

## PENALTY/BARRIER PATH-FOLLOWING IN LINEARLY CONSTRAINED OPTIMIZATION

CHRISTIAN GROSSMANN

*Institute of Numerical Mathematics*  
*Dresden University of Technology, D-01062 Dresden, Germany*

**e-mail:** grossm@math.tu-dresden.de

### Abstract

In the present paper rather general penalty/barrier path-following methods (e.g. with  $p$ -th power penalties, logarithmic barriers, SUMT, exponential penalties) applied to linearly constrained convex optimization problems are studied. In particular, unlike in previous studies [1, 11], here simultaneously different types of penalty/barrier embeddings are included. Together with the assumed 2nd order sufficient optimality conditions this required a significant change in proving the local existence of some continuously differentiable primal and dual path related to these methods. In contrast to standard penalty/barrier investigations in the considered path-following algorithms only one Newton step is applied to the generated auxiliary problems. As a foundation of convergence analysis the radius of convergence of Newton's method depending on the penalty/barrier parameter is estimated. There are established parameter selection rules which guarantee the overall convergence of the considered path-following penalty/barrier techniques.

**Keywords:** penalty/barrier, interior point methods, convex optimization.

**1991 Mathematics Subject Classification:** 49M15, 65K10, 65H10, 90C25.

## 1 Introduction

In this paper we investigate the local convergence of a general penalty/barrier path-following Newton applied to linearly constrained optimization problems

$$(1) \quad \begin{aligned} & f(x) \rightarrow \min ! \\ \text{subject to} \quad & x \in G := \{ x \in \mathbb{R}^n : a_i^T x \leq b_i, i = 1, \dots, m \}. \end{aligned}$$

Here  $a_i \in \mathbb{R}^n$  and  $b_i \in \mathbb{R}$ ,  $i = 1, \dots, m$  denote given vectors and constants, respectively, describing the linear constraint functions  $g_i(x) := a_i^T x - b_i$ ,  $i = 1, \dots, m$  of the problem. Moreover, it is assumed that the objective  $f$  is a twice continuously differentiable function.

The existence of a local path plays an essential role in penalty/barrier path-following methods. In difference to [1] here we study this task for strict local solutions of a nonlinear optimization problem with twice continuously differentiable constraint functions  $g_i$ ,  $i = 1, \dots, m$ . The main part of our study is dedicated to these more general problems. Only the direct analysis of the path-following method itself is restricted to linearly constrained problems (1).

Let  $x^*$  denote a local solution of (1) and let be satisfied the well known linear independence constraint qualification. Hence, a related dual multiplier  $y^* \in \mathbb{R}^m$  exists such that Karush-Kuhn-Tucker conditions

$$(2) \quad \begin{aligned} \nabla_x L(x^*, y^*) &= 0, \\ y^* \geq 0, \quad g(x^*) &\leq 0, \\ y^{*T} g(x^*) &= 0, \end{aligned}$$

hold, where  $L$  is the Lagrangian

$$L(x, y) := f(x) + \sum_{i=1}^m y_i g_i(x), \quad x \in \mathbb{R}^n, y \in \mathbb{R}_+^m.$$

In addition to the supposed linear independence of the gradients we assume strict complementarity

$$(3) \quad y_i^* > 0 \iff g_i(x^*) = 0$$

and

$$(4) \quad \left. \begin{aligned} & w \in \mathbb{R}^n, \quad w \neq 0, \\ & \nabla g_i(x^*)^T w = 0, \quad i \in I_o \end{aligned} \right\} \Rightarrow w^T \nabla_{xx}^2 L(x^*, y^*) w > 0,$$

where  $I_o := I_o(x^*) := \{i \in I : g_i(x^*) = 0\}$  with  $I := \{1, \dots, m\}$ . Thus, second order sufficiency optimality conditions (cf. [9]) are satisfied which

imply that  $x^*$  is the unique minimizer of (1) and that  $x^*$  is stable under perturbations. In addition, we have

$$G^0 \neq \emptyset, \quad \text{where} \quad G^0 := \{x \in \mathbb{R}^n : g_i(x) < 0, i \in I\}.$$

## 2 General penalty/barrier embeddings

Let  $\mathbb{R}_{++} := \{t \in \mathbb{R} : t > 0\}$  and let  $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$ , i.e.,  $\overline{\mathbb{R}}$  denotes the set of reals extended by  $+\infty$ . Using a parametric penalty/barrier functions  $\varphi_i(\cdot, \cdot) : \mathbb{R} \times \mathbb{R}_{++} \rightarrow \overline{\mathbb{R}}$ ,  $i = 1, \dots, m$ , we define related auxiliary problems

$$(5) \quad \begin{aligned} F(x, s) &:= f(x) + \sum_{i=1}^m \varphi_i(g_i(x), s) \rightarrow \min ! \\ \text{s.t.} \quad x &\in B_s := \{x \in \mathbb{R}^n : \varphi_i(g_i(x), s) < +\infty, i = 1, \dots, m\}, \end{aligned}$$

where  $s \in \mathbb{R}_{++}$  denotes a fixed penalty/barrier parameter.

Considered will be functions  $\varphi_i(\cdot, s)$  which for fixed  $s > 0$  are differentiable at any  $t \in \text{dom } \varphi_i(\cdot, s)$  and can be represented by

$$(6) \quad \frac{\partial}{\partial t} \varphi_i(t, s) = \psi_i\left(\frac{t}{s}\right), \quad \forall t \in \text{dom } \varphi_i(\cdot, s), \quad s > 0,$$

with some functions  $\psi_i : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ ,  $\psi_i \not\equiv 0$ ,  $i = 1, \dots, m$  which essentially satisfy:

- (i) The domain of  $\psi_i$  is either  $\mathbb{R}_{++}$  or  $\mathbb{R}$ ;
- (ii)  $\psi_i : \mathbb{R} \rightarrow \overline{\mathbb{R}}$  continuous (in the sense of  $\overline{\mathbb{R}}$ ), and  $\psi_i$  continuously differentiable in  $\text{dom } \psi_i$ ;
- (iii)  $\psi'_i(r) \geq 0$ ,  $\forall r \in \text{dom } \psi_i$ , and  $\lim_{r \rightarrow -\infty} \psi_i(r) = 0$ ;
- (iv)  $\psi'_i$  Lipschitz continuous on  $\text{dom } \psi_i$  in the sense that  $\psi'_i$  is locally Lipschitz continuous

$$|\psi'_i(\rho_1) - \psi'_i(\rho_2)| \leq L_1(r) |\rho_1 - \rho_2|, \quad \forall |\rho_1|, |\rho_2| \leq r, r \in \text{dom } \psi_i,$$

and at  $-\infty$  satisfies

$$|\psi'_i(\rho_1) - \psi'_i(\rho_2)| \leq L_2(r) \left| \frac{1}{\rho_1} - \frac{1}{\rho_2} \right|, \quad \forall \rho_1, \rho_2 \leq r < 0;$$

- (v) Either  $\lim_{r \rightarrow +\infty} \psi_i(r) = +\infty$  or  $y_i^* < \psi_i(0)$  holds;
- (vi)  $\psi_i' \circ \psi_i^{-1}$  isotone and positive on  $\text{int } \psi_i(\mathbb{R})$ ;
- (vii)  $\lim_{r \rightarrow -\infty} r^2 \psi_i'(r)$  exists.

A representation (6) with functions  $\psi_i$  that satisfy the assumed properties can be found for a wide range of penalty/barrier functions  $\varphi_i$  studied in the literature. The same representation (6), but with slightly different properties of  $\psi_i$  has recently been considered within the framework of saddle points of generalized Lagrangians based on penalty/barrier embeddings in [2, 3]. By relation (6) the penalty/barrier functions  $\varphi_i(\cdot, s)$  are well defined, up to an arbitrary function which depends exclusively on  $s$ , by its generating functions  $\psi_i$ . Since the solution of the auxiliary problem (5) is independent of additive constants, we may choose it arbitrarily. If we select this according to

$$(7) \quad \lim_{s \rightarrow 0+} \varphi_i(-1, s) = 0, \quad i = 1, \dots, m,$$

then (6) implies the well known (c.f. [9, 12, 14]) penalty/barrier property

$$(8) \quad \lim_{s \rightarrow 0+} \varphi(t, s) = \begin{cases} 0, & \text{if } t < 0, \\ +\infty, & \text{if } t > 0. \end{cases}$$

Further, we notice that the given property (i) guarantees the set  $B_s$  to be independent of the embedding parameter  $s > 0$ . Hence, in the sequel we will omit this index.

As representative examples of penalty/barrier embeddings included in our paper the following well known types are listed:

- **logarithmic barrier function**

$$(9) \quad \begin{aligned} \varphi(t, s) &= \begin{cases} -s \ln(-t), & \text{if } t < 0, \\ +\infty, & \text{if } t \geq 0, \end{cases} \quad \forall s > 0, \\ \psi(r) &= \begin{cases} |r|^{-1}, & \text{if } r < 0, \\ +\infty, & \text{if } r \geq 0. \end{cases} \end{aligned}$$

- **p-th power barrier function** ( $p > 0$  is a fixed parameter)

$$(10) \quad \begin{aligned} \varphi(t, s) &= \begin{cases} \frac{1}{p} \frac{s^{p+1}}{|t|^p}, & \text{if } t < 0, \\ +\infty, & \text{if } t \geq 0, \end{cases} \quad \forall s > 0, \\ \psi(r) &= \begin{cases} |r|^{-(p+1)}, & \text{if } r < 0, \\ +\infty, & \text{if } r \geq 0. \end{cases} \end{aligned}$$

• **p-th power loss function** ( $p \geq 2$  is a fixed parameter)

$$(11) \quad \varphi(t, s) = \frac{1}{p} s^{-(p-1)} \max^p\{0, t\}, \quad \forall s > 0, \quad \psi(r) = \max^{p-1}\{0, r\}.$$

• **exponential penalty function**

$$(12) \quad \varphi(t, s) = s \exp(t/s), \quad \forall s > 0, \quad \psi(r) = \exp(r).$$

• **smoothed exact penalty function**

$$(13) \quad \begin{aligned} \varphi(t, s) &= c \left( t + \sqrt{t^2 + s^2} \right), \quad \forall s > 0, \\ \psi(r) &= c \left( 1 + \frac{r}{\sqrt{r^2 + 1}} \right), \end{aligned}$$

with an arbitrary fixed  $c > \max y_i^*$ .

In the case of (13) a rough upper bound for the Lagrangian multipliers can be used as the required constant  $c$ . We refer to [12] for further modifications of (13) for the case if such a bound is not available.

Due to the assumed penalty/barrier properties the set  $B \subset \mathbb{R}^n$  is open. Hence, any minimizer  $x(s)$  of the auxiliary problem (5) has to satisfy the first order necessary optimality condition

$$(14) \quad \nabla f(x(s)) + \sum_{i=1}^m \psi_i(g_i(x(s))/s) \nabla g_i(x(s)) = 0.$$

Related to  $x(s) \in B \subset \mathbb{R}^n$  an approximation  $y(s) \in \mathbb{R}_+^m$  of the optimal dual multiplier  $y^*$  can be defined by

$$(15) \quad y_i(s) := \psi_i(g_i(x(s))/s), \quad i = 1, \dots, m; \quad s > 0.$$

Further, we set

$$(16) \quad x(0) := x^*, \quad y(0) := y^*.$$

In the sequel we concentrate our attention to (14) and investigate the behavior of Newton's method applied to it.

Before the convergence of a path-following Newton algorithm is studied, we show that the supposed properties of the generating functions  $\psi_i$  guarantee the local existence and the (right side) differentiability of the primal-dual trajectory  $(x(s), y(s))$  in  $[0, \bar{s}]$  with some  $\bar{s} > 0$ . In difference to [1] here we do not assume that the objective function is strongly convex. Instead, a related theory of perturbations of local solutions  $x^*$  that satisfies second order sufficiency conditions is developed. This difference in assumptions requires a significant change in the analysis of the primal-dual path. So, the main effort in the present paper is dedicated to the study of the behavior of this trajectory. Moreover, unlike in [1] here each of the constraints is treated individually by possibly different penalty/barrier embeddings.

First we provide an essential technical result about the solutions of certain parameter dependent linear systems which will be used later in the estimation of the derivatives of the local primal-dual path as well as in bounding the radius of contraction on Newton's method applied to penalty/barrier problems. A similar question is investigated in [17], but with a rather different approach.

**Lemma 1.** *Let be  $\tilde{Q} \in \mathcal{L}(\mathbb{R}^n)$  symmetric, positive definite and  $\tilde{P} \in \mathcal{L}(\mathbb{R}^s, \mathbb{R}^n)$ ,  $s \leq n$ , full rank matrices, respectively. Then some  $\delta > 0$  exists such that for any  $w \in \mathbb{R}^s$ ,  $\alpha > 0$  and matrices  $Q \in \mathcal{L}(\mathbb{R}^n)$ , symmetric and  $P \in \mathcal{L}(\mathbb{R}^s, \mathbb{R}^n)$ , satisfying*

$$(17) \quad \|\tilde{Q} - Q\| \leq \delta, \quad \|\tilde{P} - P\| \leq \delta$$

*the linear system*

$$(18) \quad (Q + \alpha PP^T)z = \alpha Pw$$

*possesses a unique solution  $z \in \mathbb{R}^n$  which can be estimated by*

$$(19) \quad \|z\| \leq c \|w\|,$$

*where  $c > 0$  is a constant independent of the parameter  $\alpha > 0$ .*

**Proof.** Since  $\tilde{Q}$  is positive definite,  $\delta > 0$  can be chosen such that any symmetric matrix  $Q$  is also positive provided  $\|\tilde{Q} - Q\| \leq \delta$ . In addition, some  $\underline{\gamma}, \bar{\gamma} > 0$  exist with

$$(20) \quad \underline{\gamma} \|v\|^2 \leq v^T Q v \leq \bar{\gamma} \|v\|^2, \quad \forall v \in \mathbb{R}^n.$$

Hence, for any  $\alpha > 0$  the system matrix in (18) is positive definite. Thus (18) possesses a unique solution  $z \in \mathbb{R}^n$ .

Because of the assumed full rank property of  $\tilde{P}$  the constant  $\delta > 0$  can be selected that (20) remains valid and some  $c_1 > 0$  exists with

$$(21) \quad \left. \begin{array}{l} P \in \mathcal{L}(\mathbb{R}^s, \mathbb{R}^n), \\ \|\tilde{P} - P\| \leq \delta \end{array} \right\} \implies \|(P^T P)^{-1} P\| \leq c_1.$$

We split the vector  $w \in \mathbb{R}^s$  according the direct sum  $\mathbb{R}^s = \mathcal{N}(P) \oplus \mathcal{R}(P^T)$  by

$$w = (I - P^T(P^T P)^{-1} P)w + P^T(P^T P)^{-1} P w.$$

Hence, the linear system (18) is equivalent to

$$(22) \quad (Q + \alpha P P^T) z = \alpha P P^T v \quad \text{with} \quad v := (P^T P)^{-1} P w.$$

Since  $Q \in \mathcal{L}(\mathbb{R}^n)$  is symmetric, positive definite and  $P P^T \in \mathcal{L}(\mathbb{R}^n)$  is symmetric, a complete system  $\{u^j\}_{j=1}^n \subset \mathbb{R}^n$  of generalized eigenvectors and eigenvalues  $\{\lambda_j\} \subset \mathbb{R}_+$  exists with

$$(23) \quad P P^T u^j = \lambda_j Q u^j, \quad j = 1, \dots, n$$

and

$$(24) \quad (u^i)^T Q u^j = \delta_{ij}, \quad i, j = 1, \dots, n,$$

where  $\delta_{ij}$  denotes Kronecker's symbol. Obviously,  $\{u^j\}_{j=1}^n \subset \mathbb{R}^n$  forms a basis in  $\mathbb{R}^n$ . Now, we expand the vectors  $v, z$  from system (22) in this basis, i.e.

$$(25) \quad v = \sum_{j=1}^n \nu_j u^j, \quad z = \sum_{j=1}^n \zeta_j u^j,$$

with coefficients  $\nu_j, \zeta_j \in \mathbb{R}$ ,  $j = 1, \dots, n$ . Then (22) – (25) result in

$$(26) \quad \zeta_j = \frac{\alpha \lambda_j}{1 + \alpha \lambda_j} \nu_j, \quad j = 1, \dots, n.$$

By Bessel's equation we have

$$\|z\|_Q^2 := z^T Q z = \sum_{j=1}^n \zeta_j^2, \quad \|v\|_Q^2 := v^T Q v = \sum_{j=1}^n \nu_j^2.$$

Hence, (26) with  $\lambda_j \geq 0$ ,  $j = 1, \dots, n$  yields

$$\|z\|_Q \leq \|v\|_Q,$$

for any  $\alpha > 0$ . Due to (20) the norms  $\|\cdot\|_Q$  are uniformly equivalent to the Euclidean norm and with (21) we obtain

$$(27) \quad \|z\| \leq \sqrt{\frac{\bar{\gamma}}{\underline{\gamma}}} c_1 \|w\|,$$

independently of the parameter  $\alpha > 0$ . ■

**Remark 1.** For the case of a fixed matrix  $P$  which is not full rank an estimate of type (27) independent of  $\alpha > 0$  is still possible by using Moore-Penrose pseudo inverse  $(P^T)^+$  of  $P^T$  instead of its representation  $(P^T P)^{-1} P$  for full rank matrices. This yields

$$(28) \quad \|z\| \leq \sqrt{\frac{\bar{\gamma}}{\underline{\gamma}}} \|(P^T)^+\| \|w\|.$$

However, this estimate is not uniform concerning perturbations of  $P$  as stated in Lemma 1 under the full rank assumption.

**Theorem 1.** *Under the made assumptions exist some  $\bar{s} > 0$ ,  $\delta > 0$  such that for any  $s \in (0, \bar{s})$  the parametric system (14) of nonlinear equations possess unique solutions  $x(s) \in B \cap U_\delta(x^*)$  and with  $y(s)$  related to  $x(s)$  by (15), (16) holds*

$$\lim_{s \rightarrow 0+} (x(s), y(s)) = (x^*, y^*).$$

*With the setting (16) the functions  $x(\cdot)$ ,  $y(\cdot)$  are continuously differentiable in  $(0, \bar{s})$ , possess a right side derivative at  $s = 0$  and these derivatives are bounded for  $s \rightarrow 0+$ .*

**Proof.** Let us consider a perturbed KKT-system

$$(29) \quad \begin{aligned} \nabla f(x) + \sum_{i \in I_0} y_i \nabla g_i(x) &= r, \\ s \psi_i^{-1}(y_i) &= g_i(x), \quad i \in I_0, \end{aligned}$$

with  $s \in \mathbb{R}$  and  $r \in \mathbb{R}^m$ . In particular, the properties of the generating functions  $\psi_i$  and strict complementarity guarantee  $y_i^* \in \text{int } \psi_i(\mathbb{R})$ ,  $i \in I_0$ . Thus  $\psi_i^{-1}(y_i^*)$ ,  $i \in I_0$  is well defined and we have

$$(30) \quad \begin{aligned} \nabla f(x^*) + \sum_{i \in I_0} y_i^* \nabla g_i(x^*) &= 0, \\ 0 \cdot \psi_i^{-1}(y_i^*) &= g_i(x^*), \quad i \in I_0, \end{aligned}$$

i.e. the optimal solution  $x^*$  and its related dual multipliers  $y_i^*$ ,  $i \in I_0$ , satisfy (29) for  $s = 0$ ,  $r = 0$ . Without loss of generality, the constraints can be numbered such that  $I_0 = \{1, \dots, m_0\}$  with some  $m_0 \leq m$ . The assumed second order sufficiency condition (4) implies that the  $(n + m_0, n + m_0)$ -matrix

$$\begin{pmatrix} \nabla^2 f(x^*) + \sum_{i \in I_0} y_i^* \nabla^2 g_i(x^*) & \nabla g_1(x^*) & \nabla g_2(x^*) & \cdot & \cdot & \nabla g_{m_0}(x^*) \\ \nabla g_1(x^*)^T & 0 & 0 & \cdot & \cdot & 0 \\ \nabla g_2(x^*)^T & 0 & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \nabla g_{m_0}(x^*)^T & 0 & 0 & \cdot & \cdot & 0 \end{pmatrix}$$

is regular (cf. [9, 14]). Now, according to the implicit function theorem some  $\delta > 0$ ,  $s_1 > 0$ ,  $\rho > 0$  exist such that for any  $|s| \leq s_1$ ,  $\|r\| \leq \rho$  there exist unique  $x(s, r) \in U_\delta(x^*)$  and  $y_i(s, r) \in U_\delta(y_i^*)$ ,  $i \in I_0$ , which satisfy

$$(31) \quad \begin{aligned} \nabla f(x(s, r)) + \sum_{i \in I_0} y_i(s, r) \nabla g_i(x(s, r)) &= r, \\ s \psi_i^{-1}(y_i(s, r)) &= g_i(x(s, r)), \quad i \in I_0. \end{aligned}$$

Moreover, these  $x(\cdot, \cdot)$  and  $y_i(\cdot, \cdot)$ ,  $i \in I_0$  are differentiable. Since the number of constraints is finite, without loss of generality we can assume that  $\delta > 0$  is selected such that

$$(32) \quad g_i(x) \leq -\sigma, \quad \forall x \in U_\delta(x^*), \quad i \in I \setminus I_0$$

holds with some  $\sigma > 0$ . By means of  $x(r, s)$  for fixed  $s \in (0, s_1]$  we define a mapping  $P : U_\rho(0) \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$  by

$$(33) \quad P r := - \sum_{i \notin I_0} \psi_i \left( \frac{g_i(x(s, r))}{s} \right) \nabla g_i(x(s, r)).$$

The supposed properties of the generating functions  $\psi_i$  and (32) guarantee that some  $s_2 \in (0, s_1]$  exists that for any fixed  $s \in (0, s_2]$  the mapping  $P$  defined by (33) satisfies

$$P r \in U_\rho(0), \quad \forall r \in U_\rho(0)$$

and is contractive on  $U_\rho(0)$ . Hence, for any  $s \in (0, s_2]$  there is a unique fixed point  $r(s) \in U_\rho(0)$  of the operator  $P$ , i.e.

$$r(s) = - \sum_{i \notin I_0} \psi_i \left( \frac{g_i(x(s, r(s)))}{s} \right) \nabla g_i(x(s, r(s))).$$

Let us abbreviate  $x(s) := x(s, r(s))$ ,  $y_i(s) := y_i(s, r(s))$ ,  $i \in I$ , where

$$(34) \quad y_i(s, r) := \psi_i \left( \frac{g_i(x(s, r))}{s} \right), \quad i \notin I_0, \quad s \in (0, s_2].$$

With the definition of  $x(s, r)$  and the properties of  $\psi_i$  we have  $x(s) \in B$ . In addition, (32) yields

$$(35) \quad \lim_{s \rightarrow 0^+} r(s) = 0.$$

With (32), (34) this implies  $\lim_{s \rightarrow 0^+} x(s) = x^*$  and  $\lim_{s \rightarrow 0^+} y(s) = y^*$ .

On the other hand, by (31), (33) we obtain

$$(36) \quad \nabla f(x(s)) + \sum_{i=1}^m \psi_i \left( \frac{g_i(x(s))}{s} \right) \nabla g_i(x(s)) = 0, \quad s \in (0, s_2].$$

With  $x(s) \in B \cap U_\delta(x^*)$  this proves the local existence of the primal path. Now, the local existence of the dual path results from the given implicit function arguments and (34).

Before we continue, let us notice that property (vii) supposed for  $\psi_i$  implies (see [1]) the existence of the limits

$$c_i := \lim_{\rho \rightarrow -\infty} \rho \psi_i(\rho).$$

Now, let us define  $r(0) := 0$  and we study

$$\begin{aligned} \lim_{s \rightarrow 0^+} \frac{r(s) - r(0)}{s - 0} &= - \lim_{s \rightarrow 0^+} \sum_{i \notin I_0} \frac{\psi_i \left( \frac{g_i(x(s))}{s} \right)}{s} \nabla g_i(x(s)) \\ &= - \sum_{i \notin I_0} \nabla g_i(x^*) \lim_{s \rightarrow 0^+} \frac{g_i(x(s))}{s} \psi_i \left( \frac{g_i(x(s))}{s} \right) \frac{1}{g_i(x(s))} \\ &= - \sum_{i \notin I_0} \frac{\nabla g_i(x^*)}{g_i(x^*)} \lim_{\rho \rightarrow -\infty} \rho \psi_i(\rho) = \sum_{i \notin I_0} c_i \frac{\nabla g_i(x^*)}{g_i(x^*)}. \end{aligned}$$

Thus,  $r$  is differentiable from the right at  $s = 0$ . As a consequence the implicit function theorem and chain rule guarantee the differentiability of the functions  $x(s) = x(s, r(s))$ ,  $y_i(s, r(s)) = y_i(s)$ ,  $i \in I_0$ , at  $s = 0$  from the right. As already shown, also  $y_i(\cdot)$ ,  $i \notin I_0$ , are differentiable from the right at  $s = 0$ , namely  $D_+ y_i(0) = 0$ ,  $i \notin I_0$ .

Now, we consider the case  $s \in (0, s_2]$ . Then  $x(s)$ ,  $y(s)$  satisfy the system

$$(37) \quad \begin{aligned} \nabla f(x(s)) + \sum_{i=1}^m y_i(s) \nabla g_i(x(s)) &= 0, \\ \psi_i \left( \frac{g_i(x(s))}{s} \right) - y_i(s) &= 0, \quad i = 1, \dots, m. \end{aligned}$$

Differentiation of system (37) w.r.t. the parameter  $s$  yields

$$(38) \quad H(s) \begin{pmatrix} \dot{x}(s) \\ \dot{y}(s) \end{pmatrix} = \begin{pmatrix} 0 \\ q(s) \end{pmatrix},$$

where  $H(s)$  denotes the  $(n+m, n+m)$ -matrix

$$H(s) := \begin{pmatrix} \nabla^2 f(x(s)) + \sum_{i=1}^m y_i(s) \nabla^2 g_i(x(s)) & \nabla g_1(x(s)) & \nabla g_2(x(s)) & \cdot & \nabla g_m(x(s)) \\ \psi'_1 \left( \frac{g_1(x(s))}{s} \right) \nabla g_1(x(s))^T & -s & 0 & \cdot & 0 \\ \psi'_2 \left( \frac{g_2(x(s))}{s} \right) \nabla g_2(x(s))^T & 0 & -s & \cdot & 0 \\ \ddots & \cdot & \cdot & \cdot & \cdot \\ \psi'_m \left( \frac{g_m(x(s))}{s} \right) \nabla g_m(x(s))^T & 0 & 0 & \cdot & -s \end{pmatrix},$$

$q(s) := (q_1(s), \dots, q_m(s))^T$ , with  $q_i(s) := s^{-1} \psi'_i \left( \frac{g_i(x(s))}{s} \right) g_i(x(s))$ , and  $\dot{x}(s)$ ,  $\dot{y}(s)$  denote the derivatives of  $x(s)$  and  $y(s)$ , respectively. These derivatives exist due to the implicit function theorem, since the matrix  $H(s)$  is regular. Next, we estimate the derivatives  $\dot{x}(s)$ ,  $\dot{y}(s)$  by using the structure and properties of  $H(s)$ . With the abbreviations

$$\begin{aligned}
R &:= \nabla^2 f(x(s)) + \sum_{i=1}^m y_i(s) \nabla^2 g_i(x(s)) = \nabla_{xx}^2 L(x(s), y(s)), \\
T_1 &:= \{\nabla g_1(x(s)), \dots, \nabla g_{m_0}(x(s))\}, \\
T_2 &:= \{\nabla g_{m_0+1}(x(s)), \dots, \nabla g_m(x(s))\}, \\
D_1 &:= \text{diag} \left\{ \psi'_1 \left( \frac{g_1(x(s))}{s} \right), \dots, \psi'_{m_0} \left( \frac{g_{m_0}(x(s))}{s} \right) \right\}, \\
D_2 &:= \text{diag} \left\{ \psi'_{m_0+1} \left( \frac{g_{m_0+1}(x(s))}{s} \right), \dots, \psi'_m \left( \frac{g_m(x(s))}{s} \right) \right\},
\end{aligned}$$

system (38) has the form

$$\begin{aligned}
(39) \quad & R \dot{x} + T_1 \dot{y}^1 + T_2 \dot{y}^2 = 0, \\
& D_1 T_1^T \dot{x} - s \dot{y}^1 = q^1, \\
& D_2 T_2^T \dot{x} - s \dot{y}^2 = q^2.
\end{aligned}$$

Here  $y^1, y^2$  and  $q^1, q^2$  denote the splittings of  $y(s)$  and  $q(s)$  respectively, adjusted to the sub-matrices of  $H(s)$ . Further, the parameter  $s$  has been omitted to simplify the notations. Eliminating  $\dot{y}^1, \dot{y}^2$  from the last equations in the system above we obtain

$$(40) \quad \dot{y}^1 = s^{-1} (D_1 T_1^T \dot{x} - q^1), \quad \dot{y}^2 = s^{-1} (D_2 T_2^T \dot{x} - q^2).$$

Hence, system (39) can be expressed in the reduced form

$$(41) \quad (R + s^{-1} T_1 D_1 T_1^T + s^{-1} T_2 D_2 T_2^T) \dot{x} = p,$$

where

$$(42) \quad p := p^1(s) + p^2(s) \quad \text{with} \quad p^k(s) := s^{-1} T_k q^k, \quad k = 1, 2.$$

With the definitions above we have

$$\begin{aligned}
p^1(s) &= \sum_{i \in I_0} s^{-2} \psi'_i \left( \frac{g_i(x(s))}{s} \right) g_i(x(s)) \nabla g_i(x(s)) \\
p^2(s) &= \sum_{i \notin I_0} s^{-2} \psi'_i \left( \frac{g_i(x(s))}{s} \right) g_i(x(s)) \nabla g_i(x(s)).
\end{aligned}$$

First, we study the behavior of  $\|p^2(s)\|$  for the limit  $s \rightarrow 0+$ . The condition  $\psi'_i(\rho) = O(|\rho|^{-2})$  for  $\rho \rightarrow -\infty$ , the convergence  $x(s) \rightarrow x^*$  and the continuous differentiability of the constraint functions  $g_i$  guarantee that a constant  $c_0 > 0$  and some  $s_0 > 0$  exist with

$$(43) \quad \|p^2(s)\| = \|T_2 s^{-1} q^2\| \leq c_0, \quad \forall s \in (0, s_0].$$

By the same arguments we obtain the boundedness of  $\|s^{-1} T_2 D_2 T_2^T\|$  for  $s \rightarrow 0+$ . So we may select  $s_0, c_0 > 0$  such that also holds

$$(44) \quad \|s^{-1} T_2 D_2 T_2^T\| \leq c_0, \quad \forall s \in (0, s_0].$$

Now, we notice that the diagonal matrices  $D_1$  can be expressed by

$$D_1 = \text{diag} \left( \psi'_i(\psi_i^{-1}(y_i(s))) \right)_{i \in I_0}.$$

Due to  $y(s) \rightarrow y^*$ , the continuity of  $\psi^{-1} \circ \psi$ , property (vi) and the strict complementarity some  $c > 0$  exists with

$$(45) \quad c \|z\|^2 \leq z^T D_1 z, \quad \forall z \in \mathbb{R}^n, \quad s \in (0, s_0].$$

This together with the supposed second order sufficient optimality condition, the continuity of the second order derivatives of the objective  $f$  and of the constraint functions  $g_i$ ,  $i = 1, \dots, m$  and  $(x(s), y(s)) \rightarrow (x^*, y^*)$  for  $s \rightarrow 0+$  guarantee that some  $\beta, \underline{\gamma}, \bar{\gamma} > 0$  and  $s_1 \in (0, s_0]$  exist such that

$$\underline{\gamma} \|z\|^2 \leq z^T (R + s^{-1} T_2 D_2 T_2^T + \beta T_1 D_1 T_1^T) z \leq \bar{\gamma} \|z\|^2, \quad \forall z \in \mathbb{R}^n, \quad s \in (0, s_0]$$

Let denote  $Q := R + s^{-1} T_2 D_2 T_2^T + \beta T_1 D_1 T_1^T$  and  $P := T_1 D_1^{1/2}$ . Then the linear system (41) can be expressed by

$$(Q + (s^{-1} - \beta) P P^T) \dot{x} = s^{-1} P D_1^{-1/2} q^1 + p^2$$

The vector  $p^2$  is bounded for  $s \rightarrow 0+$ , as shown above, and also  $D_1^{-1/2} q^1$  is bounded for  $s \rightarrow 0+$ . Thus, Lemma 1 provides some  $c > 0$  with

$$\|\dot{x}(s)\| \leq c, \quad \forall s \in (0, s_0].$$

By the arguments above, the remaining estimate for  $\|\dot{y}\|$  in the limit  $s \rightarrow 0+$  follows from here and from the representation (40).  $\blacksquare$

Theorem 1 extends the classical result of [9, 14] concerning differentiable paths to a much wider class of penalty/barrier embeddings. Differentiability properties of the primal-dual path play an important role in interior point methods (cf. [15, 18]), in particular in connection with logarithmic barriers, as well as in other primal-dual embeddings (cf. [19]).

As an immediate consequence of Theorem 1 we obtain

**Corollary 1.** *There exist some constants  $s_0 \in (0, \bar{s}]$  and  $c_L > 0$  such that*

$$(46) \quad \left. \begin{aligned} \|x(s) - x(t)\| &\leq c_L |s - t|, \\ \|y(s) - y(t)\| &\leq c_L |s - t|, \end{aligned} \right\} \quad \forall s, t \in [0, s_0].$$

### 3 Approximation of the primal-dual path

Let  $x \in B$  denote some approximation of the solution  $x(s)$  of (14). Related to  $x \in B$  we define vectors  $u(x, s), v(x, s) \in \mathbb{R}_+^m$  by

$$(47) \quad \left. \begin{aligned} u_i(x, s) &:= \psi_i(g_i(x)/s), \\ v_i(x, s) &:= s^{-1} \psi'_i(g_i(x)/s), \end{aligned} \right\} \quad \forall s > 0, \quad i = 1, \dots, m,$$

respectively. In particular, we have  $y(s) = u(x(s), s)$  which approximates the dual multiplier  $y^*$  for  $s \rightarrow 0+$ .

To simplify our further investigations in the sequel we restrict ourselves to linearly constrained optimization problems, i.e. we assume the constraint functions  $g_i$  to have a representation

$$g_i(x) = a_i^T x - b_i, \quad i = 1, \dots, m$$

with vectors  $a_i \in \mathbb{R}^n$  and reals  $b_i, i = 1, \dots, m$ . In particular, this assumption implies  $\nabla g_i(x) = a_i$  and  $\nabla^2 g_i(x) = 0, i = 1, \dots, m$ .

Next, we study the local behavior of Newton's method applied to determine elements of the primal path  $x(\cdot)$ . One step of Newton's method starting from  $x \in B$ , defines a new approximate  $\tilde{x}$  of  $x(s)$  as solution of the linear system

$$(48) \quad \nabla f(x) + \sum_{i=1}^m u_i(x, s) a_i + \left( \nabla^2 f(x) + \sum_{i=1}^m v_i(x, s) a_i a_i^T \right) (\tilde{x} - x) = 0.$$

As shown in the proof of Theorem 1, the system matrix is regular for  $x = x(s)$  provided the penalty/barrier parameter  $s > 0$  is sufficiently small.

By continuous perturbation arguments this remains true close to the primal path. So, the linear system (48) has a unique solution  $\tilde{x} \in \mathbb{R}^n$  if  $\|x - x(s)\|$  and  $s > 0$  are sufficiently small.

However, the nature of penalty/barrier methods makes the auxiliary problems (14) ill-conditioned for small parameters  $s > 0$  (cf. [16, 17, 18]). In particular, problem (14) is asymptotically incorrect for  $s \rightarrow 0+$ . This shrinks the radius of convergence of Newton's method to zero when  $s > 0$  tends to the wanted limit zero. Hence, for path-following Newton methods it is essential to derive bounds for  $\|x - x(s)\|$  which guarantee that  $\tilde{x} \in B$  and to establish an estimate of  $\|\tilde{x} - x(s)\|$  for the new iterate  $\tilde{x}$ . For any fixed penalty/barrier parameter  $s > 0$  Newton's method is locally quadratically convergent. Hence, some functions  $\delta(\cdot), \rho(\cdot) : \mathbb{R}_{++} \rightarrow \mathbb{R}_{++}$  exist such that

$$(49) \quad \|\tilde{x} - x(s)\| \leq \delta(s) \|x - x(s)\|^2, \quad \forall x \in B, \quad \|x - x(s)\| \leq \rho(s).$$

These functions  $\delta, \rho$  can be constructed by means of

**Theorem 2.** *For sufficiently small  $s > 0$  and  $x \in B$  the linear system (48) possesses a unique solution  $\tilde{x} \in \mathbb{R}^n$  and some constants  $s_0, c_\delta, c_\rho > 0$  exist such that*

$$(50) \quad \left. \begin{array}{l} x \in B, s \in (0, s_0], \\ \|x - x(s)\| \leq c_\rho s, \end{array} \right\} \implies \left\{ \begin{array}{l} x, \tilde{x} \in B, \\ \|\tilde{x} - x(s)\| \leq c_\delta s^{-1} \|x - x(s)\|^2. \end{array} \right.$$

The proof of this theorem rests similarly to [1] on the convergence analysis of Newton's method for asymptotically singular problems as occur in penalty/barrier methods (see also [5, 7, 8, 17]). Here we sketch only the essential steps and refer for further details to [1].

Newton's equation together with (14) yields

$$(51) \quad \begin{aligned} & \left( \nabla^2 f(x) + \sum_{i=1}^m v_i(x, s) a_i a_i^T \right) (\tilde{x} - x(s)) = \nabla f(x(s)) - \nabla f(x) - \nabla^2 f(x)(x(s) - x) \\ & + \sum_{i=1}^m \left( (u_i(x(s), s) - u_i(x, s)) a_i - v_i(x, s) a_i a_i^T (x(s) - x) \right). \end{aligned}$$

By Taylor's formula we have

$$(52) \quad \begin{aligned} & \nabla f(x(s)) - \nabla f(x) - \nabla^2 f(x)(x(s) - x) \\ & = \int_0^1 (\nabla^2 f(x + \tau(x(s) - x)) - \nabla^2 f(x)) d\tau (x(s) - x). \end{aligned}$$

Taking into account  $\nabla_x u_i(x, s) = v_i(x, s)a_i$  similarly it holds

$$(53) \quad \begin{aligned} & (u_i(x(s), s) - u_i(x, s))a_i - v_i(x, s) a_i a_i^T (x(s) - x) \\ &= \int_0^1 (v_i(x + \tau(x(s) - x), s) - v_i(x, s)) d\tau a_i a_i^T (x(s) - x), \end{aligned}$$

$i = 1, \dots, m$ . Next, we split the right hand side of (51) into three components  $r_j \in \mathbb{R}^n$ ,  $j = 1, 2, 3$ , according to

$$\begin{aligned} r_1 &:= \int_0^1 (\nabla^2 f(x + \tau(x(s) - x)) - \nabla^2 f(x)) d\tau (x(s) - x), \\ r_2 &:= \sum_{i \notin I_0} \left( \int_0^1 [v_i(x, s) - v_i(x + \tau(x(s) - x), s)] d\tau \right) a_i a_i^T (x - x(s)), \\ r_3 &:= \sum_{i \in I_0} \left( \int_0^1 [v_i(x, s) - v_i(x + \tau(x(s) - x), s)] d\tau \right) a_i a_i^T (x - x(s)), \end{aligned}$$

and we estimate the solutions  $d_j$  of the linear systems

$$(54) \quad \left( \nabla^2 f(x) + \sum_{i=1}^m v_i(x, s) a_i a_i^T \right) d_j = r_j, \quad j = 1, 2, 3,$$

separately. We notice that (54) behaves well for the first two cases  $r_1$  and  $r_2$  while the norm of the last vector  $r_3$  tends to infinity as  $s \rightarrow 0+$ . However, this is partially compensated by growing components  $v_i(x, s)$  which contribute to the system matrix in (54). Lemma 1 provides an adequate tool for this task. For further details of the estimations of the different parts we refer to [1, 11]. A similar approach of separating the right hand side of degenerate linear systems, but using different tools, can be found in [17].

**Remark 2.** The case of nonlinear constraints is considered in [18] for the logarithmic barrier technique. The study of the occurring ill-conditioned linear systems in Newton linearizations require to estimate a further term arising from the varying gradients.

A simple example given in [11] shows that the order of the bounds given in Theorem 2 cannot be improved. Analysis of barrier methods by self-concordance properties (cf. [10, 15, 16]) cover this fact by using norms which are implicitly dependent on the parameter  $s > 0$  and deteriorate for  $s \rightarrow 0+$ .

## 4 Path-following Newton method

The convergence estimates for Newton's method given above together with the Lipschitz continuity of the path  $x(\cdot)$  lay the foundation for a path-following algorithm with only fixed number of Newton steps at each level  $s_k$  of the penalty/barrier parameter  $s > 0$  in system (14).

### Path-following Newton algorithm

**Step 1.** Select parameters  $\varepsilon, \rho > 0, \nu \in (0, 1)$  and  $s_0 > 0$ .  
Select  $x_0 \in B$  such that

$$(55) \quad \|x_0 - x(s_0)\| \leq \rho s_0.$$

Set  $k := 0$ .

**Step 2.** Determine  $d_k \in \mathbb{R}^n$  as solution of the linear system

$$(56) \quad \nabla f(x_k) + \sum_{i=1}^m u_i(x_k, s_k) a_i + \left( \nabla^2 f(x_k) + \sum_{i=1}^m v_i(x_k, s_k) a_i a_i^T \right) d_k = 0$$

and define  $x_{k+1} := x_k + d_k$ .

**Step 3.** If  $s_k \leq \varepsilon$  then stop. Otherwise set  $s_{k+1} := \nu s_k$  and go to **Step 2** with  $k := k + 1$ .

**Remark 3.** In step 1 of the above algorithm a rather good approximation  $x_0 \in B$  of the solution  $x(s_0)$  of the auxiliary penalty/barrier problem is required. However, for relatively large parameters  $s_0 > 0$  the radius of convergence of Newton's method, as a rule, is large enough to enable a determination of  $x_0$  by Newton's method itself. In addition, damping strategies should be included to stabilize further the overall convergence of Newton's method.

**Theorem 3.** *If  $s_0 > 0, \rho > 0$  are sufficiently small and if the parameter  $\nu \in (0, 1)$  is selected such that*

$$(57) \quad c_\delta \nu^{-2} (\rho + c_L (1 - \nu))^2 \leq \rho,$$

*then the given path-following Newton algorithm generates iterates  $x_k \in B, k = 1, 2, \dots$ , which satisfy*

$$(58) \quad \|x_k - x(s_k)\| \leq \rho s_k, \quad k = 0, 1, \dots$$

Furthermore, the algorithm terminates after at most  $k^* := \lceil \ln(\varepsilon/s_0)/\ln(\nu) \rceil$  steps, where  $\lceil t \rceil$  denotes the smallest natural number greater or equal  $t$ , and it holds

$$(59) \quad \|x_{k^*} - x^*\| \leq (c_L + \rho) s_{k^*}.$$

**Proof.** We show (58) by induction.

For  $k = 0$  this inequality holds according to (55). Let it be true for some index  $k \geq 0$ . Since  $0 < \rho \leq c_\rho$  and  $s_k \leq s_0$ , Theorem 2 yields  $x_{k+1} \in B$  and

$$\|x_{k+1} - x(s_k)\|_{C(x_k, s_k)} \leq c_\delta s_k^{-1} \|x_k - x(s_k)\|_{C(x_k, s_k)}^2 \leq c_\delta \bar{\delta}^2 \rho^2 s_k.$$

The triangle inequality and Corollary 1 imply

$$\begin{aligned} \|x_{k+1} - x(s_{k+1})\| &\leq \|x_{k+1} - x(s_k)\| + \|x(s_k) - x(s_{k+1})\| \\ &\leq \underline{\delta}^{-1} \|x_{k+1} - x(s_k)\|_{C(x_k, s_k)} + \|x(s_k) - x(s_{k+1})\| \\ &\leq c_\delta \underline{\delta}^{-1} \bar{\delta}^2 \rho^2 s_k + c_L (s_k - s_{k+1}) \\ &= \nu^{-1} (\underline{\delta}^{-1} \bar{\delta}^2 c_\delta \rho^2 + c_L (1 - \nu)) s_{k+1} \leq \rho s_{k+1}. \end{aligned}$$

This completes the induction.

Since  $s_k = \nu^k s_0$ , we have  $s_k \leq \varepsilon$  for any  $k \geq \ln(\varepsilon/s_0)/\ln(\nu)$ . Finally, Corollary 1 and the local Lipschitz continuity of  $f$  result in

$$0 \leq f(x(s)) - f(x^*) \leq \rho s$$

together with (58) and Corollary 1 we complete proof.  $\blacksquare$

**Remark 4.** We notice that under the made assumption the existence of some  $\mu \in (0, 1)$  which satisfies (57) can be shown (see [11]). However, a drawback of the considered method is that only a fixed multiplier  $\mu \in (0, 1)$  is applied to generate the parameter sequence  $\{s_k\}$ . Further acceleration may be achieved by choosing an appropriate sequence  $\{\mu_k\} \subset (0, 1)$  instead of one fixed parameter  $\mu \in (0, 1)$ .

**Remark 5.** To improve stability properties the given approach can be overlaid by saddle point methods (cf. [2, 3]) as well as by augmented Lagrangian techniques (cf. [12]). The obtained approximations  $v_i(x_k, s_k)$  of Lagrangian multipliers  $y_i^*$  may be efficiently used to speed up convergence as well as to scale automatically the constraints.

### Acknowledgement

The support of this research by Deutsche Forschungsgemeinschaft (DFG) grant GR 1777/2-1 is kindly acknowledged.

### REFERENCES

- [1] D. Al-Mutairi, C. Grossmann and K.T. Vu, *Path-following barrier and penalty methods for linearly constrained problems*, Preprint MATH-NM-10-1998, TU Dresden 1998 (to appear in Optimization).
- [2] A. Auslender, R. Cominetti and M. Haddou, *Asymptotic analysis for penalty and barrier methods in convex and linear programming*, Math. Operations Res. **22** (1997), 43–62.
- [3] A. Ben-Tal and M. Zibulevsky, *Penalty/barrier multiplier methods for convex programming problems*, SIAM J. Optimization **7** (1997), 347–366.
- [4] B. Chen and P.T. Harker, *A noninterior continuation method for quadratic and linear programming*, SIAM J. Optimization **3** (1993), 503–515.
- [5] B. Chen and P.T. Harker, *A continuation method for monotone variational inequalities*, Math. Programming **69** (1995), 237–253.
- [6] R. Cominetti and J.M. Pères-Cerda, *Quadratic rate of convergence for a primal-dual exponential penalty algorithm*, Optimization **39** (1997), 13–32.
- [7] P. Deuffhard and F. Potra, *Asymptotic mesh independence of Newton-Galerkin methods via a refined Mysovskii theorem*, SIAM J. Numer. Anal. **29** (1992), 1395–1412.
- [8] J.-P. Dussault, *Numerical stability and efficiency of penalty algorithms*, SIAM J. Numer. Anal. **32** (1995), 296–317.
- [9] A.V. Fiacco and G.P. McCormick, *Nonlinear programming: Sequential unconstrained minimization techniques*, Wiley, New York 1968.
- [10] R.M. Freund and S. Mizuno, *Interior point methods: current status and future directions*, OPTIMA, Math. Programming Soc. Newsletter **51** (1996).
- [11] C. Grossmann, *Asymptotic analysis of a path-following barrier method for linearly constrained convex problems*, Optimization **45** (1999), 69–88.

- [12] C. Grossmann and A.A. Kaplan, *Straf-Barriere-Verfahren und modifizierte Lagrangefunktionen in der nichtlinearen Optimierung*, Teubner, Leipzig 1979.
- [13] M. Halická and M. Hamala, *Duality transformation functions in the interior point methods*, Acta Math. Univ. Comenianae LXV (1996), 229–245.
- [14] F.A. Lootsma, *Boundary properties of penalty functions for constrained minimization*, Philips Res. Rept. Suppl. **3** 1970.
- [15] S. Mizuno, F. Jarre and J. Stoer, *A unified approach to infeasible-interior-point algorithms via geometric linear complementarity problems*, Appl. Math. Optimization **33** (1996), 315–341.
- [16] Yu. Nesterov and A. Nemirovskii, *Interior-point polynomial algorithms in convex programming*, SIAM Studies in Appl. Math., Philadelphia 1994.
- [17] M.H. Wright, *Ill-conditioning and computational error in interior methods for nonlinear programming*, SIAM J. Opt. **9** (1998), 84–111.
- [18] S.J. Wright, *On the convergence of the Newton/Log-barrier method*, Preprint ANL/MCS-P681-0897, MCS Division, Argonne National Lab., August 1997 (revised version 1999).
- [19] H. Yamashita and H. Yabe, *An interior point method with a primal-dual  $l_2$  barrier penalty function for nonlinear optimization*, Working paper, March 1999, University of Tokyo.

Received 13 November 1999

Revised 26 January 2000