THE WILLMORE FLOW WITH SMALL INITIAL ENERGY

ERNST KUWERT & REINER SCHÄTZLE

Abstract
We consider the \(L^2\) gradient flow for the Willmore functional. In [5] it was proved that the curvature concentrates if a singularity develops. Here we show that a suitable blowup converges to a nonumbilic (compact or noncompact) Willmore surface. Furthermore, an \(L^\infty\) estimate is derived for the tracefree part of the curvature of a Willmore surface, assuming that its \(L^2\) norm (the Willmore energy) is locally small. One consequence is that a properly immersed Willmore surface with restricted growth of the curvature at infinity and small total energy must be a plane or a sphere. Combining the results we obtain long time existence and convergence to a round sphere if the total energy is initially small.

1. Introduction
For a closed, immersed surface \(f : \Sigma \to \mathbb{R}^n\) the Willmore functional (as introduced initially by Thomsen [11]) is
\[
W(f) = \int_\Sigma |A^o|^2 d\mu,
\]
where \(A^o = A - \frac{1}{2} g \otimes H\) denotes the tracefree part of the second fundamental form \(A = D^2 f^\perp\) and \(\mu\) is the induced area measure. The associated Euler-Lagrange operator is
\[
W(f) = \Delta H + Q(A^o)H.
\]
Here \(H\) is the mean curvature vector and \(Q(A^o)\) acts linearly on normal vectors along \(f\) by the formula (using summation with respect to a \(g\)-orthonormal basis \(\{e_1, e_2\}\))
\[
Q(A^o)\phi = A^o(e_i, e_j)\langle A^o(e_i, e_j), \phi \rangle.
\]
In (2) the Laplace operator $\Delta \phi = -\nabla^* \nabla \phi$ is understood with respect to the connection $\nabla_X \phi = (D_X \phi)^\perp$ on normal vector fields along $f$, where $\nabla^*$ denotes the formal adjoint of $\nabla$.

In this paper we continue our study from [5] of the $L^2$ gradient flow for (1), briefly called the Willmore flow, which is the fourth order, quasilinear geometric evolution equation

$$\partial_t f = -\mathcal{W}(f).$$

As a main result we have shown in [5] that the existence time is bounded from below in terms of the concentration of the measure $f_\#(\mu, |A|^2)$ in $\mathbb{R}^n$ at time $t = 0$. Here we study the operator (2) and the flow (4) under the assumption that $\mathcal{W}(f)$ is — either locally or globally — small. This condition is natural from the variational point of view and may be interpreted geometrically by saying that the deviation of $f$ from being round is small in an averaged sense. One of our results is:

**Theorem 5.1.** There exists $\varepsilon_0(n) > 0$ such that if at time $t = 0$ we have $\mathcal{W}(f_0) < \varepsilon_0$, then the Willmore flow exists smoothly for all times and converges to a round sphere.

The smallness assumption implies, if $\varepsilon_0$ is not too big, that $\Sigma$ is topologically a sphere and that $f$ is an embedding (see [13] for the case $n = 3$). Moreover, any sequence $f_k$ with $\mathcal{W}(f_k) \to 0$ subconverges, after appropriate translation and rescaling, to some round sphere in the sense of both Hausdorff distance and measure [8]. However the $f_k$ need not be graphs over the limit sphere, as can be seen by modifying Example 1 in [12]. At present we do not know an example ruling out the possibility of dropping the smallness condition in Theorem 5.1 entirely; in any case it is desirable to replace the number $\varepsilon_0$ by a more explicit constant.\(^1\)

The statement of the theorem was recently proved in [9] under the stronger assumption that $f_0$ is close to a round sphere in the $C^{2,\alpha}$ topology, using a center manifold analysis which gives related stability results for a couple of other flows; see [2] for an overview. Our method, which is (and has to be) entirely different, involves deriving a priori estimates from the equation satisfied by the curvature, somewhat analogous to the work of Huisken [3, 4]. However, in our problem the crucial estimates are of integral type and the smallness condition is essential in

\(^1\)Note added in proof: a numerical example of a singularity was recently contributed by U. Mayer & G. Simonett: [http://www.math.utah.edu/~mayer/math/numerics.html](http://www.math.utah.edu/~mayer/math/numerics.html).
ruling out possible concentrations related to the scale invariance of the functional.

2. Estimates for surfaces with locally small Willmore energy

Here we derive some bounds for immersed surfaces \( f : \Sigma \to \mathbb{R}^n \) depending on the \( L^2 \) norms of their curvature \( A \) and of their Willmore gradient \( W(f) = \Delta H + Q(A^0)H \), under the assumption that the \( L^2 \) norm of \( A^0 \), the tracefree part of the curvature, is locally small.

Recall the equations of Mainardi-Codazzi, Gau\ss and Ricci:

\[
\begin{align*}
\nabla_X A(Y,Z) &= (\nabla_Y A)(X,Z); \quad \nabla H = -\nabla^* A = -2\nabla^* A^0, \\
K &= \frac{1}{4}|H|^2 - \frac{1}{2}|A^0|^2, \\
R^\perp(X,Y)\phi &= A^0(e_i,X)(A^0(e_i,Y),\phi) - A^0(e_i,Y)(A^0(e_i,X),\phi). 
\end{align*}
\]

Note \( (R^\perp(X,Y)\phi,\phi) = 0 \) and in particular \( R^\perp = 0 \) for \( n = 3 \), i.e., codimension one. The Codazzi equations imply that \( \nabla A \) and \( \nabla^2 A \) can be expressed by \( \nabla A^0 \) and \( \nabla^2 A^0 \), respectively. In particular one has inequalities

\[
|\nabla A| \leq c|\nabla A^0|, \quad |\nabla^2 A| \leq c|\nabla^2 A^0|.
\]

**Lemma 2.1.** For any \( p \)-linear form \( \phi \) along \( f \) we have

\[
((\nabla^* - \nabla^* \nabla) \phi)(X_1,\ldots,X_p)
= K \phi(X_1,\ldots,X_p) + K \sum_{k=2}^p \phi(X_k,X_2,\ldots,X_1,\ldots,X_p)

- K \sum_{k=2}^p g(X_1,X_k) \phi(e_i,X_2,\ldots,e_i,\ldots,X_p)

+ R^\perp(e_i,X_1) \phi(e_i,X_2,\ldots,X_p) - (\nabla^* T)(X_1,\ldots,X_p).
\]

Here the tensor \( T \) is given by

\[
T(X_0,X_1,\ldots,X_p) = (\nabla_{X_0} \phi)(X_1,X_2,\ldots,X_p) - (\nabla_{X_1} \phi)(X_0,X_2,\ldots,X_p).
\]

**Proof.** From the proof of Lemma 2.1 in [5] we have

\[
((\nabla^* - \nabla^* \nabla) \phi)(X_1,\ldots,X_p)
= (R^p(e_i,X_1) \phi)(e_i,X_2,\ldots,X_p) - (\nabla^* T)(X_1,\ldots,X_p).
\]
Now the curvature operator $R^p$ is given by

\[
(R^p(e_i, X_1) \phi)(e_i, X_2, \ldots, X_p) = R^\perp(e_i, X_1) \phi(e_i, X_2, \ldots, X_p) \]

\[\quad - \phi(R(e_i, X_1) e_i, X_2, \ldots, X_p) \]

\[\quad - \sum_{k=2}^p \phi(e_i, X_2, \ldots, R(e_i, X_1) X_k, \ldots, X_p) \]

\[= R^\perp(e_i, X_1) \phi(e_i, X_2, \ldots, X_p) \]

\[\quad - K \phi(g(X_1, e_i) e_i, X_2, \ldots, X_p) + K \phi(g(e_i, e_i) X_1, X_2, \ldots, X_p) \]

\[\quad - K \sum_{k=2}^p \phi(e_i, X_2, \ldots, g(X_1, X_k) e_i, \ldots, X_p) \]

\[+ K \sum_{k=2}^p \phi(g(e_i, X_k) e_i, X_2, \ldots, X_1, \ldots, X_p). \]

Inserting yields the desired formula. q.e.d.

We will need three different choices for $\phi$ in (9). Taking first $\phi = A$ yields $T = 0$ and $\nabla^\ast \phi = -\nabla H$ by (5), and we get Simons’ identity ([10])

\[
\Delta A = \nabla^2 H + 2KA^\circ + R^\perp(e_i, \cdot) A(e_i, \cdot). \]

To bring this in a more useful form, let us denote by $S^\circ(B)$ the symmetric, tracefree part of any bilinear form with normal values along $f$. In particular, we have

\[
S^\circ(\nabla^2 H) = \nabla^2 H - \frac{1}{2} g(\cdot, \cdot) \Delta H - \frac{1}{2} R^\perp(\cdot, \cdot) H. \]

Now \[
\Delta \left( \frac{1}{2} g(\cdot, \cdot) H \right) = \frac{1}{2} g(\cdot, \cdot) \Delta H \quad \text{and} \quad R^\perp(e_i, X) \frac{1}{2} g(e_i, Y) H = -\frac{1}{2} R^\perp(X, Y) H, \]

which implies, using (6) and (7),

\[
\Delta A^\circ = S^\circ(\nabla^2 H) + \frac{1}{2} |H|^2 A^\circ + A^\circ * A^\circ * A^\circ. \]

Here and in the following we denote by $A * B$ any universal, linear combination of tensors obtained by tensor product and contraction from $A$ and $B$. Our second choice in (9) is $\phi = \nabla H$, where now

\[
T(X, Y) = \nabla^2_{X,Y} H - \nabla^2_{Y,X} H = R^\perp(X, Y) H. \]
Using again (5), (6) and (7), we infer
\[
\nabla^* (\nabla^2 H) = \nabla (\nabla^* \nabla H) - \frac{1}{4} |H|^2 \nabla H + A * A^\circ * \nabla A^\circ.
\]
Finally taking \( \phi = \nabla A^\circ \) in (9) yields
\[
T(X, Y, Z, V) = \left( R^2(X, Y) A^\circ \right) (Z, V) = (A * A * A)(X, Y, Z, V), \\
\nabla^* T = A * A * \nabla A^\circ.
\]
Thus we obtain from (9) and (6), (7)
\[
\nabla^* (\nabla^2 A^\circ) = \nabla (\nabla^* \nabla A^\circ) + A * A * \nabla A^\circ.
\]
We now convert (10), (11) and (12) into integral estimates.

**Lemma 2.2.** If \( f : \Sigma \to \mathbb{R}^n \) is an immersion with \( W(f) = W \) and \( \gamma \in C^1_c(\Sigma) \) satisfies \( |\nabla \gamma| \leq \Lambda \), then
\[
\int |\nabla A|^2 \gamma^2 \, d\mu \leq \frac{c}{\Lambda^2} \int |W|^2 \gamma^4 \, d\mu + c \int |A^\circ|^4 \gamma^2 \, d\mu + c \Lambda^2 \int \gamma \nabla \gamma \, d\mu.
\]

**Proof.** Multiply (10) by \( \gamma^2 A^\circ \) and integrate by parts to obtain, after applying (5),
\[
\int |\nabla A|^2 \gamma^2 \, d\mu + \frac{1}{2} \int |H|^2 |A^\circ|^2 \gamma^2 \, d\mu
\leq \frac{1}{2} \int |\nabla H|^2 \gamma^2 \, d\mu + c \int |A^\circ|^4 \gamma^2 \, d\mu + \int \gamma \nabla \gamma * A^\circ * \nabla A^\circ \, d\mu.
\]
Using the equation \( \Delta H + Q(A^\circ)H = W \) we have
\[
\frac{1}{2} \int |\nabla H|^2 \gamma^2 \, d\mu = -\frac{1}{2} \int \langle H, \Delta H \rangle \gamma^2 \, d\mu + \int \gamma \nabla \gamma * A * \nabla A^\circ \, d\mu
\]
\[
= -\frac{1}{2} \int \langle H, W \rangle \gamma^2 \, d\mu + \frac{1}{2} \int \langle H, Q(A^\circ)H \rangle \gamma^2 \, d\mu \\
+ \int \gamma \nabla \gamma * A * \nabla A^\circ \, d\mu
\leq \frac{c}{\Lambda^2} \int |W|^2 \gamma^4 \, d\mu + c \Lambda^2 \int |H|^2 \, d\mu \\
+ \frac{1}{2} \int \langle H, Q(A^\circ)H \rangle \gamma^2 \, d\mu + \int \gamma \nabla \gamma * A * \nabla A^\circ \, d\mu.
It is easy to see the inequality
\begin{equation}
0 \leq \langle Q(A^o)H, H \rangle \leq |A^o|^2 |H|^2.
\end{equation}
Furthermore we have
\begin{equation}
\int \gamma \nabla \gamma \ast A \ast \nabla A^o \, d\mu \leq \frac{1}{2} \int |\nabla A^o|^2 \gamma^2 \, d\mu + c \Lambda^2 \int [\gamma > 0] |A|^2 \, d\mu.
\end{equation}
Inserting these inequalities, absorbing and recalling (8) proves the claim.

**q.e.d.**

**Lemma 2.3.** Under the assumptions of Lemma 2.2 we have for \( \eta = \gamma^4 \)
\begin{equation}
\int |\nabla^2 H|^2 \eta \, d\mu + \int |A|^2 |\nabla A|^2 \eta \, d\mu + \int |A|^4 |A^o|^2 \eta \, d\mu \\
\leq c \int |W|^2 \eta \, d\mu + c \int (|A^o|^2 |\nabla A^o|^2 + |A^o|^6) \eta \, d\mu \\
+ c \Lambda^4 \int [\gamma > 0] |A|^2 \, d\mu.
\end{equation}

**Proof.** We start multiplying (11) by \( \eta \nabla H \) and integrating by parts. This yields

\begin{align*}
\int |\nabla^2 H|^2 \eta \, d\mu + & \frac{1}{4} \int |H|^2 |\nabla H|^2 \eta \, d\mu \\
\leq & \int |\Delta H|^2 \eta \, d\mu + \int A \ast A^o \ast \nabla A^o \ast \nabla A^o \eta \, d\mu \\
& + \int \gamma^3 \nabla \gamma \ast \nabla H \ast \nabla^2 H \, d\mu \\
\leq & c \int |W|^2 \eta \, d\mu + c \int |A^o|^4 |H|^2 \eta \, d\mu \\
& + \varepsilon \int |H|^2 |\nabla A^o|^2 \eta \, d\mu + c(\varepsilon) \int |A^o|^2 |\nabla A^o|^2 \eta \, d\mu \\
& + \frac{1}{2} \int |\nabla^2 H|^2 \eta \, d\mu + c \Lambda^2 \int |\nabla H|^2 \gamma^2 \, d\mu.
\end{align*}

Now by (13) we can estimate
\begin{equation}
\int |\nabla H|^2 \gamma^2 \, d\mu \leq \frac{c}{\Lambda^2} \int |W|^2 \eta \, d\mu + \frac{c}{\Lambda^2} \int [\gamma > 0] |A^o|^6 \eta \, d\mu + c \Lambda^2 \int [\gamma > 0] |A|^2 \, d\mu.
\end{equation}
Using the inequality $c |A^\circ|^4 |H|^2 \leq \varepsilon |H|^4 |A^\circ|^2 + c(\varepsilon) |A^\circ|^6$ and rearranging, we arrive at

\begin{align}
\int |\nabla^2 H|^2 \eta \, d\mu + \int |H|^2 |\nabla H|^2 \eta \, d\mu \\
\leq c \int |W|^2 \eta \, d\mu + c \Lambda^4 \int |A|^2 \, d\mu & \\
+ c(\varepsilon) \int |A^\circ|^2 |\nabla A^\circ|^2 \eta \, d\mu + c(\varepsilon) \int |A^\circ|^6 \eta \, d\mu & \\
+ \varepsilon \int |H|^4 |A^\circ|^2 \eta d\mu + \varepsilon \int |H|^2 |\nabla A^\circ|^2 \eta d\mu. \tag{16}
\end{align}

Next we use (10) to compute

\begin{align}
\int |H|^2 |\nabla A^\circ|^2 \eta d\mu & \\
= -\int |H|^2 \langle A^\circ, \Delta A^\circ \rangle \eta \, d\mu + \int H * \nabla H * A^\circ * \nabla A^\circ \eta \, d\mu & \\
+ \int |H|^2 A^\circ * \nabla A^\circ * \nabla \eta \, d\mu & \\
= -\int |H|^2 \langle A^\circ, \nabla^2 H \rangle + \frac{1}{2} |H|^2 A^\circ + A^\circ * A^\circ * A^\circ \eta \, d\mu & \\
+ \int H * \nabla H * A^\circ * \nabla A^\circ \eta \, d\mu + \int |H|^2 A^\circ * \nabla A^\circ * \nabla \eta \, d\mu & \\
\leq \frac{1}{2} \int |H|^2 |\nabla H|^2 \eta \, d\mu + \int H * \nabla H * A^\circ * \nabla A^\circ \eta \, d\mu & \\
+ \int |H|^2 A^\circ * \nabla A^\circ * \gamma \nabla \gamma \, d\mu & \\
- \frac{1}{2} \int |H|^4 |A^\circ|^2 \eta \, d\mu + c \int |H|^2 |A^\circ|^4 \eta \, d\mu & \\
\leq \left( \frac{1}{2} + \delta \right) \int |H|^2 |\nabla H|^2 \eta \, d\mu + c(\delta) \int |A^\circ|^2 |\nabla A^\circ|^2 \eta \, d\mu & \\
+ \delta \int |H|^2 |\nabla A^\circ|^2 \eta \, d\mu + c(\delta) \Lambda^2 \int |H|^2 |A^\circ|^2 \gamma^2 \, d\mu & \\
- \frac{1}{2} \int |H|^4 |A^\circ|^2 \eta \, d\mu + c \int |H|^2 |A^\circ|^4 \eta \, d\mu.
\end{align}
From the inequalities
\[ c \int |H|^2 |A^o|^4 \eta \, d\mu \leq \delta \int |H|^4 |A^o|^2 \eta \, d\mu + c(\delta) \int |A^o|^{6} \eta \, d\mu, \]
\[ c(\delta) \Lambda^2 \int |H|^2 |A^o|^2 \gamma^2 \, d\mu \leq \delta \int |H|^4 |A^o|^2 \eta \, d\mu + c(\delta) \Lambda^4 \int |A^o|^2 \, d\mu, \]
we see that
\[ (17) \quad (1 - \delta) \int |H|^2 |\nabla A^o|^2 \eta \, d\mu + \left( \frac{1}{2} - 2\delta \right) \int |H|^4 |A^o|^2 \eta \, d\mu \]
\[ \leq \left( \frac{1}{2} + \delta \right) \int |H|^2 |\nabla H|^2 \eta \, d\mu + c(\delta) \left( \int |A^o|^2 |\nabla A^o|^2 \eta \, d\mu + \int |A^o|^6 \eta \, d\mu + \Lambda^4 \int |A^o|^2 \, d\mu \right). \]

Adding the inequalities (16) and (17) yields
\[ \int |\nabla^2 H|^2 \eta \, d\mu + \left( \frac{1}{2} - \delta \right) \int |H|^2 |\nabla H|^2 \eta \, d\mu \]
\[ + (1 - \delta - \varepsilon) \int |H|^2 |\nabla A^o|^2 \eta \, d\mu + \left( \frac{1}{2} - 2\delta - \varepsilon \right) \int |H|^4 |A^o|^2 \eta \, d\mu \]
\[ \leq c(\delta, \varepsilon) \left( \int |A^o|^2 |\nabla A^o|^2 \eta \, d\mu + \int |A^o|^6 \eta \, d\mu + \Lambda^4 \int |A^o|^2 \, d\mu \right) \]
\[ + c \int |W|^2 \eta \, d\mu. \]

The claim of the Lemma follows by choosing \( \varepsilon = \delta = \frac{1}{8} \). q.e.d.

**Proposition 2.4.** If \( f : \Sigma \to \mathbb{R}^n \) is an immersion with \( W(f) = W \) and \( \eta = \gamma^4 \), where \( \gamma \in C^1_c(\Sigma) \) satisfies \( |\nabla \gamma| \leq \Lambda \), then
\[ \int |\nabla^2 A|^2 \eta \, d\mu + \int |A|^2 |\nabla A|^2 \eta \, d\mu + \int |A|^4 |A^o|^2 \eta \, d\mu \]
\[ \leq c \int |W|^2 \eta \, d\mu + c \Lambda^4 \int |A|^2 \, d\mu + c \int (|A^o|^2 |\nabla A^o|^2 + |A^o|^6) \eta \, d\mu. \]
Proof. Multiply (12) by $\eta \nabla A^\circ$, integrate by parts and apply (10) to get

$$
\int |\nabla^2 A^\circ|^2 \eta d\mu \leq \int |\Delta A^\circ|^2 \eta d\mu + c \int |A|^2 |\nabla A|^2 \eta d\mu \\
+ \int \gamma^3 \nabla \gamma \ast \nabla A^\circ \ast \nabla^2 A^\circ d\mu \\
\leq c \int |\nabla^2 H|^2 \eta d\mu \\
+ c \int |A|^2 |\nabla A|^2 d\mu + c \int |A|^4 |A^\circ|^2 \eta d\mu \\
+ \frac{1}{2} \int |\nabla^2 A^\circ|^2 \eta d\mu + c \Lambda^2 \int |\nabla A^\circ|^2 \gamma d\mu.
$$

The claim now follows from Lemma 2.2 and Lemma 2.3, recalling (8).

q.e.d.

We next need a multiplicative Sobolev inequality.

**Lemma 2.5.** Under the assumptions of Proposition 2.4 we have

$$
\int (|A^\circ|^2 |\nabla A^\circ|^2 + |A^\circ|^6) \eta d\mu \\
\leq c \int |A^\circ|^2 d\mu \cdot \int (|\nabla^2 A|^2 + |A|^2 |\nabla A|^2 + |A|^4 |A^\circ|^2) \eta d\mu \\
+ c \Lambda^4 \left( \int_{\gamma > 0} |A^\circ|^2 d\mu \right)^2.
$$

Proof. Recall the Michael-Simon Sobolev inequality ([7])

$$
\left( \int u^2 d\mu \right)^{\frac{1}{2}} \leq c \left( \int |\nabla u| d\mu + \int |H| |u| d\mu \right).
$$

THE WILLMORE FLOW
with \( c = c(n) \). Letting \( u = |A^0| |\nabla A^0| \gamma^2 \) we obtain

\[
\int |A^0|^2 |\nabla A^0|^2 \eta \, d\mu \leq c \left( \int |A^0| |\nabla^2 A^0| \gamma^2 \, d\mu \right)^2 + c \left( \int |A^0|^2 \gamma^2 \, d\mu \right)^2
+ c \left( \int |A| |A^0| |\nabla A^0| \gamma^2 \, d\mu \right)^2 + c \left( \int |A^0| |\nabla A^0| \gamma |\nabla \gamma| \, d\mu \right)^2
\]

\[
\leq c \int |A^0|^2 \, d\mu \int (|\nabla^2 A^0|^2 + |A|^2 |\nabla A^0|^2) \eta \, d\mu
+ c \Lambda^4 \left( \int |A^0|^2 \, d\mu \right)^2 + c \left( \int |\nabla A^0|^2 \gamma^2 \, d\mu \right)^2.
\]

In the last term, we integrate by parts to get

\[
(20) \quad \int |\nabla A^0|^2 \gamma^2 \, d\mu \leq c \int |A^0| |\nabla A^0|^2 \gamma^2 \, d\mu + c \Lambda \int |A^0| |\nabla A^0| \gamma \, d\mu
\]

\[
\leq c \left( \int |A^0|^2 \, d\mu \right)^{1/2} \left( \int |\nabla^2 A^0|^2 \eta \, d\mu \right)^{1/2}
+ \frac{1}{2} \int |\nabla A^0|^2 \gamma^2 \, d\mu + c \Lambda^2 \int |A^0|^2 \, d\mu.
\]

Absorbing and inserting proves the claimed inequality for the first term in (18). For the other term, choose \( u = |A^0|^3 \gamma^2 \) in (19) and compute

\[
\int |A^0|^6 \eta \, d\mu
\]

\[
\leq c \left( \int |A^0|^2 |\nabla A^0| \gamma^2 \, d\mu + \int |A| |A^0|^3 \gamma^2 \, d\mu + c \Lambda \int |A^0|^3 \gamma \, d\mu \right)^2
\]

\[
\leq c \left( \int |\nabla A^0|^2 \gamma^2 \, d\mu \right)^2 + c \int |A^0|^2 \, d\mu \int |A|^2 |A^0|^4 \eta \, d\mu
+ c \Lambda^4 \left( \int |A^0|^2 \, d\mu \right)^2.
\]

Combining with (20) proves the estimate for the second term on the left of (18). q.e.d.
Proposition 2.6. Let $f : \Sigma \to \mathbb{R}^n$ be an immersed surface, and let $\Lambda = \|\nabla \gamma\|_{L^\infty}$, where $\gamma$ has compact support on $\Sigma$. There exists a constant $\varepsilon_0 = \varepsilon_0(n) > 0$ such that if
\[
\int_{\{\gamma > 0\}} |A^\circ|^2 d\mu < \varepsilon_0,
\]
then we have for a constant $c = c(n) < \infty$
\[
\int\left(||\nabla^2 A|^2 + |A|^2 |\nabla A|^2 + |A|^4 |A^\circ|^2\right) \gamma^4 d\mu 
\leq c \int |W(f)|^2 \gamma^4 d\mu + c \Lambda^4 \int |A|^2 d\mu.
\]

This is an immediate consequence of Proposition 2.4 and Lemma 2.5. As a first application we deduce the following result.

Theorem 2.7 (Gap Lemma). Let $f : \Sigma \to \mathbb{R}^n$ be a properly immersed (compact or noncompact) Willmore surface, and let $\Sigma_{\varphi}(0) = f^{-1}(B_{\varphi}(0))$. If
\[
\liminf_{\varphi \to \infty} \frac{1}{\varphi^4} \int_{\Sigma_{\varphi}(0)} |A|^2 d\mu = 0, \quad \text{and}
\int_{\Sigma} |A^\circ|^2 d\mu < \varepsilon_0 = \varepsilon_0(n),
\]
then $f$ is an embedded plane or sphere.

Proof. We take $\gamma(p) = \varphi\left(\frac{1}{\varphi} |f(p)|\right)$, where $\varphi \in C^1(\mathbb{R})$ satisfies $\varphi(s) = 1$ for $s \leq \frac{1}{2}$, $\varphi(s) = 0$ for $s \geq 1$ and $\varphi \geq 0$. Then we have $\Lambda = c/\varphi$ in Proposition 2.6. Since $W(f) = 0$ by assumption, we can let $\varphi \to \infty$ and conclude $A^\circ \equiv 1$. This implies, by a standard result of differential geometry [13], that $f$ maps into a fixed, round 2-sphere or plane $S \subset \mathbb{R}^n$. As $f$ is complete, it follows that $f : (\Sigma, g) \to S$ is a global isometry.

We shall now derive an $L^\infty$ bound for $A^\circ$ from Proposition 2.6.

Lemma 2.8. For $\gamma \in C^1_c(\Sigma)$ with $|\nabla \gamma| \leq \Lambda$ and any normal $p$-form $\phi$ along $f$ we have the inequality
\[
\|\gamma^2 \phi\|_{L^\infty}^4 \leq c \|\gamma^2 \phi\|_{L^2}^2 \left[ \int\left(||\nabla^2 \phi|^2 + |H|^4 |\phi|^2\right) \gamma^4 d\mu + \lambda^4 \int_{\{\gamma > 0\}} |\phi|^2 d\mu \right].
\]
Proof. This is Lemma 4.3 in [5], except that there a bound on the second derivatives of \( \gamma \) was assumed. Letting \( \psi = \gamma^2 \phi \) we apply Theorem 5.6 in [5] to obtain

\[
\| \psi \|_{L^\infty}^6 \leq c \| \psi \|_{L^2}^2 \left( \| \nabla \psi \|_{L^4}^4 + \| H \psi \|_{L^4}^4 \right)
\]

\[
\leq c \| \psi \|_{L^2}^2 \left( \int \gamma^8 |\nabla \phi|^4 \, d\mu + \Lambda^4 \int \gamma^4 |\phi|^4 \, d\mu + \int |H|^4 |\psi|^4 \, d\mu \right).
\]

The three integrals on the right are estimated as follows (starting with the third):

\[
\int |H|^4 |\phi|^4 \, d\mu \leq \| \psi \|_{L^\infty}^2 \int |H|^4 |\phi|^2 \gamma^4 \, d\mu,
\]

\[
\Lambda^4 \int \gamma^4 |\phi|^4 \, d\mu \leq \| \psi \|_{L^\infty}^2 \Lambda^4 \int \gamma^2 \, d\mu.
\]

By partial integration, we infer

\[
\int |\nabla \phi|^2 \gamma^2 \, d\mu \leq c \int |\phi||\nabla^2 \phi| \gamma^2 \, d\mu + c\Lambda \int |\phi||\nabla \phi| \gamma \, d\mu
\]

\[
\leq \frac{c}{\Lambda^2} \int |\nabla^2 \phi|^2 \gamma^4 \, d\mu + c\Lambda^2 \int |\phi|^2 \, d\mu + \frac{1}{2} \int |\nabla \phi|^2 \gamma^2 \, d\mu.
\]

Using again integration by parts and Cauchy-Schwarz

\[
\int |\nabla \phi|^4 \gamma^8 \, d\mu \leq c \left( \int |\phi||\nabla^2 \phi|^2 |\nabla \phi| \gamma^8 \, d\mu + \Lambda \int |\phi||\nabla \phi|^3 \gamma^7 \, d\mu \right)
\]

\[
\leq c \| \psi \|_{L^\infty} \left( \int |\nabla \phi|^4 \gamma^8 \, d\mu \right)^\frac{1}{2} \left( \int |\nabla^2 \phi|^2 \gamma^4 \, d\mu \right)^\frac{1}{2}
\]

\[
+ c\Lambda \| \psi \|_{L^\infty} \left( \int |\nabla \phi|^4 \gamma^8 \, d\mu \right)^\frac{1}{2} \left( \int |\nabla \phi|^2 \gamma^2 \, d\mu \right)^\frac{1}{2}.
\]

Combining the last two inequalities, we get

\[
\int |\nabla \phi|^4 \gamma^8 \, d\mu \leq c \| \psi \|_{L^\infty} \left( \int |\nabla^2 \phi|^2 \gamma^4 \, d\mu + c\Lambda^4 \int |\phi|^2 \, d\mu \right).
\]

Inserting (22)–(24) into (21) proves the claim. q.e.d.

Combining Proposition 2.6 and Lemma 2.8, where \( \phi = A^\circ \) and \( \gamma \) is a cutoff function depending on extrinsic distance as in Theorem 2.7, we obtain the following “partial” curvature estimate.
Theorem 2.9 (Tracefree Curvature Estimate). Let $f : \Sigma \to \mathbb{R}^n$ be an immersed surface with $\Sigma_\varrho = f^{-1}(B_\varrho(x_0)) \subset \subset \Sigma$, and suppose that

$$\int_{\Sigma_\varrho} |A^\circ|^2 \, d\mu < \varepsilon_0,$$

where $\varepsilon_0 = \varepsilon_0(n) > 0$ is a fixed constant. Then

$$\|A^\circ\|_{L^\infty(\Sigma_\varrho/\varrho)} \leq c \left( \|\nabla W(f)\|_{L^2(\Sigma_\varrho)} + \frac{1}{\varrho^2} \|A\|_{L^2(\Sigma_\varrho)} \right) \|A^\circ\|_{L^2(\Sigma_\varrho)}.$$  \hspace{1cm} (25)

Assuming smallness of the full second fundamental form $A$, one easily adapts the arguments above to also prove the following:

Theorem 2.10 (Curvature Estimate). Let $f : \Sigma \to \mathbb{R}^n$ be an immersed surface, $\Sigma_\varrho = f^{-1}(B_\varrho(x_0)) \subset \subset \Sigma$ and suppose

$$\int_{\Sigma_\varrho} |A|^2 \, d\mu < \varepsilon_0,$$

where $\varepsilon_0 = \varepsilon_0(n)$ is a fixed constant. Then

$$\|A\|_{L^\infty(\Sigma_\varrho/\varrho)} \leq c \left( \|\nabla W(f)\|_{L^2(\Sigma_\varrho)} + \frac{1}{\varrho^2} \|A\|_{L^2(\Sigma_\varrho)} \right) \|A\|_{L^2(\Sigma_\varrho)}.$$  \hspace{1cm} (26)

Remark 2.11. The statements of the Theorems 2.7, 2.9 and 2.10 clearly also hold with the extrinsic distance sets $\Sigma_\varrho(x_0)$ replaced by distance sets with respect to the intrinsic distance function, since only a bound on the first derivatives of the cutoff function was needed.

3. Local estimates for the flow

We now consider solutions $f : \Sigma \times [0,T) \to \mathbb{R}^n$ to the gradient flow for the Willmore integral,

$$\partial_t f = -W(f).$$

We abbreviate $W(f) =: W$ in the following and compute first a precise
formula for the evolution of the energy density. Recall from [5]:

\begin{equation}
\partial_t (d\mu) = \langle H, W \rangle \, d\mu,
\end{equation}

\begin{equation}
\partial^\perp_t H = - \left( \Delta W + Q(A^\circ)W + \frac{1}{2}H\langle H, W \rangle \right),
\end{equation}

\begin{equation}
\partial^\perp_t A(X, Y) = -\nabla_{X,Y}^2 W + \frac{1}{2} g(X, Y) \left[ Q(A^\circ)W + \frac{1}{2}H\langle H, W \rangle \right] \\
+ \frac{1}{2}H\langle A^\circ(X, Y), W \rangle + \frac{1}{2}A^\circ(X, Y)\langle H, W \rangle \\
+ \frac{1}{2}R^\perp(X, Y)W.
\end{equation}

Here we used (2.18), (2.6) and (2.3) from [5]. Furthermore, using (2.15) in [5] we infer

\begin{equation}
\partial^\perp_t \left( \frac{1}{2} g(X, Y) H \right) = - \frac{1}{2} g(X, Y) \left( \Delta W + Q(A^\circ)W + \frac{1}{2}H\langle H, W \rangle \right) \\
+ \langle A^\circ(X, Y), W \rangle H + \frac{1}{2}g(X, Y)H\langle H, W \rangle,
\end{equation}

and subtracting this from (29) yields

\begin{equation}
\partial^\perp_t A^\circ(X, Y) = -S^\circ(\nabla^2 W) + g(X, Y)Q(A^\circ)W \\
+ \frac{1}{2}A^\circ(X, Y)\langle H, W \rangle - \frac{1}{2}H\langle A^\circ(X, Y), W \rangle.
\end{equation}

Recall that $S^\circ(\ldots)$ denotes the symmetric, tracefree component. We compute separately for $H$ and $A^\circ$. By (27) and (28)

\begin{equation}
\partial_t \left( \frac{1}{2} |H|^2 \, d\mu \right) \\
= - \left( \Delta W + Q(A^\circ)W + \frac{1}{2}H\langle H, W \rangle, H \right) \, d\mu
\end{equation}

\begin{equation}
= - (\Delta H + Q(A^\circ)H, W) \, d\mu + (\langle \Delta H, W \rangle - \langle H, \Delta W \rangle) \, d\mu
\end{equation}

\begin{equation}
= - |W|^2 \, d\mu + \nabla_{e_i}(\langle \nabla_{e_i} H, W \rangle - \langle H, \nabla_{e_i} W \rangle) \, d\mu,
\end{equation}

whence

\begin{equation}
\partial_t \left( \frac{1}{2} |H|^2 \, d\mu \right) + |W|^2 \, d\mu = (\nabla^* \alpha) \, d\mu,
\end{equation}

where $\alpha$ is the 1-form given by

\begin{equation}
\alpha(X) = \nabla_X \langle H, W \rangle - 2\langle \nabla_X H, W \rangle.
\end{equation}
In order to compute for $A^\circ$, we first have (using again (2.15) in [5]) for a $g$-orthonormal basis
\[
g(\partial_t e_i, e_j) + g(e_i, \partial_t e_j) = \partial_t (g(e_i, e_j)) - (\partial_t g)(e_i, e_j)
\]
\[
= -2\delta_j A(e_i, e_j), W)
\]
\[
= -2\langle A^\circ(e_i, e_j), W \rangle - \delta_j \langle H, W \rangle.
\]
This implies further
\[
\langle A^\circ(\partial_t e_i, e_k), A^\circ(e_i, e_k) \rangle
\]
\[
= g(\partial_t e_i, e_j) \langle A^\circ(e_i, e_k), A^\circ(e_i, e_k) \rangle
\]
\[
= -g(\partial_t e_i, e_j) \langle A^\circ(e_i, e_k), A^\circ(e_i, e_k) \rangle - \frac{1}{2} \delta_j \langle H, W \rangle
\]
\[
= -\langle A^\circ(e_i, e_k) \langle A^\circ(e_i, e_j), W \rangle, A^\circ(e_j, e_k) \rangle - \frac{1}{2} |A^\circ|^2 \langle H, W \rangle
\]
\[
= -\left\langle \frac{1}{2} g(e_j, e_k) Q(A^\circ) W, A^\circ(e_j, e_k) \right\rangle - \frac{1}{2} |A^\circ|^2 \langle H, W \rangle
\]
\[
= -\frac{1}{2} |A^\circ|^2 \langle H, W \rangle,
\]
where we used (2.5) from [5]. We use this and (30) to compute
\[
\partial_t (|A^\circ|^2 d\mu)
\]
\[
= 2\langle (\partial_t A^\circ)(e_i, e_k), A^\circ(e_i, e_k) \rangle d\mu
\]
\[
+ 2\langle A^\circ(\partial_t e_i, e_k) + A^\circ(e_i, \partial_t e_k), A^\circ(e_i, e_k) \rangle d\mu + |A^\circ|^2 \langle H, W \rangle d\mu
\]
\[
= -2\langle \nabla^2 W, A^\circ \rangle d\mu + |A^\circ|^2 \langle H, W \rangle d\mu
\]
\[
- \langle A^\circ(e_i, e_k), W \rangle \langle A^\circ(e_i, e_k), H \rangle d\mu
\]
\[
- 2|A^\circ|^2 \langle H, W \rangle d\mu + |A^\circ|^2 \langle H, W \rangle d\mu
\]
\[
= -2\langle \nabla^2 W, A^\circ \rangle d\mu - \langle Q(A^\circ) H, W \rangle d\mu
\]
\[
= \left( -2 \nabla e_i (\nabla e_j W, A^\circ(e_i, e_j)) + \langle \nabla e_i W, \nabla e_j H \rangle \right) d\mu
\]
\[
- \langle Q(A^\circ) H, W \rangle d\mu
\]
\[
= \left( -2 \nabla e_i (\nabla e_j W, A^\circ(e_i, e_j)) + \langle \nabla e_i W, \nabla e_j H \rangle \right) d\mu
\]
\[
- \langle W, \Delta H \rangle d\mu - \langle Q(A^\circ) H, W \rangle d\mu.
\]

Thus we have shown
\[
\partial_t (|A^\circ|^2 d\mu) + |W|^2 d\mu = (\nabla^* \beta) d\mu,
\]
where $\beta$ is the 1-form defined by
\[
\beta(X) = 2\langle \nabla e_j W, A^\circ(X, e_j) \rangle - \langle \nabla X H, W \rangle.
\]
Lemma 3.1. If \( f \) is a Willmore flow, then for any function \( \eta \) and \( W = W(f) \) we have:

\[
\frac{1}{2} \frac{\partial}{\partial t} \int |H|^2 \eta \, d\mu + \int |W|^2 \eta \, d\mu = \int \left( \frac{1}{2} |H|^2 \partial_t \eta - \langle H, W \rangle \Delta \eta - 2 \langle \nabla_{\nabla \eta} H, W \rangle \right) \, d\mu,
\]

\[
\frac{1}{2} \frac{\partial}{\partial t} \int |A^0|^2 \eta \, d\mu + \int |W|^2 \eta \, d\mu = \int \left( |A^0|^2 \partial_t \eta - 2 \langle A^0(e_i, e_j), W \rangle \nabla_{e_i,e_j} \eta - 2 \langle \nabla_{\nabla \eta} H, W \rangle \right) \, d\mu.
\]

Proof. Formula (35) is immediate from (31) and (32). For (36) we compute for \( \beta \) as in (34):

\[
\int \eta \nabla^* \beta \, d\mu = \int \left( 2 \langle \nabla_{e_j} W, A^0(\nabla \eta, e_j) \rangle - \langle \nabla_{\nabla \eta} H, W \rangle \right) \, d\mu
\]

\[
= - \int 2 \langle \nabla_{e_j} A^0, e_j, \nabla \eta \rangle \, d\mu
\]

\[
- \int 2 \langle A^0(\nabla_{e_j} \nabla \eta, e_j), W \rangle \, d\mu - \int \langle \nabla_{\nabla \eta} H, W \rangle \, d\mu
\]

\[
= - \int 2 \nabla_{e_i,e_j} \eta \langle A^0(e_i, e_j), W \rangle \, d\mu - \int 2 \langle \nabla_{\nabla \eta} H, W \rangle \, d\mu,
\]

which, together with (33), proves (36). q.e.d.

In controlling the energy density in time, difficulties arise because of the dependence of \( \partial_t \eta \) and \( \nabla^2 \eta \) on \( f \), and since \( W(f) \) differs from \( \Delta H \) by the term \( Q(A^0)H \). For a ball \( B_\varrho = B_\varrho(x_0) \subset \mathbb{R}^n \) and \( f : \Sigma \to \mathbb{R}^n \) we adopt as in Section 2 the notation

\[
\Sigma_\varrho(x_0) = f^{-1}(B_\varrho(x_0))
\]

and consider a cutoff function \( \tilde{\gamma} \in C^1_0(B_\varrho) \), \( \tilde{\gamma} \geq 0 \), such that

\[
|D \tilde{\gamma}| \leq \frac{c}{\varrho}, \quad |D^2 \tilde{\gamma}| \leq \frac{c}{\varrho^2}.
\]

We put \( \gamma = \tilde{\gamma} \circ f \) and observe

\[
\nabla \gamma = (D \tilde{\gamma} \circ f) \cdot Df
\]

\[
\nabla^2 \gamma = (D^2 \tilde{\gamma} \circ f) (Df, Df) + (D \tilde{\gamma} \circ f) \cdot A(\cdot, \cdot).
\]
Lemma 3.2. If $f: \Sigma \times [0,T) \to \mathbb{R}^n$ is a Willmore flow and $\eta = \gamma^4$ for $\gamma = \tilde{\gamma} \circ f$ with $\tilde{\gamma}$ as in (37), then we have for $W = W(f)$

\begin{equation}
\partial_t \int |A^0|^2 \eta \, d\mu + \frac{1}{2} \int |W|^2 \eta \, d\mu \leq \frac{c}{\varrho^2} \int |A|^2 |A^0|^2 \gamma^2 \, d\mu + \frac{c}{\varrho^2} \int |A|^2 \, d\mu \tag{39}
\end{equation}

\begin{equation}
\partial_t \int |A|^2 \eta \, d\mu + \int |W|^2 \eta \, d\mu \leq \frac{c}{\varrho^2} \int |A^0|^4 \gamma^2 \, d\mu + \frac{c}{\varrho^4} \int |A|^2 \, d\mu. \tag{40}
\end{equation}

Proof. We estimate the terms in (36) and (35). We have

$$\int \gamma^2 |\nabla H|^2 \, d\mu = - \int \gamma^2 \langle H, \Delta H \rangle \, d\mu + \frac{c}{\varrho} \int \gamma |H| |\nabla H| \, d\mu$$

$$\leq - \int \gamma^2 \langle H, W \rangle \, d\mu + c \int \gamma^2 |A^0|^2 |H|^2 \, d\mu$$

$$+ \frac{1}{2} \int \gamma^2 |\nabla H|^2 \, d\mu + \frac{c}{\varrho^2} \int |H|^2 \, d\mu. \tag{\gamma > 0}$$

As

$$\int |\nabla \eta| |\nabla H| |W| \, d\mu \leq \varepsilon \int |W|^2 \eta \, d\mu + \frac{c(\varepsilon)}{\varrho^2} \int \gamma^2 |\nabla H|^2 \, d\mu,$$

we obtain by combining

$$- \int 2 \langle \nabla \text{grad} \eta, H \rangle \, d\mu \leq \varepsilon \int |W|^2 \eta \, d\mu + \frac{c(\varepsilon)}{\varrho^2} \int \gamma^2 |A^0|^2 |H|^2 \, d\mu$$

$$+ \frac{c(\varepsilon)}{\varrho^4} \int |H|^2 \, d\mu. \tag{\gamma > 0}$$

Next using (38)

$$- \int 2 \langle A^0(e_i, e_j), W \rangle \nabla^2_{e_i, e_j} \eta \, d\mu \leq c \int |A^0| |W| \left( \frac{1}{\varrho^2} \gamma^2 + \frac{1}{\varrho} \gamma^3 |A^0| \right) \, d\mu$$

$$\leq \varepsilon \int |W|^2 \eta \, d\mu + \frac{c(\varepsilon)}{\varrho^2} \int |A^0|^4 \gamma^2 \, d\mu$$

$$+ \frac{c(\varepsilon)}{\varrho^4} \int |A^0|^2 \, d\mu. \tag{\gamma > 0}$$
Finally

\[ \int |A^0|^2 \partial_t \eta \, d\mu \leq \frac{c}{\varrho} \int |A^0|^2 |W| \gamma^3 \, d\mu \]

\[ \leq \varepsilon \int |W|^2 \gamma^4 \, d\mu + \frac{c(\varepsilon)}{\varrho^2} \int |A^0|^4 \gamma^2 \, d\mu. \]

Combining the three estimates and absorbing for \( \varepsilon > 0 \) small, we obtain (39). The estimate (40) follows analogously from (35). q.e.d.

**Lemma 3.3.** Let \( f : \Sigma \times [0, T) \to \mathbb{R}^n \) be a Willmore flow. If

\[ \int_{\Sigma_{\varrho(x_0)}} |A^0|^2 \, d\mu < \varepsilon_0 \quad \text{at some time } t \in [0, T), \]

then for a constant \( c_0 > 0 \) we have at time \( t \)

\[ \partial_t \int_{\Sigma_{\varrho(x_0)}} |A^0|^2 \gamma^4 \, d\mu + c_0 \int (|A|^2 |A| |A| + |A|^4 |A^0|^2) \gamma^4 \, d\mu \]

\[ \leq \frac{c}{\varrho^4} \int_{\Sigma_{\varrho(x_0)}} |A|^2 \, d\mu, \]

and

\[ \partial_t \int_{\Sigma_{\varrho(x_0)}} |H|^2 \gamma^4 \, d\mu + c_0 \int (|A|^2 |A| |A| + |A|^4 |A^0|^2) \gamma^4 \, d\mu \]

\[ \leq \frac{c}{\varrho^2} \int_{\Sigma_{\varrho(x_0)}} |A|^2 \, d\mu. \]

**Proof.** (42) follows by combining (39) with Proposition 2.6, after estimating

\[ \frac{c}{\varrho^2} \int |A|^2 |A^0|^2 \gamma^2 \, d\mu \leq \varepsilon \int |A|^4 |A^0|^2 \gamma^4 \, d\mu + \frac{c(\varepsilon)}{\varrho^4} \int |A^0|^2 \, d\mu. \]

For the other bound we must go back to (35), estimating the three terms on the right hand side. We have

\[ \int \frac{1}{2} |H|^2 \partial_t \eta \, d\mu = - \int \frac{1}{2} |H|^2 D \tilde{\eta} \circ f \cdot (\Delta H + Q(A^0) H) \, d\mu. \]
By Young’s inequality with 4 and 4/3, we have
\[ \frac{1}{\varrho} \int |A^{\circ}|^2 |A|^3 \gamma^3 \, d\mu \leq \varepsilon \int |A^{\circ}|^{8/3} |A|^{10/3} \gamma^4 \, d\mu + \frac{c(\varepsilon)}{\varrho^4} \int |A|^2 \, d\mu. \] (44)

Using integration by parts, we infer
\[ \int H \ast H \langle D \bar{\eta} \circ f, \Delta H \rangle \, d\mu \]
\[ \leq \frac{c}{\varrho} \int |H| |\nabla H|^2 \gamma^3 \, d\mu + \frac{c}{\varrho^2} \int |H|^2 |\nabla H| \gamma^2 \, d\mu \]
\[ \leq \varepsilon \int |H|^2 |\nabla H|^2 \gamma^4 \, d\mu + \frac{c(\varepsilon)}{\varrho^2} \int |\nabla H|^2 \gamma^2 \, d\mu \]
\[ + \frac{c(\varepsilon)}{\varrho^4 \gamma > 0} \int |H|^2 \, d\mu. \]

In the proof of Lemma 3.2, we have already shown by partial integration that
\[ \int \gamma^2 |\nabla H|^2 \, d\mu \leq \delta \varrho^2 \int |W|^2 \gamma^4 \, d\mu + \frac{c(\delta)}{\varrho^2} \int |A|^2 \, d\mu \]
\[ + \delta \varrho^2 \int |A|^2 |A^{\circ}|^4 \gamma^4 \, d\mu, \]
so that by combining we obtain
\[ \int H \ast H \langle D \bar{\eta} \circ f, \Delta H \rangle \, d\mu \]
\[ \leq \varepsilon \int (|W|^2 + |A|^2 |\nabla A|^2 + |A|^2 |A^{\circ}|^4) \, d\mu \]
\[ + \delta \varrho^2 \int |A|^2 \, d\mu. \] (45)

Thus (44) and (45) estimate the first of the three terms on the right hand side of (35). For the second we use
\[ - \int \langle H, W \rangle \Delta \eta \, d\mu \]
\[ \leq - \int \langle H, \Delta H \rangle \langle D \bar{\eta} \circ f, H \rangle \, d\mu + \frac{c}{\varrho^2} \int |H| |\Delta H| \gamma^2 \, d\mu \]
\[ + \frac{c}{\varrho} \int |A^{\circ}|^2 |A|^3 \gamma^3 \, d\mu + \frac{c}{\varrho^2} \int |A^{\circ}|^2 |A|^2 \gamma^2 \, d\mu. \]
The first integral is estimated by (45), the third integral by (44). Furthermore

\[ \frac{c}{\varrho^2} \int |H| |\Delta H| \gamma^2 \, d\mu \leq \varepsilon \int |\nabla^2 A|^2 \gamma^4 \, d\mu + \frac{c(\varepsilon)}{\varrho^4} \int |A|^2 \, d\mu \]

\[ \frac{c}{\varrho^2} \int |A^\circ|^2 |A|^2 \gamma^2 \, d\mu \leq \varepsilon \int |A^\circ|^4 |A|^2 \gamma^4 \, d\mu + \frac{c(\varepsilon)}{\varrho^4} \int |A|^2 \, d\mu. \]

The third integral on the right of (35) satisfies

\[ \int |\nabla \eta| |\nabla H| |W| \, d\mu \leq \varepsilon \int |W|^2 \gamma^4 \, d\mu + \frac{c(\varepsilon)}{\varrho^4} \int |\nabla H|^2 \gamma^2 \, d\mu \]

and the right hand side is already estimated. Thus putting things together we have shown

\[ \partial_t \left( \int \frac{1}{2} |H|^2 \eta \, d\mu \right) + \frac{3}{4} \int |W|^2 \eta \, d\mu \]

\[ \leq \varepsilon \int (|\nabla^2 A|^2 + |A|^2 |\nabla A|^2 + |A|^4 |A^\circ|^2) \eta \, d\mu + \frac{c(\varepsilon)}{\varrho^4} \int |A|^2 \, d\mu. \]

Now (43) follows from Proposition 2.6. q.e.d.

**Proposition 3.4.** Let \( f : \Sigma \times [0,T) \to \mathbb{R}^n \) be a Willmore flow with \( \int_\Sigma |A|^2 \, d\mu \leq \varkappa \). There exist constants \( \varepsilon_1 = \varepsilon_1(n) > 0 \) and \( c_1 = c(n)/\varkappa > 0 \), such that if \( \varrho > 0 \) is chosen with

\[ \int_{\varrho \geq 0} |A^\circ|^2 \, d\mu \leq \varepsilon < \varepsilon_1 \quad \text{at time } t = 0 \quad \text{for all } \Sigma_{\varrho} = \Sigma_{\varrho}(x_0) \subset \mathbb{R}^n, \]

then for any time \( 0 \leq t < t_1 = \min\{c_1 \varrho^4, T\} \) we have

\[ \int_{\varrho < 0} |A^\circ|^2 \, d\mu + \int_0^t \int_{\varrho < 0} (|\nabla^2 A|^2 + |A|^2 |\nabla A|^2 + |A|^4 |A^\circ|^2) \, d\mu \, d\tau \]

\[ \leq c(\varepsilon + \varkappa \varrho^{-4} t), \]

\[ \int_0^t \|A^\circ\|_{L^\infty(\Sigma_{\varrho})}^4 \, d\tau \leq c(\varepsilon + \varkappa \varrho^{-4} t). \]
Moreover, for $0 < \sigma \leq \varrho$ and $\tau < \min\{c_1 \sigma^4, T\}$ we then also have

$$\int_{\Sigma_{\sigma/2}(x)} |A|^2 d\mu \bigg|_{t=\tau} \leq \int_{\Sigma_{\sigma}(x)} |A|^2 d\mu \bigg|_{t=0} + c \kappa \sigma^{-4} \tau \quad \forall x \in \mathbb{R}^n. \quad (49)$$

**Proof.** Let $N = N(n)$ be the number of balls $B_{\varrho/2} \subset \mathbb{R}^n$ needed to cover $B_{\varrho} \subset \mathbb{R}^n$ and choose $\varepsilon_1 \leq \frac{\varepsilon_0}{4N}$, where $\varepsilon_0 > 0$ is as in Lemma 3.3. Assume (41) is satisfied on $[0,t]$ for all $B_{\varrho} \subset \mathbb{R}^n$, and integrate (42) to obtain using (46)

$$\int_{\Sigma_{\varrho/2}} |A^0|^2 d\mu + c_0 \int_0^t \int_{\Sigma_{\varrho/2}} (|\nabla^2 A|^2 + |A|^2 |\nabla A|^2 + |A|^4 |A^0|^2) d\mu ds \leq \varepsilon + c \kappa \varrho^{-4} t. \quad (49)$$

Assuming $t \leq c_1 \varrho^4$ where $c_1$ is chosen with $0 < c_1 \leq \frac{\varepsilon_0}{4Nc_\kappa}$, we conclude

$$\int_{\Sigma_{\varrho}} |A^0|^2 d\mu + c_0 \int_0^t \int_{\Sigma_{\varrho}} (|\nabla^2 A|^2 + |A|^2 |\nabla A|^2 + |A|^4 |A^0|^2) d\mu ds \leq N(\varepsilon + c \kappa \varrho^{-4} t) \leq N(\varepsilon_1 + c \kappa c_1) \leq \frac{\varepsilon_0}{2}.$$

It follows that (41) holds up to time $t = t_1$ for all $x_0 \in \mathbb{R}^n$, and (47) follows. In particular $\int_{\Sigma_{\varrho}} |A^0|^2 d\mu \leq c (\varepsilon_1 + \kappa c_1)$, whence a covering argument with possibly smaller $\varepsilon_1, c_1$ implies the smallness hypothesis in Theorem 2.9 for any ball $B_{2\varrho} \subset \mathbb{R}^n$ and any $t \in [0,t_1]$. Inequality (48) now follows from combining (25) with (47), again using a covering. Finally (49) is obtained by integrating (43) and (42) on $[0,t]$. q.e.d.

We next state a version of the higher order estimates obtained in [5] which is localized in time.

**Theorem 3.5** (Interior Estimates). Let $f : \Sigma \times (0,T] \rightarrow \mathbb{R}^n$ be a Willmore flow satisfying the condition

$$\sup_{0 < t \leq T} \int_{\Sigma_{\varrho}(0)} |A|^2 d\mu \leq \varepsilon < \varepsilon_0(n), \quad (50)$$

We need a version of the higher order estimates obtained in [5] which is localized in time.
where \( T \leq c(n) \varrho^4 \). Then for any \( k \in \mathbb{N}_0 \) we have at time \( t \in (0, T] \) the estimates

\[
\| \nabla^k A \|_{L^\infty(\Sigma_{\varrho/2}(0))} \leq c(k) \sqrt{\varepsilon} t^{-\frac{k+1}{4}} \\
\| \nabla^k A \|_{L^2(\Sigma_{\varrho/2}(0))} \leq c(k) \sqrt{\varepsilon} t^{-\frac{k}{4}}.
\]

**Proof.** By scaling \( f_\varrho(p, t) = \frac{1}{\varrho} f(p, \varrho^4 t) \) we may assume \( \varrho = 1 \). Using (4.13) and (4.9) from [5], see also Proposition 4.6 in [5], we obtain on \( B = B_{3/4}(0) \) the inequalities

\[
\int_0^T \int_{\Sigma_{3/4}} (|\nabla^2 A|^2 + |A|^6) \, d\mu \, dt \leq c \varepsilon, \\
\int_0^T \| A \|_{L^\infty(\Sigma_{3/4})}^4 \, dt \leq c \varepsilon.
\]

Fix a cutoff function \( \tilde{\gamma} \in C^\infty_c(\mathbb{R}^n) \) with \( \chi_{B_{1/2}} \leq \tilde{\gamma} \leq \chi_B \) and \( \| D \tilde{\gamma} \|_{L^\infty} + \| D^2 \tilde{\gamma} \|_{L^\infty} \leq c \). Also define cutoff functions in time by

\[
\chi_j(t) = \begin{cases} 
0 & \text{for } t \leq (j-1) \frac{T}{m} \\
\frac{m}{T} \left( t - (j-1) \frac{T}{m} \right) & \text{in between} \\
1 & \text{for } t \geq j \frac{T}{m},
\end{cases}
\]

where \( 0 \leq j \leq m \) and \( m \in \mathbb{N}_0 \). Note \( \chi_0 \equiv 1 \) on \([0, T]\), \( \chi_m(T) = 1 \) and

\[
0 \leq \dot{\chi}_j \leq \frac{m}{T} \chi_{j-1}.
\]

Introducing the notation \( \alpha(t) = \| A \|_{L^\infty(\Sigma_{3/4})}^4 \), \( E_j(t) = \int |\nabla^2 A|^2 \gamma^{4j+4} d\mu \) (where \( \gamma = \tilde{\gamma} \circ f \)), we have by (4.14) in [5]

\[
\frac{d}{dt} E_j(t) + \frac{1}{2} E_{j+1}(t) \leq c \alpha(t) E_j(t) + c(1 + \alpha(t)) \varepsilon.
\]

Letting \( e_j(t) = \chi_j(t) E_j(t) \) this implies, using also (55),

\[
\frac{d}{dt} e_j(t) \leq c \alpha(t) e_j(t) - \frac{1}{2} \chi_j(t) E_{j+1}(t) + c(1 + \alpha(t)) \varepsilon + \frac{m}{T} \chi_{j-1}(t) E_j(t).
\]

We now prove by induction for \( 0 \leq j \leq m \) and all \( t \in (0, T] \) the inequality

\[
e_j(t) + \frac{1}{2} \int_0^t \chi_j(s) E_{j+1}(s) \, ds \leq c(m) \frac{\varepsilon}{T^j}.
\]
For $j = 0$ this follows from assumption (50) and estimate (53). Integrating (56) on $[0, T]$ yields, for $j \geq 1$,
\[ e_j(t) + \frac{1}{2} \int_0^t \chi_j(s) E_{j+1}(s) \, ds \]
\[ \leq c \int_0^t \alpha(s) e_j(s) \, ds + c \varepsilon \int_0^t (1 + \alpha(s)) \, ds \]
\[ + \frac{m}{T} \int_0^t \chi_{j-1}(s) E_j(s) \, ds. \]

Now since $\int_0^T \alpha(t) \, dt \leq c \varepsilon$ by (54), we may apply Gronwall’s inequality to get
\[ e_j(t) + \frac{1}{2} \int_0^t \chi_j(s) E_{j+1}(s) \, ds \leq c \varepsilon + \frac{cm}{T} \frac{c(m) \varepsilon}{T^j-1} \]
\[ \leq \frac{c(m) \varepsilon}{T^j}, \]
as $T \leq c(n)$ by assumption. Thus we have at time $t = T$
\[ \int |\nabla^{2m} A|^2 \gamma^{4m+4} \, d\mu \leq \frac{c(m) \varepsilon}{T^m}. \]
The estimate for $\nabla^{2m+1} A$ follows by interpolation as in Lemma 5.1 of [5], taking $r = 1$, $p = q = 2$, $\alpha = 1$, $\beta = 0$, $s = 4m + 6$ and $t = \frac{1}{\varepsilon} \in \left[ -\frac{1}{2}, \frac{1}{2} \right]$ there and using again $T \leq c(n)$. Renaming $T$ into $t$, the $L^2$-estimate (52) is proved. Using (4.9) and (4.7) in [5], the $L^\infty$-estimate (51) follows.

\[ \text{q.e.d.} \]

4. Construction of the blowup

In this section we rescale the Willmore flow at an assumed singularity at finite or infinite time, thereby constructing a static Willmore surface as a limit. We shall need the following local area bound due to L. Simon [8].

**Lemma 4.1.** Let $f : \Sigma \to \mathbb{R}^n$ be a properly immersed surface. Then for $0 < \sigma \leq \varrho < \infty$ and $\Sigma_{\varrho} = \Sigma_{\varrho}(x_0)$ one has
\[ \frac{\mu(\Sigma_{\varrho})}{\sigma^2} \leq c \left( \frac{\mu(\Sigma_{\varrho})}{\varrho^2} + \int_{\Sigma_{\varrho}} |H|^2 \, d\mu \right). \]
In particular if $\Sigma$ is compact without boundary
\[
\frac{\mu(\Sigma)}{\sigma^2} \leq c \left( W(f) + 4\pi \chi(\Sigma) \right).
\]

The following compactness theorem, whose proof is omitted, is a localized version of the result of J. Langer [6].

**Theorem 4.2.** Let $f_j : \Sigma_j \rightarrow \mathbb{R}^n$ be a sequence of proper immersions, where $\Sigma_j$ is a two-dimensional manifold without boundary. Let
\[
\Sigma_j(R) = \{ p \in \Sigma_j : |f_j(p)| < R \}
\]
and assume the bounds
\[
\begin{align*}
\mu_j(\Sigma_j(R)) &\leq c(R) \quad \text{for any } R > 0, \\
\|\nabla^k A_j\|_{L^\infty} &\leq c(k) \quad \text{for any } k \in \mathbb{N}_0.
\end{align*}
\]
Then there exists a proper immersion $\hat{f} : \hat{\Sigma} \rightarrow \mathbb{R}^n$, where $\hat{\Sigma}$ is again a two-manifold without boundary, such that after passing to a subsequence we have a representation
\[
(57) \quad f_j \circ \varphi_j = \hat{f} + u_j \quad \text{on } \hat{\Sigma}(j) = \{ p \in \hat{\Sigma} : |\hat{f}(p)| < j \}
\]
with the following properties:
\[
\begin{align*}
\varphi_j : \hat{\Sigma}(j) &\rightarrow U_j \subset \Sigma_j \quad \text{is diffeomorphic}, \\
\Sigma_j(R) &\subset U_j \quad \text{if } j \geq j(R), \\
u_j &\in C^\infty(\hat{\Sigma}(j), \mathbb{R}^n) \quad \text{is normal along } \hat{f}, \\
\|\nabla^k u_j\|_{L^\infty(\hat{\Sigma}(j))} &\rightarrow 0 \quad \text{as } j \rightarrow \infty, \quad \text{for any } k \in \mathbb{N}_0.
\end{align*}
\]

Roughly speaking, the theorem says that on any ball $B_R(0)$ the immersion $f_j$ can be written as a normal graph $\hat{f} + u_j$ with small norm for $j$ large over a limit immersion $\hat{f}$, after suitably reparametrizing with $\varphi_j$.

Now let $f : \Sigma \times [0,T) \rightarrow \mathbb{R}^n$ be a smooth Willmore flow defined on a closed surface $\Sigma$, where $0 < T \leq \infty$. Define
\[
\varkappa(r,t) = \sup_{x \in \mathbb{R}^n} \int_{\Sigma_t(x)} |A(t)|^2 \ d\mu_t.
\]
Choose an arbitrary sequence $r_j \searrow 0$ and assume concentration in the sense that for all $j$
\[
(58) \quad t_j = \inf\{ t \geq 0 : \varkappa(r_j,t) > \varepsilon_1 \} < T,
\]
where $\varepsilon_1 = \varepsilon_0/c$ and $\varepsilon_0, c$ are the constants from Theorem 1.2 in [5]. Clearly

\[ \int_{\Sigma_{r_j}(x)} |A(t_j)|^2 d\mu_j \leq \varepsilon_1 \quad \text{for any } x \in \mathbb{R}^n. \]

On the other hand, choosing an appropriate sequence of balls at times $\tau_\nu \searrow t_j$, we find a point $x_j \in \mathbb{R}^n$ satisfying

\[ \int_{f^{-1}(B_{r_j}(x_j))} |A(t_j)|^2 d\mu_j \geq \varepsilon_1. \]

Now we rescale by considering

\[ f_j : \Sigma \times [-r_j^{-4} t_j, r_j^{-4} (T - t_j)] \to \mathbb{R}^n, \]

\[ f_j(p, t) = \frac{1}{r_j} (f(p, t_j + r_j^4 t) - x_j). \]

By the above we have $\kappa_j(1, t) \leq \varepsilon_1$ for $t \leq 0$ and

\[ \int_{f^{-1}(B_1(0))} |A_j(0)|^2 d\mu_j \geq \varepsilon_1. \]

Furthermore Theorem 1.2 of [5] yields $r_j^{-4} (T - t_j) \geq c_0$ and in fact

\[ \kappa_j(1, t) \leq \varepsilon_0 \quad \text{for } 0 < t \leq c_0. \]

We may now apply Theorem 3.5 on parabolic cylinders $B_1(x) \times (t-1, t]$ to obtain

\[ \|\nabla^k A_j(t)\|_{L^\infty} \leq c(k) \quad \text{for } -r_j^{-4} t_j + 1 \leq t \leq c_0. \]

Furthermore Lemma 4.1 yields

\[ \frac{\mu_j(t)(\Sigma R(0))}{R^2} \leq c(W(f_0) + 4\pi \chi(\Sigma)) < \infty. \]

We apply Theorem 4.2 to the sequence $f_j = f_j(\cdot, 0) : \Sigma \to \mathbb{R}^n$, thus obtaining a limit immersion $\hat{f}_0 : \hat{\Sigma} \to \mathbb{R}^n$. Let $\varphi_j : \hat{\Sigma}(j) \to U_j \subset \Sigma$ be as in (57). Then the reparametrization

\[ f_j(\varphi_j, \cdot) : \hat{\Sigma}(j) \times [c_0] \to \mathbb{R}^n \]
is a Willmore flow with initial data
\[ f_j(\varphi_j, 0) = \tilde{f}_0 + u_j : \tilde{\Sigma}(j) \to \mathbb{R}^n. \]

The flows (62) satisfy the curvature bounds (61) and have initial data converging locally in $C^k$ to the immersion $\tilde{f}_0 : \Sigma \to \mathbb{R}^n$. By standard estimates for geometric evolution equations, see (4.24)--(4.28) in [5], we deduce the locally smooth convergence
\[ f_j(\varphi_j, \cdot) \to \hat{f}, \]
where $\hat{f} : \hat{\Sigma} \times [0, c_0] \to \mathbb{R}^n$ is a Willmore flow with initial data $f_0$. But on the other hand we have
\[
\int_0^{c_0} \int_{\Sigma(j)} |W(f_j(\varphi_j, t))|^2 d\mu_{f_j(\varphi_j, \cdot)} dt \\
= \int_0^{c_0} \int_{U_j} |W(f_j(\cdot, t))|^2 d\mu_{f_j(\cdot, \cdot)} dt \\
\leq \int_{\Sigma} |A_j(c_0)|^2 d\mu_j - \int_{\Sigma} |A_j(0)|^2 d\mu \\
= \int_{\Sigma} |A(t_j + r_j^4 c_0)|^2 d\mu - \int_{\Sigma} |A(t_j)|^2 d\mu,
\]
which converges to zero as $j \to \infty$. This implies that $W(\hat{f}) \equiv 0$ which means that $\hat{f}(\cdot, t) \equiv \hat{f}_0$ is an immersed Willmore surface, which is independent of time. Furthermore (60) implies, because of the smooth convergence in (64),
\[ \int_{\hat{f}^{-1}(B_1(0))} |\hat{A}|^2 d\hat{\mu} \geq \varepsilon_1 > 0. \]
Thus $\hat{f}$ is not a union of planes.

**Lemma 4.3.** Let $\hat{f} : \hat{\Sigma} \to \mathbb{R}^n$ be the blowup constructed above. If $\hat{\Sigma}$ contains a compact component $C$, then in fact $\hat{\Sigma} = C$ and $\Sigma$ is diffeomorphic to $C$.

**Proof.** For $j$ sufficiently large, $\varphi_j(C)$ is open and closed in $\Sigma$. Hence by connectedness of $\Sigma$ we have $\Sigma = \varphi_j(C)$ and thus $\hat{\Sigma} = C$. q.e.d.

**Theorem 4.4** (Nontriviality of the Blowup). Let $\hat{f} : \hat{\Sigma} \to \mathbb{R}^n$ be the blowup of a Willmore flow as constructed above. Then none of
the components of $\hat{f}$ parametrizes a round sphere. In particular, the blowup has a component which is a nonumbilic (compact or noncompact) Willmore surface.

Proof. Otherwise, Lemma 4.3 implies that the blowup surface $\hat{f}: \hat{\Sigma} \to \mathbb{R}^n$ is an embedded round sphere, i.e., has no further components. It follows that, up to the diffeomorphism $\varphi_j: \hat{\Sigma} \to \Sigma$, the map $f_j(\cdot, 0)$ is $C^k$-close to a round sphere and therefore

$$\int_\Sigma |A^0(t_j)|^2 \, d\mu = \int_\Sigma |A_j^0(0)|^2 \, d\mu_j \to 0,$$

$$\mu(t_j)(\Sigma) = r_j^2 \mu_j(0)(\Sigma) \to 0.$$ 

This contradicts the lower area bound which will be proved in Theorem 5.2. q.e.d.

5. Small initial energy

In this section we finally prove our main result:

**Theorem 5.1** (Global Existence and Convergence for Small Initial Energy). There exists an $\varepsilon_0 = \varepsilon_0(n) > 0$ such that if at time $t = 0$ there holds

$$\mathcal{W}(f_0) = \int_\Sigma |A^0|^2 \, d\mu < \varepsilon_0,$$

then the Willmore flow exists smoothly for all times and converges exponentially to a round sphere as $t \to \infty$.

We split the proof into several steps. The first step was already used in Theorem 4.4 and is of independent interest.

**Theorem 5.2** (Area Estimate). Let $f: \Sigma \times [0, T) \to \mathbb{R}^n$ be a Willmore flow with $\mathcal{W}(f) = \int_\Sigma |A^0|^2 \, d\mu \leq \varepsilon < \varepsilon_1$, where $\varepsilon_1 = \varepsilon_1(n)$ is as in Proposition 3.4. Then

$$\int_\Sigma (|\nabla A|^2 + |A|^2 |A^0|^2) \, d\mu \leq c \varepsilon \mu_0(\Sigma).$$

**Proof.** We have

$$\frac{d}{dt} \mu(\Sigma) = -\int_\Sigma |\nabla H|^2 \, d\mu + \int_\Sigma \langle Q(A^0)H, H \rangle \, d\mu.$$
Multiplying Simons’ identity (10) by $A^\circ$ and integrating yields (cf. Lemma 2.2):

\begin{equation}
2 \int_\Sigma |\nabla A^\circ|^2 \, d\mu + \int_\Sigma |H|^2 |A^\circ|^2 \, d\mu = \int_\Sigma |\nabla H|^2 \, d\mu + \int_\Sigma A^\circ * A^\circ * A^\circ * A^\circ \, d\mu.
\end{equation}

As $\langle Q(A^\circ)H, H \rangle \leq |A^\circ|^2 |H|^2$ by (14), we obtain

\[
\frac{d}{dt} \mu(\Sigma) + 2 \int_\Sigma |\nabla A^\circ|^2 \, d\mu \leq c \int_\Sigma |A^\circ|^4 \, d\mu \\
\leq c \|A^\circ\|_{L^\infty}^4 \mu(\Sigma).
\]

From (48) with $\varrho = \infty$ we have

\[
\int_0^t \|A^\circ\|_{L^\infty}^4 \, ds \leq c \varepsilon,
\]

and the Gronwall inequality yields

\[
\mu(\Sigma) + 2 \int_0^t \int_\Sigma |\nabla A^\circ|^2 \, d\mu \, ds \leq (1 + c \varepsilon) \mu_0(\Sigma).
\]

On the other hand, by Michael-Simon

\[
\int |\nabla H|^2 \, d\mu \leq c \left( \int (|\nabla^2 H| + |H| |\nabla H|) \, d\mu \right)^2 \\
\leq c \mu(\Sigma) \int_\Sigma (|\nabla^2 H|^2 + |H|^2 |\nabla H|^2) \, d\mu.
\]

As $\langle Q(A^\circ)H, H \rangle \geq 0$ we obtain

\[
\frac{d}{dt} \mu(\Sigma) \geq -c \mu(\Sigma) \int_\Sigma (|\nabla^2 H|^2 + |H|^2 |\nabla H|^2) \, d\mu.
\]

Using (47) with $\varrho = \infty$ implies the remaining inequality in (66). In particular we obtain

\[
\int_0^t \int_\Sigma |\nabla A^\circ|^2 \, d\mu \, ds \leq c \varepsilon \mu_0(\Sigma),
\]

and

\[
\int_0^t \int_\Sigma |A^\circ|^4 \, d\mu \, ds \leq (1 + c \varepsilon) \mu_0(\Sigma) \int_0^t \|A^\circ\|_{L^\infty}^4 \, ds \\
\leq c \varepsilon \mu_0(\Sigma).
\]
Finally from (68) and Codazzi (5)

$$\int_0^t \int_\Sigma |H|^2 |A^0|^2 d\mu \, ds \leq c \int_0^t \int_\Sigma (|\nabla A^0|^2 + |A^0|^4) d\mu \, ds$$

$$\leq c \varepsilon \mu_0(\Sigma).$$

This proves the theorem. q.e.d.

**Remark.** The extrinsic diameter is bounded above and below by the diameter of the initial surface, cf. [8].

**Lemma 5.3.** Under the assumptions of Theorem 5.1 there exists a radius $r_0 > 0$ such that

$$\int_{\Sigma_{r_0}(x)} |A(t)|^2 d\mu \leq \varepsilon_1 \quad \text{for all } x \in \mathbb{R}^n, t \in [0, \infty),$$

where $\varepsilon_1 > 0$ is as in (58).

**Proof.** Otherwise, the blowup construction of Section 4 yields an immersed Willmore surface $\hat{f} : \hat{\Sigma} \to \mathbb{R}^n$ with

$$\int_{\hat{f}^{-1}(B_1(0))} |\hat{A}|^2 d\hat{\mu} \geq \varepsilon_1,$$

whereas

$$\int_\Sigma |\hat{A}^0|^2 d\hat{\mu} < \varepsilon_0.$$

By Theorem 2.7 the surface $\hat{f}$ must be a union of embedded planes and round spheres, which however contradicts the nontriviality of the blowup, Theorem 4.4. q.e.d.

**Lemma 5.4.** For any sequence $t_j \to \infty$ there exist $x_j \in \mathbb{R}^n$ and $\varphi_j \in \text{Diff}(\Sigma)$ such that, after passing to a subsequence, the immersions $f(\varphi_j, t_j) - x_j$ converge smoothly to an embedded round sphere.

**Proof.** Let $x_j = f(p, t_j)$ where $p \in \Sigma$ is arbitrary. By the previous lemma and the interior curvature estimates from Theorem 3.5, we have for $t_j \geq 1$

$$\|\nabla^k A(\cdot, t_j)\|_{L^\infty} \leq c(k).$$
Furthermore, Lemma 4.1 yields the area bound
\[
\frac{\mu(t_j)(B_R(x_j))}{R^2} \leq c(W(f(t, t_j)) + 4\pi\chi(\Sigma)).
\]
According to Theorem 4.2, there exist a properly immersed surface \(\hat{f} : \hat{\Sigma} \rightarrow \mathbb{R}^n\) and diffeomorphisms \(\varphi_j : \hat{\Sigma}(j) \rightarrow U_j \subset \Sigma\), such that (after selection of a subsequence)
\[
f(\varphi_j, t_j) - x_j \longrightarrow \hat{f}
\]
locally in \(C^k\) on \(\hat{\Sigma}\). On \(\hat{\Sigma}(j)\) we consider the Willmore flows
\[
g_j(p, t) = f(\varphi_j(p), t_j + t) - x_j \quad (t \geq -t_j).
\]
These satisfy the curvature estimates (69), and the initial data (at \(t = 0\)) converge to \(\hat{f}\). Arguing as in (64), we obtain that \(g_j\) converges locally smoothly on \(\hat{\Sigma} \times [0, \infty)\) to a Willmore flow \(g : \hat{\Sigma} \times [0, \infty) \rightarrow \mathbb{R}^n\) with initial data \(\hat{f}\). But now
\[
\int_0^1 \int_{\hat{\Sigma}(j)} |W(g_j)|^2 d\mu_{g_j} dt \leq \int_{t_j}^{t_j + 1} \int_{\Sigma} |W(f)|^2 d\mu dt \rightarrow 0 \quad \text{as } j \rightarrow \infty.
\]
Therefore we have \(W(g) \equiv 0\) which proves that \(\hat{f}\) is a Willmore surface, and Theorem 2.7 implies that \(\hat{f}\) is a union of embedded planes and round spheres. Using the upper area bound in (66) and excluding several components as in Lemma 4.3 we conclude that \(\hat{f}\) must be a round sphere, and that the subconvergence is smooth. q.e.d.

As a consequence of the above, we obtain that
\[
W(f) = \int_{\Sigma} |A^\circ|^2 d\mu \rightarrow 0 \quad \text{as } t \rightarrow \infty.
\]
Moreover, Theorem 5.2 implies the existence of the nonzero limit
\[
\omega = \lim_{t \rightarrow \infty} \mu(t)(\Sigma) \in (0, \infty).
\]
Finally we now prove exponential decay of the curvature; from this one obtains smooth convergence of \(f\) to a round sphere \(\hat{f}\) and thus Theorem 5.1 in a standard way.
Lemma 5.5. As \( t \to \infty \), the following asymptotic statements hold, where \( \lambda > 0 \) is a constant:

\[
\| \nabla^k A(t) \|_{L^\infty} \leq c_k e^{-\lambda t} \quad \text{for } k \geq 1,
\]
\[
\| A^0 \|_{L^\infty} \leq c_0 e^{-\lambda t}.
\]

Proof. For \( \omega \) as in (71), the previous Lemma implies that the sectional curvature and the mean curvature of \( f(\cdot, t) \) satisfy

\[
\left\| K(\cdot, t) - \frac{4\pi}{\omega} \right\|_{L^\infty} \to 0,
\]
\[
\left\| |H|^2 - \frac{16\pi}{\omega} \right\|_{L^\infty} \to 0 \quad \text{as } t \to \infty.
\]

In particular, we may assume after a fixed time translation that

\[ |H|^2 \geq c > 0 \quad \text{for all } t \geq 0. \]

By Lemma 3.3, we then have for all \( t \)

\[
\partial_t \int_{\Sigma} |A^0|^2 d\mu + c \int_{\Sigma} (|\nabla^2 A|^2 + |\nabla A|^2 + |A^0|^2) d\mu \leq 0,
\]

which implies

\[
\int_{\Sigma} |A^0|^2 d\mu + \int_0^\infty \int_{\Sigma} (|\nabla^2 A|^2 + |\nabla A|^2) d\mu \, d\sigma \leq c e^{-2\lambda t},
\]

for a constant \( \lambda = \lambda(n) > 0 \). From here we easily derive exponential decay of all derivatives of \( A \). Namely, letting \( \sigma \to \infty \) in Corollary 3.4 of [5], we have for \( \phi = \nabla^m A \) (\( m \geq 1 \))

\[
\frac{d}{dt} \int |\phi|^2 d\mu + \frac{3}{4} \int |\nabla^2 \phi|^2 d\mu \leq \int (P_{m+2}^2(A) + P_{m+5}^m(A)) * \phi d\mu.
\]

Using that \( A \) and all its derivatives remain bounded as \( t \to \infty \), we can estimate

\[
\int P_2^0(A) * \nabla^{m+2} A * \phi \, d\mu \leq \epsilon \int |\nabla^2 \phi|^2 d\mu + c(\epsilon) \int |\nabla^m A|^2 d\mu,
\]
\[
\int \left( \hat{P}_{3}^{m+2}(A) + P_{5}^{m}(A) \right) * \phi d\mu \leq c \sum_{j=1}^{m+1} \int |\nabla^j A|^2 d\mu.
\]
Here \( \hat{P}^{m+2}_3(A) \) denotes all terms of type \( P^{m+2}_3(A) \) that do not contain the \((m+2)\)-th derivative, and of course \( c \) is not a universal constant here. We obtain

\[
\frac{d}{dt} \int |\phi|^2 d\mu + \frac{1}{2} \int |\nabla^2 \phi|^2 d\mu \leq c \sum_{j=1}^{m+1} \int |\nabla^j A|^2 d\mu,
\]

and now by induction using (74)

\[
\|\nabla^m A\|_{L^2}^2 + \int_t^\infty \|\nabla^{m+2} A\|_{L^2}^2 ds \leq c e^{-2\lambda t}.
\]

By a Sobolev inequality, e.g., the Michael-Simon inequality, we deduce

\[
\begin{align*}
\|A^0\|_{L^\infty} & \leq c_0 e^{-\lambda t}, \\
\|\nabla^k A\|_{L^\infty} & \leq c_k e^{-\lambda t},
\end{align*}
\]

which ends the proof of Lemma 5.5 and of Theorem 5.1. q.e.d.

References


Albert-Ludwigs-Universität Freiburg
Universität Bonn