Applications of Mathematics

Radim Blaheta; Roman Kohut

Efficient inexact Newton-like methods with application to problems of the deformation theory of plasticity

Applications of Mathematics, Vol. 38 (1993), No. 6, 411–427

Persistent URL: http://dml.cz/dmlcz/104564

EFFICIENT INEXACT NEWTON-LIKE METHODS WITH APPLICATION TO PROBLEMS OF THE DEFORMATION THEORY OF PLASTICITY

Radim Blaheta and Roman Kohut, Ostrava

Summary. Newton-like methods are considered with inexact correction computed by some inner iterative method. Composite iterative methods of this type are applied to the solution of nonlinear systems arising from the solution of nonlinear elliptic boundary value problems. Two main questions are studied in this paper: the convergence of the inexact Newton-like methods and the efficient control of accuracy in computation of the inexact correction. Numerical experiments show the efficiency of the suggested composite iterative techniques when problems of the deformation theory of plasticity are solved.

Keywords: nonlinear systems, inexact Newton-like methods, composite iterations, deformation theory of plasticity

AMS classification: 65H10, 65N22, 73E99

1. Introduction

We shall be interested in efficient iterative methods for the solution of large nonlinear systems

$$A(u)u = b \tag{1.1}$$

arising from the discretization of nonlinear (quasilinear) elliptic problems. Here, $u, b \in \mathbb{R}^n$ and $A(u)$ is an $n \times n$ matrix with entries which depend on u.

For the solution of the system (1.1), we shall exploit Newton-like methods which iterations $i \longrightarrow i+1$ consist of the following steps:

$$\text{compute } \Delta^i. \quad B_i\Delta^i = r^i = b - A(u^i)u^i, \tag{1.2}$$

$$\text{put:} \quad u^{i+1} = u^i + \omega_i\Delta^i \tag{1.3}$$

where Δ^i is a correction, $0 < \omega_i \leqslant 1$ is a damping parameter and B_i is a suitable $n \times n$ matrix. We shall consider two particular choices

$$B_i = A_0, \tag{1.4}$$

$$B_i = A(u^i) \tag{1.5}$$

which give respectively the *generalized Picard* (GP) and the *secant modulus* (SM) method.

Note that we do not consider the actual Newton method for which B_i is the Jacobian, but we restrict our attention to the GP and SM methods which have the following attractive properties

- easier implementation within software for the solution of corresponding linear problems,
- sufficient efficiency for the solution of many problems with mild nonlinearity,
- possibility of application to problems with non-differentiable operators, cf. the application to the problems of the flow theory of plasticity [4].

Further, we shall study the inexact Newton-like methods for which the correction Δ^i is computed only approximately by some suitable iterative method. This approach has the following advantages:

- inexact computation of the correction can save a great deal of the computational work,
- during the iterative process zero becomes better and better initial guess for the computed correction,
- the use of iterative methods for solving linear systems itself enhances the efficiency when large linear systems are solved.

In this paper, we shall answer two main questions:

- the convergence of inexact Newton-like methods,
- the effective control of accuracy for computation of the inexact correction.

The question of convergence of the inexact Newton-like methods will be studied simultaneously for nonlinear elliptic problems and their discretization by the finite element method. We shall consider the case where the nonlinear elliptic problem can be formulated as a minimizing problem for some convex functional. Our results will be an extension of the convergence results for the Newton-like methods with the exact correction which can be found e.g. in [5] and [6].

Concerning the second question, an efficient strategy for controlling accuracy of computation of the inexact correction is described. This strategy was motivated by our numerical experiments and is partly explained by the presented theory.

Note, that our techniques are based on the linearly convergent Newton-like methods. This makes some differences from the well-known results, see e.g. [7, 8], which mostly concern the quadratically convergent Newton method.

Finally, we describe numerical experiments showing the application of the inexact Newton-like methods to the solution of problems of the deformation theory of plasticity (physically nonlinear elasticity). These numerical experiments motivate our strategy for control of accuracy for computation of the inexact correction and show the efficiency of the described methods. Note that similar techniques can be also applied for solving problems of the flow theory of plasticity, see [4].

2. Abstract nonlinear problems

Let us consider a nonlinear boundary value problems in the following weak form

$$\text{find } u \in V: \quad a(u,u,v) = b(v) \qquad \forall v \in V. \tag{2.1}$$

Above V is a Hilbert space equipped with the inner product $(\cdot,\cdot)_V$ and the norm $\|\cdot\|_V$, $a\colon V^3 \longrightarrow \mathbb{R}^1$, $a(u,\cdot,\cdot)$ is a bilinear form for any fixed $u \in V$ and $b\colon V \longrightarrow \mathbb{R}^1$ is a bounded linear functional.

Together with the boundary value problem (2.1), we shall consider its finite element approximation, i.e., the problem

$$\text{find } u_h \in V_h: \quad a(u_h,u_h,v_h) = b(v_h) \qquad \forall v_h \in V_h \tag{2.2}$$

where $V_h \subset V$ is a finite element space. We shall also consider the algebraic problem

$$\text{find } u \in \mathbb{R}^n: \quad A(u)u = b \tag{2.3}$$

which is equivalent to (2.2). It means that $A(u)$ is an $n \times n$ matrix, $b \in \mathbb{R}^n$ and

$$\langle A(u)v, w\rangle = a(u_h, v_h, w_h) \tag{2.4}$$

$$\langle b, v\rangle = b(v_h) \tag{2.5}$$

for all $u_h, v_h, w_h \in V_h$ and $u, v, w \in \mathbb{R}^n$ such that the components of u, v and w are simply the coefficients of the representation of u_h, v_h and w_h, respectively, in the given basis of V_h. $\langle\cdot,\cdot\rangle$ is the inner product in $\mathbb{R}^n$, $\langle u, v\rangle = u^T v$.

We can find a lot of examples for the abstract boundary value problem (2.1). We can mention nonlinear heat transfer, diffusion, potential flow or magnetostatic problems. But in this paper we shall be interested in only one example which will be the problem of the deformation theory of plasticity (physically nonlinear elasticity), described in detail in e.g. [5, 6]. This problem will be also briefly described in the following Section.

Considering the above mentioned examples of the abstract boundary value problem (2.1), we can suppose that

$$a(u,\cdot,\cdot)\colon V^2 \longrightarrow \mathbb{R}^1 \tag{2.6}$$

is a V-elliptic, bounded bilinear form for any fixed $u \in V$.

We can additionally suppose the existence of positive constants c and C such that

$$c\|v\|_V^2 \leqslant a(u,v,v) \qquad \forall u,v \in V, \tag{2.7}$$

$$|a(u,v,w)| \leqslant C\|v\|_V\|w\|_V \qquad \forall u,v,w \in V. \tag{2.8}$$

We can also suppose the existence of a Gâteaux differentiable functional $\varphi\colon V \longrightarrow \mathbb{R}^1$ with the following properties:

$$D\varphi(u,v) = a(u,u,v) \qquad \forall u,v \in V, \tag{2.9}$$

(2.10) *hemicontinuity*: $t \longrightarrow D\varphi(u+tv,h)$ is continuous for any $u,v,h \in V$ and $t \in \langle 0,1\rangle$,

(2.11) *strong monotonicity*: there is a constant $\alpha > 0$ such that

$$D\varphi(u+h,h) - D\varphi(u,h) \geqslant \alpha\|h\|_V^2 \qquad \forall u,h \in V.$$

(2.12) *Lipschitz continuity*: there is a constant β such that

$$|D\varphi(u+h,k) - D\varphi(u,k)| \leqslant \beta\|h\|_V\|k\|_V \qquad \forall u,h,k \in V.$$

The last property is

$$\varphi(u) - \varphi(v) \geqslant \tfrac{1}{2}a(u,u,u) - \tfrac{1}{2}a(u,v,v). \tag{2.13}$$

It is also possible to suppose the existence of the second Gâteaux differential of φ with the following properties:

$$u \longrightarrow D^2\varphi(u,h,k) \tag{2.14}$$

is a continuous mapping in V for any fixed $h,k \in V$,

$$m\|h\|_V^2 \leqslant D^2\varphi(u,h,h) \qquad \forall u,h \in V, \tag{2.15}$$

$$|D^2\varphi(u,h,k)| \leqslant M\|h\|_V\|k\|_V \qquad \forall u,h,k \in V \tag{2.16}$$

where m and M are two positive constants.

The properties (2.10) and (2.11) imply that the functional

$$\psi(u) = \varphi(u) - b(u) \tag{2.17}$$

is coercive lower semi-continuous in V. Thus this functional attains its minimum in V which is the solution of the abstract boundary value problem (2.1). Moreover, with respect to (2.11) the problem (2.1) has *unique solution.* The same holds true for the problems (2.2) and (2.3). For the proofs see e.g. [5, 6].

Finally, note that the properties (2.10) and (2.11) can be guaranteed by the existence of the second Gâteaux differential $D^2\varphi$ with the property (2.15).

3. Deformation theory of plasticity

In this Section, we shall briefly describe the problem of the deformation theory of plasticity (physically nonlinear elasticity) which is a particular example of the problem (2.1).

This problem in a domain $\Omega \subset \mathbb{R}^d$, $d = 2$ or $d = 3$, will be described by means of

$$\text{(3.1)} \qquad \textit{the displacement} \quad u = (u_1, \cdot, u_d),$$

$$\text{(3.2)} \qquad \textit{the small strain tensor} \quad e = (e_{ij}), \quad 1 \leqslant i, j \leqslant d,$$

$$\text{(3.3)} \qquad \textit{the Cauchy stress tensor} \quad \tau = (\tau_{ij}), \quad 1 \leqslant i, j \leqslant d,$$

see [6] for the details.

We shall consider isotropic material obeying the following stress-strain relation:

$$\tau_{ij} = (k - \tfrac{2}{3}\mu)e_0\delta_{ij} + 2\mu e_{ij} \tag{3.4}$$

where k is the *bulk modulus*, μ is the *shear modulus*, $e_0 = e_{11} + \ldots + e_{dd}$ is the *volumetric strain* and δ_{ij} is the Kronecker delta.

We shall suppose that

$$k = k(x, e_0), \tag{3.5}$$

$$\mu = \mu(x, \Gamma) \tag{3.6}$$

where $x \in \Omega$ is the vector of spatial coordinates and Γ denotes the intensity of shear stress,

$$\Gamma = \Gamma(e) = \sum_{i,j} (e'_{ij})^2, \quad e'_{ij} = e_{ij} - \tfrac{1}{3}e_0\delta_{ij}. \tag{3.7}$$

In our numerical experiments we shall use special hyperbolic expressions,

$$k = \frac{k_0}{1 - \alpha k_0 e_0}, \quad \mu = \frac{A}{B + \sqrt{\frac{1}{2}\Gamma}} \tag{3.8}$$

determined by the constants k_0, α, A, B. These expressions come from the mechanics of soil, see e.g. [3].

The boundary value problem of the deformation theory of plasticity with material obeying the nonlinear Hook's law (3.4), (3.5), (3.6) can be written in the form (2.1) with

$$V = \{v = (v_1, \cdot, v_d) \colon v_i \in H^1(\Omega),\ v_i = 0 \text{ on } \Gamma_0 \text{ for } i = 1, \dots, d\}, \tag{3.9}$$

$$\begin{aligned} a(u, v, w) = \int_\Omega \{&[k(e_0(u)) - \tfrac{2}{3}\mu(\Gamma(u))] \operatorname{div} v \operatorname{div} w \\ &+ 2\mu(\Gamma(u)) \sum_{ij} e_{ij}(v)\, e_{ij}(w)\} \mathrm{d}x, \end{aligned} \tag{3.10}$$

$$b(v) = \int_\Omega \sum_i f_i v_i \mathrm{d}x + \int_{\Gamma_1} \sum_i f_i v_i \mathrm{d}s \tag{3.11}$$

where V is equipped by the inner product

$$(u, v)_V = \int_\Omega \sum_i \left[u_i v_i + \sum_j \frac{\partial u_i}{\partial x_j} \frac{\partial v_i}{\partial x_j} \right] \mathrm{d}x \tag{3.12}$$

and the corresponding norm. It is assumed that $\Gamma(u) = \Gamma(e(u)), e(u) = (eij(u))$,

$$e_{ij}(u) = \tfrac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right). \tag{3.13}$$

$F = (F, \dots, F_d)$ denotes the density of the given volume force, Γ_0, Γ_1 are two disjoint parts of the boundary of Ω on which the zero displacement and the surface force with the density $f = (f_1, \dots, f_d)$ are prescribed respectively.

Theorem 1. *Let us suppose that $k = k(x)$, i.e. that the bulk modulus does not depend on the strain. Further suppose that*

$$0 < k_0 \leqslant k(x) \leqslant k_1 < \infty \qquad \textit{for all} \quad x \in \bar{\Omega}, \tag{3.14}$$

$$0 < \mu_0 \leqslant \mu(x, s) < \tfrac{3}{2} k(x) \qquad \textit{for all} \quad x \in \bar{\Omega}, s \geqslant 0. \tag{3.15}$$

Let us also suppose that k, μ are continuous and that μ is continuously differentiable with respect to s and that

$$\frac{\partial \mu}{\partial s}(x, s) \leqslant 0 \qquad \textit{for all} \quad x \in \bar{\Omega}, s \geqslant 0, \tag{3.16}$$

$$(3.17)\qquad 0 < \kappa_0 \leqslant \mu(x,s) + \frac{\partial\mu(x,s)}{\partial s}s \qquad \textit{for all} \quad x \in \bar{\Omega}, s \geqslant 0.$$

Moreover, suppose that Γ_0 is a part of boundary of Ω with a positive measure. Then all the properties (2.6)–(2.16) are fulfilled.

Proof. For the proof of this theorem see Chapter 8 of the book [6]. Note that the condition (3.16) is exploited only for the proof of the property (2.13). □

4. Inexact generalized Picard and secant modulus methods

For the solution of nonlinear problems (2.1)–(2.3), we shall consider two linearization techniques, namely inexact generalized Picard and inexact secant modulus methods.

The *inexact generalized Picard (IGP) method* can be described by the following abstract scheme:

given $u^0 \in H$, compute u^1, u^2, ...

$$(4.1)\qquad u^{i+1} = u^i + \omega_i \Delta^i,$$

$$(4.2)\qquad \Delta_i \in H: \quad \|\Delta^i - \overline{\Delta}^i\|_{a_0} \leqslant \eta_i \|\overline{\Delta}^i\|_{a_0},$$

$$(4.3)\qquad \overline{\Delta}^i \in H: \quad a_0(\overline{\Delta}^i, v) = b(v) - a(u^i, u^i, v) \qquad \forall v \in H,$$

where $H = V$ or $H = V_h$ when the nonlinear problem (2.1) or (2.2) is solved respectively. Furthermore, ω_i are suitable constants, η_i are constants from the interval $\langle 0, 1)$ and $a_0\colon H^2 \longrightarrow \mathbb{R}^1$ is a bounded symmetric positive definite bilinear form, $\|v\|_{a_0} = \sqrt{a_0(v,v)}$.

The *inexact secant modulus (ISM) method* can be described by a similar abstract scheme:

given $u^0 \in H$, compute u^1, u^2, ...

$$(4.4)\qquad u^{i+1} = u^i + \Delta^i,$$

$$(4.5)\qquad \Delta_i \in H: \quad \|\Delta^i - \overline{\Delta}^i\|_{a_i} \leqslant \eta_i \|\overline{\Delta}^i\|_{a_i},$$

$$(4.6)\qquad \overline{\Delta}^i \in H: \quad a(u^i, \overline{\Delta}^i, v) = b(v) - a(u^i, u^i, v) \qquad \forall v \in H.$$

Here H, η_i have the same meaning as above and $\|v\|_{a_i} = \sqrt{a(u^i, v, v)}$ where $a(u^i, \cdot, \cdot)$ is supposed to be a bounded symmetric positive definite bilinear form, cf. the assumption (2.6).

The IGP and ISM methods have been described in the form which is suitable for the convergence analysis, see the next Section. But our practical interest concerns the equivalent algebraic problem (2.3).

In this case, the *inexact generalized Picard method* is described by the scheme:

$$u^{i+1} = u^i + \omega_i \Delta^i, \tag{4.7}$$

$$\Delta^i \in \mathbb{R}^n: \quad \|\Delta^i - \overline{\Delta}^i\|_{A_0} \leqslant \eta_i \|\overline{\Delta}^i\|_{A_0}, \tag{4.8}$$

$$\overline{\Delta}^i \in \mathbb{R}^n: \quad A_0 \overline{\Delta}^i = r^i = b - A_i u^i \tag{4.9}$$

where $A_i = A(u^i)$, A_0 is the symmetric positive definite stiffness matrix defined by the bilinear form a_0, $\|v\|_{A_0} = \sqrt{v^T A_0 v}$.

The *inexact secant modulus method* is then described by the similar scheme:

$$u^{i+1} = u^i + \Delta^i, \tag{4.10}$$

$$\Delta^i \in \mathbb{R}^n: \quad \|\Delta^i - \overline{\Delta}^i\|_{A_i} \leqslant \eta_i \|\overline{\Delta}^i\|_{A_i}, \tag{4.11}$$

$$\overline{\Delta}^i \in \mathbb{R}^n: \quad A_i \overline{\Delta}^i = r^i = b - A_i u^i \tag{4.12}$$

where again $A_i = A(u^i)$, $\|v\|_{A_i} = \sqrt{v^T A_i v}$.

Practically, the *inexact correction* Δ^i will be obtained by an approximate solution of the correction equation (4.9) or (4.12) by some suitable (inner) iterative method. For example, the preconditioned conjugate gradient (PCG) method have been used for this ask in our numerical experiments. Note that zero can be used as a good initial guess for Δ^i. Further, under the assumptions (2.7), (2.8) the condition numbers of $A_i = A(u^i)$ are uniformly bounded so that one iteration of CG or PCG method is sufficient to give (4.8) or (4.11) with some $\eta_i < \eta < 1$.

5. Convergence analysis

Theorem 2. *Let us consider the problem* (2.1) *and let us suppose the existence of a twice Gâteaux differentiable functional* φ *such that the assumptions* (2.9), (2.14)–(2.16) *are fulfilled.*

Moreover, let us suppose that a_0 *is a symmetric bilinear form on* V *for which positive constants* m_0 *and* M_0 *exist such that*

$$m_0\|v\|_V^2 \leqslant a_0(v,v) \leqslant M_0\|v\|_V^2 \qquad \forall v \in V. \tag{5.1}$$

Then for

$$0 \leqslant \eta_i < \tfrac{1}{2}, \quad 0 < \omega' \leqslant \omega_i \leqslant \omega'' < 2\frac{m_0}{M}\frac{1-2\eta_i}{1-\eta_i} \tag{5.2}$$

the inexact generalized Picard iterations converge to the unique solution of the problem (2.1).

Under the same assumptions the same convergence statement is valid when the problems (2.2) or (2.3) are solved.

Proof. Let us define the functional Ψ as in (2.17). Then for $t > 0$, we have

$$\Psi(u^i + t\Delta^i) = \Psi(u^i) + tD\Psi(u^i, \Delta^i) + \tfrac{1}{2}t^2 D^2\Psi(u^i + \Theta t\Delta^i, \Delta^i, \Delta^i)$$

with $0 < \Theta < 1$. Using (2.16), we obtain the estimate

$$\Psi(u^i + t\Delta^i) \leqslant \Psi(u^i) + tD\Psi(u^i, \Delta^i) + \tfrac{1}{2}t^2 M\|\Delta^i\|_V^2.$$

Now consider the first differential of Ψ.

$$\begin{aligned} D\Psi(u^i, \Delta^i) &= a(u^i, u^i, \Delta^i) - b(\Delta^i) = -a_0(\overline{\Delta}^i, \Delta^i) \\ &= -a_0(\Delta^i, \Delta^i) - a_0(\overline{\Delta}^i - \Delta^i, \Delta^i). \end{aligned} \tag{5.3}$$

From the assumption (4.2), we obtain

$$\|\overline{\Delta}^i\|_{a_0} - \|\Delta^i\|_{a_0} \leqslant \eta_i\|\overline{\Delta}^i\|_{a_0}, \quad \text{i.e.} \quad \|\overline{\Delta}^i\|_{a_0} \leqslant \frac{1}{1-\eta_i}\|\Delta^i\|_{a_0}. \tag{5.4}$$

Thus,

$$\begin{aligned} \Psi(u^i + t\Delta^i) &\leqslant \Psi(u^i) - t\|\Delta^i\|_{a_0}^2 + t\|\overline{\Delta}^i - \Delta^i\|_{a_0}\|\Delta^i\|_{a_0} \\ &+ \tfrac{1}{2}t^2\frac{M}{m_0}\|\Delta^i\|_{a_0}^2 \leqslant \Psi(u^i) + P(t)\|\Delta^i\|_{a_0}^2 \end{aligned} \tag{5.5}$$

where

$$P(t) = -t + t\frac{\eta_i}{1-\eta_i} + \frac{1}{2}t^2\frac{M}{m_0} = \frac{2\eta_i - 1}{1-\eta_i}t + \frac{1}{2}t^2\frac{M}{m_0}.$$

It can be easily verified that for $\eta_i < \frac{1}{2}$ and $\omega_i \in (\omega', \omega'')$, we have $P(\omega_i) \leqslant -\varepsilon < 0$ and therefore

$$0 \leqslant \varepsilon\|\Delta^i\|_{a_0}^2 \leqslant -P(\omega_i)\|\Delta^i\|_{a_0}^2 \leqslant \Psi(u^i) - \Psi(u^{i+1}). \tag{5.6}$$

Hence, $\|\Delta'\|_{a_0} \longrightarrow 0$ for $i \longrightarrow \infty$ because the functional Ψ is bounded from below, cf. Section 2, and the sequence $\Psi(u^i)$ does not increase.

Now we shall use the strong monotonicity of $D\Psi$:

$$D\Psi(u+h, h) - D\Psi(u, h) = \int_0^1 D^2\varphi(u + th, h, h)\mathrm{d}t \geqslant m\|h\|_V^2.$$

Let u be the unique solution of (2.1), see Section 2, i.e., $D\Psi(u,v) = 0$ for all $v \in V$. Then

$$\begin{aligned} m\|u^i - u\|_V^2 &\leqslant D\Psi(u + u^i - u, u^i - u) - D\Psi(u, u^i - u) \\ &= D\Psi(u^i, u^i - u) = a(u^i, u^i, u^i - u) - b(u^i - u) \\ &= a_0(\overline{\Delta}^i, u^i - u) \leqslant M_0\|\overline{\Delta}^i\|_V\|u^i - u\|_V \end{aligned}$$

Hence,

$$\begin{aligned} \|u^i - u\|_V &\leqslant \frac{M_0}{m}\|\overline{\Delta}^i\|_V \leqslant \frac{M_0}{m\sqrt{m_0}}\|\overline{\Delta}^i\|_{a_0} \\ &\leqslant \frac{M_0}{m\sqrt{m_0}}\frac{1}{1-\eta_i}\|\overline{\Delta}^i\|_{a_0} \end{aligned} \tag{5.7}$$

and therefore $\|u^i - u\|_V \longrightarrow 0$ for $i \longrightarrow \infty$.

The above proof can be exploited also for the finite element problem (2.2), we must only replace the space V by V_h. Finally, note that the algebraic problem (2.3) is fully equivalent to the just mentioned finite element case. □

Note 1. The condition $\eta_i < \frac{1}{2}$ in (5.2) is sufficient but not necessary for the convergence. With this respect we can note that if we replace (4.2) by a similar condition

$$\|\Delta^i - \overline{\Delta}^i\|_{a_0} \leqslant \eta_i\|\Delta^i\|_{a_0} \tag{5.8}$$

then under the assumptions of Theorem 2 we obtain the convergence for

$$\eta_i < 1, \quad 0 < \omega_i < 2\frac{m_0}{M}(1 - \eta_i). \tag{5.9}$$

For the proof of this statement, it is sufficient to follow the proof of Theorem 2. The polynom $P(t)$ in (5.5) is now replaced by

$$P(t) = -t + t\eta_i + \tfrac{1}{2}t^2\frac{M}{m_0}$$

From the other hand, for $\eta_i \longrightarrow 1$ the conditions (5.9) demand very strong damping and therefore very slow convergence may be expected.

Note 2. The convergence statement of Theorem 2 remains valid also for the inexact secant modulus method if we introduce the same damping factors ω_i to (4.4). For the proof of this fact we must only replace the assumption (5.1) by (2.7) and (2.8) and follow the proof of Theorem 2 with $a_i = a(u^i, \cdot, \cdot)$ instead of a_0.

The following Theorem concerns the convergence of the inexact secant modulus method without damping.

Theorem 3. *Let us consider the problem* (2.1) *and let us suppose the existence of a Gâteaux differentiable functional* φ *such that the assumptions* (2.7)–(2.11), (2.13) *are fulfilled.*

Then for $\eta_i < 1$ *the inexact secant modulus method (4.4)–(4.6) converges.*

Under the same assumptions the same convergence statement is valid when solve the problems (2.2) *or* (2.3).

Proof. Let u be the unique solution of the problem (2.1), see Section 2, i.e.

$$D\varphi(u,v) = b(v) \qquad \forall v \in V.$$

Then from (2.11) and the condition (4.6), we obtain

$$\begin{aligned} \alpha\|u - u^i\|_V^2 &\leqslant D\varphi(u + u^i - u, u^i - u) - D\varphi(u, u^i - u) \\ &= a(u^i, u^i, u^i - u) - b(u^i - u) = -a(u^i, \overline{\Delta}^i, u^i - u) \end{aligned}$$

Hence, from (2.8) it follows that

$$\alpha\|u - u^i\|_V \leqslant C\|\overline{\Delta}^i\|_V. \tag{5.10}$$

Denote $\bar{u}^{i+1} = u^i + \overline{\Delta}^i$, then from the uniform V-elipticity (2.7) follows

$$\begin{aligned} c\|\overline{\Delta}^i\|_V^2 \leqslant a(u^i, \overline{\Delta}^i, \overline{\Delta}^i) &= a(u^i, \bar{u}^{i+1} - u^i, \bar{u}^{i+1} - u^i) \\ &= a(u^i, \bar{u}^{i+1}, \bar{u}^{i+1}) - 2a(u^i, \bar{u}^{i+1}, u^i) + a(u^i, u^i, u^i) \\ &= b(\bar{u}^{i+1}) + 2J_i(u^i) = 2J_i(u^i) - 2J_i(\bar{u}^{i+1}) \end{aligned}$$

where $J_i(v) = \frac{1}{2}a(u^i, v, v) - b(v)$. Thus

$$\|\overline{\Delta}^i\|_V^2 \leqslant \frac{2}{c}\left[J_i(u^i) - J_i(\bar{u}^{i+1})\right]. \tag{5.11}$$

For the norm $\|v\|_{a_i} = (a(u^i, v, v))^{1/2}$ we have

$$\|v - \bar{u}^{i+1}\|_{a_i}^2 = 2J_i(v) - 2J_i(\bar{u}^{i+1}).$$

Thus, the condition (4.5) can be rewritten as

$$\|u^{i+1} - \bar{u}^{i+1}\|_{a_i} \leqslant \eta_i\|\bar{u}^{i+1} - u^i\|_{a_i}$$

or

$$J_i(u^{i+1}) - J_i(\bar{u}^{i+1}) \leqslant \eta_i^2 \left[J_i(u^i) - J_i(\bar{u}^{i+1})\right].$$

This yields

$$\begin{aligned} J_i(u^i) - J_i(\bar{u}^{i+1}) &= J_i(u^i) - J_i(u^{i+1}) + J_i(u^{i+1}) - J_i(\bar{u}^{i+1}) \\ &\leqslant J_i(u^i) - J_i(u^{i+1}) + \eta_i^2 \left[J_i(u^i) - J_i(\bar{u}^{i+1})\right], \end{aligned}$$

i.e.

$$J_i(u_i) - J_i(\bar{u}^{i+1}) \leqslant \frac{1}{1-\eta_i^2} \left[J_i(u^i) - J_i(u^{i+1})\right]. \tag{5.12}$$

Now, let us define $\Psi(v) = \varphi(v) - b(v)$ as in (2.17). From (2.13), we have

$$\tfrac{1}{2}a(u^i, u^{i+1}, u^{i+1}) - \tfrac{1}{2}a(u^i, u^i, u^i) - \varphi(u^{i+1}) + \varphi(u^i) \geqslant 0$$

and therefore

$$\begin{aligned} \Psi(u^{i+1}) &= \varphi(u^{i+1}) - b(u^{i+1}) \\ &\leqslant \varphi(u^i) - b(u^i) + b(u^i) - b(u^{i+1}) + \frac{1}{2}a(u^i, u^{i+1}, u^{i+1}) - \frac{1}{2}a(u^i, u^i, u^i) \\ &= \Psi(u^i) + J_i(u^{i+1}) - J_i(u^i). \end{aligned}$$

Hence,

$$J_i(u^i) - J_i(u^{i+1}) \leqslant \Psi(u^i) - \Psi(u^{i+1}). \tag{5.13}$$

Now, putting together (5.10)–(5.13), we obtain

$$\|u - u^i\|_V \leqslant \frac{C}{\alpha} \left\{ \frac{2}{c(1-\eta_i)} \left[\Psi(u^i) - \Psi(u^{i+1})\right] \right\}^{1/2}. \tag{5.14}$$

The functional Ψ is bounded from below, cf. Section 2, and with respect to (5.11)–(5.13) the sequence $\Psi(u^i)$ does not increase. Therefore, $(\Psi(u^i) - \Psi(u^{i+1})) \longrightarrow 0$ for $i \longrightarrow \infty$ and according to (5.14) the inexact secant modulus method converges. □

6. Efficiency of composite iterations

In this section, we restrict our attention to the solution of the algebraic problem (2.3). We shall use the inexact generalized Picard or the inexact secant modulus methods (4.7)–(4.12) with the inexact correction given by solving the linear systems (4.9) or (4.12) by a suitable iterative method. In this way, we obtain *composite iterative methods.*

The efficiency of the composite iterative method will be a function of the relative accuracy of computation of the inexact corrections. Thus, we are interested in question how the value η_i influences both the *computational work* W_i and the *reduction factor* q_i of the composite iteration.

Let us denote by $\overline{\Delta}^i$ and Δ^i the exact and the inexact correction respectively and assume that

$$\|\Delta^i - \overline{\Delta}^i\| \leqslant \eta_i \|\overline{\Delta}^i\|, \qquad 0 \leqslant \eta_i < 1 \tag{6.1}$$

where $\|\cdot\|$ is a suitable norm. Furthermore, let us define the reduction factors q_i and $\bar{q}_i$ by

$$\|u^{i+1} - u\| = q_i \|u^i - u\| \tag{6.2}$$

and

$$\|\bar{u}^{i+1} - u\| = \bar{q}_i \|u^i - u\| \tag{6.3}$$

where $u^{i+1} = u^i + \omega_i \Delta^i$, $\bar{u}^{i+1} = u^i + \omega_i \overline{\Delta}^i$ and u is the exact solution of the problem (2.3). With respect to (6.1), we have

$$\begin{aligned}\|u^{i+1} - u\| &\leqslant \|u^{i+1} - \bar{u}^{i+1}\| + \|\bar{u}^{i+1} - u\| \\ &\leqslant \eta_i \|\bar{u}^{i+1} - u^i\| + \|\bar{u}^{i+1} - u\| \\ &\leqslant \eta_i \|\bar{u}^{i+1} - u\| + \eta_i \|u^i - u\| + \bar{q}_i \|u^i - u\|.\end{aligned}$$

Therefore, we obtain the estimate

$$q_i \leqslant \eta_i \bar{q}_i + \eta_i + \bar{q}_i. \tag{6.4}$$

From this estimate, we can conclude two rough recommendations:

- it will be natural to take $\eta_i \leqslant \bar{q}_i$,
- but it is not efficient to take $\eta_i \ll \bar{q}_i$.

Our numerical experiments, partly described in the following Section, shows that the above recommendations can be strengthened. We have observed that q_i is not much greater than $\bar{q}_i$ if $\eta_i < \bar{q}_i$, see e.g. Table 1 in Section 7. Thus an efficient choice of η_i will be

$$\eta_i = \xi \bar{q}_i \tag{6.5}$$

where ξ is a positive constant little less than 1, e.g. $\xi = 0.9$.

Note that in our numerical experiments we watch the Eucledian norm $|\cdot|$ of the residuals. Thus, (6.1) will be replaced by

$$|B_i\Delta^i - B_i\overline{\Delta}^i| \leqslant \eta_i |B_i\overline{\Delta}^i| \tag{6.6}$$

where $B_i = A_0$ or $B_i = A(u^i)$ for the GP or the SM method respectively. In (6.2) and (6.3), we watch

$$|A(u^i)u^i - b| \quad \text{instead of} \quad \|u^i - u\| \quad \text{etc.} \tag{6.7}$$

The recommendation (6.5) together with the fact that $\bar{q}_i$ does not change very much *in the course of the iterative process lead to the following procedure for an efficient control of η_i:*

(i) in the first iteration take η_1 sufficiently small to obtain q_1 close to $\bar{q}_1$,

(ii) in the subsequent iterations take

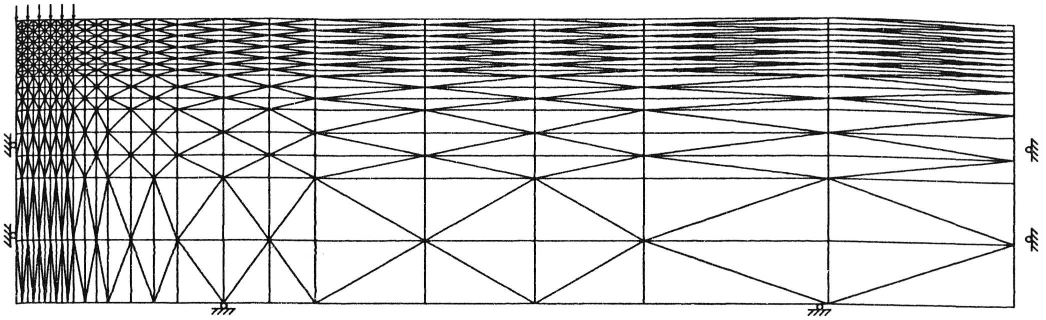
$$\eta_i = \xi q_1 \tag{6.8}$$

or

$$\eta_i = \xi \eta_{i-1} \tag{6.9}$$

supposing that $\bar{q}_i \sim \bar{q}_{i-1} \sim q_{i-1} \sim \ldots \sim q_1$.

It is also possible to repeat the steps (i) and (ii) several times during the iterative process. For example, the steps (i) and (ii) can be repeated in the moment where we observe a substantial deterioration of the convergence.



The strip footing problem—the mesh and the boundary conditions.

7. Numerical experiments

To show the behaviour of the composite iterative methods, we have solved the strip footing problem depicted in Figure 1.

The whole region of the strip footing problem consists of nonlinear elastic material with bulk and shear moduli defined by hyperbolic expressions (3.8). We shall consider two cases with different material constants:

material A: $\alpha = 0$, $k_0 = 70$, $A = .46$, $B = .01$

material B: $\alpha = 2$, $k_0 = 70$, $A = .46$, $B = .01$

In the first case, the bulk modulus is constant.

The discretization in $25 \times 19 = 475$ nodes grid with the aid of linear triangular finite elements is performed.

The discretization gives the nonlinear system of equations (2.3) which will be solved by both the inexact generalized Picard-preconditioned conjugate gradient (IGP-PCG) and the inexact secant modulus-preconditioned conjugate gradient (ISM-PCG) methods. The preconditioning for the conjugate gradient method is given by the displacement decomposition-incomplete factorization technique, see [1], [2].

The results of numerical experiments can be seen from the following tables. Note that zero initial guess was exploited in all computations.

	number of iterations								
η	1	2	3	4	5	6	7	8	9
.001	.31	.21	.45.	.46	.47	.47	.48	.48	
.01	.31	.21	.45	.46	.47	.47	.48	.48	
.1	.31	.22	.43	.45	.47	.48	.48	.48	
ADAPT	.31	.29	.38	.49	.54	.59	.43	.50	.48

Table 1. Reduction factors for the ISM-PCG method.

Table 1 shows reduction factors q_i defined by (6.2) and (6.7) for the solution of the strip footing problem with material A by ISM-PCG method. The first three rows show results corresponding to three choices of the accuracy $\eta_i = \eta$ defined in (6.6), the last row corresponds to the adaptive procedure for control of η_i exploiting the (6.9) with $\xi = 0.9$. The composite iterations are stopped when the ratio of the Eucledian norm of the residual to the Eucledian norm of the *rhs* vector is less than $\varepsilon = 0.001$.

MAT		$\eta_i = .001$	$\eta_i = .01$	$\eta_i = .1$	ADAPT
A	# COMPOSITE IT.	8	8	8	9
A	# INN. IT. (PCG)	242	180	110	78
A	COMP. WORK in WU	7 452	5 840	4 020	3 333
B	# COMPOSITE IT.	16	16	16	17
B	# INN. IT. (PCG)	452	308	159	86
B	COMP. WORK in WU	14 072	10 328	6 454	4 701

Table 2. Numbers of iterations for the composite ISM-PCG method.

Table 2 shows numbers of iterations and estimate of the computer work for solving the strip footing problem by the ISM-PCG method. The results concern both material A and material B. The work unit WU is equal to the computational work for performing the inner product with two vectors of the length equal to the dimension of the solved system. The choice of the relative accuracy for computation of the inexact correction and the stopping criterion for the composite iterations are the same as before.

MAT		$\eta_i = .1$	ADAPT
B	# COMPOSITE IT.	91	135
B	# INN. IT. (PCG)	531	199
B	COMP. WORK in WU	18 629	12 326

Table 3. Numbers of iterations for the IGP -PCG method.

Table 3 shows numbers of iterations and estimate of the computational work for solving the strip footing problem by IGP-PCG method. The results concern only the material B. The stopping criterion is the same as before.

MAT		CG	CG-DD-IF
A	# COMPOSITE IT.	8	9
A	# INNER IT.	1 685	78
A	COMP. WORK in WU	44 970	3 333

Table 4. Numbers of iterations for the composite ISM-CG and ISM-PCG methods.

Table 4 shows numbers of iterations and estimate of the computational work for solving the strip footing problem by the composite ISM-CG and ISM-PCG methods.

As the inner iterative method, we use first the conjugate gradient method without preconditioning (CG) and second the conjugate gradient method with preconditioning by displacement decomposition-incomplete factorization technique (CG-DD-IF), see [1], [2].

Finally, we would like to note that Table 1 gave a motivation for the recommendation (6.5) concerning the choice of accuracy for computation of the inexact correction. Tables 2 and 3 demonstrate efficiency of the composite iterative process including the adaptive procedure for control of the accuracy for computation of the inexact corrections. Table 4 shows the role of a good preconditioning.

For a comparison, we can note that the solution of the strip footing problem with the linear elastic material described by the bulk modulus $k = k_0$ and the shear modulus $\mu = A/B$ takes the computational work of about 1000 WU, when the described PCG iterative method is exploited for the solution of the finite element system.

References

[1] *Blaheta, R.*: Incomplete factorization preconditioning techniques for linear elasticity problems, Z. angew. Math. Mech. *71* (1991), T638-640.

[2] *Blaheta, R.*: Displacement decomposition-incomplete factorization preconditioning for linear elasticity problems, to appear in J. Numer. Lin. Alg. Appl. 1992/1993.

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

[4] *Kohut, R. and R. Blaheta*: Efficient iterative methods for numerical solution of plasticity problems, Proc. of the NUMEG'92 Conference, Prague 1992, vol. 1, pp. 129-134.

[5] *Nečas, J.*: Introduction to the theory of nonlinear elliptic equations, Teubner Texte zur Mathematik, Band 52, Leipzig, 1983.

[6] *Nečas, J. and I. Hlaváček*: Mathematical theory of elastic and elasto-plastic bodies: An introduction, Elsevier, Amsterdam, 1981.

[7] *Dembo, R.S., Eisenstat, S.C. and T. Steingang*: Inexact Newton methods, SIAM J. Numer. Anal. *19* (1982), 400-408.

[8] *Deuflhard, P.*: Global inexact Newton methods for very large scale nonlinear problems, Impact of Comp. in Science and Engng. *3* (1991), 366-393.

Authors' address: *Radim Blaheta and Roman Kohut*, Department of Applied Mathematics, Academy of Sciences of Czech Republic, Institute of Geonics, Studentská 1768, 708 00 Ostrava - Poruba, Czech Republic, fax: +42-69-44 94 52.