BIT 37:2 (1997), 296–311.

ON THE NONLINEAR DOMAIN DECOMPOSITION METHOD *

M. DRYJA[1] and W. HACKBUSCH[2]

[1] *Institute of Applied Mathematics and Mechanics, University of Warsaw Banacha 2, 02-097 Warsaw, Poland. email: m.dryja@appli.mimuw.edu.pl*

[2] *Lehrstuhl Praktische Mathematik, Christian-Albrechts-Universität zu Kiel D-24098 Kiel, Germany. email: wh@numerik.uni-kiel.de*

Abstract.

Any domain decomposition or additive Schwarz method can be put into the abstract framework of subspace iteration. We consider generalizations of this method to the nonlinear case. The analysis shows under relatively weak assumptions that the nonlinear iteration converges locally with the same asymptotic speed as the corresponding linear iteration applied to the linearized problem.

AMS subject classification: 65H10, 65N55.

Key words: Domain decomposition, nonlinear problems, preconditioner, nonlinear elliptic problems.

0 Introduction.

We describe a generalization of the subspace iteration to the nonlinear case. It reduces to the standard subspace iteration in the linear case. In this sense, it is a true generalization of the linear subspace iteration. It requires the solution of local nonlinear subproblems. The analysis shows under relatively weak assumptions that the nonlinear iteration converges locally with the same asymptotic speed as the corresponding linear iteration applied to the linearized problem. To simplify the presentation, we assume for the linear as well as for the nonlinear case that the local subproblems are solved exactly. It is straightforward to replace the exact solution by an iterative one.

The nonlinear subspace iteration presented in this paper is constructed from the linear one in the same way as the linear multi-grid method is generalized to the nonlinear one. Therefore the tools of the analysis are similar to those in the multi-grid case (cf. Hackbusch [6, §9]).

In the second part of this paper, we discuss conditions under which global convergence can be guaranteed. The proposed algorithm uses different strategies depending on whether the iterates are in a neighbourhood of the solution or outside. We remind the reader of a global convergence result for the nonlinear multi-grid method (cf. Hackbusch and Reusken [8]).

*Received December 1995. Revised September 1996.

A different approach to the solution of nonlinear problems is the use of Newton's method. The involved linear equations can be solved, e.g., by a domain decomposition method as secondary iteration. The nonlinear solver is analysed by Cai and Dryja [1] for some special cases. A convergence analysis of a similar algorithm under stronger conditions is considered in Part II of Tai [10].

1 The general setting of the problems.

1.1 The nonlinear problem.

We consider the finite dimensional nonlinear equation

$$F(x) = 0 \qquad x \in D \subset X, \ \dim X < \infty. \tag{1.1a}$$

The fact that this equation may be a finite element discretization of the nonlinear boundary value problem (see §2) will not be used in this section. We assume that there is at least one solution x^* of (1.1a) in D, which will be fixed in the following:

$$\text{there is} \quad x^* \in D \quad \text{with} \quad F(x^*) = 0. \tag{1.1b}$$

Further, we assume that the solution x^* is locally isolated:

(1.1c) $U \subset D$ is a neighbourhood of x^* such that x^* is the unique solution of (1.1a) in U;

(1.1d) the Fréchet derivative $F'(x^*)$ exists and is non-singular.

The multidimensional analogue of a difference quotient is the following operator DF. We assume that

(1.1e) a uniformly bounded linear operator $DF(x', x'') \in L(X, X)$ is defined for all $x', x'' \in X$ (i.e., $||DF(x', x'')|| \leq C$ for all $x', x'' \in X$ such that

$$F(x') - F(x'') = DF(x', x'')(x' - x'') \quad \text{and} \tag{1.1f}$$

$$||DF(x', x'') - F'(x^*)|| \to 0 \quad \text{as} \quad x', x'' \to x^*. \tag{1.1g}$$

Here, $||\cdot||$ denotes the norm of X and at the same time the induced operator norm in $L(X, X)$. The assumptions made above are very weak as shown in the following remarks.

REMARK 1.1. Since X is finite dimensional, $DF(x', x'')$ and $F'(x^*)$ are bounded. Moreover, the convergence (1.1g) implies $||DF(x', x'')|| \leq 2||F'(x^*)||$ for all x', x'' in a suitable neighbourhood $U' \subset U$ of x^*.

REMARK 1.2. Assume that $F'(x)$ exists for all $x \in U$. Then a possible choice of $DF(x', x'')$ is

$$DF(x', x'') = \int_0^1 F'(x' + t(x'' - x'))\, dt. \tag{1.2}$$

If, in addition, $F'(x)$ is continuous at x^* (i.e., $||F'(x^*) - F'(x)|| \to 0$ as $x \to x^*$), then (1.1g) holds for DF from (1.2).

REMARK 1.3. The assumption (1.1f) may even be satisfied when F is not F-differentiable in $U\{x^*\}$. As an example take any Lipschitz continuous function $F : R \to R$ with $F'(0)$ existing at 0. Then (1.1f) is valid for the difference quotient $DF(x', x'') := (F(x') - F(x''))/(x' - x'')$ if $x' \neq x''$ and $DF(x,x) := F'(0)$, otherwise. However, (1.1g) holds only in the weaker sense that as $x', x'' \to x^*$, x', x'' also have to satisfy $||x' - x^*|| + ||x'' - x^*|| \leq const||x' - x''||$.

1.2 Subspaces.

Any subspace method (Schwarz method or domain decomposition method) can be characterized by means of spaces X_κ ($\kappa \in I$, I is a finite index set), linear injective mappings

$$p_\kappa : X_\kappa \to X, \tag{1.3a}$$

and linear surjective mappings

$$r_\kappa : X \to X_\kappa. \tag{1.3b}$$

The spaces X_κ are the spaces used in the computations, while the images $p_\kappa X_\kappa$ are the subspaces of X involved in the subspace iteration method. The subspaces $p_\kappa X_\kappa$ must cover the whole space X, i.e., it is required throughout this paper that

$$X = \sum_{\kappa \in I} p_\kappa X_\kappa. \tag{1.4}$$

The standard choice of the "restriction" r_κ is

$$r := p_\kappa^T \text{ (adjoint in the Euclidean inner product).} \tag{1.5}$$

An analysis of linear subspace iterations using this notation can be found in Hackbusch [7, §11].

1.3 Nonlinear subspace iteration.

In the following we describe a nonlinear iteration which reduces to the standard subspace iteration in the linear case. In this sense, it is a true generalization of the linear subspace iteration. Let $\tilde{x}$ be a given approximation in the neighbourhood U of x^*. For each index $\kappa \in I$, we pose the following *nonlinear subspace problem*:

Find $\delta_\kappa \in X_\kappa$ such that

$$r_\kappa F(\tilde{x} - p_\kappa \delta_\kappa) = 0. \tag{1.6}$$

In (1.6) we require the exact solution of the subspace problem. In the numerical application, this will be replaced by an approximate solution analogously to the linear case (cf. [7]). Since the analysis of secondary iteration would complicate the discussion without gaining more insight, we restrict our discussion to (1.6).

In general, the solvability of (1.1a) in X does not imply solvability of (1.6) in the subspace X_κ. To establish the latter, we suppose that the operator

$$r_\kappa F'(x^*)p_\kappa : X_\kappa \to X_\kappa \tag{1.7}$$

is invertible. In the linear case, (1.7) means that the subspace problem is solvable. Usually, this assumption is not explicitly written, since the assumptions of the following remark are satisfied.

REMARK 1.4. If $F'(x^*)$ is positive definite and (1.5) holds, then $r_\kappa F'(x^*)p_\kappa$ is also positive definite, so that (1.7) holds.

The next result shows that (1.7) guarantees the local solvability of the nonlinear problems (1.6).

THEOREM 1.1. *Assume* (1.1a–g), (1.7) *and choose a sufficiently small neighbourhood* U' *of* x^* (*i.e.*, $x^* \in U' \subset U \subset X$). *Then the subspace problem* (1.6) *is uniquely solvable for all* $\tilde{x} \in U'$ *and all indices* $\kappa \in I$.

By "uniquely solvable" we mean that there is a unique continuous function $\delta_\kappa = \delta_\kappa(\tilde{x})$ *satisfying* (1.6) *for* $\tilde{x} \in U'$ *with* $\delta_\kappa(x^*) = 0$.

PROOF. Set $A_\kappa := r_\kappa F'(x^*)p_\kappa$. Since κ is fixed in the following, we write δ instead of δ_κ. The mapping $\phi = \phi_\kappa$ defined below describes a Newton-like iteration:

$$\delta^{\text{new}} := \phi(\delta^{\text{old}}, \tilde{x}) \quad \text{with} \quad \phi : X_\kappa \times U' \to X_\kappa \quad \text{defined by}$$

$$\phi(\delta, \tilde{x}) := \delta + A_\kappa^{-1} r_\kappa F(\tilde{x} - p_\kappa \delta). \tag{1.8}$$

Next, we prove that ϕ is a contraction. We write the difference as

$$\begin{aligned} \phi(\delta', \tilde{x}) - \phi(\delta'', \tilde{x}) &= \delta' - \delta'' + A_\kappa^{-1} r_\kappa [F(\tilde{x} - p_\kappa \delta') - F(\tilde{x} - p_\kappa \delta'')] \\ &= [I - A_\kappa^{-1} r_\kappa DF(\tilde{x} - p_\kappa \delta', \tilde{x} - p_\kappa \delta'')p_\kappa](\delta' - \delta''). \end{aligned}$$

Since $DF(\tilde{x} - p_\kappa \delta', \tilde{x} - p_\kappa \delta'') \to F'(x^*)$ as $\delta', \delta'' \to 0$ and $\tilde{x} \to x^*$ due to (1.1g) and since $I - A_\kappa^{-1} r_\kappa F'(x^*)p_\kappa = 0$ by definition of A_κ, the square bracket approaches 0 as $\delta', \delta'' \to 0$ and $\tilde{x} \to x^*$. Hence, the estimate

$$||I - A_\kappa^{-1} r_\kappa DF(\tilde{x} - p_\kappa \delta', \tilde{x} - p_\kappa \delta'')p_\kappa|| \le 1/2$$

holds in a suitable neighbourhood $U'' \subset U$ of x^*, i.e., for $\tilde{x} \in U''$ and $\tilde{x} - p_\kappa \delta'$, $\tilde{x} - p_\kappa \delta'' \in U''$.

For $\tilde{x} = x^*$, the fixed point of $\delta = \phi(\delta, \tilde{x})$ is $\delta = 0$. Due to the local version of Banach's fixed point theorem, (1.8) has a fixed point close to 0 as long as $\tilde{x}$ is in a neighbourhood U''' of x^*. Choosing $U' := U'' \cap U'''$, we can guarantee the existence of a fixed point of $\delta - \phi(\delta, \tilde{x})$. By definition, this is a solution of (1.6): $r_\kappa F(\tilde{x} - p_\kappa \delta) = 0$. □

The proof is similar to that of the implicit function theorem. There, however, one usually requires the existence of derivatives in a neighbourhood of x^*.

Denote the solution of (1.6) by

$$\delta_\kappa = \delta_\kappa(\tilde{x}). \tag{1.9}$$

REMARK 1.5. The function δ_κ satisfies

$$\delta_\kappa(x') - \delta_\kappa(x'') = D\delta_\kappa(x',x'')(x' - x'') \quad \text{for } x',x'' \in U' \tag{1.10}$$

(U' from Theorem 1.1) with $D\delta_\kappa$ defined by

$$D\delta_\kappa := [r_\kappa DFp_\kappa]^{-1} r_\kappa DF, \quad DF := DF(x' - p_\kappa\delta_\kappa(x'), x'' - p_\kappa\delta_\kappa(x'')). \tag{1.11}$$

The operator $D\delta_\kappa$ satisfies

$$D\delta_\kappa(x',x'') \to [r_\kappa F'(x^*)p_\kappa]^{-1} r_\kappa F'(x^*) \quad \text{as } x',x'' \to x^*.$$

PROOF. The equation

$$\begin{aligned} 0 &= 0 - 0 = r_\kappa F(x' - p_\kappa\delta_\kappa(x')) - r_\kappa F(x'' - p_\kappa\delta_\kappa(x'')) \\ &= r_\kappa DF(x' - p_\kappa\delta_\kappa(x'), x'' - p_\kappa\delta_\kappa(x''))[x' - x'' - p_\kappa(\delta_\kappa(x') - \delta_\kappa(x''))] \end{aligned}$$

yields

$$r_\kappa DF[x' - x''] = r_\kappa DFp_\kappa(\delta_\kappa(x') - \delta_\kappa(x'')).$$

Note that $r_\kappa DFp_\kappa$ is invertible because of (1.7) and (1.1g) if the neighbourhood is small enough. Applying the inverse of $r_\kappa DFp_\kappa$, we obtain (1.11). □

The solutions $\delta_\kappa = \delta_\kappa(\tilde{x})$ for each κ from (1.6) are added to form the global correction δ, defined by

$$\delta(\tilde{x}) = \delta := \sum_{\kappa\in I} p_\kappa\delta_\kappa(\tilde{x}). \tag{1.12}$$

One step of the (nonlinear) subspace iteration is defined by $x^{\text{new}} := \Phi(x^{\text{old}})$, where

$$\Phi(x) := x - \omega\delta(x) \quad (\delta(x) \text{ from } (1.12)). \tag{1.13}$$

The value of the damping parameter ω will be determined in §1.4.

1.4 The linear case.

We consider the linear problem $Ax - b = 0$ with $A := F'(x^*)$, $b = Ax^*$. This linear problem can be regarded as the linearization of the nonlinear problem (1.1a) at the true solution x^*. When we apply the subspace iteration (1.13) to the linear problem $F(x) := Ax - b = 0$ with $A := F'(x^*)$, $b = Ax^*$, we obtain a linear iteration with the iteration matrix

$$M_\omega := I - \omega\sum_{\kappa\in I} p_\kappa A_\kappa^{-1} r_\kappa A, \qquad A_\kappa = r_\kappa A p_\kappa \tag{1.14}$$

(see [7, (11.2.14b)]). The resulting additive subspace method or Schwarz iteration is well-discussed in the literature. Therefore, the analysis of the linear case
is nott Theopehe subject of this paper. Instead, we assume that the linear subspace iteration (1.14) converges with a certain contraction number ζ, i.e., $X_\kappa, p_\kappa, r_\kappa$, and ω are chosen such that

$$||M_\omega|| \leq \zeta < 1. \tag{1.15}$$

The main purpose of this chapter is to show that a similar rate locally holds for the nonlinear version (1.13). In particular, for (1.13) we use the same ω as involved in (1.15).

1.5 Quantitative analysis of the asymptotic convergence of the nonlinear iteration.

Let Φ be the nonlinear subspace iteration described in §1.3 with ω from §1.4. Moreover, we assume that the estimate (1.15) with the rate $\zeta < 1$ holds for the linear case. In the nonlinear case the norm of the iteration matrix is replaced by the contraction number with respect to the norm $||.||$. We estimate this in the following theorem.

THEOREM 1.2. *Assume* (1.1*a–g*), (1.7), (1.15) *and let ζ' be any value in the interval $(\zeta, 1)$ with ζ from* (1.15). *Then there is a neighbourhood U' of x^* so that*

$$||\Phi(x') - \Phi(x'')|| \leq \zeta' < 1 \qquad \textit{for all } x', x'' \in U', \tag{1.16}$$

i.e., the nonlinear subspace iteration Φ converges in U' and has the same asymptotic convergence rate ζ as the linear iteration applied to the linearized equations.

PROOF. a) The definition yields

$$\Phi(x') - \Phi(x'') = \left[x' - \omega \sum_{\kappa \in I} p_\kappa \delta_\kappa(x')\right] - \left[x'' - \omega \sum_{\kappa \in I} p_\kappa \delta_\kappa(x'')\right].$$

Using (1.10) for the local corrections $\delta_\kappa(x')$ and $\delta_\kappa(x'')$, we obtain

$$\Phi(x') - \Phi(x'') = \left[I - \omega \sum_{\kappa \in I} p_\kappa D\delta_\kappa(x', x'')\right](x' - x'').$$

Due to Remark 1.5, we have $\sum_{\kappa \in I} p_\kappa D\delta_\kappa(x', x'') \to \sum_{\kappa \in I} p_\kappa A_\kappa^{-1} r_\kappa F'(x^*)$ with $A_\kappa := r_\kappa F'(x^*) p_\kappa$ as $x', x'' \to x^*$. A comparison with (1.14) shows that

$$\left[I - \omega \sum_{\kappa \in I} p_\kappa D\delta_\kappa(x', x'')\right] \to M_\omega \quad \text{as } x', x'' \to x^*.$$

Therefore, (1.15) implies $\left\|[I - \omega \sum_{\kappa \in I} p_\kappa D\delta_\kappa(x', x'')]\right\| \le \zeta'$ in a neighbourhood of x^*.

b) Since ζ' can be arbitrarily close to ζ provided that the neighbourhood U' is small enough, the contraction number of Φ approaches ζ as the iteration converges to x^*. This proves that ζ is the asymptotic rate. □

2 Quasilinear elliptic problems.

In this section we illustrate the nonlinear subspace method described in §1 for quasilinear elliptic problems subject to Dirichlet boundary conditions.

The problem discussed here can also be used to illustrate the results of §3 concerning an abstract theory of a nonlinear subspace method for the minimization problem.

2.1 The differential problem.

Find $u^* \in H_0^1(\Omega)$ such that

$$b(u^*, v) = f(v) \quad \text{for all} \quad v \in H_0^1(\Omega), \tag{2.1}$$

where

$$b(u,v) = \int_\Omega \left(\sum_{i=1}^{2} a_i(x,u,\nabla u)\frac{\partial v}{\partial x_i} + a_0(x,u,\nabla u)v \right) dx$$

and

$$f(v) = \int_\Omega f v dx.$$

Here Ω is a bounded region in R^2 with Lipschitz continuous boundary.

The problem (2.1) has a unique solution under certain assumptions on the coefficients $\{a_i\}$, see Ladyzhenskaya and Ural'tseva [9]. For simplicity of presentation we consider here nonlinear problems with bounded nonlinearity (quasilinear problems). This means that the coefficients $a_i(x, p_0, p_1, p_2) = a_i(x, u, u_x, u_y)$ where $p = (p_0, p_1, p_2)$, $\partial u/\partial x_1 = u_x$, $\partial u/\partial x_2 = \partial u/\partial y = u_y$, $p' = (p_1, p_2)$ satisfy the following conditions with some constants c and C :

$$a_i \in C^1(\Omega \times R^3), \tag{A.1}$$

$$\max\left\{ |a_i| \left|\frac{\partial a_i}{\partial x_j}\right|, \left|\frac{\partial a_i}{\partial p_k}\right| \right\} \le C, \tag{A.2}$$

for $i,k = 0,1,2$, and $j = 1,2$.

In addition the operator is assumed to be strongly elliptic, i.e.

$$\sum_{i,j=0}^{2} \frac{\partial a_i(x,p)}{\partial p_j}\xi_i\xi_j \ge c\sum_{i=0}^{2}\xi_i^2 \tag{A.3}$$

for any $\xi = (\xi_0, \xi_1, \xi_2) \in R^3, \xi \neq 0$. Under these assumptions the problem (2.1) is well posed in $H_0^1(\Omega)$, see [9].

As a direct consequence of the assumptions $(A.1$–$3)$ we have the following statements which are used for an analysis of the discrete problem considered below.

The bilinear form $b(\cdot,\cdot) : H_0^1(\Omega) \times H_0^1(\Omega) \to R$ satisfies the following conditions: there exist positive constants c and C such that

$$b(u, u-v) - b(v, u-v) \geq c||u-v||_{H^1}^2 \tag{2.2}$$

for any $u, v \in H_0^1(\Omega)$, and

$$|b(u,w) - b(v,w)| \leq C||u-v||_{H^1}||w||_{H^1} \tag{2.3}$$

for any $u, v, w \in H_0^1(\Omega)$.

2.2 The discrete problem.

We discretize the problem (2.1) by the finite element method with triangular elements and continuous piecewise linear functions only (for simplicity of presentation). We assume that Ω is a polygonal region which is first divided into nonoverlapping triangular substructures Ω_i which form a coarse triangulation with a parameter H. A fine triangulation of Ω, with also triangular elements e_j and a parameter h, is obtained as refinement (several times) of the coarse triangulation. We assume that the both triangulations are quasi-uniform in the sense of finite element theory, see Ciarlet [3]. Let V_h be a finite element space defined on the fine triangulation with piecewise linear continuous functions which vanish on $\partial\Omega$, the boundary of Ω.

The discrete problem is of the following form. Find $u_h^* \in V_h$ such that

$$b(u_h^*, v_h) = f(v_h) \text{ for all } v \in V_h. \tag{2.4}$$

It follows from the properties of the bilinear form $b(\cdot,\cdot)$ (see (2.2) and (2.3)) that this problem has a unique solution. The error estimate

$$||u_h^* - u^*||_{H^1(\Omega)} \leq Ch|u^*|_{H^2(\Omega)}$$

is known provided that $u^* \in H^2(\Omega)$, see [3].

Let

$$V_h = \text{span}\{\phi_1, \ldots, \phi_n\},$$

where ϕ_i are the standard nodal basis functions and $u_h^* = \sum u_i\phi_i, u_i = u_h(x_i)$, and x_i are the nodal points. Let

$$b_i(u_1, \ldots, u_n) = b\left(\sum_{j=1}^n u_j\phi_j, \phi_j\right), \quad f_i = f(\phi_i)$$

and

$$B = (b_1, \ldots, b_n)^T, \quad \tilde{f} = (f_1, \ldots, f_n)^T.$$

Using the above notation the problem (2.4) reduces to a nonlinear problem (cf. (1.1a)) of the form

$$F(u_h) \equiv B(u_h) - \tilde{f} = 0. \tag{2.5}$$

Here we have used an isomorphism between V_h and R^n defined as follows. Any function $u_h \in V_h$ is uniquely represented by a vector, denoted also by $u_h \in R^n$ with coordinates equal to the values of $u_h \in V_h$ at interior nodal points of Ω, and vice versa. Thus $X = R^n$ and $D = U = R^n$ and, of course, the system (2.5) has a unique solution as equivalent to the problem (2.4).

The rest of this section is devoted to the solution of (2.5) using the nonlinear subspace iteration method described in §1 in terms of an additive Schwarz method, see Dryja and Widlund [4], Cai and Widlund [2], and Hackbusch [7].

2.3 The additive Schwarz method.

To obtain the additive Schwarz method as the nonlinear subspace iteration method we extend each Ω_i to a larger region $\Omega_i^{'}$, i.e., $\Omega_i \subset \Omega_i^{'}$. We assume that the overlap satisfies the following conditions: i) $\partial\Omega_i^{'}$ does not cut the elements e_j and ii) $\delta_i = dist(\partial\Omega_i, \partial\Omega_i^{'})$ is bounded from below by $H\alpha$ with $\alpha > 0$ which is independent of h (for more general case see Dryja and Widlund [5]).

The subspaces $V_\kappa, \kappa \in I = \{0, \ldots, N\}$ are defined as follows. For $\kappa = 0$, $V_0 = V_H$, a finite element space with continuous piecewise linear basis functions defined on the coarse triangulation.

For $\kappa = 1, \ldots, N$, $V_\kappa(\Omega_\kappa^{'})$ are the restriction of V_h to $\Omega_\kappa^{'}$ with zero values on $\partial\Omega_\kappa^{'}$, and $V_\kappa(\Omega)$ is an extension of $V_\kappa(\Omega_\kappa^{'})$ with zero outside $\Omega_\kappa^{'}$. The vector representation of $V_\kappa(\Omega_\kappa^{'})$ is denoted by $X_\kappa = R^{n_\kappa}$. Let $p_\kappa, \kappa = 1, \ldots, N$, be the standard prolongation operator from $V_\kappa(\Omega_\kappa^{'})$ into $V_\kappa \subset V_h$ by the extension by zero outside $\Omega_\kappa^{'}$ and $p_\kappa = r_\kappa^T$ be its matrix representation. For $\kappa = 0$, $p_0 = r_0^T$ is the standard interpolation operator (matrix form) from $V_0 = V_H$ into V_h.

The method for solving (2.5) is defined as

$$u_h^{k+1} = u_h^k - \omega\delta(u_h^k), \tag{2.6}$$

where

$$\delta(u_h^k) = \sum_{\kappa \in I} p_\kappa \delta_\kappa$$

and δ_κ is the solution of (cf.(1.6))

$$r_\kappa F(u_h^k - p_\kappa\delta_\kappa) = 0. \tag{2.7}$$

We now analyse the method (2.6). The first question is a uniqueness and existence result for the solution of (2.7). In view of Theorem 1.1 we should check the assumptions (1.1a–g) and (1.7). Under the assumptions (A.1–3) it can be shown that $F^{'}(u_h)$ exists for any $u_h \in V_h$ and it satisfies the conditions (1.1e–g). It is easy to see that the condition (1.7) is also satisfied. Thus the problems (2.7) are uniquely solvable for any $u_h^k \in V_h$, see Theorem 1.1.

To prove convergence of (2.6) by Theorem 1.2 we need to analyse the additive Schwarz method for the linear systems with matrix $A = F'(u_h^*)$, i.e.,

$$Au_h = b. \tag{2.8}$$

In the case of a symmetric matrix A the additive Schwarz method is analysed in [4], see also [7]. The symmetry of A needs additional assumptions on the coefficients $a_i(x,p)$, i.e.,

$$\frac{\partial a_i(x,p)}{\partial p_j} = \frac{\partial a_j(x,p)}{\partial p_i}, \quad i,j = 0,1,2. \tag{2.9}$$

A convergence of the method without symmetry of A is done in [2]. From these papers follows that

$$M_\omega = I - \omega \sum_{\kappa \in I} p_\kappa A_\kappa^{-1} r_\kappa A \text{ with } A_\kappa = r_\kappa A p_\kappa$$

satisfies

$$||M_\omega|| \le \zeta < 1 \tag{2.10}$$

for the some damping parameter ω. As consequence of that and Theorem 1.2 we conclude that the method (2.6) for the nonlinear problem (2.5) is convergent and it has the same asymptotic contraction rate ζ as the linearized system with $A = F'(u_h^*)$, i.e., geometrical with ζ independent of H, h and the number of substructures.

The implementation of the method (2.6) in each iteration reduces to the solution of $N+1$ subproblems which are independent, so they can be solved in parallel. For details see the papers mentioned above.

3 The case of minimization.

In §1 we presented an abstract nonlinear system and derived local nonlinear problems. In this section, we assume that the given nonlinear problem is equivalent to a minimization problem. Then the local equations turn out to be minimization problems, too. Next, we discuss global convergence. For this purpose we use conditions which can be obtained from standard convexity assumptions.

3.1 Reformulation of the problem.

Assume that a function

$$G : D \subset X \to R \tag{3.1}$$

is given and that the minimization problem

$$G(x^*) = \min_{x \in D} G(x) \tag{3.2}$$

is to be solved.

In the following we make assumptions which link the new problem to the equation from §1:

(3.3a) G has a minimum at x^*,

(3.3b) x^* is a unique solution of (3.2) in D

(3.3c) G is differentiable and $F(x) := G'(x)$ satisfies (1.1a–g).

The next theorem ensures that also the local equations (1.6) are minimization problems.

THEOREM 3.1. *Assume* (1.5). *Then the subspace problem* (1.6) *is equivalent to the following minimization problem in* X_κ*:*

$$\text{(3.4)} \qquad \textit{Find } \delta_\kappa \in X_\kappa \textit{ with} \qquad G(\tilde{x} - p_\kappa \delta_\kappa) = \min_{\delta' \in X_\kappa} G(\tilde{x} - p_\kappa \delta').$$

PROOF. The gradient of $g(\delta) := G(\tilde{x} - p_\kappa \delta)$ for $\delta \in X_\kappa$ is g' with

$$\langle g'(\delta), x\rangle = \langle -F'(\tilde{x} - p_\kappa\delta), p_\kappa x\rangle = -\langle r_\kappa F'(\tilde{x} - p_\kappa \delta), x\rangle \quad \text{for all } x \in X_\kappa.$$

Therefore, $g' = 0$ is equivalent to (1.6). □

3.2 Global solvability.

In the following, the domain of definition D is the whole linear space X. $F(x) = G'(x)$ is assumed to be continuous in X. In order to exclude stationary points of G except x^*, we require

$$\text{(3.5)} \qquad ||F(x)|| \geq \varepsilon(x) > 0 \qquad \text{for all } x \in X \backslash U,$$

where $U \subset D$ is an (open) neighbourhood. The function $\varepsilon(x)$ is required to be continuous on $X\backslash U$. Condition (3.5) ensures that outside U the function F is bounded away from zero.

Condition (3.3b) may be restricted to U (instead of D). Together with (3.5) we conclude that x^* is the unique solution of (3.2) in the whole space X.

For a real number $\rho \in R$, the level set L_ρ is defined by

$$\text{(3.6}a\text{)} \qquad L_\rho := \{x \in X : G(x) \leq \rho\}.$$

We require that

$$\text{(3.6}b\text{)} \qquad L_\rho \text{ is compact for all } \rho \in R.$$

This assumption is satisfied if $G(x)$ is unbounded for $||x|| \to \infty$.

3.3 Global convergence.

For $x \in X\backslash U$, we replace the subspace minimization problem (3.4) by the simpler gradient method:

Let $\tilde{x} \in X\backslash U$. Find $\delta_\kappa \in X_\kappa$ with $||\delta_\kappa|| = 1$ such that

$$\text{(3.7)} \qquad \langle F(\tilde{x}), p_\kappa \delta_\kappa\rangle = \max\{\langle F(\tilde{x}), p_\kappa e_\kappa\rangle : ||e_\kappa|| = 1, e_\kappa \in X_\kappa\}.$$

Assume (1.5). A solution of (3.7) can be described explicitly:

$$(3.7') \qquad \delta_\kappa = r_\kappa F(\tilde{x})/||r_\kappa F(\tilde{x})|| \quad \text{if} \quad r_\kappa F(\tilde{x}) \neq 0;$$

otherwise, any unit vector δ_κ can be chosen.

LEMMA 3.2. *Assume* (1.4), (1.5), (3.3*a*–*c*), (3.5), (3.6*a*, *b*), *define* δ_κ *by* (3.7′) *and define* $\delta := \sum_{\kappa\in I} p_\kappa \delta_\kappa$. *Then for all* ρ *there is some* $\varepsilon' > 0$ *such that*

$$(3.8) \qquad \langle F(\tilde{x}), \delta\rangle \geq \varepsilon' \qquad \textit{holds for all } \tilde{x} \in L_\rho.$$

Inequality (3.8) *states that* $-\delta$ *is a descent direction. This means that* $g(\lambda) := G(\tilde{x} - \lambda\delta)$ *has a negative derivative at* $\lambda = 0$ *which is bounded away from zero, i.e.,* $g'(0) \leq -\varepsilon' < 0$.

PROOF. Because of (1.4) and (1.5) (i.e., $X = \sum p_\kappa X_\kappa$ and $r_\kappa := p_\kappa^T$), the intersection of the kernels of r_κ is the zero space. Therefore, $||x||$ and $\sum_{\kappa\in I} ||r_\kappa x||$ are equivalent norms, in particular, there is a constant C with

$$(3.9) \qquad ||x|| \leq C \sum_{\kappa\in I} ||r_\kappa x|| \qquad \text{for all } x \in X.$$

This proves that

$$\langle F(\tilde{x}), \delta\rangle = \sum_{\kappa\in I} \langle F(\tilde{x}), p_\kappa\delta_\kappa\rangle = \sum_{\kappa\in I} \langle r_\kappa F(\tilde{x}), \delta_\kappa\rangle \underset{(3.7')}{=} \sum_{\kappa\in I} ||r_\kappa F(\tilde{x})|| \geq ||F(\tilde{x})||/C.$$

By (3.6*b*), $L_\rho \setminus U$ is also compact and it follows that $\varepsilon := \min\{\varepsilon(x) : x \in L_\rho\}$ is positive and thus $||F(\tilde{x})|| \geq \varepsilon > 0$. Combining this with the previous inequality shows (3.8) with $\varepsilon' := \varepsilon/C$. □

REMARK 3.1. If (3.7) is replaced by the subspace minimization

$$(3.10) \qquad G(\tilde{x} - \lambda p_\kappa \delta_\kappa) = \min_{\delta' \in X_\kappa} G(\tilde{x} - p_\kappa \delta') \qquad \text{with } ||\delta_\kappa|| = 1,$$

it is not obvious that $-p_\kappa\delta_\kappa$ is a local descent direction at $\tilde{x}$, i.e., $\langle F(\tilde{x}), p_\kappa\delta_\kappa\rangle > 0$ (*or* $\geq \varepsilon/C$).

In (3.10) we have introduced the factor λ to ensure $||\delta_\kappa|| = 1$ (cf. (3.7)). If $\delta'_\kappa \in X_\kappa$ is the minimizer of the right-hand side in (3.10), we have $\lambda := ||\delta'_\kappa||$.

Since F is uniformly continuous on L_ρ, there is some $\lambda_0 > 0$ such that (3.8) implies

$$(3.11) \qquad \langle F(\tilde{x} - \lambda\delta), \delta\rangle \geq \varepsilon'/2 \qquad \text{for all } 0 \leq \lambda \leq \lambda_0$$

with δ from (3.8).

Below, we propose a globally convergent algorithm each step of which consists of a local minimization problem, provided the iterate is in the neighbourhood U of the solution. Because of Remark 3.1, we replace the minimization outside U by a gradient-like method.

ALGORITHM 3.1. GLOBAL ALGORITHM

(3.12a) $\tilde{x} := x^i$ (ith iterate).

(3.12b) If $\tilde{x} \in U$, proceed as in §1 (where (1.6) may be reformulated by the equivalent minimization problem (3.4)), form $\delta := \sum_{\kappa \in I} p_\kappa \delta_\kappa$ and set $x^{i+1} := \tilde{x} - \delta$.

(3.12c) If $\tilde{x} \in X \backslash U$, compute $\delta := \sum_{\kappa \in I} p_\kappa \delta_\kappa$ as in Lemma 3.2, define the next iterate x^{i+1} by either of the following possibilities:

(3.12c_1) $\qquad x^{i+1} := \tilde{x} - \lambda_0 \delta \qquad (\lambda_0$ as in (3.11)),

(3.12c_2) $\qquad x^{i+1} := \tilde{x} - \lambda \delta$
with λ being the minimizer of $G(\tilde{x} - \lambda\delta)$ over $\lambda \geq 0$.

Since, in general, the neighbourhood U is unknown in practise, one may try to perform (3.12b). If it does not show the expected convergence speed, one switches to (3.12c). Since $G(\tilde{x} - \lambda\delta)$ is decreasing for δ small enough, it is relatively easy to find a rough approximation λ for (3.12c_2).

REMARK 3.2. Consider the case of $\tilde{x} \in X \backslash U$. In both of the cases (3.12$c_{1,2}$), G is reduced by a fixed amount:

$$G(x^{i+1}) \leq G(x^i) - \lambda_0 \varepsilon' / 2. \tag{3.13}$$

PROOF. In the case of (3.12c_1), the estimate (3.13) follows from (3.11). In the other case, the minimum value from (3.12c_2) can only be smaller than in (3.12c_1). □

THEOREM 3.3. *Assume* (1.4), (1.5), (1.15), (3.3a–c), (3.5), (3.6a, b). *Then, for any starting value* x^0 *the algorithm* (3.12) *produces a sequence* $\{x^i\}$ *converging to* x^*. *The asymptotic convergence behaviour is as described in* §1.

PROOF. Given the starting value x^0, define the value $\rho := G(x^0)$ appearing in (3.8). As long as the iterates are outside U, G decreases as described in (3.13). Therefore the iterates x^i stay in the level set L_ρ. The decrease by $\lambda^0 \varepsilon' / 2$ can happen only finitely many times. Afterwards the iterates are in U and the results of §1 are valid. In particular, they describe the asymptotic convergence behaviour. □

3.4 Convexity assumptions.

In order to justify the condition (3.5), we discuss the standard coercivity assumption for the nonlinear problem. This is the inequality

$$\langle F(x') - F(x''), x' - x'' \rangle \geq \gamma \| x' - x'' \|^\alpha \quad \text{with } \gamma > 0 \text{ for } x', x'' \in X, \tag{3.14}$$

e.g., for $\alpha > 1$. Obviously, (3.14) implies the uniqueness required in (3.3a, b).

REMARK 3.3. Condition (3.14) implies inequality (3.5) with

$$\varepsilon(x) = \gamma \|x' - x''\|^{\alpha-1} \quad \forall \ x \in X \backslash U.$$

PROOF. Since $F(x^*) = 0$, we have $\langle F(x), x - x^* \rangle \geq \gamma \|x - x^*\|^\alpha$. Therefore, for $x \in X \backslash U$, we conclude that

$$\|F(x)\| \geq |\langle F(x), x - x^* \rangle| / \|x - x^*\| \geq \gamma \|x - x^*\|^{\alpha-1} =: \varepsilon(x). \qquad \square$$

REMARK 3.4. Inequality (3.14) with $\alpha > 0$ implies (3.6*a*, *b*).

In the case of (3.14) with $\alpha > 0$, it is not necessary to use different definitions of the correction δ as in (3.12*b*) and (3.12*c*). We may change Algorithm 3.1 into

ALGORITHM 3.2. GLOBAL ALGORITHM

(3.15*a*) Set $\tilde{x} := x^i$ (*i*th iterate).

(3.15*b*) Compute δ as in §1 (where (1.6) may be replaced by the equivalent minimization problem (3.4)).

(3.15*c*) If $\tilde{x} \in U$, set $x^{i+1} := \tilde{x} - \delta$. If $\tilde{x} \in X \backslash U$, define the next iterate by either of the following possibilities:

$$(3.15d_1) \qquad x^{i+1} := \tilde{x} - \lambda_0 \delta$$

$$(3.15d_1) \qquad x^{i+1} := \tilde{x} - \lambda \delta \qquad (\lambda\text{: minimizer of } G(\tilde{x} - \lambda\delta) \text{ over } \lambda \geq 0).$$

The choice of λ_0 is discussed in the proof of the next theorem.

THEOREM 3.4. *Assume* (1.4), (1.5), (1.15), (3.3*c*), *and* (3.14) *with* $\alpha > 1$. *For sufficiently small* λ_0, *Algorithm* 3.2 *has the properties stated in Theorem* 3.3.

PROOF. a) For each κ let $\delta_\kappa \in X_\kappa$ be the solutions of (1.6) (equivalently (3.4)) and set $\delta := \sum_{\kappa \in I} p_\kappa \delta_\kappa$ and $d := (\sum_{\kappa \in I} \|p_\kappa \delta_\kappa\|^2)^{1/2}$. In parts b) and c) below we show

$$(3.16a) \qquad \|\tilde{x} - x^*\| \leq C_1 d^{1/(\alpha-1)},$$

$$(3.16b) \qquad \langle F(\tilde{x}), \delta \rangle \geq C_2 d^\alpha.$$

Since U is a neighbourhood of $x^*, \eta := \inf\{\rho \geq 0 \ : \ \{\|x^* - x\| \leq \rho\} \in U\}$ is positive. Then if $d \leq (\eta/C_1)^{\alpha-1}$, the iterate $\tilde{x}$ belongs to U and the iteration is already in the asymptotic region. Otherwise, if $d \geq (\eta/C_1)^{\alpha-1}$, the inequality (3.8) holds with $\varepsilon' := C_2(\eta/C_1)^{\alpha/(\alpha-1)}$. Hence, λ_0 can be chosen such that (3.11) is valid. Then the global convergence follows as in the proof of Theorem 3.3.

b) There are $\tilde{\delta}_\kappa \in X_\kappa$ with $\tilde{x} - x^* = \sum p_\kappa \tilde{\delta}_\kappa$ and $\sum_{\kappa \in I} ||p_\kappa \tilde{\delta}_\kappa||^2 \leq (C' ||\tilde{x} - x^*||)^2$. Let C'' be the bound of DF. Then,

$$||F(\tilde{x}) - F(\tilde{x} - p_\kappa \delta_\kappa)|| \leq C'' ||p_\kappa \delta_\kappa||.$$

Since δ_κ minimizes $G(\tilde{x} - p_\kappa \delta_\kappa)$,

$$\langle F(\tilde{x} - p_\kappa \delta_\kappa), p_\kappa x_\kappa \rangle = 0 \qquad \text{for all } x_\kappa \in X_\kappa. \tag{3.17}$$

Taking $x_\kappa = \tilde{\delta}_\kappa$, we get

$$\langle F(\tilde{x}), p_\kappa \tilde{\delta}_\kappa \rangle = \langle F(\tilde{x}) - F(\tilde{x} - p_\kappa \delta_\kappa), p_\kappa \tilde{\delta}_\kappa \rangle \leq C'' ||p_\kappa \delta_\kappa|| \, ||p_\kappa \tilde{\delta}_\kappa||.$$

Summation over κ yields

$$\begin{aligned} \langle F(\tilde{x}), \tilde{x} - x^* \rangle = \sum_\kappa \langle F(\tilde{x}), p_\kappa \tilde{\delta}_\kappa \rangle &\leq C'' \sum_\kappa ||p_\kappa \delta_\kappa|| \, ||p_\kappa \tilde{\delta}_\kappa|| \\ &\leq C'' d C' ||\tilde{x} - x^*||. \end{aligned}$$

Together with

$$\langle F(\tilde{x}), \tilde{x} - x^* \rangle = \langle F(\tilde{x}) - F(x^*), \tilde{x} - x^* \rangle \geq \gamma ||\tilde{x} - x^*||^\alpha,$$

we obtain $||\tilde{x} - x^*|| \leq (C' C'' d/\gamma)^{1/(\alpha-1)}$. Hence, (3.16a) holds with $C_1 := (C' C''/\gamma)^{1/(\alpha-1)}$.

c) For the proof of (3.16*b*) use (3.17) with $x_\kappa = \delta_\kappa$:

$$\langle F(\tilde{x}), p_\kappa \delta_\kappa \rangle = \langle F(\tilde{x}) - F(\tilde{x} - p_\kappa \delta_\kappa), p_\kappa \delta_\kappa \rangle \underset{(3.14)}{\geq} \gamma ||p_\kappa \delta_\kappa||^\alpha.$$

Summation over κ yields

$$\langle F(\tilde{x}), \delta \rangle \geq \gamma \sum_\kappa ||p_\kappa \delta_\kappa||^\alpha \geq \gamma C(\alpha) \left(\sum_\kappa ||p_\kappa \delta_\kappa||^2 \right)^{\alpha/2} = C_2 d^\alpha$$

with $C_2 := \gamma C(\alpha)$. □

REFERENCES

1. X.-C. Cai and M. Dryja, *Domain decomposition methods for monotone nonlinear elliptic problems*, Contemporary Mathematics, v. 180, pp. 21–27, D. Keyes and J. Xu, eds, AMS, 1994.

2. X.-C. Cai and O. Widlund, *Multiplicative Schwarz algorithms for nonsymmetric and indefinite elliptic problems*, SIAM J. Numer. Anal., 30 (1993), pp. 936–952.

3. P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, New York, 1978.

4. M. Dryja and O. Widlund, *An additive variant of the Schwarz alternating method for the case of many subregions*, Tech. Report 339, Courant Institute, New York, 1987.

5. M. Dryja and O. Widlund, *Domain decomposition algorithms with small overlap*, SIAM J. Sci. Comp., 15 (1994), pp. 604–620.
6. W. Hackbusch, *Multi-grid Methods and Applications*, Springer Verlag, Berlin, 1985.
7. W. Hackbusch, *Iterative Solution of Large Sparse Systems*, Springer Verlag, Berlin, 1994.
8. W. Hackbusch and A. Reusken, *Analysis of a damped nonlinear multilevel method*, Numer. Math., 55 (1989), pp. 225–246.
9. O. A. Ladyzhenskaya and N. N. Ural'tseva, *Linear and Quasilinear Elliptic Equations*, Academic Press, 1968.
10. Xue-Cheng Tai, *Parallel function decomposition and space decomposition methods*, Beijing Mathematics 1 (1995), pp. 104–152.