Error-Controlled Adaptive FEMs in Solid Mechanics
Error-Controlled Adaptive FEMs in Solid Mechanics

Edited by E. STEIN

JOHN WILEY & SONS
Chichester • New York • Brisbane • Toronto • Singapore
## Contents

4 Error Estimation and Adaptive Mesh Design for FE Models in Elasto-Plasticity ........................................ 1
  4.1 INTRODUCTION ............................................. 1
  4.2 THE DUAL-WEIGHTED-RESIDUAL APPROACH ................. 4
    4.2.1 A Model Problem ...................................... 5
    4.2.2 Evaluation of the a posteriori error bounds .......... 11
    4.2.3 Strategies for mesh adaptation ...................... 13
    4.2.4 Examples ............................................ 14
  4.3 THE VARIATIONAL FRAMEWORK ............................ 16
    4.3.1 The Hencky and Prandtl-Reuss models .............. 16
    4.3.2 Time-discretization in the Prandtl-Reuss model .... 18
    4.3.3 Treatment of almost incompressible behaviour ....... 21
    4.3.4 An elasto-plastic benchmark ....................... 21
  4.4 THE LINEAR-ELASTIC CASE ................................ 23
    4.4.1 A posteriori error analysis ......................... 23
    4.4.2 Numerical test for the linear-elastic case .......... 25
  4.5 THE STATIC HENCKY-TYPE MODEL ........................ 27
    4.5.1 A posteriori error estimates ........................ 27
    4.5.2 Decoupled displacement-pressure discretization .... 32
    4.5.3 Numerical tests for the static Hencky model ....... 33
  4.6 THE QUASI-STATIC PRANDTL-REUSS MODEL ............... 38
    4.6.1 A posteriori error analysis for spatial discretization . 38
    4.6.2 A posteriori error analysis for time discretization .. 40
    4.6.3 Numerical tests for the quasi-static Prandtl-Reuss model . 43
  4.7 CONCLUSION ............................................... 48
REFERENCES ..................................................... 48
References ..................................................... 48
Error Estimation and Adaptive Mesh Design for FE Models in Elasto-Plasticity

ROLF RANNACHER\(^{(1)}\) & FRANZ-THEO SÜTTMEIER\(^{(2)}\)

\(^{(1)}\) Institut für Angewandte Mathematik, Universität Heidelberg
INF 293, D-69120 Heidelberg, Germany
E-Mail: rannacher@iwr.uni-heidelberg.de

\(^{(2)}\) Fachbereich Mathematik, Universität Dortmund, Lehrstuhl X
Vogelpothsweg 87, D-44221 Dortmund, Germany
E-Mail: suttmeier@math.uni-dortmund.de

4.1 INTRODUCTION

This work introduces a general approach to \textit{a posteriori} error estimation and adaptive mesh design for finite element models in elasto-plasticity. This so-called "Dual-Weighted-Residual" method ("DWR" method) is based on a variational formulation of the problem and uses global duality arguments for deriving weighted \textit{a posteriori} error estimates with respect to arbitrary functionals of the error. In these estimates local residuals of the computed solution are multiplied by sensitivity factors which are obtained from a numerically computed "dual solution". The resulting local error indicators are used in a feed-back process for generating economical meshes which are tailored according to the particular goal of the computation. This method is developed here for the Hencky and Prandtl-Reuss models in perfect plasticity. It is not restricted to this kind of problem but provides a systematic way of adaptive error control for general variational problems in mechanics.
Solving complex systems of partial differential equations by discretization methods may be considered in the context of model reduction: a conceptually infinite dimensional model is approximated by a finite dimensional one. Here, the quality of the approximation depends on the proper choice of the discretization parameters, e.g., the mesh width, the polynomial degree of the trial functions, and the size of certain stabilization parameters. As the result of the computation, we obtain an approximation to the desired output quantity of the simulation and besides that certain accuracy indicators like cell-residuals. Controlling the error in such an approximation of a continuous model of a physical system requires to determine the influence factors for the local error indicators on the target quantity. Such a sensitivity analysis with respect to local perturbations of the model is common in optimal control theory and introduces the concept of a dual (or adjoint) problem.

For illustration consider a continuous model governed by a differential operator $L$ and a forcing term $f$,

$$Lu = f,$$  \hspace{0.5cm} (4.1.1)

and an approximating discrete model,

$$L_h u_h = f_h,$$  \hspace{0.5cm} (4.1.2)

depending on a parameter $h > 0$. In controlling this discretization, we have to detect the interplay of the various error propagation effects in order to achieve (i) a posteriori error control, i.e., control of the error in quantities of physical interest like point values of deflection and stress or mean stresses along parts of the boundary, etc., and (ii) solution-adapted meshing, i.e., design of economical meshes for computing these quantities with best efficiency. Our error analysis is based on the computable \textquotedblleft residual\textquotedblright $\varrho(u_h) = f - L_h u_h$ which is well defined in the context of a Galerkin finite element method.

Traditionally, a posteriori error estimation in Galerkin finite element methods is done with respect to the natural energy norm $\| \cdot \|_E$ induced by the underlying differential operator,

$$\| u - u_h \|_E \leq c_S \sup_{\| \varphi \|_E = 1} |(\varrho(u_h), \varphi)|.$$  \hspace{0.5cm} (4.1.3)

This approach has been initiated by the landmark papers of Babuska&Reinboldt [BR87], Bank&Weiser [BW85], and Zienkiewicz&Zhu [ZZ87]. For discussions and more references see the survey articles by Ainsworth&Oden [AO97] and Verfürth [Ver96]. This approach seems rather generic as it is directly based on the variational formulation of the problem and allows to exploit its coercivity properties. However, in most applications the error in the energy norm does not provide a useful bound on the error in the quantities of real physical interest. A more versatile method for a posteriori error estimation with respect to relevant error measures like $L^2$ norm over subdomains, point values, line averages, etc., is obtained by using duality arguments as common from the a priori error analysis of finite element methods \textquotedblleft Aubin-Nitsche trick\textquotedblright. This approach is particularly designed for achieving high solution accuracy at minimum computational cost. The additional work required by the evaluation of the error bounds is acceptable since, particularly in nonlinear cases, it usually amounts to only a small fraction of the total cost.
Let $J(u)$ be a quantity of physical interest derived from the solution $u$ by applying an "output functional" $J(\cdot)$. The goal is to control the error of the discretization with respect to this functional output, i.e., $E(u_h) = J(u) - J(u_h)$ in terms of the computable cell residuals $\varrho_K(u_h)$. An example is control of the total error $e_K = u - u_{MK}$ in some mesh cell $K$. By superposition, $e_K$ splits into two components, the locally produced truncation error and the globally transported pollution error, $e_K^{\text{tot}} = e_K^{\text{loc}} + e_K^{\text{trans}}$, assuming for simplicity that the underlying problem is linear. This asks for control of

- error propagation in space (global pollution effect),
- interaction of physical error sources (local sensitivity analysis).

The effect of the cell residual $\varrho_K$ on the local error $e_{K'}$, at another cell $K'$, is essentially governed by the Green’s function of the continuous problem. Capturing this dependence by numerical evaluation is the general philosophy underlying our approach to error control. In practice it is mostly impossible to determine the complex error interaction by analytical means, it rather has to be detected by computation. This automatically leads to a feed-back process in which error estimation and mesh adaptation goes hand-in-hand leading to economical discretization for computing the quantities of interest.

This approach has first been systematically developed by C. Johnson and his co-workers [EJ88, Joh93, EEH95] and was then extended by the first author and his group [BR96, BR96] to a practical feedback method, the DWR method, for mesh optimization (see also Rannacher [Ran99] for a survey of this method). The same approach has been applied in Rannacher&Suttméier [RS96] for primal as well as dual-mixed finite element methods in linear elasticity and in Rannacher&Suttméier [Sut96, RS97, RS99] to the Hencky and Prandtl-Reuss model in perfect plasticity. The DWR method yields weighted a posteriori error bounds with respect to arbitrary functionals

$$|J(u) - J(u_h)| \leq \langle \varrho(u_h), \omega(z) \rangle,$$

where the weight $\omega(z)$ is obtained by approximately solving a dual problem $L^*z = J$. The dual solution $z$ may be viewed as a generalized Green function with respect to the output functional $J(\cdot)$, and accordingly the weight $\omega(z)$ describes the effect of local variations of the residual $\varrho(u_h)$ on the error quantity $J(u) - J(u_h)$, for instance as the consequence of mesh adaptation.

**Overview:** This contribution is organized as follows:

In Section 2, we introduce the main principles of the DWR method for a posteriori error estimation using mainly material from Becker&Rannacher [BR96]. This will be at first illustrated for simple model problems. Further, we discuss strategies for mesh adaptation based on the "weighted" a posteriori error estimates.

In Section 3, following Rannacher&Suttméier [RS97], we introduce various forms of the well-known Hencky and Prandtl-Reuss models of (linear elastic) perfect plasticity which provide the basis for a finite element discretization. We consider the generic primal-mixed formulation as a (linear) variational inequality and derive from that a primal formulation as a (nonlinear) variational equation which is particularly suitable for applying the DWR method. The quasi-stationary Prandtl-Reuss model is discretized in time by the backward Euler scheme resulting in a sequence of stationary Hencky-like problems. Further, following Suttméier [Sut00], the treatment of incompressible
behaviour by introducing a pseudo-pressure is discussed. Finally, we specify a two-
dimensional elasto-plastic benchmark problem “disc with a hole” for which the
methods presented in this work will be tested.

In Section 4, we discuss the use of the DWR method for solving the Lamé-Navier
equations in linear elasticity using material from Rannacher& Suttmeier [RS96]. The
special features of goal-oriented error control is illustrated by numerical examples.

In Section 5, we extend the DWR method to static Hencky-type models and report
on extensive test calculations for the benchmark “disc with a hole”. A comparison
is also made of the primal velocity-based and the mixed velocity/pressure-based
formulation. This material is taken from Rannacher&Suttmeier [RS97] and Suttmeier
[Sut00].

In Section 6, we further extend the DWR method to incorporate also time stepping
in the quasi-stationary Prandl-Reuss model. An a posteriori error estimate is derived
for the simultaneous discretization of the loading process combined with spatial
discretization following Rannacher&Suttmeier [RS99]. Finally, the resulting algorithm
is applied to the nonstationary benchmark “disc with a hole”. This section contains
some material which has not been published before.

All numerical computations presented in this work have been done using the object-
oriented finite element software package DEAL (see Becker, Kanschat&Suttmeier
[BKS96]).

4.2 THE DUAL-WEIGHTED-RESIDUAL APPROACH

We illustrate the basic principles underlying the DWR method for error estimation by
considering perturbations of algebraic systems. At first, for given matrices $A, A_h \in \mathbb{R}^{n \times n}$ and right-hand sides $b, b_h \in \mathbb{R}^n$, consider the linear system $Ax = b$ together
with its perturbation $A_h x_h = b_h$. Let $(\cdot, \cdot)$ denote the euclidean inner product and
$\| \cdot \|$ the corresponding norm on $\mathbb{R}^n$ and $\mathbb{R}^{n \times n}$.

For estimating the error $e = x - x_h$, we use the residual $g = b - A x_h = A e$, to obtain

$$ e = A^{-1} g \Rightarrow \| e \| \leq c_s \| g \|, \quad (4.2.1) $$

with the continuous stability constant $c_s = \|A^{-1}\|$. Alternatively, we may use the
solution $z$ of the dual problem $A^* z = \| e \|^{-1} e$, to obtain

$$ \| e \| = (e, A^* z) = (b - A x_h, z) = (g, z) \leq \| g \| \| z \| \leq c^*_s \| g \|, \quad (4.2.2) $$

with the dual stability constant $c^*_s = \|A^{*-1}\|$. Of course, this approach does not yield
a new result in estimating the error in the $l_2$-norm. But it also gives us the possibility
to bound other error quantities, e.g., error moments $(\sigma, e)$ with a given $\sigma \in \mathbb{R}^d$. With
the solution of the dual problem $A^* z = \sigma$, there holds the weighted a posteriori error
estimate

$$ |(\sigma, e)| \leq \sum_{i=1}^n |g_i z_i|. \quad (4.2.3) $$

An analogous argument can also be applied in the case of a nonlinear equation. Let
$A, A_h : \mathbb{R}^n \to \mathbb{R}^n$ be (differentiable) vector functions and solve $A(x) = b$ and its
perturbation $A_h(x_h) = b_h$. Then, for the residual $\varrho = b - A(x_h)$, there holds by Taylor expansion

$$
\varrho = A(x) - A(x_h) = A'(x)e - \int_0^1 A''(x_h + se)e e s \, ds.
$$

(4.2.4)

From this, we obtain the error estimate

$$
\|e\| \leq c_s \|\varrho\| + \frac{1}{2} c_s \max_{y \in [x, x_h]} \|A''(y)\| \|e\|^2,
$$

(4.2.5)

with the stability constant $c_s = \|A'(u)^{-1}\|$. Alternatively, for a given weight vector $\sigma \in \mathbb{R}^d$, we may consider the linearized dual problem

$$
A'(x_h)^* z = \sigma.
$$

Then, observing that

$$
(\varrho, z) = (A(x) - A(x_h), z) = (A'(x_h) e, z) - \int_0^1 (A''(x_h + se)e, e')(z(1-s)) \, ds,
$$

there holds the error representation

$$
(e, \sigma) = (A'(x_h) e, z) = (\varrho, z) - \int_0^1 (A''(x_h + se)e, e')(z(1-s)) \, ds.
$$

This results in the weighted error bound

$$
|(e, \sigma)| \leq \sum_{i=1}^n |e_i z_i| + \frac{1}{2} \max_{y \in [x, x_h]} \|A''(y)\| \|e\|^2 |z|.
$$

Assuming $\|e\|^2$ to be sufficiently small, the remainder term may be neglected and we obtain a weighted error estimate analogous to the linear case. Below, we will use this duality technique for generating a posteriori error estimates for the finite element approximation of elasto-plasticity problems.

4.2.1 A Model Problem

For illustration, we consider the model problem

$$
-\Delta u = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial \Omega,
$$

(4.2.6)

on a polygonal domain $\Omega \subset \mathbb{R}^2$. Below, we will need some notation from the theory of functions spaces which is provided in the following subsection. Readers who are familiar with the usual Lebesgue- and Sobolev-space notation may want to skip this and continue with the next subsection on finite element approximation.

Function spaces notation

For an open set $Q \subset \mathbb{R}^d$, we denote by $L^2(Q)$ the Lebesgue space of square-integrable functions defined on $Q$ which is a Hilbert space provided with the scalar product and norm

$$
(v, w)_Q = \int_Q v w \, dx, \quad \|v\|_Q = \left( \int_Q |v|^2 \, dx \right)^{1/2}.
$$
Analogously, $L^2(\partial Q)$ denotes the space of square-integrable functions defined on the boundary $\partial Q$ equipped with the inner product and norm
\[ (v, w)_{\partial Q} = \int_{\partial Q} v \, w \, ds, \quad \|v\|_{\partial Q} = \left( \int_{\partial Q} |v|^2 \, ds \right)^{1/2}. \]

The Sobolev spaces $H^1(Q)$ and $H^2(Q)$ consist of those functions $v \in L^2(Q)$ which possess first- and second-order (distributional) derivatives $\nabla v \in L^2(Q)^d$ and $\nabla^2 v \in L^2(Q)^{d \times d}$, respectively. For functions in these spaces, we use the semi-norms
\[ \|v\|_Q = \left( \int_Q |v|^2 \, dx \right)^{1/2}, \quad \|\nabla v\|_Q = \left( \int_Q |\nabla v|^2 \, dx \right)^{1/2}. \]

The space $H^1(Q)$ is continuously embedded in the space $L^2(\partial Q)$, such that for each $v \in H^1(Q)$ there exists a trace $\nu_{Q\partial} \in L^2(\partial Q)$. In this notation, we do not distinguish between the function $v$ on $Q$ and its trace on $\partial Q$. Further, the functions in the subspace $H^1_0(Q) \subset H^1(Q)$ are characterized by the property $\nu_{Q\partial} = 0$. By the Poincaré inequality,
\[ \|v\|_Q \leq c \|\nabla v\|_Q, \quad v \in H^1(Q), \tag{4.2.7} \]
the $H^1$-semi-norm $\|\nabla v\|_Q$ is a norm on the subspace $H^1_0(Q)$. If the set $Q$ is the set $\Omega$ on which the differential equation is posed, we usually omit the subscript $\Omega$ in the notation of norms and scalar products, e.g., $\|v\| = \|v\|_\Omega$. All the above notation will be synonymously used for vector- or matrix-valued functions $v : \Omega \to \mathbb{R}^d$ or $\mathbb{R}^{d \times d}$.

Variational formulation and finite element approximation

The natural solution space for the boundary value problem (4.2.6) is the Sobolev space $V = H^1_0(\Omega)$ defined above. The variational formulation of (4.2.6) seeks $u \in V$, such that
\[ (\nabla u, \nabla \varphi) = (f, \varphi) \quad \forall \varphi \in V. \tag{4.2.8} \]

The finite element approximation of (4.2.8) uses finite dimensional subspaces
\[ V_h = \{ v \in V : v|_K \in P(K), \ K \in T_h \}, \]
defined on decompositions $T_h$ of $\Omega$ into triangles or quadrilaterals $K$ (cells) of width $h_K = \text{diam}(K)$; we write $h = \max_{K \in T_h} h_K$ for the global mesh width. Here, $P(K)$ denotes a suitable space of polynomial-like functions defined on the cell $K \in T_h$. We will mainly consider low-order finite elements on quadrilateral meshes where $P(K) = Q_1(K)$ consists of shape functions which are obtained as usual via a bilinear transformation from the space of bivariate functions $Q_1(K) = \text{span}\{1, x_1, x_2, x_1x_2\}$ on the reference cell $\bar{K} = [0,1]^2$ (isoparametric bilinears). Local mesh refinement or coarsening is realized by using hanging nodes as indicated in Figure 4.1. The variable corresponding to such a hanging node is eliminated from the system by linear interpolation of neighboring variables in order to preserve the conformity of the global ansatz, i.e., $V_h \subset V$ (for more details we refer to Caley & Oden [CO84] and Suttmeier [Sut96]).
Now, the discrete problem determines $u_h \in V_h$ by
\[
(\nabla u_h, \nabla \varphi_h) = (f, \varphi_h) \quad \forall \varphi_h \in V_h.
\]
(4.2.9)
The error $e = u - u_h$ of this scheme satisfies the so-called Galerkin orthogonality
\[
(\nabla e, \nabla \varphi_h) = 0, \quad \varphi_h \in V_h.
\]
(4.2.10)
In order to convert (4.2.9) into an algebraic equation, one uses the "nodal basis" \{\psi_i^h, i = 1, \ldots, n = \text{dim} V_h\} of the finite element space $V_h$. Then, the coefficient vector $x = (x_i)_{i=1}^n$ in the expansion $u_h = \sum_{i=1}^n x_i \psi_i^h$ is determined by the linear system
\[
Ax = b,
\]
with the "stiffness matrix" $A = (a_{ij})_{i,j=1}^n$ and the "load vector" $b = (b_i)_{i=1}^n$ defined by $a_{ij} = (\nabla \psi_i^h, \nabla \psi_j^h)$ and $b_i = (f, \nabla \psi_i^h)$.

A PRIORI ERROR ANALYSIS

We begin with a brief discussion of the a priori error analysis for the scheme (4.2.9). By $i_h u \in V_h$, we denote the natural nodal interpolant of $u \in C(\bar{\Omega})$ satisfying $i_h u(P) = u(P)$ at all nodal points $P$. There holds (see, e.g., Ciarlet [Cia84] and Brenner&Scott [BS94]):
\[
\|u - i_h u\|_K + h^{1/2}_K \|u - i_h u\|_{\partial K} + h_K \|\nabla (u - i_h u)\|_K \leq c_i h^2_K \|\nabla^2 u\|_K,
\]
with some interpolation constant $c_i > 0$.

(i) By the projection property of the Galerkin finite element scheme the interpolation estimate (4.2.12) directly implies the energy-error estimate
\[
\|\nabla e\| = \inf_{\varphi_h \in V_h} \|\nabla (u - \varphi_h)\| \leq c_i h^2 \|\nabla^2 u\|.
\]
(4.2.13)

(ii) Further, employing a duality argument (so-called Aubin-Nitsche trick),
\[
-\Delta z = \|e\|^{-1} e \quad \text{in } \Omega, \quad z = 0 \quad \text{on } \partial \Omega,
\]
we obtain
\[
\|e\| = (e, -\Delta z) = (\nabla e, \nabla z) = (\nabla e, \nabla (z - i_h z) \leq c_i c_s h \|\nabla e\|,
\]
(4.2.15)
where the stability constant $c_s$ is defined by the a priori bound $\|\nabla^2 z\| \leq c_s$. Together with the energy-error estimate (4.2.13), this implies the improved $L^2$-error estimate
\[
\|e\| \leq c_i c_s h^2 \|\nabla^2 u\|.
\]
(4.2.16)
(iii) On the bases of the global error estimates (4.2.13) and (4.2.16), we can also obtain error bounds for various other quantities, e.g., in $\mathbb{R}^2$:

$$|e(a)| \approx (c + \log h^{-1}) \|\nabla e\|,$$  

$$|\nabla e(a)| \approx h^{-1}\|\nabla e\|, \quad |(e, \psi)_{\partial\Omega}| \approx h^{-1/2}\|e\|.$$  

All these estimates are only sub-optimal with respect to the power of $h$. This tells us that estimating the error in functional output usually requires special effort and cannot simply be reduced to the standard energy- or $L^2$-error estimates.

A posteriori error analysis

Next, we derive a posteriori error estimates. Let $J(\cdot)$ be an arbitrary (linear) output functional defined on $V$, and $z \in V$ the solution of the corresponding dual problem

$$(\nabla \varphi, \nabla z) = J(\varphi) \quad \forall \varphi \in V.$$  

(4.2.17)

Taking $\varphi = \varepsilon$ in (4.2.17) and using the Galerkin orthogonality, we obtain after cell-wise integration by parts the error representation

$$J(\varepsilon) = (\nabla \varepsilon, \nabla z) = (\nabla \varepsilon, \nabla (z - \varphi_h))$$  

$$= \sum_{K \in \mathcal{T}_h} \left\{ (-\Delta u + \Delta u_h, z - \varphi_h)_K - (u_n \nabla u_h, z - \varphi_h)_{\partial K} \right\}$$  

(4.2.18)

At this point, we have assumed that the domain $\Omega$ is polygonal (or polyhedral) in order to ease the approximation of the boundary $\partial\Omega$. In the presence of curved parts of $\partial\Omega$ the error formula (4.2.18) contains additional terms representing the error caused by the polygonal approximation of the boundary (see, e.g., Becker&Rannacher [BR96]). In the tests presented below these terms are suppressed.

The error representation (4.2.18) can be rewritten as

$$J(\varepsilon) = \sum_{K \in \mathcal{T}_h} \left\{ (R(u_h), z - \varphi_h)_K - (r(u_h), z - \varphi_h)_{\partial K} \right\},$$  

(4.2.19)

with an arbitrary $\varphi_h \in V_h$, and cell- and edge-residuals $R(u_h)$ and $r(u_h)$, respectively, defined by

$$R(u_h)_K = f + \Delta u_h, \quad r(u_h)_{\Gamma} := \begin{cases} \frac{1}{2} n \cdot [\nabla u_h], & \text{if } \Gamma \subset \partial K \setminus \partial\Omega, \\ n \nabla u_h, & \text{if } \Gamma \subset \partial\Omega, \end{cases}$$

where $[\nabla u_h]$ denotes the jump of $\nabla u_h$ across the interelement boundaries $\Gamma$. From the error identity (4.2.19), we can infer the following result.

**Theorem 4.2.1.** For the finite element approximation of the Poisson equation (4.2.6), there holds the a posteriori error estimate

$$|J(\varepsilon)| \leq \eta_\varepsilon(u_h) := \sum_{K \in \mathcal{T}_h} \left\{ \sum_{i=1}^{2} \varrho_K^i \omega_K^i \right\},$$  

(4.2.20)

where the cell-residuals $\varrho_K^i$ and weights $\omega_K^i$ are defined by

$$\varrho_K^1 = \|R(u_h)\|_K, \quad \varrho_K^2 = h^{-1/2}_K \|r(u_h)\|_{\partial K},$$

$$\omega_K^1 = \|z - \varphi_h\|_K, \quad \omega_K^2 = h^{1/2}_K \|z - \varphi_h\|_{\partial K}.$$
The interpretation of the relation (4.2.20) is that the weights \( \omega_K \) describe the dependence of \( J(e) \) on variations of the cell residuals \( \varphi_K \),
\[
\frac{\partial J(e)}{\partial \varphi_K} \approx \omega_K \approx \| z - i_h z \|_K \approx h_K^2 \| \nabla^2 z \|_K,
\]
where we have chosen \( \varphi_h = i_h z \). We remark that in a finite difference discretization of (4.2.6) the corresponding influence factors would behave like \( \omega_K \approx \| z \|_K \). In practice the weights \( \omega_K \) have to be determined computationally. This will be discussed in the next section.

A posteriori error bounds in global norms

By the same type of argument as used above, we can also derive the traditional global error estimates in the energy and the \( L^2 \) norm.

1) Energy-error bound: First, we use the functional
\[
J(\varphi) = \| \nabla e \|^{-1}(\nabla e, \nabla \varphi)
\]
in the dual problem. For its solution \( z \in V \), there holds the a priori bound \( \| \nabla z \| \leq 1 \).

We obtain the estimate
\[
\| \nabla e \| \leq \sum_{K \in T_h} \{ \varphi_K \omega_K + \varphi_{\partial K} \omega_{\partial K} \}
\]
\[
\leq \left( \sum_{K \in T_h} \{ h_K^2 \omega_K^2 + h_K \varphi_{\partial K}^2 \} \right)^{1/2} \left( \sum_{K \in T_h} \{ h_K^{-2} \omega_K^{-2} + h_K^{-1} \varphi_{\partial K}^2 \} \right)^{1/2},
\]
with residual terms and weights as defined above. Now, we use an extension of the interpolation estimate (4.2.12) (see, e.g., Ciarlet [Cia84] and Brenner & Scott [BS94]),
\[
\left( \sum_{K \in T_h} \left\{ h_K^{-2} \| z - \tilde{i}_h z \|_K^2 + h_K^{-1} \| z - \tilde{i}_h z \|_{\partial K}^2 \right\} \right)^{1/2} \leq \tilde{c}_i \| \nabla z \|,
\]
(4.2.21)

where \( \tilde{i}_h z \in V_h \) is a modified nodal interpolation which is defined and stable on \( H^1(\Omega) \). This gives us
\[
\| \nabla e \| \leq \tilde{c}_i \left( \sum_{K \in T_h} \{ h_K^2 \omega_K^2 + h_K \varphi_{\partial K}^2 \} \right)^{1/2} \| \nabla z \|.
\]

Finally, observing the a priori bound for \( \| \nabla z \| \), we conclude the energy-error a posteriori error estimate
\[
\| \nabla e \| \leq \eta_E(u_h) = \tilde{c}_i \left( \sum_{K \in T_h} \{ h_K^2 \omega_K^2 + h_K \varphi_{\partial K}^2 \} \right)^{1/2}.
\]
(4.2.22)

2) \( L^2 \)-norm error bound: Next, we use the functional
\[
J(\varphi) = \| e \|^{-1}(e, \varphi)
\]
in the dual problem. If the (polygonal) domain $\Omega$ is convex, the dual solution $z \in V$ is in $H^2(\Omega)$ and admits the a priori bound $\|\nabla^2 z\| \leq c_a = 1$. This yields analogously as in (i) the error estimate

$$
\|e\| \leq \left( \sum_{K \in T_h} \{h_K^2 \bar{\varepsilon}^2_K + h_K^3 \bar{\varepsilon}^2_\partial K \} \right)^{1/2} \left( \sum_{K \in T_h} \{h_K^2 \bar{\omega}_K^2 + h_K^3 \bar{\omega}_\partial K^2 \} \right)^{1/2}.
$$

Now, we use the stronger version of the interpolation estimate (4.2.21),

$$
\left( \sum_{K \in T_h} \left\{ h_K^{-4} \|z - i_{hK} z\|_K^2 + h_K^{-3} \|z - i_{hK} z\|_\partial K^2 \right\} \right)^{1/2} \leq c_i \|\nabla^2 z\|,
$$

(4.2.23)

to obtain

$$
\|e\| \leq c_i \left( \sum_{K \in T_h} \{h_K^2 \bar{\varepsilon}^2_K + h_K^3 \bar{\varepsilon}^2_\partial K \} \right)^{1/2} \|\nabla^2 z\|.
$$

Finally, observing the a priori bound for $\|\nabla^2 z\|$, we conclude the $L^2$-norm a posteriori error estimate

$$
\|e\| \leq \eta_{L^2}(u_h) = c_i c_s \left( \sum_{K \in T_h} \{h_K^2 \bar{\varepsilon}^2_K + h_K^3 \bar{\varepsilon}^2_\partial K \} \right)^{1/2}.
$$

(4.2.24)

LOCAL A POSTERIORI ERROR BOUNDS BASED ON ENERGY-ERROR ESTIMATES

Finally, we briefly discuss an approach proposed by Prudhomme & Oden [PO99] for estimating the error in local functional output which is based on sharp energy-error estimates. Suppose that we have general upper and lower error bounds available of the form

$$
\eta_{\text{low}}(u_h) \leq \|\nabla e\| \leq \eta_{\text{up}}(u_h).
$$

(4.2.25)

Such estimates may be obtained by using local defect correction techniques (see, e.g., Ainsworth & Oden [AO97]). Now, let again a (linear) output functional $J(\cdot)$ be given and let $z \in V$ be the corresponding dual solution. Then, there holds the error representation

$$
J(e) = (\nabla e, \nabla \varepsilon),
$$

where $\varepsilon = z - z_h$, and $z_h \in V_h$ is the Ritz approximation of $z$ defined by

$$
(\nabla z_h, \nabla \varphi_h) = J(\varphi_h) \quad \forall \varphi_h \in V_h.
$$

By the parallelogram identity, we obtain

$$
J(e) = (\theta \nabla e, \theta^{-1} \nabla \varepsilon) = \frac{1}{2} \|\nabla(\theta e + \theta^{-1} \varepsilon)\|^2 - \frac{1}{2} \|\nabla(\theta e - \theta^{-1} \varepsilon)\|^2,
$$

with a suitable balancing parameter $\theta > 0$. In view of (4.2.25), this implies the a posteriori error estimate

$$
\hat{\eta}_{\text{low}} \leq J(e) \leq \hat{\eta}_{\text{up}},
$$

(4.2.26)
with the lower bound given by

$$\tilde{\eta}_{\text{low}} = \frac{1}{2} \eta_{\text{low}}^2 (\theta e + \theta^{-1} \varepsilon) - \frac{1}{2} \eta_{\text{up}}^2 (\theta e - \theta^{-1} \varepsilon),$$

and the upper bound given by

$$\tilde{\eta}_{\text{up}} = \frac{1}{2} \eta_{\text{up}}^2 (\theta e + \theta^{-1} \varepsilon) - \frac{1}{2} \eta_{\text{low}}^2 (\theta e - \theta^{-1} \varepsilon).$$

This estimate provides sharp upper as well as lower bounds for the functional error $J(e)$ as long as reliable bounds for the energy errors $\|\nabla (\theta e + \theta^{-1} \varepsilon)\|$ and $\|\nabla (\theta e - \theta^{-1} \varepsilon)\|$ are available. This can be achieved as long the functional $J(\cdot)$ is regular, i.e. defined on $V$, which excludes, for example, point-error evaluation. In the latter case, sharp upper and lower energy-error bounds are hard to get. Further it is not clear how to organize effective local mesh adaptation on the basis of the estimate (4.2.26). Finally, it should be emphasized that the derivation of this error bound requires the energy form to be symmetric and positive definite (i.e. to be a scalar product).

### 4.2.2 Evaluation of the a posteriori error bounds

Now, we will discuss the evaluation of the error representation (4.2.18) or the resulting error bound (4.2.20) (see Becker & Rannacher [BR96]). All strategies are based on solving the dual problem numerically. By $z_h^{(1)} \in V_h$, we denote the approximation to $z$ obtained on the current mesh by the same bilinear finite element scheme as used for $u_h$. Of course, we cannot simply replace $z$ by $z_h^{(1)}$, since then the error bound turns to zero. The remedy is to approximate $z$ with higher accuracy without investing too much work. In the following, we discuss several approaches to this problem.

1. **Approximation by higher order methods:** The dual problem is solved by using *biquadratic* finite elements on the current mesh yielding an approximation $z_h^{(2)}$ to $z$. The error estimator obtained by using this approximation in (4.2.19) is denoted by $\eta^{(1)}(u_h)$. The corresponding local cell-error indicators are

$$\eta_K^{(1)} = \left| \langle R(u_h), z_h^{(2)} - \varphi_h \rangle_K - \langle r(u_h), z_h^{(2)} - \varphi_h \rangle_{\partial K} \right|,$$

where $\varphi_h = i_h z_h^{(2)} \in V_h$ is the bilinear nodal interpolation of $z_h^{(2)}$. It is seen by theoretical analysis as well as by numerical experiments that $\eta^{(1)}(u_h)$ behaves like $\lim_{TOL \to 0} I_{\text{eff}} = 1$. However, approximating $z$ by a higher-order finite element scheme does not appear very economical in estimating the error in the low-order scheme.

2. **Approximation by higher order interpolation:** A simplification is achieved by simple patchwise *biquadratic* interpolation of the bilinear approximation $z_h^{(1)}$ on the current mesh yielding an approximation $i_h z_h^{(1)}$ to $z$. This construction requires some special care on elements with hanging nodes, in order to preserve the higher order accuracy of the interpolation process. The resulting global error estimator is denoted by $\eta^{(2)}(u_h)$. The corresponding local cell-error indicators are

$$\eta_K^{(2)} = \left| \langle R(u_h), i_h z_h^{(1)} - z_h^{(1)} \rangle_K - \langle r(u_h), i_h z_h^{(1)} - z_h^{(1)} \rangle_{\partial K} \right|.$$

In this case, we observe that $\lim \inf_{TOL \to 0} I_{\text{eff}} > 1 \sim 1 - 2$. 

3. **Approximation by difference quotients:** We use the cell-wise interpolation estimate (4.2.12) to estimate the weights \( \omega_K \) and \( \omega_0K \) by

\[
\omega_K + \omega_0K = \|z - i_h z\| K + h_K^{1/2} \|z - i_h z\|_{\partial K} \leq c h_K^2 \|\nabla^2 z\| K.
\]

The second derivative \( \nabla^2 z \) is then replaced by a suitable second-order difference quotient \( \nabla^2 z_h^{(1)} \) of the approximate dual solution \( z_h^{(1)} \in V_h \). The computation of such second-order difference quotients may be costly particularly on cell patches with hanging nodes. A useful substitute is the approximation

\[
\|\nabla^2 z\| K \approx h_K^{-1/2} \|n \cdot (\nabla z_h^{(1)})\|_{\partial K}, \tag{4.2.27}
\]

which corresponds to the jump-residual \( \partial_{\partial K} \) of \( u_h \) the evaluation of which has to be implemented anyway. The resulting global error estimator is denoted by \( \eta^{(3)}(u_h) \). It consists of local cell-error indicators of the form

\[
\eta^{(3)}_K = c h_K^{-1/2} \{ \partial K + \partial_{\partial K} \} \|n \cdot (\nabla z_h^{(1)})\|_{\partial K},
\]

where the cell-residuals are defined as above. In this case, we observe that \( \lim \inf_{\tau \to 0, L \to \infty} L > 1 \) (\( \sim 1 - 8 \)).

4. **Approximation by local residual problems:** Following the idea already used in the energy-error estimator of Bank\&Weiser [BW85] (see also Ainsworth\&Oden [AO97]), on each element \( K \) the local Neumann problems

\[
(\nabla v_K, \nabla \varphi_h)_K = (R(u_h), \varphi_h)_K - (r(u_h), \varphi_h)_{\partial K} \quad \forall \varphi_h \in V_K, \tag{4.2.28}
\]

are solved, where \( V_K = \{ q \in \hat{Q}_2(K), q \perp \hat{Q}_1(K) \} \). The related error indicator based on solving patch-wise Dirichlet problems is not considered here, since it turns out to be too expensive in practice. Using this construction, the local cell-error indicators in (4.2.20) are replaced by

\[
\eta^{(4)}_K = |(\nabla v_K, \nabla (z - \varphi_h))_K| \leq \|\nabla v_K\| K \|\nabla (z - \varphi_h)\| K.
\]

Employing again the approximation (4.2.27), we obtain localize cell-error indicators of the form \( c h_K^{-2} \|\nabla v_K\| K \|n \cdot (\nabla z_h^{(1)})\|_{\partial K} \).

**Table 4.1** Efficiency of weighted error indicators for the point-error \( J(e) = |e(0)| \).

<table>
<thead>
<tr>
<th>L</th>
<th>N</th>
<th>( \eta^{(1)}/J(e) )</th>
<th>( \eta^{(2)}/J(e) )</th>
<th>( \eta^{(3)}/J(e) )</th>
<th>( \eta^{(4)}/J(e) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4^p</td>
<td>6.667</td>
<td>12.82</td>
<td>2.066</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4^p</td>
<td>1.253</td>
<td>13.51</td>
<td>45.45</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4^p</td>
<td>1.052</td>
<td>3.105</td>
<td>9.345</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4^p</td>
<td>1.067</td>
<td>2.053</td>
<td>7.042</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4^p</td>
<td>1.003</td>
<td>1.886</td>
<td>6.667</td>
<td></td>
</tr>
</tbody>
</table>

The effectivity indices observed for these strategies in computing the point-value \( u(0) \) are listed in Table 4.1 (from Becker\&Rannacher [BR96]). The performance of
the estimator $\eta^4(u_h)$ is similar to that of $\eta^3(u_h)$ and is therefore not listed in the table. These results indicate that even for the simplest model situations, an asymptotic efficiency index $I_{\text{eff}} = 1$ is achievable only at the expense of unacceptably high cost, e.g., by approximating the dual solution using higher-order elements. We finally remark that the cell-indicators in the error bound (4.2.20) are used within the mesh adaptation process, while, for a stopping criterion, one better directly evaluates the error representation (4.2.19). In general, the transition form the error identity (4.2.19) to the error estimate (4.2.20) causes significant over-estimation of the true error.

4.2.3 Strategies for mesh adaptation

Now, we want to discuss some popular strategies for mesh adaptation based on an a posteriori error estimate of the form

$$|J(e)| \leq \eta := \sum_{K \in \mathcal{T}_h} \eta_K,$$  \hspace{1cm} (4.2.29)

with certain cell-error indicators $\eta_K = \eta_K(u_h)$ obtained on the current mesh $\mathcal{T}_h$. Suppose that a tolerance $\text{TOL}$ for the error $J(e)$ or a maximum number $N_{\text{max}}$ of mesh cells have been prescribed. Starting from an approximate solution $u_h \in V_h$ obtained on the current mesh $\mathcal{T}_h$ the mesh adaptation may be organized by one of the following strategies.

- **Error balancing strategy:** Cycle through the mesh and equilibrate the local error indicators according to $\eta_K \approx \text{TOL}/N$ with $N := \# \{ K \in \mathcal{T} \}$. This leads eventually to $\eta \approx \text{TOL}$. Since $N$ changes when a cell is refined, this strategy requires iteration with respect to $N$ on each mesh level.

- **Fixed mesh or error fraction strategy:** Order cells according to the size of $\eta_K$ and refine a certain percentage (say 20%) of cells with largest $\eta_K$ (or those which make up 20% of the estimator value $\eta$) and coarsen those cells with smallest $\eta_K$. By this strategy, we may achieve a prescribed rate of increase of $N$ or keep it constant within a time-dependent loading process.

- **Mesh optimization strategy:** Use a (proposed) representation

$$\eta := \sum_{K \in \mathcal{T}_h} \eta_K \approx \int_{\Omega} h(x)^2 \Phi(x) \, dx$$ \hspace{1cm} (4.2.30)

for directly generating a formula for an optimal mesh-size distribution:

$$h_{\text{opt}}(x) = \left( \frac{W}{N_{\text{max}}} \right)^{1/2} \Phi(x)^{-1/4}, \hspace{0.5cm} W := \int_{\Omega} \Phi(x)^{1/2} \, dx.$$ \hspace{1cm} (4.2.31)

Here, we think of a smoothly distributed mesh-size function $h(x)$ such that $h_K \approx h_{\text{opt}}$. The existence of such a representation with an $h$-independent error density function $\Phi(x) = \Phi(u(x), z(x))$ can be rigorously justified only under very restrictive conditions but is generally supported by computational experience (for details see Becker & Rannacher [BR96] and Rannacher [Ran99]).
4.2.4 Examples

Example 1: Computation of point stresses
First, we consider a highly localized error functional. As concrete example, we choose again the square domain $\Omega = (-1,1)^2$ and the output functional $J(u) = \partial_1 u(0)$. In this case, the dual solution does not exist in the sense of $H^1_0(\Omega)$, such that for practical use, we have to regularize the functional like

$$J_\varepsilon(u) = |B_\varepsilon|^{-1} \int_{B_\varepsilon} \partial_1 u \, dx = \partial_1 u(0) + O(\varepsilon),$$

where $B_\varepsilon = \{ x \in \Omega | |x| < \varepsilon \}$, and $\varepsilon = \text{TOL}$ is a suitable error tolerance. The corresponding dual solution $z$ behaves like $|\nabla^2 z(x)| \approx d(x)-3$, where $d(x) = |x| + \varepsilon$. From the general a posteriori error estimate (4.2.20), we obtain for the present case

$$|J_\varepsilon| \approx \eta_\omega(u_h) = c_i \sum_{K \in \mathcal{h}_n} h_K^3 \left\{ \varrho_T + \varrho_{\partial \Omega} \right\}. \quad (4.2.32)$$

The refinement strategy based on this a posteriori error estimate yields meshes with complexity $N_{\text{opt}} \approx \text{TOL}^{-1}$ which is better than what could be achieved on uniformly refined meshes. This predicted asymptotic behavior is well confirmed by the results shown in Table 4.2. Figure 4.2 shows the balanced mesh for $\text{TOL} = 4^{-1}$ and the approximation to the dual solution $z_\varepsilon$, with $\varepsilon = \text{TOL}$, computed on this mesh (see Becker & Rannacher [BR96]).

<table>
<thead>
<tr>
<th>TOL $^{-m}$</th>
<th>N</th>
<th>L</th>
<th>$J_\varepsilon(u_h)$</th>
<th>$\eta_\omega(u_h)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$4^{-5}$</td>
<td>940</td>
<td>9</td>
<td>4.10e-1</td>
<td>1.42e-2</td>
</tr>
<tr>
<td>$4^{-4}$</td>
<td>4912</td>
<td>12</td>
<td>4.14e-3</td>
<td>3.50e-3</td>
</tr>
<tr>
<td>$4^{-3}$</td>
<td>20980</td>
<td>15</td>
<td>2.77e-4</td>
<td>9.25e-4</td>
</tr>
<tr>
<td>$4^{-2}$</td>
<td>86740</td>
<td>17</td>
<td>5.82e-5</td>
<td>2.36e-4</td>
</tr>
</tbody>
</table>

Example 2: Computation of mean boundary stress
The second example is meant as an illustrative exercise. For problem (4.2.8) on a smoothly bounded domain $\Omega \subset \mathbb{R}^2$, we consider the functional

$$J(u) = \int_{\partial \Omega} \partial_n u \, ds \quad \left( = \int_{\Omega} f \, dx \right),$$

and ask the question: What is an optimal mesh-size distribution for computing $J(u)$?

The corresponding dual problem

$$(\nabla \varphi, \nabla z) = (1, \partial_n \varphi)_{\partial \Omega} \quad \forall \varphi \in V \cap C^1(\bar{\Omega}),$$

has a measure solution with density of the form $z \equiv 1$ in $\Omega$, $z = 0$ on $\partial \Omega$. In order to avoid the use of measures, we consider the regularized functional

$$J_\varepsilon(\varphi) = |S_{\varepsilon}|^{-1} \int_{S_{\varepsilon}} \partial_n \varphi \, ds = \int_{\partial \Omega} \partial_n \varphi \, ds + O(\varepsilon), \quad \varepsilon = \text{TOL},$$
where \( S_\varepsilon = \{ x \in \Omega, \text{dist}(x, \partial \Omega) < \varepsilon \} \). Then, the dual solution is \( z_\varepsilon = 1 \) in \( \Omega \setminus S_\varepsilon \), \( z_\varepsilon(x) = \varepsilon^{-1} \text{dist}(x, \partial \Omega) \) on \( S_\varepsilon \). This implies that

\[
J_\varepsilon(\varepsilon) \leq c_i \sum_{K \in T_\varepsilon, K \cap S_\varepsilon \neq \emptyset} h_K^2 \{| \theta K + \psi K \| \nabla^2 z_\varepsilon \| K \},
\]

i.e., there is no contribution to the error from cells in the interior of \( \Omega \). Hence, independent of the form of the forcing \( f \), the optimal strategy is to refine the elements adjacent to the boundary and to leave the others unchanged.

Figure 4.3  (Example 2) Refined mesh and approximate dual solution obtained by the weighted error estimator \( \eta_\omega(u_h) \) after 7 refinement steps.
4.3 THE VARIATIONAL FRAMEWORK

4.3.1 The Hencky and Prandtl-Reuss models

The fundamental problem in the flow theory of linear–elastic perfectly plastic material in classical notation reads (see, e.g., Duvaut & Lions [DL76])

\[
\begin{align*}
\ddot{u} - \text{div} \sigma &= f, \quad \varepsilon(\dot{u}) = A\dot{\tau} + \lambda \quad \text{in } \Omega, \\
\lambda(\tau - \sigma) &\leq 0 \quad \forall \tau \text{ with } \mathcal{F}(\tau) \leq 0, \quad \lambda\dot{\tau} = 0 \quad \text{in } \Omega, \\
\dot{u} &= 0 \quad \text{on } \Gamma_D, \quad \sigma \cdot n = g \quad \text{on } \Gamma_N,
\end{align*}
\]

where \( \sigma \) and \( u \) are the stress tensor and displacement vector, respectively. We assume a stress-free initial state \( \sigma(0) = 0 \) and \( u(0) = 0 \). This idealized model describes the deformation of an elasto-plastic body occupying a bounded domain \( \Omega \subset \mathbb{R}^d \) (\( d = 2 \) or 3) under the action of a body force \( f \) and a surface traction \( g \) along \( \Gamma_N \). Along the remaining part of the boundary, \( \Gamma_D = \partial\Omega \setminus \Gamma_N \), the body is fixed. The plastic growth is denoted by \( \lambda \), and \( \mathcal{F}(\cdot) \) is the (convex) von Mises yield function. We assume a quasi-static process, i.e., the acceleration term \( \ddot{u} \) is neglected. Accordingly, the loading is prescribed in terms of functions \( f(x, t) = a(t)f_0(x) \) and \( g(x, t) = b(t)g_0(x) \) depending on a time-like parameter \( t \in I := [0, T] \), where the functions \( a(t) \) and \( b(t) \) vary only slowly (relative to the elasto-plastic changes in the model). Further, the deformation is supposed to be small, so that the strain tensor can be written as \( \varepsilon(u) = \frac{1}{2}(\nabla u + \nabla u^T) \). The material tensor \( A \) is assumed to be symmetric and positive definite on symmetric stress tensors, with smallest and largest eigenvalues \( \alpha_- \) and \( \alpha_+ \), respectively.

We consider the case of a linear–elastic isotropic material law expressed in the form

\[
\sigma = C \varepsilon(u) := 2\mu \varepsilon^D(u) + \kappa \text{div } u I,
\]

with material dependent constants \( \mu > 0 \) and \( \kappa > 0 \). Here, \( C = A^{-1} \), and \( \tau^D := \tau - \frac{1}{3} \text{tr}(\tau) I \) is the deviatoric part of a tensor \( \tau \). The von Mises flow function has the form

\[
\mathcal{F}(\sigma) = |\sigma^D|^2 - \sigma_0^2 \leq 0.
\]

Problem (4.3.1) is to be solved by the finite element Galerkin method on adaptively optimized meshes. To this end, we have at first to formulate (4.3.1) in a variational setting. Let again \( L^2(\Omega) \) and \( H^1(\Omega) \) be the usual Lebesgue space and first-order Sobolev space as defined above. With this notation, we introduce the function spaces

\[
\begin{align*}
H := L^2(\Omega)^d, \quad V := \{ v \in H^1(\Omega)^d, \ v = 0 \text{ on } \Gamma_D \}, \\
W := L^2(\Omega)^{d \times d}_{\text{sym}}, \quad W^{\text{div}} := \{ \tau \in W, \ \text{div } \tau \in H \}, \\
W_g^{\text{div}} := \{ \tau \in W^{\text{div}}, \ \tau \cdot n = g \text{ on } \Gamma_N \}, \quad W_{f,g}^{\text{div}} := \{ \tau \in W_g^{\text{div}}, \ -\text{div } \tau = f \text{ in } \Omega \},
\end{align*}
\]

and define for any such space \( \Sigma \) the admissible subset as

\[
\Pi \Sigma := \{ \tau \in \Sigma, \mathcal{F}(\tau) \leq 0 \}.
\]

Accordingly, \((\cdot, \cdot)\) and \(\| \cdot \|\) denote the \(L^2\)-inner product and norm over \( \Omega \), and \((\cdot, \cdot)_\Gamma\) is the \(L^2\)-inner product over a boundary segment \( \Gamma \).
For the following, we assume the boundary part $\Gamma_D$ to be non-trivial such that by Korn’s inequality, there holds
\[
\|v\|_{E}^{2} = (C\varepsilon(u), \varepsilon(u)) \geq \gamma \|
abla v\|^2, \quad v \in V, \tag{4.3.2}
\]
with some constant $\gamma > 0$. The case $\Gamma_D = \emptyset$ requires some modification.

Following Duvaut & Lions [DL76], we introduce the displacement velocity $v := \dot{u}$ and state the dual-mixed formulation of (4.3.1): Find a pair $(v, \sigma) : I \rightarrow H \times \Pi W^{\text{div}}_g$, with $\sigma(0) = 0$ and $v(0) = 0$, satisfying
\[
(A\dot{\sigma}, \tau - \sigma) + (v, \text{div}(\tau - \sigma)) \geq 0, \quad \forall \tau \in \Pi W^{\text{div}}_0, \\
-(\text{div} \sigma, \varphi) = (f, \varphi) \quad \forall \varphi \in H. \tag{4.3.3}
\]

Integrating by parts in (4.3.3) leads to the primal-mixed variational formulation: Find a pair $(v, \sigma) : I \rightarrow V \times \Pi W$, with $\sigma(0) = 0$ and $v(0) = 0$, satisfying
\[
(A\dot{\sigma} - \varepsilon(v), \tau - \sigma) \geq 0, \quad \forall \tau \in \Pi W, \\
(\sigma, \varepsilon(\varphi)) = F(\varphi) \quad \forall \varphi \in V, \tag{4.3.4}
\]
where the righthand side has the form
\[
F(\varphi) := (f, \varphi) + (g, \varphi)_\Gamma.
\]

By regularization the variational inequality (4.3.4) can be transformed into an equation of the form
\[
(A\dot{\sigma}^\mu - \varepsilon(v^\mu), \tau) + \mu^{-1}(\sigma^\mu - \Pi \sigma^\mu, \tau) = 0 \quad \forall \tau \in \Pi W, \\
(\sigma^\mu, \varepsilon(\varphi)) = F(\varphi) \quad \forall \varphi \in V, \tag{4.3.5}
\]
with a small parameter $\mu > 0$. This approach is commonly used for proving existence of solutions of problem (4.3.3). In fact, it can be shown (c.f., Duvaut & Lions [DL76] and Johnson [Joh76]) that the solutions $(v^\mu, \sigma^\mu) : I \rightarrow V \times W$ converge weakly for $\mu \rightarrow 0$ to the (unique) solution $(v, \sigma) \in V \times \Pi W$ of the limiting problem (4.3.4).

In the case of instantaneous loading $f = f(x)$, $g = g(x)$, we may neglect the rate dependence in (4.3.4) obtaining a material behaviour of Hencky-type (c.f. Duvaut & Lions [DL76]). This means that a pair $(u, \sigma)$ of displacement and stress is sought satisfying
\[
(A\sigma - \varepsilon(u), \tau - \sigma) \geq 0 \quad \forall \tau \in \Pi W, \\
(\sigma, \varepsilon(\varphi)) = F(\varphi) \quad \forall \varphi \in V, \tag{4.3.6}
\]
From (4.3.6), we see that $\sigma$ is just the projection of $C\varepsilon(u) - A^{-1}\varepsilon(u)$ onto $\Pi W$. Hence, the primal-mixed formulation (4.3.6) can be written in the compact form
\[
(\sigma - \Pi C\varepsilon(u), \tau) = 0 \quad \forall \tau \in \Pi W, \\
(\sigma, \varepsilon(\varphi)) = F(\varphi) \quad \forall \varphi \in V, \tag{4.3.7}
\]
with the projection operator $\Pi : \mathbb{R}^{d\times d}_{\text{sym}} \rightarrow \{\tau \in \mathbb{R}^{d\times d}_{\text{sym}}, \mathcal{F}(\tau) \leq 0\}$ explicitly given by
\[
\Pi \tau := \begin{cases} 
\tau & \text{if } |\tau|^D \leq \sigma_0, \\
\sigma_0 |\tau|^D \tau^D + \frac{1}{2} \text{tr}(\tau) I & \text{if } |\tau|^D > \sigma_0.
\end{cases}
\]
For material with (linear) "hardening", the projection operator $\Pi$ may be assumed to be of the form

$$
\Pi_\mu \tau := \begin{cases} 
\tau & \text{if } |\tau^D| \leq \sigma_0, \\
(1 - \gamma)\sigma_0|\tau^D|^{-1}\tau^D + \gamma\tau^D + \frac{1}{2}tr(\tau)I & \text{if } |\tau^D| > \sigma_0.
\end{cases}
$$

with some "hardening" parameter $\gamma \in (0, 1)$.

Finally, eliminating the stress $\sigma$ in (4.3.7), we obtain the primal form

$$
(\Pi \varepsilon(u), \varepsilon(\varphi)) = F(\varphi) \quad \forall \varphi \in V. \quad (4.3.8)
$$

This reformulation of the static problem (4.3.6) as a nonlinear variational equation in terms of the deformation $u$ alone will also be used for the static subproblems within a load-stepping process for the quasi-static Prandtl–Reuss problem. Below, we will use linearization for nonlinear problems involving the Lipschitz-continuous function $C(\tau) := \Pi \tau$. To this end, we introduce the notation of a "derivative" of $C^\prime(\cdot)$ by

$$
C^\prime(\tau) := \begin{cases} 
I & \text{if } |\tau^D| \leq \sigma_0, \\
\sigma_0|\tau^D|^{-1}(I - |\tau^D|^2\tau^D) + \frac{1}{2}tr(\tau)I & \text{if } |\tau^D| > \sigma_0.
\end{cases} \quad (4.3.9)
$$

This notation will be used within Newton-type iterations and in the context of error estimation for expressing differences in the form

$$
C(\sigma) - C(\tau) = \int_0^1 C^\prime(\tau + s(\sigma - \tau))(\sigma - \tau)ds =: B(\sigma - \tau). \quad (4.3.10)
$$

This definition is justified since the admissible set $\{\tau \in W, \mathcal{F}(\tau) = |\tau^D|^2 - \tau_0^2 \leq 0\}$ is convex. The tensor $B \in \mathbb{R}^{d \times d \times (d \times d)}$ defined by (4.3.10) is symmetric on $\mathbb{R}_{sym}$ and positive semi-definite:

$$
E^TBF = F^TBE, \quad E^TBE \geq 0, \quad E, F \in \mathbb{R}_{sym}^{d \times d}.
$$

### 4.3.2 Time-discretization in the Prandtl-Reuss model

For discretization, the parameter interval $[0, T]$ is decomposed like

$$
0 = t_0 < t_1 < \ldots < t_M = T
$$

into subintervals $I_m := (t_{m-1}, t_m]$ of length $k_m := t_m - t_{m-1}$. Integrating in (4.3.5) over $I_m$ yields

$$
(\dot{A}\sigma^m - \dot{A}\sigma^{m-1}, \tau) + \int_{I_m} \left\{ \mu^{-1}(\sigma - \Pi\sigma, \tau) - (\varepsilon(\nu), \tau) \right\} dt = 0 \quad \forall \tau \in W, \quad (4.3.11)
$$

$$
\int_{I_m} (\sigma, \varepsilon(\varphi)) dt = \int_{I_m} F(\varphi) dt \quad \forall \varphi \in V.
$$

The integrals are approximated by quadrature formulas of the type

$$
\int_{I_m} w(t) dt = k_m \{ \alpha w^m + (1 - \alpha)w^{m-1} \}, \quad (4.3.12)
$$
with some $\alpha \in (0,1]$. The choice of $\alpha = 1$ corresponds to the backward Euler scheme, while for $\alpha = \frac{1}{2}$, we obtain the Crank-Nicolson scheme. Here, we only consider the simple Euler scheme which reads
\[
(A\sigma^m - A\sigma^{m-1} - k_m \varepsilon(v^m), \tau) + k_m \mu^{-1}(\sigma^m - \Pi \sigma^m, \tau) = 0 \quad \forall \tau \in W, \\
(\sigma^m, \varepsilon(\varphi)) = F^m(\varphi) \quad \forall \varphi \in V,
\]
(4.3.13)
with the initial values $\sigma^0 = 0$, $u^0 = 0$, and the averaged forcing terms
\[
\bar{F}^m(\varphi) := k_m^{-1} \int_{t_m} F(\varphi) \, dt.
\]
Letting now $\mu \to 0$ in the regularization results in
\[
(A\sigma^m - A\sigma^{m-1} - k_m \varepsilon(v^m), \tau) \geq 0 \quad \forall \tau \in \Pi W, \\
(\sigma^m, \varepsilon(\varphi)) = F^m(\varphi) \quad \forall \varphi \in V.
\]
(4.3.14)
Hence, in each load step, we have to solve a static problem of Hencky-type. By the same argument as used before, this can equivalently be written in form of a nonlinear variational equation,
\[
(\sigma^m - \Pi(\sigma^{m-1} + k_m C\varepsilon(v^m)), \tau) = 0 \quad \forall \tau \in W, \\
(\sigma^m, \varepsilon(\varphi)) = F^m(\varphi) \quad \forall \varphi \in V,
\]
(4.3.15)
with the projection $\Pi$ defined above. This problem can be written in terms of the displacements $u^m := u^{m-1} + k_m v^m \in V$ in the form
\[
(\sigma^m - \Pi(\sigma^{m-1} + C\varepsilon(u^m - u^{m-1})), \tau) = 0 \quad \forall \tau \in W, \\
(\sigma^m, \varepsilon(\varphi)) = F^m(\varphi) \quad \forall \varphi \in V,
\]
(4.3.16)
starting from the initial value $u^0 = 0$. We note that the stationary Hencky model (4.3.6) may be viewed as the approximation of the quasi-static Prandtl-Reuss model by one time step of length $k = 1$ starting form the initial state $\sigma^0 = 0$, $u^0 = 0$. Again, eliminating the stress $\sigma$ in (4.3.16) results in the primal formulation
\[
(C^m(\varepsilon(u^m)), \varepsilon(\varphi)) = F^m(\varphi) \quad \forall \varphi \in V,
\]
(4.3.17)
where
\[
C^m(\varepsilon(u^m)) := \Pi(\sigma^{m-1} + C\varepsilon(u^m - u^{m-1})).
\]

A PRIORI ERROR ANALYSIS FOR TIME DISCRETIZATION

The considered Prandtl-Reuss model is quasi-stationary as it does not contain any acceleration term. This should imply that the errors introduced by time discretization as well as varying spatial discretization at each time level will accumulate at most linearly. Therefore, in first-order approximation the incremental loading steps $t_{m-1} \to t_m$ may be treated separately as static problems, assuming the "starting" values $\sigma^{m-1}$ to be exact. To justify this approach, we will now derive an a priori error estimate for the time discretization based on the regularized formulation (4.3.13). To this end, we introduce the material-dependent $L^2$-norm $\| \sigma \|_A := (A\sigma, \sigma)^{1/2}$. 


Theorem 4.3.1. For the error $e^n_m := \sigma(t_m) - \sigma^m$ of the backward Euler scheme (4.3.13), there holds the a priori error estimate
\[
\max_{m=1, \ldots, M} \|e^n_m\|_A \leq \|e^n_0\|_A + T \max_{m=1, \ldots, M} \left\{ k_m \max_{I_m} \|\bar{\sigma}\|_A \right\}. \tag{4.3.18}
\]

Proof. The time-continuous solution $\bar{\sigma}^m := \sigma(t_m)$ satisfies
\[
A\bar{\sigma}^m - A\bar{\sigma}^{m-1} + k_m \mu^{-1}(\bar{\sigma}^m - \Pi \bar{\sigma}^m) - k_m \varepsilon(\dot{v}^m) = - \int_{I_m} (t - t_{m-1}) A\bar{\sigma} \, dt.
\]
Comparing this with the equation satisfied by $\sigma^m$, we find for the errors $e^n_m := \bar{\sigma}^m - \sigma^m$ and $e^n_0 := \hat{v}^m - v^m$ that
\[
Ae^n_m + k_m \mu^{-1}(e^n_m - \Pi \dot{e}^m + \Pi \sigma^m) - k_m \varepsilon(e^n_0) = Ae^{m-1}_\sigma - \int_{I_m} (t - t_{m-1}) A\bar{\sigma} \, dt.
\]
Further, there holds $\text{div } e^n_m = 0$. We introduce tensors $B^m = B^m(\bar{\sigma}^m, \sigma^m)$ by
\[
\Pi \dot{e}^m = \Pi \sigma^m = B^m e^n_\sigma
\]
(see the related definition (4.3.10)). Using this notation, the above equation can be written in the form
\[
Ae^n_m + \mu^{-1} k_m (I - B^m) e^n_m - k_m \varepsilon(e^n_0) = Ae^{m-1}_\sigma - \int_{I_m} (t - t_{m-1}) A\bar{\sigma} \, dt.
\]
We note that $\|B^m\| \leq 1$ and, consequently,
\[
((I - B^m)\tau, \tau) \geq 0, \quad \tau \in W.
\]
Multiplying this identity by $e^n_\sigma$ yields
\[
(Ae^n_m, e^n_\sigma) + \mu^{-1} k_m ((I - B^m) e^n_m, e^n_\sigma) - k_m (\varepsilon(e^n_0), e^n_\sigma)
\]
\[
= (Ae^{m-1}_\sigma, e^n_\sigma) - \int_{I_m} (t - t_{m-1}) (A\bar{\sigma}, e^n_\sigma) \, dt,
\]
and consequently, using $\text{div } e^n_m = 0$,
\[
(Ae^n_m, e^n_\sigma) \leq (Ae^{m-1}_\sigma, e^n_\sigma) - \int_{I_m} (t - t_{m-1}) (A\bar{\sigma}, e^n_m) \, dt.
\]
From this, we deduce that
\[
\|e^n_m\|_A \leq \|e^{m-1}_\sigma\|_A + k_m \int_{I_m} \|\bar{\sigma}\|_A \, dt.
\]
Finally, summing over $m = 1, \ldots, n$, we obtain
\[
\|e^n_\sigma\|_A \leq \|e^n_0\|_A + \sum_{m=1}^n k_m \int_{I_m} \|\bar{\sigma}\|_A \, dt,
\]
which implies the assertion. \qed
ADAPTIVE FEM IN ELASTO-PLASTICITY

We emphasize that the a priori estimate (4.3.18) does not explicitly depend on the regularization parameter $\mu$. Therefore, by continuity, the result carries over to the limit case $\mu \to 0$. This result shows that the error due to the incremental loading process grows at most linearly with time provided that the exact solution stays bounded. Furthermore, this time-stepping error can be controlled by balancing the local time step $k_m$ against the local bound for $\tilde{\sigma}$ on $I_m$.

The result of Lemma 4.3.1 can be extended to the dynamic version of the Prandtl-Reuss model

$$
(A \tilde{\sigma} - \varepsilon(v), \tau - \sigma) \geq 0, \quad \forall \tau \in \Pi W,
$$

$$
(\dot{v}, \varphi) + (\sigma, \varepsilon(\varphi)) = F(\varphi) \quad \forall \varphi \in V,
$$

and its backward Euler time-discretization

$$
\begin{align*}
(\sigma^m - \Pi(\sigma^{m-1} + k_m C \varepsilon(v^m)), \tau) &= 0 \quad \forall \tau \in W, \\
(v^m - v^{m-1}, \varphi) + k_m (\sigma^m, \varepsilon(\varphi)) &= k_m \tilde{F}^m(\varphi) \quad \forall \varphi \in V,
\end{align*}
$$

with the projection $\Pi$ defined above. Below, we will investigate the effect of including the acceleration term on the dynamic behaviour of the solution for a benchmark problem.

4.3.3 Treatment of almost incompressible behaviour

Usually, in the plastic zone there holds $|\Pi C \varepsilon(u)| \ll \kappa$, i.e., nearly incompressible material behaviour occurs. In this case, the relation between $\varepsilon(u)$ and $\sigma$ becomes stiff causing a poor approximation behaviour for discretizations based on (4.3.7) or (4.3.8).

One may account for this difficulty by introducing an auxiliary variable $p := -\kappa \text{div } u$ which plays the role of a pressure. Then, introducing the pressure space $Q := L^2(\Omega)$, a triple $\{u, \sigma, p\} \in V \times W \times Q$ is determined by the saddle-point problem

$$
\begin{align*}
(\sigma - \Pi C \varepsilon(u), \tau) &= 0 \quad \forall \tau \in W, \\
(\sigma, \varepsilon(\varphi)) - (p, \text{div } \varphi) &= F(\varphi) \quad \forall \varphi \in V, \\
(\text{div } u, \chi) + \kappa^{-1} (p, \chi) &= 0 \quad \forall \chi \in Q,
\end{align*}
$$

where, in the case $\Gamma_D = \partial \Omega$, the pressure space $Q$ is supplemented by the condition $(p, 1) = 0$. The well-posedness of this formulation follows from the following “inf-sup”-stability relation

$$
\sup_{u \in V} \frac{(p, \text{div } u)}{\|\varepsilon(u)\|} \geq \gamma \|p\|, \quad p \in Q,
$$

with some $\kappa$-independent constant $\gamma > 0$. The positive effect of this modification will be demonstrated by a numerical test below.

4.3.4 An elasto-plastic benchmark

Below, we will test the DWR method for a standard benchmark problem. A geometrically two-dimensional square disc with a hole is subjected to a constant boundary traction acting upon two opposite sides. No body force is applied, i.e., $f \equiv 0$. 
The plane-strain approximation is used, i.e., $\varepsilon_{ij} = 0$ $(i = 1, 2, 3)$, and perfectly plastic material behavior is assumed. In virtue of symmetry the consideration can be restricted to a quarter of the domain as shown in Figure 4.4. The height and width of the quarter corresponding to lines $15$ and $15$ are 100, and the radius of the hole is 10. The ratio of the lines 57 and 67 is 2 : 1. The material parameters are chosen as those of aluminium, $\kappa = 164.206 \text{ N/mm}^2$, $\mu = 80.193.80 \text{ N/mm}^2$, $\sigma_0 = \sqrt{2/3} 3450$ (see Lehmann & Blix [LB85]). The boundary traction is given in the form $g(t) = t g_0$, $g_0 = 100$, $t \in [0, 6]$.

![Diagram of the benchmark problem and plot of $|\sigma^0|$](image)

**Figure 4.4** Geometry of the benchmark problem and plot of $|\sigma^0|$ (plastic region black, transition zone white) computed on a mesh with $N \approx 10,000$ cells.

The two-dimensional version of this benchmark consists of two parts.

a) *Hencky case*: For the stationary Hencky model, the calculations are performed with one load step from $t = 0$ to $t = 4.5$. The quantities to be computed are:

- Displacement $u_1(P_3)$ and $u_1(P_1)$.
- Displacement $u_2(P_3)$ and $u_2(P_2)$.
- Stress $\sigma_{22}(P_2)$ and displacement $u_1(P_2)$.
- Line integral $L_1(u_2)$ of displacement $u_2$ along $\Gamma = P_4 P_5$.
- $L^2$-norm $||\sigma||_{\Omega}$ over $\Omega$.

b) *Prandtl-Reuss case*: For the quasi-stationary Prandtl-Reuss model, the calculations are performed with a sequence of load steps starting from $t = 0$ for two loading paths, the first one with linearly increasing loading and the second one with a “saw-tooth”-like loading and unloading. The quantities to be computed are:

- Value $t_f$ for which plastification sets in.
- Value $t_b$ for which static collapse occurs.
- Diagram of $t$ over $L_1(u_2)$.
- Diagram of $t$ over $\sigma_{11}(P_f)$.
- Diagram of $\sigma_{11}(P_f)$ versus $t$ over the loading and unloading path $0 \leq t \leq 18$ ("saw-tooth" curve).
4.4 THE LINEAR-ELASTIC CASE

At first, for apply the DWR method to the linear-elastic case. Setting $\Pi = \text{id}$, the \textit{primal-mixed} formulations obtained from (4.3.7) take the form

$$\sigma - C\epsilon(u), \tau) + (\sigma, \epsilon(\varphi)) = F(\varphi) \quad \forall (\varphi, \tau) \in V \times W, \quad (4.4.1)$$

where again $C := A^{-1}$.

For discretizing (4.4.1), we consider standard finite element subspaces $V_h \subset V$ and $W_h \subset W$ as described above. Here, we confine ourselves to the lowest-order approximation by continuous $P_1$- or $Q_1$-elements, i.e. linear or (iso-parametric) d-linear shape functions, for the deformations in $V_h$, while the corresponding discrete stresses in $W_h$ are chosen as discontinuous $P_0$- or likewise $Q_1$-elements, respectively.

The underlying meshes $\mathcal{T}_h$ are assumed to satisfy the usual regularity conditions (shape regularity, cf. Brenner&Scott [BS94]) and to properly match the decomposition $\partial \Omega = \Gamma_D \cup \Gamma_N$. For simplicity, we assume the boundary $\partial \Omega$ to be polygonal. Again, hanging nodes are allowed (but only one per element edge), in order to ease the refinement and coarsening process.

The discrete approximations $\{u_h, \sigma_h\} \in V_h \times W_h$ are determined by the primal-mixed system

$$(\sigma_h - C\epsilon(u_h), \eta_h) + (\sigma_h, \epsilon(\varphi_h)) = F(\varphi_h) \quad \forall (\varphi_h, \eta_h) \in V_h \times W_h, \quad (4.4.2)$$

This problem is seen to be uniquely solvable by the same arguments as used in establishing the solvability of the continuous problem. Denoting by $P_h$ the orthogonal projection onto $W_h$ ($L^2$-projection), the \textit{primal-mixed scheme} (4.4.2) may be written in the \textit{primal} form

$$(P_h C\epsilon(u_h), \epsilon(\varphi_h)) = F(\varphi_h) \quad \forall \varphi_h \in V_h. \quad (4.4.3)$$

Under the compatibility condition

$$\varphi_h \in V_h \implies C\epsilon(\varphi_h) \in W_h, \quad (4.4.4)$$

the projection $P_h$ in (4.4.3) can be dropped and we obtain the standard \textit{primal} finite element scheme. Condition (4.4.4) is satisfied, for instance, in the lowest-order case of continuous $P_1$-elements for $V_h$ and discontinuous $P_0$-elements for $W_h$, if the material tensor $C$ is piecewise constant.

Subtracting (4.4.2) from (4.4.1) yields the Galerkin orthogonality relation

$$(\epsilon_o - C\epsilon(e_o), \tau_h) + (\epsilon_o, \epsilon(\varphi_h)) = 0 \quad \forall (\varphi_h, \eta_h) \in V_h \times W_h. \quad (4.4.5)$$

for the errors $e_o := u - u_h$ and $e_o := \sigma - \sigma_h$. From this, one infers in the usual way the convergence of the approximation (see, e.g., Brenner&Scott [BS94]).

4.4.1 A \textit{a posteriori} error analysis

The \textit{a posteriori} error analysis starts with selecting a quantity $J(\{u, \sigma\})$ of physical interest which is to be controlled. This may be a point value (local average), a line integral or another locally or globally defined functional of the solution defined on
the solution space $V \times W$. Since the material tensor $C$ is positive definite, the corresponding dual problem

$$(\tau - C\varepsilon(\varphi), \pi) + (\tau, \varepsilon(z)) = J(\{(\varphi, \tau)\}) \quad \forall \{(\varphi, \tau)\} \in V \times W,$$  \hfill (4.4.6)

possesses a unique solution $\{z, \pi\} \in V \times W$. We note that (4.4.6) may also be written in its primal form

$$(C\varepsilon(\varphi), \varepsilon(z)) = J(\{(\varphi, C\varepsilon(\varphi))\}) \quad \forall \varphi \in V.$$  \hfill (4.4.7)

Taking $\{\varphi, \tau\} = \{e_u, e_\sigma\}$ in (4.4.6) and using the Galerkin orthogonality (4.4.5), we obtain the error identity

$$J(\{e_u, e_\sigma\}) = (e_u - C\varepsilon(e_u), \pi - \pi_h) + (e_\sigma, \varepsilon(z - z_h))$$  \hfill (4.4.8)

with an arbitrary pair $\{z_h, \pi_h\} \in V_h \times W_h$. Then, by element–wise integration by parts in the last term on the right-hand side, it follows that

$$J(\{e_u, e_\sigma\}) = \sum_{K \in \mathcal{T}} \left\{ (f + \text{div} \, \sigma_h, z - z_h)_{\mathcal{K}} - (r_h, z - z_h)_{\partial \mathcal{K}} \\
- (\sigma_h - C\varepsilon(u_h), \pi - \pi_h)_{\mathcal{K}} \right\},$$  \hfill (4.4.9)

with the edge residuals $r_h$ defined by

$$r_{h\Gamma} := \begin{cases} \frac{1}{2} \gamma_n [\sigma_h], & \text{if } \Gamma \subset \partial \mathcal{K} \setminus \partial \Omega, \\
\gamma_n \sigma_h, & \text{if } \Gamma \subset \Gamma_D, \\
\gamma_n \sigma_h - g, & \text{if } \Gamma \subset \Gamma_N, \end{cases}$$

where again $[\sigma_h]$ denotes the jump of $\sigma_h$ across interior cell edges $\Gamma \subset \partial \Omega$. From (4.4.9) we conclude the following theorem (Rannacher&Suttmeier [RS96]).

**Theorem 4.4.1.** Under the foregoing assumptions, the finite element scheme (4.4.2) admits the a posteriori error estimate

$$|J(\{e_u, e_\sigma\})| \leq \eta_\omega(u_h, \sigma_h) := \sum_{K \in \mathcal{T}} \left\{ \sum_{i=1}^{3} \varrho_i^K \omega_i^K \right\},$$  \hfill (4.4.10)

with the local residuals and weight factors defined by

$$\varrho_1^K = \| f + \text{div} \, \sigma_h \|_{\mathcal{K}}, \quad \omega_1^K = \| z - z_h \|_{\mathcal{K}},$$
$$\varrho_2^K = h_{\mathcal{K}}^{-1/2} r_h \|_{\partial \mathcal{K}}, \quad \omega_2^K = h_{\mathcal{K}}^{1/2} \| z - z_h \|_{\partial \mathcal{K}},$$
$$\varrho_3^K = \| \sigma_h - C\varepsilon(u_h) \|_{\mathcal{K}}, \quad \omega_3^K = \| \pi - \pi_h \|_{\mathcal{K}}.$$  \hfill (4.4.11)

The terms $\omega_3^K \varrho_3^K$ in the a posteriori error estimate (4.4.10) can be dropped if the compatibility condition (4.4.4) is satisfied. The weighting terms $z - z_h$ and $\pi - \pi_h$ can be evaluated as described above for the Poisson problem using an approximate solution of the dual problem (4.4.7).
ADAPTIVE FEM IN ELASTO-PLASTICITY

The weighted error estimate (4.4.10) will be compared against the traditional energy error estimate (see, e.g., Ainsworth & Oden [AO97] or Verfürth [Ver96])

\[ \| u - u_h \|_E \leq \eta_E := c_\varepsilon c_i \left( \sum_{K \in T} h_K^2 \varepsilon_K^2 \right)^{1/2}, \]  

(4.4.11)

with the "energy norm" \( \| \cdot \|_E := (C \varepsilon(\cdot), \varepsilon(\cdot))^{1/2} \) and the local residuals \( \varepsilon_K := \varepsilon_K^e + \varepsilon_K^p \) as defined above. The stability constant \( c_\varepsilon \) depends on the material matrix \( A \). The interpolation constant \( c_i \) is again set to \( c_i = 1 \).

4.4.2 Numerical test for the linear-elastic case

The approach described above is applied for a model problem in linear elasticity employing the two dimensional plane strain model (for more details and further examples see Rannacher & Suttmeier [RS96]). A square elastic disc with a crack is subjected to a constant boundary traction acting on half of the upper boundary (see Figure 4.5). Along the right-hand side and the lower boundary the disc is fixed and the remaining part of the boundary (including the crack) is left free. This problem is interesting as its solution develops a singularity at the tip of the crack where a strong stress concentration occurs. The material parameters are chosen as above.

![Figure 4.5 Geometry sketch of the test problem "square disc with a slit"]

We want to compute the mean normal stress over the clamped part of the boundary,

\[ J(\sigma) = \int_{\Gamma_c} n \sigma \cdot n \, ds. \]  

(4.4.12)

Since this functional is irregular, it is regularised according to

\[ J_\varepsilon(\tau) := \frac{1}{|\Gamma_c|} \int_{\Gamma_c} n \cdot \tau \cdot n \, dx, \quad \Gamma_\varepsilon = \{ x \in \mathbb{R}^2, \text{dist}(x, \Gamma_D) < \frac{1}{2}\varepsilon \}, \]  

(4.4.13)

with \( \varepsilon = TOL_L \). A reference solution \( \sigma_{\text{ref}} \) is computed on a fine mesh with about 200,000 cells for determining the relative error and the corresponding "effectivity index" (used here for expressing over-estimation) defined by

\[ E^\varepsilon := \frac{|J(\sigma_h) - J(\sigma_{\text{ref}})|}{|J(\sigma_{\text{ref}})|}, \quad I_{\text{eff}} := \frac{\eta_c(u_h, \sigma_h)}{|J(\sigma_h) - J(\sigma_{\text{ref}})|}. \]

Further, \( L \) indicates the refinement level and \( N \) the corresponding number of cells.
At first, we consider the \textit{weighted} error estimator (4.4.10) on uniformly refined meshes. The results are shown in Table 4.3.

\textbf{Table 4.3} \ The \textit{weighted} error estimator on uniformly refined meshes

<table>
<thead>
<tr>
<th>$L$</th>
<th>$N$</th>
<th>$J(\sigma_h)$</th>
<th>$E^{\text{wz}}$</th>
<th>$I_{\text{eff}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>256</td>
<td>0.017080</td>
<td>0.0283</td>
<td>1.80</td>
</tr>
<tr>
<td>2</td>
<td>1024</td>
<td>0.019528</td>
<td>0.0181</td>
<td>1.76</td>
</tr>
<tr>
<td>3</td>
<td>4096</td>
<td>0.021137</td>
<td>0.0113</td>
<td>1.70</td>
</tr>
<tr>
<td>4</td>
<td>16384</td>
<td>0.022161</td>
<td>0.0070</td>
<td>1.66</td>
</tr>
<tr>
<td>5</td>
<td>65536</td>
<td>0.022802</td>
<td>0.0043</td>
<td>1.62</td>
</tr>
<tr>
<td>$\infty$</td>
<td></td>
<td>0.0237...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Then, we compare the \textit{weighted} error estimator $\eta_w(u_h)$ and the energy-error estimator $\eta_E(u_h)$ on adaptively refined meshes. For both the \textit{fixed mesh fraction} strategy is used for mesh refinement. The results are shown in Table 4.4 and Figure 4.6. Obviously, the \textit{weighted} error estimator represents the true error quite accurately and leads to more economical meshes than the energy-error estimator.

\textbf{Table 4.4} \ Results obtained by the \textit{weighted} error estimator (left) and the \textit{energy-error} estimator (right).

<table>
<thead>
<tr>
<th>$L$</th>
<th>$N$</th>
<th>$J(\sigma_h)$</th>
<th>$E^{\text{wz}}$</th>
<th>$I_{\text{eff}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>484</td>
<td>0.019642</td>
<td>0.0180</td>
<td>1.96</td>
</tr>
<tr>
<td>3</td>
<td>1060</td>
<td>0.021138</td>
<td>0.0113</td>
<td>1.95</td>
</tr>
<tr>
<td>4</td>
<td>2113</td>
<td>0.022157</td>
<td>0.0070</td>
<td>1.96</td>
</tr>
<tr>
<td>5</td>
<td>4455</td>
<td>0.022795</td>
<td>0.0044</td>
<td>1.92</td>
</tr>
<tr>
<td>6</td>
<td>8530</td>
<td>0.023198</td>
<td>0.0027</td>
<td>1.86</td>
</tr>
<tr>
<td>7</td>
<td>15886</td>
<td>0.023428</td>
<td>0.0017</td>
<td>1.79</td>
</tr>
<tr>
<td>8</td>
<td>29947</td>
<td>0.023593</td>
<td>0.0010</td>
<td>1.79</td>
</tr>
<tr>
<td>9</td>
<td>52288</td>
<td>0.023697</td>
<td>0.0006</td>
<td>1.86</td>
</tr>
</tbody>
</table>

\textbf{Figure 4.6} \ Finest meshes obtained by the \textit{weighted} estimator $\eta_w$ (left) and the \textit{energy-error} estimator $\eta_E$ (right).
4.5 THE STATIC HENCKY-TYPE MODEL

We consider static problems of the primal-mixed form

\[
(\sigma - \Pi(\dot{\sigma} + C\varepsilon(u - \dot{u})), \tau) + k(\sigma, \varepsilon(\varphi)) = k F(\varphi),
\]

(4.5.1)

for all \( \{\varphi, \tau\} \in V \times W \), with given initial stress \( \dot{\sigma} \) and displacement \( \dot{u} \), while the parameter \( k \) is fixed and may be set to one. For later purposes, we carry \( k \) explicitly through the following analysis. Such problems occur as substeps within a quasistatic loading process for approximating the Prandtl-Reuss model. The corresponding discretized problems are

\[
(\sigma_h - \Pi(\dot{\sigma} + C(\varepsilon(u_h - \dot{u}))), \tau_h) + k(\sigma_h, \varepsilon(\varphi_h)) = k F(\varphi_h),
\]

(4.5.2)

for all \( \{\varphi_h, \tau_h\} \in V_h \times W_h \), with finite element subspaces \( V_h \subset V \) and \( W_h \subset W \) as described in the preceding section. To simplify notation, we define the nonlinear material operator

\[
C(\varepsilon(u)) := \Pi(\dot{\sigma} + C(\varepsilon(u) - C(\varepsilon(\dot{u}))),
\]

the explicit form of which has been specified above.

4.5.1 A posteriori error estimates

Following the linear case, our error analysis for the discretization (4.5.2) is based on a duality argument. Suppose that the quantity \( J(\{u, \sigma\}) \) is to be computed. We use again the notation \( e_u := u - u_h, e_\sigma := \sigma - \sigma_h \) for the discretization errors. Subtracting (4.5.2) from (4.5.1), we obtain the nonlinear Galerkin orthogonality relation

\[
(e_\sigma - C(\varepsilon(u)) + C(\varepsilon(u_h)), \tau) + k(e_\sigma, \varepsilon(\varphi)) = 0,
\]

(4.5.3)

for \( \{\varphi, \tau\} \in V_h \times W_h \). Next, we define the tensor-function \( B = B(u, u_h) \) by

\[
C(\varepsilon(u)) - C(\varepsilon(u_h)) = B(\varepsilon(e_u)).
\]

In virtue of the special form of the material operator \( C(\cdot) \), the tensor \( B \) is positive semi-definite. For representing the error \( J(\{e_u, e_\sigma\}) \), we employ the solution \( \{z, \pi\} \in V \times W \) of the linear dual problem in primal-mixed form

\[
(\pi, \pi) - (\varepsilon(\varphi), B^*\pi) + k(\tau, \varepsilon(\varphi)) = J(\{\varphi, \tau\}), \quad \forall \{\varphi, \tau\} \in V \times W.
\]

(4.5.4)

It can be shown that the dual solution is well defined by (4.5.4) as a limit of solutions to corresponding regularized problems. We note that \( B = C \) in the purely elastic regime, i.e., the dual problem (4.5.4) reduces to that of the linear-elastic case (4.4.6). Taking \( \varphi = e_u, \tau = e_\sigma \) in (4.5.4) and using the orthogonality relation (4.5.3), there holds

\[
J(\{e_u, e_\sigma\}) = (e_\sigma - B(\varepsilon(e_u)), \pi) + k(e_\sigma, \varepsilon(z))
\]

\[
= (e_\sigma - C(\varepsilon(u)) + C(\varepsilon(u_h)), \pi) + k(e_\sigma, \varepsilon(z))
\]

\[
= (e_\sigma - C(\varepsilon(u)) + C(\varepsilon(u_h)), \pi - \pi_h) + k(e_\sigma, \varepsilon(z - z_h)),
\]
with suitable approximations \( \{ z_h, \pi_h \} \in V_h \times W_h \) to \( \{ z, \pi \} \). Then, analogously as in the linear case, we obtain the error representation
\[
J(\{ e_u, e_v \}) = \sum_{K \in \mathcal{T}} \left\{ k(f + \text{div} \sigma_h, z - z_h)_K - k(r_h, z - z_h)_{\partial K} - (\sigma_h - C(\varepsilon(u_h))), \pi - \pi_h)_K \right\},
\]
with the edge residuals \( r_h \) again defined by
\[
r_h := \begin{cases} 
\frac{1}{2} n \cdot [\sigma_h], & \text{if } \Gamma \subset \partial K \setminus \partial \Omega, \\
-n \cdot \sigma_h, & \text{if } \Gamma \subset \Gamma_D, \\
n - \sigma_h - g, & \text{if } \Gamma \subset \Gamma_N.
\end{cases}
\]

From this, we conclude the following result for the static elasto-plastic case (see Rannacher&Suttmeyer [RS97]). In the purely static case, we may set \( k = 1 \).

**Theorem 4.5.1.** Under the foregoing assumptions, the primal finite element scheme (4.5.2) for the generalized Hencky model (4.5.1) admits the a posteriori error estimate
\[
|J(\{ e_u, e_v \})| \leq \eta_u(u_h, \sigma_h) := \sum_{K \in \mathcal{T}} \left\{ k \sum_{i=1}^2 \omega_K^i \varrho_K^i + \omega_K^2 \varrho_K^2 \right\},
\]
with the local residuals and weight factors defined by
\[
\begin{align*}
\varrho_K^1 &= \| f + \text{div} \sigma_h \|_K, & \omega_K^1 &= \| z - z_h \|_K, \\
\varrho_K^2 &= h_K^{-1/2} | r_h \|_{\partial K}, & \omega_K^2 &= h_K^{-1/2} | z - z_h \|_{\partial K}, \\
\varrho_K^3 &= \| \sigma_h - C(\varepsilon(u_h)) \|_K, & \omega_K^3 &= \| \pi - \pi_h \|_K.
\end{align*}
\]

**Linearization of dual problem**

In order to evaluate the a posteriori bound (4.5.5), we have to compute approximations to the dual solution \( \{ z, \pi \} \). This requires at first linearization in the tensor-coefficient \( B(u, u_h) \). The simplest possibility (used in all our computations presented below) is to replace the unknown solution \( u \) by the computed approximation \( u_h \). However, this requires some care since the function \( C(\cdot) \) is not differentiable. As guideline, we may use the representation
\[
B(\varepsilon(e_u)) := C(\varepsilon(u)) - C(\varepsilon(u_h)) = \int_0^1 C'(\varepsilon(u_h + s e_u))\varepsilon(e_u) \, ds,
\]
introduced in (4.3.10) with the derivative \( C'(\cdot) \) defined in (4.3.9). Replacing here formally \( u \) by \( u_h \) leads to \( B \approx C'(\varepsilon(u_h)) \). The resulting perturbed dual problem
\[
(\tau, \pi) - (\varepsilon(\varphi), C'(\varepsilon(u_h))^* \pi) + k(\tau, \varepsilon(z)) = J(\{ \varphi, \tau \}),
\]
for all \( \{ \varphi, \tau \} \in V \times W \), is solved by the same method as used in computing \( u_h \), yielding approximations \( \{ \tilde{z}_h, \tilde{\pi}_h \} \in V_h \times W_h \):
\[
(\tau_h, \tilde{\pi}_h) - (\varepsilon(\varphi_h), C'(\varepsilon(u_h))^* \tilde{\pi}_h) + k(\tau_h, \varepsilon(\tilde{z}_h)) = J(\{ \varphi_h, \tau_h \}),
\]
for all \( \{ \varphi_h, \tau_h \} \in V_h \times W_h \).
for all \( \{ \varphi_h, \tau_h \} \in V_h \times W_h \). The evaluation of the coefficient \( C'(\varepsilon(u_h))^{*} \) on cells in the elastic-plastic transition zone is usually done by simple numerical integration. This may appear as a rather crude approximation, but it works in practice. The reason may be that the critical situation only occurs in cells intersecting the elastic-plastic transition zone which is a lower-dimensional surface.

The weights \( \omega_K \) may then again be approximated as in the linear case described above. We emphasize that the computation of the dual solution requires to solve only linear problems and normally only amounts to a small fraction of the total cost within a Newton iteration for the nonlinear problem.

**OTHER A POSTERIORI ERROR INDICATORS**

Below, we will compare our weighted error estimator (4.5.5) against two more traditional approaches.

1) **The ZZ-approach:** The error indicator proposed by Zienkiewicz & Zhu [ZZ87] for finite element models in structural mechanics is based on the idea of higher-order stress recovery by local averaging. The element-wise error \( \| \sigma - \sigma_h \|_K \) is thought to be well represented by the auxiliary quantity \( \eta_K := \| M_h \sigma_h - \sigma_h \|_K \), where \( M_h \sigma_h \) is a local (super-convergent) approximation of \( \sigma \). The corresponding (heuristic) global error estimator reads

\[
\| \sigma - \sigma_h \| \approx \eta_{ZZ} := \left( \sum_{K \in T} \| M_h \sigma_h - \sigma_h \|_K^2 \right)^{1/2}.
\]  

(4.5.8)

For our purpose we assume the discrete stresses to be constant over each cell. One possible construction of \( M_h \sigma_h \) is the patch-wise \( L^2 \)-projection \( P_K \sigma_h \) onto the space of (bi-)linear shape functions. Here the nodal value at a point of the triangulation determining \( M_h \sigma_h \) is obtained by averaging the cell-wise constant values of \( \sigma_h \) of those cells having this point in common. For cells containing hanging nodes this process is appropriately modified.

![Figure 4.7 Stress averaging process in the ZZ-error indicator.](image)

2) **An energy error estimator:** Johnson & Hansbo [JH92] proposed an error estimator for the primal-mixed formulation of the Hencky model which is based on monotonicity properties of the energy form and, under some additional heuristic assumptions, bounds the error in the global energy norm. Let \( \Omega_h^e \) and \( \Omega_h^e \) denote the union of
elements where the discrete solution behaves elastic and plastic, respectively. Then, the estimator reads, for the case \( \bar{u} = 0, \bar{\sigma} = 0, \)
\[
\|\sigma - \sigma_h\| \approx \eta_E := c_t \left( \sum_{K \in \mathcal{T}} h_K^2 \eta_K^2 \right)^{1/2},
\]
with the local error indicators
\[
\eta_K^2 := \begin{cases} h_K^2 \max_K |R(u_h)|^2, & \text{if } K \in \Omega_h^e, \\ h_K^{-1} \max_K \left| C\varepsilon(u_h) - \mathcal{M}_h C\varepsilon(u_h) \right| \int_K |R(u_h)| \, dx, & \text{if } K \in \Omega_h^o, \end{cases}
\]
where on each element \( K \in \mathcal{T} \) the local residual is defined by \( R(u_h) := |\text{div} C\varepsilon(u_h)| + \frac{1}{2} h_K^{-1} |n - C\varepsilon(u_h)|. \)

Here, \( c_t \) is some interpolation constant usually set to one. This estimator is rather heuristic, as it relies on the assumption that the plasticification zone is already correctly captured on the current mesh. Furthermore, it is of only sub-optimal order in the plastic zone which results in mesh over-refinement in \( \Omega_h^e \) though the stresses are suspected to be rather smooth there. The ZZ-estimator (4.5.8) does not suffer from this deficiency as it essentially relies on the smoothness of \( \sigma \). Hence, we are led to modify the estimator (4.5.9) by replacing the obviously too crude a bound \( \max_K |C\varepsilon(u_h)| \) in the plastic zone by \( \max_K |C\varepsilon(u_h) - \mathcal{M}_h C\varepsilon(u_h)| \). This gives us the better (still heuristic) error indicators
\[
\eta_K^2 := \begin{cases} h_K^2 \max_K |R(u_h)|^2, & \text{if } K \in \Omega_h^e, \\ h_K^{-1} \max_K \left| C\varepsilon(u_h) - \mathcal{M}_h C\varepsilon(u_h) \right| \int_K |R(u_h)| \, dx, & \text{if } K \in \Omega_h^o. \end{cases}
\]

REALIZATION OF PROJECTION ONTO THE YIELD SURFACE

In practice, the exact evaluation of the pointwise projection \( \Pi \) becomes difficult so that the formal scheme (4.5.12) requires some modification. In our implementation, this is accomplished by carrying only information at Gaussian points (centerpoint for \( P_1 \)-elements, and \( 2 \times 2 \)-Gauss points for \( Q_1 \)-elements) at which \( \Pi \) is evaluated exactly. That is, in the purely elastic case with piecewise constant tensor \( C \), within the local mesh refinement and coarsening the projection \( P_h \) is computed exactly, i.e.,
\[
(\sigma_h, \tau) = (C(\varepsilon(u)), \tau) \quad \forall \tau \in W_h.
\]

This grid transfer process is realized in the natural way for "regular" (bisected) elements as indicated in Figure 4.8. This approach also applies when hanging nodes occur. The same process is also used in the elasto-plastic case causing an additional error in the plastic zone. This error will not be considered further in our analysis. Below, it is always assumed that the discretization is realized in its ideal form (4.5.12).
SOLUTION OF THE NONLINEAR PROBLEMS

The nonlinear algebraic problems (4.5.2) or (4.5.13) are approximated by a damped Newton iteration. Beginning with a suitable starting value $u_h^0$ (e.g., the result of a preceding load level $u_h^{n-1}$) the step $u_h^i \rightarrow u_h^{i+1}$ reads

$$(P_h C'(\varepsilon(u_h^{i-1})) \varepsilon(\delta u^i_h), \varepsilon(\varphi)) = R^{i-1}(\varphi) \quad \forall \varphi \in V_h,$$

with the derivative $C'(\cdot)$ of the function $C(\cdot)$ (see the definition (4.3.9)) and the Newton residual

$$R^{i-1}(\varphi) := \tilde{F}(\varphi) - (C'(\varepsilon(u_h^{i-1}))), \varepsilon(\varphi)).$$

This is followed by the update

$$u_h^{i+1} = u_h^i + \lambda_i \delta u^i_h.$$  

The damping parameter $0 < \lambda_i \leq 1$ is determined in the form $\lambda_i = 2^{-r}$, such that the residual is decreased. The resulting linear subproblems are solved by the CR-method with multigrid acceleration. The use of multigrid is rather natural as in the course of the mesh refinement process a sequence of nested meshes is automatically generated. Since the theme of this note is the aspect of a posteriori error estimation and mesh design, we do not go into the details of the algebraic solution techniques and refer instead to Suttmeier [Sut96].

The Newton iteration is carried on the current finite element mesh until the solution $u_h \in V_h$ and $\sigma_h := P_h C(\varepsilon(u_h)) \in W_h$ of the discrete nonlinear problem is reached with discretization accuracy. This normally requires within a hierarchical multi-level iteration only about $10 - 15$ steps even on fine meshes. Then, the mesh is adjusted according to the a posteriori error estimate (4.5.5) and the solution process is repeated until the prescribed solution accuracy TOL is achieved.
4.5.2 Decoupled displacement-pressure discretization

As mentioned above, in the plastic region the material behavior will be almost incompressible which causes stability problems of the discretization based on the formulation (4.3.15). In order to cope with this problem, we propose a stabilised finite element discretization using an auxiliary “pressure” variable (see Suttmeier [Sut00]).

We briefly describe the modification by the displacement-pressure formulation (4.3.21) described above for the pure Hencky problem (4.5.1) with $\hat{\sigma} = 0$, $\hat{u} = 0$. The finite element subspaces $V_h \subset V$ and $W_h \subset W$ are supplemented by a subspace $Q_h \subset Q$ for the discrete “pressure”. Then, the corresponding discrete problems seek $\{u_h, \sigma_h, p_h\} \in V_h \times W_h \times Q_h$ such that

$$
\begin{align*}
(\sigma_h - \Pi \varepsilon(u_h), \tau_h) + (\sigma_h, \varepsilon(\varphi_h)) - (p_h, \text{div} \varphi_h) &= F(\varphi_h), \\
(\text{div} u_h, \chi_h) + (\kappa^{-1} p_h, \chi_h) &= 0,
\end{align*}
$$

(4.5.12)

for all $\{\varphi, \tau, \chi\} \in V_h \times W_h \times Q_h$. Here, we choose $Q_h$ of “equal-order” as the “velocity space” $V_h$, i.e., it consists also of continuous, piecewise (isoparametric) bilinear functions. This discretization would be unstable due to a stiff velocity-pressure coupling, i.e., it does not satisfy the discrete analogue of the inf-sup stability condition (4.3.22). In order to avoid this difficulty, we introduce “pressure stabilization” leading to the scheme

$$
\begin{align*}
(\sigma_h - C(\varepsilon(u_h)), \tau_h) + (\sigma_h, \varepsilon(\varphi_h)) - (p_h, \text{div} \varphi_h) &= F(\varphi_h), \\
(\text{div} u_h, \varphi_h) + \kappa^{-1}(p_h, \varphi_h) + \sum_{K \in T_h} \delta_K (\nabla p_h, \nabla q_h)_K &= 0,
\end{align*}
$$

(4.5.13)

for all $\{\varphi_h, \tau_h, \varphi_h\} \in V_h \times W_h \times Q_h$. The stabilization parameter is chosen as $\delta_K = \alpha \delta K^2$, with a fixed constant $\alpha = 1$. For the so modified scheme, there holds

$$
\sup_{u_h \in \hat{V}_h} \frac{\|p_h, \text{div} \varphi_h\|}{\|\varepsilon(u_h)\|} \geq (\alpha \sum_{K \in T_h} \delta_K \|\nabla p_h\|_K^2)^{1/2} \geq \gamma \|p_h\|, \quad p_h \in Q_h,
$$

(4.5.14)

which restores the missing discrete stability. For an analysis of this approach see Franca:Stenberg [FS91]. The increase in accuracy achieved through the introduction of the extra pressure variable will be demonstrated below for the Hencky model.

**NUMERICAL TEST**

We want to demonstrate the performance of the velocity-pressure decoupling by a numerical test. To this end, we choose again a variant of the model problem “square disc with crack” (see Figure 4.9) with material values $\kappa = 2\mu = 160000$. The boundary traction is of the form $g = g_0$, with $g_0 = 100$ and $t > 0$. In the general Hencky model (4.5.1), we set $\hat{\sigma} = 0$, $\hat{u} = 0$ and take $t_{t_{\text{im}}} = 2.234$, since already for $t = 2.3$ static collapse occurs.

We consider the computation of the contour integral

$$
J_S(u) = \int_S u \cdot n ds = \int_{\Omega_S} \text{div} u \, dx,
$$

(4.5.15)
traction \( g \)

\[ S \]

**Figure 4.9** Geometry of the square disc test problem and plot of \(|\sigma|^2\) (plastic regions black) computed on a mesh with \( N \approx 64,000 \) cells.

where \( S \) is a suitable circular path around the tip of the crack. \( \Omega_S \) is the domain with boundary \( \partial \Omega_S = S \), and \( n \) the outer normal unit vector along \( S \). The results are presented in Table 4.5. It turns out that by the displacement/pressure formulation significantly higher accuracy can be achieved.

**Table 4.5** Results for computing \( J^2(u) \) on adaptive grids by the primal and the displacement/pressure discretization (reference value \( J^2(u) \approx 1.7020e - 04 \)).

<table>
<thead>
<tr>
<th>( N )</th>
<th>( u )-form</th>
<th>( u/p )-form</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>1.6760e-04</td>
<td>1.693630e-04</td>
</tr>
<tr>
<td>2000</td>
<td>1.6817e-04</td>
<td>1.695619e-04</td>
</tr>
<tr>
<td>4000</td>
<td>1.6875e-04</td>
<td>1.696680e-04</td>
</tr>
<tr>
<td>8000</td>
<td>1.6926e-04</td>
<td>1.699004e-04</td>
</tr>
<tr>
<td>16000</td>
<td>1.6963e-04</td>
<td>1.699354e-04</td>
</tr>
<tr>
<td>32000</td>
<td>1.6986e-04</td>
<td>1.700872e-04</td>
</tr>
</tbody>
</table>

4.5.3 Numerical tests for the static Hencky model

As test case for the DWR approach, we consider the plain-strain benchmark “disc with a hole” described in Section 4.3. The calculations are performed with one load step from \( t = 0 \) to \( t = 4.5 \). We recall the quantities to be computed:

- Displacement \( u_1(P_5) \) and \( u_1(P_1) \).
- Displacement \( u_2(P_5) \) and \( u_2(P_2) \).
- Stress \( \sigma_{22}(P_2) \) and displacement \( u_1(P_2) \).
- Line integral \( L_T(u_2) \) of displacement \( u_2 \) along \( \Gamma = \overline{P_5P_2} \).
- \( L^2 \)-norm \( ||\sigma||_\Omega \) over \( \Omega \).

The solutions on very fine (adapted) meshes with about 200,000 cells are taken as reference solutions \( u_{ref} \) for determining the errors \( E(u_h) \) and effectivity indices \( I_{eff} \) of the error estimator:

\[
E^{ref} := \frac{|J(u_{ref}) - J(u_h)|}{|J(u_{ref})|}, \quad I_{eff} := \frac{n\circ(u_h)}{|J(u_{ref}) - J(u_h)|}.
\]
For mesh adaptation the *fixed-mesh-fraction* strategy is used. We present the computational results in the following way: For each particular test, the corresponding table demonstrates the sharpness of the *weighted* a posteriori error bound $\eta_w(u_h)$, while the figure shows the mesh-efficiencies obtained by the weighted error estimator $\eta_w$ compared with the ZZ-error indicator $\eta_{ZZ}$ and the "energy-error" indicator $\eta_E$. Finally, the "optimized" adaptive grid generated by the weighted error estimator with about $N \approx 10000$ cells is displayed.

**The Hencky benchmark: Computation of $u_1(P_5)$**

<table>
<thead>
<tr>
<th>$N$</th>
<th>$u_1 (P_5)$</th>
<th>$E^{\text{rel}}$</th>
<th>$I_{\text{rel}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>6.5991e-02</td>
<td>7.7403e-02</td>
<td>6.4735e-01</td>
</tr>
<tr>
<td>2000</td>
<td>6.3462e-02</td>
<td>3.6121e-02</td>
<td>8.3647e-01</td>
</tr>
<tr>
<td>4000</td>
<td>6.2159e-02</td>
<td>1.4846e-02</td>
<td>1.0499e+00</td>
</tr>
<tr>
<td>8000</td>
<td>6.1554e-02</td>
<td>4.9704e-03</td>
<td>1.5502e+00</td>
</tr>
<tr>
<td>16000</td>
<td>6.1389e-02</td>
<td>2.2746e-03</td>
<td>1.7415e+00</td>
</tr>
<tr>
<td>$\infty$</td>
<td>6.1251e-02</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 4.10** Relative error for computation of $u_1(P_5)$ based on the different error indicators and an "optimal" mesh with about 10,000 cells

The following figures show optimized meshes with $N \approx 10000$ cells and the distributions of the corresponding cell-error indicators $\eta_K$ for computing the point-displacements $u_1(P_5)$ and $u_1(P_1)$. We clearly see a strong effect of the structure of the dual solution on the resulting mesh refinement in the plastic zone. For representing the value $u_1(P_5)$ with good accuracy a good resolution of the whole elastic-plastic transition zone is required, while the representation of $u_1(P_5)$ poses less requirements in this respect. The resulting significant differences in the mesh-size distributions could have hardly be predicted a priori.
Figure 4.11  Optimized meshes with about $N \approx 10000$ cells for computing the deflection $u_1(P_5)$ (left) and corresponding distribution of weights $\omega_K$ (right).

Figure 4.12  Optimized meshes with about $N \approx 10000$ cells for computing the deflection $u_1(P_5)$ (left) and corresponding distribution of weights $\omega_K$ (right).

The Hencky benchmark: Computation of $\sigma_{22}(P_2)$

Table 4.7  Results for $\sigma_{22}(P_2)$ based on the error estimator $\eta_\omega$

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\sigma_{22}(P_2)$</th>
<th>$E_{\text{ref}}$</th>
<th>$I_{\text{eff}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>5.9654e+02</td>
<td>1.4804e-01</td>
<td>2.0553e+02</td>
</tr>
<tr>
<td>2000</td>
<td>5.5653e+02</td>
<td>7.1025e-02</td>
<td>1.5028e+02</td>
</tr>
<tr>
<td>4000</td>
<td>3.3617e+02</td>
<td>3.1857e-02</td>
<td>6.6001e+01</td>
</tr>
<tr>
<td>8000</td>
<td>5.2330e+02</td>
<td>7.0786e-03</td>
<td>1.8520e+01</td>
</tr>
<tr>
<td>16000</td>
<td>5.1978e+02</td>
<td>3.0426e-04</td>
<td>2.3218e+00</td>
</tr>
<tr>
<td>$\infty$</td>
<td>5.1962e+02</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.13 Relative error for computation of $\sigma_2(P_1)$ based on the different error indicators and an "optimal" mesh with about 10,000 cells

The Hencky benchmark: Computation of the line integral $L_T(u_2)$

Table 4.8 Results for $L_T(u_2)$ based on the error estimator $\eta_w$

<table>
<thead>
<tr>
<th>N</th>
<th>$L_T(u_2)$</th>
<th>$E^{rel}$</th>
<th>$I_{eff}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>2.1760e+01</td>
<td>3.0944e-02</td>
<td>6.0564e-01</td>
</tr>
<tr>
<td>2000</td>
<td>2.2126e+01</td>
<td>1.4642e-02</td>
<td>7.7506e-01</td>
</tr>
<tr>
<td>4000</td>
<td>2.2318e+01</td>
<td>6.0949e-03</td>
<td>9.5711e-01</td>
</tr>
<tr>
<td>8000</td>
<td>2.2410e+01</td>
<td>2.0214e-03</td>
<td>1.4112e+00</td>
</tr>
<tr>
<td>16000</td>
<td>2.2433e+01</td>
<td>9.9800e-04</td>
<td>1.4671e+00</td>
</tr>
<tr>
<td>$\infty$</td>
<td>2.2454e+01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.14 Relative error for computation of $L_T(u_2)$ based on the different error indicators and an "optimal" mesh with about 10,000 cells
ADAPTIVE FEM IN ELASTO-PLASTICITY

THE HENCKY BENCHMARK: COMPUTATION OF THE NORM $\|\sigma\|_\Omega$ OVER $\Omega$

Table 4.9 Results for $\|\sigma\|_\Omega$ based on the error estimator $\eta_\omega$

<table>
<thead>
<tr>
<th>$N$</th>
<th>$|\sigma_N|_\Omega$</th>
<th>$I_{v\Omega}^{\text{rel}}$</th>
<th>$I_{\varepsilon\Omega}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>1.702970e+10</td>
<td>5.4505e-03</td>
<td>2.9646e+00</td>
</tr>
<tr>
<td>2000</td>
<td>1.70884e+10</td>
<td>2.4653e-03</td>
<td>2.6479e+00</td>
</tr>
<tr>
<td>4000</td>
<td>1.711008e+10</td>
<td>7.5642e-04</td>
<td>2.1612e+00</td>
</tr>
<tr>
<td>8000</td>
<td>1.711481e+10</td>
<td>4.7986e-04</td>
<td>1.5581e+00</td>
</tr>
<tr>
<td>16000</td>
<td>1.711965e+10</td>
<td>1.9868e-04</td>
<td>1.1285e+00</td>
</tr>
</tbody>
</table>

Figure 4.15 Relative error for computation of $\|\sigma\|_\Omega$ based on the different error indicators and an "optimal" mesh with about 10,000 cells

We close this section by some conclusions from the above numerical results for the benchmark problem "disc with a hole".

i) In all test cases the meshes generated by the DWR method are significantly more efficient for computing the different target quantities than those obtained by the competing ZZ or "energy-norm" error indicators. This effect is especially pronounced when high accuracy is required.

ii) Although the nonlinearity $C(\cdot)$ in the Hencky or Prandtl-Reuss models is locally non-differentiable, this does not affect much the linearization in deriving the a posteriori error estimate and in the Newton iteration for computing the discrete solution. This may be due to the fact that this irregularity only occurs along the elastic-plastic transition zone which is observed to be a lower-dimensional manifold.

iii) Finally, a remark on the computational costs: The evaluation of the "weighted" error estimate requires on each mesh level to solve a linearized dual problem. This amounts to the equivalent of one extra step within the Newton iteration (usually about 10 steps) on this mesh level. Therefore, the extra work for mesh adaptation makes up not more than 15% of the total work (on the optimized mesh).
4.6 THE QUASI–STATIC PRANDTL–REUSS MODEL

Now, we turn to the quasi-static Prandtl-Reuss model and consider a sequence of loading steps, $0 = t_0 < \ldots < t_m < \ldots < t_M = T$,

$$
(s^m - \Pi(s^{m-1} + k_m C \varepsilon(v^m)), \tau) + k_m(s^m, \varepsilon(\varphi)) = k_m F^m(\varphi),
$$

(4.6.1)

for all $\{\varphi, \tau\} \in V \times W$, or in regularized form

$$
(A(s^m - s^{m-1}) - k_m C \varepsilon(v^m)) - k_m \mu^{-1}(s^m - \Pi s^m, \tau)
+ k_m(s^m, \varepsilon(\varphi)) = k_m F^m(\varphi),
$$

(4.6.2)

starting from the initial stress $s^0 = 0$. The explicit dependence on the regularization parameter $\mu > 0$ is dropped. For simplicity, we restrict the following discussion to the formulation without velocity-pressure decoupling.

At each time level $t_m$, we use finite element subspaces $V_h^m \times W_h^m \subset V \times W$, satisfying the conditions stated above. We emphasize that these spaces may carry in time in the course of mesh adaptation within each time step. Then, the discrete deformation velocities and stresses $\{v_h^m, \sigma_h^m\} \in V_h^m \times W_h^m$ are defined by

$$
(A(s_h^m - s_h^{m-1}) - k_m C \varepsilon(v_h^m)) - k_m \mu^{-1}(s_h^m - \Pi s_h^m, \tau_h)
+ k_m(s_h^m, \varepsilon(\varphi_h)) = k_m F^m(\varphi_h),
$$

(4.6.3)

for all $\{\varphi_h, \tau_h\} \in V_h \times W_h$. The corresponding discrete deformations are determined recursively by $u_h^0 = u_h^{m-1} + k_m v_h^m$, starting from the initial value $u_h^0 = 0$. For solving these subproblems, we use the Newton-type procedure described above for the generalized Hencky model.

We will analyse the error in this discretization separately, first with respect to the time/load stepping and then for an adaptive discretization in space. For illustration, we consider the special case that the error is to be controlled with respect to the energy norm $\|e_\sigma^m\|_A := (Ae_\sigma^m, e_\sigma^m)^{1/2}$ corresponding to the local functionals

$$
J^m(\{\varphi, \tau\}) := \|e_\sigma^m\|_A^{-1}(A\tau, e_\sigma^m),
$$

for arguments $\{\varphi, \tau\} \in V \times W$. The functionals $J^m(\cdot)$ explicitly involve the unknown error $e_\sigma^m$. Therefore, their practical evaluation requires guesses for $e_\sigma^m$ which are usually obtained with sufficient (coarse) accuracy by extrapolating the approximations obtained on coarser meshes.

4.6.1 A posteriori error analysis for spatial discretization

Our a posteriori mesh adaptation is applied to these stationary Hencky-type problems, i.e., the finite element meshes are optimized separately within each loading step in accordance to the particular target functionals $J^m(\cdot)$. This results in a dynamic development of refinement and coarsening over the whole loading process. The mesh refinement/coarsening is based on the a posteriori error estimate of Theorem 4.5.1 applied to the present situation.

The a posteriori error bound derived in the preceding section for the static Hencky model can be applied to the incremental load step $t_{m-1} \rightarrow t_m$ only under the
assumption that the starting value $\bar{\sigma} = \sigma_{h}^{m-1}$ is exact. However, this is usually not the case in the course of the loading process. For the following analysis, we split the total error in the scheme (4.6.1) into three components,

$$\sigma(t_m) - \sigma_h^m = e_{\sigma}^m + \bar{e}_{\sigma}^m + e_{\sigma,h}^m,$$

(4.6.4)

which are defined as follows:

- $e_{\sigma}^m := \sigma(t_m) - \sigma_h^m$ is the difference at time $t_m$ between the exact solution $\sigma(t_m)$ of the regularized problem (4.3.5) and the time-discrete solution $\sigma_h^m$ determined by the corresponding backward Euler scheme (4.3.13).
- $\bar{e}_{\sigma}^m := \sigma^m - \bar{\sigma}^m$ is the difference between $\sigma^m$ and the spatially continuous backward Euler solution $\bar{\sigma}(t_m)$ corresponding to starting values $\sigma_{h}^{m-1}$.
- $e_{\sigma,h}^m := \bar{\sigma}^m - \sigma_h^m$ is the difference between $\bar{\sigma}^m$ and the fully discrete solution $\sigma_h^m$ determined by (4.6.1).

It is only the third “static” error component $e_{\sigma,h}^m$ which is covered by the a posteriori analysis of the preceding section. For the “dynamic” error $e_{\sigma}^m$ (pure time-discretization error), we recall from Section 4.3 the following a priori bound:

$$\max_{m=1,\ldots,M} \| e_{\sigma}^m \|_A \leq \| e_{\sigma}^0 \|_A + T \max_{m=1,\ldots,M} \{ k_m \max_{l_m} \| \bar{\sigma}_h \|_A \}. \tag{4.6.5}$$

For the “static” error $e_{\sigma,h}^m$, we can apply the result of Section 4.5:

$$\| e_{\sigma,h}^m \|_A \leq k_m \eta_{\sigma}^m := k_m \sum_{K \in \mathcal{T}_h} \left\{ m \sum_{i=1}^3 \theta_{K,i} \omega_{K,i}^m \right\}, \tag{4.6.6}$$

with the cell-residuals $\theta_{K,i}^m$ and weight factors $\omega_{K,i}^m$ defined by

$$\begin{align*}
\theta_{K,1}^m &:= \| f^m + \text{div} \bar{\sigma}_h^m \|_K, \\
\theta_{K,2}^m &:= h_K^{-1/2} \| r_h \|_{\text{sk}}, \\
\theta_{K,3}^m &:= k_m \| \sigma_h^m - \Pi(\sigma_h^{m-1} + k_m C \varepsilon(\tau_h)) \|_K, \\
\omega_{K,1}^m &:= \| z_h^m - z_h^m \|_K, \\
\omega_{K,2}^m &:= h_K^{1/2} \| z_h^m - z_h^m \|_{\text{sk}}, \\
\omega_{K,3}^m &:= \| \pi_h - \pi_h \|_K.
\end{align*}$$

Here, $\{ z_h^m, \pi_h^m \}$ are the solutions of the static dual problems

$$\begin{align*}
(\tau - B^m \varepsilon(\varphi), \pi_h^m) + &\ k(\tau, \varepsilon(z_h^m)) = \| e_{\sigma,h}^m \|_A^2 (A^T e_{\sigma,h}^m),
\end{align*}$$

(4.6.7)

for all $\{ \varphi, \tau \} \in V \times W$, where the tensor $B^m$ is defined by

$$\Pi(\sigma_h^{m-1} + k_m C \varepsilon(\tau_h)) - \Pi(\sigma_h^{m-1} + k_m C \varepsilon(\tau_h^m)) =: B^m \varepsilon(e_{\sigma,h}^m).$$

We collect these facts in the following theorem [see Rannacher&Suttmeyer] [RS99].

**Theorem 4.6.1.** For the scheme (4.6.1) there holds the a posteriori/a priori error estimate

$$\max_{m=1,\ldots,M} \| e_{\sigma,h}^m \|_A \leq \| e_{\sigma}^0 \|_A + (1 + T) \max_{m=1,\ldots,M} \{ \eta_{\sigma}^m + k_m \max_{l_m} \| \bar{\sigma}_h \|_A \}, \tag{4.6.8}$$

with the initial error $e_{\sigma}^0 := \sigma^0 - \sigma_h^0$ and the a posteriori error estimator $\eta_{\sigma}^m$ defined in (4.6.6).
Proof. We note that the estimates (4.6.5) and (4.6.6) remain valid also in the perfectly plastic limit \( \mu \to 0 \). Therefore, according to the error splitting (4.6.4), it remains to estimate the intermediate error component \( \tilde{e}_m = \sigma^m - \tilde{\sigma}^m \). This is done at first for the regularized formulation (4.6.2). Combining the two equations satisfied by \( \sigma^m \) and \( \tilde{\sigma}^m \), we obtain

\[
A \tilde{e}_m + k_m \mu^{-1} (\tilde{e}_m - \Pi \sigma^m + \Pi \tilde{\sigma}^m) - k_m e(\tilde{e}_m) = A (\sigma^{m-1} - \sigma_h^{m-1}).
\]  

(4.6.9)

We introduce the tensor \( \tilde{\mathcal{B}}^m = \tilde{\mathcal{B}}^m(\sigma^m, \tilde{\sigma}^m) \) through the relation

\[
\Pi \sigma^m - \Pi \tilde{\sigma}^m = : \tilde{\mathcal{B}}^m \tilde{e}_m.
\]

Then, by an analogous argument as used before in the proof of Theorem 4.3.1, based on the definiteness of \( I - \tilde{\mathcal{B}} \) and the relation \( \text{div} \tilde{e}_m = 0 \), we conclude that

\[
\| \tilde{e}_m \|_A \leq \| \sigma^{m-1} - \sigma_h^{m-1} \|_A \leq \| e^{m-1} \|_A + \| e_h^{m-1} \|_A.
\]

(4.6.10)

Summing this for \( m = 1, \ldots, M \) implies that

\[
\max_{m=1, \ldots, M} \| \tilde{e}_m \|_A \leq \| e^{0} \|_A + \sum_{m=0}^{M-1} \| e_h^{m} \|_A.
\]

(4.6.11)

Clearly, this estimate is independent of the parameter \( \mu \). Hence, the proof of the asserted estimate (4.6.8) is complete.

The mixed a posteriori/a priori error estimate (4.6.8) provides the basis for our quasi-static solution approach described above for the Prandtl-Reuss model. The spatial error is controlled adaptively in each time step while the time-stepping error is presumed to be negligible by taking \( k_m \) sufficiently small. All our test computations for the Prandtl-Reuss model presented below are based on this assumption. We note that, in the purely elastic case with a linear loading of the form \( f(t) = f_0, g(t) = g_0 \), the corresponding stress is likewise linear in \( t \), \( \sigma(t) = t \sigma_0 \). Consequently, \( \tilde{\sigma} \equiv 0 \), and our error estimate (4.6.8) indicates that there will be no time-stepping error in the incremental schemes (4.6.1) or (4.6.2). The concept of a systematic a posteriori error analysis for the general case including also the time discretization error will be briefly described in the following subsection.

4.6.2 A posteriori error analysis for time discretization

The foregoing error analysis indicates that it may be appropriate to adapt the local time/load step size \( k_m \) in accordance to the local time-regularity of the computed stress, as the accumulation of these errors seems at most linear in time. This observation is supported by a more refined a posteriori analysis of the time-stepping error following the approach described in Eriksson, Estep, Hansbo & Johnson [EEHJ95]. To this end, we embed our solution scheme into a space-time Galerkin framework and use an associated space-time duality argument. Corresponding to the sequences of values \( v_h^m, \sigma_h^m \) at times \( \{ t_m, m=1, M \} \), we introduce the piecewise constant functions

\[
\{ v_h, \sigma_h \} : [0, T] \to V \times W, \quad v_h |_{t_m} \equiv v_h^m, \quad \sigma_h |_{t_m} \equiv \sigma_h^m, \quad v_h(0) := 0, \quad \sigma_h(0) := 0.
\]
Further, for functions $w$ defined on $I$, we will use the notation
$$
 w^m := w(t_m), \quad w^m_\pm := \lim_{t \to t_m} w(t_m \pm \epsilon), \quad [w]^m := w^m_+ - w^m_-.
$$
Then, the incremental discrete problems (4.3.13) can be written in variational form as

$$
 (A[\sigma^m], \tau) + \int_{I_m} \left\{ \left( A\dot{\sigma} - \varepsilon(v_h) \right), \tau \right\} + \mu^{-1} (\sigma_h - \Pi\sigma, \tau) + (\sigma, \varepsilon(\varphi)) \right\} dt = \int_{I_m} F(\varphi) dt,
$$

for $\{ \varphi, \tau \} \in V_h \times W_h$. Here, the test functions $\{ \varphi, \tau \} \in V \times W$ are viewed as being constant over the time interval $I_m$. Notice that $\sigma_h \equiv 0$ on $I_m$. Integrating in (4.3.5) over $I_m$, we find analogous equations for the time-continuous solution $\{ \sigma, \sigma \}$,

$$
 (A[\sigma]^m, \tau) + \int_{I_m} \left\{ \left( A\dot{\sigma} - \varepsilon(v) \right), \tau \right\} + \mu^{-1} (\sigma - \Pi\sigma, \tau) + (\sigma, \varepsilon(\varphi)) \right\} dt = \int_{I_m} F(\varphi) dt,
$$

for $\{ \varphi, \tau \} \in V \times W$. Notice that $[\sigma]^m = 0$. Now, subtracting the foregoing identities yields for the error functions $e_v := v - v_h$ and $e_\sigma := \sigma - \sigma_h$ the relations

$$
 (A[e_\sigma]^m, \tau) + \int_{I_m} \left\{ \left( A\dot{e}_\sigma - \varepsilon(e_v) \right), \tau \right\} + \mu^{-1} (e_\sigma - \Pi\sigma + \Pi\sigma_h, \tau) + (e_\sigma, \varepsilon(\varphi)) \right\} dt = 0, \tag{4.6.12}
$$

for $\{ \varphi, \tau \} \in V_h \times W_h$. This identity may be interpreted as an incremental nonlinear Galerkin orthogonality relation for our discretization.

Suppose that the quantity to be controlled is the energy error $\|e_\sigma(T)\|_A$ at the final time $T$. For setting up such a space-time duality argument, we introduce the tensor-function $B = B(\sigma, \sigma_h)$ which is at each time $t \in I$ defined by

$$
 \Pi\sigma - \Pi\sigma_h := Be_\sigma.
$$

Then, the dual solution $\{ z, \pi \} : I \to V \times W$ is determined by the "backward-in-time" dual problem

$$
 -A\pi + \mu^{-1} (I - B)\pi - \varepsilon(z) = 0, \quad \text{div} \, \pi = 0, \quad \text{on} \, \ (0, T),
$$

$$
 \pi(T) = \| A e_\sigma(T) \|^{-1}_A A e_\sigma(T), \tag{4.6.13}
$$

supplemented by the appropriate boundary conditions. Integrating this over each time interval $I_m$, $m = M, \ldots, 1$, we obtain the incremental dual equations

$$
 -(\tau, A[\pi]^m) + \int_{I_m} \left\{ -(\tau, A\pi + \varepsilon(z)) + \mu^{-1} (\tau, (I - B)\pi) \right\} + (\varepsilon(\varphi), \pi) dt = 0,
$$

for all $\{ \varphi, \tau \} \in V \times W$, starting from the initial value $\pi(T)$ defined above. Notice that here again $[\pi]^m = 0$. Summing these equations for $m = 1, \ldots, M$, integrating by
parts, and taking \( \varphi = e_\nu \) and \( \tau = e_\sigma \) on each \( I_m \) results in the error representation

\[
\| e_\sigma(T) \|_A = \sum_{m=1}^{M} (A[e_\sigma]^m, \pi^m) + (e_\nu, \pi) \\
+ \int_I \left\{ (A \dot{e}_\sigma, \pi) + \mu^{-1}(e_\nu - \Pi \sigma + \Pi \sigma_h, \pi) + (e_\sigma, e(z)) \right\} dt,
\]

where the asterisk \( * \) indicates piecewise integration with respect to the subintervals \( I_m \). Using the orthogonality relation (4.6.12), we conclude that

\[
\| e_\sigma(T) \|_A = \sum_{m=1}^{M} (A[e_\sigma]^m, \pi^m - \bar{\pi}_h^m) + \int_I \left\{ (A \dot{e}_\sigma, \pi - \bar{\pi}_h) + (e_\nu, \pi - \bar{\pi}_h) \right\} dt,
\]

where \( \{ \bar{\pi}_h, \bar{\pi}_h \} : I \rightarrow V \times W \) are defined piecewise on each of the space/time elements \( Q_i^m := K \times [t_{m-1}, t_m] \) by local interpolation. Finally, recalling that \( \{ v, \sigma \} \) is the continuous solution, (4.6.15) reduces to

\[
\| e_\sigma(T) \|_A = \sum_{m=1}^{M} \sum_{K \in T_h} \left\{ - (A[e_\sigma]^m, \pi^m - \bar{\pi}_h^m) + (R^\nu, \pi - \bar{\pi}_h)_{K \times I_m} \\
+ (R^\sigma, z - \bar{z}_h)_{K \times I_m} + (r^\sigma, z - \bar{z}_h)_{K \times I_m} \right\} dt,
\]

with the domain residuals

\[
R^\nu (v_h, \sigma_h) := \dot{\sigma}_h + \mu^{-1}(\sigma_h - \Pi \sigma_h) - e(v_h), \quad R^\sigma (v_h, \sigma_h) := f + \text{div} \, \sigma_h,
\]

and the edge residuals \( r_\sigma \) as defined in Theorem 4.5.1. From (4.6.16), we deduce the following theorem.

**Theorem 4.6.2.** For the time stepping scheme (4.6.1) there holds the a posteriori error estimate

\[
\| e_\sigma(T) \|_A \leq \eta_\sigma := \sum_{m=1}^{M} k_m \sum_{K \in T_h} h_K^2 \left\{ (h_K^4 + k_m)^4 \theta_K^{m,1} \omega_K^{m,1} \right\},
\]

where the local residuals and weights are defined by

\[
\theta_K^{m,1} := h_K^{-1} k_m^{-1/2} \| R^\nu_{\sigma} \|_{K \times I_m}, \quad \theta_K^{m,2} := h_K^{-1} k_m^{-1/2} \| r^\sigma \|_{\partial K \times I_m}, \quad \theta_K^{m,3} := h_K^{-1} k_m^{-1/2} \| z^\nu \|_{\partial K \times I_m}, \quad \theta_K^{m,4} := h_K^{-1} k_m^{-1/2} \| A^{1/2} [\sigma_h]^m \|_{K},
\]

\[
\omega_K^{m,1} := h_K^{-1} k_m^{-1/2} \| R^\sigma \|_{K \times I_m}, \quad \omega_K^{m,2} := h_K^{-1} k_m^{-1/2} \| \pi^m - \bar{\pi}_h^m \|_{\partial K \times I_m}, \quad \omega_K^{m,3} := h_K^{-1} k_m^{-1/2} \| z^\nu - \bar{z}_h \|_{K \times I_m}, \quad \omega_K^{m,4} := h_K^{-1} \| A^{1/2} (\pi^m - \bar{\pi}_h^m) \|_{K},
\]

with \( \kappa := (h_K^2 + k_m)^{-1} \).
The weights $\omega_K^{(m,i)}$ have again to be evaluated by solving the dual problem approximately for a discrete dual solution $\{z_h, \pi_h\}$ as described in the preceding section.

To conclude our discussion, we have to analyse the behavior of the dual solution $\{z, \pi\}$ which determines the weights $\omega_K^{(m,i)}$. Unfortunately, the mathematical theory of problems in elasto-plasticity is not sufficiently developed to allow a complete regularity analysis. Therefore, the following argument will be largely heuristic. We note that $\|B\| \leq 1$ and, consequently,

$$((I - B)\tau, \tau) \geq 0, \quad \tau \in W. \quad (4.6.18)$$

This structural property gives us a simple bound on the dual solution. To end, we multiply the first equation in (4.6.13) by $\pi$, obtaining

$$-\frac{1}{2} \frac{d}{dt} (A\pi, \pi) + \mu^{-1}((I - B)\pi, \pi) - (\varepsilon(z), \pi) = 0.$$

Observing $\text{div} \pi = 0$ and (4.6.18) give us after integration over $[t,T]$

$$\|\pi(t)\|_A \leq \|\pi(T)\|_A = 1. \quad (4.6.19)$$

The derivation of analogous a priori bounds for space- or time-derivatives of the dual solution requires assumptions on the behavior of the quantity

$$\gamma_\mu := \mu^{-1} \int_I \|I - B\| \, dt, \quad (4.6.20)$$

for small $\mu$. Since it cannot be expected that $\gamma_\mu$ stays bounded as $\mu \to \infty$, such estimates make sense only for strictly positive regularization parameter $\mu > 0$, corresponding for example to the case of linear hardening. Therefore, it does not seem feasible to ask for analytical a priori bounds for the weights $\omega_K^{(m,i)}$, they rather have to be evaluated numerically.

### 4.6.3 Numerical tests for the quasi-static Prandtl-Reuss model

We present some results obtained with our adaptive method for the Prandtl-Reuss model. The test is done for the plane-strain benchmark problem “disc with a hole” (see Figure 4.4). We recall the quantities to be computed:

- Value $t_f$ for which plastification sets in.
- Value $t_b$ for which static collapse occurs.
- Diagram of $t$ over $L^1(u_2)$.
- Diagram of $t$ over $\sigma_{11}(P_f)$.
- Diagram of $\sigma_{11}(P_f)$ versus $t$ over the loading and unloading path $0 \leq t \leq 18$ (“saw-tooth” curve).

**Computation over the loading cycle**

First, we compute the values $t_f$ and $t_b$. To this end, we use the modified “energy-error” estimator (4.5.1) within an incremental loading process with constant loading step. The results are listed in Table 4.10.
Table 4.10 Values for $t_f$ and $t_b$ computed using fixed $k = 0.01$ and different numbers of cells (left), and with fixed $N = 8000$ and different time steps (right)

<table>
<thead>
<tr>
<th>Cells</th>
<th>$t_f$</th>
<th>$t_b$</th>
<th>$k$</th>
<th>$t_f$</th>
<th>$t_b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>3.095</td>
<td>4.655</td>
<td>0.08</td>
<td>1.880</td>
<td>4.680</td>
</tr>
<tr>
<td>2000</td>
<td>2.975</td>
<td>4.655</td>
<td>0.04</td>
<td>1.860</td>
<td>4.650</td>
</tr>
<tr>
<td>4000</td>
<td>2.485</td>
<td>4.655</td>
<td>0.02</td>
<td>1.950</td>
<td>4.665</td>
</tr>
<tr>
<td>8000</td>
<td>2.115</td>
<td>4.655</td>
<td>0.01</td>
<td>2.115</td>
<td>4.655</td>
</tr>
</tbody>
</table>

Now, we consider the pure loading process and compute the required diagrams for the quantities $L_T(u_2)$ and $\sigma_{11}(P_T)$ as function of time. First, we compute with fixed number of cells ($N = 8,000$) and different time steps $k \in \{0.01, 0.02, 0.04, 0.08\}$, to study the dependence on time discretization. Then, we compute with fixed time step $k = 0.01$ and different numbers of cells $N \in \{1000, 2000, 4000, 8000\}$, to study the dependence on space discretization. The results are shown in Figures 4.16.

![Load deflection curves](image-url)

**Figure 4.16** Load deflection curves for $L_T(u_2)$ (left) and for $\sigma_{11}(P_T)$ (right), for fixed number of cells and different time steps (upper row) and for fixed time step and different numbers of cells (lower row).
Figure 4.17 shows a sequence of zooms into adaptively refined meshes for computing $\sigma_{11}(P_t)$ over the loading path $3.6 \leq t \leq 4.2$.

![Zoom images](image)

**Figure 4.17** Zoom into dynamically adapted meshes with $N \approx 3200$ cells for computing $\sigma_{11}(P_t)$ over the loading path $3.6 \leq t \leq 4.2$.

**Computation over the full loading and unloading cycle**

We consider the dynamic case including loading and unloading. The boundary traction is now assumed as a saw-tooth function as shown in the upper graph of Figure 4.18.

![Graphs](image)

**Figure 4.18** Loading path $0 \leq t \leq 18$ and dynamic behaviour of $\sigma_{11}(P_t)$ (left) computed on time-varying meshes with about $N = 1200$ and $N = 4500$ and likewise adapted time-step distribution (right).
First, in Figure 4.18, we show the development of $\sigma_{11}(P_t)$ over the loading path $0 \leq t \leq 18$ computed with adaptively selected time steps and spatial meshes (with $N = 1200$ and $N = 4500$ cells) based on the mixed a priori/a posteriori error estimate (4.6.1). We also display the development of the time-step size distribution during the adaptation process. Notice that there is almost no difference visible anymore for the computations on $N = 1200$ and $N = 4500$ cells. From this, we conclude that on such meshes the total error is already dominated by the time-stepping error.

Next, investigate the time-step dependence of the result for $\sigma_{11}(P_t)$ displayed in Figure 4.18. To this end, we repeat the computation of $\sigma_{11}(P_t)$ on the finest mesh with $N = 4500$ cells with different (uniformly distributed) time steps $k = 0.25, ..., 0.0125$. The results are shown in Figure 4.19.

![Figure 4.19](image)

*Figure 4.19* Time behaviour of $\sigma_{11}(P_t)$ over the full loading/unloading path computed on a fixed mesh with $N = 4500$ cells with decreasing time-step sizes $k = 0.25, ..., 0.0125$ (left), and zoom into the critical part of the graph (right).

### THE EFFECT OF INCLUDING THE ACCELERATION TERM

All the results presented so far have been obtained on the basis of the quasi-static Prandtl-Reuss model. We want to investigate the effect of neglecting the acceleration term in the equilibrium law $\dot{v} - \text{div } \sigma = f$ for the present benchmark configuration. For completeness, we recall from Section 4.3 the form of a loading step (without spatial discretization) for the two versions of the model:

a) The quasi-static Prandtl-Reuss model:

$$(\sigma^m - \Pi(\sigma^{m-1} + k_m C \varepsilon(v^m)), \tau) = 0 \quad \forall \tau \in W,$$

$$(\sigma^m, \varepsilon(\varphi)) = (f^m, \varphi) \quad \forall \varphi \in V.$$

(4.6.21)

b) The dynamic Prandtl-Reuss model:

$$(\sigma^m - \Pi(\sigma^{m-1} + k_m C \varepsilon(v^m)), \tau) = 0 \quad \forall \tau \in W,$$

$$k_m^{-1}(v^m - v^{m-1}, \varphi) + (\sigma^m, \varepsilon(\varphi)) = k_m(f^m, \varphi) \quad \forall \varphi \in V.$$

(4.6.22)
In Figure 4.20, we present results obtained with the “quasi-static” scheme (a) and the “dynamic” scheme (b) on a fixed mesh with \( N = 2000 \) cells. The left picture shows the development of \( \sigma_{11}(P_t) \) over the loading/unloading path \( 0 \leq t \leq 12.8 \) computed with a fixed time step \( k = 0.025 \). The right picture shows the development of \( \sigma_{11}(P_t) \) for a loading which linearly increases until \( t = 4.5 \) after which it is frozen. The computation uses a fixed time step \( k = 0.1 \) in the static case and decreasing time steps \( k = 0.1, \ldots, 0.025 \) in the dynamic case. We see that including the dynamic term \( \dot{v} \) may affect the dynamic behaviour of the solution but also its stationary limits if the loading is kept constant after some time. Whether this effect may be viewed as significant depends on the particular characteristics of the problem considered.

The effect seen in Figure 4.20 cannot be blamed to the rough “saw-tooth” loading. This is demonstrated by Figure 4.21 which shows that the dependence of the dynamic behaviour of \( \sigma_{11}(P_t) \) on the smoothness of the loading path is rather weak.

---

**Figure 4.20** Comparison of “quasi-static” and “dynamic” computation of \( \sigma_{11}(P_t) \) for different loading paths.

**Figure 4.21** Comparison of the dynamic behaviour of \( \sigma_{11}(P_t) \) for rough “saw-tooth” loading and smooth sinusoidal loading.
4.7 CONCLUSION

In this work, we have discussed a general approach to error control and mesh adaptation for finite element approximations (the DWR method) in elasto–plasticity. By solving global dual problems a posteriori error estimates have been derived for the approximation of any given quantity of physical interest. These error bounds can be evaluated by numerically solving linearized versions of these dual problems. The main emphasis has been on the static Hencky and quasi–static Prandtl–Reuss model of perfect plasticity. The performance of the DWR method has been demonstrated for a standard two-dimensional benchmark problem “disc with a hole” including loading and unloading. In these tests the weighted error estimators prove to be asymptotically sharp and provide the basis of constructing “optimally” economical meshes. We expect that the DWR method will work equally well for the easier case of elasto-plasticity with hardening.

The DWR method is a rather universal approach to adaptivity in Galerkin finite element schemes. For a discussion from a more abstract angle, we refer to the survey paper by Ramacher [Ran99]. Other applications have been considered by Becker, Braack & Ramacher [BR95, BBR99] (viscous flows), Becker, Braack, Ramacher & Waguet [BBRW99, BR99] (chemically reactive flows), Kanschat [Kan97] (radiative transfer problems), Bangerth & Rannacher [BR99] (acoustic waves), and Becker, Kapp & Ramacher [BKR98] (optimal control).

REFERENCES


REFERENCES


