JEMS

Ernst Kuwert · Tobias Lamm · Yuxiang Li

Two-dimensional curvature functionals with superquadratic growth

Abstract. For two-dimensional, immersed closed surfaces $f : \Sigma \to \mathbb{R}^n$, we study the curvature functionals $\mathcal{E}^p(f)$ and $\mathcal{W}^p(f)$ with integrands $(1+|A|^2)^{p/2}$ and $(1+|H|^2)^{p/2}$, respectively. Here A is the second fundamental form, H is the mean curvature and we assume $p > 2$. Our main result asserts that $W^{2,p}$ critical points are smooth in both cases. We also prove a compactness theorem for $\mathcal{W}^p$-bounded sequences. In the case of $\mathcal{E}^p$ this is just Langer's theorem [16], while for $\mathcal{W}^p$ we have to impose a bound for the Willmore energy strictly below 8π as an additional condition. Finally, we establish versions of the Palais–Smale condition for both functionals.

Keywords. Curvature functionals, Palais–Smale condition

1. Introduction

Let Σ be a two-dimensional, closed differentiable manifold and $p > 2$, hence $W^{2,p}(\Sigma, \mathbb{R}^n) \subset C^{1,1-2/p}(\Sigma, \mathbb{R}^n)$ by the Sobolev embedding theorem. On the open subset $W^{2,p}_{\mathrm{im}}(\Sigma, \mathbb{R}^n)$ of immersions we consider the two functionals

$$\mathcal{E}^p(f) = \frac{1}{4} \int_\Sigma (1 + |A|^2)^{p/2} \, d\mu_g, \quad \mathcal{W}^p(f) = \frac{1}{4} \int_\Sigma (1 + |H|^2)^{p/2} \, d\mu_g.$$

Here g denotes the first fundamental form with induced measure μ_g, $A = (D^2 f)^\perp$ the second fundamental form, and H is the mean curvature vector. We prove regularity of critical points for both functionals.

Theorem 1.1. *Let $f \in W^{2,p}_{\mathrm{im}}(\Sigma, \mathbb{R}^n)$ be a critical point of $\mathcal{W}^p$ or $\mathcal{E}^p$, where $2 < p < \infty$. Then local graph representations of f are smooth.*

E. Kuwert: Mathematisches Institut, Universität Freiburg, Eckerstraße 1, 79104 Freiburg, Germany;
e-mail: ernst.kuwert@math.uni-freiburg.de

T. Lamm: Institute for Analysis, Karlsruhe Institute of Technology (KIT),
Englerstr. 2, 76131 Karlsruhe, Germany; e-mail: tobias.lamm@kit.edu

Y. Li: Department of Mathematical Sciences, Tsinghua University, Beijing 100084, P.R. China;
e-mail: yxli@math.tsinghua.edu.cn

In a graph representation, the Euler–Lagrange equations become fourth order elliptic systems, where the principal term has a double divergence structure. The systems are degenerate, in the sense that in both cases the coefficient of the principal term involves a $(p-2)$-th power of the curvature, which a priori may not be bounded. For the functional $\mathcal{W}^p(f)$, our first step towards regularity is an improvement of the integrability of H. For this we employ an iteration based on a new test function argument. More precisely, we solve the equation $L_g \varphi = |H|^{\lambda-1} H$ for appropriate $\lambda > 1$ and then insert φ as a test function. Here the operator $L_g = \sqrt{\det g}\, g^{\alpha\beta} \partial^2_{\alpha\beta}$ comes up in the principal term of the equation.

Unfortunately, the same strategy does not apply in the case of the functional $\mathcal{E}^p(f)$, since then the corresponding operator is a full Hessian and hence the equation would be overdetermined. Instead we first use a hole-filling argument to show power decay for the L^p integral of the second derivatives, and derive L^2 bounds for the third derivatives by a difference-quotient argument; these steps follow closely the ideas of Morrey [19] and L. Simon [22]. In the final critical step we adapt a Gehring type lemma due to Bildhauer, Fuchs & Zhong [7] as well as the Moser–Trudinger inequality to conclude that the solution is of class C^2. Since it is also not immediate how to modify the $\mathcal{E}^p(f)$ approach to cover the functional $\mathcal{W}^p(f)$, we decided to include both independent arguments.

As a second issue we address the existence of minimizers for the functionals. By the compactness theorem of Langer [16], sequences of closed immersed surfaces $f_k : \Sigma \to \mathbb{R}^n$ with $\mathcal{E}^p(f_k) \leq C$ subconverge weakly to an $f \in W^{2,p}_{\mathrm{im}}(\Sigma, \mathbb{R}^n)$, after suitable translation and reparametrization. In particular, we obtain the existence of a smooth $\mathcal{E}^p$ minimizer in the class of immersions $f : \Sigma \to \mathbb{R}^n$ for $p > 2$. On the other hand, boundedness of $\mathcal{W}^p(f)$ is not sufficient to guarantee the required compactness. This is easily illustrated by joining two round spheres by a shrinking catenoid neck, showing that the 8π bound in the following result is optimal.

Theorem 1.2. *Let Σ be a closed surface and $f_k \in W^{2,p}_{\mathrm{im}}(\Sigma, \mathbb{R}^n)$ be a sequence of immersions with $0 \in f_k(\Sigma)$ and*

$$\mathcal{W}^p(f_k) \leq C \quad and \quad \liminf_{k\to\infty} \frac{1}{4} \int_\Sigma |H_k|^2 \, d\mu_{g_k} < 8\pi.$$

After passing to $f_k \circ \varphi_k$ for appropriate $\varphi_k \in C^\infty(\Sigma, \Sigma)$ and selecting a subsequence, the f_k converge weakly in $W^{2,p}(\Sigma, \mathbb{R}^k)$ to an $f \in W^{2,p}_{\mathrm{im}}(\Sigma, \mathbb{R}^n)$. In particular, the convergence is in $C^{1,\beta}(\Sigma, \mathbb{R}^n)$ for any $\beta < 1 - 2/p$ and we have

$$\mathcal{W}^p(f) \leq \liminf_{k\to\infty} \mathcal{W}^p(f_k).$$

We remark that Mondino [18] recently proved the existence and partial regularity of varifolds minimizing functionals which satisfy similar growth conditions as $\mathcal{E}^p$ and $\mathcal{W}^p$ in general dimensions and codimensions.

A classical approach to the construction of harmonic maps, due to Sacks & Uhlenbeck [21], is by introducing perturbed functionals involving a power $p > 2$ of the gradient. One

motivation for our analysis is an analogous approximation for the Willmore functional

$$\mathcal{W}(f) = \frac{1}{4} \int_{\Sigma} |H|^2 \, d\mu_g = \frac{1}{4} \int_{\Sigma} |A|^2 \, d\mu_g + \pi \chi(\Sigma). \tag{1.1}$$

The Willmore functional does not satisfy a Palais–Smale type condition, since it is invariant under the group of Möbius transformations. In Section 5 we verify suitable versions of the Palais–Smale condition for the functionals $\mathcal{E}^p$ and $\mathcal{W}^p$ with $p > 2$.

Curvature functionals with nonquadratic growth appear also in the work of Bellettini, Dal Maso & Paolini [3] as well as Ambrosio & Masnou [1]. However, their focus is much different, for instance the latter paper is motivated by applications to image restoration.

2. The Euler–Lagrange equations

Here we compute in local coordinates the Euler–Lagrange equations of the functionals $\mathcal{E}^p(f)$ and $\mathcal{W}^p(f)$. For an immersed surface the fundamental forms are

$$g_{\alpha\beta} = \langle \partial_\alpha f, \partial_\beta f \rangle \quad \text{and} \quad A_{\alpha\beta} = P^\perp(\partial^2_{\alpha\beta} f).$$

Here $P^\perp$ is the projection onto the normal space given by

$$P^\perp = \mathrm{Id} - g^{\alpha\beta} \langle \partial_\alpha f, \cdot \rangle \partial_\beta f.$$

We compute further

$$|A|^2 = g^{\alpha\gamma} g^{\beta\lambda} \langle P^\perp \partial^2_{\alpha\beta} f, \partial^2_{\gamma\lambda} f \rangle \quad \text{and} \quad |H|^2 = g^{\alpha\beta} g^{\gamma\lambda} \langle P^\perp \partial^2_{\alpha\beta} f, \partial^2_{\gamma\lambda} f \rangle.$$

On the open set of $W^{2,p}$ immersions, both $\mathcal{E}^p$ and $\mathcal{W}^p$ are differentiable in the sense of Fréchet. The derivative of $\mathcal{E}^p$ is given by

$$\begin{aligned}
D\mathcal{E}^p(f)\phi &= \frac{p}{4} \int_{\Sigma} (1 + |A|^2)^{(p-2)/2} g^{\alpha\gamma} g^{\beta\lambda} \langle P^\perp \partial^2_{\alpha\beta} f, \partial^2_{\gamma\lambda} \phi \rangle \sqrt{\det g} \\
&\quad + \frac{p}{8} \int_{\Omega} (1 + |A|^2)^{(p-2)/2} \Big\langle \frac{\partial (g^{\alpha\gamma} g^{\beta\lambda} P^\perp)}{\partial p^k_\mu} \partial_\mu \phi^k \partial^2_{\alpha\beta} f, \partial^2_{\gamma\lambda} f \Big\rangle \sqrt{\det g} \\
&\quad + \frac{1}{4} \int_{\Omega} (1 + |A|^2)^{p/2} \frac{\partial \sqrt{\det g}}{\partial p^k_\mu} \partial_\mu \phi^k.
\end{aligned}$$

In particular, if $f(x) = (x, u(x))$ where $u \in W^{2,p}(\Omega, \mathbb{R}^{n-2})$, then f is a critical point of $\mathcal{E}^p$ if and only if u is a weak solution of the system

$$\partial^2_{\alpha\beta}(a^{\alpha\beta}_i(Du, D^2u)) + \partial_\alpha(b^\alpha_i(Du, D^2u)) = 0 \quad (1 \le i \le n-2), \tag{2.1}$$

where the coefficients are given by

$$a_i^{\alpha\beta}(Du, D^2u) = (1 + |A|^2)^{(p-2)/2}\sqrt{\det g}\, g^{\alpha\gamma}g^{\beta\lambda}(\delta_{ij} - g^{\mu\nu}\partial_\mu u^i\partial_\nu u^j)\partial_{\gamma\lambda}^2 u^j,$$

$$b_i^\alpha(Du, D^2u) = -\frac{1}{2}(1 + |A|^2)^{(p-2)/2}\frac{\partial(g^{\gamma\mu}g^{\lambda\nu}(\delta_{jk} - g^{\sigma\tau}p_\sigma^j p_\tau^k))}{\partial p_\alpha^i}\partial_{\gamma\lambda}^2 u^j\partial_{\mu\nu}^2 u^k\sqrt{\det g}$$

$$-\frac{1}{p}(1 + |A|^2)^{p/2}\frac{\partial\sqrt{\det g}}{\partial p_\alpha^i}.$$

For $|p| \le \Lambda$ and $V = (1 + |q|^2)^{1/2}$, where p, q are the variables corresponding to Du, D^2u, one easily checks the bounds

$$|D_q a| \le C(\Lambda)V^{p-2},$$
$$|a| + |D_p a| + |D_q b| \le C(\Lambda)V^{p-1},$$
$$|b| + |D_p b| \le C(\Lambda)V^p.$$

Moreover, the system satisfies the ellipticity condition

$$\frac{\partial a_i^{\alpha\beta}}{\partial q_{\gamma\lambda}^j}\xi_{\alpha\beta}^i\xi_{\gamma\lambda}^j \ge \lambda V^{p-2}|\xi|^2 \quad \text{where } \lambda = \lambda(\Lambda) > 0,$$

For the first variation of $\mathcal{W}^p(f)$ one obtains

$$DW^p(f)\phi = \frac{p}{4}\int_\Sigma (1 + |H|^2)^{(p-2)/2}\langle H, g^{\gamma\lambda}\partial_{\gamma\lambda}^2\phi\rangle\sqrt{\det g}$$

$$+ \frac{p}{8}\int_\Omega (1 + |H|^2)^{(p-2)/2}\left\langle\frac{\partial(g^{\alpha\beta}g^{\gamma\lambda}P^\perp)}{\partial p_\mu^k}\partial_\mu\phi^k\partial_{\alpha\beta}^2 f, \partial_{\gamma\lambda}^2 f\right\rangle\sqrt{\det g}$$

$$+ \frac{1}{4}\int_\Omega (1 + |H|^2)^{p/2}\frac{\partial\sqrt{\det g}}{\partial p_\mu^k}\partial_\mu\phi^k.$$

If we set $L_g\phi = \sqrt{\det g}\, g^{\gamma\lambda}\partial_{\gamma\lambda}^2\phi$, the first variation takes the form

$$DW^p(f)\phi = \frac{p}{4}\int_\Omega (1 + |H|^2)^{(p-2)/2}\langle H, L_g\phi\rangle + \int_\Omega B_i^\alpha(Df, D^2 f)\partial_\alpha\phi^i, \qquad (2.2)$$

where

$$B_i^\alpha(Df, D^2 f) = \frac{p}{8}(1 + |H|^2)^{(p-2)/2}\left\langle\frac{\partial(g^{\gamma\lambda}g^{\mu\nu}P^\perp)}{\partial p_\alpha^i}\partial_{\gamma\lambda}^2 f, \partial_{\mu\nu}^2 f\right\rangle\sqrt{\det g}$$

$$+ \frac{1}{4}(1 + |H|^2)^{p/2}\frac{\partial\sqrt{\det g}}{\partial p_\alpha^i}.$$

When passing to graphs we have, under the assumption $|p| \le \Lambda$,

$$|B| + |D_p B| \le C(\Lambda)V^{p-2}|q|^2 \quad \text{and} \quad |D_q B| \le C(\Lambda)V^{p-2}|q|. \qquad (2.3)$$

3. Regularity of critical points

3.1. The functional $\mathcal{W}^p$

For $\Omega \subset \mathbb{R}^2$ and $p > 2$, let $f : \Omega \to \mathbb{R}^n$ be the graph of a function $u \in W^{2,p}(\Omega, \mathbb{R}^{n-2})$. Recall from (2.2) that f is a critical point of $\mathcal{W}^p$ if and only if

$$\int_\Omega \langle \mathcal{H}, L_g \varphi \rangle + \int_\Omega B_i^\alpha (Du, D^2 u) \partial_\alpha \varphi^i = 0 \quad \text{for all } \varphi \in W_0^{2,p}(\Omega, \mathbb{R}^n). \qquad (3.1)$$

Here $\mathcal{H} = (1 + |H|^2)^{p/2-1} H$ and the functions B_α^i satisfy the bounds (2.3). We have the following result.

Theorem 3.1. *Weak solutions $u \in W^{2,p}(\Omega, \mathbb{R}^{n-2})$ of (3.1) are smooth.*

3.1.1. $W^{2,q}$-regularity.

We start by stating a regularity property for the mean curvature system. For a graph of a function $u \in W^{2,p}(\Omega, \mathbb{R}^{n-2})$, the weak mean curvature satisfies for $j = 1, \ldots, n-2$ the formula

$$g^{\alpha\beta}(\delta_{ij} - g^{\lambda\mu} \partial_\lambda u^i \partial_\mu u^j) \partial_{\alpha\beta}^2 u^i = H^{j+2}. \qquad (3.2)$$

Since $p > 2$ the left hand side may be viewed as a linear operator of the form $a_{ij}^{\alpha\beta} \partial_{\alpha\beta}^2 u^i$, where the coefficients are Hölder continuous with exponent $1 - 2/p > 0$ and the ellipticity constant is controlled by the $W^{2,p}$ norm of the function u. In particular, if we know $H \in L^q(\Omega, \mathbb{R}^n)$ for some $q \in (p, \infty)$, then standard L^q theory yields $u \in W_{\text{loc}}^{2,q}(\Omega, \mathbb{R}^{n-2})$ together with a local estimate

$$\|u\|_{W^{2,q}(\Omega')} \le C(p, q, \Lambda) (\|H\|_{L^q(\Omega)} + 1) \quad \text{if } \|u\|_{W^{2,p}(\Omega)} \le \Lambda. \qquad (3.3)$$

The dependence on the domains $\Omega' \subset\subset \Omega$ is not mentioned explicitly here.

Lemma 3.2. *Let $u \in W^{2,p}(\Omega, \mathbb{R}^{n-2})$ be a weak solution of (3.1). Then for any $\varphi \in W^{2,p}(\Omega)$ and any test function $\eta \in C_0^2(\Omega)$ we have*

$$\int_\Omega \eta \langle \mathcal{H}, L_g \varphi \rangle \le C \int_{\text{spt}\,\eta} (1 + |D^2 u|^2)^{p/2} (|\varphi| + |D\varphi|), \quad \text{where } C = C(\|\eta\|_{C^2}).$$

Proof. Expanding

$$L_g(\eta\varphi) = \eta L_g \varphi + \varphi L_g \eta + 2\sqrt{\det g}\, g^{\alpha\beta} \partial_\alpha \eta \partial_\beta \varphi,$$

we see by combining with (3.1) and (2.3) that

$$\int_\Omega \eta \langle \mathcal{H}, L_g \varphi \rangle \le C \int_\Omega (1 + |D^2 u|^2)^{p/2} (|D\eta|\, |\varphi| + |\eta|\, |D\varphi|)$$
$$+ C \int_\Omega (1 + |D^2 u|^2)^{(p-1)/2} (|D^2 \eta|\, |\varphi| + |D\eta|\, |D\varphi|).$$

This implies the lemma. $\qquad\square$

We are now ready to improve the integrability of $D^2 u$.

Theorem 3.3. *Let $u \in W^{2,p}(\Omega, \mathbb{R}^{n-2})$ be a weak solution of (3.1) where $p > 2$. Then $u \in W^{2,q}_{\mathrm{loc}}(\Omega, \mathbb{R}^{n-2})$ for any $q \in [p, \infty)$.*

Proof. Assume we already know $\|u\|_{W^{2,q}(B_r)} \leq \Lambda$ where $q \geq p$. For $|H|_A = \min(|H|, A)$ with $A > 0$ and a parameter $\lambda \in (1, q)$, we use L^q theory to obtain a solution $\varphi \in W^{2,q} \cap W^{1,q}_0(B_r, \mathbb{R}^n)$ of the linear equation

$$L_g \varphi = |H|_A^{\lambda-1} H.$$

As $1 < q/\lambda < \infty$ the function φ satisfies

$$\|\varphi\|_{W^{2,q/\lambda}(B_r)} \leq C \left\| |H|_A^{\lambda-1} H \right\|_{L^{q/\lambda}(B_r)} \leq C;$$

here and in the rest of the proof, the constant C is independent of A. By the Sobolev embedding theorem, we have for $\lambda < q/2$ the estimate

$$\|\varphi\|_{C^1(B_r)} \leq C,$$

while for $q/2 < \lambda < q$ we get instead

$$\|\varphi\|_{W^{1,s}(B_r)} \leq C \quad \text{for } s = \frac{2q}{2\lambda - q} \in [1, \infty).$$

Now Lemma 3.2 implies that

$$\int_{B_{r/2}} |H|^p |H|_A^{\lambda-1} \leq C \int_{B_r} (1 + |D^2 u|^2)^{p/2} (|\varphi| + |D\varphi|) \leq C,$$

under the condition that either $1 < \lambda < q/2$, or $q/2 < \lambda < q$ with

$$\frac{p}{q} + \frac{1}{s} \leq 1 \iff \lambda \leq 3q/2 - p.$$

Letting $A \nearrow \infty$ we get $H \in L^{p+\lambda-1}(B_{r/2})$, and so $u \in W^{2,p+\lambda-1}(B_{r/4}, \mathbb{R}^{n-2})$ from (3.3). We can now set up an iteration to get $u \in W^{2,q}_{\mathrm{loc}}(\Omega)$ for all $q < \infty$. As the initial step we choose $q = p$ and $1 < \lambda < p/2$, which brings us to $q < 3p/2 - 1$. For $p < q < 2p$ we can take $\lambda = 3q/2 - p$, improving the exponent to $3q/2 - 1$. After finitely many iterations, we arrive at some $q > 2p$. Now we continue with $q/2 < \lambda = q - p + 2 < 3q/2 - p$ and obtain the desired higher integrability. $\square$

3.1.2. $W^{1,2}_{\mathrm{loc}}$-regularity of $\mathcal{H}$. In this subsection we use difference quotient methods in order to show that $\mathcal{H} \in W^{1,2}_{\mathrm{loc}}$. For $h > 0$, $f : \Omega \to \mathbb{R}^k$ and fixed $\nu \in \{1, 2\}$ we define

$$f_h(x) = \frac{1}{h}(f(x + he_\nu) - f(x)).$$

In the following we denote by $\Phi(\cdot, h)$ an arbitrary measurable function which satisfies, for all $1 \le q < \infty$ and all $\Omega' \subset\subset \Omega$,

$$\int_{\Omega'} |\Phi(x, h)|^q \le C(q).$$

Lemma 3.4. *Let u be as in Theorem* 3.1 *and let* $\Omega' \subset\subset \Omega$. *Then, for all* $h > 0$ *small enough and all* $x \in \Omega'$,

$$|(|H|^2)_h|(x) + |H_h|(x) \le \Phi(x, h)|\mathcal{H}_h|(x).$$

Proof. We have $|\mathcal{H}|^2 = (1 + |H|^2)^{p-2}|H|^2$ and from the mean value theorem we get

$$
\begin{aligned}
(|\mathcal{H}|^2)_h(x) &= ((1 + |H|^2)^{p-2})_h(x)|H|^2(x + he_v) + (1 + |H|^2)^{p-2}(x)(|H|^2)_h(x) \\
&= (|H|^2)_h(x)\big[(p - 2)(1 + \xi(x))^{p-3}|H|^2(x + he_v) + (1 + |H|^2)^{p-2}(x)\big],
\end{aligned}
\tag{3.4}
$$

where $0 \le \xi(x) \le |H|^2(x) + |H|^2(x + he_v)$. Hence

$$|(|H|^2)_h|(x) \le |(|\mathcal{H}|^2)_h|(x).$$

On the other hand we calculate

$$(|\mathcal{H}|^2)_h(x) = (\mathcal{H})_h(x)\mathcal{H}(x + he_v) + (\mathcal{H})_h(x)\mathcal{H}(x).$$

Combining this with Theorem 3.3 proves the first estimate.

Another application of the mean value theorem yields

$$\mathcal{H}_h(x) = H_h(x)(1 + |H|^2)^{p/2-1}(x + he_v) + (p/2 - 1)(1 + \xi(x))^{p/2-2}(|H|^2)_h(x)H(x),$$

where $0 \le \xi(x) \le |H|^2 + |H|_h^2$, and hence the second estimate follows from the first one. $\qquad\square$

Corollary 3.5. *Let u be as in Theorem* 3.1 *and let* $B_r \subset \Omega$. *Then, for every* $1 < s < \infty$ *and all* $h > 0$ *small enough,*

$$\int_{B_r} \eta^{2s}|(D^2 u)_h|^s \le C \int_{B_r} \Phi(x, h)(\eta^{2s}|(\mathcal{H})_h|^s + 1), \tag{3.5}$$

where $\eta \in C_c^\infty(B_r)$ *is a smooth cut-off function.*

Proof. Using (3.2) we see that u_h solves

$$a_{ij}^{\alpha\beta}(\cdot + he_v)\partial_{\alpha\beta}^2 u_h^i = H_h^{j+2} - (a_{ij}^{\alpha\beta})_h \partial_{\alpha\beta}^2 u^i =: \tilde{H}^j(x),$$

where $a_{ij}^{\alpha\beta} = g^{\alpha\beta}(\delta_{ij} - g^{\lambda\mu}\partial_\lambda u^i \partial_\mu u^j)$.

Using Theorem 3.3, Lemma 3.4, and a standard estimate shows that for $h > 0$ small we have

$$|\tilde{H}|(x) \le \Phi(x, h)(|\mathcal{H}_h| + 1).$$

Next we use standard L^p theory in order to get, for every $1 < s < \infty$,

$$\int_{B_r} |D^2(\eta^2 u_h)|^s \le C \int_{B_r} \Phi(x, h)(|\mathcal{H}_h|^s + 1).$$

Since moreover

$$\int_{B_r} \eta^{2s} |(D^2 u)_h|^s \le C \int_{B_r} |D^2(\eta^2 u_h)|^s + C,$$

this finishes the proof of the corollary. $\qquad\square$

Now we are in a position to prove that $\mathcal{H} \in W_{\text{loc}}^{1,2}(\Omega)$.

Proposition 3.6. *Let* $u \in W^{2,p}(\Omega, \mathbb{R}^{n-2})$ *be as in Theorem* 3.1. *Then* $\mathcal{H} \in W_{\text{loc}}^{1,2}(\Omega, \mathbb{R}^n)$. *Moreover,* $H \in W_{\text{loc}}^{1,s}(\Omega, \mathbb{R}^n)$ *and* $u \in W_{\text{loc}}^{3,s}(\Omega, \mathbb{R}^n)$ *for every* $1 < s < 2$.

Proof. Taking difference quotients of equation (3.1) we get

$$\partial_{\alpha\beta}^2(\sqrt{g}\, g^{\alpha\beta} \mathcal{H})_h - \partial_\alpha(B^\alpha(Du, D^2 u))_h = 0. \tag{3.6}$$

We abbreviate $U(x) = (Du(x), D^2 u(x))$ and we use the fundamental theorem of calculus to write

$$\begin{aligned}
(f \circ U)_h (x) &= \frac{1}{h}\big(f(U(x + he_\nu)) - f(U(x))\big) \\
&= \frac{1}{h} \int_0^1 \frac{d}{dt} f\big((1 - t)U(x) + t\, U(x + he_\nu)\big)\, dt \\
&= \int_0^1 Df\big((1 - t)U(x) + t\, U(x + he_\nu)\big)\, dt \cdot U_h(x).
\end{aligned}$$

Using the notation $f^h(x) = \int_0^1 f\big((1 - t)U(x) + tU(x + he_\nu)\big)\, dt$ we thus get

$$(f \circ U)_h = \left(\frac{\partial f}{\partial q_{\lambda\mu}^j}\right)^h \partial_{\lambda\mu}^2 u_h^j + \left(\frac{\partial f}{\partial p_\lambda^j}\right)^h \partial_\lambda u_h^j,$$

and the system (3.6) takes the form

$$\partial_{\alpha\beta}^2\big((\sqrt{g}\, g^{\alpha\beta})(x + he_\nu)\mathcal{H}_h\big) + \partial_{\alpha\beta}^2\big((\sqrt{g}\, g^{\alpha\beta})_h \mathcal{H}\big) - \partial_\alpha(\tilde{B}^\alpha(Du_h, D^2 u_h)) = 0. \tag{3.7}$$

Here the coefficients are as follows:

$$\tilde{B}^\alpha(x, z, p, q) = \left(\frac{\partial B^\alpha}{\partial q_{\lambda\mu}^j}\right)^h(x) q_{\lambda\mu}^j + \left(\frac{\partial B^\alpha}{\partial p_\lambda^j}\right)^h(x) p_\lambda^j.$$

Using (2.3) we then obtain

$$|\tilde{B}(Du_h, D^2 u_h)| \le C \Phi(x, h)(|D^2 u_h| + 1). \tag{3.8}$$

For $\|Du\|_{L^\infty} < \infty$ we get

$$|(\sqrt{g}\, g^{-1})_h \mathcal{H}|(x) \le C(1 + |D^2 u|^2)^{(p-1)/2}(x)|Du_h|(x) = \Phi(x, h), \qquad (3.9)$$

and moreover the operator

$$\tilde{L}v(x) := (\sqrt{g}\, g^{\alpha\beta})(x + he_v)\partial^2_{\alpha\beta} v(x)$$

is strongly elliptic. We let $B_r \subset \Omega$ and by standard L^p theory there exists a solution $\tilde{\varphi} \in W^{2,q} \cap W^{1,q}_0(B_r, \mathbb{R}^n)$, for every $1 < q < \infty$, of

$$\tilde{L}\tilde{\varphi} = \mathcal{H}$$

satisfying

$$\|\tilde{\varphi}\|_{W^{2,q}(B_r)} \le C\|\mathcal{H}\|_{L^q(B_r)} \le C.$$

Next we let $\eta \in C^\infty_c(B_r)$ be a smooth cut-off function and we define $\varphi_h = \eta^4 \tilde{\varphi}_h$. A standard computation then shows that

$$
\begin{aligned}
\tilde{L}\varphi_h &= \eta^4 \mathcal{H}_h - (\sqrt{g}\, g^{\alpha\beta})_h(x + he_v)\partial^2_{\alpha\beta}(\eta^4 \tilde{\varphi})(x + he_v) \\
&\quad + 4\eta^2\big(\eta(\sqrt{g}\, g^{\alpha\beta})(x + he_v)\partial_\alpha \eta \partial_\beta \tilde{\varphi}_h \\
&\qquad + \eta(\sqrt{g}\, g^{\alpha\beta})(x + he_v)\partial^2_{\alpha\beta}\eta \tilde{\varphi}_h + 3(\sqrt{g}\, g^{\alpha\beta})(x + he_v)\partial_\alpha \eta \partial_\beta \eta \tilde{\varphi}_h\big) \\
&= \eta^4 \mathcal{H}_h + \eta^2 \Phi(\cdot, h)
\end{aligned}
$$

and, by using standard L^2 estimates, we conclude

$$\|\varphi_h\|_{W^{2,2}(B_r)} \le C(\|\eta^2 \mathcal{H}_h\|_{L^2(B_r)} + 1). \qquad (3.10)$$

Using φ_h as a test function in (3.7) and combining the above estimates, we conclude that for every $\delta > 0$ and h small enough,

$$\int_{B_r} \eta^4 |\mathcal{H}_h|^2 \le C \int_{B_r} \Phi(\cdot, h)(|D^2 \varphi_h| + \eta^2 |\mathcal{H}_h| + |D^2 u_h| + 1) \le \delta \int_{B_r} \eta^4 |\mathcal{H}_h|^2 + C_\delta.$$

Choosing δ small enough and letting $h \to 0$ we conclude that $\mathcal{H} \in W^{1,2}_{\text{loc}}(\Omega, \mathbb{R}^n)$ with

$$\int_K \eta^4 |D\mathcal{H}|^2 \le C \quad \text{for all } K \subset\subset \Omega.$$

Combining the estimate for $\mathcal{H}_h$ with Lemma 3.4 yields $H \in W^{1,s}_{\text{loc}}(\Omega, \mathbb{R}^n)$ for every $s < 2$. Arguing as at the beginning of this subsection we conclude that

$$u \in W^{3,s}_{\text{loc}}(\Omega, \mathbb{R}^{n-2}) \quad \text{for all } 1 < s < 2. \qquad \square$$

3.1.3. Higher regularity. In this last subsection we show the higher regularity for solutions of (3.1). We start by showing that $\mathcal{H} \in W^{1,2+\gamma}_{\text{loc}}(\Omega, \mathbb{R}^n)$ for some $0 < \gamma < 1/2$.

In order to see this, we let $B_r \subset \Omega$ and we let $\varphi_1 \in W^{2,2/(1+\gamma)} \cap W_0^{1,2/(1+\gamma)}(B_r, \mathbb{R}^n)$ be the solution of

$$\tilde{L}\varphi_1 = |\mathcal{H}_h|^\gamma \mathcal{H}_h$$

satisfying

$$\|\varphi_1\|_{W^{2,2/(1+\gamma)}(B_r)} \leq C\|\mathcal{H}_h\|_{L^2(B_r)}^{1+\gamma} \leq C.$$

Sobolev's embedding theorem yields

$$\|\varphi_1\|_{L^\infty(B_r)} + \|D\varphi_1\|_{L^{2/\gamma}(B_r)} \leq C.$$

Next we let $\eta \in C_c^\infty(B_r)$ and we define $\tilde{\varphi}_1 = \eta^4 \varphi_1$. We conclude that

$$\begin{aligned}
\tilde{L}\tilde{\varphi}_1 &= \eta^4 |\mathcal{H}_h|^\gamma \mathcal{H}_h + 8\eta^3 (\sqrt{g}\, g^{\alpha\beta})(x + he_\nu)\partial_\alpha \eta \partial_\beta \varphi_1 \\
&\quad + 4\eta^2 \varphi_1 (\sqrt{g}\, g^{\alpha\beta})(x + he_\nu)(3\partial_\alpha \eta \partial_\beta \eta + \eta \partial_{\alpha\beta}^2 \eta) \\
&=: \eta^4 |\mathcal{H}_h|^\gamma \mathcal{H}_h + \eta^2 L[\varphi_1, D\varphi_1].
\end{aligned}$$

Moreover

$$\|\tilde{\varphi}_1\|_{L^\infty(B_r)} + \|D\tilde{\varphi}_1\|_{L^{2/\gamma}(B_r)} + \|D^2\tilde{\varphi}_1\|_{L^{2/(1+\gamma)}(B_r)} \leq C.$$

Now we use $\tilde{\varphi}_1$ as a test function in (3.7) to conclude that

$$\begin{aligned}
\int_{B_r} &\eta^4 |\mathcal{H}_h|^{2+\gamma} \\
&\leq C \int_{B_r} \left(|(\sqrt{g}\, g^{-1})_h \mathcal{H}| \, |D^2\tilde{\varphi}_1| + |\tilde{B}(Du_h, D^2u_h)| \, |D\tilde{\varphi}_1| + \eta^2 |\mathcal{H}_h| \, |L[\varphi_1, D\varphi_1]| \right) \\
&\leq C \int_{B_r} \Phi(x, h)\left(|D^2\tilde{\varphi}_1| + (|D^2u_h| + 1)|D\tilde{\varphi}_1| + |\mathcal{H}_h|(|D\tilde{\varphi}_1| + 1) \right) \leq C.
\end{aligned}$$

Therefore $\mathcal{H} \in W_{\text{loc}}^{1,2+\gamma}(\Omega, \mathbb{R}^n)$. In particular this implies that $\mathcal{H} \in L_{\text{loc}}^\infty(\Omega, \mathbb{R}^{n-2})$ and hence

$$|\mathcal{H}|^2 \in W_{\text{loc}}^{1,2+\gamma}(\Omega).$$

Corollary 3.5 yields $u \in W_{\text{loc}}^{3,2+\gamma}(\Omega, \mathbb{R}^{n-2})$. By the Sobolev embedding theorem this gives

$$u \in C_{\text{loc}}^{2,\beta}(\Omega, \mathbb{R}^{n-2})$$

for some $\beta > 0$. The smoothness of solutions of (3.1) now follows from classical Schauder theory.

3.2. The functional $\mathcal{E}^p(f)$

Here we consider for $p > 2$ weak solutions $u \in W^{2,p}(\Omega, \mathbb{R}^m)$ of elliptic systems in two independent variables of the form

$$\partial_{\alpha\beta}^2 [a_{\alpha\beta}^i(\cdot, u, Du, D^2u)] + \partial_\alpha [b_\alpha^i(\cdot, u, Du, D^2u)] + c^i(\cdot, u, Du, D^2u) = 0. \quad (3.11)$$

We assume that a, b, c are C^1 functions satisfying the following ellipticity and growth conditions at all points (x, z, p, q), for $V(x, z, p, q) = (1 + |q|^2)^{1/2}$ and for constants $\lambda > 0$, $C < \infty$:

$$\frac{\partial a_{\alpha\beta}^i}{\partial q_{\lambda\mu}^j} \xi_{\alpha\beta}^i \xi_{\lambda\mu}^j \geq \lambda V^{p-2} |\xi|^2, \tag{3.12}$$

$$|a_q| \leq CV^{p-2},$$

$$|a| + |a_x| + |a_z| + |a_p| + |b_q| + |c_q| \leq CV^{p-1}, \tag{3.13}$$

$$|b| + |b_x| + |b_z| + |b_p| + |c| + |c_x| + |c_z| + |c_p| \leq CV^p.$$

As noted in Section 2, the graph function of a critical point for $\mathcal{E}^p$ satisfies a system of the required form, with suitable bounds (3.12) and (3.13). Therefore Theorem 1.1 is a consequence of the following proposition and standard higher regularity theory, for which we refer to [19].

Proposition 3.7. *Let $u \in W^{2,p}(\Omega, \mathbb{R}^m)$ be a weak solution of* (3.11), *where $p > 2$ and $\Omega \subset \mathbb{R}^2$, and assume that* (3.12) *and* (3.13) *hold. Then $u \in C_{\mathrm{loc}}^{2,\alpha}(\Omega, \mathbb{R}^m)$ for some $\alpha > 0$.*

Remark 3.8. A related regularity result, for functionals where the integrand satisfies a more general (anisotropic) ellipticity condition but depends only on the second derivatives, was proved in [6]. A crucial ingredient both in [6] and in our paper is the Gehring type lemma from Bildhauer, Fuchs & Zhong [7].

3.2.1. Growth estimate. In a first step we show a growth estimate for the L^p norm of the second derivatives of weak solutions of the system (3.11).

Lemma 3.9. *Let $p > 2$. There exist $r_0, \beta, C > 0$ such that if $u \in W^{2,p}(\Omega, \mathbb{R}^m)$ is a weak solution of the elliptic system* (3.11) *which satisfies* (3.12) *and* (3.13), *then for every $B_{2r}(x) \subset \Omega$ with $r < r_0$,*

$$\int_{B_r(x)} V^p \leq C(r/r_0)^\beta. \tag{3.14}$$

Proof. Let $r_0 > 0$. Since $u \in W^{2,p} \cap L_{\mathrm{loc}}^\infty(\Omega, \mathbb{R}^m)$, the Sobolev embedding theorem implies that $u \in C_{\mathrm{loc}}^{1,\gamma}(\Omega, \mathbb{R}^m)$ for some $\gamma > 0$. Now we choose $x_0 \in \Omega$ and we let $0 < 2r < \min\{2r_0, \mathrm{dist}(x_0, \partial\Omega)\}$. Moreover we let $A_r = B_{2r} \setminus B_r(x_0)$ and $\varphi \in C_c^\infty(B_{2r}(x_0))$ be a smooth cut-off function which satisfies

$$0 \leq \varphi \leq 1, \quad \varphi = 1 \quad \text{in } B_r(x_0), \quad \|\nabla^j \varphi\|_{L^\infty} \leq cr^{-j} \quad \forall j \in \mathbb{N}. \tag{3.15}$$

Finally, we define the linear function l_r by

$$l_r(x) = \frac{1}{|A_r|} \int_{A_r} u + (x - x_0) \cdot \frac{1}{|A_r|} \int_{A_r} Du.$$

From this definition it easily follows that

$$\|u - l_r\|_{L^\infty(B_{2r})} + r\|D(u - l_r)\|_{L^\infty(B_{2r})} \leq Cr^{1+\gamma}, \tag{3.16}$$

$$\|u - l_r\|_{L^p(A_r)} + r\|D(u - l_r)\|_{L^p(A_r)} \leq Cr^2\|D^2u\|_{L^p(A_r)}. \tag{3.17}$$

Now we choose $\varphi^4(u - l_r)$ as a test function in the weak form of (3.11) to get

$$\left| \int_\Omega \varphi^4 a^i_{\alpha\beta} \partial^2_{\alpha\beta} u^i \right| \leq C \int_\Omega \varphi^4 (|b| |D(u - l_r)| + |c| |u - l_r|)$$

$$+ Cr^{-1} \int_{A_r} \varphi^3 (|a| |D(u - l_r)| + |b| |u - l_r|)$$

$$+ Cr^{-2} \int_{A_r} \varphi^2 |a| |u - l_r| = I + II + III.$$

Using the ellipticity assumption and the bound $|a(x, z, p, 0)| \leq C$ (which follows from (3.13)) we estimate

$$q^i_{\alpha\beta} a^i_{\alpha\beta}(x, y, z, q) = q^i_{\alpha\beta} a^i_{\alpha\beta}(x, y, z, 0) + q^i_{\alpha\beta} q^j_{\gamma\delta} \int_0^1 a^i_{\alpha\beta, q^j_{\gamma\delta}}(x, y, z, tq) \, dt$$

$$\geq q^i_{\alpha\beta} a^i_{\alpha\beta}(x, y, z, 0) + \lambda |q|^2 \int_0^1 (1 + t^2|q|^2)^{(p-2)/2} \, dt \geq \tilde{\lambda} V^p - C$$

where $\tilde{\lambda} > 0$ is some number. Next we use (3.13), Hölder's inequality and (3.16) and (3.17) to obtain

$$I \leq Cr^\gamma \int_\Omega \varphi^4 V^p,$$

$$II \leq Cr^\gamma \int_{A_r} V^p + Cr^{-1} \|V\|_{L^p(A_r)}^{p-1} \|D(u - l_r)\|_{L^p(A_r)} \leq C \int_{A_r} V^p,$$

$$III \leq Cr^{-2} \left(\int_{A_r} V^p \right)^{(p-1)/p} \|u - l_r\|_{L^p(A_r)} \leq C \int_{A_r} V^p.$$

Combining all these estimates and choosing r_0 small enough we conclude that there exists a constant $C_1 > 0$ such that

$$\int_{B_r} V^p \leq C_1 \int_{A_r} V^p + Cr^2.$$

Adding $C_1 \int_{B_r} V^p$ to both sides we get

$$\int_{B_r} V^p \leq \frac{C_1}{C_1 + 1} \int_{B_{2r}} V^p + Cr^2.$$

The estimate (3.14) now follows from a standard iteration argument. □

3.2.2. Difference quotient estimates. In a second step we use the difference quotient method to show that every weak solution $u \in W^{2,p} \cap L^\infty(\Omega, \mathbb{R}^m)$ of (3.11) is in $W^{3,2}(B_{r_0/4}(x_0), \mathbb{R}^m)$ and moreover $V^{p/2} \in W^{1,2}(B_{r_0/4}(x_0))$, where $x_0 \in \Omega$ and r_0 is

as in Lemma 3.9 (in the following we allow the constants to depend on r_0). We follow closely the methods developed in [19] and [22].

Below, we use the abbreviation $U(x) = (x, u(x), Du(x), D^2u(x))$. Applying the difference quotient to (3.11) and interchanging with the derivatives yields

$$\partial^2_{\alpha\beta}[a^i_{\alpha\beta} \circ U]_h + \partial_\alpha[b^i_\alpha \circ U]_h + [c^i \circ U]_h = 0. \tag{3.18}$$

We use the fundamental theorem of calculus to write

$$[f \circ U]_h(x) = \frac{1}{h}\big(f(U(x + he_v)) - f(U(x))\big)$$

$$= \frac{1}{h}\int_0^1 \frac{d}{dt} f\big((1-t)U(x) + t\,U(x + he_v)\big)\,dt$$

$$= \int_0^1 Df\big((1-t)U(x) + t\,U(x + he_v)\big)\,dt \cdot U_h(x).$$

Using the notation $f^h(x) = \int_0^1 f\big((1-t)U(x) + tU(x + he_v)\big)\,dt$ we thus get

$$[f \circ U]_h = \left(\frac{\partial f}{\partial q^j_{\lambda\mu}}\right)^h \partial^2_{\lambda\mu} u^j_h + \left(\frac{\partial f}{\partial p^j_\lambda}\right)^h \partial_\lambda u^j_h + \left(\frac{\partial f}{\partial z^j}\right)^h u^j_h + \left(\frac{\partial f}{\partial x_v}\right)^h,$$

and the system (3.18) takes the form

$$\partial^2_{\alpha\beta}[\tilde{a}^i_{\alpha\beta}(\cdot, u_h, Du_h, D^2u_h)] + \partial_\alpha[\tilde{b}^i_\alpha(\cdot, u_h, Du_h, D^2u_h)] + \tilde{c}^i(\cdot, u_h, Du_h, D^2u_h) = 0. \tag{3.19}$$

Here the coefficients are as follows:

$$\tilde{a}^i_{\alpha\beta}(x, z, p, q) = \left(\frac{\partial a^i_{\alpha\beta}}{\partial q^j_{\lambda\mu}}\right)^h(x)q^j_{\lambda\mu} + \left(\frac{\partial a^i_{\alpha\beta}}{\partial p^j_\lambda}\right)^h(x)p^j_\lambda + \left(\frac{\partial a^i_{\alpha\beta}}{\partial z^j}\right)^h(x)z^j + \left(\frac{\partial a^i_{\alpha\beta}}{\partial x_v}\right)^h(x),$$

$$\tilde{b}^i_\alpha(x, z, p, q) = \left(\frac{\partial b^i_\alpha}{\partial q^j_{\lambda\mu}}\right)^h(x)q^j_{\lambda\mu} + \left(\frac{\partial b^i_\alpha}{\partial p^j_\lambda}\right)^h(x)p^j_\lambda + \left(\frac{\partial b^i_\alpha}{\partial z^j}\right)^h(x)z^j + \left(\frac{\partial b^i_\alpha}{\partial x_v}\right)^h(x),$$

$$\tilde{c}^i(x, z, p, q) = \left(\frac{\partial c^i}{\partial q^j_{\lambda\mu}}\right)^h(x)q^j_{\lambda\mu} + \left(\frac{\partial c^i}{\partial p^j_\lambda}\right)^h(x)p^j_\lambda + \left(\frac{\partial c^i}{\partial z^j}\right)^h(x)z^j + \left(\frac{\partial c^i}{\partial x_v}\right)^h(x).$$

In order to state bounds for these coefficients, we introduce the abbreviation

$$I_{s,h}(x) = \int_0^1 \big(1 + |(1-t)D^2u(x) + tD^2u(x+h)|^2\big)^{s/2}\,dt.$$

Using (3.12) and (3.13) we then obtain

$$|\tilde{a}(x, z, p, q)| \le C\big(I_{p-2,h}(x)|q| + I_{p-1,h}(x)(|p| + |z| + 1)\big),$$

$$|\tilde{b}(x, z, p, q)| + |\tilde{c}(x, z, p, q)| \le C\big(I_{p-1,h}(x)|q| + I_{p,h}(x)(|p| + |z| + 1)\big).$$

As above we define the linear function $l_{h,r} =: l_h$ by

$$l_h(x) = \frac{1}{|A_r|} \int_{A_r} u_h + (x - x_0) \cdot \frac{1}{|A_r|} \int_{A_r} Du_h.$$

Using again the test function $\varphi^4(u_h - l_h)$, we infer that

$$\int \tilde{a}^i_{\alpha\beta} \partial^2_{\alpha\beta} u^i_h \varphi^4$$

$$\leq C \int \left(I_{p-1,h} |D^2 u_h| + I_{p,h}(|Du_h| + |u_h| + 1) \right)(|D(u_h - l_h)| + |u_h - l_h|)\varphi^4$$

$$+ C \int \left(I_{p-2,h}|D(u_h - l_h)| + I_{p-1,h}|u_h - l_h| \right)|D^2 u_h|\varphi^3|D\varphi|$$

$$+ C \int \left(I_{p-1,h}|D(u_h - l_h)| + I_{p,h}|u_h - l_h| \right)(|Du_h| + |u_h| + 1)\,\varphi^3|D\varphi|$$

$$+ C \int \left(I_{p-2,h}|D^2 u_h| + I_{p-1,h}(|Du_h| + |u_h| + 1) \right)|u_h - l_h|\varphi^2(\varphi|D^2\varphi| + |D\varphi|^2).$$

$$(3.20)$$

On the other hand we have the ellipticity condition (using (3.12))

$$\tilde{a}^i_{\alpha\beta}(x, z, p, q)q^i_{\alpha\beta} = \int_0^1 \frac{\partial a^i_{\alpha\beta}}{\partial q^j_{\lambda\mu}}\big((1-t)U(x) + tU(x+h)\big)q^i_{\alpha\beta}q^j_{\lambda\mu}\,dt$$

$$+ \int_0^1 \frac{\partial a^i_{\alpha\beta}}{\partial p^j_\lambda}\big((1-t)U(x) + tU(x+h)\big)p^j_\lambda q^i_{\alpha\beta}\,dt$$

$$+ \int_0^1 \frac{\partial a^i_{\alpha\beta}}{\partial z^j}\big((1-t)U(x) + tU(x+h)\big)z^j q^i_{\alpha\beta}\,dt$$

$$+ \int_0^1 \frac{\partial a^i_{\alpha\beta}}{\partial x_\nu}\big((1-t)U(x) + tU(x+h)\big)q^i_{\alpha\beta}\,dt$$

$$\geq C\lambda I_{p-2,h}|q|^2 - C I_{p-1,h}\,|q|(|p| + |z| + 1).$$

Combining the two inequalities we arrive at

$$\int I_{p-2,h}|D^2 u_h|^2\varphi^4$$

$$\leq C \int I_{p-1,h}|D^2 u_h|\big(|D(u_h - l_h)|\varphi^4 + |u_h - l_h|(\varphi^4 + \varphi^3|D\varphi|) + (|Du_h| + |u_h| + 1)\varphi^4\big)$$

$$+ C \int I_{p-2,h}|D^2 u_h|\big(|D(u_h - l_h)|\varphi^3|D\varphi| + |u_h - l_h|\varphi^2(\varphi|D^2\varphi| + |D\varphi|^2)\big)$$

$$+ C \int I_{p,h}(|Du_h| + |u_h| + 1)\big(|D(u_h - l_h)|\varphi^4 + |u_h - l_h|(\varphi^4 + \varphi^3|D\varphi|)\big)$$

$$+ C \int I_{p-1,h}(|Du_h| + |u_h| + 1)\big(|D(u_h - l_h)|\varphi^3|D\varphi| + |u_h - l_h|\varphi^2(\varphi|D^2\varphi| + |D\varphi|^2)\big).$$

Using $I_{p-1,h} \leq I_{p-2,h}^{1/2} I_{p,h}^{1/2}$ and absorbing the second derivatives of u_h yields

$$\int I_{p-2,h} |D^2 u_h|^2 \varphi^4$$

$$\leq C \int I_{p,h} \big(|D(u_h - l_h)|^2 + |u_h - l_h|^2 + |Du_h|^2 + |u_h|^2 + 1 \big) \varphi^4$$

$$+ C \int I_{p,h} |u_h - l_h|^2 \varphi^2 |D\varphi|^2$$

$$+ C \int I_{p-2,h} \big(|D(u_h - l_h)|^2 \varphi^2 |D\varphi|^2 + |u_h - l_h|^2 (|D\varphi|^4 + \varphi^2 |D^2\varphi|^2) \big)$$

$$= I + II + III.$$

Before continuing we need to recall the following lemma which is essentially due to Morrey [19, Lemma 5.4.2]. In the form stated here, it can be found in [22].

Lemma 3.10. *Let $r > 0$ and let $q \geq 0$ be a function such that*

$$\int_{B_s(x) \cap B_r} q \leq c s^\gamma$$

for all $B_s(x) \subset B_{2r}$. Then for every $\varepsilon > 0$ there exists $C_\varepsilon > 0$ such that

$$\int_{B_r} q |v|^2 \leq \varepsilon r^\gamma \int_{B_r} |Dv|^2 + C_\varepsilon r^{\gamma - 2} \int_{B_r} |v|^2.$$

Next we use this lemma in order to estimate the terms *I–III* from above. By the definition of $I_{p,h}$ and Lemma 3.9 we get

$$\int_{B_s(x)} I_{p,h} \leq C s^\beta$$

for all $B_s(x) \subset B_{2r} \subset B_{r_0}$ and some $0 < \beta < 1$. Hence we can apply Lemma 3.10 and Poincaré's inequality to estimate

$$II \leq C r^{\beta-2} \| D(u_h - l_h) \|_{L^2(A_r)}^2.$$

Using the same argument we get

$$I \leq C r^\beta \int_{B_{2r}} \varphi^4 (|D^2 u_h|^2 + |D(u_h - l_h)|^2)$$

$$+ C r^{\beta-2} \int_{B_{2r}} \big(|Du_h|^2 + |D(u_h - l_h)|^2 + |u_h|^2 + |u_h - l_h|^2 + I_{p,h} \big).$$

Inserting these two estimates into the above estimate for $\int I_{p-2,h} |D^2 u_h|^2 \varphi^4$ we conclude that

$$\int I_{p-2,h}|D^2u_h|^2\varphi^4 \le Cr^\beta \int_{B_{2r}} \varphi^4(|D^2u_h|^2 + |D(u_h - l_h)|^2)$$

$$+ Cr^{\beta-2}\int_{B_{2r}}\big(|Du_h|^2 + |D(u_h - l_h)|^2 + |u_h|^2 + |u_h - l_h|^2 + I_{p,h}\big)$$

$$+ C\int I_{p-2,h}\big(|D(u_h - l_h)|^2\varphi^2|D\varphi|^2 + |u_h - l_h|^2(|D\varphi|^4 + \varphi^2|D^2\varphi|^2)\big). \quad (3.21)$$

Next we use Hölder's and Poincaré's inequalities to get

$$\int I_{p-2,h}\left(|D(u_h - l_h)|^2\frac{\chi_{A_r}}{r^2} + |u_h - l_h|^2\frac{\chi_{A_r}}{r^4}\right)$$

$$\le C\|I_{p-2,h}\|_{L^{p/(p-2)}(B_{2r})}\big(r^{-4}\|u_h - l_h\|^2_{L^p(A_r)} + r^{-2}\|D(u_h - l_h)\|_{L^p(A_r)}\big)$$

$$\le Cr^{-2}\|I_{p-2,h}\|_{L^{p/(p-2)}(B_{2r})}\|D(u_h - l_h)\|^2_{L^p(A_r)}.$$

Since $u \in W^{2,p}(\Omega)$ we know from [19, Theorem 3.6.8] that

$$I_{s,h} \to V^s \quad \text{in } L^{p/s} \quad \forall 1 \le s \le p.$$

We combine all the above estimates to get ($l_h \to \fint_{A_r} \partial_\nu u + (x - x_0)\fint_{A_r} \partial_\nu Du$) for r_0 small enough

$$\int \varphi^4 I_{p-2,h}|D^2u_h|^2 \le C(1 + r^{-2})\int_{B_{2r}} V^p + Cr^{\beta-2}\int_{B_{2r}} |D^2u|^2 + Cr^\beta.$$

In particular this estimate is true for $r = r_0/4$ and therefore we can let $h \to 0$ to conclude

$$\int_{B_{r_0/8}} V^{p-2}|D^3u|^2 \le C(1 + r_0^{-2})\int_{B_{r_0/4}} V^p + cr_0^{\beta-2}\int_{B_{r_0/4}} |D^2u|^2 + Cr_0^\beta. \quad (3.22)$$

Hence $u \in W^{3,2}(B_{r_0/8})$ and by the Sobolev embedding theorem this implies that $u \in W^{2,q}_{\text{loc}}(B_{r_0/8})$ for all $q < \infty$. Moreover the above estimate yields $V^{p/2} \in W^{1,2}(B_{r_0/8})$. Altogether this shows that we can improve Lemma 3.9 to get the estimate

$$\int_{B_r} V^p \le cr^{2-\delta} \quad (3.23)$$

for all $r \le r_0/16$ and all $\delta > 0$.

3.2.3. Higher regularity. It turns out that estimating with difference quotients and Morrey's lemma are not sufficient to get the critical L^∞ estimate for D^2u. Therefore we modify the estimates in order to apply a Gehring type lemma from [7].

Recalling (3.20) we have

$$\int \varphi^4 \tilde{a}_{\alpha\beta}^i \partial_{\alpha\beta}^2 u_h^i$$

$$\leq C \int \varphi^4 \big(I_{p-1,h} |D^2 u_h| + I_{p,h}(|Du_h| + |u_h| + 1) \big)(|D(u_h - l_h)| + |u_h - l_h|)$$

$$+ \frac{C}{r} \int \varphi^3 \big(I_{p-2,h} |D(u_h - l_h)| + I_{p-1,h} |u_h - l_h| \big)|D^2 u_h|$$

$$+ \frac{C}{r} \int \varphi^3 \big(I_{p-1,h} |D(u_h - l_h)| + I_{p,h} |u_h - l_h| \big)(|Du_h| + |u_h| + 1)$$

$$+ \frac{C}{r^2} \int \varphi^2 \big(I_{p-2,h} |D^2 u_h| + I_{p-1,h}(|Du_h| + |u_h| + 1) \big)|u_h - l_h|$$

$$= I + \cdots + IV.$$

This time we choose l_h such that $\int_{B_{2r}} (u_h - l_h) = 0$ and $\int_{B_{2r}} D(u_h - l_h) = 0$. Because of (3.23) and the strong convergence $I_{s,h} \to V^s$ in $L^{p/s}$ we have, for every $r \leq r_0/16$, every h small enough and every $\delta > 0$,

$$\int_{B_r} I_{p,h} \leq c r^{2-\delta}. \tag{3.24}$$

Now we again estimate each term separately. We start with I. By Young's inequality we get

$$I \leq \varepsilon \int \varphi^4 I_{p-2,h} |D^2 u_h|^2 + C \int \varphi^4 I_{p,h}(|Du_h|^2 + |u_h|^2 + 1)$$

$$+ C \int \varphi^4 I_{p,h}(|D(u_h - l_h)|^2 + |u_h - l_h|^2),$$

and we continue to estimate the last term with the help of Lemma 3.10, (3.24) and Poincaré's inequality by

$$\int \varphi^4 I_{p,h}(|D(u_h - l_h)|^2 + |u_h - l_h|^2) \leq C r^{2-\delta} \int_{B_{2r}} (|D^2(u_h - l_h)|^2 + |D(u_h - l_h)|^2)$$

$$+ C r^{-\delta} \int_{B_{2r}} (|D(u_h - l_h)|^2 + |u_h - l_h|^2)$$

$$\leq C r^{2-\delta} \int_{B_{2r}} |D^2 u_h|^2.$$

Next we estimate

$$II \leq \varepsilon \int \varphi^4 I_{p-2,h} |D^2 u_h|^2 + \frac{C}{r^2} \int_{B_{2r}} I_{p,h} |u_h - l_h|^2$$

$$+ \frac{C}{r} \left(\int_{B_{2r}} I_{p-2,h}^{4/3} |D^2 u_h|^{4/3} \right)^{3/4} \left(\int_{B_{2r}} |D(u_h - l_h)|^4 \right)^{1/4}.$$

The second term can be estimated as above:

$$\frac{C}{r^2} \int_{B_{2r}} I_{p,h} |u_h - l_h|^2 \le C r^{2-\delta} \int_{B_{2r}} |D^2 u_h|^2,$$

and for the third term we use the Sobolev–Poincaré inequality to get

$$\frac{C}{r} \left(\int_{B_{2r}} I_{p-2,h}^{4/3} |D^2 u_h|^{4/3} \right)^{3/4} \left(\int_{B_{2r}} |D(u_h - l_h)|^4 \right)^{1/4} \le \frac{C}{r} \left(\int_{B_{2r}} I_{p-2,h}^{4/3} |D^2 u_h|^{4/3} \right)^{3/2}.$$

III can be estimated by

$$\begin{aligned}
III &\le \frac{\gamma}{r^2} \int_{B_{2r}} (|D(u_h - l_h)|^2 + I_{p,h} |u_h - l_h|^2) \\
&\quad + C_\gamma \int_{B_{2r}} (I_{2p-2,h} + I_{p,h})(|Du_h|^2 + |u_h|^2 + 1) \\
&\le C(\gamma + r^{2-\delta}) \int_{B_{2r}} |D^2 u_h|^2 + C_\gamma \int_{B_{2r}} (I_{2p-2,h} + I_{p,h})(|Du_h|^2 + |u_h|^2 + 1).
\end{aligned}$$

Finally, using some of the estimates above, the last term is estimated as follows:

$$\begin{aligned}
IV &\le \frac{C}{r^2} \left(\int_{B_{2r}} I_{p-2,h}^{4/3} |D^2 u_h|^{4/3} \right)^{3/4} \left(\int_{B_{2r}} |u_h - l_h|^4 \right)^{1/4} \\
&\quad + \frac{\gamma}{r^4} \int_{B_{2r}} |u_h - l_h|^2 + C_\gamma \int_{B_{2r}} I_{2p-2,h}(|Du_h|^2 + |u_h|^2 + 1) \\
&\le \frac{C}{r} \left(\int_{B_r} I_{p-2,h}^{4/3} |D^2 u_h|^{4/3} \right)^{3/2} + C\gamma \int_{B_{2r}} |D^2 u_h|^2 \\
&\quad + C_\gamma \int_{B_{2r}} I_{2p-2,h}(|Du_h|^2 + |u_h|^2 + 1).
\end{aligned}$$

We also note that we have the ellipticity estimate

$$\int \varphi^4 I_{p-2,h} |D^2 u_h|^2 \le C \int \varphi^4 \tilde{a}_{\alpha\beta}^i \partial_{\alpha\beta}^2 u_h^i + C \int \varphi^4 I_{p,h}(|Du_h|^2 + |u_h|^2 + 1).$$

Combining all these estimates we get

$$\begin{aligned}
\int_{B_r} I_{p-2,h} |D^2 u_h|^2 &\le \frac{C}{r} \left(\int_{B_{2r}} I_{p-2,h}^{4/3} |D^2 u_h|^{4/3} \right)^{3/2} + C(\gamma + r^{2-\delta}) \int_{B_{2r}} |D^2 u_h|^2 \\
&\quad + C_\gamma \int_{B_{2r}} (I_{2p-2,h} + I_{p,h})(|Du_h|^2 + |u_h|^2 + 1).
\end{aligned}$$

Since $u \in W^{3,2}(B_r)$ and $V^{p/2} \in W^{1,2}(B_r)$ for $r \le r_0/16$ we can let $h \to 0$ and get

$$\int_{B_r} V^{p-2}|D^3 u|^2 \le \frac{C}{r}\left(\int_{B_{2r}} V^{4(p-2)/3}|D^3 u|^{4/3}\right)^{3/2} + C(\gamma + r_0^{2-\delta})\int_{B_{2r}} V^{p-2}|D^3 u|^2$$
$$+ C_\gamma \int_{B_{2r}} V^{2p}.$$

Defining $f = (V^{(p-2)/2}|D^3 u|)^{4/3}$, $g = V^{2(p-2)/3}$, $h = V^{4p/3}$ and $d = 3/2$ we conclude that

$$\left(\fint_{B_r} f^d\right)^{1/d} \le C\fint_{B_{2r}} fg + C(\gamma + r_0^{2-\delta})^{1/d}\left(\fint_{B_{2r}} f^d\right)^{1/d} + C_\gamma\left(\fint_{B_{2r}} h^d\right)^{1/d} \quad (3.25)$$

for all balls $B_r \subset B_{r_0/16}$. Next we need the following Gehring type lemma, which slightly generalizes Lemma 1.2 of Bildhauer, Fuchs & Zhong [7] (see also [11, Theorem 1.1]).

Lemma 3.11. *Let $d > 1$ and $\beta > 0$ be two constants. There exists $\varepsilon_0 > 0$ such that for all $\varepsilon < \varepsilon_0$ and all non-negative functions $f, g, h : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ satisfying*

$$f, h \in L_{\mathrm{loc}}^d(\Omega), \qquad e^{\beta g^d} \in L_{\mathrm{loc}}^1(\Omega)$$

and (for some constant $C > 0$)

$$\left(\fint_B f^d\right)^{1/d} \le C\fint_{2B} fg + \varepsilon\left(\fint_{2B} f^d\right)^{1/d} + C\left(\fint_{2B} h^d\right)^{1/d} \quad (3.26)$$

for all balls $B = B_r(x)$ with $B_{2r}(x) \subset\subset \Omega$, there exists $c_0 = c_0(n, d, C) > 0$ such that if

$$h^d \log^{c_0\beta}(e + h) \in L_{\mathrm{loc}}^1(\Omega),$$

then the same is true for f. Moreover, for all balls B as above we have

$$\fint_B f^d \log^{c_0\beta}\left(e + \frac{f}{\|f\|_{d,2B}}\right) \le c\left(\fint_{2B} e^{\beta g^d}\right)\left(\fint_{2B} f^d\right)$$
$$+ c\fint_{2B} h^d \log^{c_0\beta}\left(e + \frac{f}{\|f\|_{d,2B}}\right), \quad (3.27)$$

where $c = c(n, d, \beta, C) > 0$ and $\|f\|_{d,2B} = (\fint_{2B} f^d)^{1/d}$.

Proof. The proof is very similar to the one of [7, Lemma 1.2] and therefore we only comment on the differences.

We define $B_0 = 2B$ and we assume without loss of generality that

$$\int_{B_0} f^d = 1.$$

Next we define the functions $d(x) = \mathrm{dist}(x, \mathbb{R}^n \backslash B_0)$ and

$$\tilde{f}(x) = d(x)^{n/d} f, \quad \tilde{h}(x) = d(x)^{n/d} h, \quad w(x) = \chi_{B_0}(x),$$

where χ_{B_0} is the characteristic function of B_0. As in [7] it is now easy to see that because of (3.26) these new functions satisfy

$$\left(\fint_B \tilde{f}^d\right)^{1/d} \le C\fint_{2B} \tilde{f}g + C\varepsilon\left(\fint_{2B} \tilde{f}^d\right)^{1/d} + C\left(\fint_{2B} \tilde{h}^d\right)^{1/d} + C\left(\fint_{2B} w\right)^{1/d},$$

and now this inequality is true for all balls $B \subset \mathbb{R}^n$. Hence, by taking the supremum over all radii, we get (here $M(f)$ denotes the maximal function of f)

$$M(\tilde{f}^d)^{1/d} \le CM(\tilde{f}g) + C\varepsilon M(\tilde{f}^d)^{1/d} + CM(\tilde{h}^d)^{1/d} + CM(w)^{1/d}.$$

For ε_0 small enough we therefore have

$$M(\tilde{f}^d)^{1/d} \le CM(\tilde{f}g) + CM(\tilde{h}^d)^{1/d} + CM(w)^{1/d},$$

and with the help of this inequality we can copy the rest of the argument from [7, proof of Lemma 1.2] to finish the proof. $\qquad\square$

Now we want to apply this lemma to our estimate (3.25). From the previous subsection we know that

$$f^d = V^{p-2}|D^3u|^2 \in L^1_{\text{loc}}(B_{r_0/16}), \quad h^d = V^{2p} \in L^1_{\text{loc}}(B_{r_0/16}).$$

Hence it remains to check that

$$e^{\beta g^d} = e^{\beta V^{p-2}} \in L^1_{\text{loc}}(B_{r_0/16}) \tag{3.28}$$

for some constant $\beta > 0$. We actually claim that this is true for all $\beta > 0$. In order to see this, we note that by (3.22) we have

$$\int_{B_{r_0/8}} |D(V^{p/2})|^2 \le c_1(r_0).$$

Next we let $\eta \in C_c^\infty(B_{r_0/8})$ be a cut-off function such that $0 \le \eta \le 1$, $\eta(x) \equiv 1$ for all $x \in B_{r_0/16}$ and $\|D\eta\|_{L^\infty(B_{r_0/8})} \le cr_0^{-1}$. Defining $v = \eta V^{p/2}$ we get

$$\int_{B_{r_0/8}} |Dv|^2 \le cr_0^{-2}\int_{B_{r_0/8}} V^p + c\int_{B_{r_0/8}} |DV^{p/2}|^2 \le c_2(r_0).$$

Hence we see that $u = v/\sqrt{c_2(r_0)} \in H_0^1(B_{r_0/8})$ and

$$\int_{B_{r_0/8}} |Du|^2 \le 1.$$

Therefore, by the Moser–Trudinger inequality (see [23]), there exist constants $\beta_0, C = C(r_0) > 0$ such that

$$\int_{B_{r_0/16}} e^{\beta_0 V^p} \le \int_{B_{r_0/8}} e^{c_2(r_0)\beta_0 u^2} \le C.$$

In particular this implies with the help of Young's inequality that for every $\beta > 0$,

$$\int_{B_{r_0/16}} e^{\beta V^{p-2}} \le c(\beta, \beta_0) \int_{B_{r_0/16}} e^{\beta_0 V^p} \le C(r_0, \beta, \beta_0).$$

Since also

$$h^d \log^\alpha(e + h) = V^{2p} \log^\alpha(e + V^{4p/3}) \in L^1_{\text{loc}}(B_{r_0/16})$$

for every $\alpha > 0$, Lemma 3.11 implies that

$$f^d \log^\alpha(e + f) \in L^1_{\text{loc}}(B_{r_0/16})$$

for every $\alpha > 0$. Hence

$$|D^3 u|^2 \log^\alpha(e + |D^3 u|) \in L^1_{\text{loc}}(B_{r_0/16})$$

for every $\alpha > 0$. In particular, this is true for $\alpha > 1$ and therefore we can apply [10, Corollary 4.6 and Example 4.18(iv)] (see also [12, Example 5.3] for a different proof of this result) in order to conclude that

$$u \in C^2(B_{r_0/32}).$$

In particular this implies that

$$\int_{B_r} V^p \le c r^2 \quad \text{for all } r \le r_0/32. \tag{3.29}$$

In order to show the Hölder continuity of $D^2 u$ we go back to (3.21) and we estimate the last term by

$$\int I_{p-2,h}\left(|D(u_h - l_h)|^2 \frac{\chi_{A_r}}{r^2} + |u_h - l_h|^2 \frac{\chi_{A_r}}{r^4}\right) \le C \int_{A_r} |D^2 u_h|^2.$$

Inserting this estimate into (3.21), letting $h \to 0$ and using (3.29) we conclude that

$$\int_{B_r} |D^3 u|^2 \le C \int_{A_r} |D^3 u|^2 + C r^\beta \quad \text{for every } r \le r_0/32. \tag{3.30}$$

Remark 3.12. This estimate is sufficient for our purposes, but by repeating all the estimates from Subsection 3.2.2 and replacing every application of Lemma 3.9 by (3.29) one can actually improve this inequality in the sense that the term r^β on the right hand side can be replaced by r^2.

The Hölder continuity of $D^2 u$ now follows from (3.30) by another hole-filling argument. This finishes the proof of Proposition 3.7.

4. Compactness results and existence of minimizers

4.1. Compactness results

We start by quoting the fundamental compactness theorem of J. Langer (see also [8]).

Theorem 4.1 ([16]). *Let Σ be a closed surface and $p > 2$. Assume that $f_k \in W^{2,p}_{\mathrm{im}}(\Sigma, \mathbb{R}^n)$ is a sequence satisfying $0 \in f_k(\Sigma)$ for all $k \in \mathbb{N}$ and*

$$\mathcal{E}^p(f_k) \le C. \tag{4.1}$$

After replacing f_k by $f_k \circ \varphi_k$ for suitable diffeomorphisms $\varphi_k \in C^\infty(\Sigma, \Sigma)$ and passing to a subsequence, the f_k converge weakly in $W^{2,p}(\Sigma, \mathbb{R}^n)$ to an $f \in W^{2,p}_{\mathrm{im}}(\Sigma, \mathbb{R}^n)$. In particular, the convergence is in $C^{1,\beta}(\Sigma, \mathbb{R}^n)$ for any $\beta < 1 - 2/p$, and

$$\mathcal{E}^p(f) \le \liminf_{k \to \infty} \mathcal{E}^p(f_k). \tag{4.2}$$

In this section we prove Theorem 1.2, which replaces the $\mathcal{E}^p$ bound in Langer's theorem by a bound only for $\mathcal{W}^p$, under the additional assumption that the Willmore energy is bounded by 8π. Before entering the proof we include two remarks about the statement.

Remark 4.2. One can allow sequences $f_k : \Sigma_k \to \mathbb{R}^n$ in Theorem 1.2, where Σ_k are arbitrary closed oriented surfaces. In fact, a bound on the genus follows from the condition $\liminf_{k \to \infty} \mathcal{W}(f_k) < 8\pi$ by a result of Kuwert, Li & Schätzle [14].

Remark 4.3. Connecting two round spheres by a shrinking catenoid neck yields a sequence of smoothly embedded surfaces with bounded $\mathcal{W}^p$ energy and Willmore energy less than 8π. As the convergence is not in C^1, this shows that the assumption on the Willmore energy in Theorem 1.2 cannot be weakened. Similar constructions are also possible for higher genus: see Kühnel & Pinkall [13] and Simon [22].

To prove Theorem 1.2 we need the following area ratio bounds, which are immediate consequences of Simon's monotonicity identity [22].

Lemma 4.4. *Let $f : \Sigma \to \mathbb{R}^n$ be an embedded closed surface. Then*

$$\sigma^{-2} \mu_f(B_\sigma) \le C \mathcal{W}(f) \quad \text{for all } \sigma > 0. \tag{4.3}$$

Moreover, for any $p > 2$,

$$\sigma^{-2} \mu_f(B_\sigma) \le \tfrac{1}{4} \mathcal{W}(f) + C \sigma^{(p-2)/p} \quad \text{for all } \sigma > 0, \tag{4.4}$$

where the constant C depends on $\mathcal{W}^p(f)$.

Proof. By [22, (1.2)], for $0 < \sigma < \rho < \infty$ we have

$$\sigma^{-2}\mu_f(B_\sigma) \le \rho^{-2}\mu_f(B_\rho) + \tfrac{1}{4}\mathcal{W}(f)$$
$$+ \frac{1}{2}\int_{\Sigma \cap B_\rho} \rho^{-2}\langle x, H\rangle\, d\mu - \frac{1}{2}\int_{\Sigma \cap B_\sigma} \sigma^{-2}\langle x, H\rangle\, d\mu.$$

Letting $\rho \to \infty$ we conclude that for every $\sigma > 0$,

$$\sigma^{-2}\mu_f(B_\sigma) \le \tfrac{1}{4}\mathcal{W}(f) + \frac{1}{2\sigma}\int_{\Sigma \cap B_\sigma} |H|\, d\mu \le \tfrac{1}{4}\mathcal{W}(f) + C\sigma^{(p-2)/p}(\mathcal{W}^p(f))^{1/p}, \quad (4.5)$$

where we have used Hölder's inequality in the last step. □

As a second ingredient, we need the following lemma yielding an L^p estimate for the prescribed mean curvature system.

Lemma 4.5. *Let $u \in W^{2,p}(B_\varrho, \mathbb{R}^k)$, where $B_\varrho = \{x \in \mathbb{R}^2 : |x| < \varrho\}$ and $0 < \varrho < \infty$, $p \in (1, \infty)$, be a solution of the system*

$$a_{ij}^{\alpha\beta}\partial_{\alpha\beta}^2 u^i = \varphi_j \quad \text{for } j = 1, \ldots, k.$$

There is an $\varepsilon_0 = \varepsilon_0(p) > 0$ such that if

$$|a_{ij}^{\alpha\beta}(x) - \delta^{\alpha\beta}\delta_{ij}| \le \varepsilon_0 \quad \text{for all } x \in B_\varrho,$$

then for some $C = C(p) < \infty$ we have the estimate

$$\|D^2 u\|_{L^p(B_{\varrho/2})} \le C(\|\varphi\|_{L^p(B_\varrho)} + \varrho^{-1}\|Du\|_{L^p(B_\varrho)}).$$

Proof. We may assume that $\varrho = 1$ and that u has mean value zero on B_1. For $\eta \in C_0^\infty(B_1)$ satisfying $\eta = 1$ on $B_{1/2}$ and $\eta = 0$ in $B_1 \setminus B_{3/4}$, we calculate

$$a_{ij}^{\alpha\beta}\partial_{\alpha\beta}^2(\eta u^i) = \eta\varphi_j + a_{ij}^{\alpha\beta}(\partial_{\alpha\beta}^2\eta)u^i + a_{ij}^{\alpha\beta}(\partial_\alpha\eta\,\partial_\beta u^i + \partial_\beta\eta\,\partial_\alpha u^i).$$

Hence we have

$$\Delta(\eta u^j) = (\delta^{\alpha\beta}\delta_{ij} - a_{ij}^{\alpha\beta})\partial_{\alpha\beta}^2(\eta u^i) + \eta\varphi_j$$
$$+ a_{ij}^{\alpha\beta}(\partial_{\alpha\beta}^2\eta)u^i + a_{ij}^{\alpha\beta}(\partial_\alpha\eta\,\partial_\beta u^i + \partial_\beta\eta\,\partial_\alpha u^i).$$

From standard L^p estimates and the Poincaré inequality we obtain

$$\|D^2(\eta u)\|_{L^p(B_1)} \le C\varepsilon_0\|D^2(\eta u)\|_{L^p(B_1)} + C(\|\varphi\|_{L^p(B_1)} + \|Du\|_{L^p(B_1)}),$$

for a constant $C = C(p) < \infty$. This shows the desired result. □

Proof of Theorem 1.2. Let $f_k : \Sigma \to \mathbb{R}^n$ be a sequence as in the theorem. For each $q \in \Sigma$, we let $r_k(q) > 0$ be the maximal radius on which f_k is represented as a graph

over the tangent plane at q. We denote by $u_{k,q} : B_{r_k(q)} \to \mathbb{R}^{n-2}$ the corresponding graph function, obtained by choosing a suitable rigid motion. In particular

$$u_{k,q}(0) = 0 \quad \text{and} \quad Du_{k,q}(0) = 0.$$

For $\varepsilon > 0$ we define $r_k(q, \varepsilon) = \sup\{r \in (0, r_k(q)] : \|Du_{k,q}\|_{C^0(B_r)} < \varepsilon\}$ and

$$r_k = \inf_{q \in \Sigma} r_k(q, \varepsilon).$$

By compactness, the infimum is attained at some point $q_k \in \Sigma$ and we have $r_k > 0$. We will show by contradiction that

$$\liminf_{k \to \infty} r_k > 0. \tag{4.6}$$

Assuming that $r_k \to 0$ we rescale by setting

$$\tilde{f}_k : \Sigma \to \mathbb{R}^n, \quad \tilde{f}_k(p) = \frac{1}{r_k}(f_k(p) - f_k(q_k)).$$

Clearly, the $\tilde{f}_k$ have local graph representations

$$\tilde{u}_{k,q} : B_{r_k(q)/r_k} \to \mathbb{R}^{n-2}, \quad \tilde{u}_{k,q}(x) = \frac{1}{r_k}u_{k,q}(r_k x),$$

where $\tilde{u}_{k,q}(0) = 0$ and $D\tilde{u}_{k,q}(0) = 0$, and

$$\|\tilde{u}_{k,q}\|_{C^0(B_1)} + \|D\tilde{u}_{k,q}\|_{C^0(B_1)} \leq C\varepsilon.$$

From the bound $\mathcal{W}^p(f_k) \leq C$ we further infer that

$$\int_{\Sigma} |H_{\tilde{f}_k}|^p \, d\mu_{\tilde{f}_k} = r_k^{p-2} \int_{\Sigma} |H_{f_k}|^p \, d\mu_{f_k} \leq C r_k^{p-2} \to 0.$$

The prescribed mean curvature system (3.2) for the u_k fulfills the assumption of Lemma 4.5 if $\varepsilon = \varepsilon(p) > 0$ is sufficiently small. Therefore we get the L^p estimate

$$\|D^2\tilde{u}_{k,q}\|_{L^p(B_{1/2})} \leq C(\|H_{\tilde{f}_k}\|_{L^p(B_1)} + \|D\tilde{u}_{k,q}\|_{L^p(B_1)}) \leq C.$$

Moreover the monotonicity formula (4.3) yields, for $B_R = B_R(0) \subset \mathbb{R}^n$,

$$\mu_{\tilde{f}_k}(B_R) \leq C R^2 \quad \text{for any } R \in (0, \infty).$$

We now apply a localized version of Langer's compactness theorem. This is proved in [8, Thm. 1.3] assuming local curvature bounds; the necessary modifications for the case of L^p bounds are known from the compact case—see [16] or [8, Thm. 1.1]. We thus obtain a proper immersion $f_0 : \Sigma_0 \to \mathbb{R}^n$ such that (up to the choice of a subsequence) the $\tilde{f}_k$ converge to f_0 locally in $C^{1,\beta}$, for every $0 < \beta < 1 - 2/p$, up to diffeomorphisms. Weak lower semicontinuity of $\mathcal{W}^p$ implies

$$H_{f_0} = 0.$$

The Gauß–Bonnet theorem yields

$$\int_\Sigma |A_{f_k}|^2 \, d\mu_{f_k} = 4\mathcal{W}(f_k) - 4\pi\chi(\Sigma) \le C,$$

and thus we get further

$$\int_{\Sigma_0} |A_{f_0}|^2 \, d\mu_{f_0} \le \liminf_{k\to\infty} \int_\Sigma |A_{\tilde{f}_k}|^2 \, d\mu_{\tilde{f}_k} \le C.$$

Summarizing, $f_0 : \Sigma_0 \to \mathbb{R}^n$ is a properly immersed minimal surface with finite total curvature. By results of Chern & Osserman [9], f_0 admits a conformal reparametrization on a compact surface with finitely many punctures, corresponding to the ends. Moreover, each end has a well-defined tangent plane and multiplicity. Now the monotonicity formula from (4.4) implies

$$\frac{\mu_{\tilde{f}_k}(B_R)}{\pi R^2} = \frac{\mu_{f_k}(B_{r_k R})}{\pi (Rr_k)^2} \le \frac{1}{4\pi}\mathcal{W}(f_k) + C(Rr_k)^{(p-2)/p} \quad \text{for all } R > 0.$$

Letting $k \to \infty$ and then $R \to \infty$ we conclude that

$$\limsup_{R\to\infty} \frac{\mu_{f_0}(B_R)}{\pi R^2} < 2. \tag{4.7}$$

This means that f_0 has just one simple end, and is in fact a plane. We now argue that the Gauß map converges to a constant locally uniformly on $\Sigma_0 = \mathbb{R}^2$, contradicting the definition of r_k. More precisely, from the compactness theorem in [8] we know that $\tilde{f}_k \circ \phi_k \to f_0$ locally in C^1 and moreover

$$\|\tilde{f}_k \circ \phi_k - f_0\|_{C^0(U_k)} \to 0,$$

where the $U_k \subset \mathbb{R}^2$ are open sets with $U_1 \subset U_2 \subset \cdots$ and $\mathbb{R}^2 = \bigcup_{k=1}^\infty U_k$ such that $\phi_k : U_k \to \tilde{f}_k^{-1}(B_k(0))$ is diffeomorphic. Now $\tilde{f}_k(q_k) = 0$ by construction, therefore there exists a $p_k \in U_k$ with $\phi_k(p_k) = q_k$. In particular

$$f_0(p_k) = f_0(p_k) - (\tilde{f}_k \circ \phi_k)(p_k) \to 0.$$

Since f_0 is proper, we get $p_k \to p \in \mathbb{R}^2$ after passing to a subsequence. Now by the indirect assumption, there exist $x_k \in B_1(0)$ such that for all k,

$$|D\tilde{u}_{k,q_k}(x_k) - D\tilde{u}_{k,q_k}(0)| \ge \varepsilon/2 > 0.$$

Denote the corresponding point by $q_k' \in \Sigma$. Then $|\tilde{f}_k(q_k')| \le C$, and hence there are points $p_k' \in U_k$ (for k large enough) with $\phi_k(p_k') = q_k'$. This implies

$$|f_0(p_k')| \le |f_0(p_k') - (\tilde{f}_k \circ \phi_k)(p_k')| + |\tilde{f}_k(q_k')| \le C.$$

Using again the fact that f_0 is proper, we conclude after passing to a further subsequence that $p_k' \to p' \in \mathbb{R}^2$. But now $T_p f_0 = \lim_{k\to\infty} T_{p_k}(\tilde{f}_k \circ \phi_k) = \lim_{k\to\infty} T_{q_k}\tilde{f}_k$, and

analogously $T_{p'} f_0 = \lim_{k\to\infty} T_{q'_k} \tilde{f}_k$. From the indirect assumption, we obtain $T_p f_0 \neq T_{p'} f_0$, contradicting the fact that f_0 parametrizes a plane.

Given (4.6) we may finally use Lemma 4.5 with $\varrho := \inf_{k\in\mathbb{N}} r_k > 0$ to get

$$\int_{B_{\varrho/2}} |D^2 u_{k,q}|^p \leq C \quad \text{for all } q \in \Sigma, \, k \in \mathbb{N}.$$

The global area bound and a standard covering argument then imply that

$$\mathcal{E}^p(f_k) \leq C \quad \text{for all } k \in \mathbb{N}.$$

The desired conclusion now follows from Theorem 4.1. $\qquad\square$

4.2. Existence of minimizers

Combining Theorem 4.1 with our regularity from Theorem 1.1 we immediately get

Theorem 4.6. *For a closed surface Σ and $p > 2$, denote by $\alpha_\Sigma^n(p)$ the infimum of the energy $\mathcal{E}^p$ among all smooth immersions from Σ into $\mathbb{R}^n$. Then $\alpha_\Sigma^n(p)$ is attained by a smooth immersion $f : \Sigma \to \mathbb{R}^n$.*

Proof. Using mollification it is easy to see that

$$\alpha_\Sigma^n(p) = \inf\{\mathcal{E}^p(f) : f \in W_{\text{im}}^{2,p}(\Sigma, \mathbb{R}^n)\}.$$

Thus the limiting map $f \in W_{\text{im}}^{2,p}(\Sigma, \mathbb{R}^n)$ of a minimizing sequence obtained from Theorem 4.1 is a critical point of $\mathcal{E}^p$, and hence smooth after composing with a diffeomorphism by Theorem 1.1. $\qquad\square$

For any fixed immersion $f : \Sigma \to \mathbb{R}^n$ we have

$$\lim_{q\to p} \alpha_\Sigma^n(q) \leq \lim_{q\to p} \mathcal{E}^q(f) = \mathcal{E}^p(f).$$

Taking the infimum with respect to f shows that the function $\alpha_\Sigma^n : [2, \infty) \to \mathbb{R}$ is upper semicontinuous. In particular, it is right continuous since it is non-decreasing. For $\lambda > 0$ and $f : \Sigma \to \mathbb{R}^n$ fixed we also note that

$$\alpha_\Sigma^n(2) \leq \mathcal{E}^2(\lambda f) = \frac{\lambda^2}{4} \mu_g(\Sigma) + \frac{1}{4} \int_\Sigma |A|^2 \, d\mu_g.$$

Letting $\lambda \searrow 0$ and taking the infimum with respect to f shows that

$$\alpha_\Sigma^n(2) = \inf_{f:\Sigma\to\mathbb{R}^n} \frac{1}{4} \int_\Sigma |A|^2 \, d\mu_g. \tag{4.8}$$

Recall that by the Gauß equation and the Gauß–Bonnet theorem

$$\mathcal{W}(f) = \frac{1}{4} \int_\Sigma |A|^2 \, d\mu_g + \pi \chi(\Sigma). \tag{4.9}$$

The infimum of the Willmore energy among immersions of Σ into $\mathbb{R}^n$ satisfies $\beta_\Sigma^n < 8\pi$ (see [2]). Thus for $p > 2$ close to 2, we conclude for a minimizer f of $\mathcal{E}^p$ that

$$\mathcal{W}(f) \leq \mathcal{E}^p(f) + \pi\chi(\Sigma) = \alpha_\Sigma^n(p) + \pi\chi(\Sigma) < 8\pi.$$

In particular, these minimizers are embedded by the Li–Yau inequality [17].

Next we define the number $\beta_\Sigma^n(p)$ as the infimum of the energy $\mathcal{W}^p$ among all smooth immersions from Σ into $\mathbb{R}^n$. Repeating the previous discussion with β_Σ^n instead of α_Σ^n we conclude that for every sequence of immersions $f_k : \Sigma \to \mathbb{R}^n$ with $\mathcal{W}^p(f_k) \to \beta_\Sigma^n$ and $p - 2$ small enough we have

$$\mathcal{W}(f_k) \leq \mathcal{W}^p(f_k) \to \beta_\Sigma^n(p) < 8\pi.$$

Combining this estimate with Theorems 1.2 and 3.1, and arguing as in the proof of Theorem 4.6, we get

Theorem 4.7. *For every closed surface Σ there exists a number $2 < p_0 < \infty$ such that for every $2 < p < p_0$ the number $\beta_\Sigma^n(p)$ is attained by a smooth immersion $f : \Sigma \to \mathbb{R}^n$.*

The numbers $\alpha_\Sigma^n(p)$ and $\beta_\Sigma^n(p)$ depend only on the topological type of Σ. This can be refined by minimizing in regular homotopy classes of immersions $f : \Sigma \to \mathbb{R}^n$. Theorems 4.6 and 4.7 extend without any difficulties.

5. Palais–Smale condition

Here we show that for $p > 2$ the functionals $\mathcal{E}^p$ resp. $\mathcal{W}^p$ satisfy the Palais–Smale condition resp. a modified Palais–Smale condition, up to the action of diffeomorphisms on Σ. For $f \in W_{\mathrm{im}}^{2,p}(\Sigma, \mathbb{R}^n)$ and any $V \in W^{2,p}(\Sigma, \mathbb{R}^n)$ we define the norm

$$\|V\|_{W_f^{2,p}(\Sigma)} = \left(\int_\Sigma \left(|\nabla(DV)|_g^p + |DV|_g^p + |V|^p \right) d\mu_g \right)^{1/p},$$

where $g \in W^{1,p}(T^{0,2}\Sigma)$ is the metric induced by f and ∇ denotes its Levi-Civita connection, with Christoffel symbols locally in L^p. In particular, the norm is well-defined. Now set

$$\|D\mathcal{E}^p(f)\|_f = \sup\{D\mathcal{E}^p(f)V : V \in W^{2,p}(\Sigma, \mathbb{R}^n),\ \|V\|_{W_f^{2,p}(\Sigma)} \leq 1\},$$

resp.

$$\|D\mathcal{W}^p(f)\|_f = \sup\{D\mathcal{W}^p(f)V : V \in W^{2,p}(\Sigma, \mathbb{R}^n),\ \|V\|_{W_f^{2,p}(\Sigma)} \leq 1\},$$

For any diffeomorphism $\varphi \in W^{2,p}(\Sigma, \Sigma)$ we have $(f \circ \varphi)^* g_{\mathrm{euc}} = \varphi^*(f^* g_{\mathrm{euc}})$, which implies $\|V \circ \varphi\|_{W_{f \circ \varphi}^{2,p}(\Sigma)} = \|V\|_{W_f^{2,p}(\Sigma)}$ and therefore

$$\|D\mathcal{E}^p(f \circ \varphi)\|_{f \circ \varphi} = \|D\mathcal{E}^p(f)\|_f, \tag{5.1}$$

resp.

$$\|D\mathcal{W}^p(f \circ \varphi)\|_{f \circ \varphi} = \|D\mathcal{W}^p(f)\|_f. \tag{5.2}$$

Now we can formulate the main results of this section.

Theorem 5.1. *Let* $f_k \in W_{\mathrm{im}}^{2,p}(\Sigma, \mathbb{R}^n)$, $p > 2$, *be a sequence satisfying*

$$\mathcal{E}^p(f_k) \leq C \quad and \quad \|D\mathcal{E}^p(f_k)\|_{f_k} \to 0.$$

Then, after choosing a subsequence and passing to $f_k \circ \varphi_k$ *for suitable diffeomorphisms* $\varphi_k \in C^\infty(\Sigma, \Sigma)$, *the* f_k *converge strongly in* $W^{2,p}(\Sigma, \mathbb{R}^n)$ *to some* $f \in W_{\mathrm{im}}^{2,p}(\Sigma, \mathbb{R}^n)$, *and* f *is a smooth critical point of* $\mathcal{E}^p$.

Theorem 5.2. *Let* $f_k \in W_{\mathrm{im}}^{2,p}(\Sigma, \mathbb{R}^n)$, $\delta > 0$, $p > 2$, *be a sequence satisfying*

$$\mathcal{W}^p(f_k) \leq C, \quad \mathcal{W}(f_k) \leq 8\pi - \delta \quad and \quad \|D\mathcal{W}^p(f_k)\|_{f_k} \to 0.$$

Then, after choosing a subsequence and passing to $f_k \circ \varphi_k$ *for suitable diffeomorphisms* $\varphi_k \in C^\infty(\Sigma, \Sigma)$, *the* f_k *converge strongly in* $W^{2,p}(\Sigma, \mathbb{R}^n)$ *to some* $f \in W_{\mathrm{im}}^{2,p}(\Sigma, \mathbb{R}^n)$, *and* f *is a smooth critical point of* $\mathcal{W}^p$.

We recall that the Palais–Smale condition cannot hold for the Willmore functional itsself due to Möbius invariance. For sequences of critical points, i.e. Willmore surfaces, a concentration-compactness alternative was proved by Kuwert & Schätzle [15], and a full description of the bubbling was given by Bernard & Rivière [5] when the conformal type is non-degenerating. They also proved that weak limits of Palais–Smale sequences are conformal Willmore [4].

Since the arguments for the two results are very similar (thanks to Theorems 4.1 and 1.2) we only present the proof of Theorem 5.1.

Proof of Theorem 5.1. Langer's compactness theorem [16] implies that after passing to a subsequence $f_k \circ \varphi_k \to f$ in the C^1 topology and weakly in $W^{2,p}(\Sigma, \mathbb{R}^n)$, where $f \in W_{\mathrm{im}}^{2,p}(\Sigma, \mathbb{R}^n)$ and $\varphi_k \in C^\infty(\Sigma, \Sigma)$ are diffeomorphisms. It remains to see that the convergence is strong in $W^{2,p}(\Sigma, \mathbb{R}^n)$, for which it suffices to consider the local convergence of the graph representations over a disk $B_r \subset \mathbb{R}^2$. Namely, the assumption then implies that f is a critical point of $\mathcal{E}^p$ and is hence smooth by Theorem 1.1, after composing with a further diffeomorphism.

Let u_k, $u \in W^{2,p}(B_r, \mathbb{R}^{n-2})$ be the graph functions for f_k and f, respectively. Then $u_k \to u$ in $C^1(B_r)$, weakly in $W^{2,p}(B_r)$, and we can assume

$$\|Du_k\|_{C^0(B_r)} \leq L \leq 1 \quad and \quad \|u_k\|_{W^{2,p}(B_r)}^p \leq C \, \mathcal{E}^p(f_k) \leq C. \tag{5.3}$$

We let $\psi_k = \eta(u_k - u)$ where $\chi_{B_{r/2}} \leq \eta \leq \chi_{B_r}$ is a cut-off function. Clearly

$$\|\psi_k\|_{C^1(B_r)} \to 0 \quad and \quad \|\psi_k\|_{W^{2,p}(B_r)} \leq C. \tag{5.4}$$

Next we recall from (2.1) the Fréchet derivative, in a graph representation:

$$D\mathcal{E}^p(f_k)(0, \psi_k) = \int_{B_r} \left(a_i^{\alpha\beta}(Du_k, D^2u_k) \partial_{\alpha\beta}^2 \psi_k^i + b_i^\alpha(Du_k, D^2u_k) \partial_\alpha \psi_k^i \right), \qquad (5.5)$$

where

$$a_i^{\alpha\beta}(Du_k, D^2u_k) = \frac{p}{4}(1 + |A_k|^2)^{(p-2)/2} B_{ij}^{\alpha\beta,\gamma\lambda}(Du_k) \partial_{\gamma\lambda}^2 u_k^j \sqrt{\det g_k},$$

$$b_i^\alpha(Du_k, D^2u_k) = \frac{p}{8}(1 + |A_k|^2)^{(p-2)/2} \frac{\partial B_{jm}^{\gamma\lambda,\mu\nu}}{\partial p_\alpha^i}(Du_k) \partial_{\gamma\lambda}^2 u_k^j \partial_{\mu\nu} u_k^m \sqrt{\det g_k}$$
$$+ \frac{1}{4}(1 + |A_k|^2)^{p/2} \frac{\partial \sqrt{\det g_k}}{\partial p_\alpha^i}(Du_k).$$

Here $(g_k)_{\alpha\beta} = \delta_{\alpha\beta} + \langle \partial_\alpha u_k, \partial_\beta u_k \rangle$ and $B_{ij}^{\alpha\beta,\gamma\lambda}(Du_k) = g_k^{\alpha\gamma} g_k^{\beta\lambda}(\delta_{ij} - g_k^{\mu\nu} \partial_\mu u_k^i \partial_\nu u_k^j)$. We see easily that

$$a_i^{\alpha\beta}(Du_k, D^2u_k) \le C(1 + |D^2u_k|^2)^{(p-1)/2},$$
$$b_i^\alpha(Du_k, D^2u_k) \le CL(1 + |D^2u_k|^2)^{p/2},$$

and obtain, as $k \to \infty$,

$$\int_{B_r} a_i^{\alpha\beta}(Du_k, D^2u_k)\left(\partial_{\alpha\beta}^2 \psi_k^i - \eta \partial_{\alpha\beta}^2(u_k^i - u^i)\right) \to 0, \qquad (5.6)$$

$$\int_{B_r} b_i^\alpha(Du_k, D^2u_k) \partial_\alpha \psi_k^i \to 0. \qquad (5.7)$$

Now using (5.3) and (5.4) we get $\|(0, \psi_k)\|_{W_{f_k}^{2,p}(\Sigma)} \le C\|\psi_k\|_{W^{2,p}} \le C$, and hence

$$D\mathcal{E}^p(f_k)(0, \psi_k) \to 0 \quad \text{as } k \to \infty,$$

using the assumption of the theorem and (5.1). Combining this with (5.6) and (5.7), and noting that $a_i^{\alpha\beta}(Du, D^2u) \in L^p(B_r, \mathbb{R}^{n-2})'$, we conclude that

$$\int_{B_r} \eta\left(a_i^{\alpha\beta}(Du_k, D^2u_k) - a_i^{\alpha\beta}(Du, D^2u)\right) \partial_{\alpha\beta}^2(u_k^i - u^i) \to 0 \quad \text{as } k \to \infty.$$

But since $u_k \to u$ in $C^1(B_r, \mathbb{R}^{n-2})$ we also have

$$\int_{B_r} \eta\left(a_i^{\alpha\beta}(Du_k, D^2u_k) - a_i^{\alpha\beta}(Du, D^2u_k)\right) \partial_{\alpha\beta}^2(u_k^i - u^i) \to 0,$$

and by adding the last two equations we get

$$\int_{B_r} \eta\left(a_i^{\alpha\beta}(Du, D^2u_k) - a_i^{\alpha\beta}(Du, D^2u)\right) \partial_{\alpha\beta}^2(u_k^i - u^i) \to 0 \quad \text{as } k \to \infty.$$

Finally, we use the ellipticity (see (3.12)) to estimate

$$\int_{B_r} \eta \big(a_i^{\alpha\beta}(Du, D^2 u_k) - a_i^{\alpha\beta}(Du, D^2 u)\big) \partial_{\alpha\beta}^2 (u_k^i - u^i)$$

$$= \int_{B_r} \eta \int_0^1 \frac{\partial a_i^{\alpha\beta}}{\partial q_{\lambda\mu}^j}\big(Du, D^2 u + t D^2(u_k - u)\big) \partial_{\lambda\mu}^2 (u_k^j - u^j) \partial_{\alpha\beta}^2 (u_k^i - u^i)\, dt$$

$$\geq \lambda \int_{B_r} \eta \int_0^1 \big(1 + |D^2 u + t D^2(u_k - u)|^2\big)^{(p-2)/2} |D^2(u_k - u)|^2\, dt$$

$$\geq c\lambda \int_{B_{r/2}} |D^2(u_k - u)|^p.$$

In the last step we have used the elementary Lemma 19.27 from [20]. Altogether we have proved local and hence global convergence in $W^{2,p}$. $\qquad\qquad\qquad\square$

References

[1] Ambrosio, L., Masnou, S.: A direct variational approach to a problem arising in image reconstruction. Interfaces Free Bound. **5**, 63–81 (2003)   Zbl 1029.49037   MR 1959769

[2] Bauer, M., Kuwert, E.: Existence of minimizing Willmore surfaces of prescribed genus. Int. Math. Res. Notices **2003**, 553–576   Zbl 1029.53073   MR 1941840

[3] Bellettini, G., Dal Maso, G., Paolini, M.: Semicontinuity and relaxation properties of a curvature depending functional in 2D. Ann. Scuola Norm. Sup. Pisa Cl. Sci. **20**, 247–297 (1993)   Zbl 0797.49013   MR 1233638

[4] Bernard, Y., Rivière, T.: Local Palais–Smale sequences for the Willmore functional. Comm. Anal. Geom. **19**, 563–599 (2011)   Zbl 1245.53051   MR 2843242

[5] Bernard, Y., Rivière, T.: Energy quantization for Willmore surfaces and applications. Ann. of Math. **180**, 87–136 (2014)   Zbl 06316067   MR 3194812

[6] Bildhauer, M., Fuchs, M.: Higher order variational problems on two-dimensional domains. Ann. Acad. Sci. Fenn. Math. **31**, 349–362 (2006)   Zbl 1136.49027   MR 2248820

[7] Bildhauer, M., Fuchs, M., Zhong, X.: A lemma on the higher integrabilty of functions with applications to the regularity theory of two-dimensional generalized Newtonian fluids. Manuscripta Math. **116**, 135–156 (2005)   Zbl 1116.49018   MR 2122416

[8] Breuning, P.: Immersions with bounded second fundamental form. J. Geom. Anal. **25**, 1344–1386 (2015)   Zbl 1318.53060   MR 3319975

[9] Chern, S. S., Osserman, R.: Complete minimal surfaces in Euclidean n-space. J. Anal. Math. **19**, 15–34 (1967)   Zbl 0172.22802   MR 0226514

[10] Edmunds, D., Gurka, P., Opic, B.: On embeddings of logarithmic Bessel potential spaces. J. Funct. Anal. **146**, 116–150 (1997)   Zbl 0934.46036   MR 1446377

[11] Faraco, D., Koskela, P., Zhong, X.: Mappings of finite distortion: the degree of regularity. Adv. Math. **190**, 300–318 (2005)   Zbl 1075.30012   MR 2102659

[12] Kauhanen, J., Koskela, P., Maly, J.: On functions with derivatives in a Lorentz space. Manuscripta Math. **100**, 87–101 (1999) Zbl 0976.26004 MR 1714456

[13] Kühnel, W., Pinkall, U.: On total mean curvatures. Quart. J. Math. Oxford Ser. (2) **37**, 437–447 (1986) Zbl 0627.53044 MR 0868618

[14] Kuwert, E., Li, Y., Schätzle, R.: The large genus limit of the infimum of the Willmore energy. Amer. J. Math. **132**, 37–52 (2010) Zbl 1188.53057 MR 2597505

[15] Kuwert, E., Schätzle, R.: The Willmore flow with small initial energy. J. Differential Geom. **57**, 409–441 (2001) Zbl 1035.53092 MR 1882663

[16] Langer, J.: A compactness theorem for surfaces with L^p-bounded second fundamental form. Math. Ann. **270**, 223–234 (1985) Zbl 0564.58010 MR 0771980

[17] Li, P., Yau, S. T.: A new conformal invariant and its applications to the Willmore conjecture and the first eigenvalue on compact surfaces. Invent. Math. **69**, 269–291 (1982) Zbl 0503.53042 MR 0674407

[18] Mondino, A.: Existence of integral m-varifolds minimizing $\int |A|^p$ and $\int |H|^p$, $p > m$, in Riemannian manifolds. Calc. Var. **49**, 431–470 (2014) Zbl 1282.49041 MR 3148123

[19] Morrey, C. B.: Multiple Integrals in the Calculus of Variations. Grundlehren Math. Wiss. 130, Springer, New York (1966) Zbl 0142.38701 MR 0202511

[20] Palais, R.: Foundations of Global Non-Linear Analysis. W. A. Benjamin, New York (1968) Zbl 0164.11102 MR 0248880

[21] Sacks, J., Uhlenbeck, K.: The existence of minimal immersions of 2-spheres. Ann. of Math. **113**, 1–24 (1981) Zbl 0462.58014 MR 0604040

[22] Simon, L.: Existence of surfaces minimizing the Willmore functional. Comm. Anal. Geom. **1**, 281–326 (1993) Zbl 0848.58012 MR 1243525

[23] Trudinger, N.: On imbeddings into Orlicz spaces and some applications. J. Math. Mech. **17**, 473–483 (1967) Zbl 0163.36402 MR 0216286