Applied Numerical Mathematics 59 (2009) 2119–2130

Contents lists available at ScienceDirect

Applied Numerical Mathematics

www.elsevier.com/locate/apnum

# Convergence of adaptive finite element methods in computational mechanics ☆

Carsten Carstensen

*Department of Mathematics, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany*

A R T I C L E I N F O

*Article history:*
Available online 14 December 2008

*Keywords:*
Adaptive algorithm
FEM
Linear elasticity
Convexity
Hencky elastoplasticity
Convergence

A B S T R A C T

The a priori convergence of finite element methods is based on the density property of the sequence of finite element spaces which essentially assumes a quasi-uniform mesh-refining. The advantage is guaranteed convergence for a large class of data and solutions; the disadvantage is a global mesh refinement everywhere accompanied by large computational costs.

Adaptive finite element methods (AFEMs) automatically refine exclusively wherever their refinement indication suggests to do so and consequently leave out refinements at other locations. In other words, the density property is violated on purpose and the a priori convergence is not guaranteed automatically and, in fact, crucially depends on algorithmic details. The advantage of AFEMs is a more effective mesh in many practical examples accompanied by smaller computational costs; the disadvantage is that the desirable convergence property is *not* guaranteed a priori. Efficient error estimators can justify a numerical approximation a posteriori and so achieve reliability. But it is *not* theoretically justified from the start that the adaptive mesh-refinement will generate an accurate solution at all. In order to foster the development of a convergence theory and improved design of AFEMs in computational engineering and sciences, this paper describes a particular version of an AFEM and analyses convergence results for three model problems in computational mechanics: linear elastic material (A), nonlinear monotone elastic material (B), and Hencky elastoplastic material (C). It establishes conditions sufficient for error-reduction in (A), for energy-reduction in (B), and eventually for strong convergence of the stress field in (C) in the presence of small hardening.



## 1. Introduction

Three model problems (A), (B), and (C) in computational mechanics with increasing degree of complexity are addressed in this paper: linear elasticity (A), nonlinear elasticity (B), and Hencky elastoplasticity with small hardening (C) as illustrated in Fig. 1. Their a posteriori error control is an important and, over the last decades, reasonably well-established topic; cf. the books [1,4,21,29,37] plus some select references [3,8,12,14,16,24,26,30,31] in elastoplasticity. Numerical experiments reported in those references clearly provide evidence for the extreme success of adaptive finite element methods (AFEMs). This success, however, although very welcome in practice, is *not* understood in theory and hence *not* guaranteed in the forthcoming computation. The literature on the recent theory of a priori convergence of the AFEMs is still small [6,17,19,27,36].

☆ Supported by the DFG Research Center MATHEON "Mathematics for key technologies" in Berlin and the Hausdorff Institute for Mathematics in Bonn, Germany.

*E-mail address:* cc@math.hu-berlin.de.

0168-9274/$30.00 © 2009 IMACS. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.apnum.2008.12.006

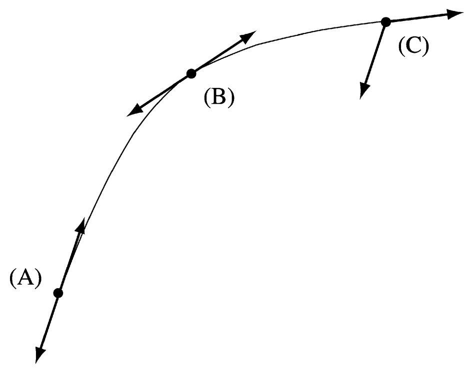


**Fig. 1.** Illustration of three standard stress–strain relations to model linear elasticity (A), (monotone) nonlinear elasticity (B), and Hencky elastoplasticity (C). The arrays indicate loading and unloading along the stress–strain path to emphasize that there are two different tangents involved in (C) modeled by a variational inequality.

This paper enlarges that theory (established for the Laplacian and the $p$-Laplacian) to problems in linear and nonlinear computational mechanics [22,32] and studies Algorithm 1.1 below with recursive loops of the form

$$\text{SOLVE} \to \text{ESTIMATE} \to \text{MARK} \to \text{REFINE}. \tag{1.1}$$

For the step SOLVE and ESTIMATE we refer to the aforementioned literature plus [2,14].

This papers follows [6,20,27] and adopts the bulk-criterion in the step MARK for edge-oriented contributions, written $\eta_E^{(\ell)}$ for any edge or face $E$. The carefully analyzed adaptive algorithm generates a sequence of strongly convergent stress approximations!

All three stationary materials (A), (B), and (C) allow for a finite element (FE) displacement $u_\ell$ and a FE stress approximation $\sigma_\ell$ to the exact displacement field $u$ and the exact stress $\sigma$ at each level $\ell$ of the AFEM.

In linear elasticity (A), an error-reduction property of the FE stress approximation in terms of the energy norm $|||\cdot|||$, that is

$$|||\sigma - \sigma_{\ell+1}|||^2 \leqslant \varrho|||\sigma - \sigma_\ell|||^2 + C\,\mathrm{OSC}_\ell^2 \quad \text{for all } \ell = 0, 1, 2, \ldots \tag{1.2}$$

can be guaranteed with some factor $\varrho < 1$. The oscillation term $\mathrm{OSC}_\ell$ in (1.2) is of higher-order when the given right-hand sides are smooth (and material properties are constant throughout the material body).

For nonlinear elasticity (B), let $E(v)$ denote the elastic energy and $\delta_\ell := E(u_\ell) - E(u)$ the difference of the discrete energy $E(u_\ell)$ to the minimal energy $E(u)$. The energy-reduction property

$$\delta_{\ell+1} \leqslant \varrho\delta_\ell + C\,\mathrm{OSC}_\ell^2 \quad \text{for all } \ell = 0, 1, 2, \ldots \tag{1.3}$$

can be guaranteed provided the energy $E$ is uniformly convex and its derivative $DE$ is Lipschitz continuous. Then, energy decrease and error decrease are equivalent and the $R$-order of convergence (up to higher-order terms) reads

$$|||\sigma - \sigma_\ell|||^2 \leqslant \varrho^\ell\delta_0 + \sum_{k=1}^{\ell-1} C\varrho^{\ell-1-k}\,\mathrm{OSC}_k^2 \quad \text{for all } \ell = 1, 2, 3, \ldots. \tag{1.4}$$

The a posteriori error control in elastoplasticity is more delicate and the convergence result covers just convergence up to higher-order terms in the sense of

$$\liminf_{\ell\to\infty} \mathrm{OSC}_\ell = 0 \quad \Rightarrow \quad \liminf_{\ell\to\infty} |||\sigma - \sigma_\ell||| = 0. \tag{1.5}$$

A typical adaptive algorithm is sketched below where the discrete stress $\sigma_\ell$ is piecewise constant with respect to a triangulation $\mathcal{T}_\ell$ with the set $\mathcal{E}_\ell$ of edges and faces in 2D and 3D, respectively. Then, for each edge or face $E \in \mathcal{E}_\ell$ of diameter $h_E := \mathrm{diam}(E)$ and with unit normal $\nu_E$, the contribution

$$\eta_E^{(\ell)} := h_E^{1/2}\left(\int_E \left|[\sigma_\ell]\nu_E\right|^2 ds\right)^{1/2} \tag{1.6}$$

accounts for the jump $[\sigma_\ell]\nu_E$ of the discrete stresses across the interior edge $E$ in the normal direction.

**Algorithm 1.1** *(Adaptive Algorithm (AFEM)).*
**Input:** Coarse shape-regular triangulation $\mathcal{T}_0$ of $\Omega$ into triangles with set of edges $\mathcal{E}_0$; $0 < \Theta < 1$.
**For** $\ell = 0, 1, 2, 3\ldots$ **do** (a)–(e):

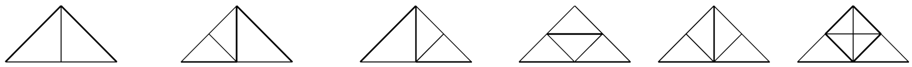


**Fig. 2.** Possible refinements of a triangle. Thick lines indicate reference edges in the sense of [6,10,13] for red-green-blue and newest-vertex bisection refinement. The reference edge indicates that edge which has to be refined if the triangle has to be refined — an information inherited in each refinement. For the understanding of this paper it suffices to know that there exists some refinements as depicted and the closure algorithm can be arranged to maintain shape-regularity.

(a) Solve the discrete problem with respect to the actual mesh $\mathcal{T}_\ell$ and corresponding FE spaces. Let $u_\ell$ denote the FE displacement and let $\sigma_\ell$ denote the discrete stress field.
(b) Compute $\eta_E^{(\ell)}$ for all edges or faces $E \in \mathcal{E}_\ell$ and $\eta_\ell := (\sum_{E\in\mathcal{E}_\ell} (\eta_E^{(\ell)})^2)^{1/2}$ as stress-error indication.
(c) Generate a set $\mathcal{M}_\ell$ of edges or faces in $\mathcal{E}_\ell$ such that

$$\Theta \eta_\ell^2 \leqslant \sum_{E\in\mathcal{M}_\ell} \big(\eta_E^{(\ell)}\big)^2. \tag{1.7}$$

(d) Control oscillations $\mathrm{OSC}_\ell$ and (possibly) add further edges to $\mathcal{M}_\ell$ to decrease $\mathrm{OSC}_{\ell+1}$.
(e) Run closure algorithm to avoid hanging nodes; refine all triangles $T$ with some edge or face $E$ in $\mathcal{M}_\ell$ with bisc5(T) and all other elements with red-green-blue or newest-vertex bisection refinement after Fig. 2. Let $\mathcal{T}_{\ell+1}$ denotes the resulting shape-regular triangulation.

**Output:** Sequence of discrete stress fields $\sigma_0, \sigma_1, \sigma_2, \ldots$ in $L^2(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}})$.

The proofs of (1.2)–(1.4) generalize techniques in [27] for a simple Poisson problem with arguments from convex analysis [32,35,39]. We adopt standard notation for Lebesgue and Sobolev spaces and norms.

## 2. Comments

Several remarks on modifications of the aforementioned algorithm as well as on empirical wisdom complement the theoretical proofs below.

### *2.1. Monitor oscillations*

One important aspect from the theoretical point of view in the convergence analysis is the focus of the oscillation terms. For globally constant volume and piecewise constant surface forces, those oscillations vanish and (d) can be canceled in the adaptive algorithm. Otherwise, the terms have to be monitored even though they may be of higher order and even if they decrease automatically. A simple way to achieve this is to compare $\mathrm{OSC}_\ell$ with $\eta_\ell$. For instance, fix some constant $c$ and if $c\eta_\ell \leqslant \mathrm{OSC}_\ell$, then red-refine elements (i.e. add all three edges to $\mathcal{M}_\ell$) such that

$$\mathrm{OSC}_{\ell+1} \leqslant \vartheta\, \mathrm{OSC}_\ell \quad \text{with some parameter } \vartheta < 1.$$

### *2.2. Bulk-criterion VS maximum-criterion for MARK*

The bulk criterion (1.7) dates back to [20]. Prior to that and still in the majority of papers nowadays, the maximum criterion is employed: Mark the element $T$ for (*red*) refinement provided

$$\Theta \max_{K\in\mathcal{T}_\ell} \eta_K \leqslant \eta_T.$$

This maximum criterion is motivated by an equi-distribution principle in 1D [5], but *not* theoretically supported in higher space dimensions. There is huge empirical evidence in the literature that the maximum criterion effectively leads to accurate approximations.

### *2.3. Computational complexity*

This paper is concerned with error reduction in terms of numbers of levels but not in terms of numbers of degrees of freedom. Coarsening led to the first proof of optimal complexity [6] for a model Poisson problem before Stevensen [33] provided optimal complexity for extremely small bulk parameter and small oscillations in AFEM. Those results are expect to hold for linear elasticity as well. The optimal strategy to involve oscillations is under current investigations [19]. The adoption for nonlinear elasticity is less clear [15] and the adoption for variational inequalities is still open. It appears that the present theory for optimal computational complexity does not fully explain the success of AFEMs in practice.

## 3. Mathematical and numerical models in elasticity

### 3.1. Strong formulation in elasticity

The material body $\Omega \subset \mathbb{R}^d$ is viewed as a bounded Lipschitz domain with polygonal or polyhedral boundary $\partial\Omega = \Gamma_D \cup \Gamma_N$ for $d=2$ or $d=3$ dimensions. It is loaded by applied volume forces $f \in L^2(\Omega;\mathbb{R}^d)$ and surface tractions $g \in L^2(\Gamma_N;\mathbb{R}^d)$ on some (relatively open) part $\Gamma_N$ of the boundary $\partial\Omega$ with exterior unit normal $\nu$. The material body is supported on the remaining closed part $\Gamma_D := \partial\Omega \setminus \Gamma_N$ of positive surface area. The displacement field is prescribed there by the Dirichlet part $u_D \in H^1(\Omega;\mathbb{R}^d)$ (piecewise affine below for simplicity).

Given a displacement field $u \in H^1(\Omega;\mathbb{R}^d)$, its (linearized) Green strain tensor reads

$$\varepsilon(v) := \big(Dv + (Dv)^T\big)/2 := \operatorname{sym} Dv.$$

In the setting of infinitely small strains, an elastic material relates the stress $\sigma$ through a $C^1$ and (strongly) monotone mapping $\mathcal{A}: \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ via

$$\sigma := \mathcal{A}\big(\varepsilon(u)\big) \quad \text{in } \Omega. \tag{3.1}$$

Details on the nonlinear situation will follow in Section 6 while, here, the focus is on the linear elastic material (A)

$$\sigma := \mathbb{C}\varepsilon(u) := \lambda \operatorname{tr}\big(\varepsilon(u)\big) I + 2\mu\varepsilon(u). \tag{3.2}$$

Two positive Lamé parameters $\lambda$ and $\mu$ define the fourth-order isotropic material tensor $\mathbb{C}$. This fourth-order tensor $\mathbb{C}$ and its inverse $\mathbb{C}^{-1}$ gives rise to the energy norm for stress fields defined by

$$\|\tau\|_{\mathbb{C}^{-1}} := \big(\tau : \mathbb{C}^{-1}\tau\big)^{1/2} \quad \text{for all } \tau \in \mathbb{R}^{d\times d}_{\mathrm{sym}}.$$

Here and throughout, $\mathbb{C}\tau$ denotes the action of $\mathbb{C}$ on $\tau$ and colon (i.e. :) denotes the scalar product of two matrices, e.g. $\tau : \gamma := \sum_{j,k=1,\dots,d} \tau_{jk}\gamma_{jk}$ for any $\tau, \gamma \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ with components $\tau_{jk}$ and $\gamma_{jk}$. The linear situation (3.2) is understood as a particular case of (3.1).

In the aforementioned notation, the *strong form* of boundary value problem (the Lamé–Navier equations in linear elasticity) reads: *Seek* $u := H^1(\Omega;\mathbb{R}^d)$ with (3.2) *and*

$$-\operatorname{div}\sigma = f \quad \text{and} \quad \sigma = \mathcal{A}\big(\varepsilon(u)\big) \quad \text{in } \Omega, \tag{3.3}$$

$$u = u_D \quad \text{on } \Gamma_D \quad \text{and} \quad \sigma\nu = g \quad \text{on } \Gamma_N. \tag{3.4}$$

### 3.2. Weak formulations in elasticity

The *weak form* consists of the (possibly nonlinear mapping) $a(u;v)$ and the linear functional $b(v)$, defined by

$$a(u;v) := \int_\Omega \mathcal{A}\big(\varepsilon(u)\big) : \varepsilon(v)\,dx = \int_\Omega \mathbb{C}\varepsilon(u) : \varepsilon(v)\,dx, \tag{3.5}$$

$$b(v) := \int_\Omega f\cdot v\,dx + \int_{\Gamma_N} g\cdot v\,ds \tag{3.6}$$

for all $u, v \in H^1(\Omega;\mathbb{R}^d)$. Then, the *weak formulation of the elastic boundary value problem* reads: *Seek* $u \in U := \{w \in H^1(\Omega;\mathbb{R}^d):\ u = u_D \text{ on } \Gamma_D\}$ *and*

$$a(u;v) = b(v) \quad \text{for all } v \in V := H^1_D\big(\Omega;\mathbb{R}^d\big) := \big\{v \in H^1\big(\Omega;\mathbb{R}^d\big) : v = 0 \text{ on } \Gamma_D\big\}.$$

### 3.3. Meshes and finite element spaces

The finite element method in level $\ell$ of the adaptive algorithm is based on a shape-regular triangulation $\mathcal{T}_\ell$ of the domain $\Omega$ in closed nondegenerating simplices. In other words, $\mathcal{T}_\ell$ has no hanging nodes and is locally quasiuniform.

Throughout this paper, $\mathcal{E}_\ell$ abbreviates the set of all edges in 2D and the set of all faces in 3D and $\mathcal{N}_\ell$ is the set of all nodes (or vertices of some triangle) in the triangulation $\mathcal{T}_\ell$.

The hat function $\varphi_z$ of some node $z \in \mathcal{N}_\ell$ is defined by the values $\varphi_z(x)$ for a node $x \in \mathcal{N}_\ell$, namely $\varphi_z(z) = 1$ and $\varphi_z(x) = 0$ for $x \in \mathcal{N}_\ell \setminus \{z\}$, followed by linear interpolation on each triangle. Then, the patch of a node is the open set $\omega_z := \{x \in \Omega:\ 0 < \varphi_z(x)\}$ which is the interior of the union of the set $\mathcal{T}(z)$ of neighboring elements,

$$\overline{\omega_z} = \bigcup \mathcal{T}(z) \quad \text{where } \mathcal{T}(z) = \{T \in \mathcal{T}:\ z \in T\}.$$

The set of free nodes $\mathcal{K}_\ell := \mathcal{N}_\ell \setminus \Gamma_D$ consists of all vertices which have a positive distance to the Dirichlet boundary $\Gamma_D$; the remaining nodes on the Dirichlet boundary are $\mathcal{N}_D := \mathcal{N}_\ell \setminus \mathcal{K}_\ell$.

For any subset $\omega$ (patch, triangle, face, or edge), let $P_k(\omega)$ denote the vector space of all algebraic polynomials viewed as real-valued functions on $\omega$ of total degree at most $k = 0, 1, 2$. Then,

$$P_k(\mathcal{T}_\ell) := \{v_h \in L^\infty(\Omega):\ \forall T \in \mathcal{T}_\ell,\ v_h|_T \in P_k(T)\}$$

denotes the piecewise polynomials of degree $\leqslant k$ and the $P_1$ finite element spaces reads

$$V_\ell := P_1(\mathcal{T}_\ell; \mathbb{R}^d) \cap V \quad \text{and} \quad U_\ell := P_1(\mathcal{T}_\ell; \mathbb{R}^d) \cap U.$$

### *3.4. Oscillations*

For the volume load $f$ and the mesh $\mathcal{T}_\ell$, let

$$\mathrm{OSC}_\ell(f) := \Bigl(\sum_{z \in \mathcal{K}_\ell} h_z^2 \int_{\omega_z} |f(x) - f_{\omega_z}|^2\, dx\Bigr)^{1/2}, \tag{3.7}$$

where $f_{\omega_z} := \int_{\omega_z} f(x)\, dx / |\omega_z|$ abbreviates the integral mean of the right-hand side $f$ over the patch $\omega_z$ of volume $|\omega_z|$ and size $h_z := \operatorname{diam}(\omega_z)$. For the surface tractions $g$, let $g_E$ abbreviate the integral mean of the right-hand side $g$ over the edge or face $E \in \mathcal{E}_N := \{E \in \mathcal{E}_\ell:\ E \subset \overline{\Gamma_N}\}$ and set

$$\mathrm{OSC}_\ell(g) := \Bigl(\sum_{E \in \mathcal{E}_N} h_E \int_E |g(x) - g_E|^2\, dx\Bigr)^{1/2}. \tag{3.8}$$

Throughout this paper we set $\mathrm{OSC}_\ell := \mathrm{OSC}_\ell(f) + \mathrm{OSC}_\ell(g)$.

### *3.5. Discrete solution*

For each level $\ell = 0, 1, 2, \ldots$, the *discrete problem of the elastic boundary value problem* reads: *Seek $u_\ell \in U_\ell$ such that*

$$a(u_\ell; v_\ell) = b(v_\ell) \quad \text{for all } v_\ell \in V_\ell.$$

For the ease of this presentation, suppose that $u_D$ is continuous and piecewise affine in the sense that

$$u_D \in \{w \in C^0(\Gamma_D; \mathbb{R}^d):\ \forall E \in \mathcal{E}_0 \text{ with } E \subset \Gamma_D, u_D|_E \in P_1(E)\}.$$

Then $u_\ell(z) = u_D(z)$ for all $z \in \mathcal{N}_D$ is equivalent to $u_\ell = u_D$ along $\Gamma_D$. Under standard assumptions there exists a unique discrete solution $u_\ell$ and one defines the error $e_\ell := u - u_\ell \in H^1(\Omega; \mathbb{R}^d)$ and its residual $R_\ell := b - a(u_\ell; \cdot) \in V^*$.

## 4. Preliminaries

This section is devoted to reliability and discrete local efficiency. In all three models (A), (B), and (C), there is a discrete stress field $\sigma_\ell$ such that the residual is a dual functional defined by

$$R_\ell(v) := \int_\Omega f \cdot v\, dx + \int_{\Gamma_N} g \cdot v\, ds - \int_\Omega \sigma_\ell : \varepsilon(v)\, dx \quad \text{for all } v \in V. \tag{4.1}$$

The energy norm, which plays a prominent role in all three models, reads

$$|||\tau||| := \Bigl(\int_\Omega \|\tau(x)\|^2_{\mathbb{C}^{-1}}\, dx\Bigr)^{1/2} \quad \text{for all } \tau \in L^2(\Omega; \mathbb{R}^{d \times d}_{\mathrm{sym}}) \tag{4.2}$$

and yields a norm for the displacements in $V$ via

$$\|v\|_V := |||\mathbb{C}\varepsilon(v)||| \quad \text{for all } v \in H^1(\Omega; \mathbb{R}^d).$$

Then, the dual norm

$$\|R_\ell\|_{V^*} := \sup_{v \in V \setminus \{0\}} \frac{R_\ell(v)}{\|v\|_V}$$

can be bounded in terms of explicit or implicit error estimators [1,4,21,37]. Estimator A typical error contribution $\eta_E^{(\ell)}$ denotes the jump of the discrete stress in normal direction (1.6) while

$$\bigl(\eta_E^{(\ell)}\bigr)^2 := h_E \int_E |g - \sigma_\ell \nu_E|^2\, ds \quad \text{for } E \in \mathcal{E}_N \quad \text{and} \quad \eta_E^{(\ell)} := 0 \quad \text{if } E \subset \Gamma_D. \tag{4.3}$$

**Theorem 4.1.** *There exist mesh-size-independent constants $C_1$ and $C_2$ such that there holds*

$$\|R_\ell\|_{V^*}^2 \leqslant C_1\,\eta_\ell^2 + C_2\,\mathrm{OSC}_\ell(f)^2.$$

**Proof.** A simple proof employs the Galerkin orthogonality $R_\ell(v) = R_\ell(v - v_\ell)$ for some discrete $v_\ell \in V_\ell$ defined in [9]. Although [9,18] study solely scalar problems, the arguments therein apply to the vectorial situation as well [38]; we omit further details. □

One key argument in the proof of (1.2) is the following assertion valid for all the three problems (A), (B), and (C) because only the discrete equilibrium enters.

**Theorem 4.2.** *There exist mesh-size-independent constants $C_3$ and $C_4$ such that there holds*

$$\sum_{E\in\mathcal{M}_\ell} \big(\eta_E^{(\ell)}\big)^2 \leqslant C_3 |\!|\!|\sigma_{\ell+1} - \sigma_\ell|\!|\!|^2 + C_4\,\mathrm{OSC}_\ell^2.$$

**Proof.** In the design of $\mathcal{T}_{\ell+1}$, the interior edge $E$ in $\mathcal{M}_\ell$ is bisected and so its midpoint $\mathrm{mid}(E) \in \mathcal{N}_{\ell+1}$ and the two neighboring triangles $T_\pm \in \mathcal{T}_\ell$ have a new inner nodes $\mathrm{mid}(T_\pm) \in \mathcal{N}_{\ell+1}$. The nodal basis functions $\phi_{\mathrm{mid}(E)}, \phi_{\mathrm{mid}(T_\pm)}$ (with respect to $\mathcal{T}_{\ell+1}$) of the two or three mentioned new nodes allow for a discrete function

$$\psi_E := \phi_{\mathrm{mid}(E)} - \alpha_+\phi_{\mathrm{mid}(T_+)} - \alpha_-\phi_{\mathrm{mid}(T_-)} \in W_0^{1,\infty}(\omega_E) \cap P_1(\mathcal{T}_{\ell+1})$$

with proper and bounded $\alpha_\pm \in \mathbb{R}$ such that

$$\int_E \psi_E\,ds = \frac{1}{2} \quad\text{and}\quad \int_{T_\pm} \psi_E\,dx = 0.$$

Then the test function $b_E := h_E[\sigma_\ell]\nu_E\,\psi_E \in V_{\ell+1}$ satisfies

$$\frac{1}{2}\big(\eta_E^{(\ell)}\big)^2 = \int_E b_E \cdot [\sigma_\ell]\nu_E\,ds = \int_{\omega_E} \sigma_\ell : Db_E\,dx.$$

An integration by parts in the last step adopts $\operatorname{div}\sigma_\ell = 0$ on each $T_\pm$. Since

$$\int_{\omega_E} \sigma_{\ell+1} : Db_E\ dx = \int_{\omega_E} f \cdot b_E\,dx$$

(there is no contribution from $g$ as, here, $E$ is an interior edge) and since $\int_{\omega_E} f_E \cdot b_E\,dx = 0$ for any constant $f_E$, this proves

$$\frac{1}{2}\big(\eta_E^{(\ell)}\big)^2 = \int_E b_E \cdot [\sigma_\ell]\nu_E\,ds = \int_{\omega_E} (\sigma_\ell - \sigma_{\ell+1}) : Db_E\,dx + \int_{\omega_E} (f - f_E)\cdot b_E\,dx.$$

Cauchy inequalities and bounds on various norms of $b_E$ eventually prove the assertion provided $E$ is an interior edge.

In the situation where $E \subset \partial\Omega$, a similar argument leads to local contributions of $\mathrm{OSC}_\ell(f) + \mathrm{OSC}_\ell(g)$; we omit the details and refer to [1,4,20,27,37] for similar arguments. □

## 5. Error reduction for linear elasticity

This section is devoted to the analysis of the error reduction (1.2) for linear elasticity (A). Throughout this section, let $u$ and $u_\ell$ solve the continuous and discrete problem and

$$|\!|\!|\sigma - \sigma_\ell|\!|\!|^2 = \|R_\ell\|_{V^*}^2. \tag{5.1}$$

**Theorem 5.1.** *There exists mesh-size-independent constants $\varrho$ and $C$ such that* (1.2) *holds.*

**Proof.** The combination of stability (5.1), Theorem 4.1, and the bulk criterion (1.7) leads to

$$|\!|\!|\sigma - \sigma_\ell|\!|\!|^2 \leqslant \frac{C_1}{\Theta} \sum_{E\in\mathcal{M}_\ell} \big(\eta_E^{(\ell)}\big)^2 + C_2\,\mathrm{OSC}_\ell^2. \tag{5.2}$$

This and the discrete local efficiency of Theorem 4.2 yield

$$|\!|\!|\sigma - \sigma_\ell|\!|\!|^2 \leqslant \frac{C_1C_3}{\Theta} |\!|\!|\sigma_{\ell+1} - \sigma_\ell|\!|\!|^2 + \left(\frac{C_1C_4}{\Theta} + C_2\right)\mathrm{OSC}_\ell^2. \tag{5.3}$$

A final essential ingredient in (A) is the Galerkin orthogonality

$$|||\sigma_{\ell+1} - \sigma_\ell|||^2 = |||\sigma - \sigma_\ell|||^2 - |||\sigma - \sigma_{\ell+1}|||^2. \tag{5.4}$$

A substitution of $|||\sigma_{\ell+1} - \sigma_\ell|||^2$ from (5.4) in (5.3) followed by some rearrangements give (1.2) with $\varrho = 1 - \Theta/(C_1 C_3)$ and $C = (C_4 + \Theta C_2/C_1)/C_3$. □

## 6. Energy reduction for nonlinear elasticity

This section is devoted to a hyperelastic material (B) where the Lipschitz continuous $\mathcal{A} := DW$ is the functional matrix of some smooth strongly convex function $W : \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}$. That means that there exist positive constants $\alpha$ and $L := \mathrm{Lip}(\mathcal{A})$ with

$$\alpha|A - B|^2 \leqslant W(B) - W(A) - \mathcal{A}(A) : (B - A) \quad \text{and} \tag{6.1}$$

$$\big|\mathcal{A}(A) - \mathcal{A}(B)\big| \leqslant L|A - B| \quad \text{for all } A, B \in \mathbb{R}^{d\times d}_{\mathrm{sym}}. \tag{6.2}$$

A typical example with (6.1), (6.2) is the nonlinear Hencky material or nonlinear Hooke's law of [28, Section 3.3] and [39, Section 62.8] where

$$W(A) := \frac{1}{2}k\big(\mathrm{tr}(A)\big)^2 + \kappa\varphi\big(|\operatorname{dev} A|^2\big)$$

for material constants $k$ and $\kappa$ and some nonlinear function $\varphi : [0, \infty) \to [0, \infty)$ with (non-displayed) smoothness and growth conditions on $\varphi$, $\varphi'$, $\varphi''$.

Throughout this section, let $u$ and $u_\ell$ denote the minimizer in the continuous and discrete setting of the energy

$$E(v) := \int_\Omega W\big(\varepsilon(v)\big)\,dx - b(v) \quad \text{for all } v \in U. \tag{6.3}$$

**Theorem 6.1.** *There exists mesh-size-independent constants $\varrho$ and $C$ such that* (1.3) *holds.*

**Proof.** The assumptions (6.1), (6.2) guarantee that $E$ is uniformly convex, $DE$ is Lipschitz continuous, and (5.1) holds in the form of

$$|||\sigma - \sigma_\ell|||^2 \leqslant C_5 \|R_\ell\|^2_{V^*}. \tag{6.4}$$

To see this, take $A := Du(x)$ and $B := Du_\ell(x)$ in (6.1), which yields an expression for almost every $x$ in $\Omega$. Thus, its integral satisfies

$$\alpha\|u - u_\ell\|^2_{H^1(\Omega)} \leqslant \int_\Omega (\sigma - \sigma_\ell) : \varepsilon(u - u_\ell)\,dx.$$

Owing to the equilibrium condition on $\sigma$, namely $f + \operatorname{div}\sigma = 0$ in $\Omega$ plus boundary conditions, its weak form yields

$$\int_\Omega \sigma : \varepsilon(u - u_\ell)\,dx = b(u - u_\ell). \tag{6.5}$$

The combination of the latter two estimates results in (6.4), namely

$$\alpha\,\|u - u_\ell\|^2_{H^1(\Omega)} \leqslant b(u - u_\ell) - \int_\Omega \sigma_\ell : \varepsilon(u - u_\ell)\,dx = R_\ell(u - u_\ell) \leqslant \|R_\ell\|_{V^*}\,\|u - u_\ell\|_V.$$

Furthermore, the residual is of the form of Theorems 4.1 and 4.2. The proof therefore might follow that of Theorem 5.1. However, one additional difficulty in the proof of (1.3) is the lack of a Galerkin orthogonality in (B).

The new argument involves (6.1) twice: First, with $A = \varepsilon(u_\ell)$ and $B = \varepsilon(u)$ and integration over $\Omega$ and with (6.5), (6.1) yields

$$\delta_\ell + \alpha\big\|\varepsilon(u - u_\ell)\big\|^2_{L^2(\Omega)} \leqslant \int_\Omega (\sigma - \sigma_\ell) : \varepsilon(u - u_\ell)\,dx.$$

Since (6.2) and Young's inequality, this is bounded from above by $\frac{\alpha}{2}\|\varepsilon(u - u_\ell)\|^2_{L^2(\Omega)} + C_6|||\sigma - \sigma_\ell|||^2$. Hence

$$\delta_\ell + \frac{\alpha}{2}\big\|\varepsilon(u - u_\ell)\big\|^2_{L^2(\Omega)} \leqslant C_6|||\sigma - \sigma_\ell|||^2. \tag{6.6}$$

Second, with $A = \varepsilon(u_{\ell+1})$ and $B = \varepsilon(u_\ell)$ and integration over $\Omega$ and substitution

$$\int_\Omega \sigma_{\ell+1} : \varepsilon(u_{\ell+1} - u_\ell)\, dx = b(u_{\ell+1} - u_\ell),$$

(6.1) yields in terms of $E(u_{\ell+1})$ and $E(u_\ell)$ that

$$\alpha \left\| \varepsilon(u_{\ell+1} - u_\ell) \right\|^2_{L^2(\Omega)} \leqslant \delta_\ell - \delta_{\ell+1}. \tag{6.7}$$

With (6.4), Theorem 4.1, and Theorem 4.2 plus the bulk criterion as in the convergence analysis of problem (A), there follow (5.2), (5.3). The combination of (5.3) with (6.6), (6.7) and (6.2) yields positive constants $C_7$ and $C$ such that

$$\delta_\ell \leqslant C_7(\delta_\ell - \delta_{\ell+1}) + C_7\, \mathrm{OSC}^2_\ell.$$

This implies (1.3) with $\varrho = 1 - 1/C_7$. □

## 7. Convergence for primal Hencky elastoplasticity

The irreversible part $p \in L^2(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}})$ of the elastoplastic strain is modeled in (C) via an additive split

$$\varepsilon(u) = \mathbb{C}^{-1}\sigma + p \tag{7.1}$$

of the total Green strain $\varepsilon(u)$ with an elastic (reversible) part $\mathbb{C}^{-1}\sigma$. The internal variables are summarized in a vector $\alpha \in L^2(\Omega; \mathbb{R}^m)$ of $m$ components and lead to the generalized stress and strains linked, in each time-step, by some variational inequality written

$$(p, \alpha) \in \partial j(\sigma, \alpha) \quad \text{or} \quad (\sigma, \alpha) \in \partial j^*(p, \alpha) \tag{7.2}$$

for some dissipation functional $j : \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ and its dual $j^*$ and subgradient $\partial j$.

The list of examples [3,7,12,22,32] includes isotropic hardening, kinematic hardening, combined isotropic and kinematic hardening, perfect plasticity, and viscoplasticity. Since internal variables and the plastic strain are local variables (with a history separately for each material point), they can be eliminated. With some given data $K \in L^\infty(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}})$, $\sigma_Y > 0$ and $\eta \geqslant 0$, this leads to (6.3) with an energy density

$$W(A) = \frac{1}{2} A : \mathbb{C}A - \frac{1}{4\mu(1+\eta)} \max\bigl\{0, \left|\mathrm{dev}(A - K)\right| - \sigma_Y\bigr\}^2 \quad \text{for all } A \in \mathbb{R}^{d\times d}_{\mathrm{sym}}. \tag{7.3}$$

The crucial parameter is $\eta \geqslant 0$. If $\eta$ is positive there hold (6.1), (6.2) and hence Theorem 6.1 applies and proves the energy-reduction property (1.3). However, the constant $\varrho$ depends on the small parameter $\eta$ in a very crucial way: $\lim_{\eta\to 0} \varrho = 1$. For $\eta = 0$, the functional analytical frame is no longer correct [23,25,34]—the displacements belong to $\mathrm{BD}(\Omega)$ rather than to $H^1(\Omega; \mathbb{R}^d)$; see [34,35] for details.

Since the following result has some overlap with the results of the energy reduction of [17], we solely outline an alternative technique for the proof of convergence based on [11].

**Theorem 7.1.** *If $u - u_\ell$ remain bounded in $V$, (1.5) holds.*

**Proof.** From [11, Proposition 7.7], we recall

$$\frac{1}{2}\left|DW(A) - DW(B)\right|^2_{\mathbb{C}^{-1}} \leqslant W(B) - W(A) - DW(A) : (B - A) \quad \text{for all } A, B \in \mathbb{R}^{d\times d}_{\mathrm{sym}}. \tag{7.4}$$

With $A = \varepsilon(u_{\ell+1})$ and $B = \varepsilon(u_\ell)$ in (7.4) there follows

$$\frac{1}{2} |||\sigma_{\ell+1} - \sigma_\ell|||^2 \leqslant \delta_\ell - \delta_{\ell+1}.$$

The point is that $\cdots \subset V_\ell \subset V_{\ell+1} \subset \cdots$ implies that $\delta_\ell$ is a Cauchy sequence and so

$$\lim_{\ell\to\infty} |||\sigma_{\ell+1} - \sigma_\ell||| = 0.$$

As before in (5.2), (5.3), one derives an upper bound of $\|R_\ell\|_{V^*}$ in terms of $|||\sigma_{\ell+1} - \sigma_\ell|||$ and $\mathrm{OSC}_\ell$. Since $\lim_{\ell\to\infty} \mathrm{OSC}_\ell = 0$, one deduces

$$\lim_{\ell\to 0} \|R_\ell\|_{V^*} = 0.$$

This and another application of (7.4) conclude the proof:

$$|||\sigma - \sigma_\ell|||^2 \leqslant \int_\Omega (\sigma - \sigma_\ell) : \varepsilon(u - u_\ell)\, dx = R_\ell(u - u_\ell)$$
$$\leqslant \|R_\ell\|_{V^*}\|u - u_\ell\|_V \leqslant C_8\|R_\ell\|_{V^*}. \qquad \square$$

The dual formulation of elastoplasticity is studied even more frequently in applied mechanics. The variables are the generalised stresses $\Sigma := (\sigma, \chi)$ in some convex set $K$, plus displacements $u$ in the continuous formulation

$$\int_\Omega \big(\big(\varepsilon(u), 0\big) - (\mathbb{A}\Sigma)\big) : (\Upsilon - \Sigma)\, dx \leqslant 0 \quad \text{for all } \Upsilon \in L^2(\Omega; K)$$

and $f + \operatorname{div}\sigma = 0$ in $\Omega$ plus the boundary conditions from (3.4). The discrete analog seeks $\Sigma_\ell \in P_0(\mathcal{T}_\ell; K)$ and $u_\ell \in V_\ell$ with

$$\int_\Omega \big(\big(\varepsilon(u_\ell), 0\big) - (\mathbb{A}\Sigma_\ell)\big) : (\Upsilon_\ell - \Sigma_\ell)\, dx \leqslant 0 \quad \text{for all } \Upsilon_\ell \in P_0(\mathcal{T}_\ell; K)$$

and

$$\int_\Omega \sigma_\ell : \varepsilon(v_\ell)\, dx = \int_\Omega f \cdot v_\ell\, dx + \int_{\Gamma_N} g \cdot v_\ell\, ds \quad \text{for all } v_\ell \in V_\ell.$$

The convergence result and its direct proof is new and innovative.

**Theorem 7.2.** *If $u - u_\ell$ remain bounded in $V$, (1.5) holds.*

**Proof.** The combination of the aforementioned variational inequalities with $\Upsilon_\ell = \Sigma_\ell^*$ defined as the piecewise integral mean of $\Sigma$ and $\Upsilon := \Sigma_\ell$ leads to

$$|||\Sigma - \Sigma_\ell|||^2 := \int_\Omega \mathbb{A}(\Sigma - \Sigma_\ell) : (\Sigma - \Sigma_\ell)\, dx$$
$$\leqslant \int_\Omega \varepsilon(u - u_\ell) : (\sigma - \sigma_\ell)\, dx + \int_\Omega (\mathbb{A}\Sigma_\ell) : (\Sigma_\ell^* - \Sigma)\, dx - \int_\Omega \varepsilon(u_\ell) : (\sigma_\ell^* - \sigma)\, dx. \tag{7.5}$$

Notice that all coefficients are (piecewise) constant and hence

$$\int_\Omega \mathbb{A}\Sigma_\ell : (\Sigma_\ell^* - \Sigma)\, dx = 0.$$

Since $\varepsilon(v_\ell) \in P_0(\mathcal{T}_\ell; \mathbb{R}^{d\times d}_{\mathrm{sym}})$,

$$\int_\Omega \varepsilon(u_\ell) : (\sigma_\ell^* - \sigma)\, dx = 0.$$

Moreover, the strong and the weak form of the equilibrium conditions result into

$$\int_\Omega \varepsilon(u - u_\ell) : (\sigma - \sigma_\ell)\, dx = R_\ell(u - u_\ell).$$

The previous three identities recast (7.5) into

$$|||\Sigma - \Sigma_\ell|||^2 = R_\ell(u - u_\ell).$$

This and the boundedness of $\|u - u_\ell\|_V$ yield with Theorem 4.1 that

$$|||\Sigma - \Sigma_\ell|||^2 \leqslant C_9\eta_\ell + C_{10}\,\mathrm{OSC}_\ell\,. \tag{7.6}$$

The aforementioned arguments also prove uniqueness of the continuous stress field. This uniqueness will be used below.

Since the oscillations vanish in the limit as $\ell \to \infty$, it remains to prove that $(\eta_\ell)_\ell$ tends to zero as well. However, there is seems no argument left to do so although the equivalent primal problem allows some refined convexity control for that. Instead, we refine the analysis from [19] and prove: There exist some $C_{11} > 0$ and $0 < \varrho < 1$ such that, for all nonnegative integers $\ell$ and $m$, it holds

$$\eta_{\ell+m}^2 \leqslant \varrho\eta_\ell^2 + C_{11}\|\sigma_{\ell+m} - \sigma_\ell\|^2_{L^2(\Omega)}. \tag{7.7}$$

Indeed, for all $E \in \mathcal{E}_\ell$ we have either $E \in \mathcal{E}_{\ell+m}$ or otherwise there exist $E_1, \ldots, E_J \in \mathcal{E}_{\ell+m}$ with $E = E_1 \cup \cdots \cup E_J$ and $J \geqslant 2$. In the second case $E \in \mathcal{E}_\ell \setminus \mathcal{E}_{\ell+m}$, for any $0 < \delta < \theta/(2-\theta)$,

$$\begin{aligned}\sum_{j=1}^{J}\eta_{E_j}^{(\ell+m)2} &= \sum_{j=1}^{J} h_{E_j}^2 \big|[\sigma_{\ell+m}]\cdot\nu_{E_j}\big|^2\\ &\leqslant \sum_{j=1}^{J} h_{E_j}^2\Big((1+\delta)\big|[\sigma_\ell\cdot\nu_{E_j}]\big|^2 + \Big(1+\frac{1}{\delta}\Big)\big|[\sigma_{\ell+m}-\sigma_\ell]\cdot\nu_{E_j}\big|^2\Big)\\ &\leqslant \frac{1+\delta}{2}\eta_E^{(\ell)2} + \Big(1+\frac{1}{\delta}\Big)\sum_{j=1}^{J} h_{E_j}^2\big|[\sigma_{\ell+m}-\sigma_\ell]\cdot\nu_{E_j}\big|^2.\end{aligned}$$

Notice that the factor 1/2 results from $J>1$ refinements (at least one bisection) of $E\in\mathcal{E}_\ell\backslash\mathcal{E}_{\ell+m}$. Therefore,

$$\sum_{\substack{E\in\mathcal{E}_{\ell+m}\\ E\subseteq\bigcup\mathcal{E}_\ell}}\eta_E^{(\ell+m)2} \leqslant \frac{1+\delta}{2}\sum_{E\in\mathcal{E}_\ell\backslash\mathcal{E}_{\ell+m}}\eta_E^{(\ell)2} + (1+\delta)\sum_{E\in\mathcal{E}_\ell\cap\mathcal{E}_{\ell+m}}\eta_E^{(\ell)2} + \Big(1+\frac{1}{\delta}\Big)\sum_{\substack{E\in\mathcal{E}_{\ell+m}\\ E\subseteq\bigcup\mathcal{E}_\ell}} h_E^2\big|[\sigma_{\ell+m}-\sigma_\ell]\cdot\nu_E\big|^2.$$

For any $E\in\mathcal{E}_{\ell+m}$ with $E\nsubseteq\bigcup\mathcal{E}_\ell$, $[\sigma_\ell]\cdot\nu_E=0$ along $E$. Hence

$$\eta_E^{(\ell+m)2} = h_E^2\big|[\sigma_{\ell+m}-\sigma_\ell]\cdot\nu_E\big|^2.$$

Therefore,

$$\eta_{\ell+m}^2 \leqslant \frac{1+\delta}{2}\sum_{E\in\mathcal{E}_\ell\backslash\mathcal{E}_{\ell+m}}\eta_E^{(\ell)2} + (1+\delta)\sum_{E\in\mathcal{E}_\ell\cap\mathcal{E}_{\ell+m}}\eta_E^{(\ell)2} + \Big(1+\frac{1}{\delta}\Big)\sum_{E\in\mathcal{E}_{\ell+m}} h_E^2\big|[\sigma_{\ell+m}-\sigma_\ell]\cdot\nu_E\big|^2.$$

Since $\sigma_{\ell+m}-\sigma_\ell$ is piecewise constant with respect to the shape regular triangulation $\mathcal{T}_{\ell+m}$,

$$h_E^2\big|[\sigma_{\ell+m}-\sigma_\ell]\cdot\nu_E\big|^2 \leqslant C_{12}\|\sigma_{\ell+m}-\sigma_\ell\|_{L^2(\omega_E)}$$

for the edge patch $\omega_E$ of $E$ in $\mathcal{T}_{\ell+m}$. Since there is only a finite overlap of all edge patches,

$$\eta_{\ell+m}^2 \leqslant \frac{1+\delta}{2}\sum_{E\in\mathcal{M}_\ell}\eta_E^{(\ell)2} + (1+\delta)\sum_{E\in\mathcal{E}_\ell\backslash\mathcal{M}_\ell}\eta_E^{(\ell)2} + C_{11}\|\sigma_{\ell+m}-\sigma_\ell\|_{L^2(\Omega)}^2.$$

The bulk criterion leads to

$$\frac{1}{2}\sum_{E\in\mathcal{M}_\ell}\eta_E^{(\ell)2} + \sum_{E\in\mathcal{E}_\ell\backslash\mathcal{M}_\ell}\eta_E^{(\ell)2} = \eta_\ell^2 - \frac{1}{2}\sum_{E\in\mathcal{M}_\ell}\eta_E^{(\ell)2} \leqslant \Big(1-\frac{\theta}{2}\Big)\eta_\ell^2$$

and so to

$$\eta_{\ell+m}^2 \leqslant (1+\delta)\Big(1-\frac{\theta}{2}\Big)\eta_\ell^2 + C_{12}\|\sigma_{\ell+m}-\sigma_\ell\|_{L^2(\Omega)}^2.$$

Since $\delta<\theta/(2-\theta)$, the resulting estimate proves the assertion (7.7).

Repeating the argument for (7.5) with $(\Sigma,u)$ replaced by $(\Sigma_\ell,u_\ell)$, we deduce

$$|||\Sigma_{\ell+m}-\Sigma_\ell|||^2 \leqslant \int_\Omega (\sigma_{\ell+m}-\sigma_\ell):\varepsilon(u_{\ell+m}-u_\ell)\,dx.$$

The discrete equilibrium leads to

$$\int_\Omega (\sigma_{\ell+m}-\sigma_\ell):\varepsilon(u_{\ell+m}-u_\ell)\,dx = \int_\Omega \big(f\cdot(u_{\ell+m}-u_\ell) - \sigma_\ell:\varepsilon(u_{\ell+m}-u_\ell)\big)\,dx.$$

Hence,

$$|||\Sigma_{\ell+m}-\Sigma_\ell|||^2 \leqslant \int_\Omega \big(f\cdot(u_{\ell+m}-u_\ell) - \sigma_\ell:\varepsilon(u_{\ell+m}-u_\ell)\big)\,dx. \tag{7.8}$$

The generalized discrete stresses are bounded and so are $u_\ell$ by assumption. With Rellich's compactness embedding, there exists some weakly convergent subsequence

$$(\Sigma_{\ell_j}) \rightharpoonup \Sigma_\infty, \quad \big(\varepsilon(u_{\ell_j})\big) \rightharpoonup \varepsilon(u_\infty) \quad \text{and} \quad \|u_\infty - u_{\ell_j}\|_{L^2(\Omega)} \to 0 \quad \text{as } j\to\infty$$

for some limit $\Sigma_\infty\in L$ and $u_\infty\in V$. Notice that $|||\cdot-\Sigma_\ell|||$ is sequentially weakly lower semicontinuous and so the above weak convergence and estimate (7.8) lead, for $m\to\infty$, to

$$|||\Sigma_\infty-\Sigma_\ell|||^2 \leqslant \int_\Omega \big(f\cdot(u_\infty-u_\ell) - \sigma_\ell:\varepsilon(u_\infty-u_\ell)\big)\,dx = R_\ell(u_\infty-u_\ell).$$

Since $V_\ell \subset \ker R_\ell$,

$$|||\Sigma_\infty - \Sigma_\ell|||^2 \leqslant R_\ell(u_\infty).$$

It is clear that, for all $\ell = 0, 1, 2, \ldots,$

$$V_\ell \subseteq \ker R_\infty \quad \text{for } R_\infty(v) := \int_\Omega \big(f \cdot v - \sigma_\infty : \varepsilon(v)\big)\,dx \quad \text{with } v \in V.$$

Hence,

$$u_\infty \in V_\infty := \overline{\bigcup_{\ell=0}^{\infty} V_\ell} \subseteq \ker R_\infty$$

and so

$$\lim_{j\to\infty} R_{\ell_j}(u_\infty) = R_\infty(u_\infty) = 0.$$

Consequently,

$$\lim_{j\to\infty} |||\Sigma_\infty - \Sigma_{\ell_j}||| = 0$$

and $(\Sigma_{\ell_j})$ is a Cauchy sequence. This and (7.7) plus some elementary analysis show that

$$\lim_{j\to\infty} \eta_{\ell_j} = 0.$$

With (7.6) and the convergence of the oscillation terms, this results in $\Sigma_\infty = \Sigma \in L^2(\Omega; K)$ and the strong convergence of *subsequences*,

$$\lim_{j\to\infty} |||\Sigma - \Sigma_{\ell_j}||| = 0.$$

Since the limit $\Sigma$ is unique and every subsequence converges contains a further subsequence with this limit, we deduce convergence of the full sequence,

$$\lim_{\ell\to\infty} |||\Sigma - \Sigma_\ell||| = 0. \qquad \square$$

## Acknowledgements

The work was supported by the Austrian Science Fund (FWF) under Projects P15274 and P16461 and by the German Research Foundation (DFG) under C13 in the Research Center MATHEON. The latest results on elastoplasticity were accomplished during a happy research stay at the Hausdorff Institute for Mathematics in Bonn, Germany, in Spring 2008.

## References

[1] M. Ainsworth, J.T. Oden, A Posteriori Error Estimation in Finite Element Analysis, Wiley-Interscience (John Wiley & Sons), New York, 2000.
[2] J. Alberty, C. Carstensen, S.A. Funken, R. Klose, Matlab implementation of the finite element method in elasticity, Computing 69 (2002) 239–263.
[3] J. Alberty, C. Carstensen, D. Zarrabi, Adaptive numerical analysis in primal elastoplasticity with hardening, Comput. Methods Appl. Mech. Engrg. 171 (1999) 175–204.
[4] I. Babuška, T. Strouboulis, The Finite Element Method and its Reliability, The Clarendon Press, Oxford University Press, 2001, xii+802.
[5] I. Babuška, M. Vogelius, Feedback and adaptive finite element solution of one-dimensional boundary value problems, Numer. Math. 44 (1984) 75–102.
[6] P. Binev, W. Dahmen, R. DeVore, Adaptive finite element methods with convergence rates, Numer. Math. 97 (2) (2004) 219–268.
[7] C. Carstensen, Domain decomposition for a non-smooth convex minimisation problem and its application to plasticity, Numer. Linear Algebra Appl. 4 (3) (1997) 1–13.
[8] C. Carstensen, Numerical analysis of primal elastoplasticity with hardening, Numer. Math. 82 (1999) 577–597.
[9] C. Carstensen, Quasi-interpolation and a posteriori error analysis in finite element method, M2AN Math. Model. Numer. Anal. 33 (1999) 1187–1202.
[10] C. Carstensen, An adaptive mesh-refining algorithm allowing for an $H^1$-stable $L^2$-projection onto Courant finite element spaces, Constr. Approx. 20 (4) (2004) 549–564.
[11] C. Carstensen, Convergence of an adaptive FEM for a class of degenerate convex minimization problems, IMA J. Numer. Anal. 28 (2008) 423–439.
[12] C. Carstensen, J. Alberty, Averaging techniques for reliable a posteriori FE-error control in elastoplasticity with hardening, Comput. Meth. Appl. Mech. Engrg. 192 (2003) 1435–1450.
[13] C. Carstensen, J. Bolte, Adaptive finite element mesh-refining algorithm for $L^2$-error control, in: W.-B. Liu, M. Ng, Z.-C. Shi (Eds.), Recent Progress in Scientific Computing, Science Press, 2007, pp. 13–39.
[14] C. Carstensen, R. Klose, Elastoviscoplastic finite element analysis in 100 lines of Matlab, J. Numer. Math. 10 (3) (2002) 157–192.
[15] C. Carstensen, Y. Kondratyuk, Optimal adaptive finite element methods for uniformly convex minimisation problems (2009), in preparation.
[16] C. Carstensen, A. Orlando, J. Hu, Framework for the a posteriori error analysis of nonconforming finite element, SIAM J. Numer. Anal. 45 (1) (2007) 68–82.
[17] C. Carstensen, A. Orlando, J. Valdman, A convergent adaptive finite element method for the primal problem of elastoplasticity, Int. J. Numer. Meth. Engrg. 67 (13) (2006) 1851–1887.

[18] C. Carstensen, R. Verfürth, Edge residuals dominate a posteriori error estimates for low order finite element methods, SIAM J. Numer. Anal. 36 (5) (1999) 1571–1587.
[19] J.M. Cascon, C. Kreuzer, R.H. Nochetto, K.G. Siebert, Quasi-optimal convergence rate for an adaptive finite element method, SIAM J. Numer. Anal. 46 (5) (2008) 2524–2550.
[20] W. Dörfler, A convergent adaptive algorithm for Poisson's equation, SIAM J. Numer. Anal. 33 (3) (1996) 1106–1124.
[21] K. Eriksson, D. Estep, P. Hansbo, C. Johnson, Computational Differential Equations, Cambridge University Press, Cambridge, 1995.
[22] W. Han, B.D. Reddy, Plasticity, Mathematical Theory and Numerical Analysis, Springer, New York, 1999.
[23] C. Johnson, Existence theorems for plasticity problems, J. Math. Pures Appl. 55 (1976) 431–444.
[24] C. Johnson, A mixed finite element method for plasticity problems with hardening, SIAM J. Numer. Anal. 14 (1977) 575–583.
[25] C. Johnson, On plasticity with hardening, J. Math. Anal. Appl. 62 (1978) 325–336.
[26] C. Johnson, P. Hansbo, Adaptive finite element methods in computational mechanics, Comput. Methods Appl. Mech. Engrg. 101 (1992) 143–181.
[27] P. Morin, R.H. Nochetto, K.G. Siebert, Convergence of adaptive finite element method, SIAM Rev. 44 (2003) 631–658.
[28] J. Necas, I. Hlavacek, Mathematical Theory of Elastic and Elasto-Plastic Bodies: An Introduction, Elsevier, New York, 1981.
[29] P. Neittaanmäki, S. Repin, Reliable Methods for Mathematical Modelling. Error Control and a Posteriori Estimates, Elsevier, New York, 2004.
[30] R. Rannacher, F.T. Suttmeier, A posteriori error control in finite element methods via duality techniques: Application to perfect plasticity, Comput. Mech. 21 (2) (1998) 123–133.
[31] R. Rannacher, F.T. Suttmeier, A posteriori error estimation and mesh adaptation for finite element models in elasto-plasticity, Comput. Methods Appl. Mech. Engrg. 176 (1–4) (1999) 333–361.
[32] J.C. Simo, T.J.R. Hughes, Computational Inelasticity, Springer-Verlag, New York, 1998.
[33] R. Stevenson, Optimality of a standard adaptive finite element method, Found. Comput. Math. 7 (2) (2007) 245–269.
[34] P.-M. Suquet, Discontinuities and plasticity, in: Nonsmooth Mechanics and Applications, in: CISM Courses, vol. 302, Springer-Verlag, New York, 1988, pp. 279–341.
[35] R. Temam, Mathematical Problems in Plasticity, Gauthier-Villars, Paris, 1985.
[36] A. Veeser, Convergent adaptive finite elements for the nonlinear Laplacian, Numer. Math. 92 (4) (2002) 743–770.
[37] R. Verfürth, A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques, Wiley–Teubner, New York, Leipzig, 1996.
[38] R. Verfürth, A review of a posteriori error estimation techniques for elasticity problems, Comput. Methods Appl. Mech. Engrg. 176 (1999) 419–440.
[39] E. Zeidler, Nonlinear Functional Analysis and its Applications III and IV, Springer-Verlag, New York, 1985 and 1988.