

Owe Axelsson

On the numerical solution of nonlinear partial differential equations on divergence form

In: Jiří Fábera (ed.): Equadiff IV, Czechoslovak Conference on Differential Equations and Their Applications. Proceedings, Prague, August 22-26, 1977. Springer-Verlag, Berlin, 1979. Lecture Notes in Mathematics, 703. pp. [5]--15.

Persistent URL: <http://dml.cz/dmlcz/702197>

## Terms of use:

© Springer-Verlag, 1979

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://project.dml.cz>

ON THE NUMERICAL SOLUTION OF NONLINEAR PARTIAL  
DIFFERENTIAL EQUATIONS ON DIVERGENCE FORM  
O. Axelsson, Göteborg

1. Introduction.

We will consider nonlinear partial differential equations on the form

$$(1.1) \quad F(u, \nabla u) = -\operatorname{div} A(u, \nabla u) + g(u, \nabla u) = 0, \quad x \in \Omega \subset \mathbb{R}^n$$

with given boundary conditions, Dirichlet conditions (for simplicity assumed to be homogeneous) on a set of measure  $> 0$  on  $\partial\Omega$ . We have

$$A^T = (A_1, \dots, A_n), \quad A_i = A_i(u, \nabla u): \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$$

$$\frac{\partial A}{\partial \xi} = \left( \frac{\partial A_i}{\partial \xi_j} \right), \quad \xi = (\xi_1, \dots, \xi_n), \quad \xi_j = \frac{\partial u}{\partial x_j}.$$

The matrix  $\frac{\partial A}{\partial \xi}$  is assumed to be uniformly positive definite (ellipticity)

$$\inf_{u, \nabla u} \xi^T \frac{\partial A}{\partial \xi} \xi \geq \rho |\xi|^2, \quad \rho > 0 \quad \forall \xi \in \mathbb{R}^n.$$

In the first part of the talk we will consider a special class of problems with so called potential operators, for which optimization (minimization) algorithms may be applied to the corresponding energy functional. In the last part of the talk the more general problem (1.1) will be dealt with by use of an embedding in a parabolic problem.

2. Potential operator problems.

Let us assume that the operator  $F$  is *potential* [1], i.e.

$$\exists f: V \rightarrow \mathbb{R} \exists$$

$$(F, \eta) = (f'(u), \eta) = (\operatorname{grad} f(u, \nabla u), \eta) \quad \forall \eta \in V \subseteq H_0^1(\Omega).$$

Then  $F \in V^*$ , the dual space of  $V$  and the variational (Galerkin) formulation of (1.1) is

$$(f'(u), \eta) = 0 \quad \forall \eta \in V.$$

( , ) is the extended scalar product in  $L^2$ . A sufficiently regular operator is potential iff its Hessian  $f''$  has a symmetric bilinear form

$$(f''\eta, \zeta) = (f''\zeta, \eta) \quad \forall \eta, \zeta \in V .$$

Here  $f''\eta = F'(u, \nabla u)\eta$ , the Gâteaux differential.

We will in particular consider potential operators on the form

$$F(u, \nabla u) = -\operatorname{div} A(\nabla u) + g(u)$$

where the matrix  $\frac{\partial A}{\partial \xi}$  is symmetric. Then, apart from an integration constant,

$$(2.1) \quad f(u) = \int_{\Omega} \left[ \int_0^u F(v, \nabla v) dv \right] dx = \int_{\Omega} \left[ \int_0^u A(\nabla v) d(\nabla v) + \int_0^u g(v) dv \right] dx$$

with

$$V = H^1(\Omega) = \{v \in H^1(\Omega); v \text{ satisfies ess.b.c's} \} .$$

In practice  $f$  often corresponds to the total energy in the system at hand. A particular example of practical importance is

$$(2.2) \quad A(\nabla u) = a(|\nabla u|^2) \nabla u .$$

Then

$$F(u, \nabla u) = - \sum_{i=1}^n \frac{\partial}{\partial x_i} \left( a(|\nabla u|^2) \frac{\partial u}{\partial x_i} \right) + g(u)$$

We also assume that the Hessian is positive definite, i.e.

$$(f''(u)\eta, \eta) \geq \delta \| \nabla \eta \|^2, \quad \delta > 0 \quad \forall \eta \in V ,$$

where  $\| \cdot \|$  is the norm in  $V$ . It is easily seen that this is satisfied if

$$\rho + \max \left( 0, \frac{1}{\mu_1} \frac{\partial g}{\partial u}(u) \right) \geq \delta > 0 ,$$

where  $\mu_1 = \mu_1(\Omega, -\Delta)$  is the smallest eigenvalue of the Laplacian operator  $(-\Delta)$  on  $V$  and  $\rho$  is the ellipticity constant. (In (2.2) we have

$$\rho = a(z) + 2 \frac{\partial a}{\partial z}(z) \cdot z, \quad z = |\nabla u|^2 .)$$

Thus  $f$  is strictly convex, proper and increasing. In practice its unique minimizer is approximated by the minimizer over a finite dimensional subspace, for instance a set of finite element functions,  $V_N \subset H_0^1(\Omega)$ , the Ritz method. This minimizer,  $\hat{u}$  may with practical efficiency be calculated by a preconditioned (scaled) conjugate

gradient method, a Newton-Kantorovich method or probably most advantageously, by a combination of these.

### 3. Algorithms for potential operator problems.

To minimize the functional  $f = f(u_1, u_2, \dots, u_N)$  over  $V_N$  we shall describe two algorithms, both of which use the Hessian matrix

$$H = \left[ \frac{\partial^2 f}{\partial u_i \partial u_j} \right].$$

In general, it is too costly to update (recalculate) this matrix frequently, so we shall give means by which this can be avoided.

#### 3.1 The Newton-Kantorovich method.

Let  $\underline{u}^\ell$  be an approximation of  $\hat{\underline{u}}$ . Then we approximate  $f$  by the quadratic functional

$$f_\ell(\underline{u}) = f(\underline{u}^\ell) + (f'(\underline{u}^\ell), \underline{u} - \underline{u}^\ell) + \frac{1}{2}(f''(\underline{u}^\ell)(\underline{u} - \underline{u}^\ell), \underline{u} - \underline{u}^\ell),$$

where the gradient and Hessian are evaluated at  $\underline{u}^\ell$ . Its minimizer, denoted by  $\underline{u}^{\ell+1}$ , satisfies

$$f'_\ell(\underline{u}) = f'(\underline{u}^\ell) + f''(\underline{u}^\ell)(\underline{u}^{\ell+1} - \underline{u}^\ell) = 0,$$

and repeating the process for  $\ell = 1, 2, \dots, \underline{u}^1$  given, we have arrived at the classical Newton-Kantorovich method for the solution of  $f'(\underline{u}) = 0$ . The quadratic convergence is assured if

$$\|\underline{u}^2 - \underline{u}^1\| < 2\delta/K$$

(see e.g. [1] and [2]). Here  $K$  is an upper bound on the second Gateaux differential,

$$|F''\eta\zeta|, \quad \|\eta\| = \|\zeta\| = 1, \quad \eta, \zeta \in V.$$

At each Newton step we do not have to assemble the Hessian matrix, as would be the case in a direct LU-factorization method. This is of importance in particular in three-dimensional problems,  $\Omega \subset \mathbb{R}^3$  (see e.g. [3]). Instead we calculate the minimizer of  $f_\ell$  by the preconditioned conjugate gradient (PCCG) method:

```

u := uℓ;  g := -f'(uℓ);
Cγ := g;  e := -γ;  δ0 := gTγ;  ε := ε0δ0;
R:  λ := -gTe/eTHe;

u := u + λe;

g := g + λHe;  Cγ := g;

δ1 := gTγ;  β := δ1/δ0;  δ̄0 := δ1;

e := -γ + βe;

IF δ1 > ε THEN GOTO R;

u := uℓ + u;

```

If  $V_N$  is spanned by  $N$  basis (or coordinate) functions  $\phi_i(x) \in H_0^1(\Omega)$  with *local support* on a "small" element (the finite element method) the matrix-vector multiplication  $He$  is calculated as a sum of its contributions from each local element. In this way  $H$  does never have to be calculated, only the local finite element matrices are calculated.

The rate of convergence of the PCCG-algorithm is linear and the number of conjugate gradient steps, i.e. number of times  $He$  is calculated, is at most

$$p = \text{int} \left[ \frac{1}{2} \sqrt{\mathcal{K}} \ln \frac{2}{\varepsilon_0} + 1 \right]$$

where  $\mathcal{K}$  is the spectral condition number, i.e. the quotient between the extreme eigenvalues, of  $C^{-1}H$ .  $C$  is usually a product of two sparse triangular matrices. In [3] it is shown that it is possible to choose  $C$  such that  $\mathcal{K} = O(N^{1/n})$ . Then  $p \sim O(N^{1/2n})$ , i.e. a small increase with the number of unknowns  $N$ .

### 3.2 Efficiency in handling the updating of the Hessian matrix.

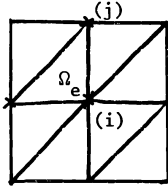
Assume for simplicity that

$$A(\nabla u) = a(|\nabla u|^2) \nabla u .$$

The corresponding local element "stiffness" matrices are

$$k_{ij}^{(e)} = \int_{\Omega_e} a(|\nabla u|^2) \nabla \phi_i(x) \nabla \phi_j(x) dx ,$$

where  $\Omega_e$  is the  $e$ 'th element. Only basis functions with a common support over the element give non-zero elements (see figure).



The corresponding part of the Hessian matrix has a similar form. The global matrices

$$K_{ij} = \sum_e k_{ij}^{(e)} \quad (\text{and } H_{ij})$$

do not have to be assembled (thereby avoiding possible cancellation of digits). We approximate

$$k_{ij}^{(e)} \simeq a(|\nabla u|_e^2) \int_{\Omega_e} \nabla \phi_i(x) \nabla \phi_j(x) dx .$$

For linear finite elements, this is exact, since then  $|\nabla u|$  is locally constant. Only the first factor have to be reevaluated at each new Newton-step. The second factor is evaluated once and for all and stored, when the finite element mesh has been generated. This is so also for so called geometrically nonlinear problems (cf. Section 5). This is done as long as the relative change in the functional is large enough. When this change is small a true gradient should be calculated during the last Newton-steps (cf. [4]). This will give more accurate approximations of the solution.

### 3.3 PCCG with restart.

An alternative to the Newton-Kantorovich method is to use a preconditioned conjugate gradient method for the minimization of the generally nonquadratic functional (2.1). Then the only change in the PCCG-algorithm is that  $\lambda$  has to be evaluated by some linesearch procedure (like Newtons modified method for one unknown variable) and the gradient  $g$  is evaluated as  $g := f'(u)$  at each c-g step. For problems with a strong nonlinearity, it may be advisable to restart the algorithm with a search along the negative gradient at every  $r$ 'th step (i.e.  $\beta := 0$  then). We observe that the number of iterations  $p = O(\sqrt{N})$  is still valid, if only  $r \geq 2$ . In the classical steepest descent method (where  $r = 1$ ) we have however  $p = O(N)$ .

### 3.4 On preconditioning.

To explain the effect of preconditioning (or scaling) we consider a functional

$$\tilde{f}(\tilde{\underline{u}}) = f(E^T \tilde{\underline{u}})$$

where the variable is transformed by a simple (i.e. triangular matrix  $E^T$ ). The minimizer of  $\tilde{f}$  is  $\tilde{\underline{u}} = E^{-T} \underline{u}$ . We get a new gradient

$$\tilde{\underline{g}} = E \underline{g}(E^T \tilde{\underline{u}})$$

and Hessian

$$\tilde{H} = E H (E^T \tilde{\underline{u}}) E^T .$$

If we work with untransformed quantities  $\underline{u} := E^T \tilde{\underline{u}}$ , then since  $\tilde{\underline{u}} := \underline{u} + \lambda \tilde{\underline{g}}$  etc, we get  $\underline{g} := E^T \tilde{\underline{g}} = E^T E \underline{g}$ . Thus the only change in the classical c-g algorithm is  $\underline{g} \rightarrow E^T E \underline{g}$ , and we arrive at the PCCG-algorithm.

Apparently the best choice of  $E$  is such that  $\mathcal{K}(\tilde{H}) \simeq 1$ . If  $E^{-1} E^{-T} \simeq H$ , this is the case. Thus we may let  $C = E^T E$  be an approximate factorization of  $H$ , (actually modified by a relaxation parameter). At each PCCG-step we have then to solve  $C \underline{\gamma} = \underline{g}$ , which is not costly since  $C$  is the product of two sparse triangular matrices.

Similar methods as described in 3.3 have been used by [5], [6], [7] among others.

### 4. Parabolic imbedding.

Consider now the problem (1.1) with

$$\rho_0 \leq \mu_1(\Omega, -\Delta)(\rho - \delta), \quad \delta > 0 ,$$

where  $\mu_1$  is the smallest eigenvalue of  $-\Delta$  on  $\Omega$  and

$$\rho_0 = \sup_{\underline{v}, \nabla \underline{v}} \left\{ \frac{1}{2} \operatorname{div} \left[ \frac{\partial A}{\partial \underline{u}}(\underline{v}, \nabla \underline{v}) + \frac{\partial \underline{g}}{\partial \xi}(\underline{v}, \nabla \underline{v}) \right] - \frac{\partial \underline{g}}{\partial \underline{u}}(\underline{v}, \nabla \underline{v}) \right\} .$$

Then it is easily seen that the operator

$$F = - \operatorname{div} A(\underline{u}, \nabla \underline{u}) + \underline{g}(\underline{u}, \nabla \underline{u})$$

is *strongly monotone*:

$$(4.1) \quad \begin{aligned} a(u, w; v) &= (A(u, \nabla u) - A(w, \nabla w), \nabla v) + (g(u, \nabla u) - g(w, \nabla w), v) \\ &\geq \delta \| \nabla(u - w) \|^2 \quad \forall u, w \in V . \end{aligned}$$

Here  $\| \cdot \|$  is the  $L_2$ -norm. The corresponding parabolic problem,

$$(4.2) \quad u_t + F(u, \nabla u) = 0, \quad t > 0, \quad u(x, 0) = u_0(x)$$

has then a unique solution for all  $t > 0$  and is asymptotically stable, that is,

$$u(x, \infty) = \lim_{t \rightarrow \infty} u(x, t)$$

exists uniquely, independently of the initial function  $u_0(x)$  (see [8]). Thus  $u(x, t)$ ,  $t$  large enough may be used as an approximation of  $u(x, \infty)$ . If one is only interested in this stationary solution, the initial function should be chosen as a smooth function satisfying the boundary conditions, so that higher order modes in the corresponding "Fourier series" have small components.

#### 4.1 Discretization error estimates.

A semi-discrete approximation of (4.2) is achieved by Galerkin's method. A variational formulation of (4.1) is

$$(u_t, v) + (A(u, \nabla u), \nabla v) + (g(u, \nabla u), v) = 0 \quad \forall v \in H_0^1(\Omega) ,$$

and the corresponding Galerkin formulation is

$$(U_t, V) + (A(U, \nabla U), \nabla V) + (g(U, \nabla U), V) = 0 \quad \forall V \in V_N \subset H_0^1(\Omega) .$$

Let  $Z \in V_N$ , for the moment be arbitrary and subtract

$(Z_t, v) + (A(Z, \nabla Z), \nabla v) + (g(Z, \nabla Z), v)$ . Let  $\mathcal{V} = U - Z$ ,  $\eta = u - Z$ . With  $v = \mathcal{V}$  we get, by monotonicity (4.1)

$$(4.3) \quad \frac{1}{2} \frac{d}{dt} (\mathcal{V}, \mathcal{V}) + \delta \| \nabla \mathcal{V} \|^2 \leq |(\eta_t, \mathcal{V})| + |a(u, Z; \mathcal{V})| .$$

We have

$$(4.4) \quad |(\eta_t, \mathcal{V})| \leq C \| \eta_t \|_{-1}^2 + \frac{\delta}{4} \| \nabla \mathcal{V} \|^2 .$$

To estimate the second term we choose for each  $t > 0$ ,  $Z$  as an elliptic projection of  $u$  such that the elliptic projection errors  $\| \eta \|$  and  $\| \eta_t \|$  are simple



to estimate and of optimal order (cf. [9], [10], [11] and [12]).

Two different approaches to this problem have been used. In [13],  $Z$  is defined as

$$\sum_{i=1}^n (A_i(u, \nabla u) - A_i(u, \nabla Z), \frac{\partial V}{\partial x_i}) + (g(u, \nabla u) - g(u, \nabla Z), V) = 0 \quad \forall V \in V_N,$$

which is still a nonlinear problem in  $Z$ . Here we use the linearized operator

$$\mathcal{A}(\omega, \nabla \omega) = \begin{bmatrix} \frac{\partial A}{\partial \xi}(\omega, \nabla \omega) & \frac{\partial A}{\partial u}(\omega, \nabla \omega) \\ \frac{\partial g}{\partial \xi}(\omega, \nabla \omega) & \frac{\partial g}{\partial u}(\omega, \nabla \omega) \end{bmatrix}$$

to define the bilinear form

$$b(u, \nabla u; \eta, V) = \int_{\Omega} [\nabla V^T, V] \mathcal{A}(u, \nabla u) \begin{bmatrix} \nabla \eta \\ \eta \end{bmatrix} dx = 0 \quad \forall V \in V_N$$

where  $\eta = u - Z$  is the elliptic projection error. Since

$$b(u, \nabla u; V, V) \geq \delta \|\nabla V\|^2 \quad \forall V \in V_N$$

$b$  is a coercive form. Then it is possible to prove the quasioptimal error estimates (see [8])

$$\|\eta\|_j = Ch^{s-j} \|u\|_s, \quad j = 0, 1, \quad 2 \leq s \leq r + 1$$

$$\|\eta_t\|_1 \leq Ch^{s-1} [\|u\|_s + \|u_t\|_s]$$

and

$$\|\eta_t\|_{-1} \leq Ch^{r_0} [\|\eta_t\|_1 + \|\eta\|_1], \quad r_0 = \min(2, r)$$

where  $C = C(u)$  and  $r$  is the degree of the continuous piecewise polynomials.

These estimates are valid for all  $t > 0$ . By Cauchy-Schwarz inequality we now get from (4.3), (4.4)

$$\frac{d}{dt} (\mathcal{V}, \mathcal{V}) + \delta \|\nabla \mathcal{V}\|^2 \leq C [\|\eta_t\|_{-1}^2 + \int_{\Omega} |\nabla \eta|^4 dx].$$

By the inverse assumptions,

$$\|\nabla V\|_{L_{\infty}^{-v}} \leq Ch^{-v} \|V\|, \quad V \in V_N$$

$$\inf_{\chi \in V_N} [\|\nabla(u - \chi)\|_{L_{\infty}^{-v}} + h^{-v} \|u - \chi\|] \leq Ch^{s-v} \|u\|_{s+\varepsilon}, \quad 2 \leq s \leq r+1, \\ 0 < \varepsilon < 1,$$

where for quasiregular elements,  $v_{\infty} = \frac{n}{2} + 1$ , we get

$$\| \nabla \eta \|_{L_\infty} \leq Ch^{s-\nu_\infty} \| u \|_{s+\varepsilon}.$$

Finally we then get

$$\frac{d}{dt}(\mathcal{V}, \mathcal{V}) + \delta \| \nabla \mathcal{V} \|^2 \leq Ch^{2r_0} \| \eta_t \|_1^2 + C[1 + \| u \|_{s+\varepsilon}] h^{2 \min(r_0, s-\nu_\infty)} \| \eta \|_1^2.$$

Thus by Gronwall's inequality

$$\| \mathcal{V}(\cdot, t) \| \leq \exp(-\frac{\delta}{2} t) \| \mathcal{V}(\cdot, 0) \| + C \sup_{t>0} [h^{r_0} \| \eta_t \|_1 + h^{\min(r_0, s-\nu_\infty)} \| \eta \|_1],$$

$$2 \leq s \leq r + 1.$$

The first term on the right hand side shows the independence of the error of the chosen initial function as  $t \rightarrow \infty$ . If  $s = r + 1$ ,  $r \geq \nu_\infty > 1$ ,  $\nu_\infty = \frac{n}{2} + 1$ , we have thus proven optimal order estimates

$$\| u - U \| \leq Ch^{r+1} \sup_{t>0} [\| u \|_{r+1} + \| u_t \|_r], \quad t > 0.$$

#### 4.2 Time-integration.

To complete the discretization we choose a simple method, the so called  $\theta$ -method for the time-integration. Let  $W$  be the corresponding approximation, then

$$(4.5) \quad (W(t+k) - W(t), V) + k[(A(\bar{W}, \nabla \bar{W}), \nabla V) + (g(\bar{W}, \nabla \bar{W}), V)] = 0 \quad \forall V \in V_N$$

where

$$\bar{W}(t) = \theta W(t) + (1 - \theta)W(t+k) \quad \text{and } k > 0 \text{ is time-integration step.}$$

If  $0 \leq \theta \leq \frac{1}{2} - |O(k)|$  one may prove the error estimate

$$\| u - W \| = O(h^{r+1}) + (\theta - \frac{1}{2})O(k) + O(k^2) = O(h^{r+1}) + O(k^2),$$

valid for all  $t > 0$  (see [8]). Finally we have to linearize (4.5) and this is again done by help of the bilinear form  $b$ . Let  $Y$  be the solution of the linearized problem

$$(4.6) \quad (Y(t+k) - Y(t), V) + kb(\tilde{Y}, \nabla \tilde{Y}; \bar{Y} - \tilde{Y}, V) \\ = -k[(A(\tilde{Y}, \nabla \tilde{Y}), \nabla V) + (g(\tilde{Y}, \nabla \tilde{Y}), V)] \quad \forall V \in V_N.$$

Here  $\tilde{Y} = Y$  or  $\tilde{Y}(t) = \theta Y(t) + (1 - \theta)\tilde{Y}(t+k)$ ,  $\tilde{\tilde{Y}}(t+k) = Y(t) + k \frac{d}{dt} Y(t)$ . The error due to linearization is  $O(k^2)$  if  $k \leq ch^{\frac{\varepsilon+\nu_\infty}{2}}$ ,  $\varepsilon > 0$  or  $O(k^4)$ , if  $k \leq ch^{\frac{\varepsilon+\nu_\infty}{4}}$ . We observe that (4.6) may be considered as a damped Newton-Kantorovich method for the numerical solution of

$$(F(U, \nabla U), V) = 0 \quad \forall V \in V_N.$$

As  $k \rightarrow \infty$  we get the (undamped) Newton-Kantorovich method. A numerical test of the above presented method is found in [4].

### 5. Applications.

Nonlinear monotone or even potential operators are found in many important practical applications. There are two classes of such examples,

- (i) Problems with nonlinear material properties
- (ii) Problems with nonlinear effects due to geometry.

Examples of the first kind are

$$(5.1) \quad a) \quad -\nabla(a(|\nabla u|^2)\nabla u) + g(u) = 0 \quad \text{in } \Omega$$

with essential and/or natural boundary conditions on different parts of the boundary  $\partial\Omega$  and are met in electromagnetic field equations and in torsion of a prismatic bar.

$$b) \quad \begin{aligned} &\nabla(a(T)\nabla T) + g(u) = 0 \quad \text{in } \Omega \\ &-\lambda \frac{\partial T}{\partial \nu} = \alpha(T - T_0) + \gamma(T^4 - T_0^4) \quad \text{on } \partial\Omega, \\ &\alpha, \lambda, \gamma \geq 0, \quad \text{a nonlinear heat convection equation.} \end{aligned}$$

Examples of the second kind are

- a) (5.1) with  $a(|\nabla u|^2) = 1/(1 + |\nabla u|^2)^{1/2}$  the minimal surface equation.
- b) (5.1) with the function  $a$  as above,  $g = Ku$  and boundary condition (see [14])  

$$\frac{\partial u}{\partial \nu} / (1 + |\nabla u|^2)^{1/2} = c \quad \text{on } \partial\Omega.$$
- c) Large displacements theories like the von Karman model for a membrane

$$f(u) = \int_{\Omega} \left\{ h \left[ (2u_x + w_x^2)^2 + (2v_y + w_y^2)^2 + 2(u_y + u_x + w_x w_y)^2 \right. \right. \\ \left. \left. + \frac{\nu}{1 - \nu} (2u_x + 2v_y + w_x^2 + w_y^2)^2 \right] + P \frac{w_x u + w_y u - w}{(1 + w_x^2 + w_y^2)^{1/2}} \right\} \partial\Omega$$

to be minimized over  $[H_0^1(\Omega)]^3$ , where  $u, v, w$  are the displacements in the  $x, y, z$ -directions, respectively.

References.

- [1] M.M. Vainberg, Variational method and method of monotone operators in the theory of nonlinear equations, Wiley, 1973.
- [2] O. Axelsson, U. Nävert, On a graphical package for nonlinear partial differential equation problems, Proc. IFIP Congress 77, B. Gilchrist (ed.), North-Holland Publ. Comp. (1977).
- [3] O. Axelsson, A class of iterative methods for finite element equations, Comp. Meth. Appl. Mech. Engrg 9(1976), 123-137.
- [4] O. Axelsson, T. Steihaug, Some computational aspects in the numerical solution of parabolic equations, in preparation.
- [5] R. Bartels, J.W. Daniel, A conjugate gradient approach to nonlinear elliptic boundary value problems in irregular regions, CNA63, Center for Numerical Analysis, Austin, Texas, 1973.
- [6] P. Concus, G.H. Golub, D. O'Leary, Numerical solution of nonlinear elliptic partial differential equations, STAN-CS-76-585, Computer Science Department, Stanford University, 1976.
- [7] D. O'Leary, Hybrid conjugate gradient algorithms, STAN-CS-76-548, Computer Science Department, Stanford University, 1976.
- [8] O. Axelsson, Error estimates for Galerkin methods for quasilinear parabolic and elliptic differential equations in divergence form, Numer. Math. 28, 1-14, (1977).
- [9] T. Dupont, Some  $L^2$  error estimates for parabolic Galerkin method in The mathematical foundations of the finite element method with applications to partial differential equations (A.K. Aziz, ed.) 491-504. Academic Press N.Y. 1972.
- [10] I. Hlaváček, On a semi-variational method for parabolic equations I, Aplikace Matematiky 17(1972), 327-351.
- [11] J. Douglas, T. Dupont, A Galerkin method for a nonlinear Dirichlet problem, Math. Comp. 29(1975), 689-696.
- [12] M.F. Wheeler, A priori  $L^2$  error estimates for Galerkin approximations to parabolic partial differential equations, SIAM J. Numer. Anal. 10(1973), 713-759.
- [13] J.E. Dendy, Galerkin's method for some highly nonlinear problems, SIAM J. Numer. Anal. 14(1977), 327-347.
- [14] H.D. Mittelmann, Numerische Behandlung nichtlinearer Randwertproblemen mit finiten Elementen, Computing 18, 67-77 (1977).

Author's address: Chalmers University of Technology, Göteborg, Sweden.