Two-dimensional curvature functionals with superquadratic growth

Abstract. For two-dimensional, immersed closed surfaces \( f : \Sigma \rightarrow \mathbb{R}^n \), we study the curvature functionals \( E_p(f) \) and \( W_p(f) \) 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 \( W^p \)-bounded sequences. In the case of \( E^p \) this is just Langer’s theorem [16], while for \( W^p \) we have to impose a bound for the Willmore energy strictly below \( 8\pi \) 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 \( \Sigma \) 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 \( W^{2,p}_\text{im}(\Sigma, \mathbb{R}^n) \) of immersions we consider the two functionals

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

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

**Theorem 1.1.** Let \( f \in W^{2,p}_\text{im}(\Sigma, \mathbb{R}^n) \) be a critical point of \( W^p \) or \( E^p \), where \( 2 < p < \infty \). Then local graph representations of \( f \) are smooth.
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 \(W^p(f)\), 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 \(L_g \phi = |H|^{\lambda - 1}H\) for appropriate \(\lambda > 1\) and then insert \(\phi\) as a test function. Here the operator \(L_g = \sqrt{\det g} g^{\alpha \beta} \partial_{\alpha \beta}\) comes up in the principal term of the equation.

Unfortunately, the same strategy does not apply in the case of the functional \(E^p(f)\), 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 \(L^2\) 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 \(E^p(f)\) approach to cover the functional \(W^p(f)\), 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 \(E^p(f_k) \leq C\) subconverge weakly to an \(f \in W^{2,p}_{\text{im}}(\Sigma, \mathbb{R}^n)\), after suitable translation and reparametrization. In particular, we obtain the existence of a smooth \(E^p\) minimizer in the class of immersions \(f : \Sigma \to \mathbb{R}^n\) for \(p > 2\). On the other hand, boundedness of \(W^p(f)\) 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 \(8\pi\) bound in the following result is optimal.

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

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

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 \(W^{2,p}_{\text{im}}(\Sigma, \mathbb{R}^k)\) to an \(f \in W^{2,p}_{\text{im}}(\Sigma, \mathbb{R}^n)\). In particular, the convergence is in \(C^{1,\beta}(\Sigma, \mathbb{R}^n)\) for any \(\beta < 1 - 2/p\) and we have

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

We remark that Mondino [18] recently proved the existence and partial regularity of varifolds minimizing functionals which satisfy similar growth conditions as \(E^p\) and \(W^p\) 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
Two-dimensional curvature functionals

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

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

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 $E^p$ and $W^p$ 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 $E^p(f)$ and $W^p(f)$. For an immersed surface the fundamental forms are

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

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

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

We compute further

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

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

$$D E^p(f) \phi = \frac{p}{4} \int_{\Sigma} (1 + |A|^2)^{(p-2)/2} g^{\alpha\gamma} g^{\beta\lambda} \langle P^\perp \partial_\alpha \partial_\beta f, \partial_\gamma \phi \rangle \sqrt{\det g}$$

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

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

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 $E^p$ if and only if $u$ is a weak solution of the system

$$\partial_{\alpha\beta}(a^\alpha_{i\beta}(Du, D^2u)) + \partial_\alpha (b^\alpha_i(Du, D^2u)) = 0 \quad (1 \leq i \leq n-2), \quad (2.1)$$
where the coefficients are given by
\[
a_{i}^{a\beta}(Du, D^{2}u) = (1 + |A|^{2})^{(p-2)/2} \sqrt{\det g} g^{a\gamma} g^{\beta\lambda} (\delta_{ij} - g^{\mu\nu} \partial_{\mu} u^{i} \partial_{\nu} u^{j}) \partial_{\gamma \lambda} \partial_{i j} u^{k}.
\]

\[
b_{i}^{\alpha}(Du, D^{2}u) = -\frac{1}{2} (1 + |A|^{2})^{(p-2)/2} \frac{\partial (g^{a\gamma} g^{\beta\lambda} (\delta_{ij} - g^{\mu\nu} \partial_{\mu} u^{i} \partial_{\nu} u^{j}))}{\partial p_{a}} \partial_{\gamma \lambda} \partial_{i j} u^{k} \sqrt{\det g}.
\]

For \(|p| \leq \Lambda\) and \(V = (1 + |q|^{2})^{1/2}\), where \(p, q\) are the variables corresponding to \(Du, D^{2}u\), one easily checks the bounds
\[
|D_{q}a| \leq C(\Lambda) V^{p-2},
\]
\[
|a| + |D_{p}a| + |D_{q}b| \leq C(\Lambda) V^{p-1},
\]
\[
|b| + |D_{p}b| \leq C(\Lambda) V^{p}.
\]

Moreover, the system satisfies the ellipticity condition
\[
\frac{\partial a_{i}^{\alpha\beta}}{\partial q_{i}^{\gamma \lambda}} \geq \lambda V^{p-2} |\xi|^{2} \quad \text{where} \quad \lambda = \lambda(\Lambda) > 0,
\]

For the first variation of \(W^{p}(f)\) one obtains
\[
D W^{p}(f) \phi = \frac{p}{4} \int_{\Omega} (1 + |H|^{2})^{(p-2)/2} (H, g^{\gamma \lambda} \partial_{\gamma \lambda} \phi) \sqrt{\det g} + \frac{p}{8} \int_{\Omega} (1 + |H|^{2})^{(p-2)/2} \left( \frac{\partial (g^{a\gamma} g^{\beta\lambda} P_{\perp})}{\partial p_{a}} \partial_{\mu} \phi \partial_{\gamma \lambda} f, \partial_{\gamma \lambda} f \right) \sqrt{\det g} + \frac{1}{4} \int_{\Omega} (1 + |H|^{2})^{p/2} \frac{\partial \sqrt{\det g}}{\partial p_{a}} \partial_{\mu} \phi f.
\]

If we set \(L_{g} \phi = \sqrt{\det g} g^{\gamma \lambda} \partial_{\gamma \lambda} \phi\), the first variation takes the form
\[
D W^{p}(f) \phi = \frac{p}{4} \int_{\Omega} (1 + |H|^{2})^{(p-2)/2} (H, L_{g} \phi) + \int_{\Omega} B_{i}^{\alpha}(Df, D^{2}f) \partial_{\alpha} \phi^{i}, \quad (2.2)
\]

where
\[
B_{i}^{\alpha}(Df, D^{2}f) = \frac{p}{8} (1 + |H|^{2})^{(p-2)/2} \left( \frac{\partial (g^{a\gamma} g^{\mu\nu} P_{\perp})}{\partial p_{a}} \partial_{\gamma \lambda} f, \partial_{\gamma \lambda} f \right) \sqrt{\det g} + \frac{1}{4} (1 + |H|^{2})^{p/2} \frac{\partial \sqrt{\det g}}{\partial p_{a}} .
\]

When passing to graphs we have, under the assumption \(|p| \leq \Lambda\),
\[
|B| + |D_{p}B| \leq C(\Lambda) V^{p-2} |q|^{2} \quad \text{and} \quad |D_{q}B| \leq C(\Lambda) V^{p-2} |q|, \quad (2.3)
\]

\[
D W^{p}(f) \phi = \frac{p}{4} \int_{\Omega} (1 + |H|^{2})^{(p-2)/2} (H, L_{g} \phi) + \int_{\Omega} B_{i}^{\alpha}(Df, D^{2}f) \partial_{\alpha} \phi^{i}.
\]
3. Regularity of critical points

3.1. The functional \( W^p \)

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 \( W^p \) if and only if

\[
\int_{\Omega} \langle H, Lg \varphi \rangle + \int_{\Omega} B^i_{\alpha}(Du, D^2 u)\partial_{\alpha} \varphi^i = 0 \quad \text{for all } \varphi \in W^{2,p}_0(\Omega, \mathbb{R}^n). \tag{3.1}
\]

Here \( H = (1 + |H|^2)^{p/2-1}H \) and the functions \( B^i_{\alpha} \) 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

\[
g^{\alpha\beta}(\delta_{ij} - g^{\lambda\mu} \partial_\lambda u^i \partial_\mu u^j)\partial^2_{\alpha\beta} u^i = H^j + 2. \tag{3.2}
\]

Since \( p > 2 \) the left hand side may be viewed as a linear operator of the form \( a^{\alpha\beta}_{ij} \partial^2_{\alpha\beta} u^i \), 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 \( u \in W^{2,q}_{loc}(\Omega, \mathbb{R}^{n-2}) \) together with a local estimate

\[
\|u\|_{W^{2,q}_0(\Omega)} \leq C(p, q, \Lambda) (\|H\|_{L^q(\Omega)} + 1) \quad \text{if } \|u\|_{W^{2,p}(\Omega)} \leq \Lambda. \tag{3.3}
\]

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

\[
\int_{\Omega} \eta \langle H, Lg \varphi \rangle \leq C \int_{\Omega} (1 + |D^2 u|^2)^{p/2} (|\varphi| + |D\varphi|), \quad \text{where } C = C(\|\eta\|_{C^2}).
\]

**Proof.** Expanding

\[
Lg(\eta \varphi) = \eta Lg \varphi + \varphi Lg \eta + 2\sqrt{\det g} g^{\alpha\beta} \partial_\alpha \eta \partial_\beta \varphi,
\]

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

\[
\int_{\Omega} \eta \langle H, Lg \varphi \rangle \leq C \int_{\Omega} (1 + |D^2 u|^2)^{p/2} (|\varphi| + |\eta| |D\varphi|) + C \int_{\Omega} (1 + |D^2 u|^2)^{(p-1)/2} (|D\eta||\varphi| + |D\eta||D\varphi|).
\]

This implies the lemma. \( \Box \)

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 \( u \in W^{2,q}_{\text{loc}}(\Omega, \mathbb{R}^{n-2}) \) 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 \( \psi \in W^{2,q}_{\text{loc}}(B_r, \mathbb{R}^n) \) of the linear equation

\[ L^q \psi = |H|^{\lambda-1}_A H. \]

As \( 1 < q/\lambda < \infty \) the function \( \psi \) satisfies

\[ \|\psi\|_{W^{2,q/\lambda}(B_r)} \leq C \|H|^{\lambda-1}_A H\|_{L^{q/\lambda}(B_r)} \leq C; \]

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

\[ \|\psi\|_{C^1(B_r)} \leq C, \]

while for \( q/2 < \lambda < q \) we get instead

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

Now Lemma 3.2 implies that

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

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

\[ \frac{p}{q} + \frac{1}{s} \leq 1 \Leftrightarrow \lambda \leq 3q/2 - p. \]

Letting \( A \not\to \infty \) we get \( H \in L^{p+\lambda-1}(B_{r/2}) \), and so \( u \in W^{2,p+\lambda-1}_{\text{loc}}(B_{r/4}, \mathbb{R}^{n-2}) \) from (3.3). We can now set up an iteration to get \( u \in W^{2,q}_{\text{loc}}(\Omega) \) 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.** \( W^{1,2}_{\text{loc}}\)-regularity of \( H \). In this subsection we use difference quotient methods in order to show that \( H \in W^{1,2}_{\text{loc}} \). For \( h > 0, f : \Omega \to \mathbb{R}^k \) and fixed \( v \in [1, 2] \) we define

\[ f_h(x) = \frac{1}{h} (f(x + he_v) - f(x)). \]
In the following we denote by $\Phi(\cdot, h)$ an arbitrary measurable function which satisfies, for all $1 \leq q < \infty$ and all $\Omega' \subset \subset \Omega$,

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

**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 $x \in \Omega'$,

$$|(|H|^2)_h(x) + |H_h|(x) \leq \Phi(x, h)|H_h|(x).$$

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

$$|(|H|^2)_h(x) = ((1 + |H|^2)^{p-2}_h(x)|H|^2(x + he_v) + (1 + |H|^2)^{p-2}_h(x)(|H|^2)_h(x)
$$

$$= (|H|^2)_h(x)\left[(p - 2)(1 + \xi(x))^{p-3}|H|^2(x + he_v) + (1 + |H|^2)^{p-2}(x)\right],$$

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

$$|(|H|^2)_h(x) \leq |(|H|^2)_h(x).$$

On the other hand we calculate

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

Combining this with Theorem 3.3 proves the first estimate.

Another application of the mean value theorem yields

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

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

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

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

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

**Proof.** Using (3.2) we see that $u_h$ solves

$$a^\alpha_{ij}(\cdot + he_v)\tilde{\partial}_\alpha a^\beta_{ij} = \tilde{H}^{i+2} - (a^\alpha_{ij})_h\tilde{\partial}_\alpha u^i =: \tilde{H}^i(x),$$

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

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

$$|\tilde{H}(x) \leq \Phi(x, h)(|H_h| + 1).$$
Next we use standard $L^p$ theory in order to get, for every $1 < s < \infty$,
\[
\int_{B_r}|D^2((\eta^2 u_h))|^s \leq C \int_{B_r} \Phi(x, h)(|\mathcal{H}_h|^r + 1).
\]
Since moreover
\[
\int_{B_r} \eta^2 |(D^2 u_h)|^s \leq C \int_{B_r} |D^2((\eta^2 u_h))|^s + C,
\]
this finishes the proof of the corollary. \qed

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

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

**Proof.** Taking difference quotients of equation (3.1) we get
\[
\partial_2^\alpha \beta (\sqrt{g} \, g^{\alpha \beta} H_h) - \partial_\alpha (\tilde{B}^a(Du, D^2 u)_h) = 0. \tag{3.6}
\]
We abbreviate $U(x) = (Du(x), D^2 u(x))$ and we use the fundamental theorem of calculus to write
\[
(f \circ U)_h(x) = \frac{1}{h} \left( f(U(x+he_v)) - f(U(x)) \right)
= \frac{1}{h} \int_0^1 \frac{d}{dt} f((1-t)U(x) + tU(x+he_v)) \, dt
= \int_0^1 Df((1-t)U(x) + tU(x+he_v)) \, dt \cdot U_h(x).
\]
Using the notation $f^h(x) = \frac{1}{h} \int_0^1 f((1-t)U(x) + tU(x+he_v)) \, dt$ we thus get
\[
(f \circ U)_h = \left( \frac{\partial f}{\partial \eta_{h}^j} \right)^h \partial_2^\alpha \beta u_h^j + \left( \frac{\partial f}{\partial p_j} \right)^h \partial_\alpha u_h^j,
\]
and the system (3.6) takes the form
\[
\partial_2^\alpha \beta ((\sqrt{g} \, g^{\alpha \beta})_h \mathcal{H}_h) + \partial_\alpha (\tilde{B}^a(Du, D^2 u)_h) - \partial_\alpha (\tilde{B}^a(Du, D^2 u)_h) = 0. \tag{3.7}
\]
Here the coefficients are as follows:
\[
\tilde{B}^a(x, z, p, q) = \left( \frac{\partial B^a}{\partial \eta_{h}^j} \right)^h (x) q_j + \left( \frac{\partial B^a}{\partial p_j} \right)^h (x) p_j.
\]
Using (2.3) we then obtain
\[
|\tilde{B}(Du_h, D^2 u_h)| \leq C \Phi(x, h)(|D^2 u_h| + 1). \tag{3.8}
\]
For \( \|Du\|_\infty < \infty \) we get

\[
|\sqrt{g}^{-1}h_{\mathcal{H}}(x)| \leq C(1 + |D^2u|^{p-1/2}(x)|Dh|) = \Phi(x, h),
\]

and moreover the operator

\[
\tilde{L}v(x) := (\sqrt{g} g^{a\beta})(x + hev) \partial^2_{a\beta}v(x)
\]

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

\[
\tilde{L}\tilde{\phi} = \mathcal{H}
\]

satisfying

\[
\|\tilde{\phi}\|_{W^{2,q}(B_r)} \leq C\|\mathcal{H}\|_{L^q(B_r)} \leq C.
\]

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

\[
\tilde{L}\phi_h = \eta^4\mathcal{H}_h - (\sqrt{g} g^{a\beta})(x + hev) \partial^2_{a\beta}(\eta^4\tilde{\phi})(x + hev)
\]

\[
+ 4\eta^2(\eta(\sqrt{g} g^{a\beta})(x + hev)\partial_a\eta\partial_\beta\tilde{\phi}_h
\]

\[
+ \eta(\sqrt{g} g^{a\beta})(x + hev)\partial^2_{a\beta}\eta \tilde{\phi}_h + 3(\sqrt{g} g^{a\beta})(x + hev)\partial_a \eta \partial_\beta \eta \tilde{\phi}_h)
\]

\[
= \eta^4\mathcal{H}_h + \eta^2\Phi(\cdot, h)
\]

and, by using standard \( L^2 \) estimates, we conclude

\[
\|\phi_h\|_{W^{2,2}(B_r)} \leq C(\|\eta^2\mathcal{H}_h\|_{L^2(B_r)} + 1).
\]

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

\[
\int_{B_r} \eta^4 |\mathcal{H}_h|^2 \leq C \int_{B_r} \Phi(\cdot, h)(|D^2\phi_h| + \eta^2|\mathcal{H}_h| + |D^2u_h| + 1) \leq \delta \int_{B_r} \eta^4 |\mathcal{H}_h|^2 + C\delta.
\]

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

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

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

\[
u \in W^{3,4}_{loc}(\Omega, \mathbb{R}^{n-2}) \quad \text{for all } 1 < s < 2.
\]

3.1.3 Higher regularity. In this last subsection we show the higher regularity for solutions of (3.1). We start by showing that \( \mathcal{H} \in W^{1,2+\gamma}_{loc}(\Omega, \mathbb{R}^n) \) for some \( 0 < \gamma < 1/2 \). □
In order to see this, we let $B_r \subset \Omega$ and we let $\varphi_1 \in W^{2,2/(1+\gamma)} \cap W^{1,2/(1+\gamma)}_0 \left( B_r, \mathbb{R}^n \right)$ be the solution of
\[
\tilde{L} \varphi_1 = |H_h|^\gamma \mathcal{H}_h
\]
satisfying
\[
\| \varphi_1 \|_{W^{2,2/(1+\gamma)}(B_r)} \leq C \| H_h \|_{L^2(B_r)}^{1+\gamma} \leq C.
\]
Sobolev’s embedding theorem yields
\[
\| \varphi_1 \|_{L^\infty(B_r)} + \| \nabla \varphi_1 \|_{L^{2/\gamma}(B_r)} \leq C.
\]
Next we let $\eta \in C^\infty_c(B_r)$ and we define $\tilde{\varphi}_1 = \eta^4 \varphi_1$. We conclude that
\[
\tilde{L} \tilde{\varphi}_1 = \eta^4 |H_h|^\gamma \mathcal{H}_h + 8 \eta^4 (\sqrt{g} g^{\alpha\beta}) (x + h e_v) \partial_\alpha \eta \partial_\beta \varphi_1 \\
+ 4 \eta^2 \varphi_1 (\sqrt{g} g^{\alpha\beta}) (x + h e_v) (3 \partial_\alpha \eta \partial_\beta \eta + \eta \partial_\alpha^2 \eta) \\
=: \eta^4 |H_h|^\gamma \mathcal{H}_h + \eta^2 L[\varphi_1, D\varphi_1].
\]
Moreover
\[
\| \tilde{\varphi}_1 \|_{L^\infty(B_r)} + \| \nabla \tilde{\varphi}_1 \|_{L^{2/\gamma}(B_r)} + \| \nabla^2 \tilde{\varphi}_1 \|_{L^{2/(1+\gamma)}(B_r)} \leq C.
\]
Now we use $\tilde{\varphi}_1$ as a test function in (3.7) to conclude that
\[
\int_{B_r} \eta^4 |H_h|^{2+\gamma} \\
\leq C \int_{B_r} \left( |(\sqrt{g}^{-1} h_{\alpha\beta})| |\nabla^2 \tilde{\varphi}_1| + |\tilde{L}(D u_h, D^2 u_h)| |\nabla \tilde{\varphi}_1| + \eta^2 |H_h| \| L[\varphi_1, D\varphi_1] \| \right) \\
\leq C \int_{B_r} \Phi(x, h) \left( |\nabla^2 \tilde{\varphi}_1| + |\nabla^2 u_h| + 1 |\nabla \tilde{\varphi}_1| + |H_h| \| (D \tilde{\varphi}_1) + 1 \| \right) \leq C.
\]
Therefore $H \in W^{1,2+\gamma}_{\text{loc}}(\Omega, \mathbb{R}^n)$. In particular this implies that $H \in L^\infty_{\text{loc}}(\Omega, \mathbb{R}^{n-2})$ and hence
\[
|H|^2 \in W^{1,2+\gamma}_{\text{loc}}(\Omega).
\]
Corollary 3.5 yields $u \in W^{3,2+\gamma}_{\text{loc}}(\Omega, \mathbb{R}^{n-2})$. By the Sobolev embedding theorem this gives
\[
u \in C^{2,\beta}_{\text{loc}}(\Omega, \mathbb{R}^{n-2})
\]
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
\[
\partial_{\alpha\beta} \left[ a_{\alpha\beta}^i (\cdot, u, Du, D^2 u) \right] + \partial_\alpha \left[ b_{\alpha}^i (\cdot, u, Du, D^2 u) \right] + c^i (\cdot, u, Du, D^2 u) = 0. \tag{3.11}
\]
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 $\lambda > 0, C < \infty$:

$$\frac{\partial a_{ij}}{\partial q^j} q^i_q q^j_{\alpha\beta} \geq \lambda V^{p-2} |q|^2, \quad (3.12)$$

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

$$|a| + |a_z| + |a_p| + |b_q| + |c_q| \leq CV^{p-1}, \quad (3.13)$$

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

As noted in Section 2, the graph function of a critical point for $E$ 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 $u \in C^{2,\alpha}_{\text{loc}}(\Omega, \mathbb{R}^m)$ 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 $r < r_0$,

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

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

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

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 \( \phi^4(u - l_r) \) as a test function in the weak form of (3.11) to get
\[ \left| \int_\Omega \phi^4 a^{ij}_a \partial_a^2 u^i \right| \leq C \int_\Omega \phi^4 (|b| |D(u - l_r)| + |c| |u - l_r|) \]
\[ + Cr^{-1} \int_{A_r} \phi^3 (|a| |D(u - l_r)| + |b| |u - l_r|) \]
\[ + Cr^{-2} \int_{A_r} \phi^2 |a| |u - l_r| = I + II + III. \]

Using the ellipticity assumption and the bound \( |a(x, z, p, 0)| \leq C \) (which follows from (3.13)) we estimate
\[ q_{ij}^a a^{ij}_a(x, y, z, q) = q_{ij}^a a^{ij}_a(x, y, z, 0) + q_{ij}^a q_{ij}^j \int_0^1 a^{ij}_{a\beta, a\delta} (x, y, z, tq) \, dt \]
\[ \geq q_{ij}^a a^{ij}_a(x, y, z, 0) + \lambda |q|^2 \int_0^1 (1 + t^2 |q|^2)^{(p-2)/2} \, dt \geq \tilde{\lambda} V^p - C \]
where \( \tilde{\lambda} > 0 \) is some number. Next we use (3.13), H"older’s inequality and (3.16) and (3.17) to obtain
\[ I \leq Cr^p \int_\Omega \phi^4 V^p, \]
\[ II \leq Cr^p \int_{A_r} V^p + Cr^{-1} \|V\|_{L^p(A_r)}^{-1} \|D(u - l_r)\|_{L^p(A_r)} \leq C \int_{A_r} V^p, \]
\[ III \leq Cr^{-2} \left( \int_{A_r} V^p \right)^{(p-1)/p} \|u - l_r\|_{L^p(A_r)} \leq C \int_{A_r} V^p. \]

Combining all these estimates and choosing \( r_0 \) small enough we conclude that there exists a constant \( C_1 > 0 \) such that
\[ \int_{B_{r_0}} V^p \leq C_1 \int_{A_r} V^p + Cr^2. \]

Adding \( C_1 \int_{B_{r_0}} V^p \) to both sides we get
\[ \int_{B_{r_0}} V^p \leq \frac{C_1}{C_1 + 1} \int_{B_{2r}} V^p + Cr^2. \]

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 \(U(x) = (x, u(x), Du(x), D^2 u(x))\). Applying the difference quotient to (3.11) and interchanging with the derivatives yields

\[
\frac{\partial^2}{\partial q^i \partial p^j} U|_h + \frac{\partial}{\partial q} [b^i \circ U|_h] + [c^j \circ U|_h] = 0. 
\]

We use the fundamental theorem of calculus to write

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

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

\[
= \int_0^1 Df((1 - t)U(x) + t U(x + he_v)) dt \cdot U_h(x).
\]

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

\[
[f \circ U]_h = \left( \frac{\partial f}{\partial q^i} \right)^h \alpha^i, (x) + \left( \frac{\partial f}{\partial p^j} \right)^h \alpha^j, (x) + \left( \frac{\partial f}{\partial z^l} \right)^h \alpha^l, (x),
\]

and the system (3.18) takes the form

\[
\frac{\partial^2}{\partial q^i \partial p^j} \alpha^i, (x) + \frac{\partial}{\partial q} [\tilde{b}^i \circ \alpha^i, (x)] + c^j \circ \alpha^j, (x) = 0.
\]

Here the coefficients are as follows:

\[
\tilde{a}^i, (x, z, p, q) = \left( \frac{\partial a^i \alpha^j}{\partial q^j} \right)^h (x) q^i, (x) + \left( \frac{\partial a^i \alpha^j}{\partial p^j} \right)^h (x) p^i, (x) + \left( \frac{\partial a^i \alpha^j}{\partial z^l} \right)^h (x) z^l, (x) + \left( \frac{\partial a^i \alpha^j}{\partial x^l} \right)^h (x),
\]

\[
\tilde{b}^i, (x, z, p, q) = \left( \frac{\partial b^i \alpha^j}{\partial q^j} \right)^h (x) q^i, (x) + \left( \frac{\partial b^i \alpha^j}{\partial p^j} \right)^h (x) p^i, (x) + \left( \frac{\partial b^i \alpha^j}{\partial z^l} \right)^h (x) z^l, (x) + \left( \frac{\partial b^i \alpha^j}{\partial x^l} \right)^h (x),
\]

\[
\tilde{c}^i, (x, z, p, q) = \left( \frac{\partial c^i \alpha^j}{\partial q^j} \right)^h (x) q^i, (x) + \left( \frac{\partial c^i \alpha^j}{\partial p^j} \right)^h (x) p^i, (x) + \left( \frac{\partial c^i \alpha^j}{\partial z^l} \right)^h (x) z^l, (x) + \left( \frac{\partial c^i \alpha^j}{\partial x^l} \right)^h (x).
\]

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

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

Using (3.12) and (3.13) we then obtain

\[
|\tilde{a}(x, z, p, q)| \leq C(I_{p-2,h}(x)|q| + I_{p-1,h}(x)(|p| + |z| + 1)),
\]

\[
|\tilde{b}(x, z, p, q)| + |\tilde{c}(x, z, p, q)| \leq C(I_{p-1,h}(x)|q| + I_{p,h}(x)(|p| + |z| + 1)).
\]
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

\[
\int \tilde{a}_{\alpha\beta} \tilde{a}_{\gamma\delta} u_h^4 \leq C \int (I_{p-1,h} |D^2 u_h| + I_{p,h}(|Du_h| + |u_h| + 1))(|D(u_h - l_h)| + |u_h - l_h|)\varphi^4
\]

\[
+ C \int (I_{p-1,h} |D(u_h - l_h)| + I_{p,h}|u_h - l_h|)|D^2 u_h|\varphi^4 |D\varphi|
\]

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

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

(3.20)

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

\[
\tilde{a}_{\alpha\beta}(x, z, p, q)q_{\alpha\beta} = \int_0^1 \frac{\partial a_{\alpha\beta}}{\partial x^\nu} ((1 - t)U(x) + tU(x + h))q_{\alpha\beta} q_{\mu\nu}^i \, dt
\]

\[
+ \int_0^1 \frac{\partial a_{\alpha\beta}}{\partial p^i} ((1 - t)U(x) + tU(x + h))p^i q_{\alpha\beta}^j \, dt
\]

\[
+ \int_0^1 \frac{\partial a_{\alpha\beta}}{\partial z^j} ((1 - t)U(x) + tU(x + h))z^j q_{\alpha\beta}^i \, dt
\]

\[
+ \int_0^1 \frac{\partial a_{\alpha\beta}}{\partial x^v} ((1 - t)U(x) + tU(x + h))q_{\alpha\beta}^v \, dt
\]

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

Combining the two inequalities we arrive at

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

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

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

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

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

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

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

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

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

$$= I + II + III.$$

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

$$\int_{B_r(x)} q \leq cs^\gamma$$

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

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

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

$$\int_{B_{r}(x)} I_{p,h} \leq Cs^\beta$$

for all $B_{r}(x) \subset B_{2r} \subset B_0$, and some $0 < \beta < 1$. Hence we can apply Lemma 3.10 and Poincaré’s inequality to estimate

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

Using the same argument we get

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

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

Inserting these two estimates into the above estimate for $\int I_{p-2,h} |D^2 u_h|^2 \phi^4$ we conclude that
\[
\int I_{p-2,h} |D^2 u_h|^2 \varphi^4 \leq C r^\beta \int_{B_{2r}} \varphi^4 (|D^2 u_h|^2 + |D(u_h - l_h)|^2) \\
+ C r^{\beta - 2} \int_{B_{2r}} (|D u_h|^2 + |D(u_h - l_h)|^2 + |u_h|^2 + |u_h - l_h|^2 + I_{p,h}) \\
+ C \int I_{p-2,h} (|D(u_h - l_h)|^2 \varphi^2 |D\varphi|^2 + |u_h - l_h|^2 (|D\varphi|^4 + \varphi^2 |D^2 \varphi|^2)). \\
\tag{3.21}
\]

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

\[
\int I_{p-2,h} \left( |D(u_h - l_h)|^2 \frac{X_{A_h}}{r^2} + |u_h - l_h|^2 \frac{X_{A_h}}{r^4} \right) \\
\leq C \|I_{p-2,h}\|_{L^p/(r'-2)(B_{r'})} \left( r^{-4} \|u_h - l_h\|_{L^p(A_{r_0})}^2 + r^{-2} \|D(u_h - l_h)\|_{L^p(A_{r_0})} \right) \\
\leq C r^{-2} \|I_{p-2,h}\|_{L^p/(r'-2)(B_{r'})} \|D(u_h - l_h)\|_{L^p(A_{r_0})}^2.
\]

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

\[ I_{s,h} \to V^s \quad \text{in} \quad L^{p/s} \quad \forall 1 \leq s \leq p. \]

We combine all the above estimates to get \( (l_h \to f_{A_h} \partial_s u + (x - x_0) f_{A_h} \partial_x Du) \) for \( r_0 \) small enough

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

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

\[
\int_{B_{r_0/8}} V^{p-2} |D^2 u|^2 \leq C (1 + r_0^{-2}) \int_{B_{r_0/4}} V^p + C r_0^{\beta - 2} \int_{B_{r_0/4}} |D^2 u|^2 + C r_0^\beta. \\
\tag{3.22}
\]

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

\[
\int_{B_r} V^p \leq C r^{2-\delta} \\
\tag{3.23}
\]

for all \( r \leq r_0/16 \) 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 \( L^\infty \) estimate for \( D^2 u \). Therefore we modify the estimates in order to apply a Gehring type lemma from [7].
Recalling (3.20) we have
\[
\int \nabla^4 \nabla^2 g^2 d^2 u_h \leq C \int \nabla^4 (I_{p-1,h} |D^2 u_h| + I_{p,h} (|Du_h| + |u_h| + 1)) (|D(u_h - l_h)| + |u_h - l_h|) \\
+ \frac{C}{r} \int \nabla^3 (I_{p-1,h} |D(u_h - l_h)| + I_{p,h} |u_h - l_h|) |D^2 u_h| \\
+ \frac{C}{r} \int \nabla^2 (I_{p-1,h} |D(u_h - l_h)| + I_{p,h} |u_h - l_h|) (|Du_h| + |u_h| + 1) \\
+ \frac{C}{r^2} \int \nabla (I_{p-2,h} |D^2 u_h| + I_{p-1,h} (|Du_h| + |u_h| + 1)) |u_h - l_h| \\
= I + \ldots + IV.
\]

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 \( I_{s,h} \to I_s \) in \( L^{p/2} \) we have, for every \( r \leq r_0/16 \), every \( h \) small enough and every \( \delta > 0 \),
\[
\int_{B_r} I_{p,h} \leq cr^{2-\delta}.
\]

Now we again estimate each term separately. We start with \( I \). By Young’s inequality we get
\[
I \leq \varepsilon \int \nabla^4 (I_{p-2,h} |D^2 u_h|^2 + I_{p,h} (|Du_h|^2 + |u_h|^2 + 1)) \\
+ C \int \nabla^3 (I_{p,h} (|D(u_h - l_h)|^2 + |u_h - l_h|^2).
\]

and we continue to estimate the last term with the help of Lemma 3.10, (3.24) and Poincaré’s inequality by
\[
\int \nabla^2 (I_{p,h} (|D(u_h - l_h)|^2 + |u_h - l_h|^2) \leq C r^{2-\delta} \int_{B_{2r}} (|D^2 (u_h - l_h)|^2 + |D(u_h - l_h)|^2) \\
+ C r^{-\delta} \int_{B_{2r}} (|D(u_h - l_h)|^2 + |u_h - l_h|^2) \\
\leq C r^{2-\delta} \int_{B_{2r}} |D^2 u_h|^2.
\]

Next we estimate
\[
II \leq \varepsilon \int \nabla^4 (I_{p-2,h} |D^2 u_h|^2 + C r^{-2} \int_{B_{2r}} I_{p,h} |u_h - l_h|^2 \\
+ \frac{C}{r} \int_{B_{2r}} (I_{p-2,h} |D^2 u_h|^4/3) \left( \int_{B_{2r}} |D(u_h - l_h)|^4 \right)^{1/4}.
\]
The second term can be estimated as above:

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

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

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

III can be estimated by

\[ III \lesssim \frac{\gamma}{r^2} \int_{B_{2r}} (|D(u_h - l_h)|^2 + I_{p,h}|u_h - l_h|^2 + \gamma \int_{B_{2r}} (I_{p-2,h} + I_{p,h})(|D u_h|^2 + |u_h|^2 + 1) \]

\[ \lesssim C(\gamma + r^{2-\delta}) \int_{B_{2r}} (|D^2 u_h|^2 + C(\gamma + r^{2-\delta}) \int_{B_{2r}} (I_{p-2,h} + I_{p,h})(|D u_h|^2 + |u_h|^2 + 1). \]

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

\[ IV \lesssim C \left( \int_{B_{2r}} I_{p-2,h}^4 |D^2 u_h|^4/3 \right)^{3/4} \left( \int_{B_{2r}} |u_h - l_h|^4 \right)^{1/4} \]

\[ + \gamma \int_{B_{2r}} |u_h - l_h|^2 + C(\gamma + r^{2-\delta}) \int_{B_{2r}} (I_{p-2,h} + I_{p,h})(|D u_h|^2 + |u_h|^2 + 1) \]

\[ \lesssim C \left( \int_{B_{2r}} I_{p-2,h}^4 |D^2 u_h|^4/3 \right)^{3/2} + C(\gamma + r^{2-\delta}) \int_{B_{2r}} (|D^2 u_h|^2 + |u_h|^2 + 1). \]

We also note that we have the ellipticity estimate

\[ \int \phi^4 I_{p-2,h}|D^2 u_h|^2 \leq C \int \phi^4 \phi_{\alpha\beta} \phi_{\alpha\beta}^2 u_h^2 + C \int \phi^4 I_{p,h}(|D u_h|^2 + |u_h|^2 + 1). \]

Combining all these estimates we get

\[ \int_{B_r} I_{p-2,h}|D^2 u_h|^2 \leq \frac{C}{r} \left( \int_{B_{2r}} I_{p-2,h}^4 |D^2 u_h|^4/3 \right)^{3/2} + C(\gamma + r^{2-\delta}) \int_{B_{2r}} |D^2 u_h|^2 \]

\[ + C(\gamma + r^{2-\delta}) \int_{B_{2r}} (I_{p-2,h} + I_{p,h})(|D u_h|^2 + |u_h|^2 + 1). \]
Since $u \in W^{3,2}(B_{r})$ and $V^{p/2} \in W^{1,2}(B_{r})$ for $r \leq r_0/16$ we can let $h \to 0$ and get

$$
\int_{B_{r}} V^{p-2} |D^3 u|^{2} \leq \frac{C}{r} \left( \int_{B_{2r}} V^{4(p-2)/3} |D^3 u|^{4/3} \right) + C(\gamma + r_0^{2-\gamma}) \int_{B_{2r}} V^{p-2} |D^3 u|^{2} + C_{1} \int_{B_{2r}} V^{2p}.
$$

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

$$
\left( \int_{B_{r}} f^{d} \right)^{1/d} \leq C \int_{B_{2r}} f g + C(\gamma + r_0^{2-\gamma}) \left( \int_{B_{2r}} f^{d} \right)^{1/d} + C_{1} \left( \int_{B_{2r}} h^{d} \right)^{1/d} \tag{3.25}
$$

for all balls $B \subset B_{r_{0}/16}$. 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

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

and (for some constant $C > 0$

$$
\left( \int_{B} f^{d} \right)^{1/d} \leq C \int_{B} f g + \varepsilon \left( \int_{B} f^{d} \right)^{1/d} + C \left( \int_{B} h^{d} \right)^{1/d} \tag{3.26}
$$

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

$$
h^{d} \log^{c_{0} \beta}(e + h) \in L_{\text{loc}}^{1}(\Omega),
$$

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

$$
\int_{B} f^{d} \log^{c_{0} \beta}(e + \frac{f}{\|f\|_{d,2B}}) \leq c \left( \int_{B} e^{\beta g^{d}} \right) \left( \int_{B} f^{d} \right) + c \int_{B} h^{d} \log^{c_{0} \beta}(e + \frac{f}{\|f\|_{d,2B}}) \tag{3.27}
$$

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

**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 $d(x) = \text{dist}(x, \mathbb{R}^{n} \setminus B_{0})$ and

$$
\tilde{f}(x) = d(x)^{n/d} f, \quad \tilde{h}(x) = d(x)^{n/d} h, \quad w(x) = \chi_{B_{0}}(x),
$$

Two-dimensional curvature functionals 3099
where $\chi_{B_0}$ 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
\[
\left( \int_{B} \tilde{f}^{d} \right)^{1/d} \leq C \int_{2B} \tilde{f} + C \epsilon \left( \int_{2B} \tilde{f}^{d} \right)^{1/d} + C \left( \int_{2B} \tilde{h}^{d} \right)^{1/d} + C \left( \int_{2B} w \right)^{1/d},
\]
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 $f$)
\[
M(f^{d})^{1/d} \leq CM(\tilde{f}g) + C \epsilon M(f^{d})^{1/d} + C M(h^{d})^{1/d} + C M(w)^{1/d}.
\]
For $\epsilon_0$ small enough we therefore have
\[
M(f^{d})^{1/d} \leq CM(\tilde{f}g) + CM(h^{d})^{1/d} + C M(w)^{1/d},
\]
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.

Now we want to apply this lemma to our estimate (3.25). From the previous subsection we know that
\[
f^{d} = V^{p-2} |D^3 u|^2 \in L^1_{\text{loc}}(B_{r_0/16}), \quad h^{d} = V^{2p} \in L^1_{\text{loc}}(B_{r_0/16}).
\]
Hence it remains to check that
\[
e^{\beta g^{d}} = e^{\beta V^{p-2}} \in L^1_{\text{loc}}(B_{r_0/16})
\]
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
\[
\int_{B_{r_0/8}} |D(V^{p/2})|^2 \leq c_1(r_0).
\]
Next we let $\eta \in C^\infty_c(B_{r_0/8})$ be a cut-off function such that $0 \leq \eta \leq 1$, $\eta(x) \equiv 1$ for all $x \in B_{r_0/16}$ and $\|D\eta\|_{L^\infty(B_{r_0/8})} \leq cr_0^{-1}$. Defining $v = \eta V^{p/2}$ we get
\[
\int_{B_{r_0/8}} |Dv|^2 \leq cr_0^{-2} \int_{B_{r_0/8}} V^{p} + c \int_{B_{r_0/8}} |DV^{p/2}|^2 \leq c_2(r_0).
\]
Hence we see that $u = v/\sqrt{c_2(r_0)} \in H^1_0(B_{r_0/8})$ and
\[
\int_{B_{r_0/8}} |Du|^2 \leq 1.
\]
Therefore, by the Moser–Trudinger inequality (see [23]), there exist constants $\beta_0, C = C(r_0) > 0$ such that
\[
\int_{B_{r_0/16}} e^{\beta_0 V^{p}} \leq \int_{B_{r_0/8}} e^{c_2(r_0)\beta_0 u^2} \leq C.
\]
In particular this implies with the help of Young’s inequality that for every $\beta > 0$,

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

Since also

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

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

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

for every $\alpha > 0$. Hence

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

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

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

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

$$\int_{I_p-2,h} (|D(u_h - l_h)|^2 + |u_h - l_h|^2 \frac{X_h}{r^4}) \leq C \int_{A_r} |D^2 u_h|^2.$$

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

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

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 $r^\beta$ 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 \( \Sigma \) be a closed surface and \( p > 2 \). Assume that \( f_k \in W^{2,p}_\text{im}(\Sigma, \mathbb{R}^n) \) is a sequence satisfying \( 0 \in f_k(\Sigma) \) for all \( k \in \mathbb{N} \) and

\[
\mathcal{E}^p(f_k) \leq C. \tag{4.1}
\]

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 \( W^{2,p}_\text{im}(\Sigma, \mathbb{R}^n) \) to an \( f \in W^{2,p}_\text{im}(\Sigma, \mathbb{R}^n) \). In particular, the convergence is in \( C^{1,\beta}(\Sigma, \mathbb{R}^n) \) for any \( \beta < 1 - 2/p \), and

\[
\mathcal{E}^p(f) \leq \liminf_{k \to \infty} \mathcal{E}^p(f_k). \tag{4.2}
\]

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 \( 8\pi \). 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 \( \Sigma_k \) 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 \( 8\pi \). 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

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

Moreover, for any \( p > 2 \),

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

where the constant \( C \) depends on \( \mathcal{W}^p(f) \).
Proof. By \cite[(1.2)]{22}, for $0 < \sigma < \rho < \infty$ we have

\[
\sigma^{-2} \mu_f(B_\sigma) \leq \rho^{-2} \mu_f(B_\rho) + \frac{1}{4} W(f) + \frac{1}{2} \int_{\Sigma \cap B_\rho} \rho^{-2}(x, H) \, d\mu - \frac{1}{2} \int_{\Sigma \cap B_\sigma} \sigma^{-2}(x, H) \, d\mu.
\]

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

\[
\sigma^{-2} \mu_f(B_\sigma) \leq \frac{1}{4} W(f) + \frac{1}{2\sigma} \int_{\Sigma \cap B_\sigma} |H| \, d\mu \leq \frac{1}{4} W(f) + C\sigma^{-2(1/p)}(W^p(f))^{1/p},
\]

where we have used Hölder’s inequality in the last step. \hfill \Box

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

\[
a_{ij}^\alpha \partial_{ij}^\alpha u = \psi_j \quad \text{for } j = 1, \ldots, k.
\]

There is an $\epsilon_0 = \epsilon_0(p) > 0$ such that if $|a_{ij}^\alpha(x) - \delta_{ij}^\alpha| \leq \epsilon_0$ for all $x \in B_\varrho$, then for some $C = C(p) < \infty$ we have the estimate

\[
\|D^2 u\|_{L^p(B_{\varrho/2})} \leq C (\|\psi\|_{L^p(B_{\varrho})} + \varrho^{-1} \|Du\|_{L^p(B_{\varrho})}).
\]

**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

\[
a_{ij}^\alpha \partial_{ij}^\alpha(\eta u^i) = \eta \psi_j + a_{ij}^\alpha (\partial_{ij}^\alpha \eta) u^i + a_{ij}^{\alpha \beta} (\partial_\alpha \eta \partial_\beta u^i + \partial_\beta \eta \partial_\alpha u^i).
\]

Hence we have

\[
\Delta(\eta u^i) = (\delta_{ij}^\alpha \partial_j^\alpha - a_{ij}^\alpha) \partial_{ij}^\alpha(\eta u^i) + \eta \psi_j
\]

\[
+ a_{ij}^{\alpha \beta} (\partial_\alpha \eta \partial_\beta u^i + \partial_\beta \eta \partial_\alpha u^i).
\]

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

\[
\|D^2(\eta u)\|_{L^p(B_1)} \leq C \epsilon_0 \|D^2(\eta u)\|_{L^p(B_1)} + C (\|\psi\|_{L^p(B_1)} + \|Du\|_{L^p(B_1)}),
\]

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

**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
\[
u_{k,q}(0) = 0 \quad \text{and} \quad D\nu_{k,q}(0) = 0.
\]

For \( \varepsilon > 0 \) we define \( r_k(q, \varepsilon) = \sup \{ r \in (0, r_k(q)) : \| D\nu_{k,q} \|_{C^0(B_r)} < \varepsilon \} \) 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
\[
\| \tilde{u}_{k,q} \|_{C^0(B_1)} + \| D\tilde{u}_{k,q} \|_{C^0(B_1)} \leq C.
\]

From the bound \( W^p(f_k) \leq C \) we further infer that
\[
\int_\Sigma |H_{\tilde{f}_k}|^p \, d\mu_{\tilde{f}_k} = r_k^{p-2} \int_\Sigma |H_{\tilde{f}_k}|^p \, d\mu_{\tilde{f}_k} \leq Cr_k^{p-2} \to 0.
\]

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 \( f_k \) converge to \( f_0 \) locally in \( C^{1,\beta} \), for every \( 0 < \beta < 1 - 2/p \), up to diffeomorphisms. Weak lower semicontinuity of \( W^p \) implies
\[
H_{f_0} = 0.
\]
The Gauß–Bonnet theorem yields 

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

and thus we get further

$$\int_{\Sigma_0} |A_{\phi}|^2 d\mu_{\phi} \leq \liminf_{k \to \infty} \int_{\Sigma} |A_{f_k}|^2 d\mu_{f_k} \leq C.$$  

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

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

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

$$\limsup_{R \to \infty} \frac{\mu_{f_0}(B_R)}{\pi R^2} < 2.$$ (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 $f_k \circ \phi_k \to f_0$ locally in $C^1$ and moreover

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

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 $f_k(q_k) = 0$ 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 $k$,

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

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

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

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 $T_{p} f_0 = \lim_{k \to \infty} T_{p_k'} (f_k \circ \phi_k) = \lim_{k \to \infty} T_{\phi_k}(f_k)$, and
analogously $T'_p f_0 = \lim_{k \to \infty} T_{q_k} f_k$. 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. \qed

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 $\Sigma$ and $p > 2$, denote by $\alpha^p_{\Sigma}(p)$ the infimum of the energy $\mathcal{E}^p$ among all smooth immersions from $\Sigma$ into $\mathbb{R}^n$. Then $\alpha^p_{\Sigma}(p)$ is attained by a smooth immersion $f : \Sigma \to \mathbb{R}^n$.

**Proof.** Using mollification it is easy to see that

$$\alpha^p_{\Sigma}(p) = \inf \{ \mathcal{E}^p(f) : f \in W^{2,p}_{\text{loc}}(\Sigma, \mathbb{R}^n) \}.$$  

Thus the limiting map $f \in W^{2,p}_{\text{loc}}(\Sigma, \mathbb{R}^n)$ 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. \qed

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

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

Taking the infimum with respect to $f$ shows that the function $\alpha^p_{\Sigma} : [2, \infty) \to \mathbb{R}$ 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

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

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

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

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

$$\mathcal{W}(f) = \frac{1}{4} \int_{\Sigma} |A|^2 \, d\mu_g + \pi \chi(\Sigma). \quad (4.9)$$
The infimum of the Willmore energy among immersions of $\Sigma$ into $\mathbb{R}^n$ satisfies $\beta_{\Sigma}^n < 8\pi$ (see [2]). Thus for $p > 2$ close to 2, we conclude for a minimizing $f$ 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 $\mathcal{W}$ among all smooth immersions from $\Sigma$ into $\mathbb{R}^n$. Repeating the previous discussion with $\beta_{\Sigma}^n$ instead of $\alpha_{\Sigma}^n$ 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 $\Sigma$ 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 $\Sigma$. 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 $\Sigma$. For $f \in W^{2,p}_{\text{im}}(\Sigma, \mathbb{R}^n)$ and any $V \in W^{2,p}(\Sigma, \mathbb{R}^n)$ we define the norm

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

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

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

resp.

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

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

$$\|D\mathcal{E}^p(f \circ \varphi)\|_{f \circ \varphi} = \|D\mathcal{E}^p(f)\|_f,$$  \hfill (5.1)
resp.
\[ \|D\!W^p(f \circ \varphi)\|_{f \circ \varphi} = \|D\!W^p(f)\|_f. \]  
(5.2)

Now we can formulate the main results of this section.

**Theorem 5.1.** Let \( f_k \in W^{2,p}_\text{im}(\Sigma, \mathbb{R}^n) \), \( p > 2 \), be a sequence satisfying
\[ E^p(f_k) \leq C \quad \text{and} \quad \|D\!E^p(f_k)\|_{f_k} \to 0. \]

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 \( f \in W^{2,p}_\text{im}(\Sigma, \mathbb{R}^n) \), and \( f \) is a smooth critical point of \( E^p \).

**Theorem 5.2.** Let \( f_k \in W^{2,p}_\text{im}(\Sigma, \mathbb{R}^n) \), \( \delta > 0 \), \( p > 2 \), be a sequence satisfying
\[ W^p(f_k) \leq C, \quad \mathcal{W}(f_k) \leq 8\pi - \delta \quad \text{and} \quad \|D\!W^p(f_k)\|_{f_k} \to 0. \]

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 \( f \in W^{2,p}_\text{im}(\Sigma, \mathbb{R}^n) \), and \( f \) is a smooth critical point of \( \mathcal{W}^p \).

We recall that the Palais–Smale condition cannot hold for the Willmore functional itself 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 & Riviére [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 \( f \in W^{2,p}_\text{im}(\Sigma, \mathbb{R}^n) \) 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 \( E^p \) 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
\[ \|Du_k\|_{C^0(B_r)} \leq L \leq 1 \quad \text{and} \quad \|u_k\|_{W^{2,p}(B_r)} \leq C \quad \text{and} \quad \|D\!E^p(f_k)\|_{f_k} \leq C, \]  
(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
\[ \|\psi_k\|_{C^1(B_r)} \to 0 \quad \text{and} \quad \|\psi_k\|_{W^{2,p}(B_r)} \leq C. \]  
(5.4)
Next we recall from (2.1) the Fréchet derivative, in a graph representation:

\[ D\mathcal{E}^p(f_k)(0, \psi_k) = \int_{B_r} (a_i^{\alpha\beta}(Du_k, D^2u_k) \partial_{\alpha\beta}^2 \psi_k^l + b_i^\alpha(Du_k, D^2u_k) \partial_\alpha \psi_k^l), \]  

(5.5)

where

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

\[ b_i^\alpha(Du_k, D^2u_k) = \frac{p}{8} (1 + |A_k|^2)^{(p-2)/2} \frac{\partial B_{ij}^{\beta\gamma, \mu\nu}}{\partial p_{ik}^\alpha}(Du_k) \partial_{\beta\gamma}^2 u_k^l \partial_{\mu\nu} u_k^m \sqrt{\det g_k}, \]

\[ + \frac{1}{4} (1 + |A_k|^2)^{p/2} \frac{\partial \sqrt{\det g_k}}{\partial p_{ik}^\alpha}(Du_k). \]

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

\[ a_i^{\alpha\beta}(Du_k, D^2u_k) \leq C (1 + |D^2u_k|^2)^{(p-1)/2}, \]

\[ b_i^\alpha(Du_k, D^2u_k) \leq C L (1 + |D^2u_k|^2)^{p/2}, \]

and obtain, as \(k \to \infty,\)

\[ \int_{B_r} a_i^{\alpha\beta}(Du_k, D^2u_k) \partial_{\alpha\beta}^2 \psi_k^l - \eta \partial_{\alpha\beta}^2 (u_k^l - u^l) \to 0, \]  

(5.6)

\[ \int_{B_r} b_i^\alpha(Du_k, D^2u_k) \partial_\alpha \psi_k^l \to 0. \]  

(5.7)

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

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

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

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

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

\[ \int_{B_r} \eta (a_i^{\alpha\beta}(Du_k, D^2u_k) - a_i^{\alpha\beta}(Du, D^2u_k)) \partial_{\alpha\beta}^2 (u_k^l - u^l) \to 0, \]

and by adding the last two equations we get

\[ \int_{B_r} \eta (a_i^{\alpha\beta}(Du, D^2u_k) - a_i^{\alpha\beta}(Du, D^2u)) \partial_{\alpha\beta}^2 (u_k^l - u^l) \to 0 \quad \text{as} \quad k \to \infty. \]
Finally, we use the ellipticity (see (3.12)) to estimate
\[
\int_{B_r} \eta (a^{\alpha \beta} (Du, D^2 u_k) - a^{\alpha \beta} (Du, D^2 u)) \partial_{\alpha \beta}^2 (u_k^i - u^i)
\]
\[
= \int_{B_r} \eta \int_0^1 \partial_{\alpha \beta}^2 (Du, D^2 u + t D^2 (u_k - u)) \partial_{\alpha \beta}^2 (u_k^i - u^i) \partial_{\alpha \beta}^2 (u_k^i - u^i) \, dt
\]
\[
\geq \lambda \int_{B_r} \eta \int_0^1 (1 + |D^2 u + t D^2 (u_k - u)|^2)^{(p-2)/2} |D^2 (u_k - u)|^2 \, dt
\]
\[
\geq c \lambda \int_{B_{r/2}} |D^2 (u_k - u)|^p.
\]

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

Acknowledgments. The authors would like to thank the referee for the detailed comments and the many helpful suggestions on the first version of this paper. This research was supported in part by the DFG Collaborative Research Center SFB/Transregio 71, and by a PIMS Postdoctoral Fellowship of the second author.

References


