begin{aligned} \int_{B_\rho(\xi)} |D^2 u_h|^2 \zeta^4 \leq 32\epsilon \int_{B_\rho(\xi)} \zeta^4 |D^2 u_h|^2 \\ + C \int_{B_\rho(\xi)} (|D\zeta|^4 + \zeta^2 |D^2\zeta|^2 + \rho^{-2}\zeta^2 |D\zeta|^2)|u_h - \ell_h|^2 \\ + C\rho^\gamma \int_{B_\rho(\xi)} ((\zeta^2 |D\zeta|^2 + \rho^{-2}\zeta^4)(1 + |Du_h|^2 + |D(u_h - \ell_h)|^2). \end{aligned}$$

Now, with $\zeta \equiv 1$ on $B_{\rho/2}(\xi)$, $|D\zeta| \leq C\rho^{-1}$, and $|D^2\zeta| \leq C\rho^{-2}$ (together with the previous restriction that $\zeta \equiv 0$ outside $B_{3\rho/4}(\xi)$), we conclude

$$\tag{3.21} \begin{aligned} \int_{B_\rho(\xi)} |D^2 u_h|^2 \zeta^4 \leq C\rho^{-4} \int_A |u_h - \ell_h|^2 \\ + C\rho^{\gamma-2} \int_{B_\rho(\xi)} (1 + |D^2 u|^2 + |Du_h|^2 + |D(u_h - \ell_h)|^2), \end{aligned}$$

where $A = B_\rho(\xi) \backslash B_{\rho/2}(\xi)$. Using the Poincaré inequality (3.13) and also (3.14) we thus have in particular that

$$\int_{B_{\rho/2}(\xi)} |D^2 u_h|^2 \leq C \int_{B_\rho(\xi)} (1 + |D^2 u|^2),$$

with C depending only on γ, n, and β (and not depending on ρ). Since this holds for $0 < |h| < \rho/4$, we can let $|h| \downarrow 0$ to deduce that $u \in W^{3,2}(B_{\rho/2}(\xi))$

and that

$$\int_{B_{\rho/2}(\xi)} |D^3u|^2 \leq C\rho^{-2} \int_{B_\rho(\xi)} (1+|D^2u|^2). \tag{3.22}$$

Also, (3.21), (3.13) and (3.14) give (after letting $|h| \downarrow 0$)

$$\int_{B_\rho(\xi)} |D^3u|^2\zeta^4 \leq C\int_A |D^3u|^2 + C\rho^{\gamma-2}\int_{B_\rho(\xi)} (1+|D^2u|^2), \tag{3.23}$$

with C independent of ρ. Since we have shown that $u \in W^{3,2}_{loc}(\mathbb{D})$, the Sobolev embedding theorem then implies $D^2u \in L^p_{loc}$ for each $p \geq 1$, and hence in particular

$$\int_{B_\rho(\xi)} (1+|D^2u|^2) \leq C_\delta \rho^{2-\delta}$$

for each $\delta > 0$, and hence (3.23) implies

$$\int_{B_{\rho/2}(\xi)} |D^3u|^2 \leq C\int_A |D^3u|^2 + C\rho^{\gamma/2}, \tag{3.24}$$

with C independent of ρ, for $|\xi| < 1-\rho$ and $\rho < 1/4$. Adding $C\int_{B_{\rho/2}(\xi)} |D^3u|^2$ to each side of the inequality (i.e., hole filling again), we thus get

$$\int_{B_{\rho/2}(\xi)} |D^3u|^2 \leq \theta\int_{B_\rho(\xi)} |D^3u|^2 + C\rho^{\gamma/2}, \tag{3.25}$$

for a fixed constant $\theta \in (0,1)$ independent of ρ, and this is valid for any $|\xi| < 1-\rho$ and $\rho < 1/4$. By iteration we thus have $\alpha > 0$ such that

$$\int_{B_{\rho/2}(\xi)} |D^3u|^2 \leq C\rho^{2\alpha}, \quad |\xi| < 1-\rho,\ \rho < 1/4,$$

so by Morrey's lemma u is $C^{2,\alpha}$ locally on $\mathbb{D}$. This completes the proof of Lemma 3.2. □

We can now show that w (as in the discussion preceeding Lemma 3.2) is actually $C^{2,\alpha}$ for some $\alpha > 0$. Recall that we already proved that w is of class $C^{1,\alpha} \cap W^{2,2}$ on the disc $L_0 \cap B_\rho(\xi)$ and that

$$\int_{L_0\cap B_\sigma(y)} |D^2w|^2 \leq C\sigma^\gamma \tag{3.26}$$

for some fixed $\gamma > 0$ and for every $y \in L_0 \cap B_\rho(\xi)$ and $\sigma < \rho$ (assuming ρ is small enough), so in view of Lemma 3.2 we will be done if we can show that (modulo a rescaling and rigid motion taking the disc $L_0 \cap B_\rho(\xi)$ to $\mathbb{D}$),

w is a weak solution of an equation of the form considered in Lemma 3.2. By construction w is a weak ($W^{2,2}$) solution of the Euler-Lagrange system for the functional

$$\mathcal{N}(w) = \int_{B_\rho(\xi)} \sum_{i,j,r,s=1}^{2} \sum_{\alpha,\beta=1}^{n-2} (\delta_{\alpha\beta} - g^{\alpha\beta}) g^{ij} g^{rs} w^\alpha_{ir} w^\beta_{js} \sqrt{g},$$

where $(g^{ij}) = (g_{ij})^{-1}$, $g_{ij} = \delta_{ij} + D_i w \cdot D_j w$, $g^{\alpha\beta} = \sum_{p,q=1}^{2} D_p w^\alpha D_q w^\beta g^{pq}$. (Notice that $\mathcal{N}(w)$ is just $\int_{\operatorname{graph} w} |A_w|^2$, where A_w is the second fundamental form of graph w, and by virtue of the Gauss-Bonnet theorem we therefore have that $\frac{1}{4}\mathcal{N}(w)$ differs from the Willmore functional of graph w by only a boundary integral; this explains why w must be a stationary point for $\mathcal{N}(w)$ (relative to variations of w which vanish in a neighbourhood of $L_0 \cap \partial B_\rho(\xi)$).) Now one checks by direct computation that the Euler-Lagrange system for the functional $\mathcal{N}(w)$ has (after a re-scaling and rigid motion taking the disc $L_0 \cap B_\rho(\xi)$ to the unit disc $\mathbb{D}$) exactly the form of the system considered in Lemma 3.2. In fact w satisfies a system as in Lemma 3.2, with $m = n-2$, $A^{ijrs}_{\alpha\beta} = \sqrt{g}(\delta_{\alpha\beta} - g^{\alpha\beta}) g^{ir} g^{js}$ and with $B^j_\alpha(x, z, p, q)$ equal to homogeneous quadratic polynomials in q (with coefficients smooth functions of x, z, p). Since w has small C^1 norm, the hypotheses (ii)–(iv) of Lemma 3.2 are easily checked, and the hypothesis (i) is satisfied by virtue of (3.26)above. We thus deduce that $w \in C^{2,\alpha}$ as required.

Completion of the proof of Theorem 3.1. Thus we have established the real analyticity of $\Sigma = \operatorname{spt} \mu$ away from the finitely many bad points $\xi_1, \ldots, \xi_P$. Next we establish the Hausdorff distance sense convergence claimed in the statement of Theorem 3.1. Indeed suppose first that there is a sequence y_j of points with $y_j \in \Sigma_j$ for each j, with $y_j \to y$ and with $\operatorname{dist}(y, \Sigma) \equiv \eta > 0$. Since Σ_j is connected, there is j_0 such that $\Sigma_j \cap \partial B_\sigma(y) \neq \emptyset$ for all $j \geq j_0$ and all $\sigma \in (\eta/4, \eta/2)$. Thus for each $N \geq 1$ and each $j \geq j_0$ we can find points $z_{j,k} \in \Sigma_j \cap \partial B_{(1+k/N)\eta/4}$. Applying (1.2) with Σ_j in place of Σ, with $z_{j,k}$ in place of 0, and with $\rho = 1/N$ we obtain

$$\pi \leq C(\rho^{-2}|\Sigma_j \cap B_\rho(z_{j,k})| + \mathcal{F}(\Sigma_j \cap B_\rho(z_{j,k}))),$$

so summing over k and using the disjointness of the $B_\rho(z_{j,k})$ we obtain, by virtue of the fact that $|\Sigma_j \cap B_\rho(z_{j,k})| \to 0$ for each $k = 1, \dots, N$,

$$N \le C \limsup_j \mathcal{F}(\Sigma_j).$$

Since N is arbitrary this says $\limsup_{j\to\infty} \mathcal{F}(\Sigma_j) = \infty$, contrary to the boundedness of $\mathcal{F}(\Sigma_j)$. Thus we have shown that the set of all possible limit points of all possible sequences $\{y_j\}$ with $y_j \in \Sigma_j$ is contained in Σ. Since the reverse inclusion is trivially satisfied, this completes the proof of the Hausdorff distance sense convergence.

Next we want to discuss $C^{1,\alpha} \cap W^{2,2}$ regularity of Σ near the bad points.

First note that (by the measure theoretic convergence of Σ_j to Σ and because $\int_{\Sigma_j} |\mathbf{H}_j|$ is bounded) a subsequence $\Sigma_{j'}$ of the Σ_j converges to Σ in the varifold sense (this is a special case of Allard's compactness theorem—see [SL2]), and hence we have the convergence of first variation. Thus, for any fixed smooth $\Phi\colon \mathbb{R}^n \to \mathbb{R}^n$ with compact support, we have

$$\lim \int_{\Sigma_{j'}} \Phi \cdot \mathbf{H}_{j'} = \int_\Sigma \Phi \cdot \mathbf{H}. \tag{3.27}$$

Now extend $\mathbf{H}$ (which is smooth on $\Sigma \backslash \{\xi_1, \dots, \xi_P\}$) to all of $\mathbb{R}^n \backslash \{\xi_1, \dots, \xi_P\}$ smoothly, and apply (3.27) with $\Phi = \zeta \mathbf{H}$, where ζ is C^∞ with compact support in $\mathbb{R}^n \backslash \{\xi_1, \dots, \xi_P\}$. After an application of the Schwarz inequality and the measure theoretic convergence of Σ_j to Σ, this gives

$$\int_{\Sigma \cap U(\xi) \backslash (\bigcup_{i=1}^P B_\rho(\xi_i))} |\mathbf{H}|^2 \le \liminf_j \int_{\Sigma_j \cap U(\xi) \backslash (\bigcup_{i=1}^P B_\rho(\xi_i))} |\mathbf{H}_j|^2 \tag{3.28}$$

for each open $U \subset \mathbb{R}^n$ and each $\rho < \frac{1}{2} \min_{i \ne j} |\xi_i - \xi_j|$. In particular

$$\int_{\Sigma \cap U} |\mathbf{H}|^2 \le \lim_{\rho \downarrow 0} \liminf_j \int_{\Sigma_j \cap U \backslash (\bigcup_{i=1}^P B_\rho(\xi_i))} |\mathbf{H}_j|^2 \quad \forall\, \rho > 0. \tag{3.29}$$

By almost the same argument, except that the weak definition of second fundamental form and the corresponding compactness theorem ([HJ]) is used in place of the first variation identity and the Allard compactness theorem, it is easily checked that

$$\int_{\Sigma \cap U} |\mathbf{A}|^2 \le \lim_{\rho \downarrow 0} \liminf_j \int_{\Sigma_j \cap U \backslash (\bigcup_{i=1}^P B_\rho(\xi_i))} |\mathbf{A}_j|^2 \quad \forall\, \rho > 0, \tag{3.30}$$

where $\mathbf{A}$, $\mathbf{A}_j$ denote the second fundamental form of Σ and Σ_j respectively.

The inequality (3.29) guarantees in particular that $\int_\Sigma |\mathbf{H}|^2 < \infty$ and the same application of the Allard compactness theorem guarantees that the first variation identity 1.1 holds for Σ. (We emphasize that 1.1 holds exactly as stated; it is not necessary that Φ vanish near the bad points.) Then we get in particular that (1.2)–(1.5) hold for all $y \in \Sigma \backslash \{\xi_1, \dots, \xi_P\}$; in fact if we pick $\xi_{i,k} \to \xi_i$ and apply this with $\xi_{i,k}$ in place of y we obtain (1.2)–(1.5) for *all* $y \in \Sigma$.

In particular Lemma 1.4 applies (without change in the proof) to Σ. As a matter of fact one easily checks that Lemma 1.4 applies to Σ even if the subsets Σ_j satisfy $\partial\Sigma_j \subset \partial B_\rho(0) \cup \{\xi_1, \dots, \xi_P\}$ rather then $\partial\Sigma_j \subset \partial B_\rho(0)$. We shall make use of this shortly.

Now let $\xi_{i,k} \in \Sigma$ with $\xi_{i,k} \to \xi_i$ and $\xi_{i,k} \neq \xi_i$ for each k. By applying (1.4) to Σ, we get

$$\pi \leq C(\rho^{-2}|\Sigma \cap B_\rho(\xi_{i,k})| + \mathcal{F}(\Sigma \cap B_\rho(\xi_{i,k}))).$$

Hence

$$\pi \leq C(\rho^{-2}|\Sigma \cap B_\rho(\xi_i)| + \mathcal{F}(\Sigma \cap B_\rho(\xi_i))). \tag{3.31}$$

Also, according to (1.2) we have

$$\int_{\Sigma \cap B_\sigma(\xi_i)} \frac{|(x-\xi_i)^\perp|^2}{|x-\xi_i|^4} < \infty,$$

so that in particular (since $|\Sigma \cap B_\sigma(\xi_i)| \geq C\sigma^2$ for sufficiently small σ by (3.31)) we have that for each $\epsilon > 0$ there is $\sigma_0 = \sigma_0(\epsilon) > 0$ such that for all $\sigma \in (0, \sigma_0)$

$$\frac{|(x-\xi_i)^\perp|}{|x-\xi_i|} \leq \epsilon \text{ on } B_\sigma(\xi_i) \backslash B_{\sigma/2}(\xi_i) \tag{3.32}$$

except for a set of measure $\leq C\epsilon\sigma^2$.

From now on we assume that $\sigma \in (0, \sigma_0(\epsilon))$ and that σ is also small enough to ensure that

$$\int_{\Sigma \cap B_\sigma(\xi_i)} |\mathbf{A}|^2 \leq \epsilon^2/4, \quad i = 1, \dots, P, \tag{3.32$'$}$$

and we also assume that $\epsilon \in (0, C^{-1})$, where C is as in (1.4) (the same as C in (3.31) above). Next take $y_i \in \partial B_{3\sigma/4}(\xi_i) \cap \Sigma$. (Notice that $\partial B_{3\sigma/4}(\xi_i) \cap \Sigma \neq \emptyset$ for σ sufficiently small, because otherwise we could apply (3.31) with $\rho \uparrow \infty$ to the component Σ_* of Σ which contains ξ_i, thus giving $\mathcal{F}(\Sigma_*) \geq C^{-1}$, contrary to (3.31).) By (3.32) we can apply the approximate graphical decomposition lemma (Lemma 2.1) to give that there is a plane L_i containing y_i and a C^1 function $u_i\colon \Omega_i \subset L_i \to L_i^\perp$, with

$$\sigma^{-1}|u_i| + |Du_i| \leq C\epsilon^{1/2(2n-3)} \tag{3.33}$$

where $\Omega_i \supset B_{\sigma/4}(y_i)\backslash \bigcup_k(d_{i,k})$, where each $d_{i,k}$ is a closed disc in L_i and $\sum_k \operatorname{diam} d_{i,k} \leq C\epsilon^{1/2}\sigma$, and where

$$B_{\theta\sigma}(y_i) \cap \Sigma = (B_{\theta\sigma}(y_i) \cap \operatorname{graph} u_i) \cup \left(\bigcup_k P_{i,k}\right), \tag{3.34}$$

where the $P_{i,k}$ are pairwise disjoint and each is diffeomorphic to the unit disc in $\mathbb{R}^2$, and

$$\sum_k \operatorname{diam} P_{i,k} \leq C\epsilon^{1/2}. \tag{3.35}$$

Notice that by (3.32) we have that the radial vector $|y_i-\xi_i|^{-1}(y_i-\xi_i)$ is almost tangent to L_i in the sense that

$$\operatorname{dist}(\xi_i, L_i) \leq C\epsilon^{1/2(2n-3)}\sigma.$$

(Recall that $y_i \in L_i$ by definition.) By virtue of (3.33), (3.34), (3.35) we can find points $y_{1,i}$, $y_{2,i} \in \Sigma$ with $\operatorname{dist}(y_{j,i}, p_{j,i}) < C\epsilon^{1/2(2n-3)}\sigma$, where $p_{1,i}, p_{2,i}$ denote the two points of $\partial B_{3\sigma/4}(\xi_i) \cap \partial B_{\theta\sigma}(y_i) \cap L_i$, and we can make a similar application of Lemma 2.1 starting with $y_{j,i}$ in place of y_i, for $j = 1, 2$. The corresponding planes $L_{j,i}$ must in this case be close to L_i in the sense that $\|L_i - L_{j,i}\| \leq C\epsilon^{1/2(2n-3)}$. We now repeat this procedure with $y_{i,j}, L_{j,i}$ in place of y_i, L_i; after a fixed number of steps, depending only on n, we then have that there is an annular region $A_i \equiv \{x \in L_i\colon (\frac{3}{4} - \frac{\theta}{2})\sigma < |x - \xi_i| < (\frac{3}{4} + \frac{\theta}{2})\sigma\} \subset L_i$ and a C^1 function $u_i\colon A_i\backslash(\bigcup_k e_{i,k}) \to L_i^\perp$ with

$$\begin{gathered} \sigma^{-1}|u_i| + |Du_i| \leq C\epsilon^{1/2(2n-3)} \\ \Sigma \cap \mathcal{A}_i = \left(\operatorname{graph} u_i \cup \left(\bigcup_k P_{i,k}\right)\right) \cap \mathcal{A}_i \\ \sum_k \operatorname{diam} P_{i,k} \leq C\epsilon^{1/2(2n-3)}\sigma, \quad P_{L_i}(P_{i,k}) = e_{i,k}, \end{gathered} \tag{3.36}$$

where P_{L_i} is the orthogonal projection onto L_i, and where

$$\mathcal{A}_i = \{x + z \colon x \in A_i, z \in L_i^\perp, |z| \le \theta\sigma/2\}.$$

Notice that the latter part of the above argument can be applied to Σ_j for j sufficiently large, assuming that $\int_{\Sigma_j \cap B_\sigma(\xi_i)\setminus B_{\sigma/4}(\xi_i)} |\mathbf{A}_j|^2 < \epsilon^2$ for all sufficiently large j; notice that in this case we cannot use (3.32)because (3.32) relied on the fact that $\mathcal{F}(\Sigma \cap B_\sigma(\xi_i))$ is small, which may not be true for Σ_j. However, in place of (3.32) we can use the Hausdorff distance sense convergence of Σ_j to Σ, which guarantees that Σ_j is in the ϵ_j neighbourhood of Σ, with $\epsilon_j \downarrow 0$; this means that for Σ_j we can take the reference plane and annular regions to be the same (i.e. L_i and A_i respectively) that we used for Σ. Thus, assuming $\int_{\Sigma_j \cap B_\sigma(\xi_i)\setminus B_{\sigma/4}(\xi_i)} |\mathbf{A}_j|^2 < \epsilon^2$, we have C^1 functions $u_{i,j} \colon A_i \setminus (\cup_k e_{i,j,k}) \to L_i^\perp$ with

$$\begin{gathered}
\sigma^{-1}|u_{i,j}| + |Du_{i,j}| \le C\epsilon^{1/2(2n-3)} \\
\Sigma_j \cap \mathcal{A}_i = \Big(\operatorname{graph} u_{i,j} \cup \Big(\bigcup_k P_{i.j,k}\Big)\Big) \cap \mathcal{A}_i \qquad (3.37)\\
\sum_k \operatorname{diam} P_{i,j,k} \le C\epsilon^{1/2(2n-3)}\sigma, \quad P_{L_i}(P_{i.j.k}) = e_{i,j,k},
\end{gathered}$$

where $\mathcal{A}_i$ is as in (3.36).

Now we claim that in fact, in place of the identity in the second line of (3.37), we have the stronger identity

(3.38)
$$\Sigma_j \cap B_{(\frac{3}{4}+\frac{\theta}{4})\sigma}(\xi_i)\setminus B_{(\frac{3}{4}-\frac{\theta}{4})\sigma}(\xi_i) = \Sigma_j \cap \mathcal{A}_i \cap B_{(\frac{3}{4}+\frac{\theta}{4})\sigma}(\xi_i)\setminus B_{(\frac{3}{4}-\frac{\theta}{4})\sigma}(\xi_i).$$

Indeed otherwise, since the discussion above (and in particular (3.36)) applies equally well with any $\widetilde{\sigma} \le \sigma$, we would have that there are two components $\Sigma^{(1)}, \Sigma^{(2)}$ of $\Sigma \cap B_{3\sigma/4}(\xi_i)\setminus\{\xi_i\}$ both containing ξ_i in their closures. This would contradict the modified version of Lemma 1.4 which applies to Σ as discussed above.

Now since $\xi_1, \dots, \xi_P$ are the only bad points of Σ, we know that for any fixed $\sigma > 0$ there is $\delta \in (0, \sigma/2)$ such that

$$\liminf_{j\to\infty} \int_{B_\delta(y)\cap\Sigma_j} |\mathbf{A}_j|^2 \le \epsilon^2 \qquad (3.39)$$

for each $y \in \Sigma \backslash (\bigcup_{i=1}^{P} B_\sigma(\xi_i))$. We assume subsequently that δ is also chosen small enough so that

$$\text{(3.40)} \quad \nabla_\Sigma |x-y| \neq 0 \text{ on } B_\sigma(y)\backslash\{y\}, \qquad \sup_{x\in\Sigma\cap B_\sigma(y)} \operatorname{dist}(x-y, T_y\Sigma) < \epsilon$$

for all $y \in \Sigma \backslash (\bigcup_{i=1}^{P} B_\sigma(\xi_i))$. Now by (3.39) we can apply Lemma 2.1 to give a plane $L_j(y)$ containing y and a smooth function $u_{j,y}$ such that

$$\text{(3.41)} \qquad \sup \delta^{-1}|u_{j,y}| + \sup |Du_{j,y}| \leq C\epsilon^{\frac{1}{2(2n-3)}}$$

$$\text{(3.42)} \qquad \left(\left(\bigcup_j P_j(y)\right) \cup \operatorname{graph} u_{j,y}\right) \cap B_{\theta\delta}(y) = \Sigma_j \cap B_{\theta\delta}(y),$$

where $P_j(y)$ are disjoint, each diffeomorphic to the closed disk in $\mathbb{R}^2$, and

$$\text{(3.43)} \qquad \sum_j \operatorname{diam} P_j(y) \leq C\epsilon^{1/2}\delta.$$

Now, from (3.40) and the fact that Σ_j converges in the Hausdorff distance sense, we have

$$\|(L_j(y)-y) - T_y\Sigma\| \leq C\epsilon^{1/2(2n-3)}$$

for j sufficiently large, and hence we can arrange that

$$\operatorname{graph} u_{j,y} = \operatorname{graph} \widetilde{u}_{j,y},$$

where $\widetilde{u}_{j,y}$ is defined over the closure of some domain $\Omega_{j,y} \subset L(y) \equiv y + T_y\Sigma$ with smooth boundary (the inner boundary components being close to circular), and where in place of (3.41) and (3.42) we still have

$$\text{(3.44)} \qquad \sup \delta^{-1}|\widetilde{u}_{j,y}| + \sup |D\widetilde{u}_{j,y}| \leq C\epsilon^{1/2(2n-3)}$$

(3.45)

$$((\bigcup_j P_j(y)) \cup \operatorname{graph} \widetilde{u}_{j,y}) \cap B_{\theta\delta}(y) = \Sigma_j \cap B_{\theta\delta}(y), \quad \sum_k \operatorname{diam} P_{j,k} \leq C\epsilon^{1/2}\delta$$

where the $P_{j,y}$ are pairwise disjoint, each diffeomorphic to the closed unit disk in $\mathbb{R}^2$, and $\partial P_{j,k}$ is a smooth Jordan curve equal to $\operatorname{graph}(\widetilde{u}_{j,y}|\gamma_k)$, where γ_k is one of the inner boundary components of $\Omega_{j,y}$.

Notice that, by virtue of the lower semi-continuity (3.30), (3.39) implies that

$$\int_{B_\delta(y)\cap\Sigma} |\mathbf{A}|^2 \leq \epsilon^2$$

for each $y \in \Sigma \backslash (\bigcup_i B_{\sigma/2}(\xi_i))$, and then the regularity theory established above in the first part of the proof of Theorem 3.1 establishes (assuming, as we do subsequently, that $\epsilon > 0$ is small enough, depending only on n) that for all such y

$$B_{\theta\delta}(y) \cap \Sigma = \operatorname{graph} u \cap B_{\theta\delta}(y)$$

where u is C^2 on $B_{\theta\delta}(y) \cap (y + T_y\Sigma)$ and

$$\delta^{-1}|u| + |Du| + \delta|D^2u| \le C\epsilon^{1/2(2n-3)}, \tag{3.46}$$

with C depending only on n. Here $\theta \in (0, \frac{1}{4})$ is a constant depending only on n.

Now, with the notation

$$S_\rho(y) = \{x + z \colon x \in \Sigma \cap B_\rho(y),\ z \in (T_x\Sigma)^\perp,\ |z| < \theta\rho/2\},\ y \in \Sigma,$$

according to (3.44), (3.45), (3.46) we have, for $y \in \Sigma \backslash (\bigcup_{i=1}^P B_\sigma(\xi_i))$,

$$S_\rho(y) \cap \Sigma_j \subset \operatorname{graph} \widetilde{u}_{j,y} \tag{3.47}$$

for a set $I_j(y)$ of $\rho \in (\theta\delta/2, \theta\delta)$ with

$$\operatorname{measure} I_j(y) \ge \theta\delta/4.$$

Also, for each $i = 1, \dots, P$, by (3.37) we can find a set $I_j(\xi_i)$ such that

$$\operatorname{measure} I_j(\xi_i) \ge \theta\sigma/2$$

and such that

$$S_\rho(\xi_i) \cap \Sigma_j \subset \operatorname{graph} u_{j,i} \tag{3.48}$$

for all $\rho \in I_j(\xi_i)$.

Now select a cover of $\Sigma \backslash (\bigcup_{i=1}^P B_{3\sigma/4}(\xi_i))$ by balls $B_{\theta\delta/2}(y_k)$, $k = 1, \dots, M$ and define $y_{M+i} = \xi_i$ for $i = 1, \dots, P$. By successively applying the selection principle of Appendix 2 we have a subsequence $\{j'\} \subset \{j\}$ and $\tau_k \in \bigcap_{j'} I_{j'}(y_k)$ for $k = 1, \dots, M + P$ such that for each $k \ne \ell$ $\partial B_{\tau_k}(y_k) \cap \Sigma$, $\partial B_{\tau_\ell}(y_\ell) \cap \Sigma$ are either disjoint or intersect transversely, and such that

$$\partial B_{\tau_k}(y_k) \cap \partial B_{\tau_\ell}(y_\ell) \cap \partial B_{\tau_m}(y_m) \cap \Sigma = \emptyset$$

for all distinct $k, \ell, m = 1, \dots, M + P$.

Then the smooth Jordan curves

$$\Gamma_k = (\Sigma\backslash \left(\bigcup_{k=M+1}^{M+P} B_{\tau_k}(y_k)\right) \cap \partial B_{\tau_\ell}(y_\ell), \quad \ell = 1, \dots, M+P$$

divide all of $\Sigma\backslash(\bigcup_{k=M+1}^{M+P}) B_{\tau_k}(y_k))$ into polygonal regions $R_1, \dots, R_Q$. For $\ell = 1, \dots, Q$, let

$$\mathcal{R}_\ell = \{x + z \colon x \in R_\ell,\, z \in (T_x\Sigma)^\perp,\, |z| \le \theta\delta/4\}.$$

Then for j sufficiently large, by (3.37) and (3.44)–(3.48) we have, writing Σ_j for $\Sigma_{j'}$, that $\Sigma_{j'} \cap \mathcal{R}_\ell$ is diffeomophic to R_ℓ, and hence

$$\Sigma_j\backslash\Big(\bigcup_{k=M+1}^{M+P} B_{\tau_k}(y_k)\Big)$$

is diffeomorphic to $\Sigma\backslash\left(\bigcup_{k=M+1}^{M+P} B_{\tau_k}(y_k)\right)$ for all sufficiently large j.

We can now construct comparison surfaces $\widetilde{\Sigma}_j$ with

$$\widetilde{\Sigma}_j\backslash(\bigcup_{k=M+1}^{M+P} B_{\tau_k}(y_k)), \quad \Sigma_j\backslash(\bigcup_{k=M+1}^{M+P} B_{\tau_k}(y_k)), \quad \Sigma\backslash(\bigcup_{k=M+1}^{M+P} B_{\tau_k}(y_k)),$$

all diffeomorphic for each j, with

$$\widetilde{\Sigma}_j \cap V_k = \Sigma_j \cap V_k \tag{3.49}$$

for some neighbourhood V_k of $\partial B_{\tau_k}(y_k)$,

$$\widetilde{\Sigma}_j\backslash\Big(\bigcup_{k=M+1}^{M+P} B_{2\tau_k}(y_k)\Big) = \Sigma\backslash\Big(\bigcup_{k=M+1}^{M+P} B_{2\tau_k}(y_k)\Big), \tag{3.50}$$

and

$$\int_{\widetilde{\Sigma}_j \cap B_{2\tau_k}(y_k)\backslash B_{\tau_k}(y_k)} |\widetilde{\mathbf{A}}_j|^2 \le C\epsilon^2. \tag{3.51}$$

Notice that then by (3.49) and the minimizing property of Σ_j we have

$$\int_{\Sigma_j\backslash(\bigcup_{k=M+1}^{M+P} B_{\tau_k}(y_k))} |\mathbf{H}_j|^2 \le \int_{\widetilde{\Sigma}_j\backslash(\bigcup_{k=M+1}^{M+P} B_{\tau_k}(y_k))} |\widetilde{\mathbf{H}}_j|^2 + \epsilon_j$$

where $\epsilon_j \downarrow 0$. However by (3.50) and (3.51) this gives

$$\int_{\Sigma_j\backslash(\bigcup_{k=M+1}^{M+P} B_{\tau_k}(y_k))} |\mathbf{H}_j|^2 \le \int_{\Sigma\backslash(\bigcup_{k=M+1}^{M+P} B_{2\tau_k}(y_k))} |\mathbf{H}|^2 + \epsilon_j + C\epsilon^2. \tag{3.52}$$

Also, since

$$(3.53)\quad \int_{\widetilde{\Sigma}_j\setminus(\bigcup_{k=M+1}^{M+P} B_{\tau_k}(y_k))} |\widetilde{\mathbf{H}}_j|^2 - \int_{\Sigma_j\setminus(\bigcup_{k=M+1}^{M+P} B_{\tau_k}(y_k))} |\mathbf{H}_j|^2 = \int_{\widetilde{\Sigma}_j\setminus(\bigcup_{k=M+1}^{M+P} B_{\tau_k}(y_k))} |\widetilde{\mathbf{A}}_j|^2 - \int_{\Sigma_j\setminus(\bigcup_{k=M+1}^{M+P} B_{\tau_k}(y_k))} |\mathbf{A}_j|^2$$

by (3.49) and the Gauss-Bonnet formula, we then have also that

$$(3.54)\quad \int_{\Sigma_j\setminus(\bigcup_{k=M+1}^{M+P} B_{\tau_k}(y_k))} |\mathbf{A}_j|^2 \le \int_{\Sigma\setminus(\bigcup_{k=M+1}^{M+P} B_{2\tau_k}(y_k))} |\mathbf{A}|^2 + \epsilon_j + C\epsilon^2.$$

Since we can do this for each $\epsilon > 0$ we thus have in particular that

$$\lim_{\sigma\downarrow 0}\limsup_{j\to\infty} \int_{\Sigma_j\setminus(\bigcup_{k=M+1}^{M+P} B_{\sigma}(y_k))} |\mathbf{H}_j|^2 \le \int_\Sigma |\mathbf{H}|^2$$

$$\lim_{\sigma\downarrow 0}\limsup_{j\to\infty} \int_{\Sigma_j\setminus(\bigcup_{k=M+1}^{M+P} B_{\sigma}(y_k))} |\mathbf{A}_j|^2 \le \int_\Sigma |\mathbf{A}|^2$$

Combining this with the lower semi-continuity (3.29), (3.30), it is then routine to establish the measure-theoretic convergence

$$(3.55)\quad \begin{aligned} |\mathbf{H}_j|^2\mathcal{H}^2 \llcorner \Sigma_j &\to |\mathbf{H}|^2\mathcal{H}^2 \llcorner \Sigma \\ |\mathbf{A}_j|^2\mathcal{H}^2 \llcorner \Sigma_j &\to |\mathbf{A}|^2\mathcal{H}^2 \llcorner \Sigma \end{aligned}$$

in the region $\mathbb{R}^n\setminus\{\xi_1,\dots,\xi_P\}$.

Next we can check that Σ has a minimizing property as follows: According to the above discussion for each ϵ, $\delta > 0$ sufficiently small there is a $\sigma \in (\delta/2, \delta)$ and a $\theta \in (0, \frac{1}{4})$ (depending only on n) with

$$(3.56)\quad \limsup \int_{\Sigma_j\cap(\bigcup_i B_{2\sigma}(\xi_i)\setminus B_\sigma(\xi_i))} |\mathbf{A}_j|^2 \le \delta^2,$$

$$(3.57)\quad \Sigma_j\setminus\Big(\bigcup_i B_\sigma(\xi_i)\Big) \text{ is diffeomorphic to } \Sigma\setminus\Big(\bigcup_i B_\sigma(\xi_i)\Big)$$

$$(3.58)\quad |\mathcal{F}(\Sigma_j\setminus\Big(\bigcup_i B_\sigma(\xi_i)\Big) - \mathcal{F}(\Sigma)| \le \delta^2$$

and such that there are C^∞ functions $u_{i,j}$ over domains $\Omega_{i,j}$ in planes $L_{i,\delta}$ with

$$(3.59)\quad \Sigma_j \cap B_{(1+\theta)\sigma}(\xi_i)\setminus B_{(1-\theta)\sigma}(\xi_i) = \Big(\operatorname{graph} u_{j,i} \cup \Big(\bigcup_k P_{i,j,k}\Big)\Big) \cap B_{(1+\theta)\sigma}(\xi_i)\setminus B_{(1-\theta)\sigma}(\xi_i)$$

$$P_{i,j',k} \cap \left(\bigcup_{i=1}^{P} \operatorname{graph} u_{j'i} | \partial L_{i,\delta} \cap B_\sigma(\xi_i) \right) = \emptyset, \tag{3.60}$$

where $\sum_k \operatorname{diam} P_{i,j,k} \leq C\delta^{1/2}$ and j' is a subsequence of j, and

$$\begin{cases} \delta^{-1}|u_{i,j}| + |Du_{i,j}| \leq C\epsilon^{\frac{1}{2(2n-3)}} \\ \int_{\Omega_{i,j} \cap B_{2\sigma}(\xi_i) \setminus B_\sigma(\xi_i)} |D^2 u_{i,j'}|^2 \leq C\delta^2. \end{cases} \tag{3.61}$$

Thus choosing sequences ϵ_j, $\delta_j \downarrow 0$ sufficiently slowly, we have a sequence $\sigma_j \in (\delta_j/2, \delta_j)$ such that (3.55)–(3.61) hold with $\epsilon_j, \delta_j, \sigma_j$ in place of ϵ, δ, σ respectively. In particular by (3.58)

$$\lim \mathcal{F}\left(\Sigma_j \setminus \left(\bigcup_{i=1}^{P} B_{\sigma_j}(\xi_i) \right) \right) = \mathcal{F}(\Sigma). \tag{3.62}$$

Now recall that for σ sufficiently small we know that $\Sigma \cap \partial B_\sigma(\xi_i)$ is a single smooth Jordan curve close in the C^1-sense to some plane $L_{i,\sigma}$ and in fact $\Sigma \cap B_{2\sigma}(\xi_i) \setminus B_{\sigma/2}(\xi_i)$ is close to the annulus $L_{i,\sigma} \cap B_{2\sigma}(\xi_i) \setminus B_\sigma(\xi_i)$ in the C^1-sense. So we can take a smooth compact surface $\widetilde{\Sigma}$ such that, for suitable points $y_1, \cdots, y_P \in \widetilde{\Sigma}$ and all sufficiently small σ, $\widetilde{\Sigma} \setminus \left(\bigcup_{i=1}^{P} B_\sigma(y_i) \right)$ is diffeomorphic to $\Sigma \setminus \left(\bigcup_{i=1}^{P} B_\sigma(\xi_i) \right)$. Thus with $\sigma = \sigma_j \downarrow 0$ sufficiently slowly (as above), for large j it is possible to replace $\widetilde{\Sigma} \cap B_{\sigma_j}(y_i)$ by a slight deformation of $\Sigma_j \cap B_{\sigma_j}(\xi_i)$ followed by a rigid motion to give $(\Sigma_j \cap B_{\sigma_j}(\xi_i))^*$, such that the composite surface

$$\widetilde{\Sigma}_j \equiv \left(\widetilde{\Sigma} \setminus \left(\bigcup_i B_{\sigma_j(y_i)} \right) \right) \cup \left(\bigcup_i \left(\Sigma_j \cap B_{\sigma_j}(\xi_i) \right)^* \right)$$

is smooth and

$$\mathcal{F}((\Sigma_j \cap B_{\sigma_j}(\xi_i))^*) \leq \mathcal{F}(\Sigma_j \cap B_{\sigma_j}(\xi_i)) + \epsilon_j, \quad \epsilon_j \downarrow 0.$$

Then note that

$$\begin{aligned} &\mathcal{F}\left(\Sigma_j \cap \left(\bigcup_i B_{\sigma_j}(\xi_i) \right) \right) + \mathcal{F}\left(\Sigma_j \setminus \left(\bigcup_i B_{\sigma_j}(\xi_i) \right) \right) = \mathcal{F}(\Sigma_j) \\ &\quad \leq \mathcal{F}(\widetilde{\Sigma}_j) + \epsilon_j \leq \mathcal{F}\left(\widetilde{\Sigma} \setminus \left(\bigcup_i B_{\sigma_j}(y_i) \right) \right) + \mathcal{F}\left(\Sigma_j \cap \left(\bigcup_i B_{\sigma_j}(\xi_i) \right) \right) + \widetilde{\epsilon}_j, \end{aligned}$$

where $\widetilde{\epsilon}_j \downarrow 0$. Thus in view of (3.62) we have

$$\mathcal{F}(\Sigma) \leq \mathcal{F}(\widetilde{\Sigma}). \tag{3.63}$$

Because of this minimizing property, we can repeat the biharmonic comparison argument for Σ (exactly as in the proof of Lemma 2.2) with balls centered at the bad points ξ_i. Thus we conclude that

$$\int_{\Sigma\cap B_\rho(\xi_i)} |\mathbf{A}|^2 \le C\rho^{2\gamma}$$

for all sufficiently small ρ, with C independent of ρ. Then combining this with the previous estimates we now have a fixed constant C such that

$$\int_{\Sigma\cap B_\rho(y)} |\mathbf{A}|^2 \le C\rho^{2\gamma}$$

for all $y \in \Sigma$ and all sufficiently small ρ. Then we deduce that Σ is a $C^{1,\alpha}\cap W^{2,2}$ surface (even in a neighbourhood of the bad points) by the same argument that we used before in the discussion of the good points. In view of the minimizing property (3.63) we can now also apply Lemma 3.2 as before to deduce that Σ is $C^{2,\alpha}$ (and hence real analytic) near the bad points.

Finally, a simple modification of the argument leading to (3.63) shows that Σ minimizes relative to all surfaces with the same genus as Σ, as claimed. This completes the proof of Theorem 3.1. □

4. Proof of the Main Fixed Genus Result in $\mathbb{R}^n$

Suppose first that $g = 1$ and let Σ_k be a sequence of embedded tori with $\mathcal{F}(\Sigma_k) \to \beta_1$. Assume we normalize (as in §3) so that $0 \in \Sigma_k$ and $|\Sigma_k| = 1$. Then by Theorem 3.1 we have a subsequence (still denoted Σ_k) and a real analytic compact embedded surface Σ of genus ≤ 1 which minimizes $\mathcal{F}$ relative to all surfaces $\tilde{\Sigma}$ of the same genus as Σ. If Σ is a sphere (genus 0) then it must be a round sphere (because only round spheres minimize $\mathcal{F}$). We are thus left with the alternatives

$$(4.1)\qquad \begin{cases} \textit{either } \Sigma \text{ is genus 1 with } \mathcal{F}(\Sigma) = \beta_1 \text{ as required} \\ \textit{or } \Sigma \text{ is a round sphere.} \end{cases}$$

Naturally the second alternative *can* occur; what we want to show is that we can make an appropriate inversion and rescaling to give a new minimizing sequence $\tilde{\Sigma}_k$ of tori for which the limit surface $\tilde{\Sigma}$ definitely satisfies the first alternative in (4.1).

As a matter of fact we shall show quite generally the following, which guarantees that, for arbitrary $g \geq 1$, we get a new minimizing sequence converging to a minimizer which is not the round sphere. (Since only round spheres minimize in the genus zero case, this will complete the existence proof.)

Lemma 4.1. *If Σ_k is any genus g minimizing sequence in the sense of* §3 *with $g \geq 1$, then there is a new genus g minimizing sequence $\widetilde{\Sigma}_k \subset B_1(0)$ converging in the sense of Theorem* 3.1 *to a minimizing surface of genus* ≥ 1.

Proof. For the moment consider an arbitrary embedded compact genus g surface $\Sigma \subset \mathbb{R}^n$, and for $y \in \mathbb{R}^n \backslash \Sigma$ let

$$d_\Sigma(y) = \operatorname{dist}(y, \Sigma)$$

and

$$S_\epsilon(y) = \{q \in \Sigma \colon |y - q| < (1 + \epsilon^3) d_\Sigma(y)\}.$$

Now there is $\epsilon_0 \in (0, \frac{1}{4})$ (independent of Σ) such that if $\epsilon \in (0, \epsilon_0)$ and $d_\Sigma(y) < \frac{1}{4}\operatorname{diam}(\Sigma)$ then $\operatorname{diam}(S_\epsilon(y)) \geq \epsilon d_\Sigma(y)$ implies that the inversion $x \mapsto d_\Sigma(y)|x - y|^{-2}(x - y)$ takes Σ to $\widetilde{\Sigma} \subset \overline{B}_1(0)$ with

$$(4.2)\quad \begin{cases} \frac{1}{4} \leq \operatorname{diam}(\widetilde{\Sigma}) \\ \text{Hausdorff distance}(\widetilde{\Sigma}, S) \geq \frac{1}{64} \text{ for any round sphere } S \subset \overline{B}_1(0). \end{cases}$$

(Because the points p, $q \in S_\epsilon(y)$ with $|p - q| \geq \epsilon d_\Sigma(y)$ map to points $\widetilde{p}$, $\widetilde{q} \in \widetilde{\Sigma}$ with $|\widetilde{p} - \widetilde{q}| \geq \epsilon/2$ and with $1 \geq |\widetilde{p}|, |\widetilde{q}| \geq (1 + \epsilon^3)^{-1}$; also since $d_\Sigma(y) < \frac{1}{4}\operatorname{diam}(\Sigma)$ there is a point $\widetilde{r} \in \widetilde{\Sigma}$ with $|\widetilde{r}| \leq \frac{1}{2}$, and one can easily check that—for ϵ sufficiently small—any round sphere in $\overline{B}_1(0)$ must be at least distance $\frac{1}{64}$ from one of the 3 points $\widetilde{p}, \widetilde{q}, \widetilde{r}$, thus giving (4.2) as required.

From now on assume $\epsilon \in (0, \frac{1}{8})$ is small enough to ensure that (4.2) holds under the stated conditions. Thus we have either there is an inversion $\widetilde{\Sigma} \subset \overline{B}_1(0)$ of Σ such that (4.2) holds, or else

$$(4.3)\quad \operatorname{diam}(S_\epsilon(y)) < \epsilon d_\Sigma(y) \quad \forall\, y \in \mathbb{R}^n \backslash \Sigma \text{ with } d_\Sigma(y) < \frac{1}{4}\operatorname{diam}(\Sigma).$$

So consider the alternative that (4.3) holds. Since genus of $\Sigma \geq 1$, there is a smooth map Γ_0 of $\mathbb{S}^{n-2}$ into $\mathbb{R}^n \backslash \Sigma$ which links Σ in the sense that Γ_0 is not homotopic, in $\mathbb{R}^n \backslash \Sigma$, to a constant map. Let $[\Gamma_0]$ denote the class of smooth maps from $\mathbb{S}^{n-2}$ into $\mathbb{R}^n \backslash \Sigma$ which are homotopic to Γ_0 in $\mathbb{R}^n \backslash \Sigma$, and let

$$\delta_0 = \sup_{\Gamma \in [\Gamma_0]} \min_{\omega \in \mathbb{S}^{n-2}} \operatorname{dist}(\Gamma(\omega), \Sigma). \tag{4.4}$$

Assume $\delta_0 < \frac{1}{16} \operatorname{diam}(\Sigma)$, and take any $\delta_1 \in ((1 - \epsilon^3/100)\delta_0, \delta_0)$. We can then by definition select $\Gamma \in [\Gamma_0]$ with $\min_{\omega \in \mathbb{S}^{n-2}} \operatorname{dist}(\Gamma(\omega), \Sigma) \equiv \delta \geq \delta_1$. Now for each y, $z \in \mathbb{R}^n \backslash \Sigma$ it follows from the triangle inequality that

$$S_{\epsilon/2}(z) \subset S_\epsilon(y) \text{ whenever } |y - z| < \tfrac{1}{4}\epsilon^3 d_\Sigma(y).$$

In particular, by (4.3), if $d_\Sigma(y) \leq \frac{1}{4} \operatorname{diam}(\Sigma)$,

$$|(z - q) - (y - p)| \leq 2\epsilon d_\Sigma(y) \tag{4.5}$$

whenever $|y - z| < \frac{1}{4}\epsilon^3 d_\Sigma(y)$ and $q \in S_{\epsilon/2}(z)$, $p \in S_{\epsilon/2}(y)$. Now let $\{\zeta_j\}_{j=1,\ldots,N}$ with support $\zeta_j \subset B_{\epsilon^3 d_\Sigma(y_j)/4}(y_j)$, $j = 1, \ldots, N$, be a partition of unity for the compact set $\Gamma(\mathbb{S}^{n-2})$ subordinate to a covering of $\Gamma(\mathbb{S}^{n-2})$ by a sub-collection $\{B_{\epsilon^3 d_\Sigma(y_j)/4}(y_j) \colon j = 1, \ldots, N\}$ of the collection $\{B_{\epsilon^3 d_\Sigma(y)/4}(y) \colon y \in \Gamma(\mathbb{S}^{n-2})\}$ of balls, the subcollection being chosen so that

(4.6)
any given point of $\Gamma(\mathbb{S}^{n-2})$ is in at most $C(n)$ of the balls $B_{\epsilon^3 d_\Sigma(y_j)/4}(y_j)$.

(Such a collection is guaranteed by the Besicovich covering lemma.) For each j select a point $q_j \in S_{\epsilon/2}(y_j)$, let

$$v_j = y_j - q_j \tag{4.7}$$

and define Γ_t (for $t \in [0, 1]$) by

$$\begin{aligned} \Gamma_t(\omega) &= \Gamma(\omega) + t \sum_{j=1}^{N} \zeta_j(\Gamma(\omega))\epsilon^3 v_j \\ &\equiv \Gamma(\omega) + \frac{t\epsilon^3}{16}(\Gamma(\omega) - q(\omega)) + \frac{t\epsilon^3}{16} \sum_{j=1}^{N} \zeta_j(\Gamma(\omega))(v_j - (\Gamma(\omega) - q(\omega))), \end{aligned} \tag{4.8}$$

where, for each $\omega \in \mathbb{S}^{n-2}$, $q(\omega)$ is any point of $S_{\epsilon/2}(\Gamma(\omega))$. Notice that by (4.5) and (4.6) for each $\omega \in \mathbb{S}^{n-2}$ we can write

$$\Gamma_t(\omega) = \Gamma(\omega) + t\frac{\epsilon^3}{16}(\Gamma(\omega) - q(\omega)) + E, \qquad |E| \leq C(n) t \epsilon^4 d_\Sigma(\Gamma(\omega))$$

at all points where $d_\Sigma(\Gamma(\omega)) \le \frac{1}{8}\operatorname{diam}\Sigma$. So assume $d_\Sigma(\Gamma(\omega)) \le \frac{1}{8}\operatorname{diam}\Sigma$. By (4.3) $|y - q(\omega)| < \epsilon d_\Sigma(\Gamma(\omega))/2$ for all $y \in S_{\epsilon/2}(\Gamma(\omega))$, and it is an easy geometric argument to show that $\Gamma(\omega) + t\frac{\epsilon^3}{16}(\Gamma(\omega) - q(\omega))$ then has distance at least $t\epsilon^3 d_\Sigma(\Gamma(\omega))/32$ from any point of $S_{\epsilon/2}(\Gamma(\omega))$. On the other hand it trivially has distance $\ge \epsilon^3 d_\Sigma(\Gamma(\omega))/2$ from any point of $\Sigma\backslash S_{\epsilon/2}(\Gamma(\omega))$. Thus $\Gamma(\omega) + t\frac{\epsilon^3}{16}(\Gamma(\omega) - q(\omega))$ has distance from Σ of at least $\ge t\epsilon^3 d_\Sigma(\Gamma(\omega))/32$ for any ω such that $d_\Sigma(\Gamma(\omega)) \le \frac{1}{8}\operatorname{diam}\Sigma$. Finally, using the above observation that the error term $|E|$ has magnitude $\le C(n)t\epsilon^4 d_\Sigma(\Gamma(\omega))$, we thus have (for small enough ϵ depending only on n), that $\Gamma_t(\omega)$ has distance at least $\ge t\epsilon^3 d_\Sigma(\Gamma(\omega))/64$ for any ω such that $d_\Sigma(\Gamma(\omega)) \le \frac{1}{8}\operatorname{diam}\Sigma$. If on the other hand $d_\Sigma(\Gamma(\omega)) > \frac{1}{8}\operatorname{diam}\Sigma$, we trivially have that $\Gamma_t(\omega)$ has distance at least $\frac{1}{8}\operatorname{diam}\Sigma - t\epsilon^3 d_\Sigma(\Gamma(\omega))/8$ from Σ. Thus if $\epsilon^3 \sup d_\Sigma(\Gamma(\omega)) < \operatorname{diam}\Sigma$ and if also $\delta_0 < \frac{1}{16}\operatorname{diam}\Sigma$, then we have shown that Γ_t is a homotopy of Γ in $\mathbb{R}^n\backslash\Sigma$, and

$$d_\Sigma(\Gamma_1(\omega)) \ge \min\{(1 + \tfrac{\epsilon^3}{64})\delta_1, \tfrac{1}{16}\operatorname{diam}\Sigma\} > \delta_0,$$

thus contradicting the definition of δ_0. Now (by composing Γ with a suitable smooth retraction), it is easily seen that we could have arranged our original choice of Γ to have the additional property that

$$\sup d_\Sigma(\Gamma(\omega)) \le 2\operatorname{diam}\Sigma,$$

and hence the above argument shows that for any Σ of genus ≥ 1 with $\delta_0 < \frac{1}{16}\operatorname{diam}(\Sigma)$, there is always an inversion $\widetilde{\Sigma} \subset \overline{B}_1(0)$ of Σ as in (4.2) above.

We can now prove the claim of the lemma. We are assuming that the sequence Σ_k converges in the sense of Theorem 1 to a round sphere, otherwise there is nothing to prove. Clearly then, if we apply the above discussion with Σ_k (with diameter bounded between fixed positive constants independent of k) in place of Σ, then (assuming k sufficiently large) we must have that $\delta_0^{(k)} \to 0$ as $k \to \infty$, where $\delta_0^{(k)}$ corresponds to δ_0 when we use Σ_k in place of Σ in the above discussion. Thus by the above discussion there is an inversion $\widetilde{\Sigma}_k \subset \overline{B}_1(0)$ of Σ_k such that (4.2) holds with $\widetilde{\Sigma}_k$ in place of $\widetilde{\Sigma}$. This completes the proof. □

Appendix A.

Here we prove the following variant of the L^2 Poincaré inequality:

Lemma A.1. *Suppose $\delta \in (0, \frac{1}{2})$ is given, and let $\Omega \subset \mathbb{D} \equiv \{x \in \mathbb{R}^2 \colon |x| < 1\}$ be a domain of the form $\Omega = \mathbb{D} \backslash E$, where E is measurable and has projection E_2 onto the y-axis of Lebesgue measure $< \delta$ and projection E_1 onto the x-axis of measure $\le \frac{1}{2}$. Then for any $f \in W^{1,2}(\Omega)$*

$$\inf_{\lambda \in \mathbb{R}} \int_\Omega |f - \lambda|^2 \le C \int_\Omega |Df|^2 + C\delta \sup |f|^2,$$

where C is an absolute constant.

Proof. Pick λ such that $\{x \colon f(x) \ge \lambda\}$ and $\{x \colon f(x) \le \lambda\}$ both have measure $\ge \frac{1}{2}|\Omega|$. First, since the projection E_1 has measure $\le \frac{1}{2}$, we can select a set $S \subset (-\frac{3}{4}, \frac{3}{4})$ of positive measure such that, for all $x_0 \in S$,

$$\int_{\ell_0 \cap \Omega} |Df(x_0, y)|^2 \, dy \le 4 \int_\Omega |Df|^2 \tag{A.1}$$

and $\ell_0 \cap E = \emptyset$, where $\ell_0 = \{(x_0, y) \colon y \in \mathbb{R}\}$. But by 1-dimensional calculus we have

$$\sup_{\ell_0 \cap \Omega} |f - f(x_0, 0)|^2 \le 4 \int_{\ell_0 \cap \Omega} |Df|^2, \quad x_0 \in S,$$

and hence, with $\lambda = f(x_0, 0)$ we have by (A.1) that

$$\sup_{\ell_0 \cap \Omega} |f - \lambda|^2 \le C \int_\Omega |Df|^2, \quad x_0 \in S. \tag{A.2}$$

On the other hand by using the calculus inequality

$$\int_a^b h^2 \le (b-a)^2 \int_a^b (h')^2,$$

valid for $h \in C^1(a, b)$ with $h = 0$ at some point of (a, b), we have, for each $y \in (-1, 1) \backslash E_2$,

$$\int_{L_y \cap \Omega} |f(x, y) - f(x_0, y)|^2 \le 4 \int_{L_y \cap \Omega} |Df|^2,$$

where $L_y = \{(x, y) \colon x \in \mathbb{R}\}$. Then by (A.2) we have

$$\int_{L_y \cap \Omega} |f(x, y) - \lambda|^2 \le 8 \int_{L_y \cap \Omega} |Df|^2 + C \int_\Omega |Df|^2,$$

and by integration over $y \in (-1, 1) \backslash E_2$ we conclude

$$\int_{\Omega \backslash p_2^{-1}(E_2)} |f - \lambda|^2 \le C \int_\Omega |Df|^2,$$

where p_2 is the projection onto the y-axis. Now by adding this to the obvious inequality

$$\int_{\Omega\cap p_2^{-1}(E_2)} |f-\lambda|^2 \le 4\sup|f|^2\delta$$

we then have the required inequality. □

Appendix B.

Here we establish the following simple selection principle, which is used in several places of the present paper.

Lemma B.1. *If $\delta > 0$, if I is a bounded interval of $\mathbb{R}$, and if $A_j \subset I$ is measurable with measure $\ge \delta$ for each $j = 1, 2, \ldots$, then there is a set $S \subset I$ of measure $\ge \delta$ such that each $x \in S$ lies in A_j for infinitely many j.*

Proof. If $\mathcal{L}$ = Lebesgue measure,

$$\mathcal{L}\{x\colon x \in A_j \text{ for any infinitely many } j\} \equiv \mathcal{L}\Big(\bigcap_{i=1}^{\infty}\bigcup_{j\ge i} A_j\Big) = \lim_{i\to\infty}\mathcal{L}\Big(\bigcup_{j\ge i} A_j\Big) \ge \delta,$$

so that $\mathcal{L}\{x\colon x \in A_j \text{ for any infinitely many } j\} \ge \delta$, as required. □

References

[BR] Bryant, R., *A duality theorem for Willmore surfaces,* J. Diff. Geometry **20** (1984), 23–53.

[GT] Gilbarg, D. and Trudinger, N., *Elliptic partial differential equations of second order,* (Second Edition), Springer-Verlag 1983.

[HJ] Hutchinson, J., *Second fundamental form for varifolds and the existence of surfaces minimizing curvature,* Indiana Univ. Math. J. **35** (1986), 45–71.

[K] Kusner, R., Personal Communication.

[L] Lawson, H. B., *Complete minimal surfaces in* $\mathbb{S}^3$, Ann. Math. **92** (1970), 335–374.

[LY] Li, P. and Yau, S. T., *A new conformal invariant and its applications to the Willmore conjecture and the first eigenvalue on compact surfaces,* Invent. Math. **69** (1982), 269–291.

[MCB] Morrey, C. B., *Multiple integrals in the calculus of variations*, Springer Verlag 1966.

[P] Pinkall, U., Personal Communication.

[SS] Schoen, R. and Simon, L., *Regularity of stable minimal hypersurfaces,* Comm. Pure and Appl. Math. **34** (1981), 741–793.

[SL1] Simon, L., *Existence of Willmore Surfaces,* Proc. Centre for Math. Anal., **10** (1985), 187–216.

[SL2] Simon, L., *Lectures on Geometric Measure Theory*, Proc. Centre for Math. Anal., **3** (1983).

[TF] Treves, F., *Basic linear partial differential equations*, Academic Press 1975.

[WJ] Wiener, J., *On a problem of Chen, Willmore, et al,* Indiana Univ. Math. J. **27** (1978), 19–35.

[W] Willmore, T., *Total curvature in Riemannian geometry,* Wiley 1982.

STANFORD UNIVERSITY, U. S. A.

RECEIVED OCTOBER 23, 1992