Willmore minimizers with prescribed isoperimetric ratio

Dissertation zur Erlangung des Doktorgrades der Fakultät für Mathematik und Physik der Albert-Ludwigs-Universität Freiburg im Breisgau

vorgelegt von Johannes Schygulla

betreut durch Prof. Dr. Ernst Kuwert

September 2011
Dekan: Prof. Dr. Kay Königsmann

1. Gutachter: Prof. Dr. Ernst Kuwert
2. Gutachter: Prof. Dr. Heiko von der Mosel

Datum der mündlichen Prüfung: 08. Dezember 2011
## Contents

1. Introduction ............................................. 1

2. Upper Bound for the Infimum ......................... 3

3. Proof of Theorem 1.1 .................................. 6

4. Proof of Theorem 1.2 .................................. 46

5. Appendix ................................................. 49
   5.1 The Monotonicity Formula ......................... 49
   5.2 The Graphical Decomposition Lemma ................. 53
   5.3 Useful Results ........................................ 54

Bibliography ................................................ 58
1 Introduction

In the spontaneous curvature model for lipid bilayers due to Helfrich [Hel73], the membrane of a vesicle is described as a 2-dimensional, embedded surface \( \Sigma \subset \mathbb{R}^3 \), whose energy is given by

\[
E(\Sigma) = \kappa_1 \int_\Sigma (H - C_0)^2 \, d\mu + \kappa_2 \int_\Sigma K \, d\mu,
\]

where \( H, K \) denote the mean curvature and Gauss curvature, \( \mu \) is the induced area measure and \( \kappa_1, \kappa_2 \) and \( C_0 \) are constants.

Restricting to surfaces of the type of the sphere, the second term reduces to the constant \( 4\pi \kappa_2 \) by the Gauss-Bonnet Theorem. Reducing further to the simplest case of spontaneous curvature \( C_0 = 0 \), the energy becomes up to a factor and an additive constant the Willmore energy

\[
W(\Sigma) = \frac{1}{4} \int_\Sigma |\vec{H}|^2 \, d\mu. \tag{1.1}
\]

According to Helfrich [Hel73], the shapes of the vesicles should be minimizers of the elastic energy \( E \) subject to prescribed area and enclosed volume. Since the Willmore energy is scaling invariant, the two constraints actually reduce to the condition that the isoperimetric ratio of the surface \( \Sigma \), given by

\[
I(\Sigma) = \frac{(6 \sqrt{\pi})^\frac{1}{4} V(\Sigma)^\frac{1}{4}}{A(\Sigma)^\frac{1}{2}}, \tag{1.2}
\]

is prescribed. Here \( A(\Sigma) \) denotes the area of \( \Sigma \) and \( V(\Sigma) \) the volume enclosed by \( \Sigma \), namely the volume of the bounded component of \( \mathbb{R}^3 \setminus \Sigma \). The normalizing constant \( (6 \sqrt{\pi})^\frac{1}{4} \) is chosen such that \( I(\Sigma) \in (0, 1] \), in particular \( I(S^2) = 1 \).

For given \( \sigma \in (0, 1] \) we denote by \( \mathcal{M}_\sigma \) the class of smoothly embedded surfaces \( \Sigma \subset \mathbb{R}^3 \) with the type of \( S^2 \) such that \( I(\Sigma) = \sigma \), and we introduce the function

\[
\beta : (0, 1] \to \mathbb{R}_+ , \quad \beta(\sigma) = \inf_{\Sigma \in \mathcal{M}_\sigma} W(\Sigma). \tag{1.3}
\]

In particular \( \mathcal{M}_1 = \{ \text{round spheres \subset \mathbb{R}^3 } \} \) and \( \beta(1) = 4\pi \).

Here we prove the following result.

**Theorem 1.1** For every \( \sigma \in (0, 1) \) there exists a surface \( \Sigma \in \mathcal{M}_\sigma \) such that

\[
W(\Sigma) = \beta(\sigma).
\]

Moreover the function \( \beta \) is continuous, strictly decreasing and satisfies

\[
\lim_{\sigma \to 0^+} \beta(\sigma) = 8\pi.
\]

It appears that so far no global existence results for the Helfrich model have been obtained. In [NT03] the authors prove existence of a one-parameter family of critical points bifurcating from the sphere. Assuming axial symmetry, several authors computed possible candidates for minimizers by solving numerically the Euler-Lagrange equations (see [BLS91], [DH76]). For example a picture one should
keep in mind is the following numerical experiment in [BLS91] for different values of $\sigma \in (0, 1]$.

![Figure 1: Seifert, Berndl, Lipowsky, Phys. Rev. A 1991.](image)

This picture suggests that the minimizers of the Willmore energy in the class $\mathcal{M}_\sigma$ converge to a doubly covered sphere as $\sigma \to 0$. This actually holds in a measure theoretic sense and is stated in the next Theorem.

**Theorem 1.2** Let $\{\sigma_k\}_{k\in\mathbb{N}} \subset (0, 1)$ such that $\sigma_k \to 0$ and $\Sigma_k \in \mathcal{M}_{\sigma_k}$ such that $\mathcal{W}(\Sigma_k) = \beta(\sigma_k)$. After translation and scaling (such that $0 \in \Sigma_k$ and $\mathcal{H}^2(\Sigma_k) = 1$) there exists a subsequence $\Sigma_k'$ which converges to a double sphere in the sense of measures, namely

$$\mu_k' \to \mu \quad \text{in } C^0_c(\mathbb{R}^3)' ,$$

where $\mu_k' = \mathcal{H}^2\Sigma_k'$ and $\mu = 2\mathcal{H}^2\partial B_r(a)$ for some $r > 0$ and $a \in \mathbb{R}^3$.

In order to prove Theorem 1.1 we adopt the methods developed by L. Simon in [Sim93], where he proved the existence of a smooth torus in $\mathbb{R}^n$ minimizing the Willmore energy among tori. This was the first study of the Willmore functional from the viewpoint of the calculus of variations and nonlinear partial differential equations. Simon also obtained existence of higher genus minimizers, provided that a certain Douglas condition holds. M. Bauer and E. Kuwert verified in [BK03] this Douglas condition, thus showing that the infimum of the Willmore functional among genus $p$ surfaces in $\mathbb{R}^n$ is attained for any $p \geq 2$.

We now briefly outline the content of this thesis. In section 2 we prove that the function $\beta$ is decreasing and that $\beta(\sigma) < 8\pi$ for all $\sigma \in (0, 1]$. This is a crucial estimate in the existence proof. In section 3 we prove Theorem 1.1. Section 4 is dedicated to the proof of Theorem 1.2, using similar techniques as in the proof of Theorem 1.1. Finally in the appendix we collect some important results we need during the proofs, as for example the Graphical Decomposition Lemma and the Monotonicity Formula proved by Simon in [Sim93].

**Acknowledgments**

This work was done within the framework of project B.3 of the DFG Collaborative Research Center SFB/Transregio 71. I would like to thank my advisor Prof. Ernst Kuwert for his support. I also would like to express my gratitude for the support I received from the DFG Collaborative Research Center SFB/Transregio 71.

Last, but definitely not least, I wish to express my gratitude to my girlfriend for everything else.
2 Upper Bound for the Infimum

In this section we prove an upper bound for the infimum of the Willmore energy in the class $M_\nu$, namely we show that $\beta(\sigma) < 8\pi$ for all $\sigma \in (0, 1]$. This estimate is crucial for the proof of Theorem 1.1 and reveals the first difference to [Sim93]. In particular Simon defines the quantity

$$\beta^a_g = \inf \{ W(\Sigma) | \Sigma \subset \mathbb{R}^n \text{ compact, embedded surface, genus } \Sigma = g \}. $$

By a construction of genus $g$ minimal surfaces $\Sigma_g \subset S^3$ due to Lawson and the conformal invariance of the Willmore functional it follows that $\beta_g^a < 8\pi$. Because of the constraint in our setting we do not have this estimate a priori.

The proof is based on attaching suitable spheres to a catenoid and the inversion of the constraint in our setting we do not have this estimate a priori.

The proof is based on attaching suitable spheres to a catenoid and the inversion of the resulting $C^{1,1}$-surfaces at a sphere together with an argument involving the Willmore flow and its properties. A reference where the authors also analyze inverted catenoids and their relation to the Willmore energy is [CV07]. For the part concerning the Willmore flow see [KS04].

Lemma 2.1 The function $\beta$ is decreasing and

$$\beta(\sigma) = \inf_{\Sigma \in M_\nu} W(\Sigma) < 8\pi \text{ for all } \sigma \in (0, 1].$$

Proof: For $a > 0$ let

$$g_a : \mathbb{R} \to \mathbb{R}^3, \quad g_a(t) = \left( a^2 \cosh \frac{t}{a^2}, 0, t \right).$$

By rotation around the $z$-axis one gets a (scaled) catenoid. Moreover define for $b \in \mathbb{R}$ and $r > 0$ the (half-)circle with center $(0, 0, b)$ and radius $r$ by

$$f_{b,r} : [b - r, b + r] \to \mathbb{R}^3, \quad f_{b,r}(t) = \left( \sqrt{r^2 - (t - b)^2}, 0, t \right).$$

Now we want to find for given $a > 0$ corresponding $b = b(a) > 0$ and $r = r(a) > 0$ such that $g_a(a) = f_{b,r}(a)$ and $g_a'(a) = f_{b,r}'(a)$. It follows that

$$a^2 \cosh \frac{1}{a} = \sqrt{r^2 - (a - b)^2} \quad \text{and} \quad \sinh \frac{1}{a} = \frac{b - a}{\sqrt{r^2 - (a - b)^2}},$$

and thus we get

$$b = a \left( 1 + a \sinh \frac{1}{a} \cosh \frac{1}{a} \right) \quad \text{and} \quad r = a^2 \cosh^2 \frac{1}{a}.$$ 

By attaching the right part of the circle to the right part of the catenoid and reflecting at the $x$-axis we get for given $a > 0$ the $C^{1,1}$-curve given by

$$h_a : [-a \left( 1 + a \left( \frac{a}{2} \left( 1 + e^2 \right) \right) \right), a \left( 1 + a \left( \frac{a}{2} \left( 1 + e^2 \right) \right) \right)] \to \mathbb{R}^3,$$

$$h_a(t) = \begin{cases} (s_a(-t), 0, t) & \text{for } t \in [-a \left( 1 + \frac{a}{2} \left( 1 + e^2 \right) \right), -a] \\ (a^2 \cosh \frac{1}{a^2}, 0, t) & \text{for } t \in [-a, a] \\ (s_a(t), 0, t) & \text{for } t \in (a, a \left( 1 + \frac{a}{2} \left( 1 + e^2 \right) \right)] \end{cases},$$

where $s_a(t) = \sqrt{a^4 \cosh^4 \frac{t}{a} - \left( t - a \left( 1 + a \sinh \frac{1}{a} \cosh \frac{1}{a} \right) \right)^2}.$
By rotating these curves around the z-axis we get embedded $C^{1,1}$-spheres $\Sigma_a \subset \mathbb{R}^3$. Since $\Sigma_a$ consists of parts of two spheres together with a piece of a catenoid, it follows that $W(\Sigma_a) < 8\pi$ for all $a > 0$. Next we invert these surfaces at the sphere $\partial B_1(e_3)$. This inversion is given by $I : \mathbb{R}^3 \to \mathbb{R}^3$, $I(x) = e_3 + \frac{1}{|x - e_3|^2}$. By this procedure and because of the invariance of the Willmore energy under inversions (with center not on the surface) we get embedded $C^{1,1}$-spheres $\tilde{\Sigma}_a = I(\Sigma_a) \subset \mathbb{R}^3$ such that $W(\tilde{\Sigma}_a) < 8\pi$ for all $a > 0$.

Observe that the isoperimetric ratio $I(\tilde{\Sigma}_a) \to 0$ as $a \to 0$. Approximating by smooth surfaces we have therefore shown that for every $\varepsilon > 0$ there exists a smooth, embedded sphere $\Sigma \subset \mathbb{R}^3$ with isoperimetric ratio $I(\Sigma) < \varepsilon$ and Willmore energy $W(\Sigma) < 8\pi$. Using Theorem 5.2 in [KS04], the Willmore flow $\Sigma_t$ with initial data
The same Willmore flow argument also yields the claimed monotonicity of the function \( \beta \). In particular let \( \sigma_0 \in (0, 1) \) and \( \varepsilon > 0 \) such that \( \beta(\sigma_0) + \varepsilon < 8\pi \). Let \( \Sigma_0 \in M_{\sigma_0} \) such that \( \mathcal{W}(\Sigma_0) \leq \beta(\sigma_0) + \varepsilon \). Again the Willmore flow \( \Sigma_t \) with initial data \( \Sigma_0 \) exists smoothly for all times and converges to a round sphere such that \( \mathcal{W}(\Sigma_t) \) is decreasing in \( t \). Therefore for every \( \sigma \in (\sigma_0, 1] \) there exists a surface \( \Sigma_{\sigma} \in M_{\sigma} \) with \( \mathcal{W}(\Sigma_{\sigma}) \leq \mathcal{W}(\Sigma_0) \leq \beta(\sigma_0) + \varepsilon \). By taking the infimum on the left hand side and letting \( \varepsilon \to 0 \), the Lemma follows. \( \square \)
3 Proof of Theorem 1.1

For proving existence of a minimizer for the Willmore energy in the class $M_\sigma$, we use the direct method in the calculus of variations, namely the concept of minimizing sequences. For given $\sigma \in (0, 1)$ let $\{\Sigma_k\}_{k \in \mathbb{N}} \subset M_\sigma$ be a minimizing sequence for the Willmore energy in the class $M_\sigma$. Since the Willmore energy and the isoperimetric ratio are invariant under translations and scalings, we may assume in view of Lemma 2.1 that for some $\delta_0 > 0$ and all $k$

$$\mathcal{H}^2(\Sigma_k) = 1, \quad 0 \in \Sigma_k, \quad \mathcal{W}(\Sigma_k) \leq 8\pi - \delta_0. \quad (3.1)$$

Using Lemma 1.1 in [Sim93], which bounds the diameter of a surface in terms of the area and the Willmore energy, we get an uniform diameter bound for $\Sigma_k$. Since $0 \in \Sigma_k$, it follows that $\Sigma_k \subset B_R(0)$ for some $R < \infty$.

Define the Radon measures $\mu_k$ in $\mathbb{R}^3$ by

$$\mu_k = \mathcal{H}^2 \llcorner \Sigma_k. \quad (3.3)$$

Since $\mu_k$ is the restriction of the 2-dimensional Hausdorff measure to a smooth, embedded surface, it follows that $\mu_k$ defines an integral, rectifiable 2-varifold. Moreover we have that $\mu_k(\mathbb{R}^3) = \mathcal{H}^2(\Sigma_k) = 1$, that the 2-density $\theta^2(\mu_k, \cdot) = 1$ for all $x \in \text{spt} \mu_k$ and that the first variation can be bounded by an universal constant, namely we get in view of (3.1) that

$$||\delta \mu_k||(\mathbb{R}^3) = \sup \left\{ -\int \langle X, \bar{H}_k \rangle d\mu_k \middle| X \in C^1_c(\mathbb{R}^3, \mathbb{R}^3), |X| \leq 1 \right\} \leq \sqrt{32\pi}.$$ 

By a compactness result for varifolds (see [Sim83]), there exists an integral, rectifiable 2-varifold $\mu$ in $\mathbb{R}^3$ with density $\theta^2(\mu, \cdot) \geq 1$ $\mu$-a.e. and weak mean curvature vector $\bar{H} \in L^2(\mu)$, such that (after passing to a subsequence)

$$\mu_k \rightharpoonup \mu \quad \text{in } C^0_c(\mathbb{R}^3)' \quad \left( \equiv \lim_{k \to \infty} \int f \, d\mu_k = \int f \, d\mu \quad \text{for all } f \in C^0_c(\mathbb{R}^3) \right), \quad (3.4)$$

$$\lim_{k \to \infty} \int \langle X, \bar{H}_k \rangle d\mu_k = \int \langle X, \bar{H} \rangle d\mu \quad \text{for all } X \in C^1_c(\mathbb{R}^3, \mathbb{R}^3), \quad (3.5)$$

$$\frac{1}{4} \int_U |\bar{H}_k|^2 d\mu_k \leq \liminf_{k \to \infty} \frac{1}{4} \int_U |\bar{H}_k|^2 d\mu_k \leq 8\pi - \delta_0 \quad \text{for all open } U \subset \mathbb{R}^3. \quad (3.6)$$

The Monotonicity formula Theorem 5.3 and (3.6) applied to $U = \mathbb{R}^3$ yield

$$\theta^2(\mu, x) \in \left[ 1, 2 - \frac{\delta_0}{4\pi} \right] \quad \text{for all } x \in \text{spt} \mu. \quad (3.7)$$

Since $\mu$ is integral, we therefore get that

$$\theta^2(\mu, x) = 1 \quad \text{for } \mu-\text{a.e. } x \in \text{spt} \mu. \quad (3.8)$$

Our candidate for a minimizer is given by

$$\Sigma = \text{spt} \mu. \quad (3.9)$$
Because of the varifold convergence and (3.2) it follows that
\[ \Sigma \subset B_{\rho}(0), \]  
and with this inclusion we get in addition that
\[ \Sigma \subset \mathbb{R}^3 \quad \text{and} \quad \mu(\mathbb{R}^3) = 1. \]

As a first consequence we prove the following Lemma.

**Lemma 3.1** After passing to a subsequence we have that
\[ \Sigma_k \to \Sigma \quad \text{in the Hausdorff distance sense.} \]

**Proof:** The proof is based on the Monotonicity formula Theorem 5.3 and is due to [Sim93], page 310. We have to show that for all \( \epsilon > 0 \) there exists a \( K \in \mathbb{N} \) such that
\[ \Sigma \subset B_{\rho}(\Sigma_k) \quad \text{and} \quad \Sigma_k \subset B_{\rho}(\Sigma) \quad \text{for all} \quad k \geq K. \]

1.) Assume that there exists a \( \epsilon > 0 \) such that for all \( K \in \mathbb{N} \) there exists a \( k \geq K \) such that \( \Sigma \setminus B_{\rho}(\Sigma_k) \neq \emptyset \). Since \( \Sigma \) is compact by (3.11), there exists a sequence of points \( x_k \in \Sigma \setminus B_{\rho}(\Sigma_k) \) such that (after passing to a subsequence) \( x_k \to x \in \Sigma \). Since \( x_k \notin B_{\rho}(\Sigma_k) \), there exists a \( K \in \mathbb{N} \) such that \( x \notin B_{\rho}(\Sigma_k) \) for all \( k \geq K \).

Let \( f \in C^0(\bar{B}_r(x)) \) such that \( 0 \leq f \leq 1, f \equiv 1 \) on \( B_{\frac{r}{2}}(x) \). Since \( \mu_k \to \mu \) in \( C^0_{\text{loc}}(\mathbb{R}^3) \), \( x \in \Sigma \) and \( B_{\frac{r}{2}}(x) \cap \Sigma_k = \emptyset \) for all \( k \geq K \), we get that
\[ 0 < \mu(B_{\frac{r}{2}}(x)) \leq \int f \, d\mu = \lim_{k \to \infty} \int f \, d\mu_k \leq \limsup_{k \to \infty} \mu_k(B_{\frac{r}{2}}(x)) = 0. \]

This is a contradiction and therefore the first inclusion is shown.

2.) Suppose that there exists a \( \epsilon > 0 \) such that for all \( K \in \mathbb{N} \) there exists a \( k \geq K \) such that \( \Sigma_k \setminus B_{\rho}(\Sigma) \neq \emptyset \). Because of (3.2) there exist points \( x_k \in \Sigma_k \setminus B_{\rho}(\Sigma) \) such that (after passing to a subsequence) \( x_k \to x \in \overline{B_{\rho}(0)} \setminus B_{\rho}(\Sigma) \).

Let \( f \in C^0(\bar{B}_{\frac{r}{2}}(x)) \) such that \( 0 \leq f \leq 1, f \equiv 1 \) on \( B_{\frac{r}{4}}(x) \). It follows that
\[ 0 = \mu(B_{\frac{r}{4}}(x)) \geq \int f \, d\mu = \lim_{k \to \infty} \int f \, d\mu_k \geq \liminf_{k \to \infty} \mu_k(B_{\frac{r}{4}}(x)). \]

Since \( x_k \to x \), we can select a \( K \in \mathbb{N} \) such that \( |x_k - x| \leq \frac{\rho}{8} \) for all \( k \geq K \). Now let \( \rho \in \left( \frac{\epsilon}{8}, \frac{\epsilon}{4} \right) \) and assume that \( \Sigma_k \setminus \partial B_{\rho}(x) = \emptyset \) for some \( k \geq K \). Since \( \Sigma_k \) is connected, it follows that either \( \text{dist}(\Sigma_k, x) > \rho \) or \( \Sigma_k \subset B_{\rho}(x) \). If \( \text{dist}(\Sigma_k, x) > \rho \), we get \( \frac{\rho}{8} < \rho < |x_k - x| \leq \frac{\rho}{8} \), a contradiction. Therefore \( \Sigma_k \subset B_{\rho}(x) \subset B_{\frac{r}{4}}(x) \).

Now let \( f \in C^0(\bar{B}_{\frac{r}{2}}(x)) \) with \( 0 \leq f \leq 1, f \equiv 1 \) on \( B_{\frac{r}{4}}(x) \). Since \( \mathcal{H}^2(\Sigma_k) = 1, \Sigma_k \subset B_{\frac{r}{4}}(x) \) and \( B_{\frac{r}{4}}(x) \cap \Sigma = \emptyset \) we get that
\[ 1 = \lim_{k \to \infty} \mu_k(B_{\frac{r}{4}}(x)) \leq \lim_{k \to \infty} \int f \, d\mu_k = \int f \, d\mu \leq \mu(B_{\frac{r}{4}}(x)) = 0. \]

Therefore for \( k \geq K \) sufficiently large we again arrive at a contradiction. It follows that there exists a \( K \in \mathbb{N} \) such that \( \Sigma_k \cap \partial B_{\rho}(x) \neq \emptyset \) for all \( \rho \in \left( \frac{\epsilon}{8}, \frac{\epsilon}{4} \right) \) and \( k \geq K \).
Therefore for all \( N \geq 1 \), all \( k \geq K \) and all \( j \in \{1, \ldots, N-1\} \) there exists a point \( z_{k,j} \in \Sigma_k \cap \partial B_{\rho_j}(x) \), where \( \rho_j = \left(1 + \frac{k}{N}\right)^2 \). The Monotonicity formula Theorem 5.3 applied to \( \Sigma_k \), \( z_{k,j} \) and \( \rho = \frac{k}{16N} \) yields (observe that \( \theta^2(\mu_k, z_{k,j}) = 1 \))

\[
1 \leq c \left( \rho^{-2} \mathcal{H}^2 \left( \Sigma_k \cap B_{\rho_j}(z_{k,j}) \right) + \mathcal{W} \left( \Sigma_k \cap B_{\rho_j}(z_{k,j}) \right) \right).
\]

Since the balls \( B_{\rho_j}(z_{k,j}) \) are pairwise disjoint, we get by summing over \( j \) that

\[
(N - 1) \leq c \left( \rho^{-2} \mathcal{H}^2 \left( \Sigma_k \cap \bigcup_{j=1}^{N-1} B_{\rho_j}(z_{k,j}) \right) + \mathcal{W} \left( \Sigma_k \cap \bigcup_{j=1}^{N-1} B_{\rho_j}(z_{k,j}) \right) \right).
\]

Since \( \bigcup_{j=1}^{N-1} B_{\rho_j}(z_{k,j}) \subset B_{\frac{3}{2}}(x) \) and \( \mathcal{W}(\Sigma_k) \leq 8\pi \), we get

\[
(N - 1) \leq c \left( \rho^{-2} \mu_k(B_{\frac{3}{2}}(x)) + 8\pi \right).
\]

By letting \( k \to \infty \) we arrive in view of (3.12) at

\[
(N - 1) \leq 8\pi c.
\]

Since \( N \) was arbitrary, we get a contradiction by choosing \( N \) sufficiently large, and the Lemma is proved. \( \square \)

In order to prove regularity, we would like to apply Simon’s Graphical Decomposition Lemma Theorem 5.5 to \( \Sigma_k \) simultaneously for infinitely many \( k \). But the most important assumption is that the \( L^2 \)-norm of the second fundamental form is locally small, which we will need simultaneously for infinitely many \( k \), but which we may not have everywhere. Therefore we define the so called bad points with respect to a given \( \varepsilon > 0 \) in the following way (they will be the points where the curvature concentrates): Define the Radon measures \( \alpha_k \) on \( \mathbb{R}^3 \) by

\[
\alpha_k = \mu_k |A_k|^2.
\]

From the Gauss-Bonnet formula and the uniform bound on the Willmore energy it follows that

\[
\alpha_k(\mathbb{R}^3) = \int_{\Sigma_k} |A_k|^2 \, d\mathcal{H}^2 = 4 \mathcal{W}(\Sigma_k) - 2 \int_{\Sigma_k} K \, d\mathcal{H}^2 = 4 \mathcal{W}(\Sigma_k) - 8\pi \leq 24\pi.
\]

By a compactness result for Radon measures (see [EG92]) there exists a Radon measure \( \alpha \) on \( \mathbb{R}^3 \) such that (after passing to a subsequence) \( \alpha_k \to \alpha \) in \( C^0_0(\mathbb{R}^3)' \). It follows that \( \text{spt} \alpha \subset \Sigma \subset \subset \mathbb{R}^3 \), and therefore \( \alpha(\mathbb{R}^3) \leq 24\pi \). We define the bad points with respect to \( \varepsilon > 0 \) by

\[
\mathcal{B}_\varepsilon = \left\{ \xi \in \Sigma \left| \alpha(\{\xi\}) > \varepsilon^2 \right. \right\}.
\]

The following properties concerning the set of bad points hold.

**Lemma 3.2** For given \( \varepsilon > 0 \) the following holds:

(i) There exist only finitely many bad points, namely \( \# \mathcal{B}_\varepsilon \leq \frac{C}{\varepsilon^2} \).

(ii) For \( \xi_0 \in \Sigma \setminus \mathcal{B}_\varepsilon \) there exists a \( \rho_0 = \rho_0(\xi_0, \varepsilon) > 0 \) such that

\[
\int_{\Sigma_k \cap B_{\rho_0}(\xi_0)} |A_k|^2 \, d\mathcal{H}^2 \leq 2\varepsilon^2 \quad \text{for } k \text{ sufficiently large.}
\]
Proof: Let \( \{\xi_1, \ldots, \xi_{p}\} \subset \mathcal{B}_c \) be distinct bad points. Since \( \alpha(\mathbb{R}^3) \leq c \), we get that \( c \geq \sum_{p=1}^{P} \alpha(\{(p,\xi_{p})\}) \geq Pe^2 \), and (i) follows. Moreover for \( \xi_0 \in \Sigma \) we have that \( \lim_{\rho \to 0} \alpha(\mathcal{B}_{\rho}(\xi_0) \setminus \{\xi_0\}) = 0 \), and therefore for \( \xi_0 \in \Sigma \setminus B_c \) there exists a \( \rho_0 = \rho_0(\xi_0, \varepsilon) > 0 \) such that \( \alpha(\mathcal{B}_{\rho_0}(\xi_0)) < 2\varepsilon^2 \). Since \( \alpha_k \to \alpha \) weakly as measures, (ii) follows. \( \square \)

Now fix \( \xi_0 \in \Sigma \setminus B_c \) and let \( \rho_0 \) as in (3.14). Let \( \xi \in \Sigma \cap \mathcal{B}_{\rho_0}(\xi_0) \). We want to apply Simon’s Graphical Decomposition Lemma Theorem 5.5 to show that in a neighborhood around the point \( \xi \) the surfaces \( \Sigma_k \) can be written as a graph with small Lipschitz norm together with some "pimples" with small diameter. This is done in exactly the same way as in [Sim93]. We just sketch this procedure: By Lemma 3.1 there exists a sequence \( \xi_k \in \Sigma_k \) such that \( \xi_k \to \xi \). By (3.14) and the Monotonicity formula applied to \( \Sigma_k \) and \( \xi_k \), the assumptions of the Graphical Decomposition Lemma Theorem 5.5 are satisfied for each given \( \rho \leq \frac{\varepsilon_0}{4} \) and infinitely many \( k \). Since \( W(\Sigma_k) \leq 8\pi - \delta_0 \), we can also apply Lemma 1.4 in [Sim93] to deduce that for \( \theta \in (\frac{1}{2}, \frac{3}{2}) \) sufficiently small, \( \tau \in (\frac{\theta}{2}, \frac{\theta}{2}) \) and infinitely many \( k \) only one of the discs \( D_{\kappa}^k \) appearing in the Graphical Decomposition Lemma can intersect the ball \( B_{\rho_k}(\xi_k) \) (see Theorem 5.5 for the notation). Moreover, by a slight perturbation from \( \xi_k \) to \( \xi \), we may assume that \( \xi \in L_k \) for all \( k \). Now the planes \( L_k \) are fixed at the common point \( \xi \), and therefore \( L_k \to L \in \xi + G_2(\mathbb{R}^3) \), which yields that we may furthermore assume that the planes, on which the graph functions are defined, do not depend on \( k \). After all we get a graphical decomposition in the following way.

**Lemma 3.3** Let \( \xi_0 \in \Sigma \setminus B_c \) and \( \rho_0 \) as in (3.14). Let \( \xi \in \Sigma \cap \mathcal{B}_{\rho_0}(\xi_0) \). It follows that for \( \varepsilon \leq \varepsilon_0 \), all \( \rho \leq \frac{\rho_0}{4} \) and infinitely many \( k \) there exist pairwise disjoint closed subsets \( P_k^1, \ldots, P_k^{N_k} \) of \( \Sigma_k \) such that

\[
\Sigma_k \cap \mathcal{B}_{\rho_k}(\xi) = D_k \cap \mathcal{B}_{\rho_k}(\xi) = \text{graph } u_k \cup \bigcup_n P_n^k \cap \mathcal{B}_{\rho_k}(\xi),
\]

where \( D_k \) is a topological disc and where the following holds:

1. The sets \( P_n^k \) are topological discs disjoint from graph \( u_k \).
2. \( u_k \in C^\infty(\overline{\Omega_k}, \mathbb{R}^3) \), where \( L \subset \mathbb{R}^3 \) is a 2-dimensional plane containing \( \xi \), and \( \Omega_k = (B_{\lambda_k}(\xi_k) \cap L) \setminus \bigcup_m d_{k,m} \). Here \( \lambda_k > \frac{\xi}{4} \) and the sets \( d_{k,m} \subset L \) are pairwise disjoint closed discs which do not intersect \( \partial B_{\lambda_k}(\xi_k) \cap L \).
3. The following inequalities hold:

\[
\sum_m \text{diam } d_{k,m} + \sum_n \text{diam } P_n^k \leq c \left( \int_{\Sigma_k \cap \mathcal{B}_{\rho}(\xi)} |A_k|^2 d\mathcal{H}^2 \right)^{\frac{1}{2}} \rho \leq c\varepsilon^{\frac{1}{2}} \rho,
\]

\[
\|u_k\|_{L^\infty(\Omega_k)} \leq c\varepsilon^{\frac{1}{2}} \rho + \delta_k, \text{ where } \delta_k \to 0,
\]

\[
\|D u_k\|_{L^\infty(\Omega_k)} \leq c\varepsilon^{\frac{1}{2}} + \delta_k, \text{ where } \delta_k \to 0.
\]
Now we leave the varifold context and define the functions
\[ \chi_k = \chi_{\Omega_k} \in BV(\mathbb{R}^3), \]
where \( \Omega_k \subset \mathbb{R}^3 \) is the open set surrounded by \( \Sigma_k \). Since \( \mathcal{H}^2(\Sigma_k) = 1 \), it follows that
\[ \| \chi_k \|_{L^1(\mathbb{R}^3)} = \frac{\sigma^3}{6 \sqrt{\pi}} \quad \text{and} \quad |D \chi_k|(U) = \mu_k(U) \leq 1 \quad \text{for every open } U \subset \mathbb{R}^3. \]
Therefore the sequence \( \chi_k \) is uniformly bounded in \( BV(\mathbb{R}^3) \), and a compactness result for BV-functions (see [EG92]) yields that (after passing to a subsequence)
\[ \chi_k \to \chi \in BV(\mathbb{R}^3) \quad \text{in } L^1(\mathbb{R}^3) \quad \text{and pointwise a.e.} \]
Since the functions \( \chi_k \) are characteristic functions, we may assume without loss of generality that \( \chi \) is the characteristic function of a set \( \Omega \subset \mathbb{R}^3 \) with
\[ L^3(\Omega) = \frac{\sigma^3}{6 \sqrt{\pi}} \quad \text{(3.15)} \]
Because of the lower semicontinuity of the perimeter on open sets and the upper semicontinuity on compact sets under convergence of measures it follows that \( |D \chi| \leq \mu \) as measures. In the end we would like to have that \( \Sigma = \partial \Omega \) is smooth. Therefore it is necessary that \( |D \chi| = \mu \) as measures, which actually holds.

**Lemma 3.4** For \( \varepsilon \leq \varepsilon_0 \) we have that \( |D \chi| = \mu \).

**Proof:** Let \( \xi_0 \in \Sigma \setminus B_\varepsilon \) and \( \rho_0 \) as in (3.14). Let \( \varepsilon \leq \varepsilon_0 \) such that Lemma 3.3 holds (now for \( \xi = \xi_0 \)) and let \( \rho \leq \frac{\rho_0}{4} \). Let \( \overline{u}_k \in C^{1,1}(B_{\rho_0}(\xi_0) \cap L, L^2) \) be an extension of

![Figure 4: Diagram for the graphical decomposition in our setting.](image-url)
Let \( u_k \) to the whole disc \( B_{\rho_k}(\xi_0) \cap L \) as in Lemma 5.6, namely \( \overline{u}_k = u_k \) in \( \Omega_k \). From the \( L^\infty \)-bounds for the function \( u_k \) and since \( \lambda_k > \frac{\rho_k}{2} \), it follows that
\[
\| \overline{u}_k \|_{L^\infty(B_{\frac{\rho_k}{2}}(\xi_0) \cap L)} \leq c \varepsilon \| \rho \| + \delta_k \leq c,
\]
\[
\| D \overline{u}_k \|_{L^\infty(B_{\frac{\rho_k}{2}}(\xi_0) \cap L)} \leq c \varepsilon \| \rho \| + \delta_k \leq c.
\]
Thus it follows that the sequence \( \overline{u}_k \) is equicontinuous and uniformly bounded in \( C^1(B_{\frac{\rho_k}{2}}(\xi_0) \cap L, L^2) \) and \( W^{1,2}(B_{\frac{\rho_k}{2}}(\xi_0) \cap L, L^2) \). Therefore there exists a function \( u \in C^{0,1}(B_{\frac{\rho_k}{2}}(\xi_0) \cap L, L^2) \) such that (after passing to a subsequence)
\[
(i) \quad \overline{u}_k \to u \quad \text{in} \quad C^0(B_{\frac{\rho_k}{2}}(\xi_0) \cap L, L^2),
\]
\[
(ii) \quad \overline{u}_k \to u \quad \text{weakly in} \quad W^{1,2}(B_{\frac{\rho_k}{2}}(\xi_0) \cap L, L^2),
\]
\[
(iii) \quad \frac{1}{\rho_k} \| u \|_{L^\infty(B_{\frac{\rho_k}{2}}(\xi_0) \cap L)} + \| D u \|_{L^\infty(B_{\frac{\rho_k}{2}}(\xi_0) \cap L)} \leq c \varepsilon \| \rho \|.
\]

In the following we assume without loss of generality that \( \xi_0 = 0, L = \mathbb{R}^2 \times \{0\} \) and that the involved graph functions are real valued functions on \( B_{\frac{\rho_k}{2}}(0) \subset \mathbb{R}^2 \), and that therefore for example graph \( u = \{(x, u(x)) \mid x \in B_{\frac{\rho_k}{2}}(0) \subset \mathbb{R}^2\} \).

Let \( g \in C^1(B_{\frac{\rho_k}{2}}(0), \mathbb{R}^3) \) with \( |g| \leq 1 \). It follows from the definition of \( |D \chi|, \chi_k \to \chi \) in \( L^1(\mathbb{R}^3) \) and Lemma 3.3 that
\[
|D \chi|(B_{\frac{\rho_k}{2}}(0)) \geq \int \chi \text{ div } g = \lim_{k \to \infty} \int \chi_k \text{ div } g
\]
\[
= \lim_{k \to \infty} \left( \int_{\text{graph } u_k \cap B_{\frac{\rho_k}{2}}(0)} \langle g, \nu_k \rangle \, d\mathcal{H}^2 + \sum_n \int_{P_n^k \cap B_{\frac{\rho_k}{2}}(0)} \langle g, \nu_k \rangle \, d\mathcal{H}^2 \right),
\]
where \( \nu_k \) denotes the outer normal to \( \partial \Omega_k = \Sigma_k \). Because of the Monotonicity formula Theorem 5.3 and the diameter estimates for the sets \( P_n^k \) we can estimate the second term on the right hand side by
\[
\left| \sum_n \int_{P_n^k \cap B_{\frac{\rho_k}{2}}(0)} \langle g, \nu_k \rangle \, d\mathcal{H}^2 \right| \leq \sum_n \mathcal{H}^2(P_n^k) \leq c \sum_n \left( \text{diam } P_n^k \right)^2 \leq c \varepsilon \rho^2.
\]
Because of the diameter estimates for the discs \( d_{k,m} \) and the \( L^\infty \)-bounds for the functions \( u_k \) and \( \overline{u}_k \), the first term on the right hand side can be estimated by
\[
\left| \int_{\text{graph } u_k \cap B_{\frac{\rho_k}{2}}(0)} \langle g, \nu_k \rangle \, d\mathcal{H}^2 \right|
\]
\[
\geq \int_{\text{graph } \overline{u}_k \cap B_{\frac{\rho_k}{2}}(0)} \langle g, \overline{\nu}_k \rangle \, d\mathcal{H}^2 - \sum_m \left| \int_{\text{graph } \overline{u}_{k,m} \cap B_{\frac{\rho_k}{2}}(0)} \langle g, \overline{\nu}_k \rangle \, d\mathcal{H}^2 \right|
\]
\[
- \sum_m \left| \int_{\text{graph } u_{k,m} \cap B_{\frac{\rho_k}{2}}(0)} \langle g, \nu_k \rangle \, d\mathcal{H}^2 \right|
\]
\[
\geq \int_{\text{graph } \overline{u}_k \cap B_{\frac{\rho_k}{2}}(0)} \langle g, \overline{\nu}_k \rangle \, d\mathcal{H}^2 - c \varepsilon \rho^2,
\]
where $\mathbf{v}_k$ denotes the outer normal to graph $\mathbf{n}_k$. Now without loss of generality we may assume that (otherwise exchange $g$ by $-g$)

$$\mathbf{v}_k(x, \mathbf{n}_k(x)) = \frac{(- D \mathbf{n}_k(x), 1)}{\sqrt{1 + |D \mathbf{n}_k(x)|^2}} \quad \text{for } x \in B_{\frac{r}{4}}(0) \subset \mathbb{R}^2.$$ 

It follows that

$$\int_{\text{graph } \mathbf{n}_k \cap B_{\frac{r}{4}}(0)} \langle g, \mathbf{v}_k \rangle \, d\mathcal{H}^2$$

$$= \int_{B_{\frac{r}{4}}(0)} \langle g(x, \mathbf{n}_k(x)), (- D \mathbf{n}_k(x), 1) \rangle \chi_{B_{\frac{r}{4}}(0)}(x, \mathbf{n}_k(x)) = \chi_{B_{\frac{r}{4}}(0)}(x, \mathbf{n}_k(x)). \quad (3.16)$$

Now we have the following convergence properties:

1.) Since $g \in C^1_c(B_{\frac{r}{4}}(0), \mathbb{R}^3)$ and since $\mathbf{n}_k \to u$ uniformly, it follows that

$$g(\cdot, \mathbf{n}_k(\cdot)) \to g(\cdot, u(\cdot)) \quad \text{uniformly on } B_{\frac{r}{4}}(0) \subset \mathbb{R}^2.$$ 

2.) We have that

$$g(\cdot, \mathbf{n}_k(\cdot))\chi_{B_{\frac{r}{4}}(0)}(\cdot, \mathbf{n}_k(\cdot)) \to g(\cdot, u(\cdot))\chi_{B_{\frac{r}{4}}(0)}(\cdot, u(\cdot)) \quad \text{in } L^\infty(B_{\frac{r}{4}}(0)).$$

To see this notice that

$$\|g(\cdot, \mathbf{n}_k(\cdot))\chi_{B_{\frac{r}{4}}(0)}(\cdot, \mathbf{n}_k(\cdot)) - g(\cdot, u(\cdot))\chi_{B_{\frac{r}{4}}(0)}(\cdot, u(\cdot))\|_{L^\infty(B_{\frac{r}{4}}(0))}$$

$$\leq \|g(\cdot, \mathbf{n}_k(\cdot)) - g(\cdot, u(\cdot))\|_{L^\infty(B_{\frac{r}{4}}(0))}$$

$$+ \|g(\cdot, u(\cdot))\chi_{B_{\frac{r}{4}}(0)}(\cdot, \mathbf{n}_k(\cdot)) - g(\cdot, u(\cdot))\|_{L^\infty(B_{\frac{r}{4}}(0))}.$$ 

Now the first term goes to 0 by 1.), and also the second term goes to 0, which can be proved in the following way:

(i) Let $x \in B_{\frac{r}{4}}(0)$ such that $\chi_{B_{\frac{r}{4}}(0)}(x, u(x)) = 1$. Thus $(x, u(x)) \in B_{\frac{r}{4}}(0)$.

Since $B_{\frac{r}{4}}(0)$ is open, we get that $(x, \mathbf{n}_k(x)) \in B_{\frac{r}{4}}(0)$ for $k$ sufficiently large, and therefore $\chi_{B_{\frac{r}{4}}(0)}(x, \mathbf{n}_k(x)) = 1$.

(ii) Let $x \in B_{\frac{r}{4}}(0)$ such that $\chi_{B_{\frac{r}{4}}(0)}(x, u(x)) = 0$, namely $(x, u(x)) \notin B_{\frac{r}{4}}(0)$.

Since $\text{spt } g \subset \subset B_{\frac{r}{4}}(0)$, it follows that $g(x, u(x)) = 0$.

Rewriting the right hand side in (3.16) yields

$$\int_{B_{\frac{r}{4}}(0)} \langle g(x, \mathbf{n}_k(x)), (- D \mathbf{n}_k(x), 1) \rangle \chi_{B_{\frac{r}{4}}(0)}(x, \mathbf{n}_k(x))$$

$$= \int_{B_{\frac{r}{4}}(0)} \langle g(x, \mathbf{n}_k(x))\chi_{B_{\frac{r}{4}}(0)}(x, \mathbf{n}_k(x)) - g(x, u(x))\chi_{B_{\frac{r}{4}}(0)}(x, u(x)), (- D \mathbf{n}_k(x), 1) \rangle$$

$$+ \int_{B_{\frac{r}{4}}(0)} \langle g(x, u(x))\chi_{B_{\frac{r}{4}}(0)}(x, u(x)), (- D \mathbf{n}_k(x), 1) - (- D u(x), 1) \rangle$$

$$+ \int_{B_{\frac{r}{4}}(0)} \langle g(x, u(x)), (- D u(x), 1) \rangle \chi_{B_{\frac{r}{4}}(0)}(x, u(x)).$$
The first term goes to 0 by 2.) above and because of the $L^\infty$-estimates for $D\overline{\eta}_k$. The second term goes to 0 since $g(\cdot, u(\cdot))\eta_{B_{\theta 0}^e(0)}(\cdot, u(\cdot)) \in L^\infty(B_{\theta 0}^e(0))$ and since $\overline{\eta}_k \rightharpoonup u$ weakly in $W^{1,2}(B_{\theta 0}^e(0))$.

Therefore we finally get that

$$|D\chi(B_{\theta 0}^e(0)) | \geq \int_{B_{\theta 0}^e(0)} \left( g(x, u(x)), (-D u(x), 1) \right) \chi_{B_{\theta 0}^e(0)}(x, u(x)) - c\epsilon \rho^2.$$ 

Now let $\eta \in C^1_c(B_{\theta 0}^e(0))$ with $|\eta| \leq 1$, and let $g = \eta \epsilon_3$. Computing the right hand side of the above inequality gives

$$|D\chi(B_{\theta 0}^e(0)) | \geq \int_{B_{\theta 0}^e(0)} \eta(x, u(x)) \chi_{B_{\theta 0}^e(0)}(x, u(x)) - c\epsilon \rho^2,$$

and letting $\eta \nearrow \chi_{B_{\theta 0}^e(0)}$ (and going back to $\xi_0$ and the plane $L$) we get that

$$|D\chi(B_{\theta 0}^e(\xi_0)) | \geq \int_{B_{\theta 0}^e(\xi_0) \cap L} \chi_{B_{\theta 0}^e(\xi_0)}(x + u(x)) - c\epsilon \rho^2.$$ 

From the $L^\infty$-bound on $u$ we get that $|\xi_0 - (x + u(x))| < \theta_0^e$ if $x \in B_{\left(\frac{1}{2}\epsilon - c\epsilon^\frac{1}{2}\right)} (\xi_0) \cap L$, and it follows that

$$\int_{B_{\theta 0}^e(\xi_0) \cap L} \chi_{B_{\theta 0}^e(\xi_0)}(x + u(x)) \geq L^2 \left( B_{\left(\frac{1}{2}\epsilon - c\epsilon^\frac{1}{2}\right)} (\xi_0) \cap L \right) = \left( \frac{\theta_0}{\epsilon} - c\epsilon \right)^2 \pi \rho^2.$$

Therefore we have finally shown that for $\xi_0 \in \Sigma \setminus B_\epsilon$ and $\rho \leq \Delta_\epsilon^e$

$$|D\chi(B_{\theta 0}^e(\xi_0)) | \geq \left( \frac{\theta_0}{\epsilon} - c\epsilon \right)^2 \pi \rho^2 - c\epsilon \rho^2. \quad (3.17)$$

In the next step we will estimate the quantity $\mu(B_{\theta 0}^e(\xi_0))$. For that let $f \in C^0_c(\mathbb{R}^3)$ such that $f \leq \chi_{B_{\theta 0}^e(\xi_0)}$. Since $\mu_k \rightharpoonup \mu$ in the weak sense we get

$$\int f \, d\mu = \lim_{k \to \infty} \int f \, d\mu_k \leq \limsup_{k \to \infty} \mu_k(B_{\theta 0}^e(\xi_0)).$$

From Lemma 3.3 it follows that

$$\mu_k(B_{\theta 0}^e(\xi_0)) = \mathcal{H}^2 \left( \text{graph } u_k \cap B_{\theta 0}^e(\xi_0) \right) + \sum_n \mathcal{H}^2 \left( P_n \cap B_{\theta 0}^e(\xi_0) \right).$$

The second term on the right hand side can be estimated as before by

$$\sum_n \mathcal{H}^2 \left( P_n \cap B_{\theta 0}^e(\xi_0) \right) \leq c\epsilon \rho^2.$$ 

The first term on right hand side can be estimated by

$$\mathcal{H}^2 \left( \text{graph } u_k \cap B_{\theta 0}^e(\xi_0) \right) \leq \int_{\Omega_h} \chi_{B_{\theta 0}^e(\xi_0)}(x + u_k(x)) \sqrt{1 + |D u_k(x)|^2} \leq \sqrt{1 + (c\epsilon^\frac{1}{2} + \delta_k)^2} \int_{\Omega_h} \chi_{B_{\theta 0}^e(\xi_0)}(x + u_k(x)).$$
From the $L^\infty$-estimates for the function $u_k$, it follows that if $x \notin B_{\frac{c_2}{c_1} \rho + \delta_k}(\xi_0) \cap L$, then $\chi_{B_{\frac{c_2}{c_1} \rho + \delta_k}(\xi_0)}(x + u_k(x)) = 0$, and therefore we get
\[
\mathcal{H}^2(\text{graph } u_k \cap B_{\frac{c_2}{c_1} \rho + \delta_k}(\xi_0)) \leq \pi \sqrt{1 + \left(\frac{\theta}{8} + c \varepsilon\right)^2} \left(\frac{\theta}{8} + c \varepsilon\right)^2.
\]
Since $\delta_k \to 0$ for $k \to \infty$, we get that
\[
\int f \, d\mu \leq \sqrt{1 + c \varepsilon} \left(\frac{\theta}{8} + c \varepsilon\right)^2 \pi \rho^2 + c \varepsilon \rho^2,
\]
and letting $f / \chi_{B_{\frac{c_2}{c_1} \rho + \delta_k}(\xi_0)}$ we conclude that
\[
\mu(B_{\frac{c_2}{c_1} \rho + \delta_k}(\xi_0)) \leq \sqrt{1 + c \varepsilon} \left(\frac{\theta}{8} + c \varepsilon\right)^2 \pi \rho^2 + c \varepsilon \rho^2.
\]
(3.18)
Since the derivative $D_\mu |D\chi(\xi_0)$ exists for $\mu$-a.e. $\xi_0 \in \Sigma$ (see [EG92]), we get for $\mu$-a.e. $\xi_0 \in \Sigma \setminus B_\varepsilon$ that
\[
D_\mu |D\chi(\xi_0) \geq \frac{(1 - c \varepsilon)^2}{\sqrt{1 + c \varepsilon}(1 + c \varepsilon)^2} \pi - c \varepsilon.
\]
Letting $\varepsilon = \frac{1}{n} \to 0$ and remembering that $|D\chi| \leq \mu$, we get
\[
D_\mu |D\chi(\xi_0) = 1 \quad \text{for } \mu\text{-a.e. } \xi_0 \in \Sigma \setminus \bigcup_n B_{\frac{1}{n}}.
\]
Since each set $B_{\frac{1}{n}}$ contains only finitely many points, and since the Monotonicity formula Theorem 5.3 yields $\mu(\{|\xi|\}) = 0$ for every $\xi \in \mathbb{R}^3$, we get that
\[
D_\mu |D\chi(\xi_0) = 1 \quad \text{for } \mu\text{-a.e. } \xi_0 \in \mathbb{R}^3.
\]
The Lemma follows from the Theorem of Radon-Nikodym. □

**Remark 3.5** Notice that the only thing we needed up to now was the bound on the Willmore energy $\mathcal{W}(\Sigma_k) \leq 8\pi - \delta_0$. We are now able to prove that
\[
\lim_{\sigma \searrow 0} \beta(\sigma) = 8\pi.
\]

**Proof:** We already know that $\beta$ is decreasing and bounded by $8\pi$. Therefore the limit exists. Let $\sigma_l \to 0$ and assume (3.19) is false. After passing to a subsequence there exists a $\delta_0 > 0$ such that $\beta(\sigma_l) \leq 8\pi - \delta_0$ for all $l \in \mathbb{N}$. Let $\Sigma_l \in M_{\varepsilon_l}$ such that $\mathcal{W}(\Sigma_l) \leq \beta(\sigma_l) + \frac{\delta_0}{2} \leq 8\pi - \frac{\delta_0}{2}$, and let $\Omega_l \subset \mathbb{R}^3$ be the open set surrounded by $\Sigma_l$. Again after scaling and translation we may assume that $\mathcal{H}^2(\Sigma_l) = 1$ and $0 \in \Sigma_l$, and that the Radon measures $\mu_l = \mathcal{H}^2(\Sigma_l)$ converge to a Radon measure $\mu$ with $\mu(\mathbb{R}^3) = 1$. On the other hand we have that the BV-functions $\chi_l = \chi_{\Omega_l}$ are uniformly bounded, and therefore converge (after passing to a subsequence) in $L^1$ to a BV-function $\chi$. Since $I(\Sigma_l) \to 0$ and $\mathcal{H}^2(\Sigma_l) = 1$, it follows that $\chi = 0$. Finally, since $\mathcal{W}(\Sigma_l) \leq 8\pi - \frac{\delta_0}{2}$, we can do exactly the same as before to get $\mu = |D\chi|$, which contradicts $\mu(\mathbb{R}^3) = 1$. Therefore (3.19) holds. □
We continue with the proof of Theorem 1.1. The main idea to prove regularity is to derive a power decay for the $L^2$-norm of the second fundamental form via constructing comparison surfaces by a cut-and-paste procedure as done in [Sim93]. But this method cannot be directly applied in our case, since the isoperimetric ratio might change by this procedure. In order to correct the isoperimetric ratio of the generated surfaces, we will apply an appropriate variation. But what is an appropriate variation in our case? Which is the quantity we have to look at? To answer this question let $\Phi : (-\varepsilon, \varepsilon) \times \mathbb{R}^3 \to \mathbb{R}^3$ be a $C^2$-variation with compact support and define $\Omega_{k,t} = \Phi_t(\Omega_k)$, $\Sigma_{k,t} = \partial \Omega_{k,t} = \Phi_t(\Sigma_k)$ and $X(t,x) = \partial_t \Phi_t(x)_{|_{t=0}$. It follows that

$$
\frac{d}{dt} I(\Sigma_{k,t})_{|_{t=0}} = \frac{I(\Sigma_k)}{3 \mathcal{H}^2(\Sigma_k)} \frac{3}{2} \int (X, \nabla H) \, d\mu_k + \frac{\mathcal{H}^2(\Sigma_k)}{2^3(\Omega_k)} \int \chi_{\Omega_k} \, \text{div} X.
$$

(3.20)

$$
= \frac{I(\Sigma_k)}{3 \mathcal{H}^2(\Sigma_k)} \int \left( X, \frac{3}{2} \nabla H - \frac{\mathcal{H}^2(\Sigma_k)}{2^3(\Omega_k)} \nu_k \right) \, d\mu_k,
$$

(3.21)

where $\nu_k$ denotes the inner normal to $\partial \Omega_k = \Sigma_k$.

It follows from (3.5), $\chi_{\Omega_k} \to \chi_\Omega$ in $L^1(\mathbb{R}^3)$, Lemma 3.4 and (3.20) that

$$
\lim_{k \to \infty} \frac{d}{dt} I(\Sigma_{k,t})_{|_{t=0}} = \frac{\sigma}{3} \int \left( X, \frac{3}{2} \nabla H - \frac{6 \sqrt{\pi}}{\sigma^3} \nu \right) \, d\mu,
$$

(3.22)

where $\nu$ is given by the equation

$$
\int \chi \, \text{div} g = - \int \langle g, \nu \rangle \, d|\nabla \chi| = - \int \langle g, \nu \rangle \, d\mu
$$

for $g \in C^1_c(\mathbb{R}^3, \mathbb{R}^3)$. The first equality follows from the Riesz Representation Theorem applied to BV-functions, and the second equality is due to Lemma 3.4.

Now if there would exist a vector field $X \in C^\infty_c(\mathbb{R}^3, \mathbb{R}^3)$ such that the right hand side of (3.22) is not equal to 0, the first variation of the isoperimetric ratio of $\Sigma_k$ would not be equal to 0 for $k$ sufficiently large, and in conclusion we would have a chance to correct the isoperimetric ratio of the generated surfaces. The next Lemma is concerned with the existence of such a vector field and relies on the fact that each surface $\Sigma \in \mathcal{M}_\sigma$ is not a round sphere.

**Lemma 3.6** There exists a vectorfield $X \in C^\infty_c(\mathbb{R}^3, \mathbb{R}^3)$ such that

$$
\int \left( X, \frac{3}{2} \nabla H - \frac{6 \sqrt{\pi}}{\sigma^3} \nu \right) \, d\mu \neq 0.
$$

**Proof:** Assume the statement is false. It follows that

$$
\hat{H}(x) = \frac{4 \sqrt{\pi}}{\sigma^3} \nu(x) \quad \text{for } \mu\text{-a.e. } x \in \Sigma.
$$

(3.23)

Now the idea of the proof is the following: We just have to show that $\Sigma$ is smooth, because then $\Sigma$ would be a smooth surface with constant mean curvature and Willmore energy smaller than $8\pi$. It follows from a Theorem of Alexandroff that $\Sigma$ is a round sphere, which contradicts our choice of $\sigma \in (0, 1)$.

To show that $\Sigma$ is smooth we just have to show that $\theta^2(\mu, x) = 1$ for every $x \in \Sigma$, because then Allard’s Regularity Theorem would yield (remember that $\hat{H} \in L^\infty(\mu)$).
now) that $\Sigma$ can be written as a $C^{1,\alpha}$-graph in a neighborhood of every point $x \in \Sigma$ that solves the constant mean curvature equation, and is therefore smooth.

Let $x_0 \in \Sigma$. Notice that, since $\chi \in BV(\mathbb{R}^3)$, $\mu$ generates an integer multiplicity, rectifiable 2-current $\mathcal{M}_\mu \in \mathcal{D}^2(\mathbb{R}^3)'$ with $\partial \mathcal{M}_\mu = 0$. Denote by $\mu_{x_0, \lambda}$ the blow-ups of $\mu$ around $x_0$. Now also the blow-ups generate integer multiplicity, rectifiable 2-currents $\mathcal{M}_{\mu_{x_0, \lambda}}$ with $\partial \mathcal{M}_{\mu_{x_0, \lambda}} = 0$. Moreover the mass of $\mathcal{M}_{\mu_{x_0, \lambda}}$ of a set $W \subset \mathbb{R}^3$ such that $W \subset B_R(0)$ is estimated in view of the Monotonicity formula by

$$\mathcal{M}_{\mu_{x_0, \lambda}}(B_R(0)) \leq \mathcal{M}_{\mu_{x_0, \lambda}}(B_{R_0}(0)) = \lambda^{-2} \mu(B_{R_0}(0)) \leq cR^2.$$  

A compactness result for integer multiplicity, rectifiable 2-currents (see [Sim83]) yields that there exists an integer multiplicity, rectifiable 2-current $\mathcal{M}_{\mu} \in \mathcal{D}^2(\mathbb{R}^3)'$ such that $\partial \mathcal{M}_{\mu} = 0$ and (after passing to a subsequence)

$$\mathcal{M}_{\mu_{x_0, \lambda}} \rightarrow \mathcal{M}_{\mu} \quad \text{for} \lambda \to 0 \text{ weakly as currents.}$$

Let $\mu_{x_0}$ be the underlying varifold.

On the other hand there exists a stationary, integer multiplicity, rectifiable 2-cone $\mu_\infty$ such that (after passing to a subsequence)

$$\mu_{x_0, \lambda} \rightarrow \mu_\infty \quad \text{for} \lambda \to 0 \text{ weakly as varifolds.}$$

Now we get the following:

1.) $\mu_{x_0} \leq \mu_\infty$: This follows from the lower semicontinuity of the mass with respect to weak convergence of currents and the upper semicontinuity on compact sets with respect to weak convergence of measures.

2.) $\theta^2(\mu_\infty, \cdot) \leq 2 - \frac{\theta_0}{\pi} \text{ everywhere:}$ Since $\mu_\infty$ is a stationary 2-cone, the Monotonicity formula yields for all $Z \in \mathbb{R}^3$ and all $B_r(0)$ such that $\mu_\infty(\partial B_r(0)) = 0$

$$\theta^2(\mu_\infty, Z) \leq \theta^2(\mu_\infty, 0) = \frac{\mu_\infty(B_r(0))}{\pi r^2} = \liminf_{\lambda \to 0} \frac{\mu_{x_0, \lambda}(B_r(0))}{\pi r^2}.$$

Now since $\frac{\mu_{x_0, \lambda}(B_r(0))}{\pi r^2} = \frac{\mu(B_r(z))}{r^2}$, it follows that $\theta^2(\mu_\infty, Z) \leq \theta^2(\mu, x_0)$, and the claim follows from (3.7).

3.) $\theta^2(\mu_\infty, \cdot) = 1 \text{ } \mu_\infty$-a.e.: This follows from 2.) since $\mu_\infty$ is integral.

4.) $\mu_{x_0} = \mu_\infty$: Choose a point $x \in \mathbb{R}^3$ such that $\theta^2(\mu_\infty, x) = 1$. By Allard’s Regularity Theorem there exists a neighborhood $U(x)$ of $x$ in which $\mu_\infty$ can be written as a $C^{1,\alpha}$-graph, which is actually smooth since $\mu_\infty$ is stationary.

Moreover we get that the convergence $\mu_{x_0, \lambda} U(x) \rightarrow \mu_\infty U(x)$ is in $C^{1,\alpha}$. Thus $\mu_{x_0, \lambda} U(x) \rightarrow \mu_\infty U(x)$ weakly. Hence for $U = \bigcup_{(x, r) \in 1} U(x)$ we have that $\mu_{x_0, \lambda} U = \mu_{x_0, \lambda} U$. Since we already know that $\theta^2(\mu_\infty, x) = 1$ for $\mu_\infty$-a.e. $x \in \mathbb{R}^3$, we get 4.)

From 4.) it follows that $\mathcal{M}_{\mu_{x_0}}$ is a stationary, integer multiplicity, rectifiable 2-current with $\partial \mathcal{M}_{\mu_{x_0}} = 0$. Since moreover $\mu_{x_0} = \mu_\infty$ is a stationary, rectifiable 2-cone, we get for all $\tau > 0$ and all $Z \in \mathbb{R}^3$ that

$$\frac{\mu_{x_0}(B_{\tau}(Z))}{\pi \tau^2} \leq \theta^2(\mu_{x_0}, 0) \leq 2 - \frac{\delta_0}{4\pi}.$$
Letting $\tau \to \infty$, we get that $\theta^2(\mu_{\tau_0}, \infty) \leq 2 - \frac{\delta_0}{4\tau}$. Using Theorem 2.1 in [KLS10], it follows that $M_{\tau_0}$ is a unit density plane, or equivalently

$$\mu_{\tau_0} = \mu_{\infty} = \mathcal{H}^2 \cdot P$$

for some $P \in G_2(\mathbb{R}^3)$.

Therefore we get for all balls $B_r(0)$ such that $\mu_{\infty}(\partial B_r(0)) = 0$

$$\theta^2(\mu, x_0) = \lim_{\lambda \to 0} \frac{\mu(B_{\lambda r}(x_0))}{\pi(\lambda r)^2} = \lim_{\lambda \to 0} \frac{\mu_{\tau_0}(B_{\lambda r}(0))}{\pi \tau^2} = \frac{\mu_{\infty}(B_{\lambda r}(0))}{\pi \tau^2} = \frac{\mathcal{H}^2 \cdot P(B_{\lambda r}(0))}{\pi \tau^2} = 1,$$

and the Lemma is proved. \(\square\)

In the next step we prove a power decay for the $L^2$-norm of the second fundamental form on small balls around the good points. This will help us to show that $\Sigma$ is actually $C^{1,\alpha} \cap W^{2,2}$ away from the bad points.

**Lemma 3.7** Let $\xi_0 \in \Sigma \setminus \mathcal{B}_\varepsilon$. If $\varepsilon \leq \varepsilon_0$, there exists a $\rho_0 = \rho_0(\xi_0, \varepsilon) > 0$ such that for all $\xi \in \Sigma \cap B_{\rho_0}(\xi_0)$ and all $\rho \leq \frac{\varepsilon_0}{4}$ we have

$$\liminf_{k \to \infty} \int_{\Sigma \cap B_{\rho_0}(\xi)} |\alpha_k| \, d\mathcal{H}^2 \leq c \rho^\alpha,$$

where $\alpha \in (0, 1)$ and $c < \infty$ are universal constants.

**Proof:** Let $\xi_0 \in \Sigma \setminus \mathcal{B}_\varepsilon$. First of all we need to localize the vectorfield of Lemma 3.6, or more precisely we want the vectorfield to vanish in a neighborhood around $\xi_0$. Therefore let $r > 0$ and $\eta_\tau \in C^\infty_c(\mathbb{R}^3)$ such that $0 \leq \eta_\tau \leq 1$ and $\eta_\tau \equiv 1$ on $B_r(\xi_0)$. Now let $X \in C^\infty_c(\mathbb{R}^3, \mathbb{R}^3)$ be the vectorfield of Lemma 3.6 and define the vectorfield $X_\tau \in C^\infty_c(\mathbb{R}^3, \mathbb{R}^3)$ by $X_\tau = (1 - \eta_\tau)X$. It follows that

$$\lim_{r \to 0} \int_{B_r(\xi_0)} \left( X, \frac{3}{2} \mathcal{H} - \frac{6}{\sigma^3} \nu \right) \, d\mu = \int \left( X, \frac{3}{2} \mathcal{H} - \frac{6}{\sigma^3} \nu \right) \, d\mu \neq 0.$$}

Therefore for $r > 0$ sufficiently small (depending only on $\xi_0 \in \Sigma \setminus \mathcal{B}_\varepsilon$) we may exchange the vectorfield $X$ by the vectorfield $X_\tau$, which vanishes in $B_r(\xi_0)$.

Now fix $r = r(\xi_0) > 0$ such that this holds. Therefore there exists a vectorfield $X \in C^\infty_c(\mathbb{R}^3, \mathbb{R}^3)$ such that

$$\int \left( X, \frac{3}{2} \mathcal{H} - \frac{6}{\sigma^3} \nu \right) \, d\mu \neq 0 \quad \text{and} \quad X \equiv 0 \text{ on } B_r(\xi_0). \quad (3.24)$$

Now let $\rho_0 > 0$ as in (3.14). We may assume without loss of generality that

$$\rho_0 < \frac{r}{2}. \quad (3.25)$$

Let $\xi \in \Sigma \cap B_{\rho_0}(\xi_0)$ and $\rho \leq \frac{\rho_0}{8}$. Notice that Lemma 3.3 holds. For $r \in \left( \frac{\rho_0}{16}, \frac{3\rho_0}{4} \right)$ define the set

$$C_r(\xi) = \left\{ x + y \mid x \in B_r(\xi) \cap L, y \in L^\perp \right\}.$$}

From the $L^\infty$-estimates for the functions $u_k$ and the diameter estimates for the sets $P_k^\perp$ it follows for $\varepsilon \leq \varepsilon_0$ and $\delta_k \leq \frac{1}{8} \rho_k^{\frac{1}{8}}$ that $D_k \cap C_r(\xi) = D_k \cap C_r(\xi) \cap B_{\rho_k}(\xi)$. Therefore

$$\Sigma_k \setminus \left( D_k \cap C_r(\xi) \right) = \Sigma_k \setminus \left( C_r(\xi) \cap B_{\rho_k}(\xi) \right) \quad \text{for} \ \varepsilon \leq \varepsilon_0 \ \text{and} \ \delta_k \leq \frac{1}{8} \rho_k^{\frac{1}{8}}.$$
Define the sets
\[ S_k(\xi) = \left\{ \tau \in \left( \frac{\theta L}{16}, \frac{3\theta L}{4} \right) \left| \partial C_\tau(\xi) \cap \bigcup_m d_{k,m} = \emptyset \right. \right\}. \]
\[ T_k(\xi) = \left\{ \tau \in S_k(\xi) \left| \int_{D_\tau \cap \partial C_\tau(\xi)} |A_k|^2 \, d\mathcal{H}^2 \leq \frac{128}{\theta \rho} \int_{D_\tau \cap C_\tau(\xi) \cup \partial C(\xi)} |A_k|^2 \, d\mathcal{H}^2 \right. \right\}. \]

The diameter estimates for the discs \( d_{k,m} \) yield \( L^1(S_k(\xi)) \geq \theta \frac{L}{64} \) for \( \varepsilon \leq \varepsilon_0 \), and then from a Fubini-type argument we get \( L^1(T_k(\xi)) \geq \theta \frac{L}{128} \). From the selection principle in [Sim93], Lemma B.1, it follows that there exists a \( \tau \in \left( \frac{\theta L}{16}, \frac{3\theta L}{4} \right) \) such that \( \tau \in T_k(\xi) \) for infinitely many \( k \).

Apply Lemma 5.6 to get a function \( w_k \in C^\infty(B_{T_k(\xi)} \cap L, L^2) \) for infinitely many \( k \) such that

(i) \( w_k = u_k \) and \( \frac{\partial w_k}{\partial \nu} = \frac{\partial u_k}{\partial \nu} \) on \( \partial B_{T_k(\xi)} \cap L \),

(ii) \( \frac{1}{\tau} ||w_k||_{L^\infty(B_{\tau}(\xi)) \cap L} \leq c \varepsilon^\frac{1}{2} + \frac{\delta_k}{\tau} \), where \( \delta_k \to 0 \),

(iii) \( \| D w_k \|_{L^\infty(B_{\tau}(\xi)) \cap L} \leq c \varepsilon^\frac{1}{2} + \delta_k \), where \( \delta_k \to 0 \),

(iv) \( \int_{B_{\tau}(\xi) \cap L} |D^2 w_k|^2 \leq c \tau \int_{\text{graph } w_k \cap B_{\tau}(\xi) \cap L} |A_k|^2 \, d\mathcal{H}^1 \).

Since graph \( w_k \subset \overline{B_{\tau}(\xi)} \) for \( \varepsilon \leq \varepsilon_0 \) and \( \delta_k \leq \frac{1}{8} \theta \frac{L}{8} \), we get from the above that

\[ \text{graph } w_k \cap \left( \Sigma_k \setminus \left( D_k \cap C(\xi) \right) \right) \subset C_\tau(\xi) \cap \overline{B_{\theta \varepsilon}(\xi)} \cap \left( \Sigma_k \setminus \left( C_\tau(\xi) \cap \overline{B_{\theta \varepsilon}(\xi)} \right) \right) = \emptyset. \]

Now define the surfaces
\[ \Sigma_k = \Sigma_k \setminus \left( D_k \cap C(\xi) \right) \cup \text{graph } w_k. \]

From the above it follows for \( \varepsilon \leq \varepsilon_0 \) and \( \delta_k \leq \frac{1}{8} \theta \frac{L}{8} \) that \( \hat{\Sigma}_k \) is a compact, embedded and connected \( C^{1,1} \)-sphere. In addition \( \hat{\Sigma}_k \) surrounds an open set \( \Omega_k \) and we have that \( \hat{\Sigma}_k \cap \mathbb{R}^3 \setminus B_{\varepsilon_0}(\xi_0) = \Sigma_k \cap \mathbb{R}^3 \setminus B_{\varepsilon_0}(\xi_0) \).

Next we compare the isoperimetric coefficients of \( \Sigma_k \) and \( \hat{\Sigma}_k \). Using the \( L^\infty \)-bounds for \( w_k \) and the Monotonicity formula we get from the definition of \( \hat{\Sigma}_k \) that

\[ |\mathcal{H}^2(\hat{\Sigma}_k) - \mathcal{H}^2(\Sigma_k)| \leq \mathcal{H}^2(\Sigma_k \cap \overline{B_{\theta \varepsilon}(\xi)}) + \mathcal{H}^2(\text{graph } w_k) \leq c \rho^2. \]  

Since \( L^3(\hat{\Omega}_k) = L^3(\Omega_k) \leq L^3(\Omega \Delta \hat{\Omega}) \) and since by construction \( \Omega \Delta \hat{\Omega} \subset \overline{B_{\theta \varepsilon}(\xi)} \), it follows that

\[ |L^3(\hat{\Omega}_k) - L^3(\Omega_k)| \leq c \rho^3. \]

Since \( \mathcal{H}^2(\Sigma_k) = 1 \) and \( L^3(\Omega_k) = \frac{c^3}{12 \sqrt{\pi}} \), we get by choosing \( \rho_0 \) smaller if necessary (smaller in an universal way) that

\[ \frac{1}{2} \leq \mathcal{H}^2(\hat{\Sigma}_k) \leq 2 \quad \text{and} \quad \frac{c^3}{12 \sqrt{\pi}} \leq L^3(\hat{\Omega}_k) \leq \frac{c^3}{3 \sqrt{\pi}}. \]
Moreover we finally get
\[ |H(\tilde{\Sigma}_k) - \sigma| = |I(\tilde{\Sigma}_k) - I(\Sigma_k)| \leq c\rho, \]
and we may assume without loss of generality that
\[ \frac{\sigma}{2} \leq I(\tilde{\Sigma}_k) \leq 2\sigma. \] (3.30)

As mentioned before, \( \tilde{\Sigma}_k \) might not have the right isoperimetric ratio and may therefore not be a comparison surface.

Let \( X \in C^\infty_c(\mathbb{R}^3, \mathbb{R}^3) \) such that (3.24) holds. We may assume that (otherwise exchange \( X \) by \(-X)\)
\[ \int \left( X \cdot \frac{3}{2} H - \frac{6}{\sigma^3} \nu \right) d\mu = c_0 > 0. \] (3.31)

Notice that the constant \( c_0 \) does not depend on \( \epsilon, \xi, \rho \) or \( k \).

Let \( \Phi \in C^\infty(\mathbb{R} \times \mathbb{R}^3, \mathbb{R}^3) \) be the flow of the vectorfield \( X \), namely
\[ \Phi_t(\cdot) = \Phi(t, \cdot) \in C^\infty(\mathbb{R}^3, \mathbb{R}^3) \] is a diffeomorphism for all \( t \in \mathbb{R} \),
\[ \Phi(0, z) = z \] for all \( z \in \mathbb{R}^3 \),
\[ \partial_t \Phi(t, z) = X(\Phi(t, z)) \] for all \( (t, z) \in \mathbb{R} \times \mathbb{R}^3 \).

Since \( X = 0 \) on \( B_r(\xi_0) \), there exists a \( T_0 = T_0(X) > 0 \) such that for all \( t \in (-T_0, T_0) \)
\[ \Phi_t = Id \] in \( B_2(\xi_0) \).

Define the sets
\[ \tilde{\Omega}_k = \Phi_t(\tilde{\Omega}_k) \quad \text{and} \quad \tilde{\Sigma}_k = \partial \tilde{\Omega}_k = \Phi_t(\tilde{\Sigma}_k). \] (3.32)

Choosing \( T_0 \) smaller if necessary (depending on \( X \)) it follows for \( t \in (-T_0, T_0) \) that
\[ \mathcal{H}^2(\tilde{\Sigma}_k) = \int_{\tilde{\Sigma}_k} J_{\xi_0} \Phi_t d\mathcal{H}^2 \quad \text{and} \quad \mathcal{L}^3(\tilde{\Omega}_k) = \int_{\tilde{\Omega}_k} \det D \Phi_t. \]

By choosing \( T_0 \) smaller if necessary (depending on \( X \)) and estimating very roughly, we get that there exists a constant \( 0 < c = c(X) < \infty \) such that for all \( t \in (-T_0, T_0) \)
\[ (i) \quad \frac{1}{c} \mathcal{H}^2(\tilde{\Sigma}_k) \leq \sup_{t \in (-T_0, T_0)} \mathcal{H}^2(\tilde{\Sigma}_k) \leq c \mathcal{H}^2(\tilde{\Sigma}_k), \]
\[ (ii) \quad \frac{1}{c} \mathcal{L}^3(\tilde{\Omega}_k) \leq \sup_{t \in (-T_0, T_0)} \mathcal{L}^3(\tilde{\Omega}_k) \leq c \mathcal{L}^3(\tilde{\Omega}_k), \]
\[ (iii) \quad \sup_{t \in (-T_0, T_0)} \left| \frac{d}{dt} \mathcal{H}^2(\tilde{\Sigma}_k) \right| + \sup_{t \in (-T_0, T_0)} \left| \frac{d}{dt} \mathcal{L}^3(\tilde{\Omega}_k) \right| \leq c, \]
\[ (iv) \quad \sup_{t \in (-T_0, T_0)} \left| \frac{d^2}{dt^2} \mathcal{H}^2(\tilde{\Sigma}_k) \right| + \sup_{t \in (-T_0, T_0)} \left| \frac{d^2}{dt^2} \mathcal{L}^3(\tilde{\Omega}_k) \right| \leq c, \]
\[ (v) \quad \sup_{t \in (-T_0, T_0)} \left| \frac{d}{dt} \int_{\tilde{\Sigma}_k} |A_k|^2 d\mathcal{H}^2 \right| \leq c. \]

The last inequality can be proved by writing \( \tilde{\Sigma}_k \) locally as a graph with small Lipschitz norm and using a partition of unity.
Now first of all it follows for the first variation of the isoperimetric coefficient of $\tilde{\Sigma}_k$, using (3.21) for $\tilde{\Sigma}_k$, that
\[
\frac{d}{dt} I(\tilde{\Sigma}_k)_{t=0} = I(\tilde{\Sigma}_k) \int_{\tilde{\Sigma}_k} \left( X, \frac{3}{2} \tilde{H}_k - \frac{\sqrt{\pi}}{\sigma^3} v_k \right) d\mathcal{H}^2.
\]

Now it follows from (3.26)-(3.28) that
\[
\left| \int_{\Sigma_k} \left( X, \frac{\sqrt{\pi}}{\sigma^3} v_k \right) d\mathcal{H}^2 \right| \leq c \left( \mathcal{H}^2(\Sigma_k) \right)^3 \mathcal{H}^2(\Sigma_k) \leq c\rho,
\]
where $c = c(X)$, and therefore (3.28) and (3.30) yield
\[
\frac{d}{dt} I(\tilde{\Sigma}_k)_{t=0} \geq I(\tilde{\Sigma}_k) \int_{\Sigma_k} \left( X, \frac{3}{2} \tilde{H}_k - \frac{6 \sqrt{\pi}}{\sigma^3} v_k \right) d\mathcal{H}^2 - c\rho.
\]

Since \( \int_{\Sigma_k} \left( X, \frac{3}{2} \tilde{H}_k - \frac{6 \sqrt{\pi}}{\sigma^3} v_k \right) d\mathcal{H}^2 \rightarrow \int \left( X, \frac{3}{2} \tilde{H} - \frac{6 \sqrt{\pi}}{\sigma^3} v \right) d\mu = c_0 > 0 \) by (3.22), it follows from (3.28) and (3.30) that there exists a constant $0 < c_0 < \infty$ independent of $e, \xi, \rho$ and $k$, such that for $k$ sufficiently large
\[
\frac{d}{dt} I(\tilde{\Sigma}_k)_{t=0} \geq c_0 - c\rho.
\]  

Moreover using the estimates (iii) and (iv) above it follows that
\[
\sup_{t \in (-T_0, T_0)} \left| \frac{d^2}{dt^2} I(\tilde{\Sigma}_k) \right| \leq c,
\]  

where $c = c(X) < \infty$ is a universal constant.

Using Taylor’s formula we get in view of (3.29) that for each $k$ there exists a $t_k$ with $|t_k| \leq c\rho$, such that
\[
I(\tilde{\Sigma}_k) = \sigma.
\]

Therefore we get by construction that $\tilde{\Sigma}_k \in M_\sigma$ is a comparison surface to $\Sigma_k$. Moreover it follows from (v) above that
\[
\left| \int_{\Sigma_k} |A_k|^2 d\mathcal{H}^2 - \int_{\Sigma_k} |A_k|^2 d\mathcal{H}^2 \right| \leq |t_k| \sup_{t \in [-t, t]} \left| \frac{d}{dt} \int_{\Sigma_k} |A_k|^2 d\mathcal{H}^2 \right| \leq c\rho.
\]

Since $\Sigma_k$ is a minimizing sequence for the Willmore functional in $M_\sigma$, and by the Gauss-Bonnet Theorem therefore a minimizing sequence for the functional $\int_\Sigma |A|^2$, we get
\[
\int_{\Sigma_k} |A_k|^2 d\mathcal{H}^2 \leq \int_{\Sigma_k} |A_k|^2 d\mathcal{H}^2 + c\rho + \epsilon_k, \quad \text{where} \ \epsilon_k \rightarrow 0.
\]

Now by definition of $\tilde{\Sigma}_k$ it follows that
\[
\int_{D_k \cap C_k(\xi)} |A_k|^2 d\mathcal{H}^2 \leq \int_{\text{graph} w_k} |A_k|^2 d\mathcal{H}^2 + c\rho + \epsilon_k.
\]

By definition of $w_k$ and the choice of $\tau$ we get
\[
\int_{\text{graph} w_k} |A_k|^2 d\mathcal{H}^2 \leq c \int_{D_k \cap C_k(\xi) \cap C_k(\xi)} |A_k|^2 d\mathcal{H}^2.
\]
Since \( \mathcal{B}_{\Theta}(\xi) \subset C_\gamma(\xi) \), we get that (remember that \( D_k \cap B_{\Theta}(\xi) = \Sigma_k \cap B_{\Theta}(\xi) \))
\[
\int_{\Sigma_k \cap B_{\Theta}(\xi)} |A_k|^2 \, d\mathcal{H}^2 \leq c \int_{\Sigma_k \cap B_{\Theta}(\xi)} |A_k|^2 \, d\mathcal{H}^2 + c\rho + \varepsilon_k.
\]

Now by adding \( c \) times the left hand side of this inequality to both sides ("hole filling"), we deduce the following:

For \( \rho \leq \frac{\rho_0}{4} \) and infinitely many \( k \) it follows that
\[
\int_{\Sigma_k \cap B_{\Theta}(\xi)} |A_k|^2 \, d\mathcal{H}^2 \leq \gamma \int_{\Sigma_k \cap B_{\Theta}(\xi)} |A_k|^2 \, d\mathcal{H}^2 + c\rho + \varepsilon_k,
\]
where \( \gamma = \frac{c}{\rho_0} \in (0, 1) \) is a fixed universal constant. If we let
\[
g(\rho) = \liminf_{k \to \infty} \int_{\Sigma_k \cap B_{\Theta}(\xi)} |A_k|^2 \, d\mathcal{H}^2,
\]
we get that
\[
g(\rho) \leq \gamma g(2\rho) + c\rho \quad \text{for all } \rho \leq \frac{\rho_0}{4}.
\]
Now in view of Lemma 5.7 it follows that
\[
g(\rho) \leq c\rho^a \quad \text{for all } \rho \leq \frac{\rho_0}{2},
\]
and the Lemma is proved.

In the next step we want to do the same as in the proof of Lemma 3.4, where we constructed a sequence of functions which converged strongly in \( C^0 \) and weakly in \( W^{1,2} \). But now with the estimate of Lemma 3.7 we will get better control on the sequence.

So let \( \xi_0 \in \Sigma \setminus \mathcal{B}_\gamma \) and \( \rho_0 > 0 \) as in Lemma 3.7. Let \( \xi \in \Sigma \cap B_{\Theta}(\xi_0) \) and define the quantity \( a_\xi(\rho) \) by
\[
a_\xi(\rho) = \int_{\Sigma \cap B_{\Theta}(\xi)} |A_k|^2 \, d\mathcal{H}^2. \tag{3.35}
\]
Notice that by the choice of \( \rho_0 > 0 \) and Lemma 3.7 we have that
\[
a_\xi(\rho) \leq c\varepsilon^2 \quad \text{and} \quad \liminf_{k \to \infty} a_\xi(\rho) \leq \min \left\{ c\varepsilon^2, c\rho^a \right\} \quad \text{for all } \rho \leq \rho_0 \frac{\rho_0}{32} \tag{3.36}
\]
Furthermore we get from Lemma 3.3 and the Monotonicity formula that
\[
\sum_m \text{diam } d_{k,m} \leq c a_\xi(\rho)^{\frac{1}{2}} \rho \leq c \varepsilon^{\frac{1}{2}} \rho \quad \text{and} \quad \sum_m \mathcal{L}^2 \left( d_{k,m} \right) \leq c a_\xi(\rho)^{\frac{1}{2}} \rho^2, \tag{3.37}
\]
\[
\sum_n \text{diam } P_n^k \leq c a_\xi(\rho)^{\frac{1}{2}} \rho \quad \text{and} \quad \sum_n \mathcal{H}^2 \left( P_n^k \right) \leq c a_\xi(\rho)^{\frac{1}{2}} \rho^2. \tag{3.38}
\]
Now for \( \varepsilon \leq \rho_0 \) we may apply the generalized Poincaré-inequality, Lemma 5.8, to the functions \( f = D_j u_k \) and \( \delta = c a_\xi(\rho)^{\frac{1}{2}} \rho \) to get a constant vector \( \eta_k \) with \( |\eta_k| \leq c \varepsilon^{\frac{1}{2}} + \delta_k \leq c \) such that
\[
\int_{\Omega \cap B_{\Theta}(\xi)} \left| D u_k - \eta_k \right|^2 \leq c \rho^2 \int_{\Omega \cap B_{\Theta}(\xi)} \left| D^2 u_k \right|^2 + c a_\xi(\rho)^{\frac{1}{2}} \rho^2 \sup_{\Omega \cap B_{\Theta}(\xi)} \left| D u_k \right|^2.
\]
Since
\[ \int_{\Omega_k \cap B_{\xi}(\ell)} |D^2 u_k|^2 \leq c \int_{\text{graph } u|_{\Omega_k \cap B_{\xi}(\ell)}} |A_k|^2 \ dH^2 \leq c \alpha_k(\rho), \]
it follows for \( \varepsilon \leq \varepsilon_0 \) that
\[ \int_{\Omega_k \cap B_{\xi}(\ell)} |D u_k - \eta_k|^2 \leq c \alpha_k(\rho)\frac{1}{\rho^2}. \tag{3.39} \]

Let again \( \overline{u}_k \in C^{1,1}(B_{\lambda_k}(\xi) \cap L, L^I) \) be an extension of \( u_k \) to the whole disc \( B_{\lambda_k}(\xi) \cap L \) as in Lemma 5.6, namely \( \overline{u}_k = u_k \) in \( \Omega_k \). We again have that
\[ \|\overline{u}_k\|_{L^\infty(B_{\lambda_k}(\xi) \cap L)} \leq c \varepsilon^{\frac{1}{2}} \rho + \delta_k \leq c, \]
\[ \|D \overline{u}_k\|_{L^\infty(B_{\lambda_k}(\xi) \cap L)} \leq c \varepsilon^{\frac{1}{2}} + \delta_k \leq c. \]

From the gradient estimates for the function \( \overline{u}_k \), since \( |\eta_k| \leq c \), from (3.37), (3.39) and the choice of \( \rho_0 > 0 \) we get that
\[ \int_{B_{\lambda_k}(\xi) \cap L} |D \overline{u}_k - \eta_k|^2 = \int_{\Omega_k \cap B_{\lambda_k}(\xi)} |D u_k - \eta_k|^2 + \sum_m \int_{\partial_m \cap B_{\lambda_k}(\xi)} |D \overline{u}_k - \eta_k|^2 \leq c \alpha_k(\rho)\frac{1}{\rho^2}, \]
and therefore in view of (3.36) (we will always write \( \alpha \) even if it might change from line to line) that
\[ \liminf_{k \to \infty} \int_{B_{\lambda_k}(\xi) \cap L} |D \overline{u}_k - \eta_k|^2 \leq \min \left\{ c \rho^{2+\alpha}, c \varepsilon^{\frac{1}{2}} \rho^2 \right\} \quad \text{for all } \rho \leq \frac{\rho_0}{32}. \tag{3.40} \]

As before there exists a function \( u_\xi \in C^{0,1}(B_{\lambda_k}(\xi) \cap L, L^I) \) such that (after passing to a subsequence)
\[
(i) \quad \overline{u}_k \rightharpoonup u_\xi \quad \text{in } C^0(B_{\lambda_k}(\xi) \cap L, L^I), \\
(ii) \quad \overline{u}_k \rightharpoonup u_\xi \quad \text{weakly in } W^{1,2}(B_{\lambda_k}(\xi) \cap L, L^I), \\
(iii) \quad \frac{1}{\rho} \|u_\xi\|_{L^\infty(B_{\lambda_k}(\xi) \cap L)} + \|D u_\xi\|_{L^\infty(B_{\lambda_k}(\xi) \cap L)} \leq c \varepsilon^{\frac{1}{2}}.
\]

Remark: Be aware that the limit function depends on the point \( \xi \), since our sequence comes (more or less) from the Graphical Decomposition Lemma, which is a local statement.

Now we get an additional information about the limit function \( u_\xi \): Since \( \eta_k \rightharpoonup \eta \) with \( |\eta| \leq c \varepsilon^{\frac{1}{2}} \) (this follows since \( |\eta_k| \leq c \varepsilon^{\frac{1}{2}} + \delta_k \leq c \)), and since \( D \overline{u}_k \rightharpoonup D u_\xi \) weakly in \( L^2(B_{\lambda_k}(\xi) \cap L) \), it follows from lower semicontinuity and (3.40) that
\[ \int_{B_{\lambda_k}(\xi) \cap L} |D u_\xi - \eta|^2 \leq \min \left\{ c \rho^{2+\alpha}, c \varepsilon^{\frac{1}{2}} \rho^2 \right\} \quad \text{for all } \rho \leq \frac{\rho_0}{32}. \tag{3.41} \]

Now we are able to show that around the good points our limit varifold is given by a graph.
Lemma 3.8 Let \( \xi_0 \in \Sigma \setminus B_\varrho \) and \( \rho_0 > 0 \) as in Lemma 3.7. For all \( \xi \in \Sigma \cap B_\varrho (\xi_0) \) and all \( \rho < \theta_{\rho \varrho} \) we have that

\[
\mu \cdot B_\varrho (\xi) = \mathcal{H}^2 \cdot (\text{graph } u_\xi \cap B_\varrho (\xi)),
\]

where \( u_\xi \in C^{0,1}(B_\varrho (\xi) \cap L, L^2) \) is as above, in particular

\[
\frac{1}{\rho} |u_\xi|_{L^\infty(B_\varrho (\xi) \cap L)} + \| D u_\xi \|_{L^\infty(B_\varrho (\xi) \cap L)} \leq c \varrho \varphi.
\]

Proof: From the definition of \( \bar{u}_\xi \) it follows for \( \rho \leq \theta_{\rho \varrho} \) that

\[
\mathcal{H}^2 \cdot (\Sigma \cap B_\varrho (\xi)) = \mathcal{H}^2 \cdot (D_\xi \cap B_\varrho (\xi))
\]

\[
= \mathcal{H}^2 \cdot (\text{graph } \bar{u}_\xi \cap B_\varrho (\xi)) + \mathcal{H}^2 \cdot (D_\xi \cap B_\varrho (\xi)) - \mathcal{H}^2 \cdot (\text{graph } \bar{u}_\xi \cap D_\xi \cap B_\varrho (\xi))
\]

\[
= \mathcal{H}^2 \cdot (\text{graph } \bar{u}_\xi \cap B_\varrho (\xi)) + \theta_\xi \rho,
\]

(3.42)

where \( \theta_\xi \) is a signed measure given by

\[
\theta_\xi = \mathcal{H}^2 \cdot (D_\xi \cap \text{graph } \bar{u}_\xi \cap B_\varrho (\xi)) - \mathcal{H}^2 \cdot (\text{graph } \bar{u}_\xi \cap D_\xi \cap B_\varrho (\xi)) = \theta_\rho - \theta_\rho.
\]

The total mass \( |\theta_\rho| \) of \( \theta_\rho \), namely \( \theta_\rho(\mathbb{R}^3) + \theta_\rho(\mathbb{R}^3) \), can be estimated in view of (3.36), (3.37) and (3.38) by

\[
\theta_\rho(\mathbb{R}^3) + \theta_\rho(\mathbb{R}^3) \leq \sum_n \mathcal{H}^2 (P_n) + \sum_m \int_{d_n} \sqrt{1 + |D \bar{u}_\xi|^2} \leq ca_\rho(\rho)^4 \rho^2.
\]

It follows from (3.36) that

\[
\lim \inf_{k \to \infty} \left( \theta_\rho(\mathbb{R}^3) + \theta_\rho(\mathbb{R}^3) \right) \leq \min \{ \epsilon \rho^{2+\alpha}, \epsilon \varrho \rho^2 \}.
\]

(3.43)

By taking limits in the measure theoretic sense we get that

\[
\mu \cdot B_\varrho (\xi) = \mathcal{H}^2 \cdot (\text{graph } u_\xi \cap B_\varrho (\xi)) + \theta_\rho,
\]

(3.44)

where \( \theta_\rho \) is a signed measure with total mass \( |\theta_\rho| \leq \min \{ \epsilon \rho^{2+\alpha}, \epsilon \varrho \rho^2 \} \). This equation holds for all \( \rho \leq \theta_{\rho \varrho} \) such that

\[
\mu \left( \partial B_\varrho (\xi) \right) = \mathcal{H}^2 \cdot \text{graph } u_\xi \left( \partial B_\varrho (\xi) \right) = 0,
\]

which holds for almost every \( \rho \leq \theta_{\rho \varrho} \).

To prove (3.44) let \( U \subset \mathbb{R}^3 \) open.

1.) Let \( \rho \leq \theta_{\rho \varrho} \) such that \( \mu \left( \partial B_\varrho (\xi) \right) = 0 \). Moreover assume that \( \mu \cdot B_\varrho (\xi) (\partial U) = 0 \). Therefore \( \mu \left( \partial (U \cap B_\varrho (\xi)) \right) = 0 \), and we get \( \mu_k (U \cap B_\varrho (\xi)) \to \mu (U \cap B_\varrho (\xi)) \). It follows that

\[
\mathcal{H}^2 \cdot (\Sigma_k \cap B_\varrho (\xi))(U) \to \mu \cdot B_\varrho (\xi)(U),
\]

(3.45)
Let $\rho \leq \theta_{\rho_{\rho_{\rho_{\rho_{\rho_{\rho_{\rho_{\rho_{T}}}}}}}}}$ such that $\mathcal{H}^2_L(\mathrm{graph} \ u_\xi(\partial B_{\xi}(\xi))) = 0$. Moreover, assume that $\mathcal{H}^2_L(\mathrm{graph} \ u_\xi \cap B_{\xi}(\xi)) (\partial U) = 0$. We have that

$$\mathcal{H}^2_L(\mathrm{graph} \ u_\xi \cap B_{\xi}(\xi)) (U) = \int_L \chi_{\mathrm{graph} \ u_\xi}(x) \sqrt{1 + |D \bar{u}_\xi(x)|^2}.$$

It follows from the $L^\infty$-bounds for $\bar{u}_k$ that

$$\left| \int_L \chi_{\mathrm{graph} \ u_\xi}(x) \sqrt{1 + |D \bar{u}_k(x)|^2} - \int_L \chi_{\mathrm{graph} \ u_\xi}(x) \sqrt{1 + |D u_\xi(x)|^2} \right| \leq c \left( \int_L \chi_{\mathrm{graph} \ u_\xi}(x) |D \bar{u}_k(x) - \eta_k(x)| + c \int_L \chi_{\mathrm{graph} \ u_\xi}(x) |\eta_k(x) - \eta| + c \int_L \chi_{\mathrm{graph} \ u_\xi}(x) |\eta - D u_\xi(x)| \right).$$

Since $\bar{u}_k \to u_\xi$ uniformly and $\mathcal{H}^2_L(\mathrm{graph} \ u_\xi(\partial (U \cap B_{\xi}(\xi))) = 0$, we first of all get that

$$\mathcal{H}^2_L(\mathrm{graph} \ u_\xi(\partial (U \cap B_{\xi}(\xi))) \to \mathcal{H}^2_L(\mathrm{graph} \ u_\xi(x)) \quad \text{for a.e. } x \in L.$$

The Dominated Convergence Theorem yields

$$\int_L \left| \chi_{\mathrm{graph} \ u_\xi}(x) \sqrt{1 + |D \bar{u}_k(x)|^2} - \int_L \chi_{\mathrm{graph} \ u_\xi}(x) \sqrt{1 + |D u_\xi(x)|^2} \right| \to 0.$$

On the other hand we have that

$$\int_L \chi_{\mathrm{graph} \ u_\xi}(x) \sqrt{1 + |D \bar{u}_k(x)|^2} \to \int_L \chi_{\mathrm{graph} \ u_\xi}(x) \sqrt{1 + |D u_\xi(x)|^2}.$$

The $L^\infty$-bound for $u_\xi$ yields that $\chi_{\mathrm{graph} \ u_\xi}(x) = 0$ if $x \notin B_{\{|1-c \varepsilon^{\frac{1}{2}}\}}(\xi) \cap L$, and we get that

$$\left( \int_L \chi_{\mathrm{graph} \ u_\xi}(x) \sqrt{1 + |D \bar{u}_k(x)|^2} \right)^{\frac{1}{2}} \leq L^2 \left( B_{\{|1-c \varepsilon^{\frac{1}{2}}\}}(\xi) \cap L \right) \leq c \rho.$$

In view of (3.40), (3.41) and since $\eta_k \to \eta$ we get that

$$\liminf_{k \to \infty} \int_L \chi_{\mathrm{graph} \ u_\xi}(x) \sqrt{1 + |D \bar{u}_k(x)|^2} \to \int_L \chi_{\mathrm{graph} \ u_\xi}(x) \sqrt{1 + |D u_\xi(x)|^2} \leq \min \left\{ c \rho^{2+\alpha}, c \varepsilon^{\frac{1}{2}} \rho^2 \right\},$$

and it follows after all that

$$\mathcal{H}^2_L(\mathrm{graph} \ u_\xi \cap B_{\xi}(\xi)) (U) = \mathcal{H}^2_L(\mathrm{graph} \ u_\xi \cap B_{\xi}(\xi)) (U) + \tilde{\partial}_k(U),$$

where $\tilde{\partial}_k$ is a signed measure with $\liminf_{k \to \infty} |\tilde{\partial}_k| \leq \min \left\{ c \rho^{2+\alpha}, c \varepsilon^{\frac{1}{2}} \rho^2 \right\}$. After passing to a subsequence, the $\tilde{\partial}_k$’s converge to some signed measure $\tilde{\partial}_\xi$ with total
mass \(|\tilde{\theta}_k| \leq \min\left\{c\rho^{2+a}, ce^\frac{1}{4}\rho^2\right\}\). Assume that \(\tilde{\theta}_k(\partial U) = 0\). Then it follows that 
\(\tilde{\theta}_k(U) \to \tilde{\theta}_c(U)\), and therefore we get
\[
\lim_{k \to \infty} \mathcal{H}^2(L(\text{graph } u_k \cap B_{\bar{\xi}}(\xi))(U) = \mathcal{H}^2(L(\text{graph } u_\xi \cap B_{\bar{\xi}}(\xi))(U) + \tilde{\theta}_c(U). \tag{3.46}
\]

3.) Since the \(\tilde{\theta}_k\)'s were signed measures such that \(\lim \inf |\tilde{\theta}_k| \leq \min\left\{c\rho^{2+a}, ce^\frac{1}{4}\rho^2\right\}\), they converge in the weak sense (after passing to a subsequence) to a signed measure \(\tilde{\theta}_c\) with total mass \(|\tilde{\theta}_c| \leq \min\left\{c\rho^{2+a}, ce^\frac{1}{4}\rho^2\right\}\). Assuming that \(\tilde{\theta}_c(\partial U) = 0\), we get \(\tilde{\theta}_k(U) \to \tilde{\theta}_c(U)\).

Now by taking limits in (3.42), it follows that
\[
\mu(B_{\bar{\xi}}(\xi))(U) = \mathcal{H}^2(L(\text{graph } u_\xi \cap B_{\bar{\xi}}(\xi))(U) + \tilde{\theta}_c(U), \tag{3.47}
\]
where \(\tilde{\theta}_c = \tilde{\theta}_c + \tilde{\theta}_c\) is a signed measure with total mass \(|\tilde{\theta}_c| \leq \min\left\{c\rho^{2+a}, ce^\frac{1}{4}\rho^2\right\}\). Notice that this equation holds for every open \(U \subset \mathbb{R}^3\) such that
\[
\mu(B_{\bar{\xi}}(\xi))(\partial U) = \mathcal{H}^2(L(\text{graph } u_\xi \cap B_{\bar{\xi}}(\xi))(\partial U) = \tilde{\theta}_c(\partial U) = \tilde{\theta}_c(\partial U) = 0.
\]

By choosing an appropriate exhaustion this equation holds for arbitrary open sets \(U \subset \mathbb{R}^3\), and (3.44) follows.

Now choose a radius \(\rho \leq \theta_{\frac{40}{64}}\) such that (3.44) holds. We take a closer look to two cases.

1.) Let \(x \in \Sigma \cap B_{\bar{\xi}}(\xi)\): Notice that by (3.6) and the choice of \(\rho_0 > 0\)
\[
\mathcal{W}(\mu, B_{\bar{\xi}}(x)) \leq \lim_{k \to \infty} \mathcal{W}(\mu, B_{\bar{\xi}}(x)) \leq c \int_{\Sigma \cap B_{\bar{\xi}}(\xi)} |\Lambda_k|^2 d\mathcal{H}^2 \leq 2\epsilon^2.
\]
Since \(\theta^2(\mu, \cdot) \geq 1\) on \(\text{spt } \mu\), it follows for \(\epsilon \leq \epsilon_0\) from Theorem 5.3 that
\[
\mu(B_{\bar{\xi}}(\xi))(\Sigma \cap B_{\bar{\xi}}(x)) = \mu(B_{\bar{\xi}}(x)) \geq c\rho^2.
\]
From (3.44), especially the bound on the total mass of \(\tilde{\theta}_c\), it follows that
\[
c\rho^2 \leq \mathcal{H}^2(\text{graph } u_\xi \cap B_{\bar{\xi}}(x)) + ce^\frac{1}{4}\rho^2.
\]
Therefore \(\mathcal{H}^2(\text{graph } u_\xi \cap B_{\bar{\xi}}(x)) > 0\) for \(\epsilon \leq \epsilon_0\), and thus \(x \in \text{graph } u_\xi\).

2.) Let \(x \in \text{graph } u_\xi \cap B_{\bar{\xi}}(\xi)\): Write \(x = z + u_\xi(z)\). If \(y \in B_{\bar{\xi}}(z) \cap L\), it follows from the estimates for \(u_\xi\) that \(y + u_\xi(y) \in B_{\bar{\xi}}(x)\) for \(\epsilon \leq \epsilon_0\). Therefore we get that
\[
\mathcal{H}^2(\text{graph } u_\xi \cap B_{\bar{\xi}}(x) \geq \int_{\mathbb{R}^3} c\rho^2.
\]
As above it follows that \(x \in \Sigma\) for \(\epsilon \leq \epsilon_0\).

After all we get for \(\epsilon \leq \epsilon_0\)
\[
\Sigma \cap B_{\bar{\xi}}(\xi) = \text{graph } u_\xi \cap B_{\bar{\xi}}(\xi) \quad \text{for all } \rho < \theta_{\frac{40}{64}}. \tag{3.48}
\]
Moreover we get that the function \( u_\xi \) does not depend on the point \( \xi \) in the following sense: Let \( x \in \Sigma \cap B_{\rho}(\xi_0) \) and \( \tau < \theta^\alpha_{\Sigma} \). Then we have that
\[
\text{graph } u_\xi \cap \left( B_{\tau}(\xi) \cap B_{\rho}(x) \right) = \text{graph } u_x \cap \left( B_{\tau}(\xi) \cap B_{\rho}(x) \right),
\]
where \( \theta \) is a signed measure with total mass smaller than \( c\tau^{2+\alpha} \).

From (3.49), (3.50) and (3.51) it follows that
\[
\text{graph } u_\xi \cap \left( B_{\tau}(\xi) \cap B_{\rho}(x) \right) \quad \text{is a weak solution of the weak mean curvature equation, namely we have}
\]
where \( \theta_x \) is a signed measure with total mass smaller than \( c\tau^{2+\alpha} \).

From (3.49), (3.50) and (3.51) it follows that
\[
\theta_\xi (B_\tau(x)) = \theta_x (B_\tau(x)),
\]
and we get a nice decay for the signed measure \( \theta_\xi \), namely
\[
\lim_{\tau \to 0} \frac{\theta_\xi (B_\tau(x))}{\tau^2} = 0.
\]

Since we already know that \( \theta^2(\mu, \cdot) \geq 1 \) on \( \Sigma \), it follows from (3.50) that
\[
\frac{\mathcal{H}^2(\text{graph } u_\xi \cap B_{\rho}(\xi))(B_\tau(x))}{\mu(B_{\rho}(\xi))(B_\tau(x))} = 1 - \frac{\theta_\xi (B_\tau(x))}{\mu(B_{\rho}(\xi))(B_\tau(x))},
\]
and by (3.52) the right hand side goes to 1 for \( \tau \to 0 \). This shows that
\[
D_{\mu, B_{\rho}(\xi)} \left( \mathcal{H}^2(\text{graph } u_\xi \cap B_{\rho}(\xi)) \right)(x) = 1
\]
for all \( x \in \Sigma \cap B_{\rho}(\xi) = \text{graph } u_\xi \cap B_{\rho}(\xi) \), and the Lemma follows from the Theorem of Radon-Nikodym.

Up to now we have shown that in a neighborhood around the good points the limit varifold \( \mu \) is given by a \( C^{0,1} \)-graph with small gradient bounded by \( c\rho^4 \). On the other hand we already know that the limit varifold has generalized mean curvature vector \( \bar{H} \in L^2(\mu) \). From the definition of the weak mean curvature vector and the graph representation it follows that the graph function solves the weak mean curvature equation, which we will prove in the next Lemma.

**Lemma 3.7** Let \( \xi_0 \in \Sigma \setminus B_\rho \) and \( \rho > 0 \) as in Lemma 3.7. Let \( \xi \in \Sigma \cap B_{\rho}(\xi_0) \) and \( \rho < \theta^\alpha_{\Sigma} \), so that we have due to Lemma 3.8 the graph representation
\[
\mu(B_{\rho}(\xi))(B_\tau(x)) = \mathcal{H}^2(\text{graph } u_\xi \cap B_{\rho}(\xi)),
\]
where \( u_\xi \in C^{0,1}(B_{\rho}(\xi) \cap L, L^2) \) satisfies
\[
\frac{1}{\rho} \| u_\xi \|_{L^2(B_{\rho}(\xi) \cap L)} + \| D u_\xi \|_{L^2(B_{\rho}(\xi) \cap L)} \leq c\rho^4.
\]
Then \( u_\xi \) is a weak solution of the weak mean curvature equation, namely we have
\[
\sum_{i,j=1}^2 \int g^{ij} \langle \partial_j \varphi, \partial_i F \rangle \sqrt{\det g} = - \int \langle \varphi, \bar{H} \circ F \rangle \sqrt{\det g}
\]
for all \( \varphi \in W^{1,2}_0(B_{\rho}(\xi) \cap L, \mathbb{R}^3) \), where \( F(x) = x + u_\xi(x) \) and \( g^{ij} = \delta_{ij} + \partial_i u_\xi \cdot \partial_j u_\xi \).
Proof: By definition of the weak mean curvature vector we have that
\[ -\int \langle X, \hat{H} \rangle \, d\mu = \int \text{div}_\mu X \, d\mu \] for all \( X \in C_c^1(\mathbb{R}^3, \mathbb{R}^3) \).

For \( X \in C_c^1(B_\xi^c(\xi), \mathbb{R}^3) \) it follows from Lemma 3.8 that
\[ \int \langle X, \hat{H} \rangle \, d\mu = \int_{\text{graph } u_\xi} \langle X, \hat{H} \rangle \, d^2\mathcal{H} = \int_{B_\xi^c(\xi) \cap L} \langle X \circ F, \hat{H} \circ F \rangle \sqrt{1 + |D u_\xi|^2}, \]
where \( F : B_\xi^c(\xi) \cap L \to \mathbb{R}^3 \) is given by \( F(x) = x + u_\xi(x) \).

For \( x \in \text{graph } u_\xi \) denote by \( \{\tau_i(x)\}_{i=1}^2 \) an ONB of \( T_x \text{graph } u_\xi \). It follows that
\[
\int \text{div}_\mu X \, d\mu = \int_{\text{graph } u_\xi} \text{div}_{\text{graph } u_\xi} X \, d^2\mathcal{H} = \sum_{i=1}^2 \int_{B_\xi^c(\xi) \cap L} \langle DX \cdot \tau_i, \tau_i \rangle \, d^2\mathcal{H}^2 \]
\[ = \sum_{i=1}^2 \int_{B_\xi^c(\xi) \cap L} \langle DX \circ F \cdot \tau_i \circ F, \tau_i \circ F \rangle \sqrt{1 + |D u_\xi|^2}. \]

Now we have that
\[
DX \circ F \cdot \tau_i \circ F = DX \circ F \cdot \sum_{k,l=1}^2 g^{kl} \langle \tau_i \circ F, \partial_k F \rangle \partial_l (X \circ F). \]

By summing over \( i = 1, 2 \) it follows that
\[
\sum_{i=1}^2 \langle DX \circ F \cdot \tau_i \circ F, \tau_i \circ F \rangle = \sum_{k,l=1}^2 g^{kl} \sum_{i=1}^2 \langle \tau_i \circ F, \partial_k F \rangle \langle \partial_l (X \circ F), \tau_i \circ F \rangle \]
\[ = \sum_{k,l=1}^2 g^{kl} \sum_{i=1}^2 \langle \partial_l (X \circ F), \langle \tau_i \circ F, \partial_k F \rangle \tau_i \circ F \rangle \]
\[ = \sum_{k,l=1}^2 g^{kl} \langle \partial_l (X \circ F), \partial_k F \rangle. \]

Therefore we get that
\[
\sum_{k,l=1}^2 \int_{B_\xi^c(\xi) \cap L} g^{kl} \langle \partial_l (X \circ F), \partial_k F \rangle \sqrt{\det g} = -\int_{B_\xi^c(\xi) \cap L} \langle X \circ F, \hat{H} \circ F \rangle \sqrt{\det g}.
\]

Let \( \varphi \in C_c^1(B_\xi^c(\xi) \cap L, \mathbb{R}^3) \). Since \( u_\xi(\xi) = 0 \) and \( |D u_\xi| \leq c e^{\frac{\xi}{2}} \), it follows for \( \epsilon \leq \epsilon_0 \) that there exists a vectorfield \( X \in C_c^1(B_\xi^c(\xi), \mathbb{R}^3) \) such that \( X \circ F = \varphi \). We get that
\[
\sum_{k,l=1}^2 \int_{B_\xi^c(\xi) \cap L} g^{kl} \langle \partial_l \varphi, \partial_k F \rangle \sqrt{\det g} = -\int_{B_\xi^c(\xi) \cap L} \langle \varphi, \hat{H} \circ F \rangle \sqrt{\det g}
\]
for all \( \varphi \in C_c^1(B_\xi^c(\xi) \cap L, \mathbb{R}^3) \), and the Lemma is proved. \( \square \)
Now since the norm of the mean curvature vector can be estimated by the norm of the second fundamental form, it follows from Lemma 3.7 and (3.6) that for all \( \xi \in B_{2\rho}^{\infty}(\xi_0) \) and all \( \rho \leq \theta \rho_0 \frac{\rho_0}{C_{\xi_0}} \)

\[
\int_{B_{\rho}(\xi)} |\tilde{H}|^2 \, d\mu \leq c \rho^\alpha.
\]

By applying Lemma 3.8 and Lemma 3.9 to \( \xi_0 \in \Sigma \setminus B_\varepsilon \), we get that

\[
\mu(B_{\rho_0}^{\infty}(\xi_0) \cap \Sigma \setminus B_\varepsilon) = \mathcal{H}^2 \left( \text{graph } u_{\xi_0} \cap B_{\rho_0}^{\infty}(\xi_0) \right),
\]

where \( u_{\xi_0} \in C^{0,1}(B_{\rho_0}^{\infty}(\xi_0) \cap \Sigma \setminus B_\varepsilon) \) is a solution of the weak mean curvature equation and satisfies

\[
\frac{1}{\rho} ||u_{\xi_0}||_{L^\infty(B_{\rho_0}^{\infty}(\xi_0) \cap \Sigma \setminus B_\varepsilon)} + ||D u_{\xi_0}||_{L^\infty(B_{\rho_0}^{\infty}(\xi_0) \cap \Sigma \setminus B_\varepsilon)} \leq c \rho^\frac{1}{2} \text{ for all } \rho \leq \theta \rho_0 \frac{\rho_0}{C_{\xi_0}}.
\]

Therefore we get for all \( \xi \in B_{\rho_0}^{\infty}(\xi_0) \) and all \( \rho \leq \theta \rho_0 \frac{\rho_0}{C_{\xi_0}} \) that

\[
\int_{\text{graph } u_{\xi_0} \cap B_{\rho}(\xi)} |\tilde{H}|^2 \, d\mathcal{H}^2 = \int_{B_{\rho}(\xi)} |\tilde{H}|^2 \, d\mu \leq c \rho^\alpha.
\]

In the next Lemma we prove that the function \( u_{\xi_0} \) is actually \( C^{1,\alpha} \cap W^{2,2} \), and that the \( L^2 \)-norm of the Hessian satisfies a certain power decay, which will be needed to prove higher regularity.

**Lemma 3.10** There exists an universal constant \( \alpha \in (0, 1) \) such that for \( \xi_0 \in \Sigma \setminus B_\varepsilon \) and \( \rho_0 > 0 \) as in Lemma 3.7

(i) \( u_{\xi_0} \in C^{1,\alpha}(B_{\rho_0}^{\infty}(\xi_0) \cap \Sigma \setminus B_\varepsilon, \cap L^1) \cap W^{2,2}(B_{\rho_0}^{\infty}(\xi_0) \cap \Sigma \setminus B_\varepsilon, \cap L^1) \),

(ii) \( \int_{B_{\rho}(\xi)} |D^2 u_{\xi_0}|^2 \leq c \tau^{2\alpha} \text{ for all } \xi \in B_{\rho_0}^{\infty}(\xi_0) \cap \Sigma \setminus B_\varepsilon \text{ and all } \tau < \theta \rho_0 \frac{\rho_0}{4096} \).

**Proof:** First of all it follows from Lemma 3.9 and a standard difference quotient argument (see [GT01], Theorem 8.8) that \( u_{\xi_0} \in W^{2,2}_{\text{loc}}(B_{\sigma_{\rho_0}}^{\infty}(\xi_0) \cap \Sigma \setminus B_\varepsilon, \cap L^1) \).

Now let \( \nu \perp L \). Since \( u_{\xi_0} \) has values in \( L^1 \), we may assume that \( u_{\xi_0} \) is real-valued. Let \( \varphi = \eta \nu \), where \( \eta \in C_{\varepsilon}^{\infty}(B_{\rho_0}^{\infty}(\xi_0) \cap L) \). Lemma 3.9 yields

\[
\sum_{i,j=1}^{2} \int_{L} g^{ij} \partial_i u_{\xi_0} \partial_j \eta \sqrt{\det g} = - \int_{L} (\hat{H} \circ F, \nu) \eta \sqrt{\det g}.
\]

Since \( u_{\xi_0} \in W^{2,2}_{\text{loc}}(B_{\rho_0}^{\infty}(\xi_0) \cap \Sigma \setminus B_\varepsilon, \cap L^1) \), we may exchange \( u_{\xi_0} \) by \( \partial_i u_{\xi_0} \) to get the following equation

\[
\sum_{i,j=1}^{2} \int_{L} g^{ij} \partial_i (\partial_i u_{\xi_0}) \partial_j \eta \sqrt{\det g} = - \sum_{i,j=1}^{2} \int_{L} \partial_i u_{\xi_0} \partial_i (g^{ij} \partial_j \eta \sqrt{\det g})
\]

\[
= - \sum_{i,j=1}^{2} \int_{L} g^{ij} \partial_i u_{\xi_0} \partial_i \partial_j \eta \sqrt{\det g} - \sum_{i,j=1}^{2} \int_{L} \partial_i (g^{ij} \sqrt{\det g}) \partial_i u_{\xi_0} \partial_j \eta
\]

\[
= \int_{L} (\hat{H} \circ F, \nu) \partial_i \eta \sqrt{\det g} - \sum_{i,j=1}^{2} \int_{L} \partial_i (g^{ij} \sqrt{\det g}) \partial_i u_{\xi_0} \partial_j \eta.
\]
Now let \( x \in B_{\rho_0 \cos \theta} (\xi_0) \cap L \) and \( \tau < \theta \frac{\rho_0}{40 \theta} \). Then we get that
\[
\sum_{i,j=1}^{2} \int_{L} g^{ij} \partial \left( \nabla u_{\xi_0} - \int_{B(\lambda \cdot L)} \partial u_{\xi_0} \right) \partial \eta \sqrt{\det g} \\
= \int_{L} (H \circ F, \nu) \partial \eta \sqrt{\det g} - \sum_{i,j=1}^{2} \int_{L} \left( g^{ij} \sqrt{\det g} \right) \partial_i u_{\xi_0} \partial_j \eta. \tag{3.56}
\]
By approximation this equation holds also for all \( \eta \in W^{1,2}_0 (B_{\rho_0 \cos \theta} (\xi_0) \cap L) \). Now let \( f \in C^1_c (B(\lambda \cdot L)) \) such that \( 0 \leq f \leq 1 \), \( f \equiv 1 \) on \( B_{\frac{\lambda}{2}} (\xi_0) \cap L \) and \( |D f| \leq \frac{\tau}{\lambda} \), and choose
\[
\eta = f^2 \left( \nabla u_{\xi_0} - \int_{B(\lambda \cdot L)} \partial u_{\xi_0} \right) \in W^{1,2}_0 (B_{\rho_0 \cos \theta} (\xi_0) \cap L).
\]
Inserting this function into (3.56) (and dropping the sum) yields
\[
\int_{L} g^{ij} \partial \left( \nabla u_{\xi_0} - \int_{B(\lambda \cdot L)} \partial u_{\xi_0} \right) \partial \eta \sqrt{\det g} \\
= \int_{L} g^{ij} \partial \left( \nabla u_{\xi_0} - \int_{B(\lambda \cdot L)} \partial u_{\xi_0} \right) 2 f^2 \partial_j \left( \nabla u_{\xi_0} - \int_{B(\lambda \cdot L)} \partial u_{\xi_0} \right) \sqrt{\det g} \\
+ \int_{L} g^{ij} \partial \left( \nabla u_{\xi_0} - \int_{B(\lambda \cdot L)} \partial u_{\xi_0} \right) f^2 \partial_j \left( \nabla u_{\xi_0} - \int_{B(\lambda \cdot L)} \partial u_{\xi_0} \right) \sqrt{\det g} \\
=: (1) + (2).
\]
Using the assumptions on \( f \), the Poincaré-inequality and the estimates for the derivative of \( u_{\xi_0} \) we can estimate (1) by
\[
\begin{align*}
(1) & \leq c \int_{B(\lambda \cdot L)} \left| g^{ij} \right| |D \nabla u_{\xi_0} - \int_{B(\lambda \cdot L)} \partial u_{\xi_0} | \sqrt{\det g} \\
& \leq c \int_{B(\lambda \cdot L)} |D \nabla u_{\xi_0}|^2.
\end{align*}
\]
Now for any \( \zeta \in \mathbb{R}^2 \) we have for \( \varepsilon \leq \varepsilon_0 \) that
\[
\sum_{i,j=1}^{2} g^{ij} \xi_i \xi_j = \frac{1}{\det g} \left( (1 + (\partial_2 u_{\xi_0})^2) \xi_1^2 - 2 (1 u_{\xi_0} \partial_2 u_{\xi_0} \xi_1 \xi_2 + (1 + (\partial_1 u_{\xi_0})^2) \xi_2^2) \right) \\
\geq \frac{1}{\det g} \left( \xi_1^2 - c e^4 |\xi_1||\xi_2| + \xi_2^2 \right) \\
\geq \frac{|\xi|^2}{2 \det g}.
\]
Using that \( f \equiv 1 \) on \( B_{\frac{\lambda}{2}} (\xi_0) \cap L \) and \( \det g = 1 + |D u_{\xi_0}|^2 \in \{ 1, 1 + c e^4 \} \subset (1, 2) \), we can estimate (2) from below by
\[
(2) = \int_{L} f^2 g^{ij} \partial_i \partial_j u_{\xi_0} \sqrt{\det g} \geq \frac{1}{4} \int_{B_{\frac{\lambda}{2}} (\xi_0) \cap L} |D \partial u_{\xi_0}|^2.
\]
After all we can estimate the left hand side of (3.56) by
\[
\sum_{i,j=1}^{2} \int_{L} g^{ij} \partial_{i} \left( \partial_{j} u_{\xi_{0}} - \int_{B_{r}(x) \cap L} \partial_{j} u_{\xi_{0}} \right) \partial_{j} \sqrt{\text{det} g} \geq \frac{1}{4} \int_{B_{2r}(x) \cap L} |D \partial_{j} u_{\xi_{0}}|^{2} - c \int_{B_{r}(x) \cap L} |D \partial_{j} u_{\xi_{0}}|^{2}. \tag{3.57}
\]

For the first term on the right hand side of (3.56) we have that
\[
\int_{L} \langle \hat{H} \circ F, v \rangle \partial_{j} \sqrt{\text{det} g} = \int_{L} \langle \hat{H} \circ F, v \rangle 2f \partial_{j} f \left( \partial_{j} u_{\xi_{0}} - \int_{B_{r}(x) \cap L} \partial_{j} u_{\xi_{0}} \right) \sqrt{\text{det} g} + \int_{L} \langle \hat{H} \circ F, v \rangle f^{2} \partial_{j}^{2} u_{\xi_{0}} \sqrt{\text{det} g} =: (3) + (4).
\]

By the assumptions on $f$ and the Poincaré-inequality we get
\[
(3) \leq \frac{c}{\tau} \int_{B_{r}(x) \cap L} |\hat{H} \circ F| \left| \partial_{j} u_{\xi_{0}} - \int_{B_{r}(x) \cap L} \partial_{j} u_{\xi_{0}} \right| \sqrt{\text{det} g} \leq c \int_{\text{graph} u_{\xi_{0}}(y \circ r_{\xi_{0}}) \cap L} |\hat{H}|^{2} + c \int_{B_{r}(x) \cap L} |D \partial_{j} u_{\xi_{0}}|^{2}.
\]

Now we have that
\[
\text{graph} u_{\xi_{0}}(y \circ r_{\xi_{0}}) \subset B_{2r}(x + u_{\xi_{0}}(x)),
\]

since for $y \in B_{r}(x) \cap L$ it follows from the $L^{\infty}$-bound for $D u_{\xi_{0}}$ that
\[
|y + u_{\xi_{0}}(y) - x - u_{\xi_{0}}(x)| \leq \left(1 + ce^{\frac{1}{\rho}}\right) |y - x| \leq 2 \tau \quad \text{for } \varepsilon \leq \varepsilon_{0}.
\]

Since $x \in B_{\frac{\varepsilon_{0}}{2048}}(\xi_{0}) \cap L$, we get that (using that $|u_{\xi_{0}}(x)| \leq ce^{\frac{1}{\rho}\rho_{0}}$)
\[
|x + u_{\xi_{0}}(x) - \xi_{0}| \leq \theta \frac{\rho_{0}}{4096} + ce^{\frac{1}{\rho}\rho_{0}} \leq \theta \frac{\rho_{0}}{2048} \quad \text{for } \varepsilon \leq \varepsilon_{0},
\]

and since $\tau < \theta \frac{\rho_{0}}{4096}$, it follows that
\[
B_{2\tau}(x + u_{\xi_{0}}(x)) \subset B_{\frac{\varepsilon_{0}}{2048}}(\xi_{0}).
\]

Therefore (3.55) yields
\[
(3) \leq c \tau^{\alpha} + c \int_{B_{r}(x) \cap L} |D \partial_{j} u_{\xi_{0}}|^{2}.
\]

Now we estimate (4). We have that
\[
(4) \leq \int_{B_{2r}(x) \cap L} |\hat{H} \circ F| \partial_{j}^{2} u_{\xi_{0}} |\sqrt{\text{det} g} + \int_{B_{r}(x) \cap L} |\hat{H} \circ F| |\partial_{j}^{2} u_{\xi_{0}}| \sqrt{\text{det} g}.
\]
Using the $\delta$-Young inequality, the first term can be estimated in view of (3.55) by

$$\int_{B_{r}(x)\cap L} |H \circ F| |\partial_i^2 u_{\delta}| \sqrt{\det g} \leq \frac{c}{\delta} \tau^p + c \delta \int_{B_{r}(x)\cap L} |D \partial_i u_{\delta}|^2.$$  

The second term can be estimated similarly (without the $\delta$-Young inequality) by

$$\int_{B_{r}(x)\cap L} |H \circ F| |\partial_i^2 u_{\delta}| \sqrt{\det g} \leq c \tau^p + c \int_{B_{r}(x)\cap L} |D \partial_i u_{\delta}|^2.$$  

It follows that

$$(4) \leq c \tau^p + \frac{c}{\delta} \tau^p + c \int_{B_{r}(x)\cap L} |D \partial_i u_{\delta}|^2 + c \delta \int_{B_{r}(x)\cap L} |D \partial_i u_{\delta}|^2,$$

and therefore the first term on the right hand side of (3.56) is estimated by

$$\int_L \langle \hat{H} \circ F, \nu \rangle \partial_{ij} \sqrt{\det g} \leq c \tau^p + c \int_{B_{r}(x)\cap L} |D \partial_i u_{\delta}|^2 + c \int_{B_{r}(x)\cap L} |D \partial_i u_{\delta}|^2. \quad (3.58)$$

Now we estimate the second term on the right hand side of (3.56), namely

$$\sum_{i,j=1}^2 \int_L \partial_i \left( g^{ij} \sqrt{\det g} \right) \partial_i u_{\delta} \partial_j \eta = \sum_{i,j=1}^2 \int_L \partial_i \left( g^{ij} \sqrt{\det g} \right) \partial_i u_{\delta} 2f \partial_j f \left( \partial_i u_{\delta} - \int_{B_{r}(x)\cap L} \partial_i u_{\delta} \right)$$

$$+ \sum_{i,j=1}^2 \int_L \partial_i \left( g^{ij} \sqrt{\det g} \right) \partial_i u_{\delta} f^2 \partial_j \partial_i u_{\delta}$$

$$= (5) + (6).$$

Now notice that

$$|\partial_i \left( g^{ij} \sqrt{\det g} \right)| \leq c \sup_{i,j} |\partial_i g^{ij}| \leq c |D u_{\delta}| \leq c |D \partial_i u_{\delta}|.$$

Therefore in view of the Poincaré-inequality and the estimates on $D u_{\delta}$ we can estimate (5) by

$$(5) \leq c \frac{\tau}{r} \int_{B_{r}(x)\cap L} |D \partial_i u_{\delta}| \left| \partial_i u_{\delta} - \int_{B_{r}(x)\cap L} \partial_i u_{\delta} \right|$$

$$\leq c \int_{B_{r}(x)\cap L} |D \partial_i u_{\delta}|^2.$$

In view of the estimates on $D u_{\delta}$, (6) can be estimated very easily by

$$(6) \leq c \tau^p \int_{B_{r}(x)\cap L} |D \partial_i u_{\delta}|^2 + c \int_{B_{r}(x)\cap L} |D \partial_i u_{\delta}|^2.$$
After all the second term on the right hand side of (3.56) is estimated by
\[
\sum_{i,j=1}^{2} \int_{L} \partial_i (g^{ij} \sqrt{\det g}) \partial_j u_{\xi_0} \partial_j \eta \\
\leq c e^{\frac{\lambda}{2}} \int_{B_{2^4}(x) \cap L} |D \partial_i u_{\xi_0}|^2 + c \int_{B_{4}(x) \cap B_{2^4}(x) \cap L} |D \partial_i u_{\xi_0}|^2.
\] (3.59)

Therefore we have finally shown that it follows from (3.56) that
\[
\frac{1}{4} \int_{B_{2^4}(x) \cap L} |D \partial_i u_{\xi_0}|^2 - c \int_{B_{4}(x) \cap B_{2^4}(x) \cap L} |D \partial_i u_{\xi_0}|^2 \\
\leq c \tau^\alpha + c \frac{\epsilon}{\delta} \tau^\alpha + c \left( \delta + \epsilon^2 \right) \int_{B_{2^4}(x) \cap L} |D \partial_i u_{\xi_0}|^2 + c \int_{B_{4}(x) \cap B_{2^4}(x) \cap L} |D \partial_i u_{\xi_0}|^2.
\]

By choosing \( \delta, \epsilon \leq \epsilon_0 \), we finally get that
\[
\int_{B_{2^4}(x) \cap L} |D \partial_i u_{\xi_0}|^2 \leq c \tau^\alpha + c \int_{B_{4}(x) \cap B_{2^4}(x) \cap L} |D \partial_i u_{\xi_0}|^2.
\] (3.60)

By adding \( c \) times the left hand side to both sides (again "hole-filling"), we get
\[
\int_{B_{2^4}(x) \cap L} |D \partial_i u_{\xi_0}|^2 \leq c \tau^\alpha + \gamma \int_{B_{4}(x) \cap L} |D \partial_i u_{\xi_0}|^2,
\]
where \( \gamma \in (0, 1) \) is an universal constant. After summing over \( l = 1, 2 \) it follows from Lemma 5.7 that there exists a \( \beta \in (0, 1) \) such that
\[
\int_{B_{2^4}(x) \cap L} |D^2 u_{\xi_0}|^2 \leq c \tau^\beta
\] (3.61)
for all \( x \in B_{\rho_0 \frac{|\xi_0|}{4096}} (\xi_0) \cap L \) and \( \tau < \theta \frac{\rho_0}{4096} \). Now it follows from Morrey’s Lemma (see [GT01], Theorem 7.19) that
\[
u_{\xi_0} \in C^{1,\frac{\alpha}{\beta}}(B_{\rho_0 \frac{|\xi_0|}{4096}} (\xi_0) \cap L, L^2),
\]
and the Lemma is proved. \( \square \)

Therefore we have up to now shown that our limit varifold \( \mu \) can be written as a \( C^{1,\alpha} \cap W^{2,2} \)-graph away from the bad points.

Now we will handle the bad points \( \mathcal{B}_\varepsilon \) and prove a similar power decay as in Lemma 3.7 for balls around the bad points. From this decay it will follow that the set of bad points is actually empty. Since the bad points are discrete and since we want to prove a local decay, we assume that there is only one bad point \( \xi \).

As mentioned in the definition of the bad points (see (3.13)), the Radon measures \( \alpha_k = \mu_k |A_k|^2 \) converge weakly to a Radon measure \( \alpha \), and it follows for all \( \varepsilon \in \mathbb{R}^3 \) that \( \alpha(B_\rho(z) \setminus \{ z \}) \to 0 \) for \( \rho \to 0 \). Therefore for given \( \varepsilon > 0 \) there exists a \( \rho_0 > 0 \) such that
\[
\alpha(B_\rho(\xi) \setminus \{ \xi \}) < \varepsilon^2 \quad \text{for all} \; \rho \leq \rho_0.
\]
Since \( a_k \to a \) in \( C^0_c(\mathbb{R}^3) \), it follows that for \( \rho < \rho_0 \) and \( k \) sufficiently large
\[
\int_{\Sigma_k \cap B_{\rho}(\xi) \setminus B_{\frac{\rho}{2}}(\xi)} |A_k|^2 \, d\mathcal{H}^2 < \epsilon^2. \tag{3.62}
\]
Moreover it follows from Theorem 5.3 applied to our minimizing sequence \( \Sigma_k \) and (3.2) that for all \( \sigma > 0 \),
\[
\int_{\mathbb{R}^3 \setminus B_{\rho\sigma}(\xi)} \left| \frac{1}{4} \bar{H}_k(x) + \frac{(x - \xi)^+}{|x - \xi|^2} \right|^2 \, d\mu_k(x) \leq \frac{1}{4\pi} \mathcal{W}(\Sigma_k) - \sigma^2(\mu_k, \xi) \leq c,
\]
where \( c \) is a universal constant independent of \( k \) and \( \sigma \). Here \( \perp \) denotes the projection onto \( T_{x_k} \Sigma_k \). Rewriting the left hand side and using Cauchy-Schwarz we get
\[
\int \chi_{\mathbb{R}^3 \setminus B_{\rho\sigma}(\xi)} \frac{|(x - \xi)^+|^2}{|x - \xi|^4} \, d\mu_k(x) \leq c,
\]
where again \( c \) is a universal constant independent of \( k \) and \( \sigma \). Now we can use the Monotone Convergence Theorem to get for \( \sigma \to 0 \) that
\[
f_k(x) = \frac{|(x - \xi)^+|^2}{|x - \xi|^4} \in L^1(\mu_k) \quad \text{and} \quad \int f_k \, d\mu_k \leq c,
\]
where \( c \) is an universal constant. Now define the Radon measures
\[
\beta_k = f_k \ll \mu_k.
\]
It follows that \( \beta_k(\mathbb{R}^3) \leq c \), and therefore (after passing to a subsequence) there exists a Radon measure \( \beta \) such that \( \beta_k \to \beta \) in \( C^0_c(\mathbb{R}^3) \). Moreover \( \beta(B_{\rho}(\xi) \setminus \{\xi\}) \to 0 \) for \( \rho \to 0 \). Therefore there exists a \( \rho_0 > 0 \) such that
\[
\beta(B_{\rho}(\xi) \setminus \{\xi\}) < \epsilon^4 \quad \text{for all} \quad \rho \leq \rho_0.
\]
Let \( \rho < \rho_0 \) and \( g \in C^0_c(\mathbb{R}^3 \setminus B_{\rho}(\xi) \setminus \{\xi\}) \) such that \( 0 \leq g \leq 1 \) and \( g \geq \chi_{B_{\rho}(\xi) \setminus B_{\frac{\rho}{2}}(\xi)} \). It follows that
\[
\int \chi_{B_{\rho}(\xi) \setminus B_{\frac{\rho}{2}}(\xi)} \frac{|(x - \xi)^+|^2}{|x - \xi|^4} \, d\mu_k(x) \leq \int g \, d\beta_k \to \int g \, d\beta \leq \beta(B_{\rho}(\xi) \setminus \{\xi\}) < \epsilon^4.
\]
Thus we get for \( k \) sufficiently large that
\[
\int_{\Sigma_k \cap B_{\rho}(\xi) \setminus B_{\frac{\rho}{2}}(\xi)} \frac{|(x - \xi)^+|^2}{|x - \xi|^4} \, d\mathcal{H}^2(x) \leq \epsilon^4.
\]
Now let \( B_k = \left\{ x \in \Sigma_k \cap B_{\rho}(\xi) \left| \frac{|x - \xi|^+}{|x - \xi|} > \epsilon \right. \right\} \). It follows from the Monotonicity formula that for \( \rho < \rho_0 \) and \( k \) sufficiently large
\[
eq \mathcal{H}^2 \left( \Sigma_k \cap B_{\rho}(\xi) \setminus B_{\frac{\rho}{2}}(\xi) \cap B_k \right) \leq \mathcal{H}^2 \left( \Sigma_k \cap B_{\rho}(\xi) \right) \sqrt{\int_{\Sigma_k \cap B_{\rho}(\xi) \setminus B_{\frac{\rho}{2}}(\xi) \cap B_k} \frac{|(x - \xi)^+|^2}{|x - \xi|^2} \, d\mathcal{H}^2(x)} \]
\[
\leq \mathcal{H}^2 \left( \Sigma_k \cap B_{\rho}(\xi) \right) \sqrt{\frac{1}{\epsilon^4} \int_{\Sigma_k \cap B_{\rho}(\xi) \setminus B_{\frac{\rho}{2}}(\xi) \cap B_k} \frac{|(x - \xi)^+|^2}{|x - \xi|^2} \, d\mathcal{H}^2(x)} \]
\[
\leq c \epsilon^2 \rho^2. \tag{3.63}
\]
Moreover by choosing \( \rho_0 \leq \frac{2}{3 \sqrt{8\pi}} \), we get for \( \rho < \rho_0 \) and \( k \) sufficiently large that
\[
\Sigma_k \cap \partial B_{\frac{3}{2}\rho}(\xi) \neq \emptyset.
\] (3.64)

To prove this notice that due to the diameter estimate in Lemma 1.1 in [Sim93] we have
\[
diam \Sigma_k \geq \sqrt{\frac{\mathcal{H}^2(\Sigma_k)}{W(\Sigma_k)}} \geq \sqrt{\frac{1}{8\pi}}.
\]

Let \( \xi_k \in \Sigma_k \) such that \( \xi_k \to \xi \). It follows that \( \Sigma_k \cap B_{\frac{3}{2}\rho}(\xi) \neq \emptyset \) for \( k \) sufficiently large. Now suppose that \( \Sigma_k \cap \partial B_{\frac{3}{2}\rho}(\xi) = \emptyset \). Since \( \Sigma_k \) is connected, we get that \( \Sigma_k \subseteq B_{\frac{3}{2}\rho}(\xi) \), and therefore \( diam \Sigma_k \leq \frac{3}{2}\rho < \frac{3}{4}\rho_0 \leq \frac{1}{\sqrt{8\pi}} \), a contradiction.

After all according to (3.62)-(3.64) the following is shown: There exists a \( \rho_0 > 0 \) such that for \( \rho < \rho_0 \) and \( k \) sufficiently large we have that
\[
\begin{align*}
(i) \quad & \int_{\Sigma_k \cap B_{\rho}(\xi) \setminus B_{\rho}(\xi)} |A_k|^2 \, d\mathcal{H}^2 < \varepsilon^2, \\
(ii) \quad & \frac{|x - \xi|^2}{|x - \xi|^2} \leq \varepsilon \quad \text{for all } x \in \left( \Sigma_k \cap B_{\rho}(\xi) \setminus B_{\rho}(\xi) \right) \setminus B_k, \\
& \text{where } B_k \subseteq \Sigma_k \cap B_{\rho}(\xi) \text{ with } \mathcal{H}^2 \left( \Sigma_k \cap B_{\rho}(\xi) \setminus B_{\rho}(\xi) \cap B_k \right) \leq c\varepsilon \rho^2 \\
& \text{and } (x - \xi)^+ = (x - \xi) - P_{\mathcal{L},\Sigma}(x - \xi), \\
(iii) \quad & \Sigma_k \cap \partial B_{\frac{3}{2}\rho}(\xi) \neq \emptyset.
\end{align*}
\]

Let \( z_k \in \Sigma_k \cap \partial B_{\frac{3}{2}\rho}(\xi) \). It follows that
\[
\int_{\Sigma_k \cap B_{\rho}(\xi)} |A_k|^2 \, d\mathcal{H}^2 \leq \int_{\Sigma_k \cap B_{\rho}(\xi) \setminus B_{\rho}(\xi)} |A_k|^2 \, d\mathcal{H}^2 \leq \varepsilon^2.
\]

The Monotonicity formula applied to \( \Sigma_k \) and \( z_k \) yields that we may apply the Graphical Decomposition Lemma to \( \Sigma_k \), \( z_k \) and infinitely many \( k \) as well as Lemma 1.4 in [Sim93], to get as in Lemma 3.3 that there exists a \( \theta \in \left( 0, \frac{1}{2} \right) \) (independent of \( k \)) and pairwise disjoint subsets \( P_1^k, \ldots, P_{N_k}^k \subset \Sigma_k \) such that
\[
\Sigma_k \cap \overline{B_{\theta \rho_k}(z_k)} = \left( \text{graph } u_k \cup \bigcup_n P_n^k \right) \cap \overline{B_{\theta \rho_k}(z_k)},
\]
where the following holds:

1. \( \{ u_k \} \) are closed topological discs disjoint from graph \( u_k \).
2. \( u_k \in C^\infty(\overline{\Sigma_k}, L^+_{\theta \rho_k}) \), where \( L_k \subset \mathbb{R}^3 \) is a 2-dim. plane such that \( z_k \in L_k \) and \( \Omega_k = (B_{\rho_k}(z_k) \cap L_k) \setminus \bigcup_m d_{k,m} \), where \( \rho_k > \frac{\varepsilon}{\rho_0} \) and where the sets \( d_{k,m} \) are pairwise disjoint closed discs in \( L_k \).
3. The following inequalities hold:
\[
\sum_m \text{diam } d_{k,m} + \sum_n \text{diam } P_n^k \leq c\varepsilon \rho.
\]
Define the perturbed 2-dim. plane \( \tilde{\Omega} \) by
\[
\|u_k\|_{L^n(\Omega_k)} \leq c e^{\frac 12} \rho + \delta_k, \quad \text{where } \lim_{k \to \infty} \delta_k = 0,
\]
and
\[
\|D u_k\|_{L^n(\Omega_k)} \leq c e^{\frac 12} + \delta_k, \quad \text{where } \lim_{k \to \infty} \delta_k = 0.
\]

In the next step we show that
\[
\text{dist}(\xi, L_k) \leq c(e^{\frac 12} + \delta_k) \rho. \tag{3.65}
\]

To prove this notice first of all that it follows from Theorem 5.3 applied to \( \Sigma_k \) and \( z_k \) and (i) above that for \( \varepsilon \leq \varepsilon_0 \)
\[
\mathcal{H}^2(\Sigma_k \cap B_{\rho_k}(z_k)) \geq c \rho^2. \tag{3.66}
\]

Now to prove (3.65) notice that
\[
\left( \text{graph } u_k \cap B_{\rho_k}(z_k) \right) \setminus B_k \neq \emptyset,
\]
where \( B_k \subset \Sigma_k \cap B_{\rho_k}(\xi) \) is the set in (ii) above. This follows from the graphical decomposition above, the diameter estimates for the sets \( P_k \), the area estimate concerning the set \( B_k \) in (ii) and (3.66).

Let \( z \in \left( \text{graph } u_k \cap B_{\rho_k}(z_k) \right) \setminus B_k \subset \left( \Sigma_k \cap B_{\rho}(\xi) \setminus B_{\rho}(\xi) \right) \setminus B_k \). It follows from (ii) that
\[
|\xi - \pi_{(z, T \Sigma_k)}(\xi)| \leq \varepsilon|z - \xi| \leq \varepsilon (|z - z_k| + |z_k - \xi|) \leq c e \rho.
\]

Define the perturbed 2-dim. plane \( L_k \) by \( L_k = L_k + (z - \pi_{L_k}(z)) \), where we have that
\[
\text{dist}(\tilde{L}_k, L_k) = |z - \pi_{L_k}(z)| \leq c e^{\frac 12} \rho \quad \text{(since } z \in \text{graph } u_k \cap B_{\rho_k}(z_k)).
\]
Now it follows from Pythagoras that
\[
|z - \pi_{L_k}(\pi_{(z, T \Sigma_k)}(\xi))|^2 \leq |z - \pi_{(z, T \Sigma_k)}(\xi)|^2 \leq |z - \xi|^2 \leq c \rho^2.
\]

Since \( z + T \Sigma_k \) can be parametrized in terms of \( D u_k(z) \) over \( L_k \), we get that
\[
|\pi_{(z, T \Sigma_k)}(\xi) - \pi_{L_k}(\pi_{(z, T \Sigma_k)}(\xi))| \leq \|D u_k\|_{L^n} |z - \pi_{L_k}(\pi_{(z, T \Sigma_k)}(\xi))| \leq c(e^{\frac 12} + \delta_k) \rho.
\]

Now (3.65) follows since
\[
\text{dist}(\xi, L_k) = |\xi - \pi_{L_k}(\xi)| \leq |\xi - \pi_{L_k}(\pi_{(z, T \Sigma_k)}(\xi))| \leq |\xi - \pi_{L_k}(\pi_{(z, T \Sigma_k)}(\xi))| + |\pi_{L_k}(\pi_{(z, T \Sigma_k)}(\xi)) - \pi_{L_k}(\pi_{(z, T \Sigma_k)}(\xi))| \leq c(e^{\frac 12} + \delta_k) \rho.
\]

Because of (3.65) we may (after translation) assume that \( \xi \in L_k \) for all \( k \), keeping the estimates for \( u_k \). Moreover we again have that \( L_k \to L = 2 \)-dimensional plane with \( \xi \in L \). Therefore for \( k \) sufficiently large we may assume that \( L_k \) is a fixed 2-dimensional plane \( L \).

Define the set
\[
T_k = \left\{ \tau \in \left( \theta \frac{\rho}{64}, \theta \frac{\rho}{\sqrt{2} \cdot 32} \right) \mid \partial B_{\tau}(z_k) \cap \bigcup_m d_{k,m} = \emptyset \right\}.
\]
It follows from the diameter estimates and the selection principle in \([Sim93]\) that for \( \varepsilon \leq \varepsilon_0 \) there exists a \( \tau \in \left( \theta \frac{\rho}{64}, \theta \frac{\rho}{\sqrt{2} \cdot 32} \right) \) such that \( \tau \in T_k \) for infinitely many \( k \).
Since \( \xi \in L \), it follows from the choice of \( \tau \) that for \( \epsilon \leq \epsilon_0 \)
\[
\partial B_{2\rho}(\xi) \cap \partial B_T(z_k) \cap L = \{ p_{1,k}, p_{2,k} \},
\]
where \( p_{1,k}, p_{2,k} \in \left( B_{2\rho}(z_k) \cap L \right) \setminus \bigcup_m d_{k,m} \) are distinct points.

Define the image points \( z_{i,k} \in \text{graph} \, u_k \) by
\[
z_{i,k} = p_{i,k} + u_k(p_{i,k}).
\]

Using the \( L^\infty \)-estimates for \( u_k \), we get for \( \epsilon \leq \epsilon_0 \) that \( \frac{\sqrt{2}}{3} \rho < |z_{i,k} - \xi| < \frac{\sqrt{2}}{5} \rho \), and therefore

\[
\int_{\Sigma \cap B_{\rho}(z_k)} |A_k|^2 \, d\mathcal{H}^2 \leq \int_{\Sigma \cap B_\rho(\xi) \setminus B_{\rho}(\xi)} |A_k|^2 \, d\mathcal{H}^2 < \epsilon^2.
\]

Therefore we can again apply the Graphical Decomposition Lemma to the points \( z_{i,k} \). Thus there exist pairwise disjoint subsets \( P_{1,k}^{i,k}, \ldots, P_{N_{ik}}^{i,k} \subset \Sigma \) such that
\[
\Sigma_k \cap \overline{B_{\rho}(z_{i,k})} = \left( \text{graph} \, u_{i,k} \cup \bigcup_n P_n^{i,k} \right) \cap \overline{B_{\rho}(z_{i,k})},
\]
where the following holds:

1. The sets \( P_n^{i,k} \) are closed topological discs disjoint from \( \text{graph} \, u_{i,k} \).

2. \( u_{i,k} \in C^\infty(\overline{\Omega}_{i,k}, L_{i,k}^1) \), where \( L_{i,k} \subset \mathbb{R}^3 \) is a 2-dim. plane such that \( z_{i,k} \in L_{i,k} \)

and \( \Omega_{i,k} = \left( B_{\lambda_{i,k}}(z_{i,k}) \cap L_{i,k} \right) \setminus \bigcup_m d_{i,k,m} \), where \( \lambda_{i,k} > \frac{\rho}{16} \) and where the sets \( d_{i,k,m} \) are pairwise disjoint closed discs in \( L_{i,k} \).

3. The following inequalities hold:
\[
\sum_m \text{diam} \, d_{i,k,m} + \sum_n \text{diam} \, P_n^{i,k} \leq c \epsilon^\frac{1}{3} \rho,
\]
\[
\|u_{i,k}\|_{L^\infty(\Omega_{i,k})} \leq c \epsilon^\frac{1}{2} \rho + \delta_{i,k}, \quad \text{where} \quad \lim_{k \to \infty} \delta_{i,k} = 0,
\]
\[
\|D \, u_{i,k}\|_{L^\infty(\Omega_{i,k})} \leq c \epsilon^\frac{1}{2} + \delta_{i,k}, \quad \text{where} \quad \lim_{k \to \infty} \delta_{i,k} = 0.
\]

Since \( \text{dist}(z_{i,k}, L) \leq c \epsilon^\frac{1}{2} \rho + \delta_k \) (this follows since \( z_{i,k} \in \text{graph} \, u_k \)), and since the \( L^\infty \)-norms of \( u_k \) and \( u_{i,k} \) are small, we may assume (after translation and rotation as done before) that \( L_{i,k} = L \).

By continuing with this procedure we get after a finite number of steps, depending not on \( \rho \) and \( k \), an open cover of \( \partial B_{\rho}(\xi) \cap L \), which also covers the set
\[
B = \left\{ x \in L \mid \text{dist} \left( x, \partial B_{\rho}(\xi) \cap L \right) < \frac{\rho}{\sqrt{2} \cdot 64} \right\},
\]
and which includes finitely many, closed discs \( d_{k,m} \) with
\[
\sum_m \text{diam} \, d_{k,m} \leq c \epsilon^\frac{1}{2} \rho.
\]
We may assume that these discs are pairwise disjoint, since otherwise we can exchange two intersecting discs by one disc whose diameter is smaller than the sum of the diameters of the intersecting discs.

Because of the diameter estimate and again the selection principle there exists a \(\tau \in \left(\theta \frac{\rho}{128}, \theta \frac{\rho}{\sqrt{2} \cdot 64}\right)\) such that
\[
\{ x \in L \mid \text{dist} \left( x, \partial B^{A_k(\epsilon)}_\rho(\xi) \cap L \right) = \tau \} \cap \bigcup_m d_{k,m} = \emptyset.
\]

Finally we get the following: There exist pairwise disjoint subsets \(P^k_n, \ldots, P^k_{N_k} \subset \Sigma_k\) such that
\[
\Sigma_k \cap \mathcal{A}(\rho) = \left( \text{graph } u_k \cup \bigcup_n P^k_n \right) \cap \mathcal{A}(\rho),
\]
where the following holds:

1. The sets \(P^k_n\) are closed topological discs disjoint from graph \(u_k\).
2. \(u_k \in C^\infty(A_k(\rho), L^\perp)\), where \(L \subset \mathbb{R}^3\) is a 2-dim. plane with \(\xi \in L\).
3. The set \(A_k(\rho)\) is given by
\[
A_k(\rho) = \left\{ x \in L \mid \text{dist} \left( x, \partial B^{A_k(\epsilon)}_\rho(\xi) \cap L \right) < \tau \right\} \setminus \bigcup_m d_{k,m},
\]
where \(\tau \in \left(\theta \frac{\rho}{128}, \theta \frac{\rho}{\sqrt{2} \cdot 64}\right)\) and where the sets \(d_{k,m}\) are pairwise disjoint closed discs in \(L\) which do not intersect \(\{ x \in L \mid \text{dist} \left( x, \partial B^{A_k(\epsilon)}_\rho(\xi) \cap L \right) = \tau \}\).
4. The set \(\mathcal{A}(\rho)\) is given by
\[
\mathcal{A}(\rho) = \left\{ x + y \in \mathbb{R}^3 \mid x \in L, \text{dist} \left( x, \partial B^{A_k(\epsilon)}_\rho(\xi) \cap L \right) < \tau, y \in L^\perp, |y| < \theta \frac{\rho}{64} \right\}.
\]
5. The following inequalities hold:
\[
\sum_m \text{diam } d_{k,m} + \sum_n \text{diam } P^k_n \leq c e^{\frac{1}{4}} \rho,
\]
\[
\|u_k\|_{L^\infty(A_k(\rho))} \leq c e^{\frac{1}{4}} \rho + \delta_k, \quad \text{where } \lim_{k \to \infty} \delta_k = 0,
\]
\[
\|D u_k\|_{L^\infty(A_k(\rho))} \leq c e^{\frac{1}{4}} \rho + \delta_k, \quad \text{where } \lim_{k \to \infty} \delta_k = 0.
\]

From the estimates for the function \(u_k\) and the diameter estimates for the sets \(P^k_n\) we also get for \(\varepsilon \leq \varepsilon_0\) and \(k\) sufficiently large that
\[
\Sigma_k \cap \mathcal{A}(\rho) \subset \left\{ x + y \in \mathbb{R}^3 \mid x \in L, \text{dist} \left( x, \partial B^{A_k(\epsilon)}_\rho(\xi) \cap L \right) < \tau, y \in L^\perp, |y| < \theta \frac{\rho}{128} \right\}.
\]

Since \(\Sigma_k \to \Sigma\) in the Hausdorff distance sense, it follows that
\[
\emptyset \neq \Sigma \cap \mathcal{A}(\rho) \subset \left\{ x + y \in \mathbb{R}^3 \mid x \in L, \text{dist} \left( x, \partial B^{A_k(\epsilon)}_\rho(\xi) \right) < \tau, y \in L^\perp, |y| < \theta \frac{\rho}{128} \right\}.
\]
Now we show that for all $\rho < \rho_0$ (after choosing $\rho_0$ smaller if necessary)
\[
\Sigma \cap \mathcal{A}(\rho) \cap B\left(\frac{\rho}{\sqrt{2}+\rho}x, \rho\right)(\xi) \setminus B\left(\frac{\rho}{\sqrt{2}}x, \rho\right)(\xi) = \Sigma \cap B\left(\frac{\rho}{\sqrt{2}+\rho}x, \rho\right)(\xi) \setminus B\left(\frac{\rho}{\sqrt{2}}x, \rho\right)(\xi).
\]
To prove this notice that due to Theorem 5.3
\[
\theta^2(\mu, x) \leq \frac{1}{4\pi} \mathcal{W}(\Sigma) \leq 2 - \frac{\delta_0}{4\pi} \text{ for all } x \in \mathbb{R}^3.
\]
Now assume that our claim is false, i.e. there exists a sequence $\rho_l \to 0$ such that
\[
\left(\Sigma \cap B\left(\frac{\rho}{\sqrt{2}+\rho}x, \rho\right)(\xi) \setminus B\left(\frac{\rho}{\sqrt{2}}x, \rho\right)(\xi) \right) \setminus \mathcal{A}(\rho_l) \neq \emptyset \text{ for all } l.
\]
Since we already know that $\Sigma$ can locally be written as a $C^{1,\alpha} \cap W^{2,2}$-graph away from the bad point $\xi$, we get that $\Sigma \cap B_{2\rho_l}(\xi)$ contains two components $\Sigma_1$ and $\Sigma_2$ such that $\Sigma_1 \cap \Sigma_2 = \{\xi\}$. Since $\Sigma_1$ can locally be written as a $C^{1,\alpha} \cap W^{2,2}$-graph in $B_{\frac{\rho_l}{\sqrt{2}}}(\xi) \setminus \{\xi\}$, we get that $\theta^2(\Sigma_1, x) = 1$ for all $x \neq \xi$, and by upper semicontinuity that $\theta^2(\Sigma_1, \xi) \geq 1$. Therefore it follows that $\theta^2(\mu, \xi) \geq \theta^2(\Sigma_1, \xi) + \theta^2(\Sigma_2, \xi) \geq 2$, a contradiction, and the claim follows.

From this and $\Sigma_k \to \Sigma$ we get for $\rho < \rho_0$ and $k$ sufficiently large that
\[
\Sigma_k \cap \mathcal{A}(\rho) \cap B\left(\frac{\rho}{\sqrt{2}+\rho}x, \rho\right)(\xi) \setminus B\left(\frac{\rho}{\sqrt{2}}x, \rho\right)(\xi) = \Sigma_k \cap B\left(\frac{\rho}{\sqrt{2}+\rho}x, \rho\right)(\xi) \setminus B\left(\frac{\rho}{\sqrt{2}}x, \rho\right)(\xi).
\]
Define the set
\[
C_k = \left\{ s \in \left(0, \frac{\rho}{\sqrt{2}+\rho}\right) \left| \right. \partial B_{\rho}(\xi) \cap L \cap \bigcup_m d_{k,m} = \emptyset \right\}
\]
The diameter estimates for the discs $d_{k,m}$ yield for $\varepsilon \leq \varepsilon_0$ that $\mathcal{L}^1(C_k) \geq \theta^2(\Sigma, \xi)$. The selection principle in [Sim93] yields that there exists a set $C \subset \left(0, \frac{\rho}{\sqrt{2}+\rho}\right)$ with $\mathcal{L}^1(C) \geq \theta^2(\Sigma, \xi)$ and such that every $s \in C$ lies in $C_k$ for infinitely many $k$.

Now define the set
\[
D_k = \left\{ s \in C \left| \right. \int_{\text{graph } u_k \cap B_{\rho}(\xi) \setminus L} |A_k|^2 \text{ d}\mathcal{H}^2 \leq \frac{4096}{\rho} \int_{\Sigma_k \cap \mathcal{A}(\rho)} |A_k|^2 \text{ d}\mathcal{H}^2 \right\}
\]
By a simple Fubini-type argument (as done before) it follows that $\mathcal{L}^1(D_k) \geq \theta(\rho) \frac{\rho}{2096}$, and again by the selection principle there exists a $s \in \left(0, \frac{\rho}{\sqrt{2}+\rho}\right)$ such that $s \in D_k$ for infinitely many $k$. It follows that $u_k$ is defined on the circle $\partial B_{\rho}(\xi) \cap L$, and that graph $u_{k|_{\rho}(\xi)} \cap L$ divides $\Sigma_k$ into two topological discs $\Sigma_1^k, \Sigma_2^k$, one of them, without loss of generality $\Sigma_1^k$, intersecting $B_{\rho}(\xi)$.

From the estimates for the function $u_k$ and the choice of $s$ we have
\[
\text{graph } u_{k|_{\rho}(\xi)} \subset \mathcal{A}(\rho) \cap B\left(\frac{\rho}{\sqrt{2}+\rho}x, \rho\right)(\xi) \setminus B\left(\frac{\rho}{\sqrt{2}}x, \rho\right)(\xi).
\]
From this inclusion and
\[
\Sigma_k \cap \mathcal{A}(\rho) \cap B\left(\frac{\rho}{\sqrt{2}+\rho}x, \rho\right)(\xi) \setminus B\left(\frac{\rho}{\sqrt{2}}x, \rho\right)(\xi) = \Sigma_k \cap B\left(\frac{\rho}{\sqrt{2}+\rho}x, \rho\right)(\xi) \setminus B\left(\frac{\rho}{\sqrt{2}}x, \rho\right)(\xi)
\]
3. Proof of Theorem 1.1

we get that

\[ \Sigma^1_k \subset B\left(\frac{x}{\rho}, \frac{1}{\rho}\right)\mu(\xi), \]

and the Monotonicity formula yields

\[ \mathcal{H}^2(\Sigma^1_k) \leq c\rho^2. \]

According to Lemma 5.6 let \( w_k \in C^\infty\left(B_{\frac{\rho}{\sqrt{2}}, \rho}\xi \right) \cap L^2 \) be an extension of \( u_k \) restricted to \( \partial B_{\sqrt{2}, \sqrt{2} + \theta}(\xi) \cap L \). In view of the estimates for \( u_k \), and therefore for \( w_k \),

we get that

\[ \\text{graph} \ w_k \subset B\left(\frac{x}{\rho}, \frac{1}{\rho}\right)\mu(\xi). \]

Now we can define the surface \( \tilde{\Sigma}_k \) by

\[ \tilde{\Sigma}_k = \Sigma_k \setminus \Sigma^1_k \cup \text{graph} \ w_k. \]

By construction we have that \( \tilde{\Sigma}_k \) is an embedded \( C^{1,1} \)-sphere, which surrounds an open set \( \tilde{\Omega}_k \subset \mathbb{R}^3 \).

The problem is again that \( \tilde{\Sigma}_k \) might not be a comparison surface. But we can do the same correction as done before in Lemma 3.7 to get for all \( \rho \leq \rho_0 \)

\[ \liminf_{k \to \infty} \int_{\Sigma_k \cap B\left(\frac{x}{\rho}, \frac{1}{\rho}\right)\mu(\xi)} |A_k|^2 \, d\mathcal{H}^2 \leq c\rho^\alpha. \]

Thus by definition of the bad points, \( \xi \) could not have been a bad point, and therefore the set of bad points is empty. Thus, according to Lemma 3.8, Lemma 3.9 and Lemma 3.10 we have shown the following:

**Lemma 3.11** If \( \epsilon \leq \epsilon_0 \), there exists an universal constant \( \alpha \in (0, 1) \), such that for every point \( \xi \in \Sigma \) there exists a radius \( \rho_\xi > 0 \), a 2-dimensional plane \( L_\xi \) containing \( \xi \) and a function

\[ u_\xi \in C^{1,\alpha}(B_{\rho_\xi}(\xi) \cap L_\xi, L^1_\xi) \cap W^{2,2}(B_{\rho_\xi}(\xi) \cap L_\xi, L^2_\xi), \]

such that the following holds for all \( \rho \leq \rho_\xi \):

(i) \[ \mu \cdot B_{\rho}(\xi) = \mathcal{H}^2(\text{graph} \ u_\xi \cap B_\rho(\xi)), \]

(ii) \[ \frac{1}{\rho} \|u_\xi\|_{L^\infty(B_{\rho}(\xi) \cap L_\xi)} + \|D_x u_\xi\|_{L^\infty(B_{\rho}(\xi) \cap L_\xi)} \leq c\rho^{\frac{1}{2}}, \]

(iii) \[ \int_{B_{\rho}(\xi) \cap L_\xi} |D^2 u_\xi|^2 \leq c\tau^{2\alpha} \text{ for all } x \in B_{\rho}(\xi) \cap L_\xi \text{ and all } \tau < \frac{\rho}{2}. \]

**Remark 3.12** Since the graph functions representing \( \mu \) are \( C^{1,\alpha} \), it now follows that the 2-density \( \tilde{\theta}^{2}(\mu, \xi) = 1 \) for all \( \xi \in \Sigma \). In addition we therefore get that

\[ \mu = \mathcal{H}^2 \Sigma \quad \text{and} \quad \mathcal{H}^2(\Sigma) = 1. \]

In the next Lemma we prove that \( \Sigma \) is a topological sphere, because then, in view of Remark 5.10, we are able to derive the Euler-Lagrange equation solved by the graph functions, which will finally yield higher regularity.

**Lemma 3.13** \( \Sigma \) is a topological sphere.
**Proof:** First we recall how the graph functions representing \( \Sigma \) were constructed (see the part after the proof of Lemma 3.7). We had according to Lemma 3.3 that for each good point \( \xi \in \Sigma \setminus B_\varepsilon \) and all \( \rho \leq \rho_\varepsilon \)

\[
\Sigma_k \cap \overline{B_\rho(\xi)} = \left( \text{graph } u_k \cup \bigcup_n P_{n,k}^\rho \right) \cap \overline{B_\rho(\xi)},
\]

where \( u_k \in C^{(0)}(\Omega_k, L^+), L \) is a 2-dim. plane with \( \xi \in L, \Omega_k = (B_\lambda(\xi) \cap L) \cup d_{k,m} \) with \( \lambda_k > \rho \), and the sets \( d_{k,m} \subset L \) are pairwise disjoint closed discs.

Then we extended the function \( u_k \) to the whole disc \( B_\lambda(\xi) \cap L \) as in Lemma 5.6 to get a function \( \overline{u}_k \in C^{1,1}(B_\lambda(\xi) \cap L, L^+) \) such that \( \overline{u}_k = u_k \) in \( \Omega_k \). We had that

\[
\| \overline{u}_k \|_{L^\infty(B_\lambda(\xi) \cap L)} \leq c \varepsilon^d \rho + \delta_k \leq c, \\
\| D \overline{u}_k \|_{L^\infty(B_\lambda(\xi) \cap L)} \leq c \varepsilon^{\frac{d}{2}} + \delta_k \leq c.
\]

It followed that (after passing to a subsequence)

\[
\overline{u}_k \to u \quad \text{in } C^0(B_\rho(\xi) \cap L, L^+), \quad \frac{1}{\rho} \| u \|_{L^\infty(B_\rho(\xi) \cap L)} + \| D u \|_{L^\infty(B_\rho(\xi) \cap L)} \leq c \varepsilon^\frac{d}{2}.
\]

In the subsequent Lemmas we proved that \( u \) is the graph function in Lemma 3.11. Observe that we now know that there are no bad points, and therefore the previous construction holds for every \( \xi \in \Sigma \).

For \( \xi \in \Sigma \) define

\[
R_\xi := \sup \left\{ \rho_\varepsilon > 0 \mid \text{Lemma 3.11 and the above holds for } \rho_\varepsilon \right\}.
\]

Since \( \Sigma \) is compact, it follows that

\[
R := \inf \left\{ R_\xi \mid \xi \in \Sigma \right\} > 0. \tag{3.73}
\]

By compactness of \( \Sigma \) there exist \( \{ \xi_1, \ldots, \xi_J \} \subset \Sigma \) such that \( \Sigma \subset \bigcup_{j=1}^J B_{\delta_j}(\xi_j) \). Since \( \Sigma_k \) converges to \( \Sigma \) in the Hausdorff distance sense, we also have for \( k \) sufficiently large that \( \Sigma_k \subset \bigcup_{j=1}^J B_{\delta_j}(\xi_j) \), and it follows from the above that

\[
\Sigma_k \cap \overline{B_{\rho(\xi_j)}} = \left( \text{graph } u_k \cup \bigcup_n P_{n,k}^\rho \right) \cap \overline{B_{\rho}(\xi_j)}.
\]

By the diameter estimates for the sets \( P_{n,k}^{\rho,j} \) and the selection principle in [Sim93], there exists a \( \xi_j \in \left( \xi, \xi_j \right) \) such that \( \partial B_{\rho}(\xi_j) \cap \bigcup_{j=1}^J P_{n,k}^{\rho,j} = \emptyset \) for all \( l \in \{1, \ldots, J\} \), namely \( \partial B_{\rho}(\xi_j) \) does not intersect any of the pimples. Of course we still have that \( \Sigma \subset \bigcup_{j=1}^J B_{\delta_j}(\xi_j) \), and also the graphical decomposition still holds in \( B_{\rho}(\xi_j) \).

First consider \( \Sigma_k \cap B_{\rho}(\xi_1) \). We replace all the pimples \( P_{n,k}^{\rho,1} \) of \( \Sigma_k \cap B_{\rho}(\xi_1) \) with the extension Lemma 5.6 as done in the proof of Lemma 3.7 by graphs of functions with small \( C^1 \)-norms defined on the discs \( d_{k,m} \). Notice that by the choice of \( \rho \) no pimple intersects \( \partial B_{\rho}(\xi_1) \), and we obtain a new embedded \( C^{1,1} \)-sphere \( \Sigma_k^1 \subset \mathbb{R}^3 \) such that

\[
\Sigma_k \setminus B_{\rho}(\xi_1) = \Sigma_k^1 \setminus B_{\rho}(\xi_1), \quad \Sigma_k \cap B_{\rho}(\xi_1) = \text{graph } \overline{u}_k \cap B_{\rho}(\xi_1). \tag{3.74}
\]
This procedure is exactly the same we described above, thus we have
\[ \overline{u}_{k,1} \to u_{\xi_1} \text{ in } C^0(B_p(\xi_1) \cap L_{\xi_1}, L_{\xi_1}^+). \]

Now consider a point \( \xi_j \) such that \( B_{2^j}(\xi_1) \cap B_{2^j}(\xi_j) \neq \emptyset \), without loss of generality we assume \( j = 2 \). Recall that
\[ \Sigma_k \cap B_p(\xi_2) = \left( \text{graph } u_k^2 \cup \bigcup_n p_{kn}^j \right) \cap B_p(\xi_2), \]

where \( u_k^2 \) was defined on an appropriate 2-plane \( L_{\xi_2} \) containing \( \xi_2 \).

Since \( \Sigma_k^1 \cap B_p(\xi_1) \cap B_p(\xi_2) = \text{graph } \overline{u}_{k,1} \cap B_p(\xi_1) \cap B_p(\xi_2) \), because of the \( C^1 \)-estimates for the graph functions \( \overline{u}_{k,1} \) and \( u_k^2 \) and the diameter estimates for the pimples, it follows that \( \Sigma_k^1 \cap B_p(\xi_1) \cap B_p(\xi_2) \) can be written as a graph defined on the plane \( L_{\xi_2} \), satisfying similar estimates. We conclude that
\[ \Sigma_k^1 \cap B_p(\xi_1) \cap B_p(\xi_2) = \text{graph } \overline{u}_{k,2} \cap B_p(\xi_1) \cap B_p(\xi_2), \]

where now the functions \( \overline{u}_{k,2} \) are defined on the planes \( L_{\xi_2} \). From (3.74), the graphical decomposition of \( \Sigma_k \cap B_p(\xi_2) \setminus B_p(\xi_1) \) and the choice of \( \rho \), we can replace the pimples inside \( B_p(\xi_2) \setminus B_p(\xi_1) \) with new graphs as done before, obtaining a new embedded \( C^{1,1} \)-sphere \( \Sigma_k^2 \subset \mathbb{R}^3 \), which can be written as a graph (without pimples) in both balls such that the corresponding graph functions converge uniformly to the graph functions representing \( \Sigma \).

Repeating this procedure finitely many times we obtain embedded \( C^{1,1} \)-spheres \( \overline{\Sigma}_k \subset \mathbb{R}^3 \) such that the following holds:

\[
\begin{align*}
(i) & \quad \overline{\Sigma}_k \subset \bigcup_{j=1}^{J} B_{2^j}(\xi_j), \quad \Sigma_k \cap B_p(\xi_j) = \text{graph } \overline{u}_{k,j} \cap B_p(\xi_j), \\
(ii) & \quad |\overline{u}_{k,j}| \leq ce^{\frac{1}{n}} \rho + \delta_k, \quad |D \overline{u}_{k,j}| \leq ce^{\frac{1}{n}} + \delta_k, \\
(iii) & \quad \overline{u}_{k,j} \to u_{\xi_j} \text{ in } C^0(B_p(\xi_j) \cap L_{\xi_j}, L_{\xi_j}^+), \\
(iv) & \quad \Sigma \subset \bigcup_{j=1}^{J} B_{2^j}(\xi_j), \quad \Sigma \cap B_p(\xi_j) = \text{graph } u_{\xi_j} \cap B_p(\xi_j), \\
(v) & \quad \overline{\Sigma}_k \to \Sigma \text{ in the Hausdorff distance sense.}
\end{align*}
\]

Next we construct for \( k \) sufficiently large a homeomorphism from \( \Sigma \) to \( \overline{\Sigma}_k \), which yields that \( \Sigma \) is a sphere. This is done by projecting \( \Sigma \) onto \( \overline{\Sigma}_k \) and follows a construction of P. Breuning developed in [Breu11].

A natural direction for the projection would be the direction of the normal along \( \Sigma \). But since \( \Sigma \) is only \( C^{1,\alpha} \), this would not yield an injective mapping. Therefore we first of all approximate \( \Sigma \) in the \( C^1 \)-topology by a smooth surface \( \Sigma_j \). We may assume that \( \Sigma_j \subset \bigcup_{j=1}^{J} B_{2^j}(\xi_j) \) and that it can be written in each ball \( B_p(\xi_j) \) as a
smooth graph \( v_j \) defined on \( B_{\rho}(\xi_j) \cap L_{\xi_j} \). By the estimates for the graph functions representing \( \Sigma \), we may assume that
\[
|D v_j| \leq c e^{\frac{1}{\rho}} + \delta,
\]
where \( \delta > 0 \) may be chosen arbitrarily small. Now the direction of projection will be the direction of the normal \( v_j \) along \( \Sigma \), namely we define \( v : \Sigma \to \mathbb{R}^3 \) in the following way: Let \( x \in \Sigma \), thus \( x = (y, u_{\xi_j}(y)) \) for some \( j \in \{1, \ldots, J\} \) and some \( y \in B_{\rho}(\xi_j) \cap L_{\xi_j} \). Then we define
\[
v(x) = v_j(y, v_j(y)) = \frac{(-D v_j(y), 1)}{\sqrt{1 + |D v_j(y)|^2}}.
\]

It follows that \( v \in C^1(\Sigma, \mathbb{R}^3) \) is well-defined. Moreover we have for \( \delta \) sufficiently small that if \( \{\tau_1(x), \tau_2(x)\} \) is an ONB for \( T_x \Sigma \), then \( \{\tau_1(x), \tau_2(x), v(x)\} \) is an ONB for \( \mathbb{R}^3 \). Using this and the compactness of \( \Sigma \), one can prove as in [DoC76] the existence of a tubular neighborhood of \( \Sigma \) in the following sense: there exists a \( \beta > 0 \) such that whenever \( x, y \in \Sigma \), the segments of the lines in direction of \( \nu \) length \( 2\beta \), centered at \( x \) and \( y \), are disjoint. Since \( \Sigma_k \to \Sigma \) in the Hausdorff distance sense, it follows that for \( k \) sufficiently large
\[
\Sigma_k \subset \bigcup_{x \in \Sigma} \left\{ x + t v(x) \mid t \in (-\beta, \beta) \right\}.
\]

From this it will follow that the projection we will define later is bijective.

Now define the function
\[
h : \Sigma \times \mathbb{R} \to \mathbb{R}^3, \quad h(x, t) = x + tv(x).
\]

1.) Let \( x \in \Sigma \cap B_{\rho}(\xi_j) \). Thus \( x = (y, u_{\xi_j}(y)) \) for some \( y \in B_{\rho}(\xi_j) \cap L_{\xi_j} \). By the definition of \( v \) it follows that
\[
h(x, t) = \left\{ y - t \frac{D v_j(y)}{\sqrt{1 + |D v_j(y)|^2}} u_{\xi_j}(y) + \frac{t}{\sqrt{1 + |D v_j(y)|^2}} \right\},
\]
and therefore \( h(x, t) \in L_{\xi_j} \) if and only if \( t = -u_{\xi_j}(y) \sqrt{1 + |D v_j(y)|^2} \). For this \( t \) we get in view of the estimates for \( u_{\xi_j} \) and (3.75) for \( \epsilon \leq \epsilon_0 \) and \( \delta \) sufficiently small that
\[
\left| y - t \frac{D v_j(y)}{\sqrt{1 + |D v_j(y)|^2}} \right| \leq \frac{\delta}{2D v_j(y)}.
\]
Therefore if we denote by \( H(x) \) the intersection of \( h(x, \mathbb{R}) \) with the plane \( L_{\xi_j} \), we have that
\[
H : \Sigma \cap B_{\rho}(\xi_j) \to B_{\frac{\delta}{2D v_j(y)}}(\xi_j) \cap L_{\xi_j}
\]
is a well-defined, continuous function.

2.) Let \( x \in \Sigma \cap B_{\rho}(\xi_j) \). Thus \( x = (y, u_{\xi_j}(y)) \) for some \( y \in B_{\rho}(\xi_j) \cap L_{\xi_j} \). Let \( z \in \Sigma_k \cap B_{\frac{\delta}{2D v_j(y)}}(\xi_j) \). Thus \( z = (w, \overline{u}_{\xi_j}(w)) \) for some \( w \in B_{\frac{\delta}{2D v_j(y)}}(\xi_j) \cap L_{\xi_j} \). Define the function \( g^k_j : \Sigma_k \cap B_{\frac{\delta}{2D v_j(y)}}(\xi_j) \times \mathbb{R} \to \mathbb{R}^3 \) by \( g^k_j(z, t) = z + tv(x) \). It follows that
\[
g^k_j(z, t) = \left\{ w - t \frac{D v_j(y)}{\sqrt{1 + |D v_j(y)|^2}} \overline{u}_{\xi_j}(w) + \frac{t}{\sqrt{1 + |D v_j(y)|^2}} \right\},
\]
and therefore $g^k(z, t) \in L_{\xi_j}$ if and only if $t = -\overline{u}_{k,j}(w) \sqrt{1 + |D v_j(y)|^2}$. For this $t$ we get in view of the estimates for $\overline{u}_{k,j}$ and (3.75) for $\varepsilon \leq \varepsilon_0$, $k$ sufficiently large and $\delta$ sufficiently small that $|w - t \frac{D v_j(y)}{\sqrt{1 + |D v_j(y)|^2}} - \xi_j| \leq \rho$. Therefore if we denote by $G^k_\lambda(z)$ the intersection of $g^k(z, R)$ with the plane $L_{\xi_j}$, we have that

$$G^k_\lambda : \Sigma_k \cap B_{\frac{\rho}{2}}(\xi_j) \to B_{\rho}(\xi_j) \cap L_{\xi_j}$$

is a well-defined, continuous function. Moreover $G^k_\lambda$ is injective, since if $G^k_\lambda(z_1) = G^k_\lambda(z_2)$ for $z_1 = (w_1, \overline{u}_{k,j}(w_1))$, $z_2 = (w_2, \overline{u}_{k,j}(w_2)) \in \Sigma_k \cap B_{\frac{\rho}{2}}(\xi_j)$, we have by definition

$$w_1 - w_2 + D v_j(y)(\overline{u}_{k,j}(w_1) - \overline{u}_{k,j}(w_2)) = 0,$$

and therefore we get in view of (3.75) and the estimates for $\overline{u}_{k,j}$ that

$$|w_1 - w_2| \leq |D v_j(y)||\overline{u}_{k,j}(w_1) - \overline{u}_{k,j}(w_2)| \leq |w_1 - w_2|,$$

which yields that $w_1 = w_2$, and therefore $z_1 = z_2$.

3.) Let $x \in \Sigma \cap B_{\frac{\rho}{2}}(\xi_j)$. Then

$$B_{\frac{\rho}{2}}(\xi_j) \cap L_{\xi_j} \subset G^k_\lambda \left( \Sigma_k \cap B_{\frac{\rho}{2}}(\xi_j) \right).$$

To prove this notice that it follows as above that for $\varepsilon \leq \varepsilon_0$, $k$ sufficiently large and $\delta$ sufficiently small that

$$G^k_\lambda(\xi_j) \in B_{\frac{\rho}{2}}(\xi_j) \cap L_{\xi_j} \quad \text{and} \quad G^k_\lambda \left( \partial(\Sigma_k \cap B_{\frac{\rho}{2}}(\xi_j)) \right) \cap \left( B_{\frac{\rho}{2}}(\xi_j) \cap L_{\xi_j} \right) = \emptyset.$$

The claim now follows from the continuity of $G^k_\lambda$. 

![Diagram for the functions $H$ and $G^k_\lambda$.](image_url)
Now let \( x \in \Sigma \cap B_{\Sigma}(\xi_j) \). Thus \( x = (y, u_{\xi_j}(y)) \) for some \( y \in B_{\Sigma}(\xi_j) \cap L_{\xi_j} \). It follows that
\[
H(x) \in B_{\Sigma}(\xi_j) \cap L_{\xi_j} \subset G^4_1 \left( \Sigma_k \cap B_{\Sigma}(\xi_j) \right),
\]
and therefore there exists a \( z \in \Sigma_k \cap B_{\Sigma}(\xi_j) \) such that \( H(x) = G^4_1(z) \). By definition the affine subspaces \( h(x, R) = x + R \nu(x) \) and \( g^k_{\xi_j}(z, R) = z + R \nu(x) \) intersect in the point \( H(x) = G^4_1(z) \) and are parallel. Thus \( z \in h(x, R) \), and we have proved that the straight line through \( x \in \Sigma \cap B_{\Sigma}(\xi_j) \) in direction \( \nu(x) \) intersects \( \Sigma_k \cap B_{\Sigma}(\xi_j) \) in at least one point.

Now assume that \( h(x, R) \) intersects \( \Sigma_k \cap B_{\Sigma}(\xi_j) \) in two points \( z_1 = (w_1, \overline{u}_{k,j}(w_1)) \), \( z_2 = (w_2, \overline{u}_{k,j}(w_2)) \). We get from the estimates for \( \overline{u}_{k,j} \) that
\[
|\overline{u}_{k,j}(w_1) - \overline{u}_{k,j}(w_2)| \leq (ce^{\ell} + \delta_k)|w_1 - w_2|.
\]
On the other hand it follows from (3.75) that
\[
|w_1 - w_2| \leq \frac{|\overline{u}_{k,j}(w_1) - \overline{u}_{k,j}(w_2)|}{1 - ce^{\ell} - \delta}
\]
For \( \varepsilon \leq \varepsilon_0 \), \( k \) sufficiently large and \( \delta \) sufficiently small we get \( w_1 = w_2 \), and therefore \( z_1 = z_2 \).

Therefore we have shown that the straight line through \( x \in \Sigma \cap B_{\Sigma}(\xi_j) \) in direction \( \nu(x) \) intersects \( \Sigma_k \cap B_{\Sigma}(\xi_j) \) in exactly one point.

Now we are able to define the desired projection \( \phi_k : \Sigma \rightarrow \overline{\Sigma}_k \): Let \( x \in \Sigma \). Therefore \( x \in \Sigma \cap B_{\Sigma}(\xi_j) \) for some \( j \in \{1, \ldots, J\} \). By the above the straight line through \( x \) in direction \( \nu(x) \) intersects \( \Sigma_k \cap B_{\Sigma}(\xi_j) \) in exactly one point \( S_k^j \in \Sigma_k \cap B_{\Sigma}(\xi_j) \). We define \( \phi_k(x) = S_k^j \) and have to show that \( \phi_k \) is well-defined. For that assume that \( x \in \Sigma \cap B_{\Sigma}(\xi_i) \cap B_{\Sigma}(\xi_j) \) for \( i, j \in \{1, \ldots, J\} \). Denote by \( S_k^i \in \Sigma_k \cap B_{\Sigma}(\xi_i) \), \( S_k^j \in \Sigma_k \cap B_{\Sigma}(\xi_j) \) the unique intersection of the straight line through \( x \) in direction \( \nu(x) \) with \( \Sigma_k \cap B_{\Sigma}(\xi_i) \), respectively \( \Sigma_k \cap B_{\Sigma}(\xi_j) \). It follows from (3.77) that
\[
|S_k^i - S_k^j| < 2\beta.
\]
By choosing \( \beta < \frac{\varepsilon}{3} \), it follows that \( S_k^i \in B_{\Sigma}(\xi_i) \). By uniqueness we get \( S_k^i = S_k^j \), and therefore \( \phi_k \) is well-defined.

Finally \( \phi_k \) is continuous by construction and bijective as mentioned before. Since \( \Sigma \) and \( \Sigma_k \) are compact subsets of \( \mathbb{R}^3 \), it follows from a basic fact in topology that \( \phi_k \) is a homeomorphism. Therefore the Lemma is proved.

Up to now we have shown that \( \Sigma \) is an embedded \( C^{1,\alpha} \cap W^{2,2} \)-sphere. By Remark 3.12 we have that \( \mathcal{H}^2(\Sigma) = 1 \), and from (3.15) it follows that \( \Sigma \) surrounds an open set \( \Omega \subset \mathbb{R}^3 \) with volume \( \mathcal{L}^3(\Omega) = \frac{\sigma^3}{6 \sqrt{\varepsilon}} \). It follows that \( \Sigma \) has the correct isoperimetric ratio \( I(\Sigma) = \sigma \). Now (3.6) and Remark 5.10 yield
\[
\mathcal{W}(\Sigma) \leq \inf \left\{ \mathcal{W}(\Sigma) \mid \Sigma \text{ is an embedded } C^1 \cap W^{2,2} \text{-sphere with } I(\Sigma) = \sigma \right\}.
\]
On the other hand we have that the first variation of the isoperimetric ratio of \( \Sigma \) is not equal to 0 as shown in Lemma 3.6. Therefore there exists a Lagrange multiplier.
3. Proof of Theorem 1.1

$\lambda \in \mathbb{R}$ such that for all $\phi \in C_c^\infty((-\varepsilon, \varepsilon) \times \mathbb{R}^3, \mathbb{R}^3)$ with $\phi(0, \cdot) = 0$

$$\frac{d}{dt} \left( \mathcal{W}(\phi_t(\Sigma)) - \lambda \mathcal{I}(\phi_t(\Sigma)) \right)_{|t=0} = 0. \quad (3.82)$$

This Euler-Lagrange equation differs from the equation in [Sim93] just by the additional term $\lambda \frac{d}{dt} \mathcal{I}(\phi_t(\Sigma))_{|t=0}$ coming from the constraint, which actually is a lower order term. Since the graph functions representing $\Sigma$ satisfy due to Lemma 3.11 the same conditions as in [Sim93], we can carry out the same bootstrap argument as in [Sim93] to get that $\Sigma$ is actually smooth. Therefore $\Sigma$ is a minimizer of the Willmore energy in the class $\mathcal{M}_{\sigma}$, and the existence part of Theorem 1.1 is proved.

Last but not least we have to show that the function $\beta$ is continuous and strictly decreasing. For that let $0 < \sigma_0 < 1$. Choose according to the above $\Sigma_0 \in \mathcal{M}_{\sigma_0}$ such that $\mathcal{W}(\Sigma_0) = \beta(\sigma_0)$. As in section 2, the Willmore flow $\Sigma_t$ with initial data $\Sigma_0$ exists smoothly for all times and converges to a round sphere. By a result of Bryant in [Bry84], which states that the only Willmore spheres with Willmore energy smaller than $8\pi$ are round spheres, it follows that $\mathcal{W}(\Sigma_t)$ is strictly decreasing in $t$. Therefore for every $\sigma \in (\sigma_0, 1]$ there exists a surface $\Sigma \in \mathcal{M}_{\sigma}$ with $\mathcal{W}(\Sigma) < \mathcal{W}(\Sigma_0) = \beta(\sigma_0)$, and therefore $\beta(\sigma) < \beta(\sigma_0)$. To prove the continuity notice that the first variation of the isoperimetric ratio of $\Sigma_0$ is not equal to 0 by Lemma 3.6. As in Lemma 3.7, where we corrected the isoperimetric ratio by applying a suitable variation, we can change the isoperimetric ratio of $\Sigma_0$ a little bit, in fact make it a little larger, without changing the $L^2$-norm of the second fundamental form, or equivalently by Gauss-Bonnet the Willmore energy, too much. Therefore we get a new surface $\Sigma \in \mathcal{M}_{\sigma}$ for a slightly larger $\sigma$ such that $|\mathcal{W}(\Sigma) - \mathcal{W}(\Sigma_0)|$ is small. Finally we get from the monotonicity of $\beta$ that

$$|\beta(\sigma) - \beta(\sigma_0)| = \beta(\sigma) - \beta(\sigma_0) \leq |\mathcal{W}(\Sigma) - \mathcal{W}(\Sigma_0)|.$$

This shows that $\beta$ is continuous, and Theorem 1.1 is now completely proved. $\square$
4 Proof of Theorem 1.2

In this section we prove the convergence to a double sphere stated in the introduction in Theorem 1.2. For that let $\sigma_k \in (0, 1)$ such that $\sigma_k \to 0$. Choose according to Theorem 1.1 surfaces $\Sigma_k \in \mathcal{M}_{\sigma_k}$ such that $\mathcal{W}(\Sigma_k) = \beta(\sigma_k) \leq 8\pi$. After scaling and translation we may assume that $0 \in \Sigma_k$ and $\mathcal{H}^2(\Sigma_k) = 1$. As in section 3 it follows (after passing to a subsequence) that

$$
\mu_k = \mathcal{H}^2|_{\Sigma_k} \to \mu \ \text{in} \ C_c^0(\mathbb{R}^3),
$$

where $\mu$ is an integral, rectifiable 2-varifold with compact support, weak mean curvature vector $\bar{H}_\mu \in L^2(\mu)$, and such that $\theta^2(\mu, \cdot) \geq 1$ $\mu$-a.e. and

$$
\mathcal{W}(\mu) \leq \liminf_{k \to \infty} \mathcal{W}(\Sigma_k) = \liminf_{k \to \infty} \beta(\sigma_k) = 8\pi.
$$

The last equation follows from Theorem 1.1. Moreover we get as in section 3 that

$$
\Sigma_k \to \text{spt } \mu \ \text{in the Hausdorff distance sense.}
$$

Define again the bad points $B_k$ with respect to a given $\varepsilon > 0$ as in (3.13).

As before there exist only finitely many bad points, and for every $\xi_0 \in \text{spt } \mu \setminus B_\varepsilon$ there exists a $\rho_0 = \rho_0(\xi_0, \varepsilon) > 0$ such that

$$
\int_{\xi_0 \setminus B_{\rho_0}(\xi_0)} |\mathcal{H}^2| \leq 2\varepsilon^2
$$

for infinitely many $k$.

Let $\xi_0 \in \text{spt } \mu \setminus B_\varepsilon$ and choose a sequence $\xi_k \in \Sigma_k$ such that $\xi_k \to \xi_0$. For $k$ sufficiently large we may apply the Graphical Decomposition Lemma to $\Sigma_k, \xi_k$ and $\rho < \frac{\rho_0}{2}$ to get for $\varepsilon \leq \varepsilon_0$ that there exist pairwise disjoint closed subsets $P^k_1, \ldots, P^k_{N_k}$ of $\Sigma_k$ such that

$$
\Sigma_k \cap B_{\frac{\rho}{2}}(\xi_k) = \left( \bigcup_{j=1}^{J_k} \text{graph } u^k_j \cup \bigcup_{n=1}^{N_k} P^k_n \right) \cap B_{\frac{\rho}{2}}(\xi_k),
$$

where the sets $P^k_n$ are topological discs disjoint from graph $u^k_j, u^k_j \in C^\infty(\overline{\Omega_{k,j}}, L^2_{k,j})$, $\Omega_{k,j} = \left( B_{\lambda_{k,j}}(\xi_k) \cap L_{k,j} \right) \setminus \bigcup_{m=1}^{M_{k,j}} d_{k,j,m}$ with $\lambda_{k,j} > \frac{\rho}{2}$, $L_{k,j}$ is a 2-dim. plane and the sets $d_{k,j,m}$ are pairwise disjoint closed discs in $L_{k,j}$, and such that we have the estimates

$$
\sum_m \text{diam } d_{k,j,m} + \sum_n \text{diam } P^k_n \leq c\varepsilon^2 \rho \quad \text{and} \quad \frac{1}{\rho} \|u^k_j\|_{L^\infty(\Omega_{k,j})} + \|Du^k_j\|_{L^\infty(\Omega_{k,j})} \leq c\varepsilon^2.
$$

We claim that for given $\theta \in (0, 1)$ and $k$ sufficiently large (depending on $\rho, \theta$)

$$
\text{graph } u^k_j \cap B_{\rho \theta}(\xi_k) \neq \emptyset \ \text{for at least two } j \in \{1, \ldots, J_k\}. \quad (4.1)
$$

Suppose this is false. Notice that at least one graph has to intersect $B_{\rho \theta}(\xi_k)$, since $\xi_k \in \Sigma_k$ and because of the diameter estimates for the $P^k_n$. After passing to a subsequence we may assume that

$$
\Sigma_k \cap B_{\rho \theta}(\xi_k) = \left( \text{graph } u_k \cup \bigcup_{n=1}^{N_k} P^k_n \right) \cap B_{\rho \theta}(\xi_k)
$$

for all $k \in \mathbb{N}$.
and \( B_{l_0}(\xi_0) \subset B_{l_0}(\xi_k) \) for all \( k \). Since we are now again in the situation of having only one graph, we may assume that the graph function \( u_k \) is defined on a subset of a fixed 2-dimensional plane \( L \) containing \( \xi_0 \) as done before in the proof of Theorem 1.1. Let \( \chi_k = \chi_{\Omega_k} \), where \( \Omega_k \) is the open set surrounded by \( \Sigma_k \). Since the isoperimetric ratio \( I(\Sigma_k) \to 0 \), it follows that \( \chi_k \to 0 \) in \( L^1 \). Let \( g \in C_c^1(B_{l_0}(\xi_0), \mathbb{R}^3) \). We get that

\[
\int_{\Sigma_k} \langle g, v_k \rangle \, d\mathcal{H}^2 = \int \chi_k \text{div} g \to 0, \tag{4.2}
\]

where \( v_k \) is the outer normal to \( \partial \Omega_k = \Sigma_k \). By assumption we have

\[
\int_{\Sigma_k} \langle g, v_k \rangle \, d\mathcal{H}^2 = \int_{\text{graph } u_k \cap B_{l_0}(\xi_0)} \langle g, v_k \rangle \, d\mathcal{H}^2 + \sum_n \int_{\mu_n \cap B_{l_0}(\xi_0)} \langle g, v_k \rangle \, d\mathcal{H}^2.
\]

The Monotonicity formula and the diameter estimates yield that the second term is bounded by \( c\rho^2 \). Choose \( g = \varphi \nu \), where \( \varphi \in C^1_c(B_{l_0}(\xi_0)) \) such that \( \varphi \geq \chi_{B_{l_0}(\xi_0)} \) and \( \nu \perp L \) with \( |\nu| = 1 \). By exchanging \( \varphi \) by \( -\varphi \) if necessary we get

\[
\int_{\text{graph } u_k \cap B_{l_0}(\xi_0)} \langle g, v_k \rangle \, d\mathcal{H}^2 \geq \int_{\Omega_k} \chi_{B_{l_0}(\xi_0)}(x + u_k(x)).
\]

The diameter estimates for the discs \( d_{k,m} \) and the bounds for \( u_k \) yield for \( \varepsilon \leq \varepsilon_0 \) and \( k \) sufficiently large that \( \int_{\Sigma_k} \langle g, v_k \rangle \, d\mathcal{H}^2 \geq c\rho^2 - c\rho^2 \). In view of (4.2) we arrive for \( \varepsilon \leq \varepsilon_0 \) at a contradiction.

Now let \( \rho < \frac{\varphi_0}{2} \) such that \( \mu(\partial B_{\rho}(\xi_0)) = 0 \), and therefore \( \mu_k(B_{\rho}(\xi_0)) \to \mu(B_{\rho}(\xi_0)) \). Let \( \delta, \theta \in (0, \frac{1}{2}) \). For \( k \) sufficiently large we may assume that \( B_{(1-\delta)\rho}(\xi_k) \subset B_{\rho}(\xi_0) \), and by (4.1)

\[
\text{graph } u^k_1 \cap B_{(1-\delta)\rho}(\xi_k) \neq \emptyset \quad \text{and} \quad \text{graph } u^k_2 \cap B_{(1-\delta)\rho}(\xi_k) \neq \emptyset.
\]

Let \( x^k_j \in \text{graph } u^k_1 \cap B_{(1-\delta)\rho}(\xi_k) \). In view of the diameter estimates for the sets \( P^k_n \) we get

\[
\mu_k(B_{\rho}(\xi_k)) \geq \sum_{j=1}^2 \int_{\Omega_{k,j}} \chi_{B_{(1-\delta)\rho}(\xi^k_j)}(x + u^k_j(x)) - c\rho^2.
\]

Since \( x^k_j \in \text{graph } u^k_1 \cap B_{(1-\delta)\rho}(\xi_k) \), we have that \( x^k_j = \xi^k_j + u^k_1(\xi^k_j) \) with \( \xi^k_j \in L_{k,j} \) such that \( |x^k_j - \pi_{\Omega_k}(\xi^k_j)| \leq \rho(1 - \delta)\rho^2 \). Therefore \( B_{(1-\theta)(1-\delta)\rho}(\xi^k_j) \subset B_{\rho}(\pi_{\Omega_k}(\xi^k_j)) \).

Moreover it follows from the bounds for \( u^k_2 \) that \( \chi_{B_{(1-\rho)(1-\delta)\rho}(\xi^k_j)}(x + u^k_2(x)) = 1 \) if \( |x - z^k_j| < \frac{(1-\theta)(1-\rho)}{1 + c\varepsilon} \rho \). Therefore after all we get in view of the diameter estimates for the discs \( d_{k,m,j} \) that

\[
\mu_k(B_{\rho}(\xi_k)) \geq 2 \left( \frac{(1-\theta)(1-\delta)}{1 + c\varepsilon} \right)^2 \pi \left( \frac{\rho}{2} \right)^2 - c\rho^2 \geq \frac{3}{2} \pi \left( \frac{\rho}{2} \right)^2,
\]

for \( \varepsilon \leq \varepsilon_0 \) and \( \delta, \theta \) sufficiently small. Thus for all \( \xi_0 \in \text{spt } \mu \setminus B_{\rho} \)

\[
\mu(B_{\rho}(\xi_0)) \geq \frac{3}{2} \pi \left( \frac{\rho}{2} \right)^2.
\]

Now since the density exists everywhere by Theorem 5.3, since \( \mu \) is integral and since \( \mu(B_{\rho}) = 0 \) (which follows from the Monotonicity formula), we have shown
Since $\theta^2(\mu, \cdot) \geq 2 \mu$-a.e. Since $\mathcal{W}(\mu) \leq 8\pi$, the Monotonicity formula in Theorem 5.3 yields $2 \leq \theta^2(\mu, \cdot) \leq \frac{1}{4\pi} \mathcal{W}(\mu) \leq 2 \mu$-a.e., and therefore

$$\theta^2(\mu, \cdot) = 2 \quad \mu\text{-a.e.} \quad \text{and} \quad \mathcal{W}(\mu) = 8\pi.$$ 

Now define the new varifold

$$\bar{\mu} = \frac{1}{2}\mu.$$ 

It follows that $\bar{\mu}$ is a rectifiable 2-varifold in $\mathbb{R}^3$ with compact support $\text{spt} \bar{\mu} = \text{spt} \mu$ and weak mean curvature vector $\bar{H}_{\bar{\mu}} = \bar{H}_\mu \in L^2(\bar{\mu})$, such that $\theta^2(\bar{\mu}, \cdot) = 1 \bar{\mu}$-a.e. and $\mathcal{W}(\bar{\mu}) = 4\pi$. The next Lemma yields that $\bar{\mu}$ is a round sphere in the sense that $ar{\mu} = \mathcal{H}^2 \cap \partial B_r(a)$ for some $r > 0$ and $a \in \mathbb{R}^3$. Therefore $\mu$ is a double sphere as claimed and Theorem 1.2 is proved.

**Lemma 4.1** Let $\mu \neq 0$ be a rectifiable 2-varifold in $\mathbb{R}^3$ with compact support and weak mean curvature vector $\bar{H} \in L^2(\mu)$ such that

(i) $\theta^2(\mu, x) = 1$ for $\mu$-a.e. $x \in \mathbb{R}^3$,

(ii) $\mathcal{W}(\mu) = \frac{1}{4} \int |\bar{H}|^2 \, d\mu \leq 4\pi$.

Then $\mu$ is a round sphere, namely $\mu = \mathcal{H}^2 \cap \partial B_r(a)$ for some $r > 0$ and $a \in \mathbb{R}^3$.

**Proof:** From Theorem 5.3 it follows that the density exists everywhere and that $\theta^2(\mu, x) \geq 1$ for all $x \in \text{spt} \mu$. But then Theorem 5.3 yields

$$\mathcal{W}(\mu) = 4\pi \quad \text{and} \quad \theta^2(\mu, x) = 1 \quad \text{for all} \quad x \in \text{spt} \mu.$$  (4.3)

Since $\mu \neq 0$ and $\text{spt} \mu$ compact, it follows from Theorem 5.3 that there exists a $R > 0$ such that $\text{spt} \mu \setminus B_R(x) \neq \emptyset$ for all $x \in \text{spt} \mu$. Let $x_0 \in \text{spt} \mu$. Since $\text{spt} \mu$ is compact, it follows from Theorem 5.3 that $|\bar{H}(x)| = 4 \frac{(x-x_0)^2}{|x-x_0|^2} \leq \frac{8}{R}$ for $\mu$-a.e. $x \in \text{spt} \mu \setminus B_R(x_0)$. On the other hand by choosing $x_1 \in \text{spt} \mu \setminus B_R(x_0)$ it follows that $|\bar{H}(x)| \leq \frac{8}{R}$ for $\mu$-a.e. $x \in \text{spt} \mu \setminus B_R(x_1)$. Since $B_R(x_0) \cap B_R(x_1) = \emptyset$, it follows that $|\bar{H}(x)| \leq \frac{8}{R}$ for $\mu$-a.e. $x \in \text{spt} \mu$, and therefore $\bar{H} \in L^\infty(\mu)$. Using Allard’s Regularity Theorem (Theorem 24.2 in [Sim83]) we see that $\text{spt} \mu$ can locally be written as a $C^{1,\alpha}$-graph $u$ for some $\alpha \in (0, 1)$. As in Lemma 3.9 it follows that $u$ is a weak solution of

$$\sum_{i,j=1}^2 \partial_j \left( \sqrt{\det g} \, g^{ij} \partial_i F \right) = \sqrt{\det g} \, \bar{H} \circ F,$$

where $F(x) = x + u(x)$, $g_{ij} = \delta_{ij} + \partial_i \mu \cdot \partial_j \mu$. Since $\bar{H} \in L^p(\mu)$ for every $p \geq 1$, it follows from a standard difference quotient argument (as for example in [GT01], Theorem 8.8) that $u \in W^{2,p}$ for every $p \geq 1$, and therefore

$$\int_{B_R} |D^2 u|^2 \leq C\rho^p.$$  (4.4)

From a classical result of Willmore [Wil82] and an approximation argument we get

$$\mathcal{W}(\mu) = 4\pi \leq \inf_{\text{smooth} \, \Sigma} \mathcal{W}(\Sigma) = \inf_{\Sigma \in C^1 \cap W^{2,2}} \mathcal{W}(\Sigma).$$

Therefore $\mu$ solves the Euler-Lagrange equation (3.82) (but with $\lambda = 0$, since we do not have any constraints), and it follows similarly that $\text{spt} \mu$ is smooth. Because of (4.3) and due to Willmore [Wil82], $\text{spt} \mu$ must be a round sphere. □
5 Appendix

5.1 The Monotonicity Formula

Following L. Simon [Sim93], we derive here a Monotonicity formula for rectifiable 2-varifolds in $\mathbb{R}^3$ with square integrable weak mean curvature vector as in [KS04]. For this let $\mu$ be a rectifiable 2-varifold with weak mean curvature vector $\mathbf{H} \in L^2(\mu)$. By definition a rectifiable varifold $\mu$ has weak mean curvature vector $\mathbf{H} \in L^1_{\text{loc}}(\mu)$, if

$$\int \text{div}_\mu \phi \, d\mu = - \int \langle \phi, \mathbf{H} \rangle \, d\mu \quad \text{for all } \phi \in C^1_c(\mathbb{R}^3, \mathbb{R}^3),$$

(5.1)

where $\text{div}_\mu \phi(x) = \sum_{i=1}^2 \langle D_i \phi(x), e_i \rangle$ for an orthonormal basis $\{e_1, e_2\}$ of $T_x \mu$.

We use the notation

$$\mathcal{W}(\mu, E) = \frac{1}{4} \int_E |\mathbf{H}|^2 \, d\mu \quad \text{for } E \subset \mathbb{R}^3 \text{ Borel.}$$

**Lemma 5.1** Assume that $\mathbf{H}(x) \perp T_x \mu$ for $\mu$-a.e. $x \in \mathbb{R}^3$. For $x_0 \in \mathbb{R}^3$ and $r > 0$ consider the function

$$g_{x_0}(r) = \frac{\mu(B_r(x_0))}{\pi r^2} + \frac{1}{4\pi} \mathcal{W}(\mu, B_r(x_0)) + \frac{1}{2\pi r^2} \int_{B_r(x_0)} \langle x - x_0, \mathbf{H}(x) \rangle \, d\mu(x).$$

Then $g_{x_0}$ is increasing, and more precisely we have for $0 < \sigma < \rho$ that

$$g_{x_0}(\rho) - g_{x_0}(\sigma) = \int_{B_\rho(x_0) \setminus B_\sigma(x_0)} \frac{1}{4} |\mathbf{H}(x) + \frac{(x - x_0)^t}{|x - x_0|^2}|^2 \, d\mu(x).$$

(5.2)

Here $(x - x_0)^t = (x - x_0) - P_x(x - x_0)$, where $P_x : \mathbb{R}^3 \to T_x \mu$ is the orthogonal projection onto $T_x \mu$.

**Proof:** For $\phi(x) = \varphi(x)(x - x_0)$ with $\varphi \in C^1_c(\mathbb{R}^3)$ we compute

$$\text{div}_\mu \phi(x) = 2\varphi(x) + \langle \nabla \varphi(x), x - x_0 \rangle,$$

where $\nabla \varphi(x) = P_x(\text{grad} \varphi(x))$ and $P_x$ is the orthogonal projection onto $T_x \mu$. Let $0 < \sigma < \rho$ and put

$$\varphi(x) = \begin{cases} \sigma^{-2} - \rho^{-2} & \text{if } |x - x_0| \leq \sigma, \\ |x - x_0|^{-2} - \rho^{-2} & \text{if } \sigma < |x - x_0| < \rho, \\ 0 & \text{else.} \end{cases}$$

Since $\langle \nabla |x - x_0|^2, x - x_0 \rangle = -2\frac{(x - x_0)^t}{|x - x_0|^t} = -2|x - x_0|^2 + 2\frac{|(x - x_0)^t|^2}{|x - x_0|^t}$, we get that

$$\text{div}_\mu \phi(x) = \begin{cases} 2\left(\frac{\sigma^{-2} - \rho^{-2}}{|x - x_0|^t}\right) & \text{if } |x - x_0| \leq \sigma, \\ 2\left(\frac{|(x - x_0)^t|^2}{|x - x_0|^t} - \rho^{-2}\right) & \text{if } \sigma < |x - x_0| < \rho, \\ 0 & \text{else.} \end{cases}$$
Inserting a smooth approximation into the first variation formula (5.1), we obtain for $\mu(\partial B_\rho(x_0)) = \mu(\partial B_\rho(x_0)) = 0$, namely for a.e. $0 < \sigma < \rho$, the identity

$$\int_{B_\sigma(x_0)} (\sigma^{-2} - \rho^{-2}) \, d\mu + \int_{B_\rho(x_0), B_\sigma(x_0)} \frac{|(x - x_0)^2|}{|x - x_0|^4} \, d\mu(x)$$

$$= \frac{1}{2} (\sigma^{-2} - \rho^{-2}) \int_{B_\rho(x_0)} \langle x - x_0, \vec{H}(x) \rangle \, d\mu(x)$$

$$- \frac{1}{2} \int_{B_\rho(x_0), B_\sigma(x_0)} \frac{|(x - x_0)^2|}{|x - x_0|^2} \langle x - x_0, \vec{H}(x) \rangle \, d\mu(x).$$

Using that $\vec{H}$ is perpendicular, we get by rearranging that

$$\frac{\mu(B_\sigma(x_0))}{\sigma^2} + \frac{1}{2\sigma^2} \int_{B_\rho(x_0)} \langle x - x_0, \vec{H}(x) \rangle \, d\mu(x)$$

$$+ \int_{B_\rho(x_0), B_\sigma(x_0)} \frac{|(x - x_0)^2|}{|x - x_0|^4} \, d\mu(x) + \frac{1}{2} \int_{B_\rho(x_0), B_\sigma(x_0)} \frac{(x - x_0)^2}{|x - x_0|^2} \vec{H}(x) \, d\mu(x)$$

$$= \frac{\mu(B_\rho(x_0))}{\rho^2} + \frac{1}{2\rho^2} \int_{B_\rho(x_0)} \langle x - x_0, \vec{H}(x) \rangle \, d\mu(x).$$

Equation (5.2) follows by adding $\frac{1}{\pi} W(\mu, B_\rho(x_0))$ and dividing by $\pi$. Finally, approximating arbitrary $0 < \sigma < \rho$ by appropriate sequences, the Lemma follows in full generality.

The following result is standard if the weak mean curvature vector belongs to $L^p_{\text{loc}}(\mu)$ for some $p > 2$, see [Sim83].

**Theorem 5.2** Assume that $\vec{H}(x) \perp T_{x}\mu$ for $\mu$-a.e. $x \in \mathbb{R}^3$. Then

$$\theta^2(\mu, x) = \lim_{\rho \to 0} \frac{\mu(B_\rho(x))}{\pi \rho^2}$$

exists for all $x \in \mathbb{R}^3$ and the function $\theta^2(\mu, \cdot)$ is upper semicontinuous.

**Proof:** Putting

$$h_{\rho}(r) = \frac{1}{2\pi r^2} \int_{B_\rho(x_0)} \langle x - x_0, \vec{H}(x) \rangle \, d\mu(x)$$

and using the Cauchy-Schwarz inequality, we can estimate

$$|h_{\rho}(r)| \leq \left( \frac{\mu(B_\rho(x_0))}{\pi \rho^2} \right)^\frac{1}{2} \left( \frac{1}{\pi} W(\mu, B_\rho(x_0)) \right)^\frac{1}{2}. \tag{5.3}$$

Moreover it follows for $\epsilon > 0$ that

$$|h_{\rho}(r)| \leq \epsilon \frac{\mu(B_\rho(x_0))}{\pi \rho^2} + \frac{1}{4\pi \epsilon} W(\mu, B_\rho(x_0)).$$

Using Lemma 5.1 we obtain

$$\frac{\mu(B_\sigma(x_0))}{\pi \sigma^2} \leq \frac{\mu(B_\rho(x_0))}{\pi \rho^2} + \frac{1}{4\pi} W(\mu, B_\rho(x_0)) + h_{\rho}(\rho) - h_{\rho}(\sigma)$$

$$\leq (1 + \epsilon) \frac{\mu(B_\rho(x_0))}{\pi \rho^2} + \epsilon \frac{\mu(B_\rho(x_0))}{\pi \sigma^2} + C_\epsilon W(\mu, B_\rho(x_0)).$$
which implies after adapting constants
\[
\frac{\mu(B_{\rho}(x_0))}{\pi \sigma^2} \leq (1 + \varepsilon) \frac{\mu(B_\rho(x_0))}{\pi \rho^2} + C_\varepsilon \mathcal{W}(\mu, B_\rho(x_0)).
\] (5.4)

It follows that \( \limsup_{\rho \to 0} \frac{\mu(B_\rho(x_0))}{\sigma^2} < \infty \), and hence we get in view of (5.3) that
\[
\lim_{\rho \to \infty} h_{x_0}(r) = 0.
\]

Lemma 5.1 implies that \( \theta^2(\mu, x_0) \) exists, and more precisely that
\[
\theta^2(\mu, x_0) = \lim_{\rho \to \infty} g_{x_0}(r).
\] (5.5)

Furthermore we see
\[
\theta^2(\mu, x_0) \leq (1 + \varepsilon) \frac{\mu(B_\rho(x_0))}{\pi \rho^2} + C_\varepsilon \mathcal{W}(\mu, B_\rho(x_0)).
\] (5.6)

For a sequence \( x_j \to x_0 \) we obtain from (5.6)
\[
\frac{\mu(B_{\rho_j}(x_0))}{\pi \rho^2} \geq \limsup_{j \to \infty} \frac{\mu(B_{\rho_j}(x_j))}{\pi \rho^2} \geq \frac{1}{1 + \varepsilon} \limsup_{j \to \infty} \left( \theta^2(\mu, x_j) - C_\varepsilon \mathcal{W}(\mu, B_{\rho_j}(x_j)) \right) \geq \frac{1}{1 + \varepsilon} \limsup_{j \to \infty} \theta^2(\mu, x_j) - C_\varepsilon' \mathcal{W}(\mu, B_{2\rho_j}(x_0)).
\]

Letting first \( \rho \to 0 \) and then \( \varepsilon \to 0 \) proves the upper semicontinuity. \( \square \)

Now we put
\[
\theta^2(\mu, \infty) = \limsup_{\rho \to \infty} \frac{\mu(B_{\rho}(0))}{\pi \rho^2} \quad \text{and} \quad \theta^2(\mu, \infty) = \liminf_{\rho \to \infty} \frac{\mu(B_{\rho}(0))}{\pi \rho^2}.
\]

For \( x_0 \in \mathbb{R}^3 \) and \( r > d := |x_0| \) we have that
\[
\left(1 - \frac{d^2}{r}\right) \frac{\mu(B_{r-d}(0))}{\pi (r-d)^2} \leq \frac{\mu(B_r(0))}{\pi r^2} \leq \left(1 + \frac{d^2}{r}\right) \frac{\mu(B_{r+d}(0))}{\pi (r+d)^2},
\]
which implies that for \( x_0 \in \mathbb{R}^3 \)
\[
\theta^2(\mu, \infty) = \limsup_{\rho \to \infty} \frac{\mu(B_{\rho}(x_0))}{\pi \rho^2} \quad \text{and} \quad \theta^2(\mu, \infty) = \liminf_{\rho \to \infty} \frac{\mu(B_{\rho}(x_0))}{\pi \rho^2}.
\]

Now let us assume that \( \theta^2(\mu, \infty) < \infty \). Taking the lim inf in (5.4) as \( \rho \to \infty \) yields
\[
\frac{\mu(B_\rho(x_0))}{\pi \sigma^2} \leq (1 + \varepsilon) \theta^2(\mu, \infty) + C_\varepsilon \mathcal{W}(\mu),
\]
in particular \( \theta^2(\mu, \infty) < \infty \). Furthermore we estimate for \( 0 < \sigma < \rho \)
\[
|h_{x_0}(\rho)| \leq \frac{1}{2\pi} \int_{B_{\rho}(x_0)} \left| \nabla \right| d\mu + \left( \frac{\mu(B_{\rho}(x_0))}{\pi \rho^2} \right)^{\frac{1}{2}} \left( \frac{1}{\pi} \mathcal{W}(\mu, B_\rho(x_0) \setminus B_\sigma(x_0)) \right)^{\frac{1}{2}}.
\]
Letting first $\rho \to \infty$ and then $\sigma \to \infty$ yields
\[
\lim_{\rho \to \infty} h_{x_0}(\rho) = 0.
\]
Therefore the density $\theta^2(\mu, \infty)$ exists and is finite by Lemma 5.1. Moreover we have
\[
\lim_{\rho \to \infty} g_{x_0}(\rho) = \theta^2(\mu, \infty) + \frac{1}{4\pi} W(\mu).
\]
(5.7)

Finally we get the following summarizing result.

**Theorem 5.3** Assume that $\vec{H}(x) \perp T_x\mu$ for $\mu$-a.e. $x \in \mathbb{R}^3$. Then the density
\[
\theta^2(\mu, x) = \lim_{\rho \to 0} \frac{\mu(\mathcal{B}_\rho(x))}{\pi \rho^2}
\]
exists for all $x \in \mathbb{R}^3$ and the function $\theta^2(\mu, \cdot)$ is upper semicontinuous. Moreover if $\theta^2(\mu, \infty) = 0$, then we have for all $x_0 \in \mathbb{R}^3$ and all $0 < \sigma < \rho$
\[
\mu(\mathcal{B}_\rho(x_0)) \leq c \rho^2, \tag{5.8}
\]
\[
\theta^2(\mu, x_0) \leq c \left( \frac{\mu(\mathcal{B}_\rho(x_0))}{\pi \rho^2} + W(\mu, \mathcal{B}_\rho(x_0)) \right), \tag{5.9}
\]
\[
\int_{\mathcal{B}_\rho(x_0) \setminus \mathcal{B}_\sigma(x_0)} \left| \frac{1}{4} \vec{H}(x) + \frac{(x - x_0)^2}{|x - x_0|^2} \right|^2 \, d\mu(x) \leq \frac{1}{4\pi} W(\mu) - \theta^2(\mu, x_0), \tag{5.10}
\]
where $\perp$ denotes the projection onto $T_x\mu$.

**Proof:** Inequality (5.8) follows by letting $\rho \to \infty$ in (5.4), while (5.10) follows from Lemma 5.1 together with (5.5) and (5.7). □

**Remark 5.4** Brakke proved in chapter 5 of [Br78] that $\vec{H}$ is perpendicular for any integral varifold with locally bounded first variation. Therefore the statements of this section apply to integral varifolds with square integrable weak mean curvature vector.
5.2 The Graphical Decomposition Lemma

Before stating Graphical Decomposition Lemma of Simon, we recall a result of Langer. In [Lan85] he proved a compactness result for surfaces with \( \|A\|_{L^2} \leq \Lambda \) for \( p > 2 \), using that the surfaces are represented as \( C^1 \)-bounded graphs over discs of radius \( r = r(p, \Lambda) > 0 \). This relies on the Sobolev embedding \( W^{2,p}(U) \hookrightarrow C^{1,\alpha}(\overline{U}) \) for \( U \subset \mathbb{R}^2 \) and some \( \alpha > 0 \), which does not hold for \( p = 2 \). The Graphical Decomposition Lemma of Simon is concerned with the borderli ne case \( p = 2 \) and states the existence of an "almost" graphical representation if \( \|A\|_{L^2} \leq \Lambda \), and is proved in [Sim93].

**Theorem 5.5** Let \( \Sigma \subset \mathbb{R}^n \) be a smooth surface. For given \( \xi \in \Sigma \) and \( \rho > 0 \) let

\[
\begin{align*}
(i) \quad & \partial \Sigma \cap \overline{B_\rho(\xi)} = \emptyset, \\
(ii) \quad & \mathcal{H}^2(\Sigma \cap \overline{B_\rho(\xi)}) \leq \beta \rho^2 \quad \text{for some } \beta > 0, \\
(iii) \quad & \int_{\Sigma \cap \overline{B_\rho(\xi)}} |A|^2 \, d\mathcal{H}^2 \leq \varepsilon^2.
\end{align*}
\]

Then there exists a \( \varepsilon_0 = \varepsilon_0(n,\beta) > 0 \) such that if \( \varepsilon \leq \varepsilon_0 \), there exist pairwise disjoint closed subsets \( P_1, \ldots, P_N \) of \( \Sigma \) such that

\[
\Sigma \cap \overline{B_\varepsilon(\xi)} = \left( \bigcup_{j=1}^J \text{graph } u_j \right) \cup \left( \bigcup_{n=1}^N P_n \right) \cap \overline{B_\varepsilon(\xi)},
\]

where the following holds:

1. The sets \( P_n \) are topological discs disjoint from \( \text{graph } u_j \).

2. \( u_j \in C^\infty(\overline{\Omega_j}, \mathbb{R}^2) \), where \( \Omega_j \subset \mathbb{R}^n \) is a 2-dim. plane and the set \( \Omega_j \) is given by \( \Omega_j = (B_{\lambda_j}(\pi_{L_j}(\xi)) \cap L_j) \setminus \bigcup_{j,m} d_{jm}, \) where \( \lambda_j > \frac{\varepsilon}{\rho} \), the sets \( d_{jm} \subset L_j \) are pairwise disjoint closed discs which do not intersect \( \partial B_{\lambda_j}(\pi_{L_j}(\xi)) \cap L_j \), and \( \pi_{L_j} \) denotes the orthogonal projection onto \( L_j \).

3. Let \( \tau \in \overline{\Omega_j} \) such that \( \Sigma \cap \partial B_\varepsilon(\xi) \) is transversal and \( \partial B_\varepsilon(\xi) \cap \left( \bigcup_{n=1}^N P_n \right) = 0 \). Denote by \( \{\Sigma_l\}_{l=1}^L \) the components of \( \Sigma \cap B_\varepsilon(\xi) \) such that \( \Sigma_l \cap \overline{B_\varepsilon(\xi)} \neq \emptyset \). It follows (after renumeration) that

\[
\Sigma_l \cap \overline{B_\varepsilon(\xi)} = D_{\tau,l} = \left( \text{graph } u_l \cup \bigcup_{n=1}^N P_n \right) \cap \overline{B_\varepsilon(\xi)},
\]

and that \( D_{\tau,l} \) is a topological disc.

4. The following inequalities hold:

\[
\begin{align*}
\sum_{m=1}^M \text{diam } d_{jm} & \leq c(n) \left( \int_{\Sigma \cap \overline{B_\rho(\xi)}} |A|^2 \, d\mathcal{H}^2 \right)^{\frac{1}{2}} \rho \leq c(n) \varepsilon^2 \rho, \\
\sum_{n=1}^N \text{diam } P_n & \leq c(n,\beta) \left( \int_{\Sigma \cap \overline{B_\rho(\xi)}} |A|^2 \, d\mathcal{H}^2 \right)^{\frac{1}{2}} \rho \leq c(n,\beta) \varepsilon^2 \rho, \\
\frac{1}{\rho} \|u_j\|_{L^\infty(\Omega_j)} + \|Du_j\|_{L^\infty(\Omega_j)} & \leq c(n) \varepsilon \frac{1}{\rho^{\alpha-\varepsilon}}.
\end{align*}
\]
5.3 Useful Results

In this section we state some useful results we need for the proof of Theorem 1.1. Lemma 5.6 is an extension result adapted to the cut-and-paste procedure we use and was proved in [Schy08].

Lemma 5.6 Let \( L \) be a 2-dim. plane in \( \mathbb{R}^n, x_0 \in \mathbb{L} \) and \( u \in C^\infty (U, L^1) \), where \( U \subset \mathbb{L} \) is an open neighborhood of \( L \cap \partial B_r(x_0) \). Moreover let \( |D u| \leq c \) on \( U \). Then there exists a function \( w \in C^\infty (B_r(x_0), L^1) \) such that

\[
(i) \quad w = u \quad \text{and} \quad \frac{\partial w}{\partial y} = \frac{\partial u}{\partial y} \quad \text{on} \quad \partial B_r(x_0),
\]

\[
(ii) \quad \frac{1}{\rho} \| w \|_{L^\infty(B_r(x_0))} \leq c(n) \left( \frac{1}{\rho} \| u \|_{L^\infty(\partial B_r(x_0))} + \| D u \|_{L^\infty(\partial B_r(x_0))} \right).
\]

\[
(iii) \quad \| D w \|_{L^\infty(B_r(x_0))} \leq c(n) \| D u \|_{L^\infty(\partial B_r(x_0))},
\]

\[
(iv) \quad \int_{B_r(x_0)} |D^2 w|^2 \leq c(n) \rho \int_{\text{graph } u_{|_{\partial B_r(x_0)}}} |A|^2 \, d\mathcal{H}^1.
\]

Proof: After translation and rotation we may assume that \( x_0 = 0 \) and \( L = \mathbb{R}^2 \times \{0\} \). Moreover we may assume that \( \rho = 1 \), the general result follows by scaling.

Let \( \phi \in C^\infty (B_1(0)) \) be a cutoff-function such that \( 0 \leq \phi \leq 1, \phi = 1 \) on \( B_1(0) \), \( \phi = 0 \) on \( B_1(0) \setminus B_2(0) \) and \( |D \phi| + |D^2 \phi| \leq c(n) \), and define the function \( w_1 \in C^\infty (B_1(0)) \) by

\[
w_1(x) = (1 - \phi(x)) u \left( \frac{x}{|x|} \right) + \phi(x) \int_{\partial B_1(0)} u.
\]

It follows that

\[
w_1 = u, \quad \frac{\partial w_1}{\partial y} = 0 \quad \text{on} \quad \partial B_1(0),
\]

\[
\| w_1 \|_{L^\infty(B_1(0))} \leq c(n) \| u \|_{L^\infty(\partial B_1(0))}, \quad \| D w_1 \|_{L^\infty(B_1(0))} \leq c(n) \| D u \|_{L^\infty(\partial B_1(0))}.
\]

Using the Poincaré-inequality we also get

\[
\int_{B_1(0)} |D^2 w_1|^2 \leq c(n) \| u \|_{W^{2,2}(\partial B_1(0))}^2.
\]

Next let \( w_2 \in C^\infty (B_1(0)) \) be the unique solution of the elliptic boundary value problem given by

\[
\Delta w_2 = 0 \quad \text{in} \quad B_1(0), \quad w_2 = \frac{\partial u}{\partial y} \quad \text{on} \quad \partial B_1(0).
\]

The solution \( w_2 \) is explicitly given by

\[
w_2(x) = \frac{1}{2\pi} \int_{\partial B_1(0)} \frac{1 - |x|^2}{|x - y|^2} \frac{\partial u}{\partial y}(y) \, dy.
\]

Using standard estimates it follows that

\[
\| w_2 \|_{L^\infty(B_1(0))} \leq \| D u \|_{L^\infty(\partial B_1(0))}, \quad |D w_2(x)| \leq \frac{6}{1 - |x|^2} \| D u \|_{L^\infty(\partial B_1(0))},
\]

\[
\| w_2 \|_{W^{1,2}(B_1(0))}^2 \leq c(n) \left( \| D u \|_{L^2(\partial B_1(0))}^2 + \| D^2 u \|_{L^2(\partial B_1(0))}^2 \right).
\]
Next let \( w_3 \in C^\infty(\overline{B_1(0)}) \) be given by
\[
  w_3(x) = \frac{1}{2} (|x|^2 - 1) w_2(x).
\]

It follows that
\[
  w_3 = 0, \quad \frac{\partial w_3}{\partial \nu}(x) = \frac{\partial u}{\partial \nu}(x) \quad \text{on } \partial B_1(0),
\]
\[
  \|w_3\|_{L^\infty(B_1(0))} \leq c \|w_2\|_{L^\infty(B_1(0))} \leq c \|D u\|_{L^\infty(\partial B_1(0))},
\]
\[
  \|D w_3\|_{L^\infty(B_1(0))} \leq c \|D u\|_{L^\infty(\partial B_1(0))}.
\]

Moreover
\[
  \Delta w_3(x) = w_2(x) + x \cdot D w_2(x) \quad \text{in } B_1(0).
\]

Using again standard estimates it follows that
\[
  \int_{B_1(0)} |D^2 w_3|^2 \leq c \left( \|D u\|^2_{L^2(\partial B_1(0))} + \|D^2 u\|^2_{L^2(\partial B_1(0))} \right).
\]

Finally define \( w \in C^\infty(\overline{B_1(0)}) \) by
\[
  w(x) = w_1(x) + w_3(x).
\]

The properties of \( w_1 \) and \( w_3 \) yield
\[
  w = u, \quad \frac{\partial w}{\partial \nu} = \frac{\partial u}{\partial \nu} \quad \text{on } \partial B_1(0),
\]
\[
  \|w\|_{L^\infty(B_1(0))} \leq c \left( \|w\|_{L^\infty(\partial B_1(0))} \right),
\]
\[
  \|D w\|_{L^\infty(B_1(0))} \leq c \|D u\|_{L^\infty(\partial B_1(0))},
\]
\[
  \int_{B_1(0)} |D^2 w|^2 \leq c \|D u\|^2_{L^2(\partial B_1(0))}.
\]

By subtracting an appropriate linear function from \( w \), using again the Poincaré-inequality and the assumption \( |D u| \leq c \), we can get a better estimate for the \( L^2 \)-norm of \( D^2 w \), namely
\[
  \int_{B_1(0)} |D^2 w|^2 \leq c \int_{\partial B_1(0)} |D^2 u|^2 \leq c \int_{\text{graph } u_{|\partial B_1(0)}} |A|^2,
\]
and the Lemma is proved. \( \square \)

The second Lemma is a decay result we need to get a power decay for the \( L^2 \)-norm of the second fundamental form.

**Lemma 5.7** Let \( g : (0, b) \to [0, \infty) \) be a bounded function such that
\[
  g(x) \leq \gamma g(2x) + cx^\alpha \quad \text{for all } x \in \left( 0, \frac{b}{2} \right),
\]
where \( \alpha > 0 \), \( \gamma \in (0, 1) \) and \( c \geq 0 \) is a constant. There exists a \( \beta \in (0, 1) \) and a constant \( C = C(\gamma, \alpha, b, \|g\|_{L^\infty(0,b)}) \) such that
\[
  g(x) \leq C x^\beta \quad \text{for all } x \in (0, b).
\]
Proof: First of all we may assume that \( \gamma \neq \left( \frac{1}{2} \right) ^{\alpha} \), since otherwise we may choose a larger \( \gamma \in (0, 1) \) for which the assumptions still hold. Now let \( \beta \in (0, \min(1, \alpha)) \) such that \( \gamma \leq \left( \frac{1}{2} \right) ^{\beta} \), which is possible since \( \gamma \in (0, 1) \). Now let \( x \in \left( \frac{1}{2}, b \right) \) and \( m \in \mathbb{N} \). It follows by induction that

\[
g(2^{-m}x) \leq \gamma^{m}g(x) + \sum_{j=0}^{m-1} c y^{j} \left( 2^{j-m}x \right) ^{\alpha}.
\]

Now we have that, using \( \gamma \neq \left( \frac{1}{2} \right) ^{\alpha} \),

\[
\sum_{j=0}^{m-1} c y^{j} \left( 2^{j-m}x \right) ^{\alpha} = \frac{c}{2^{m}\alpha} \left( \frac{1 - 2^{m}y^{m}}{1 - 2^{m}\gamma} \right) \leq cy^{m} + c \left( 2^{-m}x \right) ^{\alpha} \leq cy^{m} + c \left( 2^{-m}x \right) ^{\beta},
\]

and therefore it follows that (since \( g \) is bounded)

\[
g(2^{-m}x) \leq cy^{m} + c \left( 2^{-m}x \right) ^{\beta} \quad \text{for all } x \in \left( \frac{1}{2}, b \right) \text{ and } m \in \mathbb{N}.
\]

Especially for \( m = 0 \) it follows for \( x \in \left( \frac{1}{2}, b \right) \) that

\[
g(x) \leq c + cx^{\beta} \leq 2c \left( \frac{x}{b} \right) ^{\beta} \leq c x^{\beta} \leq cx^{\beta}.
\]

Now for \( m \geq 1 \) let \( l_{m} = \left( 2^{-(m+1)}b, 2^{-m}b \right) \). For \( y \in l_{m} \) there exists \( x \in \left( \frac{1}{2}, b \right) \) such that \( y = 2^{-m}x \), and therefore we get (notice that \( \gamma^{m} \leq \left( \frac{1}{2} \right) ^{\beta} y^{\beta} \))

\[
g(y) \leq cy^{m} + cy^{\beta} \leq cy^{\beta}.
\]

Therefore the Lemma is proved. \( \Box \)

Next we state a generalized Poincaré-inequality proved by Simon in [Sim93].

**Lemma 5.8** Let \( \mu > 0 \), \( \delta \in \left( 0, \frac{1}{2} \right) \) and \( \Omega = B_{\mu}(0) \setminus E \), where \( E \) is measurable with \( L^{1}(p_{1}(E)) \leq \frac{\mu}{\delta} \) and \( L^{1}(p_{2}(E)) \leq \delta \) where \( p_{1} \) is the projection onto the x-axis and \( p_{2} \) is the projection onto the y-axis. Then for any \( f \in C^{1}(\Omega) \) there exists a point \( (x_{0}, y_{0}) \in \Omega \) such that

\[
\int_{\Omega} |f - f(x_{0}, y_{0})|^{2} \leq C \mu^{2} \int_{\Omega} |Df|^{2} + C \delta \mu \sup_{\Omega} |f|^{2},
\]

where \( C \) is an absolute constant.

The last statement is an approximation result and shows that it does not matter if we minimize among smooth or \( C^{1} \cap W^{2,2} \)-spheres.

**Lemma 5.9** For \( \sigma \in (0, 1) \) define the sets

\[
\mathcal{M}_{\sigma} = \left\{ \Sigma \subset \mathbb{R}^{3} \left| \Sigma \text{ is a smoothly embedded sphere with } I(\Sigma) = \sigma \right. \right\},
\]

\[
\overline{\mathcal{M}}_{\sigma} = \left\{ \Sigma \subset \mathbb{R}^{3} \left| \Sigma \text{ is an embedded } C^{1} \cap W^{2,2} \text{-sphere with } I(\Sigma) = \sigma \right. \right\}.
\]

Let \( \varepsilon > 0 \) be arbitrary. Then for every \( \Sigma \in \overline{\mathcal{M}}_{\sigma} \) there exists a \( \Sigma \in \mathcal{M}_{\sigma} \) such that

\[
|W(\Sigma) - W(\overline{\Sigma})| \leq \varepsilon.
\]
Proof: Let $\Sigma \in M_{r}$. By writing $\Sigma$ locally as $C^1 \cap W^{2,2}$-graphs and approximating (in $C^1 \cap W^{2,2}$) by smooth graphs, it follows that for every $\epsilon > 0$ there exists a smoothly embedded sphere $\Sigma \subset \mathbb{R}^3$ such that

$$|W(\Sigma) - W(\bar{\Sigma})| \leq \epsilon \quad \text{and} \quad |I(\Sigma) - I(\bar{\Sigma})| \leq \epsilon. \quad (5.11)$$

For $\epsilon > 0$ sufficiently small we may assume without loss of generality that

$$I(\Sigma) \in (0, 1). \quad (5.12)$$

After scaling we may moreover assume that $H^2(\Sigma) = 1$.

Let $\Phi : (-\epsilon, \epsilon) \times \mathbb{R}^3 \to \mathbb{R}^3$ be a $C^2$-variation with compact support and define $\Sigma_t = \Phi_t(\Sigma)$ and $X(x) = \partial_t \Phi_t(x)|_{t=0}$. It follows as in (3.21) that

$$\frac{d}{dt} I(\Sigma_t)|_{t=0} = \frac{1}{3} \int_{\Sigma} \left( X, \frac{3}{2} \hat{H} - \frac{6 \sqrt{\pi}}{I(\Sigma)} \nu \right) dH^2,$$

where $\hat{H}$ is the mean curvature vector and $\nu$ the inner normal of $\Sigma$.

By (5.12) it follows as in Lemma 3.6 that there exists a vectorfield $X \in C^\infty(\mathbb{R}^3, \mathbb{R}^3)$ such that

$$\int_{\Sigma} \left( X, \frac{3}{2} \hat{H} - \frac{6 \sqrt{\pi}}{I(\Sigma)} \nu \right) dH^2 \neq 0.$$

Let $\Phi \in C^\infty(\mathbb{R} \times \mathbb{R}^3, \mathbb{R}^3)$ be the flow of the vectorfield $X$ and again $\Sigma_t = \Phi_t(\Sigma)$.

Now we may assume without loss of generality that

$$\frac{d}{dt} I(\Sigma_t)|_{t=0} = c_0 > 0. \quad (5.13)$$

For given $T > 0$ we can moreover estimate the second derivative of the isoperimetric ratio by

$$\sup_{t \in (-T, T)} \left| \frac{d^2}{dt^2} I(\Sigma_t) \right| \leq c. \quad (5.14)$$

Using Taylor’s formula it follows from (5.11), (5.13) and (5.14) that there exists a $t_0 \in (-c\epsilon, c\epsilon)$ such that

$$I(\Sigma_{t_0}) = I(\bar{\Sigma}),$$

and therefore $\Sigma_{t_0} \in M_{r}$. Now we also have that the first variation of the Willmore energy can be estimated by an universal constant, namely for given $T > 0$ we have

$$\sup_{t \in (-T, T)} \left| \frac{d}{dt} W(\Sigma_t) \right| \leq c.$$

It follows from (5.11) that

$$|W(\Sigma_{t_0}) - W(\bar{\Sigma})| \leq |t_0| \sup_{t \in (-\bar{t}_0, t_0)} \left| \frac{d}{dt} W(\Sigma_t) \right| + \epsilon \leq c\epsilon,$$

and the Lemma follows. \qed

Remark 5.10: It follows from Lemma 5.9 that

$$\inf_{\Sigma \in M_{r}} W(\Sigma) = \inf_{\Sigma \in \overline{M}_{r}} W(\Sigma).$$
References


