

# THE AUBIN–NITSCHKE TRICK FOR SEMILINEAR PROBLEMS

HANNE HARDER

**ABSTRACT.** The Aubin–Nitsche trick is a common tool to show  $L^2$ -error estimates for discretizations of  $H^1$ -elliptic linear partial differential equations arising for example as Euler–Lagrange equations of a quadratic energy functional [Cia78]. The technique itself is linear: for quasilinear problems it is not applicable. We generalize the Aubin–Nitsche trick to a class of minimization problems closely related to semi-linear partial differential equations.

In textbooks on the numerical analysis of partial differential equations [Cia78, Bra92], the Aubin–Nitsche trick is usually presented after establishing discretization error bounds in  $O(h^m)$  for the minimization of  $H^1$ -elliptic energies by  $m$ -th order Lagrangian finite elements, where  $h$  is a mesh width parameter. While by Poincaré’s inequality one automatically obtains  $L^2$ -error estimates in  $O(h^m)$ , this is not optimal as the  $L^2$ -interpolation error is of the better order  $O(h^{m+1})$ . The Aubin–Nitsche trick is then introduced as a tool to obtain optimal  $L^2$ -error estimates from the  $H^1$ -error estimates under mild additional regularity assumptions.

In these arguments the energy is quadratic, i.e., of the type  $\mathfrak{J}(v) = \frac{1}{2}a(v, v) + f(v)$ , where  $a(\cdot, \cdot) : H \times H \rightarrow \mathbb{R}$  denotes an  $H$ -elliptic scalar product on some subspace  $H \subset H^1(\Omega)$ , and  $f \in H'$  is a linear map. The  $H^1$ -error estimates can be generalized to nonlinear problems [Cia78, Ch. 5]. The Aubin–Nitsche trick, however, relies on the linear concept of Galerkin orthogonality. To obtain optimal  $L^2$ -error estimates for nonlinear problems one option is to deform the nonlinear problem to a linear one and use a method of continuity argument [DR80]. In the case of only mildly nonlinear problems, in particular semi-linear ones, this technique is not needed. Instead, we propose a new proof that replaces addition by integration in the nonlinear setting and estimate additional terms.

Linearity and semi-linearity are concepts that refer to the Euler–Lagrange equations associated to energy problems obtained by setting the first variation of the energy to zero. In this context, it is more feasible to work with properties of the energy directly. The concept of “mildly nonlinear” we will use is a bound on the third variation of the energy. We will call such energies predominantly quadratic. One example is  $\mathfrak{J}(v) := \int_{\Omega} |Dv(x)|^2 dx + \int_{\Omega} \psi(u(x)) dx$ , where  $\psi$  denotes a nonlinear potential.

## CONTENTS

1. The Aubin–Nitsche Trick for Quadratic Energies	2
2. Semi-linearity and Predominantly Quadratic Energies	4
3. Galerkin Orthogonality and the Adjoint Problem	6
4. $L^2$ -Error Estimate	6
References	8

# 1. THE AUBIN–NITSCHKE TRICK FOR QUADRATIC ENERGIES

We will briefly summarize the basic tools we need from standard theory and the recall the Aubin–Nitsche-Trick for quadratic energies.

In the following  $\Omega \subset \mathbb{R}^d$  will denote an open subset with piecewise Lipschitz boundary  $\partial\Omega$ . Further  $W^{k,p}(\Omega, \mathbb{R}^n)$  will denote the standard Sobolev space with the usual abbreviation  $H^k(\Omega, \mathbb{R}^n) := W^{k,2}(\Omega, \mathbb{R}^n)$  [Wlo87].

**1.1.  $H^1$ -Ellipticity and  $H^1$ -Discretization Error Bounds.** We consider the minimization of energies  $\mathfrak{J}$  in  $H \subset W_\phi^{1,2}(\Omega, \mathbb{R}^n)$ , where  $\phi$  denotes suitable boundary data:

$$(1) \quad u \in H : \quad \mathfrak{J}(u) \leq \mathfrak{J}(v) \quad \forall v \in H.$$

To bound the error of discrete approximations to minimizers of  $\mathfrak{J}$ , we need the concept of  $W^{1,2}$ -ellipticity.

**Definition 1.1.** Let  $\mathfrak{J} : H \rightarrow \mathbb{R}$  be twice continuously differentiable, and let  $\delta^2 \mathfrak{J}$  denote the second variation of  $\mathfrak{J}$ . We say that  $\mathfrak{J}$  is

(a)  $W^{1,2}$ -coercive, if there exists a constant  $\lambda > 0$  such that for all  $v \in H$  and  $V \in W_0^{1,2}(\Omega, \mathbb{R}^n)$  we have

$$(2) \quad \lambda \|V\|_{W^{1,2}}^2 \leq \delta^2 \mathfrak{J}(v)(V, V),$$

(b)  $W^{1,2}$ -bounded, if there exists a constant  $\Lambda > 0$  such that for all  $v \in H$  and for all  $V, W \in W_0^{1,2}(\Omega, \mathbb{R}^n)$  we have

$$(3) \quad |\delta^2 \mathfrak{J}(v)(V, W)| \leq \Lambda \|V\|_{W^{1,2}} \|W\|_{W^{1,2}},$$

(c)  $W^{1,2}$ -elliptic, if (a) and (b) hold.

In order to obtain a finite-dimensional approximation of  $H$ , we assume that we have a conforming grid  $G$  on  $\Omega$ , i.e., a partition into polytopes such that the closures intersect in common faces.

**Definition 1.2.** We say that a conforming grid  $G$  for the domain  $\Omega \subset \mathbb{R}^d$  is of width  $h$  and order  $m$ , if for each element  $T_h$  of  $G$  there exists a  $C^\infty$ -diffeomorphism  $F_h : T_h \rightarrow T$  to a reference element  $T \subset \mathbb{R}^d$  that scales with  $h$  of order  $m$ , i.e.,

$$c h^{-d} \leq \|\det(DF_h)\|_{L^\infty} \leq C h^{-d}, \quad \|\partial_\alpha F_h\|_{L^\infty} \leq C h \quad \forall \alpha = 1, \dots, d,$$

and

$$|F^{-1}|_{W^{k,\infty}} \leq C h^k \quad \forall k = 0, \dots, m.$$

Let  $S_{h;\phi}^m \subset H \cap C(\Omega, \mathbb{R}^n)$  be a finite-dimensional approximation space for a grid  $G$  on  $\Omega$  of width  $h$  and order  $m$ . Note that this requires that the boundary data  $\phi$  can be represented exactly in  $S_{h;\phi}^m$ . This requirement may be waived and replaced by a standard approximation argument for boundary data [Cia78, Ch. 4].

Consider the discrete approximation of (1)

$$(4) \quad u_h \in S_{h;\phi}^m : \quad \mathfrak{J}(u_h) \leq \mathfrak{J}(v_h) \quad \forall v_h \in S_{h;\phi}^m.$$

In order to control the error between  $u$  and  $u_h$ , we need standard approximation conditions for the discrete space  $S_{h;\phi}^m$ .

The first condition consists of an estimate for the best approximation error in  $S_{h;\phi}^m$  [Cia78].

**Condition 1.1.** Let  $kp > d$ ,  $m \geq k - 1$ , and  $u \in W_\phi^{k,p}(\Omega, \mathbb{R}^n)$ . For small enough  $h$  let there exist a map  $u_I \in S_{h;\phi}^m$  and constants  $C_1, C_2$  with

$$(5) \quad |u_I|_{W^{l,q}(\Omega, \mathbb{R}^n)} \leq C_1 |u|_{W^{l,q}(\Omega, \mathbb{R}^n)}$$

for all  $k - \frac{d}{p} \leq l \leq k$  and  $q \leq \frac{pd}{d-p(k-l)}$  such that  $u_I$  fulfills on each element  $T_h \in G$  the estimate

$$(6) \quad \|u - u_I\|_{L^p} + h |u - u_I|_{W^{1,p}} \leq C_2 h^k |u|_{W^{k,p}}.$$

The second condition is generally known as an inverse estimate.

**Condition 1.2.** On a grid  $G$  of width  $h$  and order  $m$ , under the additional assumption that  $F_h^{-1} : T \rightarrow T_h$  scales with order 2 for every  $T_h \in G$ , let there exist a constant  $C_3$ , such that

$$(7) \quad \|v_h\|_{W^{1,p}(T_h, \mathbb{R}^n)} \leq C_3 h^{-d \max\{0, \frac{1}{q} - \frac{1}{p}\}} \|v_h\|_{W^{1,q}(T_h, \mathbb{R}^n)}$$

for any  $v_h \in S_h^m(T_h, M)$  and for any  $p, q \in [1, \infty]$ .

Note that the discrete functions in  $S_{h;\phi}^m$  are globally only of  $C^0 \cap W^{1,2}$ -smoothness. Whenever we consider higher Sobolev norms, we implicitly define them as grid dependent, i.e.,

$$(8) \quad |u|_{W^{k,p}(\Omega, \mathbb{R}^n)} := \left( \sum_{T_h \in G} |u|_{W^{k,p}(T_h, \mathbb{R}^n)}^p \right)^{\frac{1}{p}}.$$

By summation over the elements of  $G$ , estimates like (6) and (7) carry over to global grid-dependent norms. For  $H^1$ -elliptic energies, Céa's Lemma with Condition 1.1 yields the following  $W^{1,2}$ -error estimate [Cia78].

**Theorem 1.3.** Let  $2(m+1) > d$ , and  $m \geq 1$ . Assume that  $u \in W_\phi^{m+1,2}(\Omega, M)$  is a minimizer of an  $H^1$ -elliptic  $\mathfrak{J} : H \rightarrow \mathbb{R}$ . Then the discrete minimizer

$$u_h := \arg \min_{v_h \in S_h^m} \mathfrak{J}(v_h)$$

fulfills the a priori error estimate

$$(9) \quad \|u - u_h\|_{W^{1,2}} \leq C_4 h^m |u|_{W^{m+1,2}}.$$

**1.2. The Aubin–Nitsche Trick For Quadratic Energies.** The purpose of the Aubin–Nitsche trick is to show that for  $W^{1,2}$ -elliptic minimization problems the  $L^2$ -discretization error is in  $O(h^{m+1})$ .

We recall the Aubin–Nitsche lemma for the approximation of a quadratic minimization problem in  $H = H_0^1(\Omega, \mathbb{R})$  by Lagrangian finite elements. For an elliptic bilinear form  $a(\cdot, \cdot)$  and given  $f \in H^{-1}$  consider the energy  $J(v) = \frac{1}{2}a(v, v) - (f, v)$ , the variational equalities

$$(10) \quad u \in H : \quad a(u, v) = (f, v) \quad \forall v \in H,$$

$$(11) \quad u_h \in S_{h;0}^m : \quad a(u_h, v_h) = (f, v_h) \quad \forall v_h \in S_{h;0}^m,$$

and the adjoint problem

$$(12) \quad w \in H : \quad a(v, w) = (g, v) \quad \forall v \in H,$$

where  $g := u - u_h$ . We assume  $H^2$ -regularity of the adjoint problem, i.e.,  $|w|_{H^2} \leq C \|g\|_{L^2}$ . The subtraction of equations (10) and (11) with the same test function  $v_h \in S_{h;0}^m \subset H$  yields the concept of Galerkin orthogonality, i.e.,

$$(13) \quad a(u - u_h, v_h) = 0 \quad \forall v_h \in S_{h;0}^m.$$

Using Galerkin orthogonality and the  $H^1$ -ellipticity of  $a(\cdot, \cdot)$ , we can then estimate

$$\begin{aligned} \|u - u_h\|_{L^2}^2 &= (g, u - u_h) = a(u - u_h, w) = a(u - u_h, w - w_I) \\ &\leq \Lambda \|u - u_h\|_{H^1} \|w - w_I\|_{H^1} \\ &\leq Ch^m |u|_{H^k} h |w|_{H^2} \\ &\leq Ch^{m+1} |u|_{H^2} \|u - u_h\|_{L^2}. \end{aligned}$$

Galerkin orthogonality  $a(u - u_h, w_I) = 0$  is the essential tool used here. It allows to incorporate an approximation of  $w$  and thus leads to a better order estimate than the  $H^1$ -error. It is at first glance a purely linear concept that is verified by adding the equations (10) and (11) for the same discrete test function, a technique that does not work for nonlinear energies. We will circumvent the need for addition by integration and show that semi-linearity of the Euler–Lagrange equation will then be sufficient to obtain equivalent error estimates.

## 2. SEMI-LINEARITY AND PREDOMINANTLY QUADRATIC ENERGIES

We now introduce the concept of predominantly quadratic energies. Energies with semi-linear Euler–Lagrange equations fall in this category. At the same time, this property is exactly what we need for the  $L^2$ -error bounds. Let the energy functional  $\mathfrak{J} : W^{1,2}(\Omega, \mathbb{R}^n) \rightarrow \mathbb{R}$  be given by

$$\mathfrak{J}(v) := \int_{\Omega} L(Dv, v, x) \, dx,$$

where  $L : \mathbb{R}^{n \times d} \times \mathbb{R}^n \times \Omega \rightarrow \mathbb{R}$ ,  $(p, z, x) \mapsto L(p, z, x)$ , is a smooth Lagrangian, and  $D$  denotes the (weak) differentiation operator of a function from  $\Omega \subset \mathbb{R}^d$  to  $\mathbb{R}^n$ .

We calculate the first variation of  $\mathfrak{J}$  at a function  $u \in W^{1,2}(\Omega, \mathbb{R}^n)$  in direction  $V \in W_0^{1,2}(\Omega, \mathbb{R}^n)$ :

$$\begin{aligned} \delta \mathfrak{J}(u)(V) &= \left. \frac{d}{d\tau} \right|_{\tau=0} \mathfrak{J}(u + \tau V) \\ &= \int_{\Omega} \partial_p L(Du, u, x) \cdot DV + \partial_z L(Du, u, x) \cdot V \, dx \\ &= \int_{\Omega} (-D_x \partial_p L(Du, u, x) + \partial_z L(Du, u, x)) \cdot V \, dx, \end{aligned}$$

where  $\partial_p, \partial_z, \partial_x$  denote partial differentiation of  $L$  with respect to the corresponding variables. Setting the variation to zero yields the corresponding system of Euler–Lagrange equations (cf. [Eva98])

$$(14) \quad - \sum_{i=1}^d \frac{d}{dx_i} \left( \partial_{p_i^k} L(Du, u, x) \right) + \partial_{z^k} L(Du, u, x) = 0 \quad \text{in } \Omega \quad (k = 1, \dots, n).$$

On the other hand a semi-linear system of partial differential equations

$$a_{ij}(x) D^{ij} u(x) + a_0(Du(x), u(x), x) = 0$$

is characterized by the independence of the coefficients  $a_{ij}$  of the solution  $u$ . Thus, the Euler–Lagrange equation (14) is semi-linear if the operator  $\partial_p^2 L$  is independent of  $u$ , i.e.,

$$\partial_p^2 L(Du, u, x) = \partial_p^2 L(x),$$

and accordingly

$$\partial_p^3 L(Du, u, x) = 0, \quad \text{and} \quad \partial_z \partial_p^2 L(Du, u, x) = 0.$$

These vanishing third order derivatives appear in the third variation of  $\mathfrak{J}$ . In general, the second variation of  $\mathfrak{J}$  reads

$$\begin{aligned} \delta^2 \mathfrak{J}(u)(V, W) &= \int_{\Omega} \partial_p^2 L(Du, u, x)(DV, DW) + \partial_z \partial_p L(Du, u, x)(DV, W) \\ &\quad + \partial_p \partial_z L(Du, u, x)(V, DW) + \partial_z^2 L(Du, u, x)(V, W) dx, \end{aligned}$$

and the third variation is

$$\begin{aligned} \delta^3 \mathfrak{J}(u)(V, W, U) &= \int_{\Omega} \partial_p^3 L(Du, u, x)(DV, DW, DU) + \partial_z \partial_p^2 L(Du, u, x)(DV, DW, U) \\ &\quad + \partial_p \partial_z \partial_p L(Du, u, x)(DV, W, DU) + \partial_z^2 \partial_p L(Du, u, x)(DV, W, U) \\ &\quad + \partial_p^2 \partial_z L(Du, u, x)(V, DW, DU) + \partial_z \partial_p \partial_z L(Du, u, x)(V, DW, U) \\ &\quad + \partial_p \partial_z^2 L(Du, u, x)(V, W, DU) + \partial_z^3 L(Du, u, x)(V, W, U) dx. \end{aligned}$$

If (14) is semi-linear, the third variation reduces to

$$\begin{aligned} (15) \quad \delta^3 \mathfrak{J}(u)(V, W, U) &= \int_{\Omega} \partial_z^2 \partial_p L(Du, u, x)(DV, W, U) + \partial_z \partial_p \partial_z L(Du, u, x)(V, DW, U) \\ &\quad + \partial_p \partial_z^2 L(Du, u, x)(V, W, DU) + \partial_z^3 L(Du, u, x)(V, W, U) dx. \end{aligned}$$

Thus, semi-linearity of the Euler–Lagrange equation necessarily implies the dependence of each term of the third variation on at most one direction gradient.

We turn this observation into a definition. More generally, we will consider predominantly quadratic energies. By this we mean the following:

**Definition 2.1.** Let  $q > \max\{d, 2\}$  and  $\mathfrak{J} : H \rightarrow \mathbb{R}$  be an energy functional. We say that  $\mathfrak{J}$  is predominantly quadratic with respect to  $q$  if  $\mathfrak{J}$  is  $C^3$ , and for any  $u \in H \cap W^{1,q}(\Omega, \mathbb{R}^n)$ ,  $U \in W^{2,2}(\Omega, \mathbb{R}^n)$ , and  $V \in W^{1,2} \cap W^{o,r}(\Omega, \mathbb{R}^n)$  with either  $(o, r) = (1, 2)$ , or  $o = 0$  and  $r \leq d$ , there exists a constant  $C_5$  possibly depending on  $\|u\|_{W^{1,q}}$  such that

$$(16) \quad |\delta^3 \mathfrak{J}(u)(U, V, V)| \leq C_5 \|U\|_{W^{2,2}} \|V\|_{W^{1,2}} \|V\|_{W^{o,r}}.$$

**Example.** We have seen in (15) that as long as the Lagrangian  $L$  is smooth enough and its third variations are bounded in  $L^q$  in terms of  $\|u\|_{W^{1,q}}$ , the leading term of the third variation of the corresponding energy will have a bound of the form

$$|\delta^3 \mathfrak{J}(u)(U, V, V)| \leq C \left( \int_{\Omega} (|DV||V||U| + |V|^2|DU|)^{\frac{q}{q-1}} dx \right)^{1-\frac{1}{q}},$$

if we assume semi-linearity of the Euler–Lagrange system. Thus, such an energy is predominantly quadratic (Hölder’s inequality).

**Example.** The leading term of the third variation of the energy for a typical quasi-linear equation, e.g., the minimal surface energy for graphs  $\mathfrak{J}(u) = \int_{\Omega} \sqrt{1 + |Du|^2} dx$ , has the form

$$|\delta^3 \mathfrak{J}(v)(U, V, V)| \leq C \left( \int_{\Omega} (|U||DV|^2 + |DU||DV||V|)^{\frac{q}{q-1}} dx \right)^{1-\frac{1}{q}}.$$

For  $d < 4$  such an energy is predominantly quadratic with respect to  $q = \infty$ , but not in general.

### 3. GALERKIN ORTHOGONALITY AND THE ADJOINT PROBLEM

We consider the variational formulations of the problems (1) and (4)

$$(17) \quad u \in H : \quad \delta \mathfrak{J}(u)(V) = 0 \quad \forall V \in W_0^{1,2}(\Omega, \mathbb{R}^n),$$

$$(18) \quad u_h \in S_{h;\phi}^m : \quad \delta \mathfrak{J}(u_h)(V_h) = 0 \quad \forall V_h \in S_{h;0}^m.$$

These correspond to (10) and (11) in the linear setting. By inserting a discrete test function into (17) we obtain by the fundamental theorem of calculus (replacing subtraction in the linear setting)

$$(19) \quad 0 = \delta \mathfrak{J}(u_h)(V_h) - \delta \mathfrak{J}(u)(V_h) = \int_0^1 \delta^2 \mathfrak{J}(\Gamma(t))(V_h, u_h - u) dt,$$

where  $\Gamma(t) = (1-t)u + tu_h$ . This is a nonlinear generalization of Galerkin orthogonality. Note that for a quadratic energy  $\delta^2 \mathfrak{J}$  is independent of the function  $\Gamma(t)$  and we recover the standard notion of Galerkin orthogonality (13).

We now define a nonlinear generalization of the adjoint problem (12) featuring in the Aubin–Nitsche trick. For nonlinear energies the adjoint problem is essentially a linearization of problem (17) [DR80] with a right hand side that is given by the difference of the solutions  $u$  and  $u_h$  to (17) and (18), respectively:

$$(20) \quad W \in W_0^{1,2}(\Omega, \mathbb{R}^n) : \quad \delta^2 \mathfrak{J}(u)(W, V) = -(V, u_h - u)_{L^2} \quad \forall V \in W_0^{1,2}(\Omega, \mathbb{R}^n).$$

Note that as long as the operator  $\partial_p^2 L(x)$  is in  $W^{1,q}$ ,  $\partial_p \partial_z L(Du, u, x)$  is in  $L^q$ ,  $\partial_z^2 L(Du, u, x)$  is in  $L^{\frac{\max\{q,4\}}{2}}$ , and  $u_h - u$  is in  $L^2$  for  $q > \max\{2, d\}$ , standard regularity results for linear elliptic systems [LU68] yield that the adjoint problem is  $H^2$ -regular, i.e., that the solution  $W$  fulfills

$$(21) \quad \|W\|_{W^{2,2}(\Omega, \mathbb{R}^n)} \leq C \|u - u_h\|_{L^2(\Omega, \mathbb{R}^n)}.$$

### 4. $L^2$ -ERROR ESTIMATE

We will now combine the nonlinear Galerkin orthogonality (19) with the standard estimate for  $H^1$ -elliptic energies (9) to show that a higher order estimate for the  $L^2$ -error for predominantly quadratic energies can be obtained by the Aubin–Nitsche trick analogous to the linear setting described in Section 1.2.

**Theorem 4.1.** *Let  $m \in \mathbb{N}$ , and  $2(m+1) > d$ . Assume that  $u \in W_\phi^{m+1,2}(\Omega, \mathbb{R}^n)$  is a minimizer of an elliptic energy  $\mathfrak{J}$  that is predominantly quadratic with respect to  $q > \max\{2, d\}$  with  $q \leq 2d$  if  $d - 2m = 1$ . Let  $u_h$  be a (local) minimizer of  $\mathfrak{J}$  in  $S_{h;\phi}^m$  fulfilling (9). Finally, suppose that the adjoint problem (20) is  $H^2$ -regular, i.e., that its solution  $W$  fulfills (21). Then there exists a constant  $C_6$ , such that*

$$(22) \quad \|u - u_h\|_{L^2(\Omega, \mathbb{R}^n)} \leq C_6 h^{m+1},$$

where  $C_6$  depends nonlinearly on  $\|u\|_{W^{1,q}}$  and  $|u|_{W^{m+1,2}}$ .

*Proof.* We insert  $V := u_h - u$  into (20), and obtain

$$\|u - u_h\|_{L^2}^2 = -\delta^2 \mathfrak{J}(u)(W, u_h - u),$$

where  $W \in W_0^{2,2}(\Omega, \mathbb{R}^n)$  is the solution of (20). Let  $W_I \in S_{h;0}^m$  be an approximation of  $W$  in the sense of Condition 1.1. As  $u_h$  is a local minimizer in  $S_{h;\phi}^m$ , generalized Galerkin

orthogonality (19) holds, so that for  $\Gamma(t) = (1-t)u + tu_h$

$$\begin{aligned}
 \|u_h - u\|_{L^2(\Omega, \mathbb{R}^n)}^2 &= -\delta^2 \mathfrak{J}(u)(W, u_h - u) + \int_0^1 \delta^2 \mathfrak{J}(\Gamma(t))(W_I, u_h - u) dt \\
 &= \int_0^1 \int_0^t \frac{1}{t} \frac{d}{ds} \delta^2 \mathfrak{J}(\Gamma(s))(sW_I + (t-s)W, u_h - u) ds dt \\
 &= \int_0^1 \int_0^t \frac{1}{t} \delta^3 \mathfrak{J}(\Gamma(s))(sW_I + (t-s)W, u_h - u, u_h - u) ds dt \\
 (23) \quad &+ \int_0^1 \int_0^t \frac{1}{t} \delta^2 \mathfrak{J}(\Gamma(s))(W_I - W, u_h - u) ds dt.
 \end{aligned}$$

The second integral in (23) is estimated using the ellipticity assumption (3):

$$\int_0^1 \int_0^t \frac{1}{t} \delta^2 \mathfrak{J}(\Gamma(s))(W_I - W, u_h - u) ds dt \leq \Lambda \|W_I - W\|_{W^{1,2}(\Omega, \mathbb{R}^n)} \|u_h - u\|_{W^{1,2}(\Omega, \mathbb{R}^n)}.$$

Using Condition 1.1 on  $W_I$ , the  $H^1$ -error bound (9), and the  $H^2$ -regularity (21), we obtain

$$\begin{aligned}
 \int_0^1 \int_0^t \frac{1}{t} \delta^2 \mathfrak{J}(\Gamma(s))(W_I - W, u_h - u) ds dt &\leq C h^{m+1} |W|_{W^{2,2}(\Omega, \mathbb{R}^n)} \|u\|_{W^{m+1,2}(\Omega, \mathbb{R}^n)} \\
 &\leq C h^{m+1} \|u\|_{W^{m+1,2}(\Omega, \mathbb{R}^n)} \|u_h - u\|_{L^2(\Omega, \mathbb{R}^n)}.
 \end{aligned}$$

In order to estimate the first integral term in (23) we use that  $\mathfrak{J}$  is predominantly quadratic

$$\begin{aligned}
 |\delta^3 \mathfrak{J}(\Gamma(s))(sW_I + (t-s)W, u_h - u, u_h - u)| \\
 \leq C(\|\Gamma(s)\|_{W^{1,q}}) \|u_h - u\|_{W^{1,2}} \|u_h - u\|_{W^{o,r}} (s\|W_I\|_{W^{2,2}} + (t-s)\|W\|_{W^{2,2}}).
 \end{aligned}$$

Using again Condition 1.1 on  $W_I$ , the  $H^1$ -error bound, and the  $H^2$ -regularity, we obtain

$$\begin{aligned}
 \int_0^1 \int_0^t \frac{1}{t} \delta^3 \mathfrak{J}(\Gamma(s))(sW_I + (1-s)W, u_h - u, u_h - u) ds dt \\
 \leq C \int_0^1 \int_0^t C(\|\Gamma(s)\|_{W^{1,q}}) ds dt \|u_h - u\|_{W^{1,2}} \|u_h - u\|_{W^{o,r}} \|u_h - u\|_{L^2} \\
 \leq C(\max\{\|u\|_{W^{1,q}}, \|u_h\|_{W^{1,q}}\}) h^m |u|_{W^{m+1,2}} \|u_h - u\|_{W^{o,r}} \|u_h - u\|_{L^2}.
 \end{aligned}$$

Note that we can assume that

$$\|u_h\|_{W^{1,q}} \leq C(\|u\|_{W^{1,q}} + |u|_{W^{m+1,2}})$$

for  $h$  small enough, as we can use the inverse estimate in Condition 1.2 on  $(u_h - u_I) \in S_{h;0}^m$ , Condition 1.1 on  $u_I$ , and the  $H^1$ -error bound to estimate

$$\begin{aligned}
 \|u_h\|_{W^{1,q}} &\leq \|u_h - u_I\|_{W^{1,q}} + \|u_I\|_{W^{1,q}} \\
 &\leq h^{-d(\frac{1}{2} - \frac{1}{q})} \|u_h - u_I\|_{W^{1,2}} + C\|u\|_{W^{1,q}} \\
 &\leq Ch^{m-d(\frac{1}{2} - \frac{1}{q})} |u|_{W^{m+1,2}} + C\|u\|_{W^{1,q}}.
 \end{aligned}$$

Note that  $m - d(\frac{1}{2} - \frac{1}{q}) \geq 0$ , if we assume  $q \leq 2d$  in the case  $d - 2m = 1$ . If  $o = 1$  and  $r = 2$ , then we can use the  $H^1$ -error bound again to obtain

$$\begin{aligned}
 \int_0^1 \int_0^t \frac{1}{t} \int_0^1 \delta^3 \mathfrak{J}(\Gamma(s))(sW_I + (1-s)W, u_h - u, u_h - u) ds dt \\
 \leq C(\|u\|_{W^{1,q}}, |u|_{W^{m+1,2}}) h^{2m} \|u_h - u\|_{L^2}
 \end{aligned}$$

with  $2m \geq m + 1$ .

If instead  $o = 0$  and  $r \leq d$ , then either we are in the same situation as before, or  $d \geq 4$  and  $\frac{2d}{d-2} \leq r \leq d$ . In that case  $L^p$ -interpolation with  $\varepsilon = h$  yields

$$\begin{aligned} \|u_h - u\|_{L^r} &\leq h \|u_h - u\|_{L^\infty} + h^{1 - \frac{r(d-2)}{2d}} \|u_h - u\|_{L^{\frac{2d}{d-2}}} \\ &\leq h (\|u\|_{L^\infty} + \|u_h\|_{L^\infty}) + C h^{m+1 - \frac{r(d-2)}{2d}} |u|_{W^{m+1,2}}. \end{aligned}$$

As  $(m+1) \geq \frac{d}{2}$  and  $d \geq r$ , we have  $m+1 - \frac{r(d-2)}{2d} \geq 1$ . Thus, we obtain also for this case

$$\begin{aligned} \int_0^1 \int_0^t \frac{1}{t} \int_0^1 \delta^3 \mathfrak{J}(\Gamma(s)) (sW_t + (1-s)W, u_h - u, u_h - u) \, ds \, dt \\ \leq C (\|u\|_{W^{1,q}}, |u|_{W^{m+1,2}}) h^{m+1} \|u_h - u\|_{L^2}. \end{aligned}$$

This yields the assertion.  $\square$

## REFERENCES

- [Bra92] Dietrich Braess. *Finite Elemente - Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie*. Springer, 1992.
- [Cia78] Philippe G. Ciarlet. *The finite element method for elliptic problems*. Elsevier, 1978.
- [DR80] Manfred Dobrowolski and Rolf Rannacher. Finite element methods for nonlinear elliptic systems of second order. *Math. Nachr.*, 94:155–172, 1980.
- [Eva98] Lawrence C. Evans. *Partial Differential Equations*. AMS, 1998.
- [LU68] Olga A. Ladyzhenskaya and Nina N. Ural'tseva. *Linear and Quasilinear Elliptic Equations*. AP, 1968.
- [Wlo87] Joseph Wloka. *Partial Differential Equations*. Cambridge University Press, 1987.

HANNE HARDERING, TECHNISCHE UNIVERSITÄT DRESDEN, INSTITUT FÜR NUMERISCHE MATHEMATIK, D-01062 DRESDEN, GERMANY  
*E-mail address:* hanne.hardering@tu-dresden.de