

CH. GROSSMANN

A. A. KAPLAN

**On the rate of convergence of sequential  
unconstrained minimization techniques**

*RAIRO. Analyse numérique*, tome 17, n° 3 (1983), p. 267-292

[http://www.numdam.org/item?id=M2AN\\_1983\\_\\_17\\_3\\_267\\_0](http://www.numdam.org/item?id=M2AN_1983__17_3_267_0)

© AFCET, 1983, tous droits réservés.

L'accès aux archives de la revue « RAIRO. Analyse numérique » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>

## ON THE RATE OF CONVERGENCE OF SEQUENTIAL UNCONSTRAINED MINIMIZATION TECHNIQUES (\*)

by Ch. GROßMANN <sup>(1)</sup>, A. A. KAPLAN <sup>(2)</sup>

Communicated by P. J. LAURENT

*Résumé.* — Dans cet article, sur la base des relations étroites entre les solutions des problèmes auxiliaires de la méthode de pénalité et le comportement du problème original à l'égard des perturbations un principe relativement général va être dérivé pour déterminer l'ordre de convergence de certaines méthodes successives pour la minimization non contrainte. L'accès présenté va être appliqué aux méthodes de pénalité et aux méthodes de centres. De plus à l'aide de l'ordre de convergence certaines règles pour le choix des paramètres des méthodes de pénalité régularisées vont être données.

*Summary.* — Basing on the close relation between solutions of the auxiliary problems arising in sequential unconstrained minimization techniques and the behaviour of the primal problem subject to perturbations in the right hand sides of the inequality constraints in this paper we derive a quite general technique for estimating the rate of convergence of sequential unconstrained minimization methods. The given approach is applied to penalty methods and methods of centers. Furthermore parameter selection rules for regularized penalty methods are founded by means of the given estimations for the rate of convergence.

### 1. INTRODUCTION

The transformation of a given nonlinear programming problem

$$f_0(x) \rightarrow \min ! \quad \text{subject to } x \in X, \quad f_i(x) \leq 0, \quad i = 1, \dots, m \quad (1)$$

into a sequence of unconstrained minimization problems

$$T(x, y^k) \rightarrow \min ! \quad \text{subject to } x \in X \quad (2)$$

forms and effective tool for handling constrained optimization problems

---

(\*) Received in May 1982.

<sup>(1)</sup> Technische Universität Dresden, Sektion Mathematik, DDR-8027-Dresden, Mommsenstr. 13, R.D.A.

<sup>(2)</sup> Sibirian Branch of the Soviet Academy, of Sciences, Institute of Mathematics, 630090, Novosibirsk, U.R.S.S.

(see e.g. [4, 6, 9, 20, 26]). Thereby estimations for the rate of convergence play an essential role as well for theoretical investigations as for the control of the parameters in practical applications of the related method.

First results on the rate of convergence in the convex case are given by Bittner [2] and Eremin [3] for the logarithmic barrier method and for the quadratic penalty method respectively. Poljak [24] derives estimations of the local rate of convergence for quadratic penalties in nonconvex problems by means of the implicit function theorem. Using continuous parameter imbedding Fiacco/McCormick [4] prove differentiable trajectories of solutions to exist for some specific penalty methods under additional conditions. This leads to the rate of convergence of the related methods. On the base of the Kuhn-Tucker-conditions Mifflin [22] introduces a quite general technique for getting convergence bounds of nonlinear programming algorithms and he applies this in [23] to methods of centers. Another approach using directly parameters of the given optimization problem (1) to estimate the rate of convergence was proposed by Kaplan [14, 15]. Quantitative convergence bounds in methods of exterior centers are derived in [6, 19, 21] by means of different techniques. The rate of convergence of augmented Lagrangian methods has been investigated by Gol'shtejn/Tret'jakov [5], Bertsekas [1], Kort/Bertsekas [17], Rockafellar [25, 26] and Skarin [27] e.g.

Basing on the close relation between solutions of (2) and the behaviour of the problem (1) subject to perturbations in the right hand side of the inequality constraints (for the case of augmented Lagrangians compare [26]) in this paper we derive a general technique for estimating the rate of convergence of sequential unconstrained minimization methods.

In the sequel we only investigate solvable nonlinear programming problems (1) with a closed subset  $X \subseteq R^n$  and continuous functions  $f_i : X \rightarrow R^1$ ,  $i = 0, 1, \dots, m$ . To short our notation set  $v = \begin{pmatrix} v_0 \\ \underline{v} \end{pmatrix}$  with  $\underline{v} = (v_1, \dots, v_m)^T$  for any  $v \in R^{m+1}$  and we define  $f(x) = (f_0(x), f_1(x), \dots, f_m(x))^T$ . Thus,  $\underline{f}(x)$  especially denotes  $\underline{f}(x) = (f_1(x), \dots, f_m(x))$ .

Let be selected an arbitrary set  $Y$  of parameters and a generating function  $E : Y \times R^{m+1} \rightarrow \tilde{R} := R^1 \cup \{ + \infty \}$ . We define a related auxiliary function  $T : X \times Y \rightarrow \bar{R} := \tilde{R} \cup \{ - \infty \}$  by

$$T(x, y) = \inf \{ E(y, v) \mid v \geq f(x) \} \quad \text{for any } x \in X, \quad y \in Y. \quad (3)$$

Now, most of the sequential unconstrained minimization techniques (SUMT) can be represented by the following scheme :

*step 1* : Select a starting point  $y^1 \in Y$ . Set  $k := 1$ .

step 2 : Determine a solution  $x^k$  of the auxiliary problem (2).

step 3 : Compute a new parameter  $y^{k+1} \in Y$  and set  $k := k + 1$ . Go to step 2.

The various special algorithms we get from the general scheme by specifying the parameter set  $Y$ , the generating function  $E$  and by an appropriate selection of the updating rule defining the sequence  $\{y^k\} \subset Y$  of parameters in step 3 of SUMT (see [4, 9, 20] e.g.).

2. DUALITY AND ESTIMATIONS VIA COMPARISON PROBLEMS

Let us define the set

$$Q = \{ v \in R^{m+1} \mid \exists x \in X \text{ with } f(x) \leq v \} \tag{4}$$

characterizing the given optimization problem (1). Now, we introduce a sequence of comparison problems

$$E(y^k, v) \rightarrow \min ! \quad \text{s.t. } v \in Q \tag{5}$$

related to the auxiliary problems (2) of the algorithm under consideration. Let denote  $\chi : R^m \rightarrow \bar{R}$  the primal function (or optimal value function) of the problem (1), that means

$$\chi(u) = \inf \{ f_0(x) \mid x \in X, f(x) \leq u \} \quad \text{for any } u \in R^m .$$

Directly from the definitions of the function  $\chi$  and of the set  $Q$  we get

$$Q \subset \text{epi } \chi \quad \text{and} \quad \bar{Q} = \overline{\text{epi } \chi} . \tag{6}$$

Thus the comparison problem (5) can't be solved directly. However, the close relation between the problems (2) and the problems (5) established in the following lemma forms an effective base for the investigation of sequential unconstrained minimization techniques via the comparison problems (see [7, 8]).

LEMMA 1 : For any parameter  $y \in Y$  the equality

$$\inf \{ T(x, y) \mid x \in X \} = \inf \{ E(y, v) \mid v \in Q \} \tag{7}$$

holds. If there exist some solution  $x(y)$  of the auxiliary problem

$$T(x, y) \rightarrow \min ! \quad \text{s.t. } x \in X \tag{8}$$

and some solution  $v(y)$  of the related problem

$$E(y, v) \rightarrow \min ! \quad \text{s.t.} \quad v \geq f(x(y)) \quad (9)$$

then  $v(y)$  also solves the comparison problem

$$E(y, v) \rightarrow \min ! \quad \text{s.t.} \quad v \in Q. \quad (10)$$

If furthermore  $v(y)$  is  $Q$ -regular, that means

$$v(y) - \varepsilon(1, 0, \dots, 0)^T \notin Q \quad \text{for any} \quad \varepsilon > 0, \quad (11)$$

then  $x(y)$  forms a solution of the perturbed problem

$$f_0(x) \rightarrow \min ! \quad \text{s.t.} \quad x \in X, \quad \underline{f}(x) \leq \underline{v}(y). \quad (12)$$

For any  $y \in Y$  we denote

$$\tau(y) = \sup \left\{ t \in \mathbb{R}^1 \mid E(y(t, 0, \dots, 0)^T) \leq \inf_{v \in Q} E(y, v) \right\}. \quad (13)$$

To develop a general duality theory the set  $Y$  and the function  $E$  are assumed to satisfy the following condition : (V). For any  $y \in Y$  and arbitrary  $v \in Q$  not being  $Q$ -regular the inequality  $E(y, v) > \inf_{w \in Q} E(y, w)$  holds.

As shown in [8] this results in the weak duality estimation

$$\sup \{ \tau(y) \mid y \in Y \} \leq \inf \{ f_0(x) \mid x \in X, \underline{f}(x) \leq 0 \}. \quad (14)$$

Thus the problem

$$\tau(y) \rightarrow \sup ! \quad \text{s.t.} \quad y \in Y \quad (15)$$

can be considered as a dual problem to (1). If especially

$$Y = \mathbb{R}_+^m \quad \text{and} \quad E(y, v) = v_0 + y^T \underline{v}$$

are chosen then the related auxiliary function  $T$  equals the ordinary Lagrangian  $L(x, y) = f_0(x) + y^T \underline{f}(x)$  of (1) and (15) coincides with the well known Lagrange dual problem.

We remark that property (V) automatically holds if the generating function  $E$  has the following structure

$$E(y, v) = v_0 + e(y, \underline{v}) \quad \text{for any} \quad y \in Y, \quad v \in \mathbb{R}^{m+1}. \quad (16)$$

Whereby  $e : Y \times \mathbb{R}^m \rightarrow \tilde{\mathbb{R}}$  denotes a given function.

The aim of this paper consists in constructing a set  $A \subset R^{m+1}$  containing the set  $Q$  and such that the problems (approximated comparison problems)

$$E(y, v) \rightarrow \min ! \quad \text{s.t.} \quad v \in A$$

can be solved explicitly. If similar to  $\tau$  we define

$$\tau_A(y) = \sup \left\{ t \in R^1 \mid E(y, t, 0, \dots, 0)^T \leq \inf_{v \in A} E(y, v) \right\} \quad (17)$$

then  $Q \subset A$  results in

$$\tau_A(y) \leq \tau(y) \quad \text{for any} \quad y \in Y. \quad (18)$$

LEMMA 2 : Let  $(x^*, u^*)$  be a saddle point of the Lagrangian  $L$  related to (I). Then the set

$$Q^* = \{ v \in R^{m+1} \mid v_0 + u^{*T} v \geq \chi(0) \}$$

contains the characteristic set  $Q$ .

Proof : Let denote  $\underline{L}(u) = \inf_{x \in X} L(x, u)$  and  $\bar{L}(x) = \sup_{u \in R_+^m} L(x, u)$ . Since  $(x^*, u^*)$  forms a saddle point of the Lagrangian  $L$  we get

$$\bar{L}(x^*) = \underline{L}(u^*). \quad (19)$$

Furthermore the equality

$$\chi(0) = f_0(x^*) = \bar{L}(x^*) \quad (20)$$

holds (see [9] e.g.). Let denote

$$K(x, u, w) = \begin{cases} w_0 + u^T w, & \text{if } w \geq f(x) \\ + \infty, & \text{otherwise.} \end{cases}$$

Then we get

$$L(x, u) = \inf_{w \in R^{m+1}} K(x, u, w) \quad \text{for any } x \in X, \quad u \in R_+^m.$$

This results in

$$\begin{aligned} \underline{L}(u^*) &= \inf_{x \in X} L(x, u^*) = \inf_{x \in X} \inf_{w \in R^{m+1}} K(x, u^*, w) = \inf_{w \in R^{m+1}} \inf_{x \in X} K(x, u^*, w) \\ &= \inf_{\underline{w} \in R^m} \inf_{x \in X, f(x) \leq \underline{w}} \{ f_0(x) + u^{*T} \underline{w} \} \\ &= \inf_{\underline{w} \in R^m} \{ \chi(\underline{w}) + u^{*T} \underline{w} \} \leq \chi(\underline{v}) + u^{*T} \underline{v} \quad \text{for any } \underline{v} \in R^m. \end{aligned} \quad (21)$$

Let be  $v \in Q$ . Using (6) we get  $\chi(\underline{v}) \leq v_0$ . With (19)-(21) this leads to the wanted inequality

$$\chi(0) \leq v_0 + u^{*T} \underline{v} \quad \text{for any } v \in Q. \quad \blacksquare$$

As a trivial consequence of lemma 2 we get the well known estimation

$$\chi(0) \leq f_0(x) + u^{*T} \underline{f}(x) \quad \text{for any } x \in X. \quad (22)$$

If the functions  $E(y, \cdot)$  are convex for any fixed  $y \in Y$  and (1) forms a convex programming problem, i.e. the set  $X$  and the functions  $f_i$ ,  $i = 0, 1, \dots, m$  are supposed to be convex, then to each solution  $v^k$  of the comparison problem (5) a related  $t^k \in R_+^{m+1}$  exists such that

$$t^{kT}(v - v^k) \geq 0 \quad \text{for any } v \in Q.$$

With  $v = \begin{pmatrix} \chi(0) \\ 0 \end{pmatrix} \in Q$  this results in

$$t_0^k(\chi(0) - v_0^k) \geq \underline{t}^{kT} \underline{v}^k.$$

Moreover, if  $t_0^k \neq 0$  then we get the lower bound

$$\chi(0) \geq v_0^k + \frac{1}{t_0^k} \underline{t}^{kT} \underline{v}^k \quad (23)$$

for the optimal value  $\chi(0)$  of the given problem (1). The inequality (23) extends the duality bounds known from penalty methods ([2, 4] e.g.) to more general sequential unconstrained minimization methods. We remark that the estimations proposed by Mifflin [22] are closely related to the inequalities (22), (23).

Now, we proceed in getting the rate of convergence for some specific methods by explicitly solving the approximated comparison problems with  $A = Q^*$ .

3. PENALTY METHODS

In this chapter we derive convergence bounds for some special penalty techniques. Thereby a penalty method is characterized by an explicitly given sequence  $\{ y^k \} \subset Y$  of parameters and the typical penalty property

$$\lim_{k \rightarrow \infty} T(x, y^k) = \begin{cases} 0, & \text{if } \underline{f}(x) < 0 \\ + \infty, & \text{if } \underline{f}(x) \not\leq 0. \end{cases}$$

For further properties and details of penalty methods, especially general convergence theorems, the interested reader is referred to [4, 9] e.g.

In the sequel in our paper we suppose the Lagrangian related to the primal problem (1) to possess a saddle point  $(x^*, u^*) \in X \times R_+^m$ . Basing on the lemmata 1, 2 and on the relations (16)-(18), now, we underestimate the generalized dual value  $\tau(y^k)$  in some methods.

First let us investigate the  $p$ -th order loss function generated by the function

$$E(y, v) = v_0 + \sum_{i=1}^m y_i |v_i|^p, \quad p > 1. \tag{24}$$

Thereby we set  $Y = \text{int } R_+^m$  and  $p$  denotes some fixed parameter. From (3) we get the related auxiliary function

$$T(x, y) = f_0(x) + \sum_{i=1}^m y_i \max^p \{ 0, f_i(x) \}. \tag{25}$$

**THEOREM 1 :** *Let be defined  $T$  by (25). Then for any  $y \in Y$  the inequalities*

$$\chi(0) - (p - 1) \sum_{i=1}^m y_i^{\frac{1}{1-p}} \left( \frac{u_i^*}{p} \right)^q \leq \inf_{x \in X} T(x, y) \leq \chi(0) \tag{26}$$

with  $\frac{1}{p} + \frac{1}{q} = 1$  hold.

*Proof :* Due to lemma 1, the definition (13) of the dual function and (24) we have

$$\tau(y) = \inf_{x \in X} T(x, y) \quad \text{for any } y \in Y. \tag{27}$$

Furthermore  $E$  possesses the structure (16). Thus we get the right inequality of (26) from (27) and the weak duality estimation (14).



Lemmata 1 and 2 lead to

$$\inf_{v \in Q^*} E(y, v) \leq \inf_{v \in Q} E(y, v) = \inf_{x \in X} T(x, y). \tag{28}$$

Now, we solve explicitly the linearly constrained problem

$$E(y, v) \rightarrow \min ! \quad \text{s.t.} \quad v_0 + u^{*T} \underline{v} \geq \chi(0). \tag{29}$$

Due to (24) the optimal value of this problem equals the optimal value of

$$\chi(0) + \sum_{i=1}^m [y_i |v_i|^p - u_i^* v_i] \rightarrow \min ! \quad \text{s.t.} \quad \underline{v} \in R^m. \tag{30}$$

Using the convexity and separability of the objective function we get

$$\tilde{v}_i(y) = \left( \frac{u_i^*}{p y_i} \right)^{\frac{1}{p-1}}, \quad i = 1, \dots, m$$

for the optimal solution  $\underline{\tilde{v}}(y)$  of the problem (30). With

$$\tilde{v}_0(y) = \chi(0) - u^{*T} \underline{\tilde{v}}(y)$$

this leads to

$$\begin{aligned} \inf_{v \in Q^*} E(y, v) &= E(y, \tilde{v}(y)) = \chi(0) + \sum_{i=1}^m \left[ y_i \left( \frac{u_i^*}{p y_i} \right)^{\frac{p}{p-1}} - u_i^* \left( \frac{u_i^*}{p y_i} \right)^{\frac{1}{p-1}} \right] \\ &= \chi(0) - (p - 1) \sum_{i=1}^m y_i^{\frac{1}{1-p}} \left( \frac{u_i^*}{p} \right)^q. \end{aligned}$$

Combining with (28) we get the left inequality in (26). ■

We remark that the estimation

$$\chi(0) - \frac{1}{4r} \|u^*\|^2 \leq \inf_{x \in X} \left\{ f_0(x) + r \sum_{i=1}^m \max^2 \{ 0, f_i(x) \} \right\} \leq \chi(0)$$

given by Eremin [3] is contained in theorem 1 with  $p=2$  and  $y_i=r, i=1, \dots, m$ .

Now, we consider the exponential penalty function

$$T(x, y) = f_0(x) + \sum_{i=1}^m \frac{y_i + m}{y_i} \exp(y_i f_i(x)). \tag{31}$$

$Y = \text{int } R_+^{2m}$  (see [13] e.g.).

**THEOREM 2 :** *Let the auxiliary function  $T$  be given by (31). Then the estimation*

$$\chi(0) + \sum_{i \in I_+} \frac{u_i^*}{y_i} \left[ 1 - \ln \left( \frac{u_i}{y_{i+m}} \right) \right] \leq \inf_{x \in X} T(x, y) \leq \chi(0) + \sum_{i=1}^m \frac{y_{i+m}}{y_i}$$

for any  $y \in Y$  holds. Thereby denotes  $I_+ = \{ i \in \{ 1, \dots, m \} \mid u_i^* > 0 \}$ .

*Proof :* The right inequality is a consequence of the weak duality and

$$\tau(y) = \inf_{x \in X} T(x, y) - \sum_{i=1}^m \frac{y_{i+m}}{y_i}.$$

Similar to the proof of theorem 1 we estimate the optimal value of the problem

$$\chi(0) + \sum_{i=1}^m \left[ \frac{y_{i+m}}{y_i} \exp(y_i v_i) - u_i^* v_i \right] \rightarrow \inf ! \quad \text{s.t.} \quad \underline{v} \in R^m. \quad (32)$$

This can be carried out componentwise. Let be  $u_i^* = 0$ . Then holds

$$\frac{y_{i+m}}{y_i} \exp(y_i v_i) - u_i^* v_i \geq 0 \quad \text{for any} \quad v_i \in R^1.$$

If  $u_i^* > 0$  then the related component of (32) is minimized at

$$\tilde{v}_i(y) = \frac{1}{y_i} \ln \left( \frac{u_i^*}{y_{i+m}} \right).$$

Thus we get the optimal value

$$\chi(0) + \sum_{i \in I_+} \frac{u_i^*}{y_i} \left[ 1 - \ln \left( \frac{u_i^*}{y_{i+m}} \right) \right] \quad \text{and the wanted inequality.} \quad \blacksquare$$

A well known disadvantage of the exponential penalty function (31) consists in the rapid growth of the exponential function. To overcome this Kaplan [16] proposed the function

$$T(x, y) = f_0(x) + \sum_{i=1}^m y_i (f_i(x) + \sqrt{f_i^2(x) + y_{i+m}}) \quad (33)$$

with  $Y = \text{int } R_+^{2m}$ . If additionally the parameters  $y_i, y_{i+m}$  are adjusted according to  $y_{i+m} = y_i^{-2-\theta}$  with some fixed  $\theta \geq 0$  from (33) we get

$$T(x, y) = f_0(x) + \sum_{i=1}^m y_i (f_i(x) + \sqrt{f_i^2(x) + y_i^{-2-\theta}}) \quad (34)$$

the method earlier investigated in [14, 15]. The advantage of (33) consists in avoiding  $y_i, i = 1, \dots, m$  to tend to infinity. This results in a technique being more numerically stable than (34).

**THEOREM 3 :** *Let the auxiliary function be given by (33). Then the estimation*

$$\chi(0) + \sum_{i=1}^m \sqrt{u_i^* y_{i+m}(2 y_i - u_i^*)} \leq \inf_{x \in X} T(x, y) \leq \chi(0) + \sum_{i=1}^m y_i \sqrt{y_{i+m}}$$

holds for any  $y \in Y$  with  $y_i \geq \frac{1}{2} u_i^*, i = 1, \dots, m$ .

*Proof :* Using  $\tau(y) = \inf_{x \in X} T(x, y) - \sum_{i=1}^m y_i \sqrt{y_{i+m}}$  and (14) we get the second inequality. The first one we get by determining the infimal value of

$$\chi(0) + \sum_{i=1}^m [y_i(v_i + \sqrt{v_i^2 + y_{i+m}}) - u_i^* v_i] \rightarrow \inf ! \quad \text{s.t. } \underline{v} \in R^m \quad (35)$$

and the inequality (28).

If  $u_i^* = 0$  then  $y_i(v_i + \sqrt{v_i^2 + y_{i+m}}) - u_i^* v_i \geq 0$  for any  $v_i \in R^1$ . With  $\lim_{v_i \rightarrow -\infty} y_i(v_i + \sqrt{v_i^2 + y_{i+m}}) = 0$  this leads to

$$\inf_{v_i \in R^1} [y_i(v_i + \sqrt{v_i^2 + y_{i+m}}) - u_i^* v_i] = 0 \quad (36)$$

in this case.

Now, let be  $y_i > \frac{1}{2} u_i^* > 0$ . Then differentiating the  $i$ -th component of the objective function in (35) we get the necessary and due to the convexity also sufficient condition

$$y_i \left( 1 + \frac{\tilde{v}_i(y)}{\sqrt{\tilde{v}_i^2(y) + y_{i+m}}} \right) = u_i^*$$

for the related minimizer  $\tilde{v}_i(y)$ . This results in

$$\tilde{v}_i(y) = (u_i^* - y_i) \sqrt{\frac{y_{i+m}}{2 y_i - u_i^*}} \quad (37)$$

and

$$\min_{v_i \in R^1} [y_i(v_i + \sqrt{v_i^2 + y_{i+m}}) - u_i^* v_i] = \sqrt{u_i^* y_{i+m}(2 y_i - u_i^*)}. \quad (38)$$

With  $E(y, v) = v_0 + \sum_{i=1}^m y_i(v_i + \sqrt{v_i^2 + y_{i+m}})$  and (36), (38) we get

$$\inf_{v \in Q^*} E(y, v) = \chi(0) + \sum_{i=1}^m \sqrt{u_i^* y_{i+m}(2 y_i - u_i^*)}.$$

Using (28) this results in the wanted inequality. ■

Now, let us discuss the influence of the parameter  $y \in Y$  in the auxiliary function (33) and in the related duality estimations. If the parameter sequence  $\{y^k\} \subset Y$  is selected such that

$$\lim_{k \rightarrow \infty} y_i^k = +\infty \quad \text{and} \quad \lim_{k \rightarrow \infty} y_i^k \sqrt{y_{i+m}^k} = 0, \quad i = 1, \dots, m$$

then the sequence  $\{T(x, y^k)\}$  uniformly approximates the linear loss penalty function

$$f_0(x) + \sum_{i=1}^m 2 y_i^k \max \{0, f_i(x)\}$$

in the sense that the difference uniformly tends to zero. On the other hand it is well known that any solution of an auxiliary problem

$$f_0(x) + \sum_{i=1}^m \tilde{y}_i \max \{0, f_i(x)\} \rightarrow \min ! \quad \text{s.t.} \quad x \in X \tag{39}$$

also solves the nonlinear programming problem (1) if  $\tilde{y}_i > u_i^*$ ,  $i = 1, \dots, m$  holds (see [9] e.g.). Now, if we select

$$y_i^k = \frac{1}{2} \tilde{y}_i \quad \text{and} \quad \lim_{k \rightarrow \infty} y_{i+m}^k = 0, \quad i = 1, \dots, m$$

then the method (2), (33) approximates the exact penalty technique (39). A more general approach to the approximation of (39) including also regularization techniques was given in [16].

Now, let us investigate the quasi-barrier method proposed by Hamala [11]. As the parameter set  $Y$  we choose  $Y = \text{int } R_+^m$  and the generating function  $E$  is defined by

$$E(y, v) = \begin{cases} v_0 - \sum_{i=1}^m y_i(-v_i)^p, & \text{if } \underline{v} \leq 0 \\ +\infty, & \text{otherwise.} \end{cases} \tag{40}$$

Thereby  $p \in (0, 1)$  denotes some fixed parameter. Then according to (3) we get the related auxiliary function

$$T(x, y) = \begin{cases} f_0(x) - \sum_{i=1}^m y_i (-f_i(x))^p, & \text{if } f(x) \leq 0 \\ + \infty, & \text{otherwise.} \end{cases}$$

**THEOREM 4 :** *Let be  $T$  defined via (3), (40) and let  $u^* > 0$ . Then the estimation*

$$\chi(0) + \frac{p-1}{p} p^{\frac{1}{1-p}} \sum_{i=1}^m \left( \frac{y_i}{u_i^{*p}} \right)^{\frac{1}{1-p}} \leq \inf_{x \in X} T(x, y) \leq \chi(0)$$

for any  $y \in Y$  holds.

*Proof :* The point  $\tilde{v}_i(y) = - \left( \frac{p y_i}{u_i^*} \right)^{\frac{1}{1-p}}$  minimizes the function  $\varphi_i(v_i) = - u_i^* v_i - y_i (-v_i)^p$  subject to  $v_i \in (-\infty, 0]$ . Similary to the proofs of the previous theorems with

$$\varphi_i(\tilde{v}_i(y)) = \frac{p-1}{p} p^{\frac{1}{1-p}} \left( \frac{y_i}{u_i^{*p}} \right)^{\frac{1}{1-p}}$$

we get the inequality stated above. ■

*Remark :* Since  $-(-\sigma)^p$  is not bounded from below for  $\sigma \leq 0$  we used the condition  $u^* > 0$  to guarantee the boundness of  $\inf_{x \in X} T(x, y)$ . The condition  $u^* > 0$  can be relaxed if the set  $X$  is bounded e.g. Because of  $\text{dom } \chi \subset \{ u \geq \hat{u} \}$  with  $\hat{u}_i = \min_{x \in X} f_i(x), i = 1, \dots, m$  we have

$$\chi(0) - \sum_{i \in I_+} y_i \hat{u}_i^p + \frac{p-1}{p} p^{\frac{1}{1-p}} \sum_{i \in I_+} \left( \frac{y_i}{u_i^{*p}} \right)^{\frac{1}{1-p}} \leq \min_{x \in X} T(x, y) \leq \chi(0)$$

in this case.

Let us consider the barrier technique generated by

$$E(y, v) = \begin{cases} v_0 + \sum_{i=1}^m y_i (-v_i)^{-p}, & \text{if } v < 0 \\ + \infty, & \text{otherwise,} \end{cases} \tag{41}$$

$Y = \text{int } R_+^m$  and  $p > 0$  fixed. Because of  $e(y, 0) = +\infty$  the dual value

$\tau(y)$  equals  $-\infty$  and can't generate any upper bound to  $\inf_{x \in X} T(x, y)$ . Similary to the previous theorems some lower bounds can be proved by means of the lemmata 1, 2.

**THEOREM 5 :** *Let the auxiliary function  $T$  be defined by (3), (41). Then the estimation*

$$\chi(0) + \frac{p + 1}{p} p^{\frac{1}{p+1}} \sum_{i=1}^m u_i^{*\frac{1}{p+1}} y_i^{\frac{1}{p+1}} \leq \inf_{x \in X} T(x, y)$$

for any  $y \in Y$  holds.

In the method generated by (41) as well as in other barrier methods the technique presented in chapter 2 fails to give upper bounds for the value  $\inf_{x \in X} T(x, y)$ . A possible way to overcome these troubles consists in the application of appropriate perturbations. Similary to (13) we define for any  $y \in Y, u \in \mathbb{R}^m$  the value

$$\hat{\tau}(y, u) = \sup \left\{ t \in \mathbb{R}^1 \mid E\left(y, \begin{pmatrix} t \\ u \end{pmatrix}\right) \leq \inf_{v \in Q} E(y, v) \right\}.$$

The function  $\hat{\tau}$  generalizes the dual function  $\tau$  such that

$$\tau(y) = \hat{\tau}(y, 0) \quad \text{for any } y \in Y$$

holds.

Let  $E$  possess the structure (16) and let exist some  $u \in \mathbb{R}^m, u < 0$  with  $\chi(u) < +\infty$ . Then, from lemma 1 we get the estimation

$$\hat{\tau}(y, u) \leq \chi(u)$$

and furthermore

$$\hat{\tau}(y, u) = \inf_{x \in X} T(x, y) - e(y, u).$$

This results in the upper bound

$$\inf_{x \in X} T(x, y) \leq \chi(u) + e(y, u).$$

This inequality can be useful applied to get convergence bounds as exemplified in the following theorem.

**THEOREM 5a :** *Let the given optimization problem be convex and let exist some  $\tilde{x} \in X$  with  $f(\tilde{x}) < 0$ . Let the auxiliary function  $T$  be defined by (3), (41). Then for any  $\varepsilon \in (0, 1]$ ,  $y \in Y$  the estimation*

$$\begin{aligned} \chi(0) + \frac{p+1}{p} p^{\frac{1}{p+1}} \sum_{i=1}^m u_i^{*\frac{p}{p+1}} y_i^{\frac{1}{p+1}} &\leq \inf_{x \in X} T(x, y) \leq \\ &\leq (1 - \varepsilon) \chi(0) + \varepsilon f_0(\tilde{x}) + \sum_{i=1}^m y_i \varepsilon^{-p} (-f_i(\tilde{x}))^{-p}. \end{aligned}$$

If we chose  $y_i^k = \sigma_k \rightarrow +0$ ,  $i = 1, \dots, m$  and  $\varepsilon_k = \sigma_k^\alpha$  with some  $\alpha > 0$  then optimal asymptotic bounds we get if  $\max \{ \alpha, 1 - \alpha p \}$  is maximal. This holds for  $\alpha = \frac{1}{p+1}$ .

In a similar way the theorems 2, 3 can be refined and we get also asymptotic bounds.

Up to now in this chapter we only estimated the generalized dual value  $\tau(y)$  or the infimal value  $\inf_{x \in X} T(x, y)$ . Now, we outline a way to get also bounds for the value  $f_0(x(y))$  of the objective function  $f_0$  of (1) at the minimizers  $x(y)$  of the auxiliary problems

$$T(x, y) \rightarrow \min ! \quad \text{s.t. } x \in X .$$

Let us assume that the set  $X$  and the functions  $f_i$ ,  $i = 0, 1, \dots, m$  are convex. Furthermore let be the generating function  $E$  of the typ (16) with some function  $e$  being convex and differentiable with respect to  $\underline{v}$  on its effective domain for any fixed parameter  $y \in Y$ .

Let denote  $v^k$  an optimal solution of the comparison problem

$$E(y^k, v) \rightarrow \min ! \quad \text{s.t. } v \in Q$$

and we set

$$w^k = \nabla_v e(y^k, \underline{v}^k) . \tag{42}$$

From the necessary optimality condition we get

$$t^{kT}(v - v^k) \geq 0 \quad \text{for any } v \in Q$$

with  $t^k = \begin{pmatrix} 1 \\ w^k \end{pmatrix}$ . Using (23) this results in

$$\chi(0) \geq f_0(x^k) + w^{kT} \underline{v}^k, \quad k = 1, 2, \dots . \tag{43}$$

Due to the convexity of  $e(y, \cdot)$  and (42) the point  $\underline{v}^k$  solves the unconstrained problem

$$e(y^k, \underline{v}) - w^{kT} \underline{v} \rightarrow \min ! \quad \text{s.t.} \quad \underline{v} \in R^m. \tag{44}$$

Similary for fixed  $w \in \text{int } R_+^m$  we consider a sequence of problems

$$e(y^k, \underline{v}) - w^T \underline{v} \rightarrow \min ! \quad \text{s.t.} \quad \underline{v} \in R^m. \tag{45}$$

We assume that the properties of the function  $e$  guarantee its solvability and denote by  $\underline{v}^k(w)$  a related solution. If some monotonicity of  $w^T \underline{v}^k(w)$  is known then  $w^{kT} \underline{v}^k$  can be replaced by  $w^T \underline{v}^k(w)$  in (43) to get upper bounds of  $f_0(x^k)$ .

As an example we investigate the penalty function (34). The related function  $e$  is given by

$$e(y, \underline{v}) = \sum_{i=1}^m y_i (v_i + \sqrt{v_i^2 + y_i^{-2-\theta}}), \quad \theta \geq 0.$$

Let be  $w \in \text{int } R_+^m$  fixed then from (37) with  $y_{i+m} = y_i^{-2-\theta}$ ,  $i = 1, \dots, m$  we get the solution  $\underline{v}^k(w)$  of (45) from

$$v_i^k(w) = (w_i - y_i^k) [w_i(2 y_i^k - w_i)]^{-\frac{1}{2}} (y_i^k)^{-1-\frac{\theta}{2}}, \quad i = 1, \dots, m$$

if  $k$  is large enough. This results in

$$- w^T \underline{v}^k(w) = \sum_{i=1}^m (y_i^k)^{-\frac{1+\theta}{2}} w_i^{\frac{1}{2}} \left(1 - \frac{w_i}{y_i^k}\right) \left(2 - \frac{w_i}{y_i^k}\right)^{-\frac{1}{2}}. \tag{46}$$

We remark that this equality also holds if  $w_i = 0$  for some  $i$  and the  $v_i^k(w)$  denotes an arbitrary real number.

Any accumulation point of the sequence  $\{x^k, w^k\}$  can be shown (see [9] e.g.) to be a saddle point of the Lagrangian related to (1). Let be  $\{x^k, w^k\}$  bounded. Without loss of generality we can assume that  $\lim_{k \rightarrow \infty} w^k = u^*$  holds.

Let denote  $w(\epsilon) \in \text{int } R_+^m$  the vector defined by

$$w_i(\epsilon) = u_i^* + \epsilon, \quad i = 1, \dots, m$$

where  $\epsilon > 0$  is an arbitrary positive number. According to (46) and  $\lim_{k \rightarrow \infty} y_i^k = +\infty$  some index  $k_1(\epsilon)$  exists such that

$$- w(\epsilon)^T \underline{v}^k(w(\epsilon)) \geq - w^T \underline{v}^k(w) \quad \text{for any} \quad k \geq k_1(\epsilon) \quad \text{and} \quad 0 \leq w \leq w(\epsilon) \tag{47}$$



holds. From  $\lim_{k \rightarrow \infty} w^k = u^*$  and  $w(\varepsilon) > u^*$  we get an integer  $k_2(\varepsilon)$  with

$$w^k \leq w(\varepsilon) \quad \text{for any } k \geq k_2(\varepsilon).$$

Using (43), (46), (47) this results in

$$f_0(x^k) - \chi(0) \leq \sum_{i=1}^m \frac{(y_i^k)^{-\frac{1+\theta}{2}} (u_i^* + \varepsilon)^{\frac{1}{2}} \left(1 - \frac{u_i^* + \varepsilon}{y_i^k}\right)}{\left(2 - \frac{u_i^* + \varepsilon}{y_i^k}\right)^{1/2}}$$

for any  $k \geq \max \{k_1, k_2\}$ .

Now, the continuity of the right hand side with respect to  $\varepsilon \rightarrow 0$  leads to

$$\overline{\lim}_{k \rightarrow \infty} r_k^{\frac{1+\theta}{2}} [f_0(x^k) - \chi(0)] \leq \frac{1}{\sqrt{2}} \sum_{i=1}^m \sqrt{u_i^*}$$

where  $r_k = \min_{1 \leq i \leq m} y_i^k$  denotes.

To get asymptotic lower bounds we remember that  $v^k \in Q$ ,  $k = 1, 2, \dots$  due to lemma 1 holds. Using lemma 2 we get

$$f_0(x^k) - \chi(0) \geq -u^{*T} \underline{v}^k \quad \text{for any } k$$

and with (46) the inequalities

$$f_0(x^k) - \chi(0) \geq \sum_{i=1}^m \frac{(y_i^k)^{-\frac{1+\theta}{2}} u_i^{*\frac{1}{2}} \left(1 - \frac{u_i^*}{y_i^k}\right)}{\left(2 - \frac{u_i^*}{y_i^k}\right)^{\frac{1}{2}}}, \quad k = 1, 2, \dots$$

This results in

$$\lim_{k \rightarrow \infty} s_k^{\frac{1+\theta}{2}} [f_0(x^k) - \chi(0)] \geq \frac{1}{\sqrt{2}} \sum_{i=1}^m \sqrt{u_i^*}$$

where denotes  $s_k = \max_{1 \leq i \leq m} y_i^k$ .

The technique applied above to get convergence bounds for the sequence  $\{f_0(x^k)\}$  related to the function (34) can be used in the same way to establish the rate of convergence of other penalty methods. For the function (25) as well as for the function defined by (3), (41) this has been done in [8].

4. PARAMETER SELECTION IN REGULARIZED PENALTY TECHNIQUES

A possible modification of penalty methods to improve the numerical stability as well as to force the convergence consists in introducing an additional regularization term in the sense of Tihonov [28]. The principle structure of the generated subproblems can be described by

$$T(x, y^k) + p_k \| x \|^2 \rightarrow \min ! \quad \text{s.t. } x \in X \tag{48}$$

where  $T$  denotes an auxiliary function of penalty typ as considered in chapter 3. Regularized subproblems (48) in penalty methods are considered in [10], [12], [13], [18], [29] e.g. To get convergence results the penalty parameters  $y^k$  and the regularization parameters  $p_k$  are to be adjusted in an appropriate manner. In this chapter we apply the convergence bounds for  $\left\{ \inf_{x \in X} T(x, y^k) \right\}$  proved in 3 to derive parameter selection rules for regularized methods. In the following theorem some condition using dual informations to control the sequences  $\{ y^k \}$ ,  $\{ p_k \}$  will be given.

**THEOREM 6 :** *Let  $T$  denote an auxiliary function generated by means of a function  $E$  of the typ (16) and let denote  $\{ y^k \}$  a sequence such that the properties of a penalty technique (see 4, 9 e.g.) and  $e(y^k, 0) < + \infty$  hold. Furthermore let be  $\lim_{k \rightarrow \infty} p_k = 0$  and*

$$\lim_{k \rightarrow \infty} \frac{1}{p_k} [\chi(0) - \tau(y^k)] = 0. \tag{49}$$

*Then any sequence  $\{ z^k \}$  of solutions of the regularized problems (48) is bounded and each accumulation point of  $\{ z^k \}$  solves (1).*

*Proof :* Let be  $z^k$  some solution of (48). Then the inequalities

$$\begin{aligned} \inf_{x \in X} T(x, y^k) + p_k \| z^k \|^2 &\leq T(z^k, y^k) + p_k \| z^k \|^2 \\ &\leq T(x, y^k) + p_k \| x \|^2 \quad \text{for any } x \in X \end{aligned} \tag{50}$$

hold. Especially with  $x = x^*$  where  $x^*$  denotes an arbitrary solution of (1) we get

$$\inf_{x \in X} T(x, y^k) + p_k \| z^k \|^2 \leq T(x^*, y^k) + p_k \| x^* \|^2. \tag{51}$$

Since  $x^*$  forms a solution of (1) we have  $f(x^*) \leq \begin{pmatrix} \chi(0) \\ 0 \end{pmatrix}$ .

Now, (3) leads to

$$T(x^*, y^k) = \inf_{v \geq f(x^*)} E(y^k, v) \leq \chi(0) + e(y^k, 0). \quad (52)$$

Furthermore (7), (13) and (16) result in

$$\tau(y^k) = \inf_{x \in X} T(x, y^k) - e(y^k, 0). \quad (53)$$

Combining (51)-(53) we get

$$\|z^k\|^2 \leq \|x^*\|^2 + \frac{1}{p_k} [\chi(0) - \tau(y^k)] \quad \text{for any } k.$$

Now, by condition (49) the sequence  $\{z^k\}$  is bounded and each of the related accumulation point  $z^*$  satisfies

$$\|z^*\| \leq \|x^*\|. \quad (54)$$

Because of  $\{T(\cdot, y^k)\}$  forms a sequence of penalty functions and  $\lim_{k \rightarrow \infty} p_k = 0$  holds the sequence  $\{F_k\}$  with

$$F_k(x) = T(x, y^k) + p_k \|x\|^2 \quad \text{for any } x \in X, \quad k = 1, 2, \dots$$

also forms a penalty function sequence (see [4], [9] e.g.). Thus each accumulation point  $z^*$  of  $\{z^k\}$  solves the nonlinear programming problem (1). Because of (54) the point  $z^*$  forms a norm minimal solution of (1) since  $x^*$  denoted an arbitrary solution of (1). ■

*Remarks :* If (1) forms a convex programming problem then the related norm minimal solution (Euclidean norm)  $x^N$  is unique. With theorem 6 this results in

$$\lim_{k \rightarrow \infty} z^k = x^N.$$

Furthermore the existence and uniqueness of the solutions  $z^k$  of the regularized subproblems (48) is guaranteed if the functions  $T(\cdot, y^k)$  are convex and lower semicontinuous.

If in convex programming the functions  $T(\cdot, y^k)$  are convex and differentiable then the stopping criterion

$$\|\nabla_x T(z_e^k, y^k) + 2p_k z_e^k\| \leq \varepsilon_k, \quad k = 1, 2, \dots$$

can be used to determine approximate solutions  $z_\varepsilon^k$  of the subproblems (48). Choosing  $\{\varepsilon_k\}$  such that  $\varepsilon_k > 0$  and  $\lim_{k \rightarrow \infty} \frac{\varepsilon_k}{p_k} = 0$  hold then by the strong convexity of  $F_k$  we get

$$\lim_{k \rightarrow \infty} \|z_\varepsilon^k - z^k\| = 0$$

and therefore  $\lim_{k \rightarrow \infty} z_\varepsilon^k = x^N$  holds.

On the base of duality bounds given in chapter 3 we derive parameter selection rules for some specific penalty methods.

Let us consider the auxiliary function (24) being generated by means of  $E(y, v) = v_0 + \sum_{i=1}^m y_i |v_i|^p, p > 1$ . Then (26), (27) result in

$$0 \leq \chi(0) - \tau(y) \leq (p - 1) \sum_{i=1}^m y_i^{\frac{1}{1-p}} \left(\frac{u_i^*}{p}\right)^q \tag{55}$$

with  $\frac{1}{p} + \frac{1}{q} = 1$ .

Let denote  $r_k = \min_{1 \leq i \leq m} y_i^k$ . Using (55) we get

$$\lim_{k \rightarrow \infty} p_k r_k^{q-1} = +\infty$$

as a condition being sufficient for (49). This is just the same condition as in [30, theorem 3] derived there directly without duality bounds.

Now, let us consider the exponential penalties defined in (31). Using theorem 2 and

$$\tau(y) = \inf_{x \in X} T(x, y) - \sum_{i=1}^m \frac{y_{i+m}}{y_i}$$

in this case we get

$$0 \leq \chi(0) - \tau(y) \leq \sum_{i=1}^m \frac{y_{i+m}}{y_i} - \sum_{i \in I_+} \frac{u_i^*}{y_i} \left[ 1 - \ln \left( \frac{u_i^*}{y_{i+m}} \right) \right]. \tag{56}$$

Similar to [13] let us set

$$y_i^k = t_k^2; \quad y_{i+m}^k = s_k t_k^2, \quad i = 1, \dots, m \tag{57}$$

with some sequences  $\{s_k\}$ ,  $\{t_k\}$  satisfying

$$s_k > 0, \quad t_k > 0, \quad s_k t_k \geq 1 \quad \text{for } k = 1, 2, \dots \quad (58)$$

and  $\lim_{k \rightarrow \infty} t_k = +\infty$ ,  $\lim_{k \rightarrow \infty} s_k \sqrt{t_k} = 0$ .

If we select

$$p_k = s_k \sqrt{t_k}, \quad k = 1, 2, \dots \quad (59)$$

then condition (49) holds. Indeed, (57)-(59) result in

$$\lim_{k \rightarrow \infty} \frac{y_{i+m}^k}{p_k y_i^k} = 0, \quad \lim_{k \rightarrow \infty} p_k y_i^k = +\infty \quad \text{and} \quad \lim_{k \rightarrow \infty} \frac{\ln y_{i+m}^k}{p_k y_i^k} = 0,$$

$i = 1, \dots, m$ . With (56) this guarantees (49) to hold.

It is to remark that (59) differs from the rule proposed in [13] because of relaxing the condition

$$\{x \in X \mid \underline{f}(x) < 0\} \neq \emptyset. \quad (60)$$

On the base of (50) and (60) the condition (49) can be replaced by

$$\lim_{k \rightarrow \infty} \frac{1}{p_k} \left[ T(\tilde{x}^k, y^k) - \inf_{x \in X} T(x, y^k) \right] = 0. \quad (61)$$

Thereby  $\{\tilde{x}^k\} \subset X$  denotes an arbitrary but appropriate sequence with  $\underline{f}(\tilde{x}^k) < 0$ ,  $k = 1, 2, \dots$  and  $\lim_{k \rightarrow \infty} \tilde{x}^k = x^N$ .

Let the set  $X$  and the functions  $f_i$  be convex. Furthermore let  $\tilde{x} \in X$  denote a point with  $\underline{f}(\tilde{x}) < 0$ . Now, we define  $\{\tilde{v}^k\}$ ,  $\{\tilde{x}^k\}$  by

$$\begin{aligned} \tilde{v}^k &= \lambda_k f(\tilde{x}) + (1 - \lambda_k) \begin{pmatrix} \chi(0) \\ 0 \end{pmatrix}, \quad k = 1, 2, \dots \\ \tilde{x}^k &= \lambda_k \tilde{x} + (1 - \lambda_k) x^N \end{aligned}$$

where  $\{\lambda_k\} \subset (0, 1]$  denotes some sequence tending to zero.

From  $\begin{pmatrix} \chi(0) \\ 0 \end{pmatrix} \geq f(x^N)$  and the convexity of  $f$  we get

$$\tilde{v}^k \geq f(\tilde{x}^k) \quad \text{for any } k = 1, 2, \dots$$

Now, (3) and (31) lead to

$$\begin{aligned} T(\tilde{x}^k, y^k) &= \inf_{v \geq f(\tilde{x}^k)} E(y^k, v) \leq \inf_{v \geq \tilde{v}^k} E(y^k, v) \\ &= \lambda_k \tilde{v}_0 + (1 - \lambda_k) \chi(0) + \sum_{i=1}^m \frac{y_{i+m}^k}{y_i^k} \exp(y_i^k \lambda_k \tilde{v}_i). \end{aligned}$$

Using (57), (58) and theorem 2 this results in

$$\begin{aligned} 0 &\leq T(\tilde{x}^k, y^k) - \inf_{x \in X} T(x, y^k) \\ &\leq \lambda_k (\tilde{v}_0 - \chi(0)) + \sum_{i=1}^m s_k \exp(t_k^2 \lambda_k \tilde{v}_i) - \sum_{i \in I_+} \frac{u_i^*}{t_k^2} \left[ 1 - \ln \left( \frac{u_i^*}{s_k t_k^2} \right) \right]. \end{aligned}$$

If we choose  $p_k = \frac{s_k}{t_k}$  like in [13] and  $\lambda_k = t_k^{-3/2}$  then (61) holds. Therefore the related regularized method converges.

It is to remark that the approach presented here simplifies the proof of convergence and shows the natural interaction between the rate of convergence and parameter selection rules of regularized techniques.

If the auxiliary function  $T$  is given by (33) then we have

$$\tau(y) = \inf_{x \in X} T(x, y) - \sum_{i=1}^m y_i \sqrt{y_{i+m}} \tag{62}$$

and theorem 3 results in

$$0 \leq \chi(0) - \tau(y) \leq \sum_{i=1}^m [y_i \sqrt{y_{i+m}} - \sqrt{u_i^* y_{i+m} (2 y_i - u_i^*)}].$$

Let be  $y_i^k = r_k$  and  $y_{i+m}^k = r_k^{-2-\theta}$ ,  $i = 1, \dots, m$ ;  $k = 1, 2, \dots$  with some sequence  $\{ r_k \}$  of positive reals tending to infinity. Then condition (49) can be forced by

$$\lim_{k \rightarrow \infty} p_k r_k^{\theta/2} = + \infty.$$

Similarity to the exponential penalties on the base of (61) the parameter selection rule can be refined if the Slater condition (60) holds. In the same way regularized barrier techniques can be derived from theorem 5.

## 5. METHODS OF CENTERS

Let the nonlinear programming problem (1) be convex and let the Slater condition (60) hold. We consider the methods of centers generated by the functions

$$E(y, v) = \begin{cases} \sum_{i=0}^m (y_i - v_i)^{-p}, & \text{if } v < y, \\ +\infty, & \text{otherwise,} \end{cases} \quad p > 0 \quad (63)$$

or

$$E(y, v) = \begin{cases} -\sum_{i=0}^m \ln(y_i - v_i), & \text{if } v < y \\ +\infty, & \text{otherwise} \end{cases} \quad (64)$$

respectively.

Thereby denotes  $Y = \{y \in R^{m+1} \mid \underline{y} = 0, y_0 > \chi(0)\}$ .

Starting with an arbitrary  $y^1 \in Y$  in the method of centers (see [9], [20], e.g.) the sequences  $\{x^k\}$  and  $\{y^k\}$  are mutually generated according to

$$y_0^{k+1} = f_0(x^k), \quad k = 1, 2, \dots \quad (65)$$

whereby  $x^k$  denotes some solution of the related auxiliary problem (2).

**THEOREM 7 :** *Let  $\{x^k, y^k\}$  denote a sequence generated by a method of centers with the function (63) or (64) and let  $\{x^k\}$  converge to  $x^*$ . Furthermore let the optimal Lagrange multiplier  $u^*$  at  $x^*$  be unique. Then holds*

$$\lim_{k \rightarrow \infty} \frac{y_0^{k+1} - \chi(0)}{y_0^k - \chi(0)} = \frac{\sum_{i=1}^m (u_i^*)^{\frac{p}{p+1}}}{1 + \sum_{i=1}^m (u_i^*)^{\frac{p}{p+1}}}$$

or

$$\frac{p^+}{1 + p^+} \leq \liminf_{k \rightarrow \infty} \frac{y_0^{k+1} - \chi(0)}{y_0^k - \chi(0)} \leq \overline{\lim}_{k \rightarrow \infty} \frac{y_0^{k+1} - \chi(0)}{y_0^k - \chi(0)} \leq \frac{p^0}{1 + p^0}$$

with  $p^+ = \text{card } I_+$ ,  $p^0 = \text{card } \{i \mid f_i(x^*) = 0\}$  if the functions (63) or (64) respectively are used.

*Proof :* Due to the strict convexity and monotonicity of  $E(y, \cdot)$  and due to lemma 1 we get

$$v^k = f(x^k), \quad k = 1, 2, \dots \quad (66)$$

for the solutions  $v^k$  of the related comparison problems (5). The properties of  $E$  result in  $f(x^k) < y^k, k = 1, 2, \dots$

Let us define

$$t^k = \nabla_v E(y^k, v^k), \quad k = 1, 2, \dots \tag{67}$$

Then  $t^k \neq 0$  for any  $k$  holds. We set  $w^k = \frac{1}{t^k} t^k, k = 1, 2, \dots$

Now, using (66), (67) we get

$$w_i^k = \frac{[y_0^k - f_0(x^k)]^{p+1}}{[-f_i(x^k)]^{p+1}}, \quad i = 1, \dots, m; \quad k = 1, 2, \dots$$

and this leads to

$$-f_i(x^k) = [y_0^k - f_0(x^k)] (w_i^k)^{-\frac{1}{p+1}}, \quad i = 1, \dots, m; \quad k = 1, 2, \dots \tag{68}$$

Thereby the function (64) is included with  $p = 0$ . According to (22), (23) the inequalities

$$-w^{kT} \underline{f}(x^k) \leq f_0(x^k) - \chi(0) \leq -u^{*T} \underline{f}(x^k), \quad k = 1, 2, \dots$$

hold. With (68) we get

$$\begin{aligned} [y_0^k - \chi(0) + \chi(0) - f_0(x^k)] \sum_{i=1}^m (w_i^k)^{\frac{p}{p+1}} &\leq f_0(x^k) - \chi(0) \leq \\ &\leq [y_0^k - \chi(0) + \chi(0) - f_0(x^k)] \sum_{i \in I_+} u_i^* (w_i^k)^{-\frac{1}{p+1}}, \quad k = 1, 2, \dots \end{aligned}$$

Due to (65) this results in

$$\frac{\sum_{i=1}^m (w_i^k)^{\frac{p}{p+1}}}{1 + \sum_{i=1}^m (w_i^k)^{\frac{p}{p+1}}} \leq \frac{y_0^{k+1} - \chi(0)}{y_0^k - \chi(0)} \leq \frac{\sum_{i \in I_+} u_i^* (w_i^k)^{-\frac{1}{p+1}}}{1 + \sum_{i \in I_+} u_i^* (w_i^k)^{-\frac{1}{p+1}}}, \quad k=1, 2, \dots \tag{69}$$

From the theory of the methods of centers we know  $x^*$  to form a solution of (1) the uniqueness of  $u^*$  leads to

$$\lim_{k \rightarrow \infty} w^k = u^*$$

(see [4] e.g.). Now, (69) proves the wanted inequalities. ■



*Remarks* : The idea used in the proof of theorem 7 is similar to [22], [23] and shows the close relation between the inequalities (22), (23) and the estimations given by MIFFLIN.

If the optimal multiplier  $u^*$  is not unique the theorem 7 remains valid with  $u_i^* = \overline{\lim}_{k \rightarrow \infty} w_i^k$ ,  $i = 1, \dots, m$ . Assumption (6) guarantees  $\{w^k\}$  to be bounded.

## 6. CONCLUDING REMARKS

In this paper we sketched the technique of deriving convergence bounds *via* comparison problems and related estimations. Here we investigated the penalty methods and the methods of centers more in detail. Basing on the same idea convergence bounds of the methods of exterior centers (see [6]) as well as of the augmented Lagrangian methods (see [8]) are available also. Furthermore, starting from the close relation between sequential unconstrained minimization techniques and the behaviour of the optimal value of the primal problem subject to perturbations in the right hand side of the inequality constraints new concepts of updating rules for the parameters in augmented Lagrangian methods can be derived such that the related method superlinearly converges.

It should be mentioned that computational results showed a good coincidence between the theoretical convergence bounds and numerical test results (compare [9]).

If the user is interested in the inequalities derived in the chapters 3 and 5 from the quantitative point of view and not only qualitative then the optimal Lagrange multipliers play an essential role. In general these multipliers are not available. By means of the sequential unconstrained minimization techniques approximations of the Lagrange multipliers are generated and the magnitude of the multipliers can be estimated. In convex programming problems satisfying the Slater-condition (60) upper bounds of each component of the Lagrange multipliers are available. Basing on this estimations we get a close relation to the convergence bounds derived by Kaplan [14, 15].

Finally let us remark that the principle sketched in this paper can be applied to more general problems also, for instance in partially ordered Hilbert-spaces (for augmented Lagrangians see [26]). In this case, however the componentwise optimization used in chapter 3 e.g. to solve the approximated comparison problems has to be replaced by the investigation of the generating functional  $E$  under one linear inequality constraint.

## REFERENCES

- [1] D. P. BERTSEKAS, *On penalty and multiplier methods for constrained minimization*, SIAM J. Control Optimization 14 (1976) 2, 216-235.
- [2] L. BITTNER, *Eine Verallgemeinerung des Verfahrens des logarithmischen Potentials von Frisch für nichtlineare Optimierungsprobleme*, Aus : Prékopa, A. (ed.) : Colloque on appl. math. to econ. Budapest, 1965, 43-53.
- [3] I. I. EREMIN, *Metod strofov v vypuklom programirovanii*, Dokl. AN SSSR 173 (1967) 4, 748-751.
- [4] A. V. FIACCO, G. P. MCCORMICK, *Nonlinear programming : sequential unconstrained minimization techniques*, Wiley, New York, 1968.
- [5] E. G. GOL'STEJN, N. V. TRET-JAKOV, *Modificirovannye funkcii Lagranža*. Ekon. Mat. Metody 10 (1974) 3, 568-691.
- [6] C. GROßMANN, *Rates of convergence in methods of exterior centers*, Math. OF Statist. Optimization 9 (1978) 3, 373-388.
- [7] C. GROßMANN, *Common properties of nonlinear programming algorithms basing on sequential unconstrained minimizations*. Proc. 7-th Sum. School « Nonlin. Anal », 1979. Akademie Verlag, 1981, 107-117.
- [8] C. GROßMANN, *Sensitivitätsanalyse als einheitlicher Zugang zu den Verfahren der sukzessiven unrestringierten Minimierung*, Diss. B, TU Dresden, 1979.
- [9] C. GROßMANN, A. A. KAPLAN, *Strafmethoden und modifizierte Lagrangefunktionen in der nichtlinearen Optimierung*, Teubner-Text, Teubner Verlag, Leipzig, 1979.
- [10] C. GROßMANN, G. SCHÖNIGER, *Sensitivität und Anwendbarkeit von Straf-Barrieremethoden*, ZAMM 57 (1977), 255-264.
- [11] M. HAMALA, *A general approach to interior point methods with linear parameter for mathematical programming*, TRITA-MAT-1978-20, Royal Inst. Techn., Stockholm.
- [12] J. HARTUNG, *A stable interior penalty method for convex extremal problems*, Numer. Math. 29 (1978), 149-158.
- [13] J. HARTUNG, *On exponential penalty function methods*, Math. OF Statist. Optimization.
- [14] A. A. KAPLAN, *O skorosti shodimosti metoda strofov*, Dokl. AN SSSR 229 (1976) 2, 288-291.
- [15] A. A. KAPLAN, *K analizu nekotoryh ocenok skorosti shodimosti metoda štrafov*, Optimizacija 21 (1978).
- [16] A. A. KAPLAN, *Methods of nonlinear programming applied to solving variational inequalities*, Internat. Conf. « Mathematical Programming », Eisenach, 1980.
- [17] B. W. KORT, D. P. BERTSEKAS, *Combined primal-dual methods for convex programming*, SIAM J. Control Optimization 14 (1976), 268-294.
- [18] M. KOVAČ, *O regularizaciji nekorrektnyh ekstrema'nyh zadač s ispol'zovaniem metoda bar'ernyh funkcii*, Čisl. Analiz. Budapest, 1978, 62-78.
- [19] S. KURCZYUSZ, J. MAJCHRZAK, *Morrison-type algorithms for constrained optimization*.
- [20] F. A. LOOTSMA, *Boundary properties of penalty functions for constrained minimization*, Philips Res. Rept., Suppl. 3, 1970.

- [21] F. A. LOOTSMA, *Convergence rates of quadratic exterior penalty function methods for solving constrained minimization problems*, Philips Res. Rept. 29 (1974), 1-12.
- [22] R. MIFFLIN, *Convergence bounds for nonlinear programming algorithms*, Math. Programming 8 (1975) 3, 251-271.
- [23] R. MIFFLIN, *Rates of convergence for a method of centers algorithm*, JOTA 18 (1976) 2, 199-228.
- [24] B. T. POLJAK, *O skorosti shodivosti metoda štrafnih funkcij*, ŽBMMF 11 (1971) 1, 3-11.
- [25] R. T. ROCKAFELLAR, *A dual approach to solving nonlinear programming problems by unconstrained optimization*, Math. Programming 5 (1973), 354-373.
- [26] R. T. ROCKAFELLAR, *Augmented Lagrangians and applications of the proximal point algorithm in convex programming*, Math. Operations Res. 1 (1976) 2, 97-116.
- [27] V. D. SKARIN, *Ob odnoj modifikaciji metoda štrafnih funkcij v vypoklom programmirovanii*, Trudy Instituta Mat. Meh. Sverdlovsk 5 (1973), 51-62.
- [28] A. N. TIHONOV, F. P. VASIL'EV, *Metody rešenija nekorrektnyh ekstremal'nyh zadač*, Banach Centers Publications, Vol. 3, Warszawa, 1978, 197-342.
- [29] F. P. VASIL'EV, *O regularizacii nekorrektnyh ekstremal'nyh zadač*, Dokl. AN SSSR 241 (1978) 5, 1001-1004.
- [30] F. P. VASIL'EV, M. KOVAČ, *O regularizacii nekorrektnyh ekstremal'nyh zadač v sočetanii so štrafnymi funkcijami obščego vida*, Problemy Vyčisl. Mat. i Sistemnogo Programirovanija, Budapest, 1980, 19-41.