

A globally convergent version of the method of moving asymptotes

C. Zillober

Mathematisches Institut, Universität Bayreuth, D-95440 Bayreuth, Germany

Abstract The method of moving asymptotes (MMA) which is known to work excellently for solving structural optimization problems has one main disadvantage: convergence cannot be guaranteed and in practical use this fact sometimes leads to unsatisfactory results. In this paper we prove a global convergence theorem for a new method which consists iteratively of the solution of the known MMA-subproblem and a line search performed afterwards.

1 Introduction

In the last few years, the concept of convex approximations has caused more and more interest in structural optimization. A structural optimization problem written in the form

$$\min_{\mathbf{x} \in X} f(\mathbf{x}) \quad (\mathbf{x} \in \mathbb{R}^n), \quad \text{s.t.} \quad h_j(\mathbf{x}) \leq 0, \quad j = 1 \dots M, \quad (1)$$

where $X := \{\mathbf{x} \mid \underline{x}_i \leq x_i \leq \bar{x}_i, i = 1 \dots n\}$ is replaced by a sequence of easier to solve, convex, separable subproblems which approximate the original problem. The functions f and h_j ($j = 1 \dots M$) are assumed to be continuously differentiable and the feasible region is assumed to be non-empty.

The most general of these methods, the Method of Moving Asymptotes (MMA), is nowadays implemented in many software systems (see e.g. Hörnlein and Schittkowski 1992). The experience of users is that if MMA converges, then it approximates a solution rapidly by a sequence of steadily improved designs, but the major drawback is that it diverges in quite a few of applications.

This was the reason for the investigations which resulted in this paper. By adding a line search subject to a function measuring the global convergence, the behaviour of the method can be stabilized without losing the known advantages.

In the following section, we will describe the optimization methods CONLIN and MMA and outline some of their main features. In Section 3 we will explain the new method and formulate the most important results including a proof of global convergence. Section 4 contains one possibility to overcome difficulties concerning the solvability of subproblems. The numerical behaviour is illustrated by some examples in the final section.

2 CONLIN and MMA

Using the idea of using reciprocal variables, Fleury and Braibant (1986) developed the optimization method CONLIN (convex linearization). An approximation of a function

is defined by separate linearization for each component depending on the sign of the partial derivative at the expansion point. If the sign is positive, then the linearization is performed with respect to the original variable, if the sign is negative, then it is subject to the inverse variable, leading to a convex approximation of the original function.

The Method of Moving Asymptotes (MMA), however, is a generalization of CONLIN. Svanberg (1987) proposed a linearization with respect to substituted variables

$$\frac{1}{U_i - x_i} \quad \text{and} \quad \frac{1}{x_i - L_i},$$

respectively, where U_i and L_i are some chosen parameters.

Definition 2.1. Let g be a continuously differentiable function on X . An MMA approximation \tilde{g} of g is defined by

$$\tilde{g}(\mathbf{x}) = g(\mathbf{x}^0) + \sum_{+} \frac{\partial g}{\partial x_i} \Big|_{\mathbf{x}^0} \left[\frac{(U_i - x_i^0)^2}{U_i - x_i} - (U_i - x_i^0) \right] - \sum_{-} \frac{\partial g}{\partial x_i} \Big|_{\mathbf{x}^0} \left[\frac{(x_i^0 - L_i)^2}{x_i - L_i} - (x_i^0 - L_i) \right],$$

where \sum_{+} (\sum_{-}) means summation over all components

i where the partial derivative $\frac{\partial g}{\partial x_i}$ at the expansion point \mathbf{x}^0 is non-negative (negative); \tilde{g} is defined on $D_{\tilde{g}} := \{\mathbf{x} \mid \max(L_i, \underline{x}_i) < x_i < \min(U_i, \bar{x}_i), i = 1 \dots n\}$.

It is easy to verify that \tilde{g} is a first-order approximation of g , i.e.

$$\tilde{g}(\mathbf{x}^0) = g(\mathbf{x}^0) \quad \text{and} \quad \nabla \tilde{g}(\mathbf{x}^0) = \nabla g(\mathbf{x}^0),$$

and \tilde{g} is convex and separable, where g stands for the objective or constraint function, respectively.

Remarks

- The CONLIN method is obtained by letting $L_i = 0$ and $U_i \rightarrow \infty$ ($i = 1 \dots n$), and
- L_i and U_i are asymptotes for \tilde{g} .

The algorithm proposed by Svanberg (1987) can be outlined as follows.

Step 0: Choose a starting point \mathbf{x}^0 , let $k := 0$.

Step 1: Compute $f(\mathbf{x}^k)$, $h_j(\mathbf{x}^k)$, $\nabla f(\mathbf{x}^k)$, $\nabla h_j(\mathbf{x}^k)$, $j = 1 \dots M$.

Step 2: Define a subproblem, replacing f , h_j by \tilde{f} , \tilde{h}_j , $j = 1 \dots M$, according to (2.1).

Step 3: Solve this subproblem and let its solution be denoted by \mathbf{x}^{k+1} . Let $k := k + 1$ and goto Step 1.

We neglect here certain details, for example, the choice of the asymptotes, which are in general updated in each iteration. When applying this algorithm one sometimes may observe difficulties concerning the objective function. It is possible that the approximation is linear or almost linear leading to useless subproblems. Therefore, Svanberg (1993) suggested to append a term to the approximation of the objective function which guarantees strict convexity of the objective function. We use here the following mode, where we first rewrite the original approximation,

$$\tilde{f}(\mathbf{x}) = \bar{f} + \sum_{+} \frac{c_i}{U_i - x_i} - \sum_{-} \frac{c_i}{x_i - L_i}$$

Definition 2.2. For the (continuously differentiable) objective function of (1) we use the approximation

$$\tilde{f}(\mathbf{x}) = \bar{f} + \sum_{+} \left[\frac{c_i}{U_i - x_i} + \frac{\epsilon(x_i - x_i^0)^2}{U_i - x_i} \right] - \sum_{-} \left[\frac{c_i}{x_i - L_i} - \frac{\epsilon(x_i - x_i^0)^2}{x_i - L_i} \right],$$

where ϵ is some positive number.

This approximation remains separable and is now strictly convex with two major advantages.

- The subproblem has a unique solution (if it has at least one).
- A very efficient dual approach for solving the subproblem is applicable (cf. Fleury 1989; Svanberg 1987; Zillober 1992).

To simplify the notation, we shall use the following abbreviation.

Definition 2.3. By $SP(\mathbf{x}^k)$ (subproblem) we state the optimization problem

$$\min \tilde{f}(\mathbf{x}) \quad \text{subject to} \quad \tilde{h}_j(\mathbf{x}) \leq 0, \quad j = 1 \dots M, \quad \mathbf{x} \in X',$$

where \mathbf{x}^k is the expansion point and the functions are chosen according to (2.1) and (2.2) and $X' := \{\mathbf{x} \mid \max[\underline{x}_i, x_i^k - \omega(x_i^k - L_i)] \leq x_i \leq \min[\bar{x}_i, x_i^k + \omega(U_i - x_i^k)]\}$, ($\omega \in]0, 1[$ fixed).

We note that always $X' \subseteq X$ and all $\mathbf{x} \in X'$ are bounded away from the asymptotes. Next, we state a result which is very important in identifying a solution.

Lemma 2.4. \mathbf{x}^* is a stationary point of (1) if and only if \mathbf{x}^* is a stationary point of $SP(\mathbf{x}^*)$.

Proof: Zillober (1992).

Of course the choice of asymptotes is a crucial point for the behaviour of the algorithms discussed in this paper, but it is not our central scope. For theoretical purposes we now restrict the set of possible choices.

Definition 2.5. A strategy for the choice of asymptotes is called continuous, if for any sequence $\{\mathbf{x}^k\} \rightarrow \mathbf{x}$ we have $L_i(\mathbf{x}^k) \rightarrow L_i(\mathbf{x})$ and $U_i(\mathbf{x}^k) \rightarrow U_i(\mathbf{x})$, respectively, for all $i = 1 \dots n$ and for all $\mathbf{x}, \mathbf{x}^k \in X$.

By the notation $L_i(\mathbf{x}^k)$, $U_i(\mathbf{x}^k)$ we mean the asymptotes resulting from the evaluation of the chosen strategy at the point \mathbf{x}^k . These asymptotes may depend on the current iteration point, additionally on previous iteration points as the strategy in Section 5, or may be independent of the iteration point (note that, for example, the CONLIN method results, if $L_i = 0$ and $U_i = +\infty$ are chosen for all $i = 1 \dots n$). In

the theorems stated later we will assume that the strategy for the choice of the asymptotes fulfills (2.5). The reason is that (together with other assumptions) the feasible regions of the subproblems then vary continuously with respect to the expansion point. Tests with different strategies showed that global convergence of the new method is observed by applying non-continuous strategies. In other words, the continuity of the strategy for the choice of the asymptotes is sufficient for global convergence, but not necessary.

For the CONLIN method, Nguyen *et al.* (1987) gave a convergence proof but only for the case that (1) consists of concave functions, which is of less practical interest. They indicated furthermore by some examples that a generalization to non-concave functions is not possible. Zillober (1992) showed that a similar convergence analysis could be applied to MMA. Using the flexibility of the asymptotes which influence the curvature of the approximations, it is possible to omit the concavity assumption. However, the resulting restriction on the choice of the asymptotes leads to a very slow numerical convergence of the algorithm so that the theoretical improvement does not result in a more efficient algorithm. This was the motivation to look for another way to prove convergence of the method without losing the good behaviour of the original method. These results are reported in the next section.

3 Sequential convex programming

It is well-known that by adding a line search with respect to a function measuring the global convergence of an algorithm, the behaviour of an optimization method is improved. Generally, such a line search needs additional evaluations of the original functions of (1). Therefore most people reject this idea to globalize MMA, since its numerical performance is excellent even without a stabilization. In this section, we prove that MMA, together with a line search subject to an augmented Lagrange function, leads to a globally convergent optimization method. The examples in Section 5 are chosen to illustrate this fact. However, first we rewrite (1) in order to obtain a simplified notation in this section,

$$\min f(\mathbf{x}) \quad \text{s.t.} \quad h_j(\mathbf{x}) \leq 0, \quad j = 1 \dots m, \quad (2)$$

where $m = M + 2n$, i.e. we also write the box constraints in the form $h_j(\mathbf{x}) \leq 0$. Now we introduce the augmented Lagrange function.

Definition 3.1. The augmented Lagrange function $\Phi_r : \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ associated to (2) is for a fixed parameter $r > 0$ defined by

$$\Phi_r(\mathbf{x}) = f(\mathbf{x}) +$$

$$\sum_{j=1}^m \begin{cases} u_j h_j(\mathbf{x}) + \frac{r}{2} h_j^2(\mathbf{x}), & \text{if } -\frac{u_j}{r} \leq h_j(\mathbf{x}) \\ -\frac{u_j^2}{2r}, & \text{otherwise} \end{cases}$$

This function is also used in the general purpose optimization method SQP (Sequential Quadratic Programming) by Schittkowski (1981) as a merit function and is known to work well for stabilizing the method.

The motivation to choose the augmented Lagrange merit function results from the following two well-known statements.

- (1) A point $(\mathbf{x}^*, \mathbf{u}^*)$ is stationary for Φ_r if and only if $(\mathbf{x}^*, \mathbf{u}^*)$ is stationary for (2).
 (2) Under some regularity conditions there is an $\bar{r} > 0$ such that \mathbf{x}^* is a global minimizer for $\bar{\Phi}_r(\mathbf{x}) := \Phi_r(\mathbf{x}, \mathbf{u}^*)$ for all $r \geq \bar{r}$.

To simplify the notation we use the following definitions.

Definition 3.2. Let $\mathbf{u} \in \mathbb{R}^m$ be a vector of multipliers and $\mathbf{x} \in X$ be an arbitrary point. Then

- $\bar{\mathbf{u}} := (\bar{u}_1, \dots, \bar{u}_m)^T$, where $\bar{u}_j = \begin{cases} u_j, & \text{if } -\frac{u_j}{r} \leq h_j(\mathbf{x}) \\ 0, & \text{otherwise} \end{cases}$,
- $\hat{h}(\mathbf{x}) := (h_1(\mathbf{x}), \dots, h_m(\mathbf{x}))^T$,
- $\bar{h}(\mathbf{x}) := (\bar{h}_1(\mathbf{x}), \dots, \bar{h}_m(\mathbf{x}))^T$, where $\bar{h}_j(\mathbf{x}) = \begin{cases} h_j(\mathbf{x}), & \text{if } -\frac{u_j}{r} \leq h_j(\mathbf{x}) \\ 0, & \text{otherwise} \end{cases}$,
- $\hat{h}(\mathbf{x}) := (\hat{h}_1(\mathbf{x}), \dots, \hat{h}_m(\mathbf{x}))^T$, where $\hat{h}_j(\mathbf{x}) = \begin{cases} h_j(\mathbf{x}), & \text{if } -\frac{u_j}{r} \leq h_j(\mathbf{x}) \\ -\frac{u_j}{r}, & \text{otherwise} \end{cases}$,
- $A(\mathbf{x}) := (\nabla h_1(\mathbf{x}), \dots, \nabla h_m(\mathbf{x}))^T$,
- $J := \{j \mid 1 \leq j \leq m; -\frac{u_j}{r} \leq h_j(\mathbf{x})\}$,
- $K := \{j \mid 1 \leq j \leq m; j \notin J\}$.

Hence, the gradient of the augmented Lagrange function is

$$\nabla \Phi_r(\mathbf{x}, \mathbf{u}) = \left\{ \begin{array}{c} \nabla f(\mathbf{x}) + A(\mathbf{x})[\bar{\mathbf{u}} + r\bar{h}(\mathbf{x})] \\ \hat{h}(\mathbf{x}) \end{array} \right\}.$$

Before we formulate the algorithm we state some auxiliary results.

Lemma 3.3. Let us consider $SP(\mathbf{x}^k)$ where \mathbf{x}^k is an arbitrary element of X . Then

- (a) $h_j(\mathbf{x})$ and $f(\mathbf{x})$ are uniformly bounded on X for all $j = 1 \dots m$;
 (b) $\nabla h_j(\mathbf{x})$ and $\nabla f(\mathbf{x})$ are uniformly bounded on X for all $j = 1 \dots m$;
 (c) $\bar{h}_j(\mathbf{x})$ and $\bar{f}(\mathbf{x})$ are uniformly bounded on X' for all $j = 1 \dots m$;
 (d) $\bar{\nabla} h_j(\mathbf{x})$ and $\bar{\nabla} \bar{f}(\mathbf{x})$ are uniformly bounded on X' for all $j = 1 \dots m$;

- (e) $\xi^k := \min \left\{ \min_{i=1 \dots n} \left[2\epsilon \frac{(U_i^k - x_i^k)^2}{(U_i^k - L_i^k)^3} \right], \min_{i=1 \dots n} \left[2\epsilon \frac{(x_i^k - L_i^k)^2}{(U_i^k - L_i^k)^3} \right] \right\} (> 0)$ is a lower bound such that $\frac{\partial^2 \bar{f}}{\partial x_i^2} \Big|_{\mathbf{x}} \geq \xi^k$ for all $\mathbf{x} \in D_{\bar{f}} (i = 1 \dots n)$.

- (f) There are $\bar{\xi}$ and $\xi > 0$ such that $\bar{\xi} \geq \xi^k \geq \xi$ for all k .

Corollary 3.4. Let \mathbf{y}^{k+1} be the optimal point of $SP(\mathbf{x}^k)$. Then

- (a) $\nabla f(\mathbf{x}^k)^T(\mathbf{x}^k - \mathbf{y}^{k+1}) \geq \bar{f}(\mathbf{x}^k) - \bar{f}(\mathbf{y}^{k+1}) + \eta^k \|\mathbf{y}^{k+1} - \mathbf{x}^k\|^2$, where $\eta^k := \frac{\xi^k}{2}$ and $\|\cdot\|$ denotes the Euclidean norm throughout.
 (b) If \mathbf{x}^k is feasible for (2) then $\nabla f(\mathbf{x}^k)^T(\mathbf{x}^k - \mathbf{y}^{k+1}) \geq \eta^k \|\mathbf{y}^{k+1} - \mathbf{x}^k\|^2$.

- (c) There are $\bar{\eta}$ and $\eta > 0$ such that $\bar{\eta} \geq \eta^k \geq \eta$ for all k .

Lemma 3.5. If $\mathbf{x} \in X$ and $0 \leq u_j \leq u_{\max}$ for all $j = 1 \dots m$, then $\Phi_r(\mathbf{x}, \mathbf{u})$ is uniformly bounded from below for all $r \geq 1$.

The proofs for these statements are straightforward and omitted here. The interested reader is referred to the doctoral dissertation by Zillober (1992).

Now we formulate the SCP (Sequential Convex Programming) algorithm in order to show the similarity to the SQP-method.

Step 0 : Choose $\mathbf{x}^0 \in X$, $\mathbf{u}^0 \geq 0$, $0 < c < 1$ (e.g. 0.001), $0 < \psi < 1$ (e.g. 0.5), $r > 0$ (e.g. 1), let $k := 0$.

Step 1 : Compute $f(\mathbf{x}^k)$, $\nabla f(\mathbf{x}^k)$, $h_j(\mathbf{x}^k)$, $\nabla h_j(\mathbf{x}^k)$, $j = 1 \dots m$.

Step 2 : Compute L_i^k and U_i^k ($i = 1 \dots n$) by some scheme; define $\bar{f}(\mathbf{x})$, $\bar{h}_j(\mathbf{x})$, $j = 1 \dots m$ [cf. (2.1) and (2.2)].

Step 3 : Solve $SP(\mathbf{x}^k)$; let $(\mathbf{y}^{k+1}, \mathbf{v}^{k+1})$ be the solution, where \mathbf{v}^{k+1} denotes the corresponding vector of Lagrange multipliers.

Step 4 : If $\mathbf{y}^{k+1} = \mathbf{x}^k$ stop; $(\mathbf{x}^k, \mathbf{u}^k)$ is the solution.

Step 5 : Let $\mathbf{s}^k := (\mathbf{x}^k - \mathbf{y}^{k+1}, \mathbf{v}^{k+1})$, $\delta^k := \|\mathbf{y}^{k+1} - \mathbf{x}^k\|$, η^k as defined in (3.4).

Step 6 : Compute $\Phi_r(\mathbf{x}^k, \mathbf{u}^k)$, $\nabla \Phi_r(\mathbf{x}^k, \mathbf{u}^k)$, $\nabla \Phi_r(\mathbf{x}^k, \mathbf{u}^k)^T \mathbf{s}^k$.

Step 7 : If $\nabla \Phi_r(\mathbf{x}^k, \mathbf{u}^k)^T \mathbf{s}^k < \frac{\eta^k (\delta^k)^2}{4}$ let $r := 10r$ and goto Step 6; otherwise compute the smallest $j \in \mathbb{N}_0$, such that $\Phi_r \left[\begin{array}{c} \mathbf{x}^k \\ \mathbf{u}^k \end{array} - \psi^j \mathbf{s}^k \right] \leq \Phi_r(\mathbf{x}^k, \mathbf{u}^k) - c\psi^j \nabla \Phi_r(\mathbf{x}^k, \mathbf{u}^k)^T \mathbf{s}^k$ (Armijo); let $\sigma^k := \psi^j$.

Step 8 : Let $(\mathbf{x}^{k+1}, \mathbf{u}^{k+1}) := (\mathbf{x}^k, \mathbf{u}^k) - \sigma^k \mathbf{s}^k$, $k := k + 1$, goto Step 1.

The major difficulties in proving global convergence for the new method are to show that the search direction defined in Step 5 of the algorithm is a descent direction for Φ_r and that the resulting sequence of penalty parameters is bounded. To prepare the proofs we proceed as follows.

For $\mathbf{x} \in X'$ we rewrite the approximating functions

$$\bar{f}(\mathbf{x}) = \bar{f}(\mathbf{x}^k) + \bar{\nabla} \bar{f}(\mathbf{x}^k)^T(\mathbf{x} - \mathbf{x}^k) + R_{\bar{f}}(\mathbf{x}),$$

$$\bar{h}_j(\mathbf{x}) = \bar{h}_j(\mathbf{x}^k) + \bar{\nabla} \bar{h}_j(\mathbf{x}^k)^T(\mathbf{x} - \mathbf{x}^k) + R_{\bar{h}_j}(\mathbf{x}),$$

where $R_{\bar{f}}, R_{\bar{h}_j} : \mathbb{R}^n \rightarrow \mathbb{R}$ are continuously differentiable.

Since we have convex approximations we conclude $R_{\bar{f}}(\mathbf{x}), R_{\bar{h}_j}(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in X'$.

Since the approximations are of first-order for objective function and non-box-constraints, and the box-constraints can be rewritten as $\bar{h}_j(\mathbf{x}) := h_j(\mathbf{x}) + \kappa_j$ with $\kappa_j \geq 0$ constant, we further write:

$$\bar{f}(\mathbf{x}) = f(\mathbf{x}^k) + \nabla f(\mathbf{x}^k)^T(\mathbf{x} - \mathbf{x}^k) + R_{\bar{f}}(\mathbf{x}),$$

$$\bar{h}_j(\mathbf{x}) = h_j(\mathbf{x}^k) + \nabla h_j(\mathbf{x}^k)^T(\mathbf{x} - \mathbf{x}^k) + R_{\bar{h}_j}(\mathbf{x}), \quad j = 1 \dots m.$$

Next we define: $\bar{h}(\mathbf{x}) := [\bar{h}_1(\mathbf{x}), \dots, \bar{h}_m(\mathbf{x})]^T$ and $R_{\bar{h}}(\mathbf{x}) := [R_{\bar{h}_1}(\mathbf{x}), \dots, R_{\bar{h}_m}(\mathbf{x})]^T$. Hence $\bar{h}(\mathbf{x}) = h(\mathbf{x}^k) + A(\mathbf{x}^k)^T(\mathbf{x} - \mathbf{x}^k) + R_{\bar{h}}(\mathbf{x})$.

For the optimal solution \mathbf{y}^{k+1} of $SP(\mathbf{x}^k)$ we define and conclude as follows.

Corollary 3.6. Let $\Delta\mathbf{x}^k := \mathbf{y}^{k+1} - \mathbf{x}^k$. Then

- (a) $A(\mathbf{x}^k)^T \Delta\mathbf{x}^k = \bar{h}(\mathbf{y}^{k+1}) - h(\mathbf{x}^k) - R_{\bar{h}}(\mathbf{y}^{k+1})$,
 (b) $R_{\bar{h}}(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in X'$,
 (c) $\bar{\mathbf{u}} + r\bar{h}(\mathbf{x}) \geq 0$,
 (d) $\bar{h}(\mathbf{y}^{k+1}) \leq 0$ since \mathbf{y}^{k+1} is feasible for $SP(\mathbf{x}^k)$.

A result from perturbation theory together with (2.5) tells us that $\Delta\mathbf{x}^k$ is a continuous function with respect to \mathbf{x}^k . This is the main reason for introducing the qualification (2.5).

The next lemma shows that the elements of the sequence $(\mathbf{x}^k, \mathbf{u}^k)_{k=0,1,2,\dots}$ are all elements of a predefined compact set.

Lemma 3.7. Let the sequence $(\mathbf{x}^k, \mathbf{u}^k)_{k=0,1,2,\dots}$ be produced by SCP, all subproblems be solvable and gradients of active constraints at the optimal points of $SP(\mathbf{x}^k)$ be linear independent as well as those of $SP(\mathbf{x}^*)$ where \mathbf{x}^* is any possible accumulation point of $(\mathbf{x}^k)_{k=0,1,2,\dots}$. Then all points \mathbf{x}^k are in X and all \mathbf{u}^k are in a compact set U .

Proof: Zillober (1992).

Remark. The property $\mathbf{x}^k \in X$ is a trivial conclusion of the definition of X' .

Now we are able to state the first main result.

Theorem 3.8. Let the assumptions of (3.7) be valid; hence there is a u_{\max} with $u_j^k, v_j^{k+1} \leq u_{\max}$ for all $k \geq 0$ and $j = 1 \dots m$. Let $\mathbf{x}^k \in X$ and $\mathbf{u}^k \geq 0$ be given where \mathbf{x}^k is not a stationary point or (2); \mathbf{s}^k, η^k and δ^k are defined as in the algorithm and let the choice of asymptotes be continuous.

(1) Then there is a penalty parameter $\bar{r}^k > 0$ such that \mathbf{s}^k is a direction of descent for all $r \geq \bar{r}^k$ for the augmented Lagrange function Φ_r , i.e.

$$\nabla\Phi_r(\mathbf{x}^k, \mathbf{u}^k)^T \mathbf{s}^k \geq \frac{\eta^k (\delta^k)^2}{4} \text{ for all } r \geq \bar{r}^k.$$

(2) For each fixed $\delta > 0$ there is a finite \bar{r}^δ such that for all $(\mathbf{x}^k, \mathbf{u}^k)$ with $\delta^k \geq \delta$ we have

$$\nabla\Phi_r(\mathbf{x}^k, \mathbf{u}^k)^T \mathbf{s}^k \geq \frac{\eta^k (\delta^k)^2}{4} \geq \frac{\eta\delta^2}{4} \text{ for all } r \geq \bar{r}^\delta.$$

Proof. $\nabla\Phi_r(\mathbf{x}^k, \mathbf{u}^k)^T \mathbf{s}^k =$

$$-\nabla f(\mathbf{x}^k)^T \Delta\mathbf{x}^k - [\bar{\mathbf{u}}^k + r\bar{h}(\mathbf{x}^k)]^T A(\mathbf{x}^k)^T \Delta\mathbf{x}^k -$$

$$\hat{h}(\mathbf{x}^k)^T (\mathbf{v}^{k+1} - \mathbf{u}^k) = -\nabla f(\mathbf{x}^k)^T \Delta\mathbf{x}^k - [(\bar{\mathbf{u}}^k +$$

$$r\bar{h}(\mathbf{x}^k)]^T [\bar{h}(\mathbf{y}^{k+1}) - h(\mathbf{x}^k) - R_{\bar{h}}(\mathbf{y}^{k+1})] -$$

$$\hat{h}(\mathbf{x}^k)^T (\mathbf{v}^{k+1} - \mathbf{u}^k) \text{ cf. (3.6a)} = -\nabla f(\mathbf{x}^k)^T \Delta\mathbf{x}^k -$$

$$[\bar{\mathbf{u}}^k + r\bar{h}(\mathbf{x}^k)]^T \bar{h}(\mathbf{y}^{k+1}) + [\bar{\mathbf{u}}^k + r\bar{h}(\mathbf{x}^k)]^T h(\mathbf{x}^k) +$$

≤ 0

$$[\bar{\mathbf{u}}^k + r\bar{h}(\mathbf{x}^k)]^T R_{\bar{h}}(\mathbf{y}^{k+1}) - \hat{h}(\mathbf{x}^k)^T (\mathbf{v}^{k+1} -$$

≥ 0

$$\mathbf{u}^k) \text{ cf. (3.6b, c, d)} \geq -\nabla f(\mathbf{x}^k)^T \Delta\mathbf{x}^k + (\bar{\mathbf{u}}^k)^T h(\mathbf{x}^k) +$$

$$r\bar{h}(\mathbf{x}^k)^T h(\mathbf{x}^k) - \hat{h}(\mathbf{x}^k)^T (\mathbf{v}^{k+1} - \mathbf{u}^k) = -\underbrace{\nabla f(\mathbf{x}^k)^T \Delta\mathbf{x}^k}_{(a)} +$$

$$\underbrace{\sum_{j \in J} u_j^k h_j(\mathbf{x}^k)}_{(b)} + r \underbrace{\sum_{j \in J} h_j^2(\mathbf{x}^k)}_{(c)} - \underbrace{\sum_{j \in J} h_j(\mathbf{x}^k) (v_j^{k+1} - u_j^k)}_{(d)} +$$

$$\underbrace{\sum_{j \in K} \frac{u_j^k}{r} (v_j^{k+1} - u_j^k)}_{(e)} \text{ cf. (3.2).}$$

Using the last inequality we first prove part (2) of the theorem. For this purpose, we will show that critical terms are sufficiently small for large penalty parameters.

Let \mathbf{x}^k be feasible. By (3.4b) we then have

$$-\nabla f(\mathbf{x}^k)^T \Delta\mathbf{x}^k \geq \eta^k \|\Delta\mathbf{x}^k\|^2 = \eta^k (\delta^k)^2 > 0.$$

If $J = \emptyset$: (b), (c) and (d) are empty sums,

$$|(e)| = \left| \frac{1}{r} \sum_{j \in K} u_j^k (v_j^{k+1} - u_j^k) \right| \leq \frac{1}{r} m u_{\max}^2 \text{ independent}$$

of the size of K . Hence we have for $r \geq r_1 :=$

$$m \frac{4u_{\max}^2}{\eta\delta^2} : |(e)| \leq \frac{\eta\delta^2}{4} \text{ This means } \nabla\Phi_r(\mathbf{x}^k, \mathbf{u}^k)^T \mathbf{s}^k \geq$$

$$3 \frac{\eta^k (\delta^k)^2}{4} \text{ for all } r \geq r_1.$$

If $J \neq \emptyset$:

We have (c) ≥ 0 . Similar to the above is $|(e)| \leq \frac{\eta\delta^2}{4}$ for all $r \geq r_1$.

(b): Assume $|(b)| \geq \frac{\eta\delta^2}{8}$ (otherwise there is nothing to show!)

$$|(b)| \leq u_{\max} \sum_{j \in J} |h_j(\mathbf{x}^k)| \Rightarrow \sum_{j \in J} |h_j(\mathbf{x}^k)| \geq \frac{\eta\delta^2}{8u_{\max}}.$$
 This sum consists of at most m elements and is essentially not

influenced by those elements where $|h_j(\mathbf{x}^k)| \leq \frac{1}{m} \frac{\eta\delta^2}{8u_{\max}}$.

Since $j \in J$ we have for all j with $|h_j(\mathbf{x}^k)| > \frac{1}{m} \frac{\eta\delta^2}{8u_{\max}}$:

$$-\frac{u_j^k}{r} \leq h_j(\mathbf{x}^k) < -\frac{1}{m} \frac{\eta\delta^2}{8u_{\max}} \text{ and } -\frac{u_j^k}{r} \geq -\frac{u_{\max}}{r}.$$

However, for $r > r_2 := m \frac{8u_{\max}^2}{\eta\delta^2}$ this j is in K and does therefore not increase (b).

(d): Analogously either $|h_j(\mathbf{x}^k)| \leq \frac{1}{m} \frac{\eta\delta^2}{8u_{\max}}$, or $j \in K$ for

$$r > r_2 \text{ such that } |(d)| < \frac{\eta\delta^2}{8}.$$

For each feasible \mathbf{x}^k we conclude

$$\nabla\Phi_r(\mathbf{x}^k, \mathbf{u}^k)^T \mathbf{s}^k \geq \frac{\eta^k (\delta^k)^2}{2} \text{ for all } r > \max\{r_1, r_2\}.$$

Now we assume that \mathbf{x}^k is infeasible. Therefore there is at least one j with $h_j(\mathbf{x}^k) > 0$. This j is an element of J for all positive r and is a non-box constraint because of the definition of X' .

Corollary (3.4a) tells us that (a) $\geq \bar{f}(\mathbf{x}^k) - \bar{f}(\mathbf{y}^{k+1}) + \eta^k (\delta^k)^2$. Since \mathbf{x}^k is now infeasible we can no longer conclude that $\bar{f}(\mathbf{x}^k) - \bar{f}(\mathbf{y}^{k+1}) \geq 0$. However, there is $\beta_1 > 0$, such that for all \mathbf{x}^k where $\|\mathbf{x}^k - \mathbf{y}^{k+1}\| \geq \delta$ we have:

$$\max_{j=1 \dots m} h_j(\mathbf{x}^k) \leq \beta_1 \Rightarrow \bar{f}(\mathbf{x}^k) - \bar{f}(\mathbf{y}^{k+1}) \geq -\frac{\eta \delta^2}{4}; \beta_1 \text{ does not depend on } \mathbf{x}^k (\in X) \text{ since the feasible region is compact.}$$

These conditions yield (a) $\geq 3 \frac{\eta^k (\delta^k)^2}{4}$. In other words, compared with (3.4b) a weaker condition is also valid for expansion points \mathbf{x}^k , which are "slightly infeasible". Now we have to consider the other terms,

$$|(d)| \leq u_{\max} \sum_{j \in J} |h_j(\mathbf{x}^k)| \leq \frac{\eta \delta^2}{8} \Leftrightarrow \sum_{j \in J} |h_j(\mathbf{x}^k)| \leq \frac{\eta \delta^2}{8 u_{\max}}$$

This condition is fulfilled, if (e.g.) all $|h_j(\mathbf{x}^k)| \leq \frac{1}{m} \frac{\eta \delta^2}{8 u_{\max}} =: \beta_2$. If $h_j(\mathbf{x}^k) < 0$ it is valid for $r > r_2$, since this j then belongs to K .

(b): Since $u_j^k \geq 0$ all $j \in J$ where $h_j(\mathbf{x}^k) \geq 0$ are positive elements in (b) and can therefore be neglected; i.e. with the same arguments as above we have for $r > r_2$: (b) $\geq -\frac{\eta \delta^2}{8}$.

Now let $\beta := \min\{\beta_1; \beta_2\}$ and assume $h_j(\mathbf{x}^k) \leq \beta$ for all j . Since (c) ≥ 0 and |(e)| $\leq \frac{\eta \delta^2}{4}$ for $r \geq r_1$ we conclude: $\nabla \Phi_r(\frac{\mathbf{x}^k}{\mathbf{u}^k})^T \mathbf{s}^k \geq 3 \frac{\eta^k (\delta^k)^2}{4} - \frac{\eta \delta^2}{8} - \frac{\eta \delta^2}{8} - \frac{\eta \delta^2}{4} \geq \frac{\eta^k (\delta^k)^2}{4}$ for all $r > \max\{r_1, r_2\}$.

Otherwise, there is at least one j with $h_j(\mathbf{x}^k) > \beta$. Generally we have:

$$|(a)| \leq \|\nabla f(\mathbf{x}^k)\| \|\Delta \mathbf{x}^k\| \leq \max_{\mathbf{x}^k \in X} \|\nabla f(\mathbf{x}^k)\| \max_{\mathbf{x}^k \in X} \|\Delta \mathbf{x}^k\| =: \gamma_1 \quad (3.3b)$$

$$|(b)| \leq m u_{\max} \max_{\mathbf{x}^k \in X} h_j(\mathbf{x}^k) =: \gamma_2, \quad (3.3a)$$

$$|(d)| \leq m u_{\max} \max_{\mathbf{x}^k \in X} h_j(\mathbf{x}^k) = \gamma_2,$$

$$|(e)| \leq m u_{\max}^2 =: \gamma_3 \text{ for } r \geq 1,$$

$$\frac{\eta^k (\delta^k)^2}{4} \leq \frac{1}{4} \eta \max_{\mathbf{x}^k \in X} \|\Delta \mathbf{x}^k\|^2 =: \gamma_4. \quad (3.4)$$

In our situation we have in addition (c) $\geq r \beta^2$. $\Rightarrow \nabla \Phi_r(\frac{\mathbf{x}^k}{\mathbf{u}^k})^T \mathbf{s}^k \geq -\gamma_1 - \gamma_2 - \gamma_3 + r \beta^2 \geq \frac{\eta^k (\delta^k)^2}{4}$ for $r \geq r_3 \geq \frac{\gamma_1 + \gamma_2 + \gamma_3 + \gamma_4}{\beta^2}$ independent of $\mathbf{x}^k (\in X)$

This yields $\nabla \Phi_r(\frac{\mathbf{x}^k}{\mathbf{u}^k})^T \mathbf{s}^k \geq \frac{\eta^k (\delta^k)^2}{4}$ for $r > \bar{r}^\delta := \max\{r_1, r_2, r_3\}$, which proves the second part of the theorem.

Replacing δ by δ^k and η by η^k we can show the existence of corresponding r_1^k, r_2^k (instead of r_1, r_2 , independent of k) for a fixed $(\frac{\mathbf{x}^k}{\mathbf{u}^k})$ where \mathbf{x}^k is feasible, with the same argumentation and hence of \bar{r}^k . However, with these arguments we

can no longer guarantee the boundedness of \bar{r}^k since δ^k can be arbitrarily small.

For infeasible \mathbf{x}^k there is at least one $j \in \{1, \dots, m\}$ where $h_j(\mathbf{x}^k) = \tilde{\beta} > 0$. Therefore the condition is fulfilled with $r \geq \frac{\eta^k (\delta^k)^2}{4} + \frac{\gamma_1 + \gamma_2 + \gamma_3}{\tilde{\beta}^2}$. Since $\tilde{\beta}$ might be arbitrarily small we cannot conclude the boundedness of the penalty parameter in this case also without further arguments.

The next theorem shows us conditions which guarantee the boundedness of the penalty parameter also in the case of $\delta^k \rightarrow 0$.

Theorem 3.9. Let the assumptions of (3.7) be valid and assume a continuous choice of asymptotes. For $\delta^k \neq 0$ we define: $\alpha^k := \frac{\|\mathbf{u}^k - \mathbf{v}^{k+1}\|^2}{(\delta^k)^2}$. Let $(\frac{\mathbf{x}^k}{\mathbf{u}^k})$, where \mathbf{x}^k is not a stationary point for (2), be determined by the SCP-algorithm and fulfill the following two conditions:

- (a) $j \in J$ if and only if $\tilde{h}_j(\mathbf{y}^{k+1}) = 0$ ($j = 1 \dots m$) (i.e. subproblems and the original problem identify the same set of active constraints),
- (b) there is an $\alpha \in \mathbb{R}$ independent of k , such that $\alpha^k \leq \alpha < \infty$

Then there is a $\delta_\tau > 0$, such that

$$\nabla \Phi_\tau(\frac{\mathbf{x}^k}{\mathbf{u}^k})^T \mathbf{s}^k \geq \frac{\eta^k (\delta^k)^2}{4} \text{ for all } (\frac{\mathbf{x}^k}{\mathbf{u}^k}) \text{ with } \delta^k \leq \delta_\tau,$$

$$\text{where } \tau := \min \left\{ 10^j : j \in \mathbb{N}_0, 10^j \geq \frac{2\alpha}{\eta} \right\}.$$

Proof. Since we know the correct set of active constraints we have $u_j^k = v_j^{k+1} = 0$ for all $j \in K$ because of the stationarity conditions of $SP(\mathbf{x}^k)$. Therefore: $\mathbf{u}^k = \bar{\mathbf{u}}^k, \mathbf{v}^{k+1} = \bar{\mathbf{v}}^{k+1}, \tilde{h}_j(\mathbf{x}^k) = \bar{h}_j(\mathbf{x}^k)$. Moreover, $\tilde{h}_j(\mathbf{x}^k) = h_j(\mathbf{x}^k)$ is valid also for active box-constraints.

Using Taylor's formula and the sufficient differentiability of the approximating functions in X' , we can write:

$$\tilde{f}(\mathbf{x}) = f(\mathbf{x}^k) + \nabla f(\mathbf{x}^k)^T (\mathbf{x} - \mathbf{x}^k) + \frac{1}{2} (\mathbf{x} - \mathbf{x}^k)^T \nabla^2 \tilde{f}(\mathbf{x}^k) (\mathbf{x} - \mathbf{x}^k) + \tilde{f}_R(\mathbf{x}), \quad \tilde{h}_j(\mathbf{x}) = h_j(\mathbf{x}^k) + \nabla h_j(\mathbf{x}^k)^T (\mathbf{x} - \mathbf{x}^k) +$$

$$\frac{1}{2} (\mathbf{x} - \mathbf{x}^k)^T \nabla^2 \tilde{h}_j(\mathbf{x}^k) (\mathbf{x} - \mathbf{x}^k) + \tilde{h}_{j,R}(\mathbf{x}),$$

where $\tilde{f}_R, \tilde{h}_{j,R}$ ($j \in J$): $\mathbb{R}^n \rightarrow \mathbb{R}$ are differentiable like \tilde{f} and \tilde{h}_j . Considering Lagrange's remainder we conclude

$$\tilde{f}_R(\mathbf{x}), \tilde{h}_{j,R}(\mathbf{x}) = O\left(\|\mathbf{x} - \mathbf{x}^k\|^3\right).$$

Hence, we write the gradients

$$\nabla \tilde{f}(\mathbf{x}) = \nabla f(\mathbf{x}^k) + \nabla^2 \tilde{f}(\mathbf{x}^k) (\mathbf{x} - \mathbf{x}^k) + \nabla \tilde{f}_R(\mathbf{x}),$$

$$\nabla \tilde{h}_j(\mathbf{x}) = \nabla h_j(\mathbf{x}^k) + \nabla^2 \tilde{h}_j(\mathbf{x}^k) (\mathbf{x} - \mathbf{x}^k) + \nabla \tilde{h}_{j,R}(\mathbf{x}),$$

$$\text{where } \|\nabla \tilde{f}_R(\mathbf{x})\|, \|\nabla \tilde{h}_{j,R}(\mathbf{x})\| = O\left(\|\mathbf{x} - \mathbf{x}^k\|^2\right).$$

Using these equations the stationarity conditions for $SP(\mathbf{x}^k)$ are

$$\begin{aligned}
\text{(i)} \quad & \nabla f(\mathbf{x}^k) + \nabla^2 \tilde{f}(\mathbf{x}^k)(\mathbf{y}^{k+1} - \mathbf{x}^k) + \nabla \tilde{f}_R(\mathbf{y}^{k+1}) = \\
& - \sum_{j=1}^m v_j^{k+1} \left[\nabla h_j(\mathbf{x}^k) + \nabla^2 \tilde{h}_j(\mathbf{x}^k)(\mathbf{y}^{k+1} - \mathbf{x}^k) + \right. \\
& \quad \left. \nabla \tilde{h}_{j,R}(\mathbf{y}^{k+1}) \right], \\
\text{(ii)} \quad & v_j^{k+1} \left[h_j(\mathbf{x}^k) + \nabla h_j(\mathbf{x}^k)^T (\mathbf{y}^{k+1} - \mathbf{x}^k) + \frac{1}{2} (\mathbf{y}^{k+1} - \right. \\
& \quad \left. \mathbf{x}^k)^T \nabla^2 \tilde{h}_j(\mathbf{x}^k) (\mathbf{y}^{k+1} - \mathbf{x}^k) + \tilde{h}_{j,R}(\mathbf{y}^{k+1}) \right] = 0, \quad j = \\
& \quad 1 \dots m \\
\text{(iii)} \quad & h_j(\mathbf{x}^k) + \nabla h_j(\mathbf{x}^k)^T (\mathbf{y}^{k+1} - \mathbf{x}^k) + \frac{1}{2} (\mathbf{y}^{k+1} - \\
& \quad \mathbf{x}^k)^T \nabla^2 \tilde{h}_j(\mathbf{x}^k) (\mathbf{y}^{k+1} - \mathbf{x}^k) + \tilde{h}_{j,R}(\mathbf{y}^{k+1}) \leq 0, \quad j = \\
& \quad 1 \dots m, \\
\text{(iv)} \quad & v_j^{k+1} \geq 0, \quad j = 1 \dots m; \\
\text{(i)} \quad & \text{yields} \quad -\nabla f(\mathbf{x}^k)^T \Delta \mathbf{x}^k = (\Delta \mathbf{x}^k)^T \nabla^2 \tilde{f}(\mathbf{x}^k) \Delta \mathbf{x}^k + \\
& \nabla \tilde{f}_R(\mathbf{y}^{k+1})^T \Delta \mathbf{x}^k + (\mathbf{v}^{k+1})^T A(\mathbf{x}^k)^T \Delta \mathbf{x}^k + \\
& \left\{ \sum_{j=1}^m v_j^{k+1} \left[\nabla^2 \tilde{h}_j(\mathbf{x}^k) \Delta \mathbf{x}^k + \nabla \tilde{h}_{j,R}(\mathbf{y}^{k+1}) \right] \right\}^T \Delta \mathbf{x}^k.
\end{aligned}$$

From the definition of \tilde{h}_j , we know

$$A(\mathbf{x}^k)^T \Delta \mathbf{x}^k = \tilde{h}(\mathbf{y}^{k+1}) - h(\mathbf{x}^k) - H(\Delta \mathbf{x}^k) - \tilde{h}_R(\mathbf{y}^{k+1}),$$

where

$$\tilde{h}_R(\mathbf{y}^{k+1}) := [\tilde{h}_{1,R}(\mathbf{y}^{k+1}), \dots, \tilde{h}_{m,R}(\mathbf{y}^{k+1})]^T,$$

and

$$H(\Delta \mathbf{x}^k) := \left[\frac{1}{2} (\mathbf{y}^{k+1} - \mathbf{x}^k)^T \nabla^2 \tilde{h}_1(\mathbf{x}^k) (\mathbf{y}^{k+1} - \mathbf{x}^k), \dots, \right.$$

$$\left. \frac{1}{2} (\mathbf{y}^{k+1} - \mathbf{x}^k)^T \nabla^2 \tilde{h}_m(\mathbf{x}^k) (\mathbf{y}^{k+1} - \mathbf{x}^k) \right]^T.$$

If $j \in J$, then

$$|h_j(\mathbf{x}^k)| = |\tilde{h}_j(\mathbf{x}^k)| = |\tilde{h}_j(\mathbf{x}^k) - \tilde{h}_j(\mathbf{y}^{k+1})|$$

$$[\text{since } \tilde{h}_j(\mathbf{y}^{k+1}) = 0] =$$

$$|h_j(\mathbf{x}^k) - h_j(\mathbf{x}^k) - \nabla \tilde{h}_j(\mathbf{x}^k + \theta(\mathbf{y}^{k+1} - \mathbf{x}^k))^T (\mathbf{y}^{k+1} - \mathbf{x}^k)|,$$

$$(\text{where } \theta \in [0; 1]) =$$

$$|-\nabla \tilde{h}_j(\mathbf{x}^k + \theta(\mathbf{y}^{k+1} - \mathbf{x}^k))^T (\mathbf{y}^{k+1} - \mathbf{x}^k)| \leq L \|\mathbf{y}^{k+1} - \mathbf{x}^k\|,$$

with $L \geq 0$ (3.3d).

Now we are able to derive the necessary inequality

$$\begin{aligned}
& \nabla \Phi_L(\mathbf{x}^k)^T \mathbf{s}^k = \\
& -\nabla f(\mathbf{x}^k)^T \Delta \mathbf{x}^k - (\bar{\mathbf{u}}^k + \underline{r}\bar{h}(\mathbf{x}^k))^T A(\mathbf{x}^k)^T \Delta \mathbf{x}^k - \\
& \hat{h}(\mathbf{x}^k)^T (\mathbf{v}^{k+1} - \mathbf{u}^k) = (\Delta \mathbf{x}^k)^T \nabla^2 \tilde{f}(\mathbf{x}^k) \Delta \mathbf{x}^k + \\
& \nabla \tilde{f}_R(\mathbf{y}^{k+1})^T \Delta \mathbf{x}^k + (\mathbf{v}^{k+1})^T A(\mathbf{x}^k)^T \Delta \mathbf{x}^k + \\
& \quad O[(\delta^k)^3]
\end{aligned}$$

$$\begin{aligned}
& \underbrace{\sum_{j=1}^m v_j^{k+1} (\Delta \mathbf{x}^k)^T \nabla^2 \tilde{h}_j(\mathbf{x}^k) \Delta \mathbf{x}^k +}_{\geq 0 \text{ (convexity)}} \\
& \underbrace{\sum_{j=1}^m v_j^{k+1} \nabla \tilde{h}_{j,R}(\mathbf{y}^{k+1})^T \Delta \mathbf{x}^k - (\bar{\mathbf{u}}^k +}_{O[(\delta^k)^3]}
\end{aligned}$$

$$\begin{aligned}
& \underline{r}\bar{h}(\mathbf{x}^k))^T A(\mathbf{x}^k)^T \Delta \mathbf{x}^k - \hat{h}(\mathbf{x}^k)^T (\mathbf{v}^{k+1} - \mathbf{u}^k) \geq
\end{aligned}$$

$$\begin{aligned}
& (\Delta \mathbf{x}^k)^T \nabla^2 \tilde{f}(\mathbf{x}^k) \Delta \mathbf{x}^k + (\mathbf{v}^{k+1})^T A(\mathbf{x}^k)^T \Delta \mathbf{x}^k - \\
& \geq \eta^k (\delta^k)^2 \text{ (3.4), (3.3e)} \\
& [\bar{\mathbf{u}}^k + \underline{r}\bar{h}(\mathbf{x}^k)]^T A(\mathbf{x}^k)^T \Delta \mathbf{x}^k - \hat{h}(\mathbf{x}^k)^T (\mathbf{v}^{k+1} - \mathbf{u}^k) + \\
& \quad = \bar{h}(\mathbf{x}^k)^T (\mathbf{v}^{k+1} - \mathbf{u}^k)
\end{aligned}$$

$$O[(\delta^k)^3] \geq \eta^k (\delta^k)^2 + \underbrace{(\mathbf{v}^{k+1})^T \tilde{h}(\mathbf{y}^{k+1}) - (\mathbf{v}^{k+1})^T \tilde{h}(\mathbf{x}^k)}_{=0 \text{ (ii)}} -$$

$$(\mathbf{v}^{k+1})^T H(\Delta \mathbf{x}^k) - \underbrace{(\mathbf{v}^{k+1})^T \tilde{h}_R(\mathbf{y}^{k+1})}_{O[(\delta^k)^3]} -$$

$$[\bar{\mathbf{u}}^k + \underline{r}\bar{h}(\mathbf{x}^k)]^T \tilde{h}(\mathbf{y}^{k+1}) + [\bar{\mathbf{u}}^k + \underline{r}\bar{h}(\mathbf{x}^k)]^T \tilde{h}(\mathbf{x}^k) +$$

$$= 0, \text{ cf. (a)}$$

$$[\bar{\mathbf{u}}^k + \underline{r}\bar{h}(\mathbf{x}^k)]^T H(\Delta \mathbf{x}^k) + \underbrace{[\bar{\mathbf{u}}^k + \underline{r}\bar{h}(\mathbf{x}^k)]^T \tilde{h}_R(\mathbf{y}^{k+1})}_{O[(\delta^k)^3]} -$$

$$\bar{h}(\mathbf{x}^k)^T (\mathbf{v}^{k+1} - \mathbf{u}^k) + O[(\delta^k)^3] \geq \eta^k (\delta^k)^2 - \underbrace{(\mathbf{v}^{k+1})^T \tilde{h}(\mathbf{x}^k)}_{= (\mathbf{v}^{k+1})^T \tilde{h}(\mathbf{x}^k)} -$$

$$(\mathbf{v}^{k+1})^T H(\Delta \mathbf{x}^k) + \underbrace{(\bar{\mathbf{u}}^k)^T \tilde{h}(\mathbf{x}^k)}_{= (\mathbf{u}^k)^T \tilde{h}(\mathbf{x}^k)} + \underline{r} \bar{h}(\mathbf{x}^k)^T \tilde{h}(\mathbf{x}^k) +$$

$$= \bar{h}(\mathbf{x}^k)^T \tilde{h}(\mathbf{x}^k)$$

$$\underbrace{(\bar{\mathbf{u}}^k)^T H(\Delta \mathbf{x}^k)}_{= (\mathbf{u}^k)^T H(\Delta \mathbf{x}^k)} + \underline{r} \bar{h}(\mathbf{x}^k)^T H(\Delta \mathbf{x}^k) - \bar{h}(\mathbf{x}^k)^T (\mathbf{v}^{k+1} - \mathbf{u}^k) +$$

$$O[(\delta^k)^3] = \eta^k (\delta^k)^2 + \underline{r} \bar{h}(\mathbf{x}^k)^T [\bar{h}(\mathbf{x}^k) + H(\Delta \mathbf{x}^k)] -$$

$$(\mathbf{v}^{k+1})^T [2\bar{h}(\mathbf{x}^k) + H(\Delta \mathbf{x}^k)] + (\mathbf{u}^k)^T [2\bar{h}(\mathbf{x}^k) + H(\Delta \mathbf{x}^k)] +$$

$$O[(\delta^k)^3] = \eta^k (\delta^k)^2 + \underline{r} [\bar{h}(\mathbf{x}^k) + H(\Delta \mathbf{x}^k)]^T [\bar{h}(\mathbf{x}^k) +$$

$$H(\Delta \mathbf{x}^k)] + 2(\mathbf{u}^k - \mathbf{v}^{k+1})^T [\bar{h}(\mathbf{x}^k) + H(\Delta \mathbf{x}^k)] - \underline{r} H(\Delta \mathbf{x}^k)^T$$

$$[\bar{h}(\mathbf{x}^k) + H(\Delta \mathbf{x}^k)] - (\mathbf{u}^k - \mathbf{v}^{k+1})^T H(\Delta \mathbf{x}^k) + O[(\delta^k)^3] =$$

$$\eta^k (\delta^k)^2 + \underbrace{\left\| \sqrt{\underline{r}} [\bar{h}(\mathbf{x}^k) + H(\Delta \mathbf{x}^k)] + \frac{1}{\sqrt{\underline{r}}} (\mathbf{u}^k - \mathbf{v}^{k+1}) \right\|^2}_{\geq 0} -$$

$$\frac{\|\mathbf{u}^k - \mathbf{v}^{k+1}\|^2}{\underline{r}} - \underbrace{\underline{r} H(\Delta \mathbf{x}^k)^T \bar{h}(\mathbf{x}^k)}_{O[(\delta^k)^3]} - \underbrace{\underline{r} H(\Delta \mathbf{x}^k)^T H(\Delta \mathbf{x}^k)}_{O[(\delta^k)^4]} -$$

$$\underbrace{(\mathbf{u}^k - \mathbf{v}^{k+1})^T H(\Delta \mathbf{x}^k)}_{O[(\delta^k)^3] \text{ cf. (b)}} + O[(\delta^k)^3] \geq \eta^k (\delta^k)^2 -$$

$$\frac{\|\mathbf{u}^k - \mathbf{v}^{k+1}\|^2}{\underline{r}} + O[(\delta^k)^3] \geq \frac{1}{2} \eta^k (\delta^k)^2 + \frac{1}{2} \eta^k (\delta^k)^2 -$$

$$\frac{\alpha (\delta^k)^2}{\underline{r}} + O[(\delta^k)^3], \text{ cf (b)} = \frac{1}{2} \eta^k (\delta^k)^2 +$$

$$\underbrace{\left(\frac{1}{2} \eta - \frac{\alpha}{\underline{r}} \right) (\delta^k)^2 + O[(\delta^k)^3]}_{\geq 0} \geq \frac{1}{2} \eta^k (\delta^k)^2 + O[(\delta^k)^3].$$

Now let $g(\frac{x^k}{u^k})$ be the (rest-) function which is $O[(\delta^k)^3]$.

Thus, there is an $M_r \in \mathbb{R}$ (r is fixed!), such that $\left| \frac{g(\frac{x^k}{u^k})}{(\delta^k)^3} \right| < M_r$.

Our goal is: $\left| g(\frac{x^k}{u^k}) \right| \leq \frac{1}{4} \eta^k (\delta^k)^2$. This is fulfilled with $\delta^k \leq \frac{\eta}{4M_r} =: \delta_r = \text{const.} > 0$, such that $\nabla \Phi_r(\frac{x^k}{u^k})^T s^k \geq \frac{\eta^k (\delta^k)^2}{4}$ for all $(\frac{x^k}{u^k})$ with $\delta^k \leq \delta_r$.

Corollary 3.10. Let the sequence $(\frac{x^k}{u^k})_{k=0,1,2,\dots}$ be produced by SCP. If the assumptions of (3.8) and (3.9) are valid, then the corresponding sequence of penalty parameters is bounded, i.e. there is an $\bar{r} < \infty$, such that $r^k = \bar{r}$ for all $k \geq \bar{k} \geq 0$.

Proof. We consider r and δ_r of (3.9). Using (3.8,2) there is an $\bar{r}^\delta < \infty$ such that for $\delta := \delta_r$ we have

$$\nabla \Phi_r(\frac{x^k}{u^k})^T s^k \geq \frac{\eta^k (\delta^k)^2}{4} \text{ for all } r \geq \bar{r}^\delta \text{ and all } (\frac{x^k}{u^k}) \text{ with } \delta^k \geq \delta.$$

If $\bar{r}^\delta \leq r$ then the assertion follows directly since then r is the required \bar{r} .

If $\bar{r}^\delta > r$ we have to make some additional considerations. By the assumptions of (3.9) we have for the function $\Psi : [1, \infty[\rightarrow \mathbb{R}$ with $\Psi(r) := \nabla \Phi_r(\frac{x^k}{u^k})^T s^k : \Psi$ is one-dimensional and continuously differentiable with $\Psi'(r) = -\bar{h}(x^k)^T A(x^k)^T \Delta x^k = \text{const.}$ [cf. first row of the main inequality in the proof of (3.9) using $\hat{h}(x^k) = \bar{h}(x^k)$]. That is Ψ is linear and since $\lim_{r \rightarrow \infty} \Psi(r) > 0$ [cf. (3.8,1)] we conclude

that $\Psi'(r) \geq 0$. This means that $\Psi(\bar{r}^\delta) \geq \Psi(r) \geq \frac{\eta^k (\delta^k)^2}{4}$ for all $(\frac{x^k}{u^k})$ where $\delta^k \leq \delta$. That is the penalty parameter also does not need to be enlarged in these points and $\bar{r} = \bar{r}^\delta$.

Now we are able to formulate the main result of this paper.

Theorem 3.11. Let $x^0 \in X$ and the sequence $(\frac{x^k}{u^k})_{k=0,1,2,\dots}$ produced by SCP fulfill the assumptions of (3.8) and (3.9). Then the sequence either terminates at a stationary point, or it has at least one accumulation point and each accumulation point is a stationary point for (2).

Proof. X is compact and convex. Since each subproblem is uniquely solvable, its solution yields y^{k+1} , which is different from x^k as long as x^k is not a stationary point (2.4). Hence the search direction s^k does not vanish. Theorem (3.8) tells us then that we can find an $r > 0$ such that s^k is a direction of descent for the augmented Lagrange function Φ_r . The step size procedure (Step 7 of the algorithm) results in a point $(\frac{x^{k+1}}{u^{k+1}})$ where $x^{k+1} \in]x^k, y^{k+1}]$ and $u^{k+1} \in]u^k, v^{k+1}]$. Using (3.7), $(\frac{x^k}{u^k})_{k=0,1,2,\dots}$ is therefore a sequence in the compact set (X, U) . Hence, $(\frac{x^k}{u^k})_{k=0,1,2,\dots}$ has at least one accumulation point $(\frac{x^*}{u^*}) \in (X, U)$.

If a stationary point is an element of the sequence, then the algorithm terminates because of (2.4) and Step 4 of the

algorithm. Otherwise an infinite sequence is produced.

Let us now assume that there is an accumulation point $(\frac{x^*}{u^*})$ which is not stationary.

By (3.10) we have a finite \bar{r} , such that

$$\nabla \Phi_r(\frac{x^k}{u^k})^T s^k \geq \frac{\eta^k (\delta^k)^2}{4} \text{ for all } r \geq \bar{r} \text{ and } k \geq 0.$$

This means that reaching some iteration index the penalty parameter does not need to be enlarged and is therefore constant. Let us now consider this remaining sequence with constant penalty parameter \bar{r} . Hence, in Step 7 of the algorithm we always have the same function $\Phi_{\bar{r}}$. The remaining sequence now contains a subsequence which converges to $(\frac{x^*}{u^*})$.

Let $\|\Delta x^*\| =: 2\delta$; $\delta > 0$ because of (2.4). Since Δx^k is a continuous function of x^k there is a $\bar{k} \in \mathbb{N}_0$ such that $\|\Delta x^k\| \geq \delta$ for all elements of the subsequence with $k \geq \bar{k}$, i.e. for infinitely many iteration points. For these elements of the subsequence we conclude now that the scalar product of the gradient of the augmented Lagrange function and the search direction is uniformly bounded away from zero, i.e.

$$\nabla \Phi_{\bar{r}}(\frac{x^k}{u^k})^T s^k \geq \frac{\eta \delta^2}{4}.$$

Standard arguments of optimization tell us further that this fact leads to a sequence of step sizes (σ^k) which is bounded away from zero by a $\underline{\sigma} > 0$. This means (cf. Step 7 of the algorithm)

$$\Phi_{\bar{r}}(\frac{x^k}{u^k}) - \Phi_{\bar{r}}\left[\left(\frac{x^k}{u^k}\right) - \sigma^k s^k\right] \geq c \sigma^k \nabla \Phi_{\bar{r}}(\frac{x^k}{u^k})^T s^k \geq c \underline{\sigma} \frac{\eta \delta^2}{4},$$

for infinitely many iteration points, yielding $\lim_{k \rightarrow \infty} \Phi_{\bar{r}}(\frac{x^k}{u^k}) = -\infty$, which is a contradiction to (3.5). Hence the assumption is wrong and the assertion proved.

At the end of this section we want to prove a result which is weaker than (3.11) because it makes no use of (3.9). This means that we do not assume $\alpha^k \leq \alpha$ any longer, but allow an unbounded sequence of α^k .

Theorem 3.12. Let $x^0 \in X$ and the sequence $(\frac{x^k}{u^k})_{k=0,1,2,\dots}$ produced by SCP fulfill the assumptions of (3.8). Then the sequence either terminates at a stationary point, or it has at least one accumulation point and at least one accumulation point is a stationary point for (2).

Proof.

- (1) There is at least one accumulation point similar to the first part of the proof of (3.11).
- (2) If the sequence of penalty parameters is bounded, then the proof of (3.11) is transferable and the therein cited stronger results hold. Therefore we assume a non-bounded sequence of penalty parameters.
- (3) There is an infinite subsequence of $(\frac{x^k}{u^k})_{k=0,1,2,\dots}$ with $\delta^k \rightarrow 0$; otherwise the sequence of parameters would be bounded because of (3.8,2).
- (4) Since this subsequence is in the compact set (X, U) it has at least one accumulation point where $\delta^k = 0$. Using (2.4), this is a stationary point.

4 Artificial variables

One important assumption in the theorems (3.8) and (3.9) is the solvability of the subproblems, i.e. non-empty feasible regions. This cannot be ensured in advance and in fact it is sometimes not fulfilled, especially in the first iterations, when we are still far away from the solution. However, there are various techniques to overcome this situation (see e.g. Fleury and Braibant 1986; Svanberg 1987; Schittkowski 1983). In this section we briefly describe one possibility and state a result which preserves the convergence of the method.

We look at the situation, where there is at least one $j \in \{1, \dots, m\}$ with $h_j(\mathbf{x}^k) > 0$. Otherwise \mathbf{x}^k is feasible and therefore $SP(\mathbf{x}^k)$ is solvable.

Definition 4.1. The auxiliary problem $AP_\rho(\mathbf{x}^k, 1)$ is defined by

$$\min_{\mathbf{x}, \mu} \left\{ \bar{f}(\mathbf{x}, \mu) \mid \bar{h}_j'(\mathbf{x}, \mu) \leq 0, j = 1 \dots m, 0 \leq \mu_j \leq 1, j \in M \right\},$$

where $M := \{j \in \{1, \dots, m\} \mid h_j(\mathbf{x}^k) > 0\}$, $\bar{f}(\mathbf{x}, \mu) := \bar{f}(\mathbf{x}) + \frac{1}{2} \sum_{j \in M} \rho_j \mu_j^2$,

$$\bar{h}_j'(\mathbf{x}, \mu) := \begin{cases} \bar{h}_j(\mathbf{x}) - \mu_j \bar{h}_j(\mathbf{x}^k), & \text{if } j \in M \\ \bar{h}_j(\mathbf{x}), & \text{if } j \notin M \end{cases} \quad \text{and}$$

ρ_j ($j \in M$) are some positive parameters.

We note that $(\mathbf{x}^k, 1)$ is feasible for $AP_\rho(\mathbf{x}^k, 1)$ and that μ_j ($j \in M$) are additional primal variables. The functions of the auxiliary problem remain convex and separable.

The solution of $AP_\rho(\mathbf{x}^k, 1)$ still yields a descent direction for the augmented Lagrange function.

Theorem 4.2. Let \mathbf{x}^k and $\mathbf{u}^k \geq 0$ be given, where \mathbf{x}^k is not feasible for $SP(\mathbf{x}^k)$. Let the gradients of functions \bar{h}_j belonging to active constraints of $AP_\rho(\mathbf{x}^k, 1)$ in the solution point $(\mathbf{y}^{k+1}, \mathbf{v}^{k+1}, \mu^{k+1})$ be linear independent. If $\mu_j^{k+1} < 1$ for all $j \in M$ then there is a penalty parameter $\bar{r} > 0$, such that $\mathbf{s}^k := \begin{pmatrix} \mathbf{x}^k - \mathbf{y}^{k+1} \\ \mathbf{u}^k - \mathbf{v}^{k+1} \end{pmatrix}$ is a descent direction for the augmented Lagrange function Φ_r for all $r \geq \bar{r}$, i.e.

$$\nabla \Phi_r(\mathbf{x}^k, \mathbf{u}^k)^T \mathbf{s}^k \geq \frac{\eta^k (\delta^k)^2}{4} \quad \text{for all } r \geq \bar{r}.$$

Proof. Zillober (1992).

Now we have to answer the question, whether the solvability condition was replaced by another strong condition ($\mu_j < 1$ for all $j \in M$), or if this condition is weaker.

Lemma 4.3. Let the gradients of \bar{h}_j belonging to active constraints of $AP_\rho(\mathbf{x}^k, 1)$ be linear independent at each feasible point. Then there is for arbitrarily fixed $0 < \bar{\mu}_j < 1$ a parameter $\bar{\rho}_j < \infty$, such that $\mu_j^{k+1} \leq \bar{\mu}_j$ for all $\rho_j \geq \bar{\rho}_j$ ($j \in M$) in the optimal point $(\mathbf{y}^{k+1}, \mathbf{v}^{k+1}, \mu^{k+1})$ of $AP_\rho(\mathbf{x}^k, 1)$.

Proof. Zillober (1992).

The lemma tells us that an iterative process is needed to obtain a suitable solution of the auxiliary problem since a sufficient parameter ρ is not known in advance.

5 Examples

By the following 2 examples we compare the convergence behaviour of SCP to that of the MMA method. The efficiency of

MMA in the case of convergence is well-known. The examples are computed using the Finite-Element-System LAGRANGE (Knepe *et al.* 1987). The examples analysed are called "tbdyn" and "aplate2". In the table below we show the weight of the structure in each iteration followed by the maximum constraint violation and, in the case of SCP, the step size (σ^k in Step 7 of the algorithm). We neglect the units and further details since they are not essential for our purposes. The asymptotes are chosen identical for both methods due to a slightly modified scheme by Svanberg (1987). Let k be the iteration index; then for all $i = 1 \dots n$ we choose as follows:

$$k = 0, 1 : L_i^k = \underline{x}_i - 0.1(\bar{x}_i - \underline{x}_i), \quad U_i^k = \bar{x}_i + 0.1(\bar{x}_i - \underline{x}_i),$$

$$k = 2, 3, \dots : \text{if } \text{sign}(x_i^k - x_i^{k-1}) = \text{sign}(x_i^{k-1} - x_i^{k-2}) :$$

$$L_i^k = x_i^k - \frac{x_i^{k-1} - L_i^{k-1}}{t}, \quad U_i^k = x_i^k + \frac{U_i^{k-1} - x_i^{k-1}}{t}.$$

$$\text{If } \text{sign}(x_i^k - x_i^{k-1}) \neq \text{sign}(x_i^{k-1} - x_i^{k-2}) :$$

$$L_i^k = x_i^k - t(x_i^{k-1} - L_i^{k-1}), \quad U_i^k = x_i^k + t(U_i^{k-1} - x_i^{k-1}),$$

where $t \in]0, 1[$ is a parameter. We chose $t = 0.7$.

5.1 Example "tbdyn"

We consider the famous 10-bar truss and minimize the weight of the structure subject to one major constraint. There is a lower bound on the smallest eigenfrequency. Additionally, there are lower and upper bounds on the design variables. The initial weight is 3630. The optimal weight is 2185.19. The constraint is active at the optimum.

It.	SCP			MMA		
0	3630	/	0	/	-	3630 / 0
...		
3	1857.00	/	0.72	/	1	1857.00 / 0.72
4	1857.01	/	0.70	/	0.10	1857.10 / 0.79
5	1856.62	/	0.69	/	0.16	1856.75 / 0.72
6	1855.89	/	0.68	/	0.1	1852.95 / 0.79
7	1855.85	/	0.68	/	0.03	1855.94 / 0.72
8	1860.44	/	0.67	/	0.04	1847.27 / 0.79
9	1865.56	/	0.65	/	0.04	1853.22 / 0.73
10	1883.42	/	0.59	/	0.11	1842.03 / 0.79
11	1905.00	/	0.52	/	0.13	1847.78 / 0.73
12	1924.59	/	0.46	/	0.13	1960.98 / 0.73
13	2127.57	/	0.12	/	1	1844.64 / 0.73
14	2119.35	/	0.11	/	0.1	1953.24 / 0.72
...		
36	2185.19	/	$2.5 \cdot 10^{-7}$	/	1	1924.78 / 0.53

The first 3 iterations are equivalent (i.e. $\sigma^k = 1$ in Step 7 of the algorithm). MMA is cycling between designs that are very different from the optimal one. SCP overcomes this situation and finally converges to the optimum.

5.2 Example "aplate2"

This is a minimum weight problem of a plate with 49 finite elements, 7 design variables, 49 stress constraints, 2 displacement constraints, i.e. 51 major restrictions and lower and upper bounds on the variables. The initial weight is (scaled) 5.282, the optimal weight is 3.217.

It.	SCP			MMA				
0	5.282	/	0	/	-	5.282	/	0
...		
7	3.159	/	1.36	/	1	3.159	/	1.36
8	3.148	/	1.11	/	0.1	3.046	/	95.29
9	3.162	/	0.61	/	0.47	3.093	/	38.75
10	3.174	/	0.41	/	0.34	3.122	/	15.23
11	3.177	/	0.37	/	0.1	3.145	/	5.92
12	3.180	/	0.33	/	0.1	3.175	/	4.45
13	3.210	/	$4.14 \cdot 10^{-2}$	/	1	3.227	/	1.70
14	3.211	/	$3.9 \cdot 10^{-2}$	/	0.06	3.267	/	0.49
15	3.217	/	$1.07 \cdot 10^{-3}$	/	1	3.277	/	0.06
16	3.217	/	$8.46 \cdot 10^{-7}$	/	1	3.258	/	2.20
17						3.088	/	78.04
18						3.111	/	17.04
19						3.143	/	7.23
						...		

The first 7 iterations are equivalent. MMA also cycles in this example, but close to the optimal solution without approximating it. SCP, however, converges directly to the optimal design.

Finally we must mention that in many other test cases SCP and MMA produce the same sequence of iteration points. Sometimes the sequences differ but both methods converge to the optimal solution and in some examples MMA diverges, contrary to SCP. Two of them have been shown in this section.

6 Conclusion

The method we described in this paper has the advantage of guaranteed global convergence in contrast to the original MMA method. Numerical experience shows that this is not only a theoretical property, but has significant practical impact. The SCP method is a stabilized version of MMA and solves problems for which MMA cycles.

In our opinion, more attention should be paid to the choice of the asymptotes in the future. These parameters are essential for the behaviour of MMA or SCP. First tests

showed significant differences between some preliminary versions.

Acknowledgement

This work was supported by the German Science Foundation (DFG) under grant Schi 173/6-1.

References

- Fleury, C. 1989: An efficient dual optimizer based on convex approximation concepts. *Struct. Optim.* 1, 81-89
- Fleury, C.; Braibant, V. 1986: Structural optimization - a new dual method using mixed variables. *Int. J. Num. Meth. Eng.* 23, 409-428
- Hörnlein, H.; Schittkowski, K. (eds.) 1992: *Software systems for structural optimization*. Basel: Birkhäuser Verlag
- Knepe, G.; Krammer, J.; Winkler, F. 1987: Structural optimization of large scale problems using MBB LAGRANGE. Presented at 5th World Congress and Exhibition on FEM (held in Salzburg)
- Nguyen, V.H.; Strodiot, J.J.; Fleury, C. 1987: A mathematical convergence analysis for the convex linearization method for engineering design optimization. *Eng. Opt.* 11, 195-216
- Schittkowski, K. 1981: The nonlinear programming method of Wilson, Han and Powell with an augmented Lagrangian type line search function. *Numerische Mathematik* 38, 83-114
- Schittkowski, K. 1983: On the convergence of a Sequential Quadratic Programming method with an augmented Lagrangian line search function. *Math. Operationsforschung u. Statistik, Ser. Optimization* 14, 197-216
- Svanberg, K. 1987: The method of moving asymptotes - a new method for structural optimization. *Int. J. Num. Meth. Eng.* 24, 359-373
- Svanberg, K. 1993: The method of moving asymptotes (MMA) with some extensions. In: Rozvany, G.I.N. (ed.) *Optimization of large structural systems* (Proc. NATO ASI held in Berchtesgaden, Germany, 1991), pp. 555-566. Dordrecht: Kluwer
- Zillober, C. 1992: *Eine global konvergente Methode zur Lösung von Problemen aus der Strukturoptimierung*. Doctoral Dissertation, Technische Universität München

Received Dec. 4, 1992

Revised manuscript received Apr. 22, 1993