Multigrid Preconditioned Minimization Algorithms for Elastoplastic Type Problems*

Johanna Kienesberger

Institute of Analysis and Computational Mathematics

Department of Computational Mathematics and Optimization

Johannes Kepler University, Linz, Austria

Abstract

In this report, we give a brief review of the general theory of plasticity. Furthermore, we establish a useful connection between plasticity and a special kind of nonstandard, nonlinear heat conduction. The finite element approximation of both problems should have the same features.
This enables us to develop and test multigrid preconditioned minimization algorithms for the simpler finite element equation derived from the nonlinear heat conduction sample problem before using the algorithm for the more complicated plasticity problem in a forthcoming paper.

Keywords

Elastoplasticity, finite element method, optimization methods

1 Introduction

Elastoplasticity is an interesting subject in the broad field of mechanics of solids. The simpler case to consider is the one of pure elasticity, since the problem is linear as the stress depends linearly on the strain (material linear) and the strain depends linearly on the displacements (geometric linear). Adding a term describing the plastic part of the strain leads to some material nonlinearity. Now, the set of admissible stresses is determined by a yield function additionally depending on the hardening. Furthermore, this yield function characterizes the plastic behavior: Isotropic hardening, kinematic hardening, viscoplasticity or perfect plasticity. One important law to describe the time development is the Prandtl-Reuß normality law.
The dual formulation of the problem is equivalent to a minimization problem depending

*Work supported by the Austrian Fonds zur Förderung der wissenschaftlichen Forschung under grant SFB013/F1306

only on the displacement vector u and the plastic part p of the strain. Instead of the elasticity problem we consider the following scalar problem sharing the same effects. A very simplified case reads as follows:

$$\begin{aligned} f(u,p) &= \min_{v,q} f(v,q) \\ \text{with } f(v,q) &:= \int_{\Omega} \left(\frac{1}{2}|\nabla v - q|^2 + |q| + \frac{H}{2}|q|^2 \right) dx \end{aligned}$$

This problem is equivalent to a modified heat equation and the basis for numerical computations with FEPP, a finite element code in C++, see [11].
For the theory of elasticity and plasticity and their analytical background (variational inequalities) see e. g. [6], [7], [9], [14], and [15]. Multigrid literature can be found in e. g. [8], and [10]. For the theory of finite elements see e. g. [2], [3], and [5].
Section 2 gives an introduction to elastoplasticity in general, presents Sobolev spaces of abstract functions, and defines the problem and its dual formulation. The special case of isotropic hardening is also treated. Section 3 presents the simplified case of the heat equation and the connection with the mechanical background. The algorithms handling the problem of the "elastoplastic" heat equation, line search and a modified Newton's algorithm, can be found in Section 4, whereas the numerical results are presented in the last section.

2 Introduction to elastoplasticity

2.1 Basic equations

We consider the following problem: A deformed body Ω in $\mathbb{R}^n$ $(n = 2, 3)$ with Lipschitz continuous boundary has to satisfy certain conditions: Firstly, the stress field $\sigma \in L^2(\Omega, \mathbb{R}^{n\times n})$ has to fulfill, accordingly to the basic theorem of Cauchy,

$$\sigma = \sigma^T \quad \text{in } \Omega, \tag{1}$$

$$\operatorname{div} \sigma + b = 0 \quad \text{in } \Omega, \tag{2}$$

with b being a vector field of given body forces.
Secondly, the linearized Cauchy Green strain tensor is obtained by using the displacement vector $u \in H_0^1(\Omega)^n$:

$$\varepsilon(u) = \frac{1}{2}\left(\nabla u + (\nabla u)^T\right) \quad \text{a. e. in } \Omega. \tag{3}$$

In the case of small deformations the strain is split additively into two parts:

$$\varepsilon(u) = \mathbb{A}\sigma + p \qquad \text{a. e. in } \Omega \tag{4}$$

Here, $\mathbb{A}\sigma$ denotes the elastic, p the plastic part. The linear, symmetric, positive definite mapping $\mathbb{A}$ from $\mathbb{R}^{n\times n}$ to $\mathbb{R}^{n\times n}$ describes the linear elasticity, thus $\mathbb{C} = \mathbb{A}^{-1}$ is the elasticity tensor.
Purely elastic material behavior is characterized by $p = 0$, but for describing plasticity we need another material law in order to determine p. There are restrictions on the stress variables described by a dissipation functional φ, which is convex, and non-negative, but may also attain $+\infty$. The first restriction is

$$\varphi(\sigma, \alpha) < \infty \qquad \text{a. e. in } \Omega. \tag{5}$$

This inequality means that the hardening parameter α controls the set of admissible stresses. The pair (σ, α) is called generalized stresses and certain values are called admissible if $\varphi(\sigma, \alpha) < \infty$.
The time development of p and α is given by the Prandtl-Reuß normality law which states that for all other generalized stresses (τ, β) there holds:

$$\dot{p} : (\tau - \sigma) - \dot{\alpha} : (\beta - \alpha) \leq \varphi(\tau, \beta) - \varphi(\sigma, \alpha) \qquad \text{a. e. in } \Omega, \tag{6}$$

where, $\dot{p}$ denotes the time derivative of p, i. e. $\dot{p} = \frac{\partial p}{\partial t}$, and : is the scalar product of matrices such that $A : B = \sum_{i,j=1}^{n} A_{ij} \cdot B_{ij}$ for all $A, B \in \mathbb{R}^{n\times n}$.

2.2 Sobolev spaces of abstract functions

For the exact mathematical definition of our problem, we need Sobolev spaces of abstract functions defined on some (quasi) time interval $(0, T)$ with values in some separable Hilbert space.
Let us first introduce the abstract L^p-space

$$L^p(0, T; H) := \{f : (0, T) \to H; f \text{ measureable and } ||f(.)||_H \in L^p(0, T)\}$$

with some $p \in (0, \infty)$ and the natural norm

$$||f||_{L^p(0,T;H)} := ||\, ||f(.)||_H \,||_{L^p(0,T)} = \left(\int_0^T ||f(t)||_H^p dt \right)^{\frac{1}{p}},$$

where $||.||_H$ denotes the norm in the Hilbert space H. The abstract Sobolev spaces are defined in a similar way:

$$\begin{aligned} &W^{m,p}(0, T; H) := \\ &\qquad \{f \in L^p(0, T; H) : f \ m\text{-times differentiable and } f^{(1)}, f^{(2)}, \dots, f^{(m)} \in L^p(0, T; H)\} \end{aligned}$$

with norm

$$||f||_{W^{m,p}(0,T;H)} := ||(f^{(1)}, f^{(2)}, \dots, f^{(m)})||_{L^p(0,T;H)},$$

where $f^{(i)}$ denotes the i-th derivative of the function f.
We note that in the following, H is usually a Sobolev space of vector functions with respect to the spatial variable, e. g. $H = H_0^1(\Omega)^n \times L^2(\Omega, \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \times \mathbb{R}^m$ in Problem 1 below. A more detailed introduction to Sobolev spaces can be found in [1] or [17].

2.3 Definition of the problem

Now we are able to define the problem $(\mathbf{P})$ for $b(0) = 0$:

Problem 1 $(\mathbf{P})$ *We seek* $(u, p, \sigma, \alpha) = (u, (p, \sigma, \alpha)) \in W^{1,2}(0, T;\; H_0^1(\Omega)^n \times L^2(\Omega, \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \times \mathbb{R}^m))$ *with the initial condition* $(u, p, \sigma, \alpha)(0) = 0$ *such that (1) to (6) are satisfied.*

There exists a given yield function Φ for the dissipation functional φ, such that the admissible generalized stresses are characterized by

$$\Phi(\sigma, \alpha) \le 0 \qquad \text{in } \Omega. \tag{7}$$

To guarantee the convexity of the set of admissible stresses, the function Φ itself has to be convex. Then, φ is the characteristic functional of the set of admissible generalized stresses, i. e.,

$$\varphi(\sigma, \alpha) = \begin{cases} 0 & \text{if } \Phi(\sigma, \alpha) \le 0 \\ \infty & \text{if } \Phi(\sigma, \alpha) > 0 \end{cases}$$

Now we can simplify (5) and (6) to:

$$\begin{aligned} \Phi(\sigma, \alpha) &\le 0 \\ \dot{p} : (\tau - \sigma) - \dot{\alpha} : (\beta - \alpha) &\le 0 \end{aligned}$$

2.3.1 Isotropic hardening

In the case of isotropic hardening the dimension m of the hardening parameter α is 1, i. e. α is a scalar. The yield function Φ is defined by:

$$\Phi(\sigma, \alpha) := |\operatorname{dev} \sigma| - \sigma_y(1 + \alpha H)$$

Here $\sigma_y > 0$ is the initial yield stress and $H > 0$ the modulus of hardening. dev is the deviation and defined by $\operatorname{dev} A := A - \operatorname{tr}(A) \cdot I_n$, where $\operatorname{tr}(A) := \sum_{i=1}^{n} A_{ii}$ is the trace of a matrix.
There must hold $\alpha \ge 0$, as otherwise $\Phi(\sigma, \alpha) = +\infty$. This implies that accordingly to (7) the stresses would not be admissible.

2.4 Dual formulation

The time dependent variational inequality (6) is solved with an implicit time discretization, for example generalized midpoint rules like Crank-Nicholson or backward Euler schemes. An implicit Euler discretization for the weak formulation leads to the next problem definition:

Problem 2 ($\mathbf{P}^*_{\Delta t}$) *Let $H \subseteq H^1_0(\Omega)^n, L^{n\times n}_{sym} \subseteq L^2(\Omega, \mathbb{R}^{n\times n}_{sym})$ and $L^m \subseteq L^2(\Omega, \mathbb{R}^m)$ be closed subspaces. We seek $(u, p, \sigma, \alpha) \in H \times L^{n\times n}_{sym} \times L^{n\times n}_{sym} \times L^m$, such that for given $u_0 \in H$; $p_0, \sigma_0 \in L^{n\times n}_{sym}$ and $\alpha_0 \in L^m$ at some time step t_0 the following conditions are satisfied for $t_1 = t_0 + \Delta t$:*

$$\int_\Omega \sigma : \varepsilon(v)\, dx = \int_\Omega b\, v\, dx \quad \forall v \in H \tag{8}$$

$$\int_\Omega \{(p - p_0) : (\tau - \sigma) - (\alpha - \alpha_0) : (\beta - \alpha)\}\, dx \leq \Delta t \int_\Omega \varphi(\tau, \beta)\, dx - \int_\Omega \varphi(\sigma, \alpha)\, dx \tag{9}$$

for all (τ, β).

The inequality (9) reads as

$$\frac{1}{\Delta t}(p - p_0, -\alpha + \alpha_0) \in \partial\varphi(\sigma, \alpha) \tag{10}$$

The sub-differential $\partial\varphi(b)$ is defined by

$$a \in \partial\varphi(b) \Leftrightarrow \int_\Omega a : (c - b)\, dx \leq \int_\Omega \varphi(c)\, dx - \int_\Omega \varphi(b)\, dx$$

Since φ is convex, the above equation is equivalent to:

$$(\sigma, \alpha) \in \partial\varphi^*(\frac{1}{\Delta t}(p - p_0, -\alpha + \alpha_0)) \tag{11}$$

φ^* is the dual functional of φ, which is computed by the Fenchel transformation:

$$\varphi^*(y) := \sup_x \{y : x - \varphi(x)\}$$

Using $\sigma = \mathbb{C}\,(\varepsilon(u) - p)$, we can omit the variable σ and we obtain the simplified problem $\mathbf{P}_{\Delta t}$:

Problem 3 ($\mathbf{P}_{\Delta t}$) *Find an $(u, p, \alpha) \in H \times L^{n\times n}_{sym} \times L^m$, such that for all (v, q, β) the following conditions are satisfied:*

$$\int_\Omega \mathbb{C}[\varepsilon(u) - p] : \varepsilon(v) dx = \int_\Omega b\, v\, dx$$

$$\int_\Omega \{\mathbb{C}[\varepsilon(u) - p] : (\Delta t\, q - p + p_0) + \alpha : (-\Delta t \beta + \alpha_0 - \alpha)\} dx \leq$$

$$\leq \Delta t \int_\Omega \varphi^*(q, \beta) dx - \Delta t \int_\Omega \varphi^*(\frac{p - p_0}{\Delta t}, \frac{\alpha_0 - \alpha}{\Delta t}) dx$$

Problem $(\mathbf{P}_{\Delta t})$ is the stationary condition of a minimizer in the following minimization problem $(\mathbf{M})$, and a minimizer of f of Problem $(\mathbf{M})$ is a solution of $(\mathbf{P}_{\Delta t})$:

Problem 4 $(\mathbf{M})$ *Find the minimizer* $(u, p, \alpha) \in H \times L^{n\times n}_{sym} \times L^m$ *of*

$$\begin{aligned} f(u,p,\alpha) &:= \frac{1}{2}\int_\Omega \mathbb{C}[\varepsilon(u)-p] : (\varepsilon(u)-p)dx + \frac{1}{2}\int_\Omega |\alpha|^2 dx \\ &+ \Delta t \int_\Omega \varphi^*(\frac{p-p_0}{\Delta t}, \frac{\alpha_0-\alpha}{\Delta t})dx - \int_\Omega b\, u\, dx \end{aligned}$$

2.4.1 Computation of the dual functional φ^* in the case of isotropic hardening

This subsection was mainly taken from [4]. The case of isotropic hardening is characterized by:

$$\begin{aligned} \varphi(\sigma,\alpha) &= \begin{cases} 0 & \text{if } \Phi(\sigma,\alpha) \le 0, \\ \infty & \text{if } \Phi(\sigma,\alpha) > 0, \end{cases} \\ \text{with } \Phi(\sigma,\alpha) &= |\operatorname{dev}\sigma| - \sigma_y(1+\alpha H). \end{aligned}$$

The dual functional will be computed for values from the set of admissible stresses, i. e. $\varphi(\sigma,\alpha)=0$:

$$\varphi^*(A,B) = \sup_{|\operatorname{dev}\sigma| \le \sigma_y(1+\alpha H)} (\sigma : A + \alpha B)$$

The pair $(c\cdot I_n, 0)$ is element of the set of admissible stresses for all $c \in \mathbb{R}$. The scalar product $\sigma : A$ (and the supremum, too) is finite only if $\operatorname{tr} A = 0$ for all matrices A. Then, $\operatorname{dev} A = A$.
The other admissible pair considered is $(\lambda A, \alpha)$ with $\lambda = \frac{\sigma_y(1+H\alpha)}{|A|}$. Here, the scalar product is

$$\begin{aligned} \sigma : A + \alpha B &= \lambda A : A + \alpha B = \lambda |A|^2 + \alpha B \\ &= \sigma_y(1+H\alpha)|A| + \alpha B \\ &= \sigma_y|A| + \alpha(\sigma_y H|A| + B) \end{aligned}$$

Since $(\sigma : A + \alpha B) \le \varphi^*(A,B)$, $(\sigma_y H|A| + B)$ must be ≤ 0 in order to obtain a finite supremum.
These calculations prove:

$$\varphi^*(A,B) = \begin{cases} \sigma_y|A| & \text{if } \operatorname{tr} A = 0 \wedge (\sigma_y H|A| + B) \le 0 \\ \infty & \text{if } \operatorname{tr} A \ne 0 \vee (\sigma_y H|A| + B) > 0 \end{cases}$$

The minimization with respect to α in problem $(\mathbf{M})$ concerns only the term $\frac{1}{2}\int_\Omega |\alpha|^2$ under the restriction $(\sigma_y H|A| + B) \le 0$ with $A = \frac{p-p_0}{\Delta t}$ and $B = \frac{\alpha_0-\alpha}{\Delta t}$. The unique solution is

$\alpha = \alpha_0 + \sigma_y H|p - p_0|$. So we obtain a simplified minimization problem:
Find the minimizer (u,p) of

$$\begin{aligned} f(u,p) &:= \tfrac{1}{2}\int_\Omega \mathbb{C}[\varepsilon(u)-p] : (\varepsilon(u)-p)dx + \tfrac{1}{2}\int_\Omega (\alpha_0 + \sigma_y H|p-p_0|)^2 dx \\ &\quad + \int_\Omega \sigma_y |p-p_0| dx - \int_\Omega b\, u\, dx \end{aligned} \tag{12}$$

under the side restriction $\mathrm{tr}(p - p_0) = 0$. The uniqueness of the minimizer follows from the properties of the dual functional, see [6].

3 Heat equation

One of the easiest examples for applying the theory is a scalar diffusion problem, e. g. a heat equation. From a physical point of view it seems that the mechanical problem of the last chapter and the heat equation presented here have hardly anything in common, except maybe that they are both differential equations. But both problems are based on the same formulas with only the variables having different physical interpretations:

mechanics	heat equation (scalar)
displacement u	temperature u
strain $\varepsilon(u)$	temperature gradient ∇u
stress σ	heat flux p

The basic differential equation of heat conduction reads as follows:

$$\mathrm{div}\, p = f, \tag{13}$$

where f are given heat sources and the flux p has a special form:

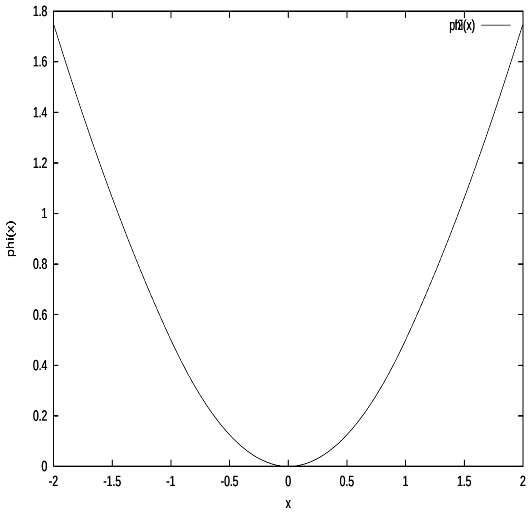
$$p = \lambda(|\nabla u|)\, \nabla u \tag{14}$$

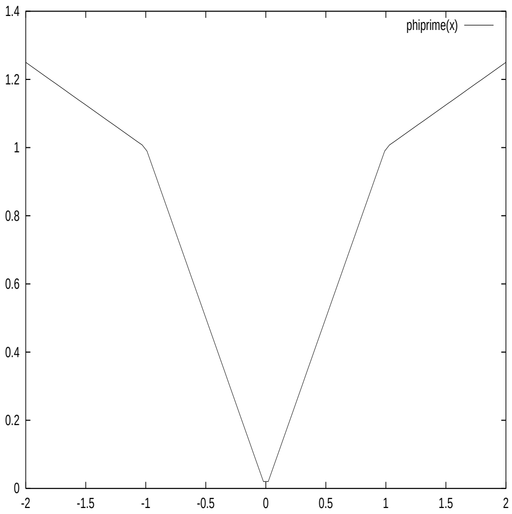
$$\text{with } \lambda(|\nabla u|) = \frac{\varphi'(|\nabla u|)}{|\nabla u|} \tag{15}$$

The function φ is defined as follows:

$$\varphi(x) = \begin{cases} \frac{1}{2}x^2 & \text{if } |x| \leq 1 \\ |x| - \frac{1}{2} + \frac{1}{2}\, c\, (|x|-1)^2 & \text{if } |x| \geq 1 \end{cases}$$

The pictures show φ and φ' for $c = 0.25$:





The minimization problem now has the form:

$$
\begin{aligned}
J(u) &= \min_{v \in H} J(v) \\
J(v) &= \int\limits_{\Omega} \varphi(|\nabla v|) dx - \int\limits_{\Omega} f\ v\ dx
\end{aligned}
$$

Proof: u being the minimizer is equivalent to the derivative of the objective being equal to zero:

$$
\begin{aligned}
0 &= \langle J'(u), v \rangle \\
&= \int\limits_{\Omega} \varphi'(|\nabla u|) \frac{\langle \nabla u, \nabla v \rangle}{|\nabla u|} dx - \int\limits_{\Omega} f\ v\ dx \\
&= \int\limits_{\Omega} p\ \nabla v\ dx - \int\limits_{\Omega} f\ v\ dx
\end{aligned}
$$

This leads to

$$
\operatorname{div} p = f
$$

We still have to show that the mechanical problem leads to a special kind of heat equation:
The scalar version of the mechanical problem with isotropic hardening writes for $\mathbb{C} = I_n$, $\sigma_y = 1$, $\alpha_0 = p_0 = 0$ and due to the scalar case $\varepsilon(u) = \nabla u$ as follows:

$$
\min_{u,p} \int\limits_{\Omega} \left\{ \frac{1}{2} |\nabla u - p|^2 + |p| + \frac{H}{2} |p|^2 - f\ u \right\} dx \tag{16}
$$

This is equivalent to:

$$\min_u \left\{ \min_p \left\{ \int_\Omega (\frac{1}{2}|\nabla u - p|^2 + |p| + \frac{H}{2}|p|^2) dx \right\} - \int_\Omega f\ u\ dx \right\}$$

We may interchange the inner minimization and the integration, since the argument doesn't depend on a derivative of p. Now, we are able the minimize the argument pointwise a. e. in $x \in \Omega$. For $\nabla u(x)$ given, find $p(x)$, such that

$$\frac{1}{2}|\nabla u(x) - p(x)|^2 + |p(x)| + \frac{H}{2}|p(x)|^2 = \min! \tag{17}$$

for all $x \in \Omega$ a. e.. We may assume, that p points in the direction of ∇u, i. e. we assume $p = \tilde{c}\,\nabla u$. So we obtain a minimization problem concerning $\tilde{c}$:

$$\min_{\tilde{c}\geq 0} \frac{1}{2}(1-\tilde{c})^2|\nabla u|^2 + \tilde{c}|\nabla u| + \frac{H}{2}\tilde{c}^2|\nabla u|^2$$

The minimum is at $\tilde{c} = \frac{1}{1+H}(1 - \frac{1}{|\nabla u|})$, which corresponds to $p = \frac{1}{1+H}(\nabla u - \frac{\nabla u}{|\nabla u|})$. Using this p in (17) we get the value of the minimum:

$$\frac{1}{1+H}\left(\frac{H}{2}|\nabla u|^2 + |\nabla u| - \frac{1}{2}\right)$$

This function is the same as $\varphi(|\nabla u|)$, if we set $c = \frac{H}{1+H}$. Now, we have the connection between the mechanical and the heat equation problem.
From setting the constant $c = 0$ it follows that $H = 0$, which is the case of perfect plasticity. The material (or in our case: the heat) flows without any hardening.

4 Algorithms

The necessary first oder condition on a minimizer $u \in H$ of

$$\min_{v\in H} J(v)$$

reads as

$$A(u) = f. \tag{18}$$

Multiplying (18) with a test function v leads to

$$\langle A(u), v\rangle = \langle f, v\rangle$$

which yields

$$\int_\Omega \varphi'(\nabla u)\frac{<\nabla u, \nabla v>}{|\nabla u|}\ dx = \int_\Omega f\ v\ dx.$$

The iterations are terminated if either the number of iterations is larger than maxit, which is chosen as 40, or the condition $|J(u^{n+1}) - J(u^n)| < eps$, e. g. $eps = 10^{-14}$, is fulfilled.

4.1 Preconditioned steepest descent method

Using preconditioned steepest descent method the iteration looks as follows:

$$u^{n+1} = u^n + \tau C^{-1}(f - A(u)), \tag{19}$$

where, C is the preconditioning operator, and τ the damping parameter.
This method faces a problem: If u^n is domain of φ with linear behavior (in the case of perfect plasticity) or nearly linear behavior if the constant c is small, it takes a lot of iterations to find the minimum. A better approach is to compute a quadratic function φ^n for each u^n, such that φ^n takes its minimum at 0 and φ and φ^n coincide at u^n as well as the derivatives of both functions. Searching the minimum for φ^n with a starting point in the domain of linear behavior, we assume that φ^n has the following form, as it is a symmetric quadratic function:

$$\varphi^n(x) = \frac{1}{2}ax^2 + b$$

For $x \leq 1$ φ and φ^n are identical, if $x > 1$ we have to evaluate and equate the derivatives: $ax = 1 + c(x-1)$. Thus, we obtain $a = \frac{1+c(x-1)}{x}$.
The following, more generalized lemma resulted from [13]:

Theorem 1 *Let V be a Hilbert space with scalar product $(.,.)_V$ and norm $||.||_V$, $J : V \to \mathbb{R}$ a functional and $||.||_{\bar{A}_u}$ a family of uniformly equivalent Hilbert space norms. Using the quadratic functionals*

$$\begin{aligned} \bar{J}_u(v) &:= J(u) + \langle J'(u), v-u\rangle + \frac{1}{2}||u-v||^2_{\bar{A}_u} &(20)\\ \underline{J}_u(v) &:= J(u) + \langle J'(u), v-u\rangle + \frac{\alpha}{2}||u-v||^2_{\bar{A}_u} &(21)\end{aligned}$$

with $0 < \alpha \leq 1$ we can state

$$\underline{J}_u(v) \leq J(v) \leq \bar{J}_u(v) \qquad \forall u, v \in V \tag{22}$$

Note, that $\underline{J}_u(u) = J(u) = \bar{J}_u(u)$. Then, the iteration

$$u^{k+1} = argmin \bar{J}_{u^k}(v) \tag{23}$$

converges to the unique minimum u^* of $J(.)$.

Proof: From (22) we see that

$$\min_v \underline{J}_{u^k}(v) \leq J(u^{k+1}) \leq \bar{J}_{u^k}(u^{k+1})$$

The Hilbert space norm $||.||_{\bar{A}_u}$ induces a scalar product $\bar{A}_u(.,.)$. From (23) we obtain $\bar{J}_{u^k}(u^k) = \min_w \bar{J}_{u^k}(w)$. Its necessary first order condition yields $\langle \bar{J}'_{u^k}(u^{k+1}), v\rangle = 0$ for all $v \in V$, implying

$$\langle J'(u^k), v\rangle + \bar{A}_{u^k}(u^{k+1} - u^k, v) = 0 \qquad \forall v \in V.$$

Using the definition of $\bar{J}_u(v)$ we obtain the descent in the k-th step

$$\begin{aligned}\bar{J}_{u^k}(u^{k+1}) - \bar{J}_{u^k}(u^k) &= J(u^k) + \langle J'(u^k), u^{k+1} - u^k\rangle + \frac{1}{2}||u^{k+1} - u^k||^2_{\bar{A}_{u^k}} - J(u^k) \\ &= \langle J'(u^k), u^{k+1} - u^k\rangle + \frac{1}{2}||u^{k+1} - u^k||^2_{\bar{A}_{u^k}} \\ &= -\bar{A}_{u^k}(u^{k+1} - u^k, u^{k+1} - u^k) + \frac{1}{2}||u^{k+1} - u^k||^2_{\bar{A}_{u^k}} \\ &= -\frac{1}{2}||u^{k+1} - u^k||^2_{\bar{A}_{u^k}}\end{aligned}$$

Differentiating the definition of $\underline{J}_{u^k}$ and setting it equal to zero at its minimum v^*, we get

$$\langle \underline{J}'_{u^k}(v^*), v\rangle = \langle J'(u^k), v\rangle - \alpha\bar{A}_{u^k}(u^k - v^*, v) = 0 \qquad \forall v \in V$$

Then, the descent of the lower bound writes as

$$\begin{aligned}\min_v \underline{J}_{u^k}(v) - \underline{J}_{u^k}(u^k) &= \langle J'(u^k), v^* - u^k\rangle + \frac{\alpha}{2}||u^k - v^*||_{\bar{A}_{u^k}} \\ &= -\alpha\bar{A}_{u^k}(u^k - v^*, u^k - v^*) + \frac{\alpha}{2}||u^k - v^*||_{\bar{A}_{u^k}} \\ &= -\frac{\alpha}{2}||u^k - v^*||_{\bar{A}_{u^k}} \\ &= -\frac{1}{2\alpha}||u^{k+1} - u^k||_{\bar{A}_{u^k}}\end{aligned} \tag{24}$$

The last equality is valid, since (comparing both derivatives we have computed) $\bar{A}_{u^k}(u^{k+1} - u^k, v) = \alpha\bar{A}_{u^k}(v^* - u^k, v)$ for all v, thus $\alpha(v^* - u^k) = u^{k+1} - u^k$. Recapitulating these considerations, we obtain linear q-convergence of the functional:

$$\begin{aligned}0 \le J(u^{k+1}) - J(u^*) &\overset{(22)}{\le} \bar{J}_{u^k}(u^{k+1}) - J(u^*) \\ &\overset{(20)}{=} J(u^k) - \frac{1}{2}||u^{k+1} - u^k||^2_{\bar{A}_{u^k}} - J(u^*) \\ &\overset{(24)}{=} J(u^k) + \alpha(\min_v \underline{J}_{u^k}(v) - \underline{J}^k_u(u^k)) - J(u^*) \\ &\overset{(22)}{\le} J(u^k) + \alpha(\min_v J(v) - J(u^k)) - J(u^*) \\ &= (1-\alpha)(J(u^k) - J(u^*))\end{aligned}$$

Due to the uniform norm equivalence we have linear r-convergence: $||u^{k+1} - u^k||^2_V \le c(1-\alpha)^k$ and thus convergence in V. The existence of u^* and v^* is given by the continuity and coercivity of the problems.

4.2 Line search and modified Newton's algorithm

The Newton iteration is modified in the following way:

$$u_i^{n+1} := u^{n+1} = u^n + \frac{1}{2^i}H^{-1}(f - A(u)), \tag{25}$$

where, H is the Hessian and i controls the step size as follows: $u^{n+1} := u_j^{n+1}$ with $C = H$, $\tau = \frac{1}{2^i}$, and $j = \operatorname{argmin}\{i \in \mathbb{N}_0 \mid J(u_i^{n+1}) < J(u^n)\}$.
It is necessary to combine Newton's algorithm with the line search to obtain convergence. Note, that for computing the Hessian we need φ''.

5 FE-Discretization and numerical results

5.1 FE-Discretization

Now, let V_h be a finite dimensional sub-space of V. Instead of the function u, we now search a function $u_h \in V_h$, defined by $u_h := \sum_{i=1}^{N_h} u^{(i)} p^{(i)}$. Here, $p^{(i)}$ are the standard finite element basis functions of V_h, N_h is the number of basis functions, and $u^{(i)}$ the coefficients to determine.
The most computational effort is applying the preconditioning matrix C^{-1}. The performance of the algorithm depends strongly on the preconditioners C. The application of C^{-1} shall be fast and the parameter α of (21) should be close to 1. These requirements are satisfied by multigrid methods [8].
In FEPP, finding a solution works as follows: Starting on the first level with a triangular grid, we compute the solution. On the second level, the grid from the previous level is refined either uniformly or adaptively. The solution is now computed by means of the suggested method and using multigrid preconditioning.
For both algorithms we use $u = 0$ as initial solution on the first level. For the other levels the interpolated solution of the previous level provides a good initial guess.

5.2 Preconditioned steepest descent method

Based on the considerations in the last chapters about this method, the assembling of the element matrix A_e is modified as follows:

$$A_e = \sum_{ip} w_i |det\ J| (\nabla p)^T M \nabla p$$

where ip are the integration points, w_i the weights, p the form functions, J is here the Jacobian of the mapping from the reference element to the actual element and M is a diagonal matrix with entries either 1 for $|\nabla u| \leq 1$ or $\frac{1+c(|\nabla u|-1)}{|\nabla u|}$ for $|\nabla u| > 1$.
Then, the matrix A is assembled from the modified element matrices.
Two different concepts were chosen to determine the preconditioning matrix C: Firstly, C was calculated as the preconditioning matrix for a linearized A, as if solving Laplace's equation where the flux depends linearly on the temperature derivative. This approach failed, as the number of CG-iterations exploded and the iterations themselves were not stable. This outcome was no big surprise, because using a preconditioning matrix not fitting the problem usually is not a good idea.
The second concept was calculating C as the preconditioning matrix for the non-linear matrix A. In this case, the CG-iterations were stable.

5.3 Newton's algorithm

To compute the Hessian we need φ'', which reads for the dimension $n = 2$ as:

$$\varphi''(x) = \begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{if } |x| \leq 1 \\ cI + (1-c)\frac{1}{|x|} \begin{pmatrix} 1 - \frac{x_1^2}{|x|^2} & -\frac{x_1 x_2}{|x|^2} \\ -\frac{x_1 x_2}{|x|^2} & 1 - \frac{x_2^2}{|x|^2} \end{pmatrix} & \text{if } |x| \geq 1 \end{cases}$$

In the part of the program where the Hessian is assembled, each element matrix is computed by:

$$A^T = \sum_{ip} w_i |det\ J| (\nabla p)^T \varphi'' \nabla p$$

where ip are the integration points, w_i the weights, p the form functions and J is the Jacobian of the mapping from the reference element to the actual element.

5.4 L-shaped domain

As a test case for the multigrid solver we chose an L-shaped domain, which is turned upside down. On the right boundary a Neumann condition and on the lower boundary a homogenous Dirichlet condition have to be satisfied. The Neumann condition is a given flux with intensity 0.8, in order to fulfill the safe-load assumption.
We chose the constant c very small, i. e. $c = 10^{-6}$, which means that there is hardly any hardening and that we are quite near to the case of perfect plasticity.
Using adaptive refinement we obtain the grid shown in figure 1 (here for level 7): The nu-

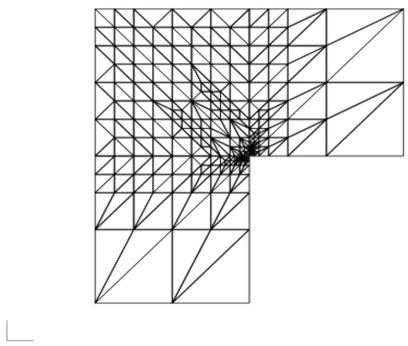


Figure 1: Adaptively refined grid

meric solutions obtained by using the two different algorithms are nearly indistinguishable,

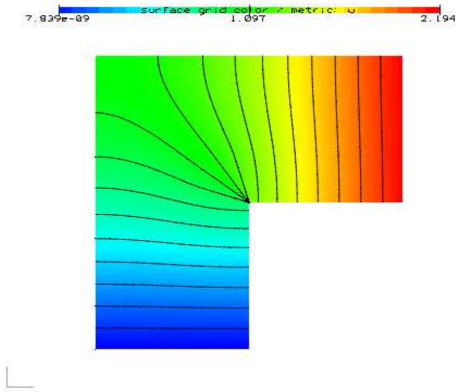


Figure 2: Solution u

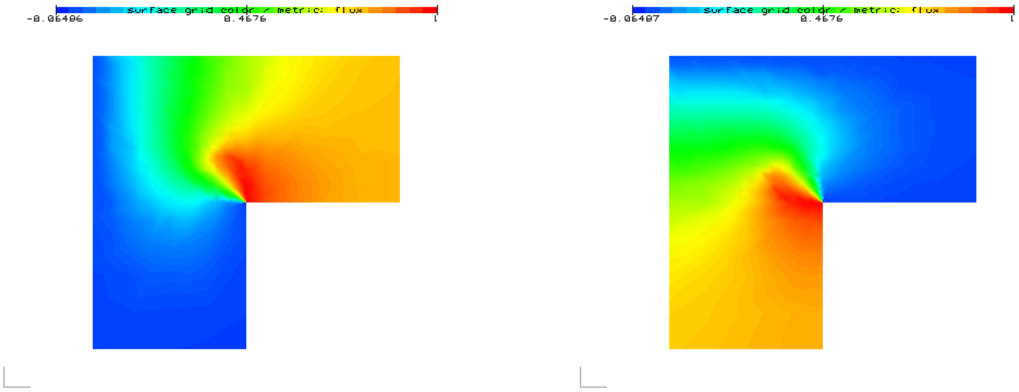


Figure 3: First and second component of the flux

but their CPU times aren't. In fact they obviously differ as it will be shown in figure 5. Figure 2 shows the solution u on level 15, with iso-lines. The numerical flux is symmetric in the two coordinates, shown in figure 3. It is quite evident, that the main difficulties occur at the sharp corner. Note that the numerical flux is computed as $p = \frac{\varphi'(|\nabla u|)}{|\nabla u|}\nabla u$. The plasticity domain (that's where $|\nabla u| > 1$) is the red area in figure 4. The CPU times of the two algorithms are almost the same for the first levels, because the modified Newton's algorithm takes more time to compute H^{-1} in each iteration, but fewer iterations. The maximum number of iterations is 40 in our example, which is never reached by the Newton algorithm, but on the other side the preconditioned steepest descent method needs less than 40 iterations only on the second level. Thus, the steepest descent method hardly ever satisfies the truncation condition, such that for higher levels the CPU times are much lower, see figure 5. The number of Newton iterations is really good, as you can see in table

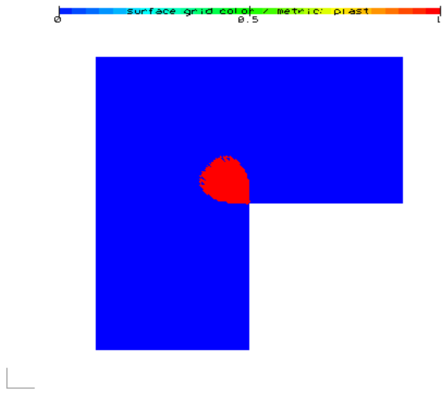


Figure 4: Plasticity domain (red)

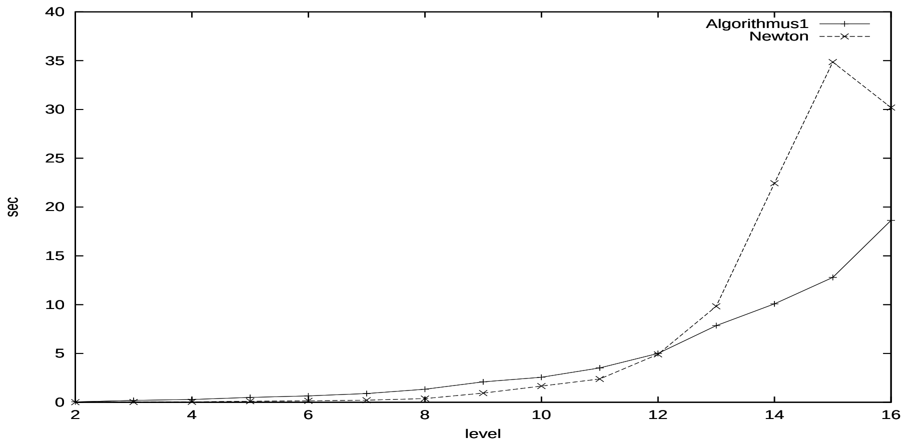


Figure 5: CPU times

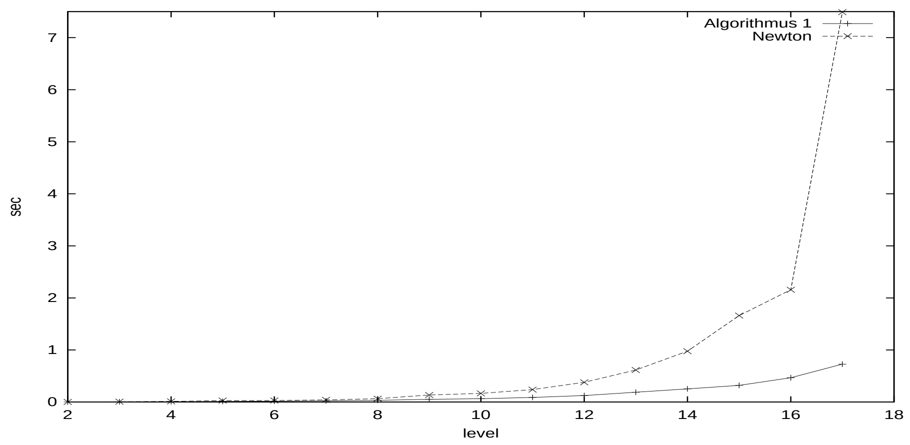


Figure 6: CPU time per iteration

1.

level	2	3	4	5	6	7	8	9	10
iterations	4	4	4	4	5	5	6	7	10
deg of freedom	21	51	75	124	161	221	335	509	626

level	11	12	13	14	15	16
iterations	10	13	16	23	21	14
deg of freedom	837	1220	1891	2309	3239	4584

Table 1

Still the CPU times increase dramatically for higher levels, as mentioned above. In figure 6 the times each iteration takes on the different levels are displayed for both algorithms. Both algorithms have nearly identical convergence properties. It only makes a difference whether adaptively or uniformly refined grids are used: Figure 7 shows the logarithmic error of the Ritz value according to the logarithmic number of unknowns (in the system of equations).

5.5 Quarter of a ring

In the second example we consider a quarter of a ring to see what happens, if there are no corners with interior angle larger than π any more. The left figure in figure 8 shows the solutions with the same boundary conditions as for the L-shaped domain. The right figure shows the plasticity domain.

Remarks

Still, there's a lot of work to do: Implementing the steady state mechanical problem in 2D, then in 3D. The next steps would be making the problem quasi static, trying to find error

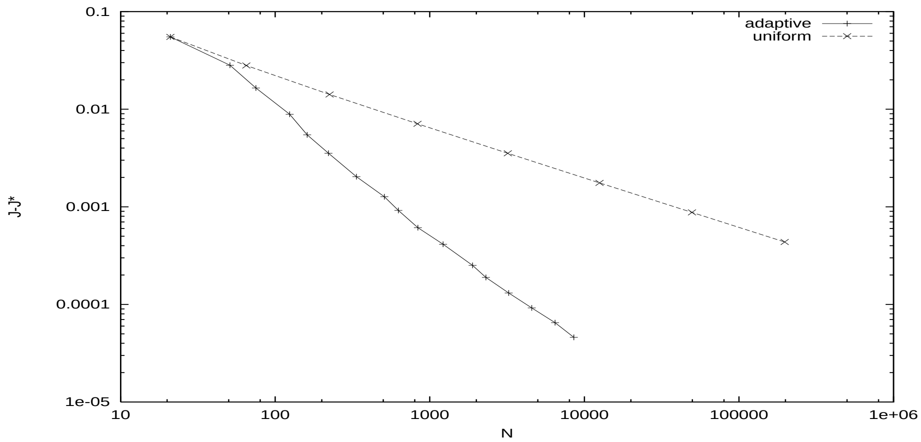


Figure 7: Convergence

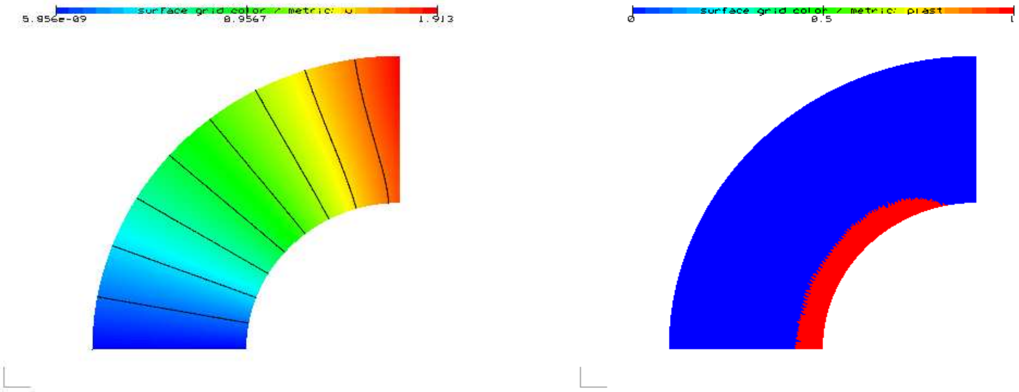


Figure 8: Solution and plasticity domain (red) of a quarter of a ring

estimates and proving convergence.

[]

References

[1] R. A. Adams, *Sobolev spaces*, Academic Press, New York, 1976.

[2] D. Braess, *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*, Cambridge University Press, 1997.

[3] S. Brenner and L. R. Scott, *The mathematical theory of finite element methods*, Springer, Berlin, Heidelberg, New York, 1994.

[4] C. Carstensen, *Domain decomposition for a nonsmooth convex minimization problem and its application to plasticity*, NLAA **4** (1997), no. 3, 1–13.

[5] P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, 1978.

[6] I. Ekeland and R. Temam, *Convex analysis and variational problems*, North-Holland, Amsterdam, 1976.

[7] R. Glowinski, J.-L. Lions, and R. Trémolières, *Numerical analysis of variational inequalities*, North-Holland Publishing Company, Amsterdam - New York - Oxford, 1981.

[8] W. Hackbusch, *Multigrid methods and applications*, Springer, Berlin, Heidelberg, New York, 1985.

[9] W. Han and B. D. Reddy, *Computational plasticity: Mathematical theory and numerical analysis*, Springer, Berlin, Heidelberg, New York, 1999.

[10] M. Jung, U. Langer, A. Meyer, W. Queck, and M. Schneider, *Multigrid preconditioners and their applications*, Third Multigrid Seminar, Biesenthal 1988 (G. Telschow, ed.), Report R-MATH-03/89, 1989, pp. 11–52.

[11] M. Kuhn, U. Langer, and J. Schöberl, *Scientific computing tools for 3D magnetic fiels problems*, The Mathematics of Finite Elements and Applications X (J. R. Whiteman, ed.), Elsevier, 2000, pp. 239 – 258.

[12] U. Langer, *Skriptum zur Vorlesung Numerik I (Gleichungen)*, Institut für Mathematik, JKU Linz, 1995.

[13] J. Schöberl, *private communication*, Linz, 1999.

[14] J. C. Simo and T.J.R. Hughes, *Computational inelasticity*, Springer, Berlin, Heidelberg, New York, 1998.

[15] K. Washizu, *Variational methods in elasticity and plasticity*, Pergamon, Oxford, 1975.

[16] C. Wieners, *Theorie und Numerik der Prandtl-Reuß-Plastizität*, 1999, habilitation thesis, Heidelberg University.

[17] E. Zeidler, *Nonlinear functional analysis and its applications*, vol. 2,3,4, Springer, New York, 1990,1985,1988.