Appendix 1

Performance analysis of Newton type and adaptive continuation solvers for nonlinear problems in geomechanics

R. Blaheta and P. Byczanski
Institute of Geonics AS CR, Ostrava, Czech Republic

HIPERGEOS Report, IGAS Ostrava, 2002

SUMMARY

The paper describes special variants of the damped Newton and continuation methods. The solvers based on these methods are implemented and shown to be efficient and robust when solving nonlinear elasticity problems.

1 Introduction

In many applications of mathematical modelling, it is necessary to consider nonlinear effects and solve large nonlinear systems arising from discretization of the boundary value problems. Such situations are quite common in geomechanics, where it may be necessary to consider nonlinear behaviour of soils and rocks, see e.g. [13], [14], [10]. For the solution of the nonlinear systems, which arise from the FE analysis of the above problems, we need solvers, which are efficient, robust and parameter free or only weekly depending on a few control parameters.

In this paper, we shall seek such solvers in two classes of methods, which are both based on linearization. Namely, we shall consider damped inexact Newton methods with a suitable damping strategy and continuation methods with an adaptive procedure for defining the size of continuation steps. In both cases, we shall exploit the condition of sufficient accuracy of the linearization, which enables to construct the solvers fulfilling the above requirements. The considered solvers was firstly introduced in [7] and [8].

The efficiency and robustness of the described methods is tested on the solution of nonlinear soil mechanics problems, which belong among elasticity problems with nonlinear behaviour of the material.

2 Inexact Newton method

Let us consider the solution of a nonlinear system in the following form

$$F(u) = b, \; u, b \in R^n, \tag{2.1}$$

where $F : R^n \to R^n$ is a nonlinear mapping. We shall assume that F has the Jacobian,

$$DF(u) = \left[\frac{\partial F_i}{\partial u_j}(u)\right] \quad i, j = 1, \ldots, n. \tag{2.2}$$

which represents the Gateaux derivative of F. For good performance of the following methods, we shall assume that the partial derivatives of F are continuous and DF represents the Frechet derivative of F.

The Newton method for solving (2.1) is based on linearization. Let u^i be an approximation to the solution. Then the nonlinear mapping F can be linearized in the vicinity of u^i to compute the correction v^i and a new better approximation $u^{i+1} = u^i + v^i$. More exactly, v^i is computed from the equation

$$F(u^i + v^i) \sim F(u^i) + DF(u^i)v^i = b.$$

This leads to the following basic algorithm.

Newton method (N)

Let u^0 be given.
for $i = 0, 1 \ldots$ **until** convergence **do**

$$\text{compute } v^i : \quad DF(u^i)v^i = b - F(u^i) \tag{2.3}$$

$$\text{let} : \quad u^{i+1} = u^i + v^i \tag{2.4}$$

end

Note 2.1 *It can be expected that the linearization gives good correction and that the Newton iterations converge to the solution of (2.1) if the initial guess will be sufficiently close to the solution. This assertion can be proved under suitable assumptions concerning the nonlinear mapping F.*

Note 2.2 *We can also prove that the convergence is self accelerating, superlinear or quadratic, see the following definition.*

Definition 2.1 *Let $u^i \to u^*$ and let $c > 0$, $\alpha > 0$ be constants such that*

$$\| u^{i+1} - u^* \| \leq c \| u^i - u^* \|^{\alpha}.$$

Then we say that the convergence is

- *linear if* $\alpha = 1$,

- *superlinear if* $\alpha > 1$,
- *quadratic if* $\alpha = 2$.

Note 2.3 *The condition of convergence (stopping criterion) takes usually the form*

$$\| b - F(u^i) \| \leq \varepsilon \| b \| , \tag{2.5}$$

where $\varepsilon = \varepsilon_R$ *is a relative error tolerance (*b *is assumed to be nonzero). For* $b = 0$ *, we have to use an absolute error tolerance* $\varepsilon = \varepsilon_A$ *and the condition*

$$\| F(u^i) \| \leq \varepsilon . \tag{2.6}$$

The conditions (2.5) and (2.6) can be combined, e.g. as

$$\| b - F(u^i) \| \leq \max\{\varepsilon_R \| b \| , \varepsilon_A\} . \tag{2.7}$$

Instead of residuals $b - F(u^i)$ *, it is also possible to check the differences* $u^{i+1} - u^i$*, which are related to the residuals through (2.3), (2.4).*

The correction v^i can be computed from (2.3) only approximately, e.g. by using some inner iterations. This leads to the so called inexact Newton method.

Inexact Newton method (IN)

Let u^0 be given
for $i = 0, 1, \ldots$ **until** convergence **do**

$$\text{compute } v^i : \quad DF(u^i)v^i \sim b - F(u^i) \tag{2.8}$$

$$\text{let} : \quad u^{i+1} = u^i + v^i \tag{2.9}$$

end

The accuracy of computation of v^i in (2.8), can be controlled by the following condition

$$\| DF(u^i)v^i - b + F(u^i) \| \leq \varepsilon_i \| b - F(u^i) \| . \tag{2.10}$$

For obtaining (linear) convergence, it is possible to take $\varepsilon_i = \varepsilon_0 < 1$. But for obtaining the superlinear convergence, it is necessary to choose ε_i decreasing to zero (so called forcing sequence). For example, we can take

$$\varepsilon_i = \bar{\varepsilon} \| b - F(u^i) \| .$$

3 Damped inexact Newton method

The (inexact) Newton method generally converges if the initial guess is in some neighbourhood of the exact solution of the nonlinear equation (local convergence). The size of this neighbourhood depends on the amount of nonlinearity.

For an initial guess, which is not sufficiently good, the convergence can be saved by a proper damping of the computed corrections. This leads to the following algorithm.

Damped Inexact Newton method (DIN).

Let u^0 be given,
for $i = 0, 1, \ldots$ **until** convergence **do**

$$\text{compute } v^i : \qquad DF(u^i)v^i \sim b - F(u^i) \tag{3.1}$$

$$\text{let:} \qquad u^{i+1} = u^i + \lambda_i v^i, \quad \lambda_i \in (0, 1\rangle \tag{3.2}$$

enddo

In this algorithm, the vectors v_i determined in (3.1) are not corrections but rather Newton directions, which serve for determination of the corrections by a line search. The use of damping with $\lambda_i < 1$ is usually needed only in the beginning of the iterative process, before we reach the region of convergence.

Various procedures can be used for determining the damping parameters. We can use

$$\lambda_i = \max\{\lambda \in (0, 1\rangle \; : \; \varphi(\lambda) \leq 0\}, \tag{3.3}$$

where

$$\varphi(\lambda) = \| b - F(u^i + \lambda v^i) \| - (1 - \alpha\lambda) \| b - F(u^i) \| \tag{3.4}$$

or

$$\varphi(\lambda) = \| F(u^i + \lambda v^i) - F(u^i) - \lambda DF(u^i)v^i \| - \eta\lambda \| DF(u^i)v^i \| . \tag{3.5}$$

The computing of λ_i from (3.3) and (3.4) is known as Armijo's rule. The second condition prevents to use the linearization if the departure from linear behaviour overcome η part of the value of the linearization itself. This gives more clear meaning to the parameter η in comparison with the parameter α from the Armijo's condition.

It is possible to show that the condition (3.5), which can be called the condition of the sufficient accuracy of the linearization, is stronger than (3.4) in the following sense.

Lemma 3.1 : *Let* $r^i = b - F(u^i)$ *and* $DF(u^i)v^i = \tilde{r}^i$, *where* $\| r^i - \tilde{r}^i \| \leq \varepsilon_0 \| r^i \|$ *and let* $\lambda \in (0,1\rangle$ *be such that*

$$\| F(u^i + \lambda v^i) - F(u^i) - \lambda DF(u^i) \| \leq \eta\lambda \ \| DF(u^i)v^i \| \ . \tag{3.6}$$

Then

$$\| b - F(u^i + \lambda v^i) \| \leq (1 - \alpha\lambda) \| b - F(u^i) \| \ . \tag{3.7}$$

for $\alpha = 1 - (1 + \varepsilon_0)\eta - \varepsilon_0$.

Proof. From (3.6), we get $\| -r^{i+1} + r^i - \lambda\tilde{r}^i \| \leq \ \eta\lambda \| \tilde{r}^i \|$

$$\begin{aligned} \| r^{i+1} \| &\leq \| r^i - \lambda\tilde{r}^i \| + \eta\lambda \| \tilde{r}^i \| \\ &\leq (1 - \lambda) \| r^i \| + \lambda\varepsilon_0 \| r^i \| + \eta\lambda(1 + \varepsilon_0) \| r^i \| \ . \end{aligned}$$

□

In the following, we shall use the condition (3.5) with $\eta = 1/2$, which can be shown to be optimal in the case of exact computation of the Newton direction v^i, see Appendix.

Practically, the damping parameters are determined only approximately, e.g. by means of the following procedures:

a) *backtracking*

```
for    λ = 1, 1/2, 1/4, ...
       if  φ(λ) ≤ 0    then
           λ_i = λ
           STOP
       end-if
end
```

b) *p-backtracking* with $p \geq 1$, which is the corrected backtracking for determination of λ_i with the precision to p valid binary digits. For $p = 1$ we get the simple backracking.

```
for    λ = 1, 1/2, 1/4, ...
       if  φ(λ) ≤ 0    then
           λ_min = λ
           λ_max = 2λ
       end-if
end
```

for $\quad k = 1, \dots, p-1$

$\quad \mu = \frac{1}{2}(\lambda_{\min} + \lambda_{\max})$

$\quad$ **if** $\varphi(\mu) \le 0$ **then** $\lambda_{\min} = \mu$

$\quad$ **else** $\lambda_{\max} = \mu$ **end-if**

end

$\lambda_i = \lambda min$

c) use polynomial interpolation of φ for determining λ_i. For example $\varphi(0), \varphi(\frac{1}{2})$ and $\varphi(1)$ can be used for constructing quadratic polynomial and computing λ_i, see e.g. [12].

In Section 6 we shall test the damped Newton method with the damping parameters computed adaptively with the aid of (3.3) and (3.5) and p-backtracking. Thus this algorithm depends on two parameters η and p, we shall write DN(η, p).

4 Adaptive continuation method

In this section, we shall describe another numerical method, which is based on linearization. Firstly, we shall imbed the nonlinear problem (2.1) into somewhat more complex problem

$$\text{find } u(t) : F(u(t)) = (1-t)b_0 + tb \quad \forall t \in \langle 0, 1 \rangle \tag{4.1}$$

where $u(0) = u^0$ is prescribed, $b^0 = F(u^0)$.

Now, we can solve the problem (4.1) approximately in successive continuation steps. In each step, we start from the initial state (t_i, u^i), where u^i is assumed to be close to $u(t_i)$, and then we shall continue to the next state (t_{i+1}, u^{i+1}), where

$$t_{i+1} = t_i + \tau, \ u^{i+1} = u^i + w\,. \tag{4.2}$$

If τ is given, then w can be computed by using the linearization,

$$F(u^i + w) \sim F(u^i) + DF(u^i)w = (1 - t_{i+1})b^0 + t_{i+1}b\,. \tag{4.3}$$

Assuming that $F(u^i) \sim (1-t_i)b^0 + t_i b$, the increment w can be computed from

$$DF(u^i)w = \tau(b - b^0), \ \tau \le 1 - t_i\,. \tag{4.4}$$

Moreover, the value of τ can be determined adaptively from the condition of sufficient accuracy of the linearization, which was already used in the previous section. This leads to the following algorithm.

Adaptive continuation (AC) algorithm

Given u^0 ,
$b^0 = F(u^0)$,
$t_0 = 0$
for $i = 0, 1 \ldots$ **until** $t_i = 1$ **do**

$$\text{compute } z: \quad DF(u^i)z = (1-t_i)(b-b^0) \tag{4.5}$$

$$\text{find}: \quad \kappa_i = \max\{\kappa \in (0,1\rangle : \varphi(\kappa) \leq 0\} \tag{4.6}$$

$$\text{let}: \quad \tau = \kappa_i(1-t_i), \quad t_{i+1} = t_i + \tau$$

$$w = \kappa_i z, \quad u^{i+1} = u^i + w$$

end

where

$$\varphi(\kappa) = \| F(u^i + \kappa z) - F(u^i) - \kappa DF(u^i)z \| - \eta\kappa \| DF(u^i)z \| . \tag{4.7}$$

The continuation parameter κ_i can be determined approximately by the same procedures as described in the previous section. It is also possible to use inexact version of the above algorithm, with z computed from (4.5) only approximately.

In derivation of (4.4), we assumed that $F(u^i)$ is close to $(1-t_i)b^0 + t_i b$. This can not be true, especially after many continuation steps or when using inexact version of the AC method. For this reason, we can equipped the AC algorithm with the DIN correction, see the following algorithm.

Corrected adaptive continuation (CAC) algorithm

1. Given u^0,
   $b^0 = F(u^0)$,
   $t_0 = 0$,
2. **for** i=0,1, ... **until** $t_i = 1$ **do**

   (a) perform DIN iterations starting from u^i to get $\bar{u}^i$ such that

   $$\| F(\bar{u}^i) - (1-t_i)b^0 - t_i b \| \leq \varepsilon_i \| (1-t_i)b^0 + t_i b \| \tag{4.8}$$

   (b) compute z : $DF(\tilde{u}^i)z = (1-t_i)(b-b^0)$
   with the last Jacobian evaluated in the previous step

   (c) find $\kappa_i \sim \max\{\kappa \in (0,1\rangle : \varphi(\kappa) \leq 0\}$

   (d) let $t_{i+1} = t_i + \kappa_i(1-t_i)$, $u^{i+1} = \bar{u}^i + \kappa_i z$

   end

3. perform DIN iterations starting from u^i to get the final approximation with the required tolerance ε.

end

In the following special variant of CAC, we perform just one Newton iteration in step (i), i.e.

(a) compute $v^i \,:\, DF(u^i)v^i \,=\, (1-t_i)b^0 + t_i b - F(u^i)$
$\bar{u}^i \,=\, u^i + v^i$

(b) compute $z \,:\, DF(u^i)z \,=\, (1-t_i)(b-b^0)$

(c) find κ_i

(d) let $t_{i+1} \,=\, t_i + \kappa_i\,(1-t_i), \quad u^{i+1} \,=\, \bar{u}^i + \kappa_i\, z\,.$

Such correction can be very cheap, as in both steps (i) and (ii) we solve linear systems with the same matrix and additionally these systems in (a) and (b) can be solved in parallel. Similar procedure (without adaptation) was introduced in [2]

The following variant of the CAC method is used in Section 6.

CAC^* **algorithm** with parameters η, p

1. Given u^0
   $b^0 \,=\, F(u^0)$
   $t_0 \,=\, 0$

2. **for** i=0, 1, ... **until** $t_i = 1$ **do**

   (a) compute $v^i \,:\, DF(u^i)v^i \,=\, (1-t_i)b_0 + t_i b - F(u^i)$
   (b) determine damping parameter λ_i by using (3.3) and (3.5) with η and the p-backtracking ,
   (c) $\bar{u}^i \,=\, u^i + \lambda_i\, v^i$
   (d) if $\lambda_i \,<\, 1$ then $u^i = \bar{u}^i$ and **goto** (a)
   (e) compute $z \,:\, DF(u^i)z \,=\, (1-t_i)(b-b^0)$
   (f) find $\kappa_i \,\sim\,$ max $\{\kappa \in \langle 0, 1\rangle \,:\, \varphi(\kappa) \,\leq\, 0\}$ by using φ from (4.7) with η and p-backtracking,
   (g) let $t_{i+1} \,=\, t_i + \kappa_i(1-t_i), \quad u^{i+1} \,=\, \bar{u}^i + \kappa_i z\,.$

   enddo

3. perform DN iterations starting from the last iteration u^i to get the final approximation with the required tolerance ε.

end

5 Physically nonlinear elasticity

Let us consider deformation of a body $\Omega \subset R^3$ under a given loading and assume that the problem can be described by the displacement $u : \Omega \to R^3$, the small strain tensor $\varepsilon : \Omega \to S$ and the Cauchy stress tensor $\sigma : \Omega \to S$, see e.g. [15] for more details. Futher, we shall assume a nonlinear stress–strain relation, which can be written as

$$\sigma(x) = T(x, \varepsilon(x)) \qquad \text{for all } x \in \Omega\,. \tag{5.1}$$

The relation (5.1) can be abbreviated by omitting the dependence on x as

$$\sigma = T(\varepsilon)\,. \tag{5.2}$$

Moreover, we shall consider a special case of the relation (5.2), namely

$$T(\varepsilon) = D(\varepsilon)\varepsilon, \;\; D(\varepsilon) = 3K(\hat{\varepsilon})\hat{E} + 2G(\bar{\varepsilon})\bar{E} \tag{5.3}$$

where $\hat{E}$ and $\bar{E}$ are operators, which give the volumetric and deviatoric parts of the tensors from S, respectively. This means that

$$\hat{E}\nu = (\hat{\nu}_{ij}), \;\; \hat{\nu}_{ij} = \tfrac{1}{3}(\nu_{11} + \nu_{22} + \nu_{33})\delta_{ij}\,, \tag{5.4}$$

$$\bar{E}\nu = \nu - \hat{E}\nu \tag{5.5}$$

for any $\nu \in S$. The values $\hat{\varepsilon}$ and $\bar{\varepsilon}$ in (5.3) characterize the magnitude of the volumetric and deviatoric part of ε. We take

$$\hat{\varepsilon} = \varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33} \quad \text{and} \quad \bar{\varepsilon} = \| \bar{E}\varepsilon \|^2\,. \tag{5.6}$$

The norm in (5.6) is defined as

$$\| \nu \| = [\nu, \nu]^{1/2}, \;\; [\nu, \xi] = \sum_{ij} \nu_{ij}\,\xi_{ij} \quad \forall \nu, \xi \in S\,. \tag{5.7}$$

The function K describes the value of the *bulk modulus* and G gives the value of the *shear modulus.*

Lemma 5.1 *Let K continuous in R^1 and differentiable in $R^1\backslash\{0\}$, G be continuous in $\langle 0,\infty)$ and differentiable in $(0,\infty)$. Then T has the Gateaux derivative T' which can be written as*

$$
\begin{aligned}
T'(\varepsilon)\nu &= D(\varepsilon)\nu + D_1(\varepsilon,\nu) + D_2(\varepsilon,\nu)\,, && (5.8)\\
D_1(\varepsilon,\nu) &= \begin{cases} 0 & \textit{if } \hat{E}\varepsilon = 0\,,\\ 3K'(\hat{\varepsilon})\hat{\nu}\,\hat{E}\varepsilon & \textit{if } \hat{E}\varepsilon \neq 0\,,\end{cases} && (5.9)\\
D_2(\varepsilon,\nu) &= \begin{cases} 0 & \textit{if } \bar{E}\varepsilon = 0,\\ 2G'(\bar{\varepsilon})\,\|\,\bar{E}\varepsilon\,\|^{-1}\,[\bar{E}\varepsilon,\bar{E}\nu]\,\bar{E}\varepsilon & \textit{if } \bar{E}\varepsilon \neq 0\,.\end{cases} && (5.10)
\end{aligned}
$$

Proof. Let $\varepsilon,\nu\in S$. Then the derivative in the direction ν is

$$
\begin{aligned}
dT(\varepsilon,\nu) &= \lim_{t\to 0} t^{-1}[T(\varepsilon+t\nu) - T(\varepsilon)]\\
&= \lim_{t\to 0} t^{-1}[D(\varepsilon+t\nu)(\varepsilon+t\nu) - D(\varepsilon)\varepsilon]\\
&= \lim_{t\to 0} D(\varepsilon+t\nu)\nu + \lim_{t\to 0} t^{-1}[D(\varepsilon+t\nu) - D(\varepsilon)]\varepsilon\\
&= D(\varepsilon)\nu + \lim_{t\to 0} 3t^{-1}\left[K(\hat{\varepsilon}+t\hat{\nu}) - K(\hat{\varepsilon})\right]\bar{E}\varepsilon \,+\, \lim_{t\to 0} 2t^{-1}\left[G(\overline{\varepsilon+t\nu}) - G(\bar{\varepsilon})\right]\bar{E}\varepsilon
\end{aligned}
$$

The first term $D(\varepsilon)\,\nu$ follows from continuity of $t\to D(\varepsilon+t\nu)$. The expressions (5.9) and (5.10) clearly hold in the cases of $\hat{E}\varepsilon = 0$ and $\bar{E}\varepsilon = 0$, respectively. Otherwise, they follow from the rule of derivation of composite functions. As $\eta\mapsto dT(\varepsilon,\eta)$ is a linear mapping, T has the Gateaux derivative given by (5.8). □

Note 5.1 *If K, G are continuously differentiable, then T' is the Frechet's derivative.*

The *boundary value problem* on Ω, with boundary conditions

$$
\begin{aligned}
u &= 0 \quad \text{on} \quad \Gamma_0 && (5.11)\\
\sum_j \sigma_{ij}\,n_j &= g_i \quad \text{on} \quad \Gamma_1, \quad i = 1,2,3 && (5.12)
\end{aligned}
$$

and the volume loading with the density f, can be formulated as

$$
\text{find } u\in V:\ a(u,v) = b(v) \quad \forall v\in V \tag{5.13}
$$

where

$$
a(u,v) = \int_\Omega [T(\varepsilon(u)),\,\varepsilon(v)]\,d\Omega\,, \quad b(v) = \int_\Omega \langle f,v\rangle\,d\Omega + \int_{\Gamma_1}\langle g,v\rangle\,d\Gamma \tag{5.14}
$$

$$V \;=\; \{v \in [H^1(\Omega)]^3 \,:\, v = 0 \quad \text{on } \Gamma_0\,\}\,. \tag{5.15}$$

For the finite element discretization, we shall use a finite element space $V_h \subset V$ with the basis $\{\phi_i\}_{i=1}^n$. The solution of the discrete problem can be then written in the form

$$u_h \;=\; \sum_1^n u_i\,\phi_i \tag{5.16}$$

where $\underline{u} = (u_i)$ is the solution of the algebraic problem

$$\text{find } \underline{u} \in R^n \,:\, F(\underline{u}) = \underline{b}\,. \tag{5.17}$$

Above $\underline{b} \;=\; (b(\phi_i))$ and for $\underline{v} \;=\; (v_i) \in R^n$, $v_h = \sum v_i\phi_i$,

$$F(\underline{v}) \;=\; A(\underline{v})\underline{v}, \quad A(\underline{v}) \;=\; [A_{ij}(\underline{v})], \tag{5.18}$$

$$A_{ij}(\underline{v}) \;=\; \int_\Omega [D(\varepsilon(v_h))\,\varepsilon(\phi_i),\, \varepsilon(\phi_j)]\, d\Omega. \tag{5.19}$$

The value $F(\underline{v})$ is usually computed in an element by element way. Moreover, the computation on the element level is equivalent to computation of stress by using (5.2) and determination of nodal forces which are equivalent to the computed stress.

Under the assumption of Lemma 5.1, F has the Jacobian $DF(\underline{v}) = [J_{ij}(\underline{v})]$,

$$J_{ij}(\underline{v}) \;=\; \int_\Omega \left[T'(\varepsilon(v_h))\,\varepsilon(\phi_i),\, \varepsilon(\phi_j)\right]\, d\Omega\,. \tag{5.20}$$

Plane strain problems

The plane strain problems is characterized by the deformation ε such that

$$\varepsilon_{i3} \;=\; \varepsilon_{3i} \;=\; 0 \qquad \text{for} \quad i = 1,2,3.$$

The plane strain can be represented by the reduced tensor $\varepsilon \in S_2$,

$$\varepsilon \;=\; \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} \\ \varepsilon_{12} & \varepsilon_{22} \end{bmatrix}.$$

The previous definitions of the volumetric and deviatoric part of ε give tensors, which, however, do not belong to S_2. For computation of $\bar{\varepsilon}$, we shall use the norm of such 3×3 strain deviator, but in (5.8), (5.9), (5.10), (5.19), (5.20) etc. we can use the proper reduced 2×2 tensors.

6 Numerical experiments

For testing the nonlinear system solvers, we shall use the finite element discretization of the nonlinear elasticity problem shown in Figure 1. This plane strain problem represents a concrete wall in a nonlinear soil. The boundary conditions consist from zero normal displacements and zero tangential stresses with an exception of the top side, where a load P acts on the concrete wall. The problem is discretized by the linear triangular finite elements.

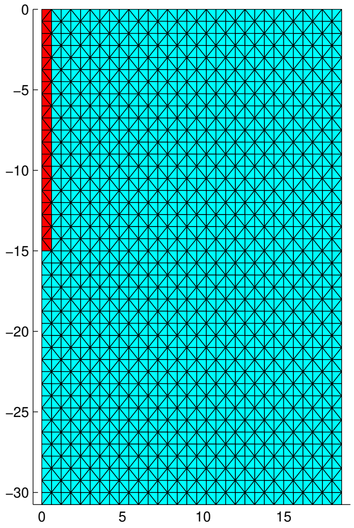

Fig. 1: The test problem.

The concrete is characterized by the Young modulus $E = 31.5$ GPa, the Poisson ratio $\nu = 0.2$ and the density 2 500 kgm^{-3}. The soil has density 1 800 kgm^{-3} and its nonlinear behaviour is described by the functions

$$\begin{aligned} K(\hat{\varepsilon}) &= \frac{K_0}{1-\hat{\varepsilon}/m}, \quad K_0 = 41.417 \text{ MPa}, \\ G(\bar{\varepsilon}) &= \frac{A}{B+\bar{\varepsilon}}, \quad A = 65.8 \text{ KPa}, B = 0.0094. \end{aligned}$$

In the case of $m = \infty$ the material description corresponds to the behaviour of clay presented in [11]. The parameter m is considered to be equal to 0.1, 0.01 and 0.001. The decrease of m is used for increasing the nonlinearity in the model problem.

The function K will be defined and differentiable on the interval $(-\infty, m)$, the function G will be defined on $\langle 0, \infty)$ and differentiable on $(0, \infty)$. The condition $\hat{\varepsilon} < m$ is checked in all iterations. If a new iteration $u^i + v$ violates this condition, then v is damped to get an admissible result. Such action is marked by the exclamation mark in Tables 1 and 3.

The nonlinear systems arising from the discretization of the described test problem are solved by two methods:

1. Damped Newton DN(η, p) method with the parameter η from (3.5) and computation of the damping parameters by p-backtracking. As default values, we shall use $\eta = 1/2$ and $p = 3$.

2. Corrected adaptive continuation method CAC*(η, p) with the parameter η from (3.5) and (4.7) and computation of the damping parameters by p-backtracking. As default values, we shall again use $\eta = 1/2$ and $p = 3$.

The convergence of the above methods for the load $P = 1.5$ MPa can be seen in Figure 2. More information about the values of damping and continuation parameters and computational work can be seen in Tables 1 and 2.

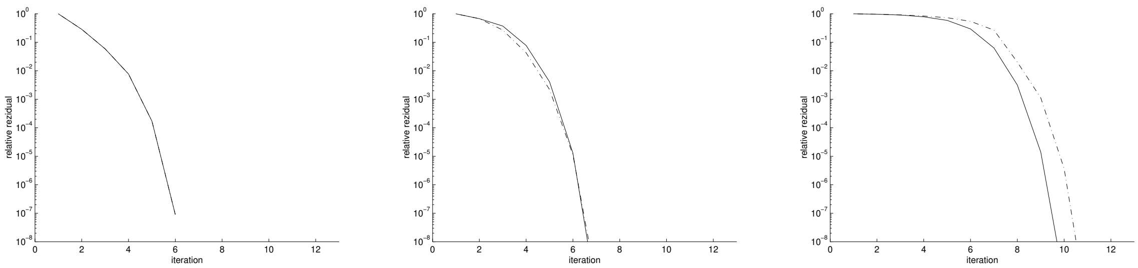

Fig. 2: The convergence of the DN (dot-and-dashed line) and $ACA*$ (solid line) methods in the case of solving the test problem with P=1.5 MPa and soil parameters m=0.1(left), m=0.01 (middle) and m=0.001 (right).

	m=0.1			m=0.01			m=0.001		
	DN	CAC*		DN	CAC*		DN	CAC*	
	λ	λ	t	λ	λ	t	λ	λ	t
1	1.000	1.000	1.000	!0.313	1.000	!0.313	!0.031	1.000	!0.031
2	1.000	1.000	1.000	0.875	1.000	1.000	!0.047	1.000	!0.092
3	1.000	1.000	1.000	1.000	1.000	1.000	!0.078	1.000	!0.205
4	1.000	1.000	1.000	1.000	1.000	1.000	!0.125	1.000	!0.404
5	1.000	1.000	1.000	1.000	1.000	1.000	!0.250	1.000	0.702
6				1.000	1.000	1.000	0.625	1.000	1.000
7							1.000	1.000	1.000
8							1.000	1.000	1.000
9							1.000	1.000	1.000
10							1.000		

Tab. 1: The history of the convergence of $DN(1/2; 3]$ and $CAC^*(1/2; 3)$ for the test problem with P=1.5 MPa.

	m=0.1				m=0.01				m=0.001			
	NI	NFN	NDF	NSS	NI	NFN	NDF	NSS	NI	NFN	NDF	NSS
DN	5	11	5	5	6	20	6	6	10	53	10	10
CAC*	5	11	5	5	6	18	6	7	9	49	9	14

Tab. 2: The computational requirements for the test problem with P=1.5 MPa. Notation: NI = number of iterations, NFN = number of evaluations of the mapping F, NDF = number of evaluations of the Jacobian DF, NSS = number of solutions of the systems with the Jacobian matrix

The same tests were performed for the test problem with the load $P = 3$ MPa, the results can be seen from Fig. 3 and Tables 3 and 4.

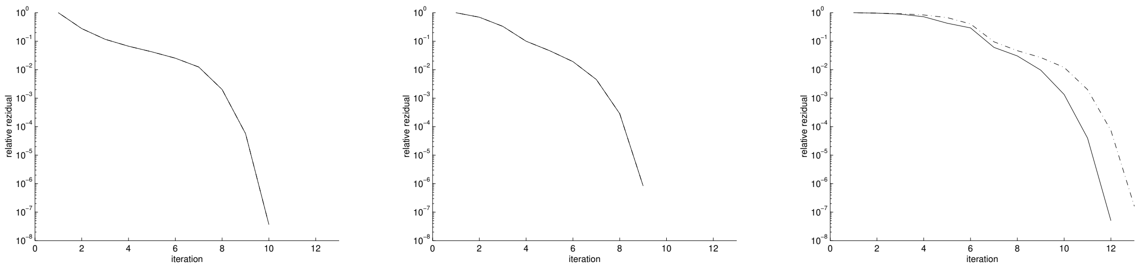

Fig. 3: The convergence of the DN (dot-and-dashed line) and $ACA*$ (solid line) methods in the case of solving the test problem with P=3 MPa and soil parameters m=0.1(left), m=0.01 (middle) and m=0.001 (right).

	m=0.1			m=0.01			m=0.001		
	DN	CAC*		DN	CAC*		DN	CAC*	
	λ	λ	t	λ	λ	t	λ	λ	t
1	1.000	1.000	1.000	!0.313	1.000	!0.313	!0.031	1.000	!0.031
2	1.000	1.000	1.000	1.000	1.000	1.000	!0.047	1.000	!0.107
3	0.750	0.750	1.000	1.000	1.000	1.000	!0.094	1.000	!0.274
4	0.500	0.500	1.000	1.000	1.000	1.000	!0.188	1.000	!0.592
5	0.625	0.625	1.000	1.000	1.000	1.000	!0.438	1.000	1.000
6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
8	1.000	1.000	1.000	1.000	1.000	1.000	0.750	1.000	1.000
9	1.000	1.000	1.000				1.000	1.000	1.000
10							1.000	1.000	1.000
11							1.000	1.000	1.000
12							1.000	1.000	1.000

Tab. 3: The history of the convergence of $DN(1/2;3]$ and $CAC^{*}(1/2;3)$ for the test problem with P=3 MPa.

	m=0.1				m=0.01				m=0.001			
	NI	NFN	NDF	NSS	NI	NFN	NDF	NSS	NI	NFN	NDF	NSS
DN	9	28	9	9	8	21	8	8	12	57	12	12
CAC*	9	28	9	9	8	22	8	9	11	49	11	15

Tab. 4: The computational requirements for the test problem with P=3 MPa.

Now we shall test the sensitivity of the DN and CAC^* method to the choice of the parameters η and p. For this purpose, we shall solve the test problem with the load $P = 3$ MPa. The results, which are summarized in Tables 5 - 7, should be compared with the results from Table 4. In Tables 4 - 7, we can relative insensivity of the considered methods to the solver parameters. An exception is the case, when η is too small.

	m=0.1				m=0.01				m=0.001			
	NI	NFN	NDF	NSS	NI	NFN	NDF	NSS	NI	NFN	NDF	NSS
DN	10	25	10	10	10	26	10	10	15	55	15	15
CAC*	10	25	10	10	10	28	10	12	12	48	12	18

Tab. 5: The computational requirements of solving the test problem with P=3 MPa by DN and CAC^* methods with $\eta = 1/2$ and $p = 1$.

	m=0.1				m=0.01				m=0.001			
	NI	NFN	NDF	NSS	NI	NFN	NDF	NSS	NI	NFN	NDF	NSS
DN	13	59	13	13	11	47	11	11	15	85	15	15
CAC*	13	59	13	13	11	49	11	13	14	82	14	20

Tab. 6: The computational requirements of solving the test problem with P=3 MPa by DN and CAC^* methods with $\eta = 1/3$ and $p = 3$.

	m=0.1				m=0.01				m=0.001			
	NI	NFN	NDF	NSS	NI	NFN	NDF	NSS	NI	NFN	NDF	NSS
DN	8	20	8	8	8	21	8	8	12	51	12	12
CAC*	8	20	8	8	8	22	8	9	11	49	11	15

Tab. 7: The computational requirements of solving the test problem with P=3 MPa by DN and CAC^* methods with $\eta = 2/3$ and $p = 3$.

The last test shows the efficiency of the AC algorithm, which does not include the Newton correction step, see Section 4. The adaptive determination of the size of continuation steps uses the condition (4.7) and computation of the damping parameters by p-backtracking. The test was performed with the values $\eta = 1/2$ and $p = 3$.

	m=0.1				m=0.01				m=0.001			
	NI	NFN	NDF	NSS	NI	NFN	NDF	NSS	NI	NFN	NDF	NSS
CAC	9	28	9	9	9	28	9	9	13	65	13	13

Tab. 8: The computational requirements of solving the test problem with P=3 MPa by AC method with $\eta = 1/2$ and $p = 3$.

7 Conclusions

This paper describes damped Newton and adaptive continuation solvers and show their efficiency and robustness when solving nonlinear elasticity problems arising from the soil mechanics. The robustness is important e.g. for computation the load - settlement dependence, see the curve on Fig. 4. In all of the performed experiments, we observed superlinear convergence.

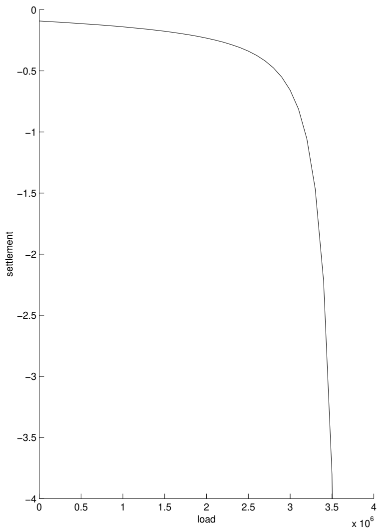

Fig. 4: The load - settlement curve.

References

[1] O. Axelsson: On Newton type solvers for nonlinear boundary value problems, in Geomechanics'96, A.A. Balkema, Rotterdam 1997

[2] O. Axelsson, Globally convergent continuation methods for non-linear equations, Proc. Conf. Large Scale Scientific Computations of Engineering and Environmental Problems, H. Griebel, O.P. Iljev, S.D. Margenov and P.S. Vassilevski eds., Vieweg Verlag 1998, 3-17.

[3] O. Axelsson, On Newton type continuation methods. Communications in Applied Analysis 4(2000), 575-595

[4] O. Axelsson, A. Padiy, On a two-level Newton-type procedure applied for solving nonlinear elasticity problems, Int. Journal for Numer. Meth. Engrg. 49(2000), 1479-1493

[5] R. Blaheta: Elasticity with nonlinear behaviour of material, manuscript, 1996

[6] R. Blaheta: Multi-level Newton methods for nonlinear problems with applications to elasticity, HIPERGEOS, technical report, Ostrava 1997 Available at http://www-math.sci.kun.nl/math/Copernicus

[7] P. Byczanski: Solution of nonlinear systems of equations: "postup01", Proceedings of the conference Contemporary mathematical methods in engineering 2000, TU Ostrava 2000 (in Czech)

[8] P. Byczanski: Solution of nonlinear systems of equations: "postup08", Proceedings of the conference MATLAB 2000, Prague 2000 (in Czech)

[9] J.E. Dennis, R.B. Schnabel: Numerical Methods for Unconstrained Optimization and Nonlinear Equations, SIAM Philadelphia 1996

[10] C.S. Desai, H.J. Siriwardane: Constitutive laws for engineering materials with emphasis on geologic materials, Prentice Hall, Englewood Cliffs, NJ 1984

[11] Asen Georgiev, Angel Baltov and Svetozar Margenov: HIPERGEOS Benchmark Problems Related to Bridge Engineering Applications, COPERNICUS CP 94-0820 Project Report, MOST-XXI, Sofia 1996

[12] C.T. Kelley: Iterative Methods for Linear and Nonlinear Equations, SIAM Philadelphia 1995

[13] D.J. Naylor, G.N. Pande: Finite Elements in Geotechnical Engineering, Pineridge Press, Swansea 1981

[14] G.N. Pande, G. Beer, J.R. Williams: Numerical Methods in Rock Mechanics, J. Willey, Chichester 1990

[15] J. Nečas, I. Hlaváček: Mathematical Theory of Elastic and Elasto-Plastic Bodies, Elsevier, Amsterdam 1981

[16] W.C. Rheinboldt: Methods for Solving Systems of Nonlinear Equations, Second Edition, SIAM Philadelphia 1998

8 Appendix: Convergence of the DN(η, p) method

In this section, we investigate the convergence of the damped Newton method with exact computation of the Newton corrections and computation of the damping parameters with the use of (3.3) and (3.5) with the parameter η and p-backtracking.

Theorem 8.1 *Let us assume that $DF(u)$ is defined and nonsingular for all $u \in R^n$ and let m, K be two constants such that*

$$\begin{aligned} \| DF(u)^{-1} \| &\leq m && \forall u, v \in R^n, && (8.1)\\ \| DF(u) - DF(v) \| &\leq K \| u - v \| && \forall u, v \in R^n . && (8.2) \end{aligned}$$

Let u^0 be an initial guess, $r^0 = b - F(u^0) \neq 0$ be the initial residual. Then the damped Newton method $DN(\eta, p)$, where $\eta \in (0,1)$, $p \geq 1$, produces the sequence of damping parameters λ_i, iterates u^i and residuals r^i for which

$$\lambda_i \geq \bar{\lambda} = \min\left\{1\,;\frac{\eta}{Km^2\|r^0\|}\right\} \tag{8.3}$$

$$\| r^{i+1} \| \leq [1 - (1-\eta)\bar{\lambda}]^i \| r^0 \| \; . \tag{8.4}$$

Proof. Let $u^i \in R^n$, $\Delta = DF(u^i)^{-1} r^i$, $\lambda \in R^1$. Then

$$\begin{aligned} F(u^i + \lambda\Delta) - F(u^i) - \lambda DF(u^i)\Delta &= \textstyle\int_0^1 [DF(u^i + t\lambda\Delta) - DF(u^i)]\, \lambda\Delta dt \\ \| F(u^i + \lambda\Delta) - F(u^i) - \lambda DF(u^i)\Delta \| &\leq \tfrac{1}{2} K\lambda^2 \| \Delta \|^2 \\ &\leq \tfrac{1}{2} K\lambda^2 m^2 \| DF(u^i)\Delta \|^2 = \tfrac{1}{2} K\lambda^2 m^2 \| r^i \|^2 \end{aligned} \tag{8.5}$$

Now assume that $u^{i+1} = u^i + \lambda\Delta$, where $\lambda \in (0, 1\rangle$ fulfils the condition

$$\| F(u^{i+1}) - F(u^i) - \lambda DF(u^i)\Delta \| \leq \eta\lambda \| DF(u^i)\Delta \| \; .$$

Then

$$\begin{aligned} \| -r^{i+1} + r^i - \lambda r^i \| &\leq \eta\lambda \| r^i \| , \\ \| r^{i+1} \| - (1-\lambda) \| r^i \| &\leq \eta\lambda \| r^i \| , \\ \| r^{i+1} \| &\leq [1 - (1-\eta)\lambda] \| r^i \| \; . \end{aligned} \tag{8.6}$$

Now, we are ready to prove (8.3) and (8.4) by induction.

Let $i = 0$, then

$$\| F(u^0 + \lambda\Delta) - F(u^0) - \lambda DF(u^0)\Delta \| \leq \frac{1}{2} K\lambda^2 m^2 \| r^0 \|^2 \leq \eta\lambda \| r^0 \|$$

for $\lambda \leq \frac{1\eta}{Km^2\|r^0\|}$. Thus by $p-$backtracking, we have to found $\lambda_0 \geq \bar{\lambda}$. Then (8.4) immediately follows from (8.6) .

For $i > 0$, the induction assumptions imply $\| r^i \| \leq \| r^0 \|$. Then

$$\begin{aligned}\| F(u^i+\lambda\Delta)-F(u^i)-\lambda DF(u^i)\Delta \| &\leq \tfrac{1}{2}Km^2\lambda^2 \| DF(u^i)\Delta \|^2 \\ &\leq \tfrac{1}{2}Km^2\lambda^2 \| DF(u^i)\Delta \| \| r^0 \| \leq \eta\lambda \| DF(u^i) \|\end{aligned}$$

for $\lambda < \frac{2\eta}{Km^2\|r^0\|}$. Thus by $p-$backtracking, we have to found $\lambda_i \geq \bar{\lambda}$ and (8.4) follows from (8.6).

□

Conclusion 8.1 *Let the assumptions (8.1) , (8.2) be valid. Then for any $\eta \in (0,1)$, $p \geq 1$ the $DN(\eta,p)$ method converges. Moreover, the estimates (8.3), (8.4) imply that the optimal value of η be 1/2 in the case that some damping is necessary.*

Proof. We shall show that the produced sequence of iterations u^i fulfils the Bolzano-Cauchy condition and therefore that it is convergent, $u^i \to u^*$. As F is a continuous mapping and $F(u^i) \to b$, it holds that $F(u^*) = b$.

We have,

$$u^{i+1} - u^i \;=\; \lambda_i DF(u^i)^{-1} r^i .$$

thus

$$\| u^{i+1} - u^i \| \leq \; m \| r^i \| \leq \; mq^i \| r^0 \|$$

where $q = 1-(1-\eta)\bar{\lambda} \in \langle 0,1)$. As a consequence,

$$\begin{aligned}\| u^{i+k} - u^i \| &\leq \sum_{l=1}^{k} \| u^{i+l} - u^{i+l-1} \| \leq \\ &\leq m \| r^0 \| \sum_{l=1}^{k} q^{i+l-1} \;=\; q^i \frac{1-q^k}{1-q} m \; \| r^0 \|\end{aligned}$$

from which we get the Bolzano-Cauchy condition for the sequence u^i.

Now, we shall prove the second assertion. The damping is necessary if $\frac{2\eta}{Km^2\|r^0\|} \; < \; 1$. Then (8.3) and (8.4) give

$$\| r^{i+1} \| \leq q(\eta) \| r^i \|, \quad \text{where } q(\eta) = 1 - \frac{(1-\eta)\eta}{Km^2 \| r^0 \|} .$$

The value $q(\eta)$ is minimal for $\eta \; = \; 1/2$.

□

Note 8.1 *For the case of F arising from the FE solution of nonlinear elasticity problems, the assumptions of the theorem 8.1 can be verified under reasonable assumptions. For a special cases, see [15] and [4].*