THE WILLMORE FUNCTIONAL ON LAGRANGIAN TORI:
ITS RELATION TO AREA
AND EXISTENCE OF SMOOTH MINIMIZERS

WILLIAM P. MINICOZZI II

1. INTRODUCTION

Given a compact surface, $\Sigma$, in $\mathbb{R}^4$, the Willmore functional is defined to be

$$F(\Sigma) = \frac{1}{4} \int_{\Sigma} |H|^2 \, d\mu,$$

where $H$ is the mean curvature vector of $\Sigma$, and $d\mu$ is the two-dimensional area measure of $\Sigma$. This functional is invariant under conformal transformations of the ambient space; in particular, it is fixed under dilations. Willmore (see [14]) showed that this is bounded below by $4\pi$ with equality only for the round spheres. Note that the Gauss-Bonnet theorem implies that the Willmore functional on tori is equivalent to the $L^2$ norm of the full second fundamental form.

It is conjectured that the Clifford (product) torus achieves the minimum of the Willmore functional amongst all immersed tori. Since the Clifford torus is lagrangian, it is natural to conjecture that it minimizes the Willmore functional amongst the smaller class of lagrangian tori. Simon (see [8]) has proven the existence and regularity of embedded tori which minimize the Willmore functional in arbitrary codimension. We will obtain the corresponding result in the lagrangian setting.

It is also conjectured that the Clifford torus minimizes area in its Hamiltonian isotopy class. This can be seen as a generalization of the isoperimetric inequality in the plane (which says that the circle minimizes length in its Hamiltonian isotopy class). We will show that, modulo an existence and regularity result for the area problem, this conjecture would follow from the Willmore conjecture for lagrangian tori. Furthermore, if the full Willmore conjecture holds, then, again modulo the existence and regularity result, the Clifford torus...
would be the unique minimizer of area (up to isometry). We also show that the lagrangian cylinder (a round $S^1$ crossed with an interval) minimizes area in its Hamiltonian isotopy class.

In order to obtain the smooth lagrangian torus minimizing the Willmore functional, we use the direct method in the calculus of variations to produce a candidate, and then prove regularity. Specifically, we take a minimizing sequence of lagrangian tori with Willmore integral close to the infimum (which we normalize using the conformal invariance) and we extract a measure-theoretic limit by going to a subsequence. We get some limit measure, $\mu$; it is the support of $\mu$ which we will prove is a smooth lagrangian torus achieving the infimum of the Willmore functional. Because the sequence converges measure-theoretically and has bounded Willmore integral, we necessarily have convergence in the Hausdorff distance sense as well. The uniform bound on the Willmore integral also implies that the mean curvature can blow up in a neighborhood of only finitely many bad points in the limit set. The local regularity proof consists of three stages.

First we show that the limit surface is $C^{1,\alpha} \cap W^{2,2}$; we do this initially only away from the above bad points. Away from the bad points, we can apply Simon's graphical decomposition lemma to get the surfaces as mostly graphs of gradients (with some small exceptional sets). We then construct a competing sequence by gluing in the graphs of the gradients of triharmonic functions; then a triharmonic comparison lemma and the minimizing property of the original sequence imply an $L^2$ decay estimate on the second fundamental form. We then use this estimate and the Poincare inequality to show convergence of the graphs with the Arzela-Ascoli theorem. This then implies multiplicity one varifold convergence of the sequence locally, $C^{1,\alpha}$ regularity of the limit away from bad points, and also that the limit is lagrangian. The $W^{2,2}$ regularity follows since $\Sigma$ is a weak solution to the prescribed mean curvature equation; finally, we use hole-filling to get a decay estimate on the Sobolev norm.

The second stage of the proof consists of showing that any graphical $C^{1,\alpha} \cap W^{2,2}$ weak solution must, in fact, be smooth. We do this in two steps; first we add one order of differentiability to prove that the solution is actually in $C^{2,\gamma} \cap W^{3,2}$, and then it is fairly standard to show, via iterative difference quotient arguments, that the surface must actually be $C^{\infty}$. To get the $W^{3,2}$ result, we uniformly bound the $L^2$ norm of the second derivatives of the difference quotients (which then converge to the third derivatives with a corresponding estimate) by choosing appropriate Hamiltonian variations in the weak form of the Euler-Lagrange equation (which is sixth order and nonlinear). The Holder continuity then follows from a decay estimate on the $W^{3,2}$ norm and Morrey's lemma.

The final stage of the regularity argument involves using the established convergence and Allard's compactness theorem for varifolds to prove lower semicontinuity of the Willmore functional with respect to this convergence, and to justify the application of our previous arguments directly to the limit surface.
The paper is organized as follows. In Section 2 we review background material from geometric analysis and from symplectic geometry. In Section 3 we discuss the relationship between the Willmore functional and area. In Section 4 we state the main existence and regularity results for the lagrangian Willmore problem. Finally, in Sections 5 through 8 we give the proofs.

Acknowledgment. The author wishes to thank Professor Richard M. Schoen for his guidance and support as well as many illuminating discussions.

2. Preliminaries

2.1. Symplectic geometry. Throughout this paper we will be working in $\mathbb{R}^4$, with coordinates $x_i, y_i, i = 1, 2,$ and the standard metric,

$$ds^2 = \sum_{i=1,2} dx_i^2 + dy_i^2,$$

and symplectic structure,

$$\omega = \sum_{i=1,2} dx_i \wedge dy_i.$$

Given two vectors, $u$ and $v$, we will write $(u, v)$ for the inner product of $u$ and $v$. Notice that $\omega(u, v) = (Ju, v)$, where $J$ is the standard complex structure on $\mathbb{R}^4$. We can write $J = d(\alpha)$, where $\alpha = L - y_j dx_j$.

Definition 2.1. Let $\Sigma$ be a surface in $\mathbb{R}^4$, with tangent and normal bundles, $T\Sigma$ and $N\Sigma$ respectively. Then $\Sigma$ is lagrangian if any of the following equivalent conditions are satisfied:

1. $\omega$ restricted to $\Sigma$ is zero;
2. $J(T\Sigma) = N\Sigma$;
3. $\alpha$ restricted to $\Sigma$ is closed.

As a result $\alpha$ determines a cohomology class on each lagrangian surface; if this class is trivial, we say that the surface is exact. It follows from Gromov's work [3] that if $\Sigma$ is a closed surface, then $\alpha$ is a nonvanishing class. Notice that if $\Sigma$ is given as a graph, $y_i = f_i(x_1, x_2)$, then condition (3) implies that

$$y_i = \frac{\partial g}{\partial x_i},$$

for some function $g$ on the $x_1 - x_2$ plane. We will refer to $g$ as the potential or generating function.

Notice that the above definition implies that $J$ gives an isometry between the tangent and normal bundles of a lagrangian surface, $\Sigma$. Furthermore, since $J$ is parallel (i.e., $\mathbb{R}^4$ is Kahler), this isometry commutes with the connection. Therefore, we can convert sections of the normal bundle into sections of the tangent bundle with $J$. For example, if $H$ is the mean curvature, then $JH$ is a tangential vector field; moreover, $H$ is parallel with respect to the normal
connection if and only if $JH$ is parallel with respect to the induced connection on $\Sigma$. Using the metric in the standard way, we can associate $JH$ with a one-form on $\Sigma$; then the Codazzi equations imply that this form is closed (i.e., in normal coordinates the gradient of $JH$ is symmetric). Hence, the mean curvature gives rise to a cohomology class, which turns out to be the Maslov class.

If $a_{ijk}$ denotes the components of the second fundamental form (transformed into a purely tangential tensor, as above) in normal coordinates for $\Sigma$ a lagrangian surface, we will show that the $a_{ijk}$ are fully symmetric. We have the definition

$$a_{ijk} = \langle D_{e_j} e_k, Je_i \rangle,$$

where the $e_i$ form an orthonormal frame for $T\Sigma$ at a point. It is standard that $a_{ijk} = a_{jik}$, so we need only show that $a_{ijk} = a_{ikj}$. Since we are using the induced connection and $J$ is parallel, we get

$$a_{ijk} = -(JD_{e_j} e_k, e_i).$$

Since $J^2 = -1$, and $J$ is orthogonal, this becomes

$$a_{ijk} = \langle D_{e_i} e_k, Je_j \rangle;$$

and, hence, $a_{ijk} = a_{ikj}$ as desired.

For general surfaces in $\mathbb{R}^4$, the oriented Gauss map goes to $S^2 \times S^2$; the first factor may be identified with the set of orthogonal complex structures on $\mathbb{R}^4$. The lagrangian restriction then forces the first factor of the Gauss map to go into the equator; we will call this map $\theta$. We then get a cohomology class, $d\theta$, on each lagrangian surface. It is a basic fact that this class is dual to the mean curvature under the isomorphism coming from the symplectic form. Specifically, $H = JD\theta$, where $D$ here is the gradient on the surface. Notice that this implies that a surface which is both lagrangian and minimal must then be holomorphic with respect to some orthogonal complex structure, $K$. As a consequence, note that the equations for a minimal lagrangian surface are overdetermined.

For our purposes, the basic example of a lagrangian surface is the Clifford torus, $S^1 \times S^1$, which is minimal in the sphere of radius $\sqrt{2}$. In fact, every embedded oriented lagrangian surface without boundary must be a torus topologically (this follows from a simple characteristic class argument). This fact will be important later, as it will allow us to eliminate the possibility of a change in the topological type for a minimizing sequence, as long as we can show that embeddedness is preserved in the limit.

If we wish to preserve the symplectic structure (and hence take lagrangian surfaces to lagrangian surfaces), we must restrict ourselves to Hamiltonian isotopies. These are isotopies generated by Hamiltonian vector fields; namely, the space of allowable variations consists of vector fields of the form

$$X = JDf,$$
where $f$ is an ambient function. Note that there is another group of transformations which take lagrangian surfaces to lagrangian surfaces: namely the conformally symplectic transformations. These are generated by vector fields which send the symplectic form to a multiple of itself (the basic example is the radial, or position, vector field which generates the dilations).

We will now use the above to derive the Euler-Lagrange equation for the area functional on lagrangian surfaces. We saw above that, as a tangential form, $H$ is closed; we will now show that for critical points for area it is also co-closed (and hence harmonic). Given a Hamiltonian variation, $J D f$, the first variation of area must vanish; hence

$$\int \langle H, J D f \rangle = 0.$$ 

Since $J$ is skew-symmetric, we can put it on the left and integrate by parts to get

$$\int f \text{div}(J H) = 0,$$

for all smooth functions $f$ on $\Sigma$. We conclude that $\text{div}(J H) = 0$; i.e., $H$ is co-closed.

2.2. Geometric preliminaries. We will also need some facts from geometry. These may all be found with proofs in [8]. In the following, $\Sigma$ will be a smooth surface, possibly with boundary $\partial \Sigma$. We will always assume that $\Sigma$ contains the origin. Letting $B_\rho$ denote the ball of radius $\rho$ centered at the origin, and $A_{\sigma, \rho}$ the annulus $B_\rho \setminus B_\sigma$, then define

$$\Sigma_\rho \equiv \Sigma \cap B_\rho,$$

$$\Sigma_{\sigma, \rho} \equiv \Sigma \cap A_{\sigma, \rho}.$$

We will let $|\Sigma|$ denote the area of $\Sigma$. Finally, let $\text{diam}(\Sigma)$ denote the diameter of $\Sigma$ as a subset of Euclidean space.

The first variation identity states that

$$(1) \quad \int_{\Sigma} \text{div}_\Sigma(V) = \int_{\partial \Sigma} \eta \cdot V - \int_{\Sigma} V \cdot H,$$

where $H$ is, as always, the mean curvature, $V$ is any $C^1$ vector field, $\text{div}_\Sigma$ is the divergence with respect to $\Sigma$, and $\eta$ is the outward pointing unit conormal for $\partial \Sigma$. This identity remains valid (by an approximation argument) even if the vector field is merely Lipschitz on $\Sigma$.

When $\Sigma$ has no boundary, by choosing appropriate variation vector fields, Simon gets the following identity:

$$(2) \quad \sigma^{-2} |\Sigma_\rho| + \int_{\Sigma_{\sigma, \rho}} \left| \frac{1}{4} H + \frac{X^\perp}{|X|^2} \right|^2 = \rho^{-2} |\Sigma_\rho| + \frac{1}{4} F(\Sigma_\rho) + \frac{1}{2} \int_{\Sigma_\rho} \rho^{-2} X \cdot H - \frac{1}{2} \int_{\Sigma_\rho} \sigma^{-2} X \cdot H.$$
Letting $\sigma$ go to zero, we get

$$\pi \text{Mult}(\Sigma, 0) + \int_{\Sigma} \left| \frac{1}{4} H + \frac{X^\perp}{|X|^2} \right|^2 = \rho^{-2} |\Sigma_p| + \frac{1}{4} F(\Sigma_p) + \frac{1}{2} \int_{\Sigma_p} \rho^{-2} X \cdot H,$$

where $\text{Mult}(\Sigma, 0)$ is the multiplicity of the surface at the origin. If the surface is smooth and embedded there, this is one; if the surface has a self-intersection there, then this is at least two. Now taking the limit as $\rho \to \infty$, we get

$$\pi \text{Mult}(\Sigma, 0) + \int_{\Sigma} \left| \frac{1}{4} H + \frac{X^\perp}{|X|^2} \right|^2 = \frac{1}{4} F(\Sigma).$$

Throwing away the second (positive) term on the left, this shows that $F(\Sigma) \geq 4\pi$, and that if $\Sigma$ is not embedded, then $F(\Sigma) \geq 8\pi$.

We now derive the pseudo-monotonicity estimates. If we apply Cauchy-Schwarz to the right-hand side of equation (2), and throw away the positive term on the left, we get

$$\sigma^{-2} |\Sigma_p| \leq C \left( \rho^{-2} |\Sigma_p| + F(\Sigma_p) \right),$$

and letting $\sigma \to 0$, for $0$ a regular point,

$$\pi \leq C \left( \rho^{-2} |\Sigma_p| + F(\Sigma_p) \right).$$

When $\Sigma$ does have boundary, we get that

$$\pi \text{Mult}(\Sigma, 0) + \int_{\Sigma} \left| \frac{1}{4} H + \frac{X^\perp}{|X|^2} \right|^2 = \frac{1}{4} F(\Sigma) + \frac{1}{2} \int_{\partial \Sigma} \eta \cdot \frac{X}{|X|^2}.$$

We will use a number of consequences of the preceding calculations. These lemmas are due to Simon (see [8]), and their proofs may be found there.

**Lemma 2.2.** If $\partial \Sigma = \emptyset$, where $\Sigma$ is compact and connected, then

$$\sqrt{\frac{|\Sigma|}{F(\Sigma)}} \leq \text{diam} \Sigma \leq C \sqrt{|\Sigma| F(\Sigma)},$$

where $C$ is a constant independent of $\Sigma$.

**Lemma 2.3.** When $\Sigma$ is connected and its closure is compact, then

$$\text{diam} \Sigma \leq C \left( \int_{\Sigma} |A| + \sum_j \text{diam} \Gamma_j \right),$$

where $|A|$ is the norm of the second fundamental form of $\Sigma$ and the $\Gamma_j$ are the connected components of the boundary.
Corollary 1. Given $\theta \in (0, 1)$, there is some $\gamma = \gamma(n, \theta)$ such that if $\int_{\Sigma} |A| < \theta \rho$, if $\sum_j \text{diam} \Gamma_j \leq \gamma \rho$, if $\Sigma \cap \partial B_{\rho} \neq \emptyset$, and if $\Sigma \cap \partial B_{\theta \rho} \neq \emptyset$, then $|\Sigma_{\rho}| \geq C \rho^2$.

The next lemma quantifies the earlier remark that $F(\Sigma) < 8\pi$ guarantees embeddedness; it may be viewed as a generalization of a lemma of Li and Yau [5].

Lemma 2.4. Suppose that $\Sigma$ is a compact surface without boundary, $\partial B_{\rho}$ intersects $\Sigma$ transversely, and $\Sigma_{\rho}$ contains disjoint subsets $\Sigma_1, \Sigma_2$ with $\Sigma_j \cap \hat{B}_{\theta \rho} \neq \emptyset$, $\partial \Sigma_j \subset \partial B_{\rho}$, and $|\partial \Sigma_j| \leq \beta \rho$, where $\theta \in (0, \frac{1}{2})$ and $\beta > 0$. Then

$$F(\Sigma) \geq 8\pi - C \beta \theta,$$

where $C$ is independent of $\Sigma$, $\beta$, $\theta$.

We will need one further lemma from [8]. This lemma asserts that in balls where the $L^2$ norm of the second fundamental form is small, the surface is given by a union of discs, where each disc is almost the graph of a function with small Lipschitz norm. In the lagrangian case, it will follow that, locally, the surface is close to the graph of the gradient of a function (where the gradient has small Lipschitz norm).

Lemma 2.5. For any $\beta > 0$, there is an $\epsilon_0 > 0$ (depending on $\beta$) such that if $\epsilon \in (0, \epsilon_0]$, if $\partial \Sigma \cap \overline{B}_\rho = \emptyset$, if $0 \in \Sigma$, if $|\Sigma_{\rho}| \leq \beta \rho^2$, and if $\int_{\Sigma} |A|^2 \leq \epsilon^2$, then the following holds:

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

$$\sum_j \text{diam} P_j \leq C \epsilon^\frac{1}{2} \rho$$

and

$$\Sigma \cap B_\epsilon \setminus \left( \bigcup_{i=1}^{N} P_i \right) = \left( \bigcup_{i=1}^{M} \text{graph} u_i \right) \cap B_\epsilon,$$

where each $u_i \in C_\infty^\infty(\Omega_i, L_i^1)$, $L_i$ is a plane in $\mathbb{R}^4$, $\Omega_i$ is a smooth bounded subdomain of $L_i$, and

$$\sup_{\Omega_i} \rho^{-1} |u_i| + \sup_{\Omega_i} |Du_i| \leq C \epsilon^\frac{1}{6}.$$

Furthermore, $\forall \sigma \in (\frac{\rho}{4}, \frac{\rho}{2})$ such that $\Sigma \cap \partial B_\sigma$ and $\partial B_\sigma \cap \bigcup_{i=1}^{M} P_j = \emptyset$, then

$$\Sigma_\sigma = \bigcup_{i=1}^{M} D_{\sigma, i},$$

where $D_{\sigma, i}$ is a disc, and $D_{\sigma, i} \setminus \text{graph}(u_i) \text{ is a union of discs}.$

Notice that the lemma says that $\Sigma_\sigma$ is a union of discs with smooth boundary contained in $\partial B_\sigma$, and each disc can be written as a Lipschitz graph with small
gradient in addition to some “pimples” $P_j$, which are small in every sense. First, the pimples have small total diameter; second, they have small total curvature (in the $L^2$ sense). Furthermore, it follows from the earlier pseudo-monotonicity estimates, equation (5), that they have small area:

$$\sum_j |P_j| \leq C \rho^2.$$ 

Note further that if the surfaces are lagrangian, we can take the planes, $L_i$, to be lagrangian.

3. WILLMORE IMPLIES PLATEAU

In this section, we will show the relationships between several conjectures. Precisely, it is conjectured that the Clifford torus minimizes area in its Hamiltonian isotopy class; we show that if a smooth minimizer to this problem exists, then this conjecture follows from the Willmore conjecture for lagrangian tori. This relationship was one of the original motivations for this work.

Schoen and Wolfson ([7]) have proven an existence and regularity theorem for area minimizing lagrangian surfaces with fixed boundary. For this reason, it seems likely that either there exists a smooth area minimizer in the Hamiltonian isotopy class of the Clifford torus, or the infimum of area over the class is zero (so a minimizing sequence degenerates). We can show that if the area is going to zero under a sequence of Hamiltonian isotopies, then the Willmore functional is blowing up.

**Lemma 3.1.** Suppose that $\Sigma_j$ are Hamiltonian isotopic to the Clifford torus, $T$, and that $|\Sigma_j| \to 0$. Then $F(\Sigma_j) \to \infty$.

**Proof.** Given a subset, $S$, of $\mathbb{R}^4$, we define the cylindrical capacity to be the infimum of $\pi r^2$ such that there exists a symplectic map, $G$, with

$$G(S) \subset \{ x_i, y_i : x_i^2 + y_i^2 \leq r^2 \}.$$

First note that the cylindrical capacity, $c$, is a symplectic invariant; hence, $c(\Sigma_j) = c(T)$, which is a nonzero constant (see [3] for details). Therefore, since small spheres can be translated inside the above cylinders, diam $\Sigma_j$ must be bounded below (in fact, it’s always at least 1). Now apply Lemma 2.2 to conclude that $\forall \Sigma \in C$, where $C$ is the Hamiltonian isotopy class of the Clifford torus, we must have

$$|\Sigma|F(\Sigma) \geq k,$$

for some constant $k$ independent of $\Sigma$. The lemma follows. $\square$

As mentioned previously, the Euler-Lagrange equation for minimizing area among lagrangian surfaces is $\Delta_H H = 0$, where $H$ is the mean curvature and $\Delta_H$ is the Hodge Laplacian. It follows that any lagrangian surface with $H$ parallel is critical with respect to area. We prove below that the cylinder $C = S^1 \times ([a, b], 0)$ minimizes area in its Hamiltonian isotopy class (we already
knew that it was a critical point). This example illustrates just how poorly behaved these lagrangian minimal surfaces can be (for example, we can't have monotonicity of density); in the minimal surface case, as $|b - a|$ increases, the minimizing surface would get narrower and eventually pinch into two disjoint disks.

**Proposition 1.** The lagrangian cylinder $\tilde{C}$ minimizes area in its Hamiltonian isotopy class.

**Proof.** Given any competing lagrangian surface $\tilde{C} = G(\tilde{C})$, where $G$ is a symplectic map fixing the boundary $\partial \tilde{C}$, we use the $x_2$ coordinate function to slice $\tilde{C}$ and apply the co-area formula. We get that

$$|\tilde{C}| \geq \int_{t=a}^{t=b} L(\tilde{C} \cap \{x_2 = t\}) \, dt,$$

where $L$ is the length (actually one-dimensional Hausdorff measure). Since $\tilde{C}$ is connected, the slices above are nonempty; by Sard's theorem, almost every slice is transverse (and hence one dimensional). Given a transverse slice, $\Gamma$, then Stoke's theorem yields that

$$\int_{\Gamma} \alpha = \int_{S^1 \times (a, 0)} \alpha = \pi$$

since $\tilde{C} \cap \{x_2 \leq t\}$ is lagrangian.

Note further that on the hyperplane $\{x_2 = t\}$ the form $\alpha$ evaluated on a curve $\Gamma$ just measures the area (with multiplicity) of the projection to the $x_1 - y_1$ plane. Now let $S$ be the surface of least area spanning $\Gamma$; it follows that $|S| \geq \pi$.

Now the isoperimetric inequality implies that $L(\Gamma) = L(\partial S) \geq 2\pi$. Therefore,

$$|\tilde{C}| \geq 2\pi(b - a),$$

which is the area of the cylinder. \(\Box\)

**Proposition 2.** Let $C$ be the Hamiltonian isotopy class of the Clifford torus. If $\exists \Sigma$, some smooth critical point with respect to area in $C$, then $|\Sigma| = 2F(\Sigma)$.

**Proof.** First note that $[\alpha]$ is fixed under Hamiltonian isotopy (since these preserve $\omega = d\alpha$). Second, since $[d\theta]$ is an integral class, it is also fixed under Hamiltonian isotopy (since it must change continuously). Since these two classes are equal on the Clifford torus, they are then equal on every element of $C$. Hence, $\alpha - d\theta$ is an exact form on $\Sigma$; write $\alpha - d\theta = df$, where $f$ is a smooth function on $\Sigma$. Now by the tubular neighborhood theorem, we can use a cutoff in the normal directions and extend $f$ to an ambient function, $\hat{f}$, whose normal derivative vanishes along $\Sigma$.

Consider the first variation of area from the Hamiltonian vector field $JD\hat{f}$. Since $\Sigma$ is critical, this must be zero. Therefore,

$$0 = \int_{\Sigma} \langle H, H + R \rangle \, d\mu = \int_{\Sigma} |H|^2 \, d\mu - \int_{\Sigma} \text{div}(R) \, d\mu,$$
since \( \alpha \sim JR^\perp \). Given any \( n \)-dimensional submanifold of any Euclidean space, it follows from a simple calculation that the divergence of the radial vector field is \( n \). It follows that the radial variation for surfaces always gives twice the area; we get that

\[
|\Sigma| = 2F(\Sigma). \quad \Box
\]

Since the above proposition is true for all smooth critical points of area, it also applies to a smooth area minimizer (if one exists).

**Theorem 1.** If there exists a smooth area minimizer in \( C \), then the Willmore conjecture for lagrangian tori would imply that the Clifford torus, \( T \), minimizes area in \( C \). Furthermore, if, in addition, the full Willmore conjecture is true, then \( T \) is the unique area minimizer (up to isometry).

**Proof.** If \( \Sigma \) is the area minimizer, then the proposition implies that

\[
|\Sigma| = 2F(\Sigma) \geq 2F(T) = 4\pi^2.
\]

Finally note that we have equality for \( \Sigma \) equal to the Clifford torus.

Furthermore, if the full Willmore conjecture holds, then \( \Sigma \) achieves the minimum for \( F \), and hence satisfies the Willmore equation (see [13] for a derivation; note that Weiner defines the mean curvature to be \( \frac{1}{2}H \)):

\[
\Delta H - \frac{1}{2}|H|^2 H + \tilde{A}(H) = 0,
\]

where \( \Delta \) is the Laplacian on the normal bundle and \( \tilde{A} \) is the operator defined by Simons in [10].

We will now show that if \( \Sigma \) is a smooth lagrangian surface critical for area which satisfies equation (9), then it must be a product torus (and hence the Clifford torus). This follows from Hoffman's classification of constant Gauss curvature and parallel mean curvature surfaces in \( S^3 \) (see [4]). Note that Chen (see [1]) has shown that if \( \Sigma \subset R^4 \) has parallel mean curvature vector, then it has constant mean curvature in an affine three-space or a round \( S^3 \). Since \( \Sigma \) is closed and lagrangian, it must have constant mean curvature in some three-sphere.

Since \( \Sigma \) is critical for area, we get that

\[
\Delta_H H = 0,
\]

where \( \Delta_H \) is essentially the Hodge Laplacian on one-forms (with some canonical isomorphisms induced by \( J \) and the metric). Combining equation (10) and the Bochner formula, equation (9) becomes

\[
KH - \frac{1}{2}|H|^2 H + \tilde{A}(H) = 0.
\]
Next, since the Gauss equation implies that $|H|^2 = |A|^2 + 2K$, equation (11) yields

$$\frac{1}{2} |A|^2 H = \tilde{A}(H). \tag{12}$$

At this point we will utilize the full symmetry of $A$ for lagrangian surfaces. Work at a point $p \in \Sigma$ and choose a frame such that $H = |H|(1,0)$. Then, if $a_{ijk}$ gives the components of $A$ in these coordinates (where, as discussed in Section 2, we have used $J$ to transform $A$ into a purely tangential tensor), equation (12) yields

$$\frac{1}{2} a_{ijk}^2 h_k = a_{ijk} a_{ijl} h_l \tag{13},$$

where $h_k$ gives the components of $H$. Taking the inner product of equation (13) with $H$, and then dividing by $|H|^2$, we see that

$$a_{ijk}^2 = 2a_{ij1}. \tag{14}$$

Note that $|H|$ does not vanish since $H$ is a harmonic one-form, and the space of harmonic one-forms on a surface is a conformal invariant (and in the flat model space the harmonic one-forms are parallel and hence nonvanishing). Using the symmetry now of $a_{ijk}$, this implies that

$$a_{111} = a_{122}. \tag{15}$$

Given this, we use the Gauss equation to compute $K$; we get that $K = -2a_{112}^2 \leq 0$. However, Gauss-Bonnet implies that $\int K = 0$; we conclude that $K = 0$ identically. The Bochner formula now yields that $H$ is parallel. \(\square\)

Another consequence of Proposition 2 is the following.

**Corollary 2.** If there exists a smooth surface, $\Sigma$, in $C$ with $|\Sigma| \leq 8\pi$, then there does not exist a smooth area minimizer in $C$.

**Proof.** If a smooth area minimizer, $\Sigma$, did exist, then it’s area would be less than or equal to $8\pi$; hence, by Proposition 2, $F(\Sigma) \leq 4\pi$. Only the round spheres satisfy this; this is a contradiction. \(\square\)

4. THE MAIN THEOREM

In this section, we will state the main theorem, as well as a different version of the theorem which the methods of the proof will easily extend to cover.

**Theorem 2.** There exists a smooth, embedded, lagrangian torus, $\Sigma$, in $\mathbb{R}^4$, with the standard metric and symplectic form, which minimizes the Willmore functional amongst all such tori. Furthermore, this $\Sigma$ may be realized as the limit of a minimizing sequence $\{\Sigma_k\}$.

Several remarks are in order. First, smooth means real analytic. Second, the Clifford torus gives us an example with $F(T) = 2\pi^2$, and it’s always true
that $F(\Sigma) \geq 4\pi$; hence the minimum is between $4\pi$ and $2\pi^2$. Furthermore, embeddedness comes for free, since a double point would force $F(\Sigma)$ to be at least $8\pi$. Finally, topological considerations force any oriented embedded lagrangian surface in $\mathbb{R}^4$ to be a torus; in [8], when Simon proved existence and regularity in the standard case, extra care needed to be taken to ensure that the limit surface was a torus.

In Proposition 3, we will show that, whenever the Willmore functional is uniformly bounded and we have convergence as measures, we actually get convergence in Hausdorff distance. As a consequence, we will be able to work with marked tori (and be able to make sense of the markings under limits).

**Definition 4.1.** Let $C_\alpha$ denote the class of embedded, marked, lagrangian tori in $\mathbb{R}^4$ with $\alpha$ evaluated on the generators for homology yielding $2\pi$ and $2\pi r$.

Notice that the product torus with radii $1$ and $r$ would then be in $C_\alpha$; in fact, anything Hamiltonian isotopic to this product torus is in that class. Furthermore, these classes are equivariant with respect to rescalings; that is, we can dilate, hit it with a Hamiltonian isotopy, and then undilate and we will remain in the class at the end.

**Theorem 3.** For each $r$ such that $(r + r^{-1}) < 8\pi^{-1}$, there exists a smooth, embedded, lagrangian torus, $\Sigma \in C_r$, which minimizes the Willmore functional over the class $C_r$. Furthermore, this $\Sigma$ may be realized as the limit of a minimizing sequence $\{\Sigma_k\}$.

Note that the restriction on $r$ allows us to prove embeddedness for the limit surface. With $r$ as above, the product torus in $C_r$ has Willmore functional less than $8\pi$. Of course, the class, $C_1$, of the Clifford torus satisfies this constraint.

After proving Theorem 2, we will explain how to use the very strong convergence of the minimizing sequence to extend the proof to cover Theorem 3.

5. CONVERGENCE AND $C^{1,\alpha}$ REGULARITY

We will show that given a minimizing sequence, we can extract a convergent subsequence which, away from some bad points, is a $C^{1,\alpha}$ surface. Choose a minimizing sequence $\Sigma_j$ such that

$$F(\Sigma_j) \leq \inf_{\Sigma \in C} F(\Sigma) + \epsilon_j,$$

where $\epsilon_j \to 0$, and each surface is a smooth lagrangian torus. Because $F$ is conformally invariant, we can normalize the $\Sigma_j$ such that $|\Sigma_j| \equiv 1$, and $0 \in \Sigma_j$.

**Proposition 3.** Given $\Sigma_j$ as above, there exists a subsequence which converges as measures to a limit measure $\mu$; furthermore, the subsequence converges to $\text{spt}(\mu)$ in the Hausdorff distance sense.

The proof of this is due to Simon [8]; however, for completeness, we will go through it anyway. Note that it is not necessary that the Willmore functional
be approaching the infimum, but merely that it is uniformly bounded; if we did not have convergence in Hausdorff distance, the supremum of \( F(\Sigma_j) \) would have to be infinite.

**Proof.** First note that the normalization and Willmore bound imply, by Lemma 2.2, that all the \( \Sigma_j \) are contained in a fixed compact subset of \( \mathbb{R}^4 \). Hence, the area bound implies that the measures

\[
\mu_j \equiv \text{Hausdorff measure}|_{\Sigma_j}
\]

have a weakly convergent subsequence—which we will still call \( \Sigma_j \)—which converges to a limit measure \( \mu \). Therefore, given any continuous function, \( f \), on \( \mathbb{R}^4 \), we get

\[
\lim_{j \to \infty} \int_{\Sigma_j} f \to \int_{\mathbb{R}^4} f \, d\mu.
\]

Let \( \Sigma \) denote the support of \( \mu \), \( \text{spt}(\mu) \). Suppose that \( \Sigma_j \) does not converge to \( \Sigma \) in Hausdorff distance; then \( \exists y_j \in \Sigma_j \) such that \( y_j \to y \) and \( \text{dist}(y, \Sigma) = \eta > 0 \). Since all of the surfaces are connected, \( \exists j_0 \) such that

\[
\Sigma_j \cap B_\sigma(y) \neq \emptyset \quad \forall j \geq j_0 \quad \text{and} \quad \sigma \in \left( \frac{\eta}{4}, \frac{\eta}{2} \right).
\]

Therefore, \( \forall N \geq 1 \) and \( j \geq j_0 \),

\[
\exists z_{j, k} \in \Sigma_j \cap B_{\rho_k}(y), \quad k = 1, \ldots, N,
\]

where \( \rho_k = \left(1 + \frac{k}{N}\right) \frac{\eta}{4} \). Now apply the pseudo-monotonicity estimate, equation (5), to \( \Sigma_j \) about the points \( z_{j, k} \) with \( \rho = \frac{1}{N} \left( \frac{\eta}{4} \right) \). It follows that

\[
\pi \leq C \left( \rho^{-2} |\Sigma_j \cap B_\rho(z_{j, k})| + F(\Sigma_j \cap B_\rho(z_{j, k})) \right).
\]

Observing that the \( B_\rho(z_{j, k}) \) are pairwise disjoint, and that there are \( N \) of them, we conclude that

\[
N \pi \leq C \left( \sum_k \rho^{-2} |\Sigma_j \cap B_\rho(z_{j, k})| + F(\Sigma_j \cap A_{\frac{\eta}{4}}, y(z_{j, k})) \right).
\]

The key fact is now that \( y \) is not in the support of the limit measure; hence, the limit of the densities near \( y \) must be zero. Therefore, the first term on the right tends to zero. Taking the limit,

\[
N \pi \leq \limsup F(\Sigma_j).
\]

Since this is true for any finite value of \( N \), this contradicts the uniform bound on the Willmore functional, and the proposition follows. \( \square \)

It is this set \( \Sigma = \text{spt}(\mu) \) that we will show is the limit surface. This amounts to proving regularity for \( \Sigma \), and showing that the above convergence is good enough to both make \( \Sigma \) lagrangian and force it to achieve the infimum of
the Willmore functional (essentially show that the functional is lower semi-
continuous with respect to our convergence).

We will prove regularity in several stages. First, we will show that, away from
a finite set of bad points, \( \Sigma \) is a \( C^{1,\alpha} \cap W^{2,2} \) surface. In Section 6 we will
then show that it is actually \( C^{2,\alpha} \cap W^{3,2} \). From there higher regularity is fairly
standard. Finally, we will eliminate the bad points.

Before we start the first stage, we will prove a comparison lemma which will
allow us to construct a competing sequence in the proof of Proposition 4.

**Lemma 5.1.** Let \( \Sigma \subset \mathbb{R}^4 \) be a smooth embedded lagrangian surface, \( y \in \mathbb{R}^4 \), \( L \)
a lagrangian plane containing \( y \), \( U \subset L \) an open neighborhood of \( L \cap \partial B_\rho(y) \),
\( u \in C^\infty(U) \), and
\[
\text{graph}(Du) \subset \Sigma, \quad |D^2 u| \leq 1.
\]
Then \( \exists w \in C^\infty(L \cap B_\rho(y)) \) such that
\[
\begin{align*}
\Delta^3 w &= 0 & \text{on } L \cap B_\rho(y), \\
Dw &= Du, \quad D^2 w = D^2 u & \text{on } L \cap \partial B_\rho(y),
\end{align*}
\]
and \( C \) independent of \( \Sigma \) and \( \rho \) such that
\[
\int_{L \cap B_\rho(y)} |D^3 w|^2 \leq C \rho \int_{\Gamma} |A|^2 d\mathcal{H}^1,
\]
where \( \Gamma \) is the graph of \( Du \) on the boundary, \( A \) is the second fundamental form
of \( \Sigma \), and \( \mathcal{H}^1 \) is the arclength measure on \( \Gamma \).

**Proof.** For ease of exposition, rescale such that \( \rho = 1 \). Let \( \gamma \) be the projection
to \( L \) of \( \Gamma \). The key fact here is that the triharmonic \( w \) minimizes \( \int |D^3 w|^2 \)
over all functions with the above boundary values. By the Sobolev space trace
lemma (see, for example, [12]),
\[
\int |D^3 w|^2 \leq C \left( |w|_{W^{2,\frac{1}{2}}}^2 + |Du|_{W^{2,\frac{1}{2}}}^2 + |D^2 w|_{W^{2,\frac{1}{2}}}^2 \right),
\]
where the fractional Sobolev norms are computed along \( \gamma \). Trivially bounding
these norms in terms of higher Sobolev norms, and applying this to \( w - q \),
where \( q \) is a quadratic polynomial, we get that
\[
\int |D^3 w|^2 \leq C \int_{\gamma} \left( |w - q|^2 + |Dw - Dq|^2 + |D^2 w - D^2 q|^2 + |D^3 w|^2 \right) d\mathcal{H}^1.
\]
By choosing \( q \) appropriately, we can bound the first three terms on the right by
some constant times the last term (the Wirtinger inequality); we conclude that
\[
\int |D^3 w|^2 \leq C \int_{\gamma} |D^3 w|^2 d\mathcal{H}^1 \equiv C \int_{\gamma} |D^3 u|^2 d\mathcal{H}^1.
\]
Since \( |D^2 u| \leq 1 \) on \( \gamma \), we get that \( |D^3 u|^2 \leq C|A|^2 \). This completes the proof
of the lemma. \( \square \)
Proposition 4. The limit surface, $\Sigma$, defined above is a $C^{1,\alpha} \cap W^{2,2}$ surface away from a finite set of points $\xi_1, \ldots, \xi_N$.

Proof. Fix $\epsilon > 0$—this will be taken small enough to apply the graphical decomposition lemma later. A point $\xi \in \Sigma$ is called a bad point relative to $\epsilon$ if

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

For future reference, define the quantity $\Psi(\rho, \xi)$ to be

$$\Psi(\rho, \xi) \equiv \lim_{k \to \infty} \int_{\Sigma_k \cap B_\rho(\xi)} |A_k|^2.$$ 

Since $\int_{\Sigma_k} |A_k|^2$ is bounded, there can only be finitely many bad points for each $\epsilon > 0$; denote the bad points by $\xi_1, \ldots, \xi_N$.

Now let $\xi$ be any point in $\Sigma \setminus \{\xi_1, \ldots, \xi_N\}$; then $\exists \rho(\xi, \epsilon) > 0$ such that

$$\lim_{\rho \to 0} \left( \liminf_{k \to \infty} \int_{\Sigma_k \cap B_\rho(\xi)} |A_k|^2 \right) \leq \epsilon^2 \quad \forall \rho < \rho.$$ 

Therefore, there are infinitely many values of $k$ such that:

$$\int_{\Sigma_k \cap B_\rho(\xi)} |A_k|^2 \leq \epsilon^2 \quad \forall \rho < \rho,$$

and we may as well suppose that this subsequence was our original sequence (for ease of notation).

We can now apply the graphical decomposition lemma, Lemma 2.5, to each $\Sigma_k \cap B_\rho(\xi)$. We conclude that:

$\exists$ lagrangian planes $L_k$ containing $\xi$, functions $u_k \in C^\infty(\Omega_k)$, pimples $P_{k,j}$, and disks $D_{1}^{(k)}$ such that

$$\rho^{-1}|u_k| + |D u_k| \leq C \epsilon^\frac{1}{2},$$

$$\sum_j \text{diam } P_{k,j} \leq C \epsilon^\frac{1}{2} \rho,$$

$$\sum_j |P_{k,j}| \leq C \epsilon \rho^2,$$

where each $P_{k,j}$ is diffeomorphic to a closed disk, and $\sigma \in (\theta^\frac{\rho}{2}, \theta \rho)$ for some $\theta \in (0, 1)$ independent of $k$. In fact, this is true $\forall \sigma \in T$, where the measure of $T$ is at least $\theta^\frac{\rho}{8}$ and we have again gone to a subsequence.

Note that we used Lemma 2.4 to conclude that, as long as $\theta$ was small enough, only one disk $D_{1}^{k}$ intersected $B_{\theta \rho}(\xi)$.

Now, because $\Sigma_k \cap B_{\theta \rho}(\xi)$ is topologically trivial, the form $\alpha$ restricted to it must be exact. Therefore, each $(\text{graph } u_k)$ is an exact lagrangian graph (a priori
there is some worry, because of the pimples); we conclude that each \( u_k \) is, in fact, the gradient of a potential. That is,

\[
u_k \equiv Dw_k.
\]

Moreover, since the potential is only defined up to a constant, we can arrange that

\[
\rho^{-2}|w_k| + \rho^{-1}|Dw_k| + |D^2w_k| \leq C\epsilon^\theta.
\]  

Given a \( \sigma \in T \), solve for the triharmonic functions, \( w_k \), on \( B_\sigma(\xi) \cap L_k \) which match up with \( \Sigma_k \) on the boundary, \( \Gamma_k \). It follows from Lemma 5.1 that

\[
\int_{B_\sigma(\xi) \cap L_k} |D^3w_k|^2 \leq C\sigma \int_{\Gamma_k} |A_k|^2.
\]

If we let \( \bar{A}_k \) denote the second fundamental form of \( \text{graph} Dw_k \), then

\[
\int_{\text{graph} Dw_k} |\bar{A}_k|^2 \leq C\sigma \int_{\Gamma_k} |A_k|^2,
\]

since the second fundamental form of a graph is always dominated by a dimensional constant times the norm of its Hessian.

We will now construct a competing sequence, by patching in these triharmonic graphs. Define

\[
\tilde{\Sigma}_k \equiv (\Sigma_k - D_1^k) \cup \text{graph} Dw_k.
\]

By virtue of the fact that \( \{\Sigma_k\} \) is a minimizing sequence for \( F \), we conclude that

\[
F(\tilde{\Sigma}_k) \geq F(\Sigma_k) - \epsilon_k, \quad \epsilon_k \to 0.
\]

Since these surfaces agree except on the triharmonic patch, it follows from equations (17) and (18) that

\[
\int_{D_1^k} |A_k|^2 \leq C\sigma \int_{\partial D_1^k} |A_k|^2 + \delta_k, \quad \delta_k \to 0.
\]

Recall that we could have done this for any \( \sigma \in T \) where \( T \subset (\theta \rho, \theta\rho) \) and has measure at least \( \theta^p \). Hence, we can arrange that

\[
\theta \rho \int_{\partial D_1^k} |A_k|^2 \leq 8 \int_{D_1^k \backslash B_{\theta \rho}} |A_k|^2.
\]

Therefore, equations (19) and (20) imply that:

\[
\int_{\Sigma_k \cap B_{\theta \rho}} |A_k|^2 \leq C \int_{\Sigma_k \cap A_\rho} |A_k|^2 + \delta_k, \quad \delta_k \to 0,
\]
for some $\sigma \in (\theta_x, \theta \rho)$; note that $A_{\sigma}$ denotes the annulus $B_{\sigma} \setminus B_{\frac{1}{2}}$, and all balls are centered at $\xi$. We now use hole-filling to get the key decay estimate on the second fundamental form:

\[
\int_{\Sigma_k \cap B_{\theta \frac{1}{2}}} |A_k|^2 \leq \gamma \int_{\Sigma_k \cap B_{\theta \rho}} |A_k|^2 + \delta_k,
\]

where $\gamma \in (0, 1)$.

Taking limits, we can now rewrite equation (22) in terms of the function $\Psi$:

\[
\Psi(\frac{\theta \rho}{2}, \xi_0) \leq \gamma \Psi(\theta \rho, \xi_0),
\]

where $\gamma < 1$ is independent of $\rho$ and $\xi$. Furthermore, equation (23) is valid for all $\xi_0 \in \Sigma \cap B_{\rho_0}(\xi)$. Define $\rho_0 \equiv \frac{\theta \rho(\xi, \sigma)}{2}$. If we iterate (23), we get some $\alpha \in (0, 1)$ such that

\[
\Psi(\sigma, \xi_0) \leq C \left(\frac{\sigma}{\rho_0}\right)^\alpha \Psi(\rho(\xi), \xi),
\]

for all $\xi_0$ and $\sigma$ as above.

Now fix some $\xi_0 \in \Sigma \setminus \{\xi_1, \ldots, \xi_N\}$ and set $\rho_0 \equiv \frac{\theta \rho(\xi_0, \sigma)}{2}$. Take any $\xi \in \Sigma \cap B_{\rho_0}(\xi_0)$ and $\rho \in (0, 1)$. Define

\[
\alpha_k = \alpha_k(\rho, \xi) = \int_{\Sigma_k \cap B_\rho(\xi)} |A_k|^2;
\]

note that this is bounded by $\epsilon^2$ by construction. Furthermore, equation (24) implies that, for infinitely many $k$,

\[
\alpha_k(\rho, \xi) \leq 2\Psi(\rho, \xi) \leq C \left(\frac{\rho}{\rho_0}\right)^\alpha \alpha_k(\rho_0, \xi_0).
\]

So, once we fix $\xi_0$ and $\rho_0$, we get that (again for $k$ large):

\[
\alpha_k(\rho, \xi) \leq C \left(\frac{\rho}{\rho_0}\right)^\alpha.
\]

Recall the conclusion of the graphical decomposition lemma; we had the potentials $\psi_k$ defined on $\Omega_k \subset L_k$ with estimates (16). We can now extend these functions to all of $L_k$ and preserve the estimates by the Whitney extension theorem (see Stein, [11], for example). We will still denote the extensions by $u_k$ and their gradients by $u_k$. We will now show convergence of these graphs with the Arzela-Ascoli theorem.

By virtue of Simon's version of the Poincare inequality (see [8], appendix A), given a function $f$,

\[
\inf_{x \in L_k} \int |f(x) - \lambda|^2 \leq C \rho^2 \int |Df|^2 + C \sqrt{\alpha_k}(\sup |f|^2) \rho^2,
\]
where we have used the diameter bound on the pimples from the graphical decomposition lemma. Applying the above to $f = D^{ij} w_k$, we get constants, $\eta^k = (\eta^k_{ij})$, such that

$$\int_{B_p(\xi)} |D^2 w_k - \eta^k|^2 \leq C \rho^2 \left( \int_{B_p(\xi)} |D^3 w_k|^2 + \sqrt{\alpha_k(p, \xi)} \right).$$

Using the bounds from (26) and the above, we get that

$$\int_{B_p(\xi)} |D^2 w_k - \eta^k|^2 \leq C \rho^{2+\alpha}. \tag{27}$$

By compactness of the Grassmann manifold, $G_{2,4}$, we can take a subsequence such that $L_k \to L$ and $\eta^k \to \eta$, where $L$ is a lagrangian plane. Given the estimates (16) and (27), it then follows from the Arzela-Ascoli theorem that graph $Dw_k$ converges to graph $Dw$ in Hausdorff distance, and that

$$\rho^{-2} |w| + \rho^{-1} |Dw| + |D^2 w| \leq C \epsilon^{1/6} \tag{28}$$

and

$$\int_{B_p(\xi)} |D^2 w - \eta|^2 \leq C \rho^{2+\alpha}, \tag{29}$$

by the lower semi-continuity of energy.

If we interpret the above in measure-theoretic terms, we see that for every point $\xi \in \Sigma \cap B_{\rho_0}(\xi_0)$ and all $\rho \in (0, \rho_0)$, we have

$$\mathcal{H}^2|_{\Sigma \cap B_\rho(\xi)} = \mathcal{H}^2|_{\text{graph } Dw_k \cap B_\rho(\xi)} + \Theta_k, \tag{30}$$

where $\mathcal{H}^2$ is two-dimensional area measure and where $\Theta_k$ is a signed measure with small mass, $|\Theta_k| \leq C \rho^{2+\alpha}$. Furthermore, taking the limit, we get

$$\mu|_{B_\rho(\xi)} = \mathcal{H}^2|_{\text{graph } Dw \cap B_\rho(\xi)} + \Theta, \tag{31}$$

with $\xi$ and $\rho$ as above, and

$$|\Theta| \leq C \rho^{2+\alpha}. \tag{32}$$

It follows from (31) that at $\xi$ the density of $\mu$ is given by

$$\text{Dens}(\mu, \xi) = \pi + \text{Dens}(\Theta, \xi).$$

However, the decay estimate, equation (32), implies that $\text{Dens}(\Theta, \xi) = 0$. We conclude that $\Sigma$ has a multiplicity one tangent plane at $\xi \in \Sigma \cap B_{\rho_0}(\xi_0)$. Furthermore, the energy decay on $Dw$ implies a Hölder estimate on the normal
space to $\Sigma$:

$$|N(y_1) - N(y_2)| \leq C|y_1 - y_2|^\gamma.$$ 

That is, $\Sigma \cap B_{\rho_0}(\xi_0)$ is locally a $C^{1, \frac{1}{2}}$ surface, which, given $\delta > 0$, can be given in some small neighborhood, $U$, by graph $Dw$ where

$$Dw \in C^{1, \frac{1}{2}}(U \cap L_0),$$

$$\sup_{U \cap L_0} |D^2 w| \leq \delta,$$

where $L_0$ is the tangent plane at $\xi_0$. Since $\Sigma$ is locally the graph of a gradient over a lagrangian plane, $\Sigma$ must be a lagrangian surface. This would also follow from the multiplicity-one varifold convergence.

It remains to show that the limit is locally in $W^{2,2}$ and that the mean curvature exists in the limit. To these ends, first note that in the small balls, $B_{\rho}(\xi)$, $\Sigma$ is the multiplicity-one varifold limit of the $\Sigma_k$, and that

$$\int_{\Sigma_k \cap B_{\rho}(\xi)} |H_k|^2 \leq C \rho^\gamma;$$

we conclude that $\Sigma$ has generalized mean curvature vector $H$ such that

$$\int_{\Sigma \cap B_{\rho}(\xi)} |H|^2 \leq C \rho^\gamma.$$ 

Once we have the estimate (35), we get that $Dw$ is a $C^1$ weak solution to the prescribed mean curvature equation. We can then use standard difference quotient arguments (see, for example, [2], Theorem 8.8) to conclude that $Dw \in W^{2,2}$ and, again by hole-filling, get the estimate

$$\int_{L_0 \cap B_{\rho}(\xi)} |D^3 w|^2 \leq C \rho^{2\alpha},$$

for some $\alpha > 0$ and for $\xi \in \Sigma \setminus \{\xi_1, \ldots, \xi_N\}$.

Notice that Morrey's lemma and the decay estimate, equation (36), also imply the Holder continuity. We conclude that, away from the bad points, $\Sigma$ is a $C^{1, \alpha} \cap W^{2,2}$ lagrangian surface. □

6. ADDITIONAL REGULARITY

In order to get additional local regularity for the limit surface, we will prove a general regularity theorem for a certain type of sixth-order nonlinear partial differential equation. We will be working on $D \equiv \{x \in \mathbb{R}^2 | |x| < 1\}$, and the results here rely heavily on the fact that we are in dimension two.
Theorem 4. Let $\beta, \gamma, L > 0$ and $u = Dw \in W^{2,2}(D, \mathbb{R}^2) \cap C^{1,\gamma}(D, \mathbb{R}^2)$ such that $\sup_{D}|u| + \sup_{D}|D\!u| \leq 1$ with

$$\int_{D \cap B_p(\xi)} |D^2 u|^2 \leq \beta \rho^{2\gamma} \quad \forall \rho < 1 \quad \text{and} \quad \xi \in D,$$

where $w$ is a weak solution to the sixth-order quasilinear equation:

$$D^\alpha \left(A_{\alpha\alpha}(x, u, Du)D^{\alpha}w\right) + D^{ij} \left(B_{ij}(x, u, Du, D^2 u)\right) + D^k \left(B^0_k(x, u, Du, D^2 u)\right) = 0$$

(with $\alpha, \beta$ multi-indices of order three), where $A(x, z, p), B(x, z, p, q)$ for $|z| + |p| \leq 1$ satisfy:

(A) Given a three-tensor, $\zeta_{\alpha}, \sum_{\alpha} \bar{\zeta}_{\alpha} \zeta_{\alpha} A_{\alpha\alpha} \geq L^{-1}|\zeta|^2$.

(B) $|A| \leq L, \ |D_{(x, z, p)} A| \leq L$.

(C) At any point $(x, z, p, q)$, we have

$$|B| + |D_{(x, z, p)} B| \leq L(1 + |q|^3)$$

and

$$|D_q B| \leq L(1 + |q|).$$

Then there exists some $\alpha > 0$ such that $u \in W^{3,2}_{\text{loc}}(D) \cap C^{2,\alpha}(D)$. In fact, there exists some constant $C > 0$ (depending on $\beta, \gamma$, and $L$) such that

$$\int_{B_p(\xi)} |D^3 u|^2 \leq C \rho^{2\alpha} \quad \forall \ \xi \in D \ s.t. \ \text{dist}(\xi, \partial D) > 2\rho.$$

The proof of this theorem relies upon using difference quotients, and picking specific test functions to put into the weak form of the equation. Recall the following operators:

$$\delta_h f(x) = h^{-1} (f(x + he) - f(x))$$

and

$$\delta_h f(x) = h^{-1} (f(x) - f(x - he)),$$

where $h \neq 0, x \in D_{|h|} \equiv \{x \in D \ | \ \text{dist}(x, \partial D) > |h|\}$, and $e = (1, 0)$ or $(0, 1)$. We have the following easy consequences of the definitions:

(37) $\delta_h (f g)(x) = (\delta_h f(x)) g(x + he) + f(x) \delta_h g(x),$  

(38) $\int_D g \delta_h f = -\int_D f \delta_h g$, where $\text{spt}(fg) \subset D_{|h|}$.

We will also use a lemma of Morrey (see [6], whose proof may be found there).
Lemma 6.1. Given a function \( q \geq 0 \) on \( B_\rho \subset \mathbb{R}^2 \), such that
\[
\int_{B_\rho \cap B_\rho(y)} q \leq \beta \sigma^\gamma \quad \forall y \in B_\rho \text{ and } \sigma \in (0, \rho],
\]
then \( \forall \epsilon > 0 \exists C = C(\beta, \gamma, \epsilon) \) such that
\[
\int_{B_\rho} q|v|^2 \leq \epsilon \rho^\gamma \int_{B_\rho} |Dv|^2 + C \rho^{\gamma-2} \int_{B_\rho} |v|^2.
\]

We will now begin the proof of the theorem.

Proof. Given that \( w \in W^{3,2} \cap C^{2,\gamma} \), we can write the weak form of the equation as
\[
(39) \quad \int_D A_{\alpha \beta} D^\alpha w D^\beta f - \int_D B_{ij} D^{ij} f + \int_D B_0 D^k f = 0,
\]
where this is true \( \forall f \in W^{3,2}_0(D) \), and the coefficient matrices \( A, B, B_0 \) depend on \( x, u, Du, \) and \( D^2 u \) as described in the statement of the theorem.

This is true for all such functions; we will now pick a specific type of test function \( f \). Fixing temporarily \( h \in (0, \frac{\rho}{4}) \), and a cut-off function, \( \xi \in C_0^\infty(D) \), with \( \text{spt}(\xi) \subset B_{\frac{3\rho}{4}} \), we take
\[
f = \delta_h (\xi \bar{w}_h),
\]
where \( \bar{w}_h \) denotes \( \delta_h w - q \) (\( q \) is the quadratic polynomial which approximates \( \delta_h w \) over the annulus \( B_\rho \setminus B_{\frac{3\rho}{4}} \)). Note that we subtract off this quadratic term precisely in order to allow us to use the Poincare inequality on the annulus for \( \bar{w}_h, D\bar{w}_h = \bar{u}_h, \) and \( D\bar{u}_h \). Note further that \( D^2 \bar{u}_h = D^2 u_h \), which is precisely the term that we must bound to prove the theorem. The Poincare inequality (scaled appropriately) implies that:
\[
(40) \quad \int_A |\bar{w}_h|^2 \leq C \rho^2 \int_A |\bar{u}_h|^2,
\]
\[
(41) \quad \int_A |\bar{u}_h|^2 \leq C \rho^2 \int_A |D\bar{u}_h|^2,
\]
\[
(42) \quad \int_A |D\bar{u}_h|^2 \leq C \rho^2 \int_A |D^2 u_h|^2,
\]
where \( A \) here denotes the annulus \( B_\rho \setminus B_{\frac{3\rho}{4}} \).

Furthermore, we get additional estimates on \( \bar{w}_h \) and \( \bar{u}_h \) from the fundamental theorem of calculus and our estimates on \( w \) and \( u \). First, since \( w \in C^2 \), we know that:
\[
(43) \quad w_h(x) = \int_0^1 D_\xi w(x + she) \, ds
\]
and
\[
(44) \quad u_h(x) = \int_0^1 D_\xi u(x + she) \, ds.
\]
As a result, the sup bounds on $Dw$ and $D^2w$ give sup bounds on $w_h$, $\bar{w}_h$, $u_h$, and $\bar{u}_h$. In addition, the $L^2$ bounds on $D^3w$ yield bounds on $Du_h$ and $D\bar{u}_h$; namely:

\begin{equation}
\int_{B_{\delta}(y)} |Du_h|^2 \leq C\sigma^{2\gamma}
\end{equation}

and

\begin{equation}
\int_{B_{\delta}(y)} |D\bar{u}_h|^2 \leq C\sigma^{2\gamma},
\end{equation}

for any point $y \in D$ with $B_{\delta}(y) \subset D$.

We now put this specific test function into the weak form of the equation, (39), and use the integration by parts formula for difference quotient operators, (37). Letting $a$ denote the translation of $A$,

$$a_{\alpha\beta}(x) = A_{\alpha\beta}(x + he, u(x + he), Du(x + he)),$$

and noting that the difference quotient operators commute with differentiation, we get that

\begin{equation}
\int_{B_p} a_{\alpha\beta} D^\alpha w_h D^\alpha(\xi \bar{w}_h) + \delta_h \left( A_{\alpha\beta}(x, u, Du) \right) D^\alpha w D^\alpha(\xi \bar{w}_h)
\end{equation}

\begin{equation}
- \delta_h \left( B_{ij}(x, u, Du, D^2u) \right) D^{ij}(\xi \bar{w}_h) + \delta_h \left( B^0_{ij}(x, u, Du, D^2u) \right) D^k(\xi \bar{w}_h)
\end{equation}

\begin{equation}
= 0.
\end{equation}

At this point, we will get pointwise estimates on the coefficient matrices in our transformed equation. First, note that the sup bounds on $A$ translate into sup bounds on $a$; likewise, the ellipticity condition on $A$ implies one on $a$. Next, we will use condition (B) in the statement of the theorem to get a bound on $|\delta_h(A)|$. Since the first derivatives of $A$ in all three variables are bounded above by $L$, the mean value theorem and triangle inequality yield:

\begin{equation}
|\delta_h (A(x, u, Du))| \leq h^{-1} |A(x + he, u(x + he), Du(x + he)) - A(x, u, Du)|
\end{equation}

\begin{equation}
\leq h^{-1} \left[ |A(x + he, u(x + he), Du(x + he)) - A(x + he, u(x + he), Du(x))| + |A(x + he, u(x + he), Du(x)) - A(x, u(x), Du(x))| \right]
\end{equation}

\begin{equation}
\leq h^{-1} L \left[ |Du(x + he) - Du(x)| + |u(x + he) - u(x)| + |h| \right]
\end{equation}

\begin{equation}
\leq C \left[ 1 + |Du_h| \right].
\end{equation}

Note that in the last line we used the sup bound on $u_h$. Note further that the generic constant $C$ is independent of $h$. 

Similarly, using the bounds on $B$ we get that
\begin{equation}
|\delta_h(B)| \leq C \left[ (1 + |D^2 u|^2) (1 + |Du_h|) + (1 + |D^2 u|) (|D^2 u_h|) \right],
\end{equation}
and the same bound for $|\delta_h B^0|$. 

Next we will group the terms in equation (47) by the number of derivatives applied to the cut-off function $\xi$. We get an equation of the form:
\begin{equation}
\int_{B_p} \xi(I) + (D\xi)(II) + (D^2\xi)(III) + (D^3\xi)(IV) = 0,
\end{equation}
where the tensors $I, II, III,$ and $IV$ are given by:
\begin{align*}
I &= a_{\alpha a} D^a w_h D^a w - (\delta_h A)_{\alpha a} D^a w D^a w_h - (\delta_h B)_{ij} D^i w_h + (\delta_h B^0) \cdot \bar{u}_h, \\
II &= a(D^2 u_h, D\bar{u}_h) + (\delta_h A) (D^2 u, D\bar{u}_h) - (\delta_h B) \cdot (\bar{u}_h) + (\delta_h B^0) \bar{w}_h, \\
III &= a(D^2 u_h, \bar{u}_h) + (\delta_h A) (D^2 u, \bar{u}_h) - (\delta_h B) \cdot (\bar{w}_h), \\
IV &= a(D^2 u_h, \bar{w}_h) + (\delta_h A) (D^2 u, \bar{w}_h),
\end{align*}
where in $II, III,$ and $IV$ we have adopted the obvious notation which we will continue to use. For example, $II$ is a one-tensor and
\[ D\xi \cdot a(D^2 u_h, D\bar{u}_h) = (a_{\alpha, ijk}) (D^\alpha w_h) (D^i \bar{w}_h) (D^k \xi). \]

The next stage of the proof involves getting a handle on each of these terms. We will use the definitions above, the bounds on the coefficient matrices, equations (48) and (49), and the Cauchy-Schwarz inequality repeatedly. In particular, at one point while working on $I$ when we get a term involving $|D^2 u_h|$, we will use the absorbing (or Peter-Paul) inequality to get:
\[ 2f|D^2 u_h| \leq \varepsilon^2 |D^2 u_h| + \frac{1}{\varepsilon^2} f^2. \]

First we will deal with $I$; rewrite $I$ as $I' + a(D^3 w_h, D^3 w_h)$. Then, it follows from (51) that:
\begin{equation}
|I'| \leq |\delta_h A||D^2 u||D^2 u_h| + |\delta_h B||D\bar{u}_h| + |\delta_h B^0||\bar{u}_h|.
\end{equation}
Introducing the bounds on the coefficient matrices, it follows that:
\begin{align*}
|I'| &\leq C \left[ (1 + |Du_h|)|D^2 u||D^2 u_h| + (1 + |D^2 u|^2)(1 + |Du_h|)(|\bar{u}_h| + |D\bar{u}_h|) \\
&\quad + (1 + |D^2 u|)(|D^2 u_h|)(|\bar{u}_h| + |D\bar{u}_h|) \right] \\
&\leq 2\varepsilon^2 |D^2 u_h|^2 + C \left[ (1 + |D^2 u|^2)(1 + |Du_h|^2 + |D\bar{u}_h|^2) \right],
\end{align*}
where the second inequality uses the absorbing inequality and several applications of the Cauchy-Schwarz inequality.
In exactly the same manner, using the definitions, (51), the bounds on the coefficient matrices, (48) and (49), and the Cauchy-Schwarz inequality, we get the following bounds on \( II, III, \) and \( IV \):

\[
|II| \leq C \left[ |D^2 u_h| |D u_h| + (1 + |D u_h| |D^2 u| |D u_h| \\
+ (1 + |D^2 u|^2)(1 + |D u_h|) (|u_h| + |\omega_h|) \\
+ (1 + |D^2 u| |D^2 u_h| (|u_h| + |\omega_h|)) \right],
\]

(53)

\[
|III| \leq C \left[ |D^2 u_h| |u_h| + (1 + |D u_h| |D^2 u| |u_h| \\
+ (1 + |D^2 u|^2)(1 + |D u_h| + (1 + |D^2 u| |D^2 u_h|) |\omega_h|) \right],
\]

(54)

\[
|IV| \leq C \left[ |D^2 u_h| |\omega_h| + (1 + |D u_h| |D^2 u| |\omega_h|) \right].
\]

(55)

Now we will choose \( \xi = \psi^6 \), where

\[
\psi \in C^\infty, \operatorname{spt}(\psi) \subset B_{\frac{|h|}{4}},
\]

where \( |h| < \frac{\xi}{4} \). Note that the purpose of this is to ensure that the top derivatives of \( \xi \) can be split, via Cauchy-Schwarz, with one of the terms giving \( \xi \) itself. This will allow us to absorb bad terms back into the left-hand side of our inequalities.

Using the ellipticity condition and equation (50), we now have that

\[
\int_{B_{\rho}} \psi^6 |D^2 u_h| \leq L \left[ \int_{B_{\rho}} \psi^6 (I') + (D \psi^6)(II) + (D^2 \psi^6)(III) + (D^3 \psi^6)(IV) \right].
\]

(56)

Using our prior work on bounding \( I, II, III, \) and \( IV \), we will now attack each of the four terms on the right in equation (56). The general scheme is to bound the \( L^2 \) norm of \( D^2 u_h \) on some fixed ball by some fixed constant independent of \( |h| \); hence, we will conclude that the difference quotients converge in \( L^2 \) to a weak derivative.

It follows from equation (6) that:

\[
\int \psi^6 |I'| \leq 2\varepsilon^2 \int \psi^6 |D^2 u_h| + C \varepsilon \int \psi^6 (1 + |D^2 u|^2)(1 + |D u_h|^2 + |D u_h|^2).
\]

(57)

Next, equation (53) yields

\[
\left| \int (D \psi^6)(II) \right| \leq C \int \psi^5 |D \psi| \left[ |D^2 u_h| |D u_h| + (1 + |D u_h|) |D^2 u| |D u_h| \\
+ (1 + |D u_h|)(1 + |D^2 u|^2)|u_h| + (1 + |D u_h|)(1 + |D^2 u|^2)|\omega_h| \\
+ |D^2 u_h|(1 + |D^2 u|)|u_h| + |D^2 u_h|(1 + |D^2 u|)|\omega_h| \right].
\]
Now, on each of the three terms above including a $|D^2 u_h|$ we apply the absorbing inequality. On the two terms in the middle line, we factor out a $(1 + |D^2 u|^2)$ and apply Cauchy-Schwarz to the rest. Finally, on the other term in the first line, we isolate the $\psi^2 |D\psi| |Du_h|$ and apply Cauchy-Schwarz. Keeping in mind that our constants now depend on $\epsilon$, we have:

$$\left| \int (D\psi^6)(II) \right| \leq \epsilon^2 \int \psi^6 |D^2 u_h|^2 + C \int \psi^6 (1 + |D^2 u|^2)(1 + |Du_h|^2)$$

$$+ C \int \psi^4 |Du|^2 \left[ (1 + |D^2 u|^2)(|u_h|^2 + |\psi|^2) + |D\psi_h|^2 \right].$$

In exactly the same fashion, we can bound the other two terms in equation (56):

$$\left| \int (D^2 \psi^6)(III) \right| \leq \epsilon^2 \int \psi^6 |D^2 u_h|^2 + C \int \psi^6 (1 + |D^2 u|^2)(1 + |Du_h|^2)$$

$$+ C \int (\psi^2 |D\psi|^4 + \psi^4 |D^2 \psi|^2) \left[ (1 + |D^2 u|^2)|\psi|^2 + |Du_h|^2 \right].$$

$$\int (D^3 \psi^6)(IV) \leq \epsilon^2 \int \psi^6 |D^2 u_h|^2 + C \int \psi^6 |D^2 u|^2 (1 + |Du_h|^2)$$

$$+ C \int (|D\psi|^6 + \psi^2 |D^2 \psi|^2 |D\psi|^2 + \psi^6 |D^3 \psi|^2)|\psi|^2.$$

Now choosing an $\epsilon$ small enough, we can absorb all of the bad terms onto the left-hand side. Putting it all together, we get:

$$\int \psi^6 |D^2 u_h|^2 \leq C \int \psi^6 (1 + |D^2 u|^2)(1 + |Du_h|^2)$$

$$+ C \int \psi^4 |Du|^2 \left[ (1 + |D^2 u|^2)(|u_h|^2 + |\psi|^2) + |D\psi_h|^2 \right]$$

$$+ C \int (\psi^2 |D\psi|^4 + \psi^4 |D^2 \psi|^2) \left[ (1 + |D^2 u|^2)|\psi|^2 + |Du_h|^2 \right]$$

$$+ C \int (|D\psi|^6 + \psi^2 |D^2 \psi|^2 |D\psi|^2 + \psi^6 |D^3 \psi|^2)|\psi|^2.$$

We can arrange that the cut-off function $\psi$ is identically one in the ball $B_{\frac{\epsilon}{2}}$ and that

$$|D^{(j)} \psi| \leq C \rho^{-j}.$$ 

Note that with this choice, $|D\psi|$ has its support on the annulus, $A$. 

With this choice of \( \psi \), we can bound the terms in equation (61) which don’t include \( |D^2 u|^2 \). First, trivially we get that

\[
\int_D \psi^4 |D\psi|^2 |D\bar{u}_h|^2 \leq C \rho^{-2} \int_A |D\bar{u}_h|^2.
\]

Similarly, we get

\[
\int_D (\psi^2 |D\psi|^4 + \psi^4 |D^2 \psi|^2) \left[ |\bar{u}_h|^2 \right] \leq C \rho^{-4} \int_A |\bar{u}_h|^2 \leq C \rho^{-2} \int_A |D\bar{u}_h|^2,
\]

where the second inequality follows from the scaled version of the Poincare inequality. We repeat this procedure, this time applying the Poincare inequality twice to get

\[
\int_D (|D\psi|^6 + \psi^2 |D^2 \psi|^2 |D\psi|^2 + \psi^4 |D^3 \psi|^2) |\bar{u}_h|^2 \leq C \rho^{-2} \int_A |D\bar{u}_h|^2.
\]

Note that by equation (45) these quantities are all controlled independent of \( |h| \). It remains for us to bound the terms involving \( |D^2 u|^2 \); this is where that lemma of Morrey comes in. Note that

\[
q = (1 + |D^2 u|^2)
\]

satisfies the hypotheses of Lemma 6.1.

First, taking \( v = \psi^2 |D\psi||\bar{u}_h| \), and applying the lemma and the Poincare inequality, we get that

\[
\int \psi^4 |D\psi|^2 (1 + |D^2 u|^2) |\bar{u}_h|^2 \leq C \rho^{-2} \int_A |\bar{u}_h|^2,
\]

which is bounded by the terms already there (for example, in equation (64)).

Similarly, taking

\[
v = \psi^2 |D\psi||\bar{u}_h|,
\]

and applying the lemma and the Poincare inequality, we get that

\[
\int \psi^4 |D\psi|^2 (1 + |D^2 u|^2) |\bar{u}_h|^2 \leq C \rho^{-2} \int_A |D\bar{u}_h|^2,
\]

which is also bounded by the terms already there.

At this point, using (62) to (66), we update equation (61) to get

\[
\int \psi^6 |D^2 u|^2 \leq C \int \psi^6 (1 + |D^2 u|^2) (1 + |Du_h|^2 + |D\bar{u}_h|^2)
\]

\[
+ C \int (\psi^2 |D\psi|^4 + \psi^4 |D^2 \psi|^2) (1 + |D^2 u|^2) |\bar{u}_h|^2
\]

\[
+ C \rho^{-2} \int_A |D\bar{u}_h|^2.
\]

Applying Lemma 6.1 with

\[
v = |\bar{u}_h|(|\psi|D\psi|^2 + \psi^2 |D^2 \psi|),
\]

we can bound the second line in (67) by the the third line.
To take care of the first term, we apply Lemma 6.1 twice. First, taking
\[ v = \psi^3 (1 + |Du_h|^2)^{\frac{1}{2}}, \]
and then taking
\[ v = \psi^3 (1 + |Du_h|^2)^{\frac{1}{2}}, \]
we get that
\[ (68) \quad \int_D \psi^6 (1 + |D^2 u|^2)(1 + |Du_h|^2 + |Du_h|^2) \]
\[ \leq \epsilon \int \psi^6 |D^2 u_h|^2 + C \rho^{-2} \int |Du_h|^2 + C \rho^{-2} \int (1 + |Du_h|^2 + |Du_h|^2). \]

Choosing \( \epsilon \) small enough, and noting that \( \psi \) is identically one on \( B_{\frac{r}{2}} \), it follows from (67) and (68) that
\[ (69) \quad \int_{B_{\frac{r}{2}}} |D^2 u_h|^2 \leq C \rho^{-2} \int |Du_h|^2 + C \rho^{-2} \int (1 + |Du_h|^2 + |Du_h|^2). \]

However, by (45) the right-hand side is bounded independent of \( |h| \). Since we could have done this on any ball in \( D \), it follows that \( D^2 u_h \in L^2_{\text{loc}} \), and taking the limit as \( h \) goes to zero, \( u \in W^{3,2}_{\text{loc}} \).

Therefore, we can take the limit as \( h \to 0 \) of equation (69) to get:
\[ (70) \quad \int_{B_{\frac{r}{2}}} |D^3 u|^2 \leq C \int |D^3 u|^2 + C \rho^{-2} \int (1 + |D^2 u|^2), \]
where we applied the Poincare inequality to the first term on the right-hand side.

Since \( u \in W^{3,2}_{\text{loc}} \), the Sobolev theorem yields that \( |D^2 u| \in L^p_{\text{loc}} \forall p \geq 1 \). Therefore,
\[ \int_{B_{\frac{r}{2}}} (1 + |D^2 u|^2) \leq C \rho^{2-\delta}; \]
letting \( \delta = \frac{\rho}{2} \), equation (70) becomes:
\[ (71) \quad \int_{B_{\frac{r}{2}}} |D^3 u|^2 \leq C \int |D^3 u|^2 + C \rho^{\frac{\rho}{2}}. \]

By hole filling (adding \( C \) times the term on the right to both sides), we conclude that
\[ (72) \quad \int_{B_{\frac{r}{2}}} |D^3 u|^2 \leq \theta \int_{B_{\frac{r}{2}}} |D^3 u|^2 + C \rho^{\frac{\rho}{2}}, \]
where \( \theta \in (0, 1) \). By iterating this inequality we conclude that \( \exists \alpha > 0 \) such that
\[
\int_{B_\rho} |D^3 u|^2 \leq C \rho^{2\alpha},
\]
for any such ball in \( D \). Therefore, equation (73) implies that \( u \in C^{2, \alpha}_{(loc)} \), as desired. \( \square \)

7. FULL REGULARITY

Given the results of Theorem 4, it follows that the solution, \( w \in C^{3, \alpha} \cap W^{4,2} \), is actually \( C^\infty \), with estimates that then yield real analyticity. This is proven by the same difference quotient methods of Section 6; however, these estimates are much easier as a result of the stronger estimates on the coefficient matrices in the weak equation.

Recall that the coefficient matrices are \( A(x, z, p), B(x, z, p, q) \), where the \( B \) are homogeneous quadratic polynomials in the \( q \) and everything is smooth in \( x, z, p \). However, now that \( w \in C^{3, \alpha} \), it follows that \( u, Du, \) and \( D^2 u \) are all bounded. Therefore, all \( x, z, p, q \) derivatives of \( A \) and \( B \) are bounded in the sup norm.

Given \( w \in C^{3, \alpha} \cap W^{4,2} \), a weak solution to the equation in Theorem 4, we can choose a test function \( \delta_h \xi \), where \( \xi \) is a test function. Then use equation (37) to integrate by parts to get a new equation:
\[
\int D^\alpha \xi \left[ a_{\alpha \beta} D^\beta w_h + \delta_h (A_{\alpha \beta}) D^\beta w \right] - D^{ij} \xi \left[ \delta_h (B_{ij}) \right] + D^k \xi \delta_h (B^0_k) = 0,
\]
where we have used the notation from Section 6. Since \( w \in W^{4,2} \), all of the above terms are bounded, independent of \( |h| \), so we can take the limit as \( h \to 0 \); we get the following:
\[
\int D^\alpha \xi \left[ a_{\alpha \beta} D^\beta w' + (\partial_x A_{\alpha \beta} + u' \partial_z A_{\alpha \beta} + Du' \partial_p A_{\alpha \beta}) D^\beta w \right]
- D^{ij} \xi \left[ \partial_x (B_{ij}) + u' \partial_z (B_{ij}) + Du' \partial_p (B_{ij}) + D^2 u' \partial_q (B_{ij}) \right]
+ D^k \xi \left[ \partial_x (B_k) + u' \partial_z (B_k) + Du' \partial_p (B_k) + D^2 u' \partial_q (B_k) \right] = 0.
\]

Now choose the test function \( \delta_h \xi \) in equation (75) and integrate by parts to get
\[
\int D^\alpha \xi \left[ a_{\alpha \beta} D^\beta w' + (\partial_x A_{\alpha \beta} + u' \partial_z A_{\alpha \beta} + Du' \partial_p A_{\alpha \beta}) D^\beta w_h \right]
+ \delta_h (\partial_x A_{\alpha \beta} + u' \partial_z A_{\alpha \beta} + Du' \partial_p A_{\alpha \beta}) D^\beta w
- D^{ij} \xi \delta_h \left[ \partial_x (B_{ij}) + u' \partial_z (B_{ij}) + Du' \partial_p (B_{ij}) + D^2 u' \partial_q (B_{ij}) \right]
+ D^k \xi \delta_h \left[ \partial_x (B_k) + u' \partial_z (B_k) + Du' \partial_p (B_k) + D^2 u' \partial_q (B_k) \right] = 0,
\]
where we have the following bounds on the new coefficient matrices:

\begin{align}
|a_{\alpha\beta}| & \leq C \text{ and } a \text{ is elliptic,} \\
|\delta_h \partial_x A| & \leq C, \\
|u' \delta_h \partial_x A| & \leq C, \\
|u' \delta_h \partial_x A| & \leq C,
\end{align}

and

\begin{align}
|Du' \delta_h \partial_p A| & \leq C(1 + |D^2 u_h|).
\end{align}

Similarly, for the \( B \) terms, we get

\begin{align}
|\delta_h (B\text{-terms})| & \leq C(1 + |D^2 u_h| + |D^3 u_h|).
\end{align}

The \( D^3 u_h \) term will have to be absorbed, but all the other terms are in \( L^2 \) because \( u \in W^{3,2} \).

Now we choose a specific test function, as in Section 6,

\[ \xi = \psi^4 w_h', \]

with \( \psi \) a cutoff function having support in \( B_1 \), \( \psi \equiv 1 \) on \( B_{\frac{1}{2}} \), and

\[ |D^k \psi| \leq C \quad \forall k \leq 4, \]

where \( w_h' \) is the derivative of \( w_h \), which was defined in the proof of Theorem 4. Plugging this test function into the weak equation, (76), and using the bounds, (77) to (82), we get

\begin{align}
\int \psi^4 |D^3 w_h'|^2 \leq & C \int \psi^2 (1 + |D^2 w_h'|)|D^3 w_h'|
+ C \int \psi^4 [1 + |D^3 w_h|]|D^3 w_h'|
+ C \int \psi^4 [1 + |D^4 w_h|]^2
+ C \int [\psi^2 + \psi^4 |D^4 u_h|][1 + |D^3 w_h| + |D^3 u_h|],
\end{align}

where we have also used the fact that \( w \in C^{3,\alpha} \) to get sup bounds on up to third derivatives of \( w \). Since \( w \in W^{4,2} \), the fourth derivatives of \( w \) are in \( L^2 \); the only other terms appearing above are fifth derivatives of \( w \), and it is these that we must bound in \( L^2 \).

We now apply the absorbing inequality to bring the stray fifth derivative terms on the right back to the left of inequality (83), and we use Cauchy-Schwarz...
repeatedly to get that

\[ \int_{B_{1/2}} |D^4 w_h|^2 \leq C. \]

Hence \( w \in W^{5,2} \) and we can take the limit as \( h \to 0 \) in inequality (83). We now use hole filling and the Poincare inequality to repeat the techniques in the proof of Theorem 4 to get a decay estimate on the \( L^2 \) norm of the fifth derivative. Finally, by Morrey’s lemma, this decay implies that \( w \in C^{4,\alpha} \).

At this point, we have sup bounds on all derivatives up to the fourth, and \( L^2 \) bounds up to the fifth derivatives, so we can repeat the above procedure to get bounds on the sixth derivatives (and then decay which gives sup bounds on the fifth by Morrey, etc.). To get real analyticity, just repeat the above procedure to get bounds on each successive derivative.

8. Regularity everywhere

It remains to be shown that \( \Sigma \) is smooth at the bad points, \( \xi_1, \ldots, \xi_N \), and that \( F(\Sigma) \) achieves the infimum. It is clear that once we know that \( \Sigma \) is \( C^{1,\alpha} \cap W^{2,2} \) there, the regularity theory of Theorem 4 and Section 7 will apply.

This part of the proof is a straightforward analog of Simon’s proof in [8], and will only be sketched briefly. First, Allard’s compactness theorem (see [9]), implies that a subsequence of the \( \Sigma_j \) converges as varifolds, and hence we get convergence of first variation. This, coupled with the measure-theoretic convergence and the boundedness of \( F(\Sigma_j) \), implies that \( F(\Sigma) < \infty \). We can now apply the graphical decomposition lemma directly to \( \Sigma \), as well as the first variation identities and their consequences (see Section 2). From here, Simon proves that a sequence with these properties has lower semi-continuity of the Willmore functional. We conclude immediately that \( \Sigma \) is a lagrangian submanifold with \( F(\Sigma) \) achieving the infimum. It is therefore embedded, and hence, must be a torus. Once we know that it is minimizing, we can use the triharmonic comparison lemma, Lemma 5.1, and get the decay estimate which allowed us to prove regularity in Proposition 4.

Finally, we will show how to extend this proof to cover Theorem 3. First, note that our proof of existence and convergence, Proposition 3, only required boundedness of \( |\Sigma_j| \) and \( F(\Sigma_j) \), and that embeddedness just needed \( F(\Sigma_j) \) small; this is satisfied by the classes \( C_r \), as long as \( r \) is in the specified range. Second, the classes \( C_r \) are invariant under Hamiltonian isotopy, so the specific variations used in deriving the Euler-Lagrange equations are still allowable; hence, the higher regularity theory carries over exactly. Third, the classes \( C_r \) are preserved when we glue in a triharmonic disk (as in the proof of the \( C^{1,\alpha} \cap W^{2,2} \) result). Finally, since \( \alpha \) is an ambient form and we have convergence in Hausdorff distance, the classes will be preserved under our convergence.
LAGRANGIAN WILLMORE TORI

REFERENCES


ABSTRACT. In this paper we prove an existence and regularity theorem for lagrangian tori minimizing the Willmore functional in Euclidean four-space, $\mathbb{R}^4$, with the standard metric and symplectic structure. Technical difficulties arise because the Euler-Lagrange equation for this problem is a sixth-order nonlinear partial differential equation. This research was motivated by a study of the seemingly unrelated Plateau problem for lagrangian tori, and in this paper we illustrate this connection.

DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY, STANFORD, CALIFORNIA 94305-2125

COURANT INSTITUTE OF MATHEMATICAL SCIENCES, 251 MERCER STREET, NEW YORK, NEW YORK 10012

E-mail address: minicozz@cims.nyu.edu