4400 University Drive
Fairfax, Virginia 22030

George Mason University

**Modified Interior Distance Functions
(Theory and Methods)**

**Roman A. Polyak
Department of Operations Research
George Mason University
Fairfax, Virginia 22030**

Department of Operations Research
School of Information Technology
and Engineering

Modified Interior Distance Functions
(Theory and Methods)

Roman A. Polyak
Department of Operations Research
George Mason University
Fairfax, Virginia 22030

Modified Interior Distance Functions
(Theory and Methods)

Roman A. Polyak[*]
Department of Operations Research
George Mason University, Fairfax VA 22030
email: rpolyak@vmsi.gmu.edu

Abstract. In this paper we introduced and developed the theory of Modified Interior Distance Functions(MIDFs).

The MIDF is a Classical Lagrangian (CL) for a constrained optimization problem, which is equivalent to the initial one and can be obtained from the latter by monotone transformation both the objective function and constraints.

In contrast to the Interior Distance Functions (IDFs), which played a fundamental role in Interior Point Methods (IPMs), the MIDFs are defined on an extended feasible set and along with center, have two extra tools, which control the computational process: the barrier parameter and the vector of Lagrange multipliers.

The extra tools allow to attach to the MIDFs very important properties of Augmented Lagrangeans. One can consider the MIDFs as Interior Augmented Lagrangeans. It makes MIDFs similar in spirit to Modified Barrier Functions (MBFs), although there is a fundamental difference between them both in theory and methods.

Based on MIDFs theory, Modified Center Methods (MCMs) have been developed and analyzed.

The MCMs find an unconstrained minimizer in primal space and update the Lagrange multipliers, while both the center and the barrier parameter can be fixed or updated at each step.

The MCMs convergence was investigated, and their rate of convergence was estimated. The extension of the feasible set and the special role of the Lagrange multipliers allow to develop MCMs, which produce, in case of nondegenerate constrained optimization, a primal and dual sequences that converge to the primal-dual solutions with linear rate, even when both the center and the barrier parameter are fixed. Moreover, every Lagrange multipliers update shrinks the distance to the primal dual solution by a factor $0 < \gamma < 1$ which can be made as small as one wants by choosing a fixed interior point as a "center" and a fixed but large enough barrier parameter.

The numerical realization of MCM leads to the Newton MCM (NMCM). The approximation for the primal minimizer one finds by Newton Method followed by the Lagrange multipliers update.

Due to the MCM convergence, when both the center and the barrier parameter are fixed, the condition of the MIDF Hessians and the neighborhood of the primal minimizer where Newton method is "well" defined remains stable.

It contributes to both the complexity and the numerical stability of the NMCM.

---

1. Introduction.

In the mid 60s, P. Huard [BuiH66], [Huar67a] and [Huar67b] introduced Interior Distance Functions (IDFs) and developed Interior Center Methods (ICMs) for solving constrained optimization problems. Later these functions, as well as Interior Center Methods, were intensively studied by A. Fiacco and G. McCormick [FiacM68], K. Grossman and A. Kaplan [GrosK81], R. Mifflin [Miff76], and E. Polak [PolE71], just to mention a few.

It was found [PolE71] that there are close connections not only between the IDFs and the Barrier Functions [FiacM68], but also between ICMs and methods of feasible directions [Zout60], [ZPP63].

The ICMs consist of finding at each step a central (in a sense) point of the Relaxation Feasible Set (RFS) and updating it in accordance with the objective function level that has already been attained. The RFS is the intersection of the feasible set with the Relaxation (level) set of the objective function at the attained level. In the Classical ICM the "center" is sought as a minimum of the IDF.

Interest in the IDFs, as well as in the Barrier Functions (BFs), grew dramatically in connection with the well known developments in mathematical programming during the last ten years since N. Karmarkar published his projective scaling method [Kar84]. In fact, his potential function is an IDF and his method is a Center Method, which produces centers of spheres, which belong to the feasible polytop. The concept of centers has a long and interesting history.

In the 60s, concurrently with P. Huard's ICM, the Gravity Center Method was independently developed by A. Levin [Lev65] and D. Newman [New65], the Affine Scaling (ellipsoid centers) Method by I. Dikin [Dik67], and the Chebyshev Center Method by S. Zuchovitsky, R. Polyak and M. Pimak [ZPP69]. The Affine Scaling Method, which one can view as a method of feasible direction with special direction normalization, was rediscovered in 1986 independently by E. Barnes [Barn86] and R. Vanderbei, M. Maketon and B. Freedman [VanMF86] as a simplified version of Karmarkar's method.

In the 70s, N. Shor [Sh70] and independently D. Yudin and A. Nemirovsky [YuN76] developed the ellipsoid method, which generates centers of ellipsoids with minimal volume circumscribed around some convex sets. Using this method, L. Khachiyan [Kh79] was the first to prove in 1979 the polynomial complexity of the Linear Programming problem. His result had a great impact on the complexity theory, but numerically the ellipsoid method appeared to be not efficient. It is interesting to note that the rate of convergence, which was established by I. Dikin [Dik74] for the Affine Scaling Method, in case of nondegenerate linear programming problems, is asymptotically much better than the rate of convergence of the ellipsoid method and numerically, as it turned out, the Affine Scaling Method is much more efficient [AdRVK89].

The concept of centers became extremely popular in the 80s. Centering and reducing the cost are two basic ideas that are behind the developments in the Interior Point Methods (IPMs) for the last ten years. Centering means to stay away from the boundary. A successful answer to the main question: how far from the boundary one should stay, was given by Sonnevend [Son85] (see also [JarSS88]) through the definition of the analytic center of a polytop. The analytic center is a unique minimizer of the Interior Distance Function. The central path - curve, which is formed by the analytic centers, plays a very important role in the IPM developments. It was brilliantly shown in the paper by C. Gonzaga [Gon92].

Following the central path J. Renegar [Ren88] obtained the first path-following algorithm

with $O(\sqrt{n}L)$ number of iterations against $O(nL)$ of the N. Karmarkar's method.

Soon afterwards, C. Gonzaga [Gon88] and P. Vaidya [Vaid87] described algorithms based on the centering ideas with overall complexity $O(n^3L)$ arithmetic operations, which is the best known result so far.

In the course of the 30 years history of center methods it became clear that both the theoretical importance and the practical efficiency of the center type methods depends very much on the "quality" of the center and on the cost to compute the center or its approximation.

The center was and still is the main tool to control the computational process in a wide variety of center methods in general and in IPMs in particular.

However, still there is a fundamental question, which has to be answered: **how consistent the main idea of center methods - to stay away from the boundary with the main purpose of constrained optimization - to find a solution on the boundary.**

In this paper we will try to address this issue. The purpose of this paper is to introduce the Modified Interior Distance Functions (MIDFs) and to develop their theory. Based on this theory, we are going to develop the Modified Center Methods (MCMs), to investigate their convergence and to establish their rate of convergence.

The MIDFs are particular realizations of the Nonlinear Rescaling Principle [Pol86], which consists of transforming a constrained optimization problem into an equivalent one and using the Classical Lagrangean for the equivalent problem for both theoretical analysis and numerical methods. In the case of MIDFs, we transform both the objective functions and the constraints by monotone transformations. The constraints transformation is parametrized by a positive parameter. The MIDF, which is a Classical Lagrangean for the equivalent problem, has properties that make it substantially different from both IDF as well as Classical Lagrangean for the initial problem.

Instead of one tool (the centers), which controls the process in the IDF, the MIDF has three tools: the center, the barrier parameter and the vector of Lagrange multipliers. Two extra tools provide the MIDF with very important properties.

The barrier parameter not only allows to retain the convexity of the MIDFs when both the objective function and the constraints are convex, it also allows to "convexify" the MIDFs in the case when the objective function and/or the constraints are not convex but the second order optimality conditions are satisfied. The barrier parameter is also crucial for the rate of convergence of the MCMs.

The other critical extra tool is the vector of Lagrange multipliers. It allows to attach to the MIDFs nice local properties of Augmented Lagrangeans [Ber82], [GoIT89], [Hes69], [Man75], [Pow69], [Rock74] and provides them with new important features.

One can consider MIDFs as Interior Augmented Lagrangeans. However, in addition to the nice local Augmented Lagrangean properties, the MIDFs possess important global properties, which manifests itself when the Lagrange multipliers are fixed and one changes the barrier parameter and/or the center to approach the solution.

What is most important, the MIDFs are defined and keep smoothness of the order of the initial functions on the extension of the feasible set.

The special MIDF's properties allows to develop MCMs, which produce the primal-dual sequences that converge to the primal-dual solution, even when both the center and the barrier parameter are fixed. Moreover, under nondegeneracy assumptions the primal and dual sequences

converge to the primal-dual solution with linear rate.

So the main driver in MCMs is the vector of Lagrange multipliers rather than the center or the barrier parameter. It allows not only to stay as close to the boundary as one wants, it makes possible for the primal minimizer to be even outside of the feasible set. At the same time, in constrast to the IDF, the solution of the constrained optimization problem is always inside of the level set of the MIDF.

So at the final stage, the MCMs are closer to the multipliers methods rather than to IPM. As far as the initial stage is concerned, when the vector of Lagrange multipliers is fixed, then one can take advantage of the global-self concordance [NesN94] properties of the IDF, which guarantee a very reasonable convergence by following the central path, which one obtains by changing the barrier parameter or the center.

The numerical realization of the MCMs leads to the Newton MCM. In contrast to the IDF, the MIDF Hessian not only exists on an extended feasible set but due to the MCMs convergence, when both the center and the barrier parameter are fixed, the condition number of the MIDF Hessian is stable and so the neighborhoods of the primal minimizers, where the Newton method is "well" defined [Sm86].

It contributes to substantial reduction of the number of Newton steps per Lagrange multipliers update from step to step. Every update shrinks the distance to the primal-dual solution by a fixed factor $0 < \gamma < 1$, which depends on the input data and the size of the problem. It can be made as small as one wants even when both the "center" and the barrier parameter are fixed, but the parameter is large enough.

The paper is organized as follows. After the statement of the problem, we discuss the IDF's properties and introduce the MIDFs. Then we establish the basic MIDF's properties at the primal-dual solution and compare them with the correspondent IDF's properties. Then we prove the basic theorem, which is the foundation for the MCMs and their convergence. We describe the MCMs and analyze their convergence.

The MIDFs have some common features with Modified Barrier Functions [Pol92], but there are fundamental differences between them as well. We illustrate the differences using a few small examples.

Then we describe the Newton MCM, which is a numerical realization of the MCM.

We conclude the paper by considering dual problems, that are based on MIDF. They have some distinctive features, which we will discuss briefly.

4

1. Problem Formulation and Basic Assumptions. Let $f_0(x)$ and $-f_i(x)$, $i = 1,...,m$ be convex, C^2-function in $\mathbb{R}^n$ and there exists

$$x^* = \text{argmin } \{f_0(x) / x \in \Omega \} \qquad (1.1)$$

where $\Omega = \{x : f_i(x) \geq 0, \ i = 1,...,m\}$.

We will assume that Slater condition holds, i.e.

$$\exists \ x^0 : f_i(x^0) > 0, \ i = 1,...,m \qquad (1.2)$$

So the Karush-Kuhn-Tucker's (K-K-T's) optimality conditions hold true, i.e. there exists a vector $u^* = (u_1^*,...,u_m^*) \geq 0^m$ such that

$$L'_x(x^*u^*) = f'_0(x^*) - \sum_{i=1}^{m} u_i^* f'_i(x^*) = 0^n, \ f_i(x^*)u_i^* = 0, \ i = 1,...,m, \qquad (1.3)$$

where $L(x,u) = f_0(x) - \sum_{i=1}^{m} u_i f_i(x)$ is the Lagrange function for (1.1) and $f'_i(x) = \text{grad} f_i(x)$, $i = 0,...,m$, are row-vectors. Let $I^* = \{i : f_i(x^*) = 0\} = \{1,...,r\}$ is the active constraints set and $r \leq n$.

We consider the vector-function $f(x) = (f_1(x),...,f_m(x))$, the vector-function of active constraints $f_{(r)}(x) = (f_1(x),...,f_r(x))$ and the vector-function of passive constraints $f_{(m-r)}(x) = (f_{r+1}(x),...,f_m(x))$.

We also consider their Jacobians $f'(x) = J(f(x))$, $f'_{(r)}(x) = J(f_{(r)}(x))$, $f'_{(m-r)}(x) = J(f_{(m-r)}(x))$, diagonal matrices $U = [\text{diag } u_i]_{i=1}^m$, $U_r = [\text{diag } u_i]_{i=1}^r$ with entries u_i, $i = 1, ..., m$ and Hessians

$$f''_i(x) = \left| \frac{\partial f_i}{\partial x_s \partial x_t} \right|_{\substack{s = 1,...,n \\ t = 1,...,n}} \qquad i = 0,1,...,m,$$

5

of the objective function and constraints. The sufficient regularity condition

$$\text{rank } f'_{(r)}(x^*) = r, \quad u_i^* > 0, \quad i \in I^* \tag{1.4}$$

together with the sufficient condition for the minimum x^* to be isolated

$$(L''_{xx}(x^*, u^*) z, z) \geq \lambda(z, z), \quad \lambda > 0 \quad \forall z \neq 0 : f'_{(r)}(x^*) z = 0^r \tag{1.5}$$

comprise the standard second order optimality condition, which we will assume in this paper.

We shall use the following assertion, which is a slight modification of the Debreu theorem (see [Pol 92]).

<u>Assertion 1</u>   Let A be a symmetric $n \times n$ matrix, B be an $r \times n$ matrix, and

$$U = [\text{diag } u_i]_{i=1}^{r} : R^r \to R^r, \quad \text{where } u = (u_1, \dots, u_r) > 0^r \text{ and let}$$

$$(A y, y) \geq \lambda (y, y), \quad \lambda > 0, \quad \forall y : B y = 0^r$$

Then there exists $k_0 > 0$ such that for any $0 < \mu < \lambda$ the following inequality

$$((A + k B^T U B) x, x) \geq \mu(x, x), \quad \forall x \in R^n$$

holds true whenever $k \geq k_0$.

2. Interior Distance Function. Let $y \in \text{int } \Omega$ and $\alpha = f_0(y)$, we consider the Relaxation Feasible Set (RFS) on the level α: $\Omega(\alpha) = \Omega \cap \{x : f_0(x) \leq \alpha\}$ and an interval $T = \{\tau : \alpha < \tau < \alpha^* = f_0(x^*)\}$. The Classical IDFs $F(x, \alpha)$ and $H(x, \alpha) : \Omega(\alpha) \times T \to \mathbb{R}^1$ are defined by formulas

$$F(x, \alpha) = -m \ln(\alpha - f_0(x)) - \sum_{i=1}^{m} \ln f_i(x) \quad ; \quad H(x, \alpha) = m (\alpha - f_0(x))^{-1} + \sum_{i=1}^{m} f_i^{-1}(x)$$

Let us assume that $\ln t = -\infty$ and $t^{-1} = \infty$ for $t \leq 0$, the Classical Interior Center Methods

6

(ICMs) consists of finding the "center" of the RFS by solving the following unconstrained optimization problem

$$\hat{x} = \hat{x}(\alpha) = \text{argmin} \{ F(x, \alpha) / x \in \mathbb{R}^n \}$$

and updating the objective function level α, i.e., replacing α by $\hat{\alpha} = f_0(\hat{x})$. Due to

$$x \rightarrow \partial \Omega(\alpha) \rightarrow F(x, \alpha) \rightarrow \infty \text{ the new center } \hat{x}(\alpha) \in \text{int } \Omega(\alpha) \subset \Omega \text{ for any } \alpha \in T.$$

Moreover, if the IDF possess the self-concordance properties (see [NesN 94]) the central trajectory $\{\hat{x}(\alpha), \alpha \in T\}$ has some very special features (see [Ren 88] and [Gon 92]).

Starting at a point close to the central trajectory - "warm" start - for a particular $\alpha \in T$ and using Newton step for solving the system

$$F'_x(x, \alpha) = 0^n$$

in x following by a "careful" α update, one can guarantee that the new approximation will be again a "warm" start and the gap between the current level $\alpha = f_0(x)$ and the optimal level $\alpha^* = f(x^*)$ will be reduced by a factor $0 < q_a < 1$, which is dependent only on the size of the problem.

However along with these nice properties the IDFs have their well known drawbacks. Neither the IDFs $F(x, \alpha)$ and $H(x, \alpha)$ nor their derivatives exist at the solution. Both $F(x, \alpha)$ and $H(x, \alpha)$ grow infinitely when $\hat{x}(\alpha)$ approaches the solution.

All constraints contribute equally to IDFs and one can obtain the optimal Lagrange multipliers only in the limit when $\hat{x}(\alpha) \rightarrow x^*$. What is particularly important for nonlinear constrained optimization is the fact that the condition number of the IDF Hessians vanishes when the process approaches the solution. Let's consider this issue briefly, using $F(x, \alpha)$. Keeping in mind the boundness of the RFS $\Omega(\alpha)$ one can guarantee that the unconstrained minimizer $\hat{x} \in \text{int } \Omega(\alpha)$ exists and

7

$$F'_x(\hat{x}, \alpha) = \frac{m}{\alpha - f_0(\hat{x})} f'_0(\hat{x}) - \sum_{i=1}^{m} \frac{f'_i(\hat{x})}{f_i(\hat{x})} = 0^n \qquad (2.1)$$

or

$$f'_0(\hat{x}) - \sum_{i=1}^{m} \frac{\alpha - f_0(\hat{x})}{m f_i(\hat{x})} f'_i(\hat{x}) = 0^n \qquad (2.2)$$

We define

$$\hat{u}_i = \hat{u}_i(\alpha) = (\alpha - f_0(\hat{x}))(m f_i(\hat{x}))^{-1}, \quad i = 1, \dots, m \qquad (2.3)$$

and consider the vector of Lagrange multipliers $\hat{u} = \hat{u}(\alpha) = (\hat{u}_i(\alpha), i = 1, \dots, m)$, then (2.2) can be rewritten as follows:

$$L'_x(\hat{x}, \hat{u}) = f'_0(\hat{x}) - \sum_{i=1}^{m} \hat{u}_i f'_i(\hat{x}) = f'_0(\hat{x}) - \hat{u} f'(\hat{x}) = 0^n$$

Also $\hat{u}_i f_i(\hat{x}) = (\alpha - f_0(\hat{x})) m^{-1}$, $i = 1, \dots, m$. So $\sum_{i=1}^{m} \hat{u}_i f_i(\hat{x}) = \alpha - f_0(\hat{x})$.

Under the uniqueness assumptions (1.4) - (1.5) we have

$$\lim_{\alpha \to \alpha^*} \hat{u}(\alpha) = u^*, \qquad \lim_{\alpha \to \alpha^*} \hat{x}(\alpha) = x^*$$

Let's consider the Hessian $F''_{xx}(x, \alpha)$. We obtain

$$F''_{xx}(\hat{x}, \alpha) = f''_0(\hat{x}) - \sum_{i=1}^{m} \frac{\alpha - f_0(\hat{x})}{m f_i(\hat{x})} f''_i(\hat{x}) - \sum_{i=1}^{m} \left[\left(\frac{\alpha - f_0(\hat{x})}{m f_i(\hat{x})} \right)' \right]^T f'_i(\hat{x})$$

Further, for any $i = 1, \dots, m$, we have $\left(\dfrac{\alpha - f_0(x)}{m f_i(x)} \right)' = m^{-1} \left[\dfrac{-f'_0(x) f_i(x) - f'_i(x)(\alpha - f_0(x))}{f_i^2(x)} \right]$

8

Therefore

$$-\sum_{i=1}^{m}\left[\left(\frac{\alpha-f_0(\hat{x})}{m f_i(\hat{x})}\right)'\right]^T f'_i(\hat{x}) = \sum_{i=1}^{m}\frac{(f'_0(\hat{x}))^T}{m f_i(\hat{x})}f'_i(\hat{x}) + \sum_{i=1}^{m}\frac{\alpha-f_0(\hat{x})}{m f_i(\hat{x})}\frac{1}{f_i(\hat{x})}(f'_i(\hat{x}))^T f'_i(\hat{x})$$

$$= \frac{1}{\alpha-f_0(\hat{x})}(f'_0(\hat{x}))^T\sum_{i=1}^{m}\frac{\alpha-f_0(\hat{x})}{m f_i(\hat{x})}f'_i(\hat{x}) + \sum_{i=1}^{m}\frac{\alpha-f_0(\hat{x})}{m f_i(\hat{x})}\frac{1}{f_i(\hat{x})}(f'_i(\hat{x}))^T f'_i(\hat{x})$$

$$= \frac{1}{\alpha-f_0(\hat{x})}(f'_0(\hat{x}))^T\sum_{i=1}^{m}\hat{u}_i f'_i(\hat{x}) + \sum_{i=1}^{m}\frac{\hat{u}_i}{f_i(\hat{x})}(f'_i(\hat{x}))^T f'_i(\hat{x})$$

Let $D(x) = [\ \mathrm{diag}\ f_i(x)\ ]_{i=1}^{m}$, $U(\alpha) = [\ \mathrm{diag}\ u_i(\alpha)\ ]_{i=1}^{m}$, then for the Hessian $F''_{xx}(\hat{x},\alpha)$

we obtain

$$F''_{xx}(\hat{x},\alpha) = L''_{xx}(\hat{x},\hat{u}) + (f'(\hat{x}))^T U(\alpha) D^{-1}(\hat{x}) f'(\hat{x}) + \frac{1}{\alpha-f_0(\hat{x})}[f'_0(\hat{x})]^T\sum_{i=1}^{m}\hat{u}_i(\alpha) f'_i(\hat{x})$$

In view of $\hat{x} = \hat{x}(\alpha) \to x^*$, $\hat{u} = \hat{u}(\alpha) \to u^*$ we obtain

$$L''_{xx}(\hat{x},\hat{u}) \to L''_{xx}(x^*,u^*),\ [f'_0(\hat{x})]^T\sum_{i=1}^{m}\hat{u}_i f'_i(\hat{x}) \to f_0^T(x^*)f'_0(x^*),$$

$$f'(\hat{x}) \to f'(x^*),\ \hat{U} \to U^*,\ D(\hat{x}) \to D(x^*)$$

9

Therefore

$$F''_{xx}(\hat{x},\alpha) \approx L''_{xx}(x^*,u^*) + (f'_{(r)}(x^*))^T \hat{E}(\alpha) f'_{(r)}(x^*) + (\alpha - f_0(\hat{x}))^{-1} f'^T_0(x^*) f'_0(x^*) \quad (2.3)$$

where $\hat{E}(\alpha) = [\,\text{diag}\,\hat{e}_i(\alpha)\,]^r_{i=1}$ and

$$\lim_{\alpha \to \alpha^*} \hat{e}_i(\alpha) = \lim_{\alpha \to \alpha^*} \hat{u}_i(\alpha)\, f_i^{-1}(\hat{x}(\alpha)) = +\infty \quad (2.4)$$

The mineigval $F''_{xx}(\hat{x},\alpha)$ is defined by the first two terms (2.3), therefore in view of (2.4) due to the Assertion 1 with $A = L''_{xx}(x^*,u^*)$ and $B = f'_{(r)}(x^*)$ there exists $\mu > 0$:

$$\text{mineigval}\ F''_{xx}(\hat{x},\alpha) = \mu$$

At the same time due to (2.4) we have maxeigval $F''_{xx}(\hat{x},\alpha) \to \infty$ when $\alpha \to \alpha^*$. Therefore the condition number of the Interior Distance Functions Hessians vanishes when $\hat{x}$ approaches x^*. The consequences of the ill -conditioning is much more substantial in nonlinear optimization than in Linear Programming. In case of LP the term $L''_{xx}(x, u)$ in the expression for the IDF Hessian disappears and by rescaling one can practically eliminate the ill-conditioning effect, at least, when the problem is not degenerate.

In nonlinear optimization the situation is completely different and the ill conditioning was and still is an important issue both in theory and practice. To eliminate the ill conditioning of the IDF we will introduce the Modified Interior Distance Functions.

3. Modified Interior Distance Functions We consider a vector $y \in \text{int}\ \Omega$ and $\Delta\ (y, x) =$

10

$f_0(y) - f_0(x) > 0$, then the Relaxation Feasible Set (RFS):

$$\Omega(y) = \{ x : f_i(x) \geq 0, i = 1,...,m \; ; \; \Delta(y,x) > 0 \}$$

The problem (1.1) is equivalent to

$$x^* = \text{argmin} \, \{ f_0(x) \, / \, x \in \Omega(y) \} \tag{3.1}$$

It is easy to see that for any $k > 0$

$$\Omega(y) = \{ x : k^{-1} [\ln(kf_i(x) + \Delta(y,x)) - \ln \Delta(y,x)] \geq 0 \quad i = 1,...,m; \; \Delta(y,x) > 0 \} .$$

Therefore the problem (3.1) is equivalent to the following problem:

$$x^* = \text{argmin} \, \{ -\ln \Delta(y,x) \, / \, x \in \Omega(y) \} \tag{3.2}$$

Assuming $\ln t = -\infty$ for $t \leq 0$ we define the MIDF $F(x, y, u, k)$:

$\mathbb{R}^n \times \text{int} \, \Omega \times \mathbb{R}_+^m \times \mathbb{R}_{++}^1 \to \mathbb{R}^1$ as a Classical Lagrangean for the equivalent problem (3.2):

$$F(x,y,u,k) = (-1 + k^{-1} \sum_{i=1}^{m} u_i) \ln \Delta(y,x) - k^{-1} \sum_{i=1}^{m} u_i \ln(kf_i(x) + \Delta(y,x)) \tag{3.3}$$

The MIDF $F(x, y, u, k)$ corresponds to the IDF $F(x, \alpha)$. To define the MIDF, which corresponds to $H(x, \alpha)$, we first note that for any $k > 0$

$$\Omega(y) = \{ x : k^{-1} [(kf_i(x) + \Delta(y,x))^{-1} - \Delta^{-1}(y,x)] \leq 0 , \; i = 1,...,m , \; \Delta(y,x) > 0 \}$$

Therefore the problem (1.1) is equivalent to

$$x^* = \text{argmin} \, \{ \Delta^{-1}(y,x) \, / \, x \in \Omega(y) \} \tag{3.4}$$

Assuming $t^{-1} = \infty$ for $t \leq 0$ we define the MIDF $H(x,y,u,k) : \mathbb{R}^n \times \text{int} \, \Omega \times \mathbb{R}_+^m \times \mathbb{R}_{++}^1 \to \mathbb{R}^1$

as a Classical Lagrangean for the equivalent problem (3.4):

$$H(x,y,u,k) = (-1 + k^{-1} \sum_{i=1}^{m} u_i) \Delta^{-1}(y,x) + k^{-1} \sum_{i=1}^{m} u_i (kf_i(x) + \Delta(y,x))^{-1} \tag{3.5}$$

11

The MIDF (3.5) corresponds to the P. Huard's IDF $H(x, \alpha)$. Both $F(x, y, u, k)$ and $H(x, y, u, k)$ are Classical Lagrangeans for problems equivalent to (1.1), which we obtained by monotone transformation both the objective function and the constraints.

Finally, the MIDF $Q(x,y,u,k) : \mathbf{R}^n \times \text{int}\,\Omega \times \mathbf{R}_+^m \times \mathbf{R}_{++}^1 \to \mathbf{R}^1$, which is defined by formula

$$Q(x,y,u,k) = (\Delta(y,x))^{-1+k^{-1}\Sigma u_i} \cdot \prod_{i=1}^{m} (kf_i(x) + \Delta(y,x))^{-k^{-1}u_i}$$

corresponds to the potential function

$$Q(x,\alpha) = (\alpha - f_0(x))^{-m} \prod_{i=1}^{m} f_i^{-1}(x).$$

So, we have $F(x, y, u, k) = \ln Q(x, y, u, k)$ and all basic facts about $F(x, y, u, k)$ remain true for $Q(x, y, u, k)$, therefore we will not consider the MIDF $Q(x, y, u, k)$ further in this paper.

There is a fundamental difference between the Classical and Modified Interior Distance Functions. First we are going to show the difference at the local level - in the neighborhood of the primal-dual solution. In the next section, we will consider the local MIDFs properties.

4. Local MIDFs Properties In contrast to the IDFs, the MIDFs are defined at the solution, they do not grow infinitely when the primal approximation approaches the solution and under the fixed optimal Lagrange multipliers, one can obtain the primal solution by solving one smooth unconstrained optimization problem.

Proposition 1. For any $k > 0$ and any $y \in \text{int}\,\Omega$, the following relations are taking place.

P1. $F(x^*,y,u^*,k) = -\ln\Delta(y,x^*)$ i.e. $f_0(x^*) = f_0(y) - \exp(-F(x^*,y,u^*,k))$

and

$$H(x^*,y,u^*,k) = \Delta^{-1}(y,x^*) \quad \text{i.e.} \quad f_0(x^*) = f_0(y) - H^{-1}(x^*,y,u^*,k)$$

The property P1 follows immediately from the definition of MIDFs and the complementary

12

conditions for the K-K-T's pair (x^*, u^*):

$$u_i^* f_i(x^*) = 0, \quad i = 1,...,m.$$

The fact that the MIDF's value at (x^*, u^*) coincides with the optimal objective function value for the equivalent problem independently on both the center $y \in \text{int } \Omega$ and barrier parameter $k > 0$ indicates that one can approach the solution by means other than those, which have been traditionally used in the IPM developments.

<u>Proposition 2</u>. For any $k > 0$ and any $y \in \text{int } \Omega$, the following relations are taking place.

P2. $\qquad F'_x(x^*, y, u^*, k) = \Delta^{-1}(y, x^*) L'_x(x^*, u^*) = 0$

and

$$H'_x(x^*, y, u^*, k) = \Delta^{-2}(y, x^*) L'_x(x^*, u^*) = 0$$

The proposition 2 immediately follows from the definition of MIDFs and K-K-T's conditions. If $k > \Sigma u_i^*$ the unconstrained minimizer of $F(x, y, u^*, k)$ or $H(x, y, u^*, k)$ in x is a solution of the convex programming problem (1.1), i.e., the following property is taking place.

<u>Proposition 3</u>

P3. $\qquad x^* = \text{argmin } \{ F(x, y, u^*, k) / x \in \mathbb{R}^n \} = \text{argmin } \{ H(x, y, u^*, k) / x \in \mathbb{R}^n \}$

In other words, the knowledge of the optimal Lagrange multipliers allows us to solve the problem (1.1) by solving one unconstrained optimization problem. Therefore if $F(x, y, u, k)$ is strongly convex in x and we know a good approximation u for the vector u^*, then $\hat{x} = \hat{x}(y, u, k) = $ argmin $\{ F(x, y, u, k) / x \in \mathbb{R}^n \}$ is a good approximation for x^* while both the "center" $y \in \text{int } \Omega$ and $k > \Sigma u_i$ are fixed.

If by using $\hat{x}$ we can improve the approximation u, then it is possible to develop a method where the convergence is due to the Lagrange multipliers update rather than due to the center or the

13

barrier parameter update.

Our goal is to develop such a method, but first we will try to understand under what conditions the MIDFs $F(x, y, u, k)$ and $H(x, y, u, k)$ will be strongly convex in x when both y and $k > 0$ are fixed.

The following proposition is the first step in this direction.

<u>Proposition 4</u>. If $f_i(x) \in C^2$, $i = 0, 1,..., m$, then for any fixed $y \in \text{int } \Omega$, $k > 0$ and any KKT's pairs (x^*, u^*) the following is true:

P4
$$F''_{xx}(x^*, y, u^*, k) = \Delta^{-1}(y, x^*)[L''_{xx}(x^*, u^*) +$$

$$\Delta^{-1}(y, x^*)(k(f'_{(r)}(x^*))^T U_r^* f'_{(r)}(x^*) - (f'_{(r)}(x^*))^T u_{(r)}^{*T} u_{(r)}^* f'_{(r)}(x^*))]$$

and

$$H''_{xx}(x^*, y, u^*, k) = \Delta^{-2}(y, x^*)[L''_{xx}(x^*, u^*) +$$

$$\Delta^{-1}(y, x^*)(k(f'_{(r)}(x^*))^T U_r^* f'_{(r)}(x^*) - (f'_{(r)}(x^*))^T u_{(r)}^{*T} u_{(r)}^* f'_{(r)}(x^*))]$$

The proof is given in the Appendix A1. We are now ready to prove the first basic statement.

<u>Theorem 1</u>. If $f_i(x) \in C^2$, $i = 0, 1,..., m$, then for the convex programming problem (1.1) the following statements are true:

1) for any fixed $y \in \text{int } \Omega$ and $k > \Sigma u_i^*$ the function $F(x, y, u, k)$ is strongly convex in the neighborhood x^* if one of the functions $f_0(x)$ or $-f_i(x)$, $i = 1,..., r$ is strongly convex or sufficient regularity conditions (1.4) are taking place and $r = n$;

2) if none of $f_0(x)$ and $-f_i(x)$, $i = 1,..., r$ are strongly convex and $r < n$, but the second order optimality conditions (1.4) - (1.5) are fulfilled then there exist $k_0 > 0$ large enough that for any

14

fixed $y \in$ int Ω and any fixed $k > \Delta(y, x^*) k_0 + \Sigma u_i^*$ there exist such that $\mu > 0$ and $M < + \infty$ that the following is true :

P5.   a) mineigval $F''_{xx}(x^*, y, u^*, k) \geq \Delta^{-1}(y, x^*) \mu$

b) maxeigval $F''_{xx}(x^*, y, u^*, k) \leq \Delta^{-1}(y, x^*) M$

<u>Proof</u>   1) Using P4 for any $v \in \mathbf{R}^a$ we obtain

$$F''_{xx}(x^*, y, u^*, k)v,v) = \Delta^{-1}(y, x^*)[L''_{xx}(x^*, u^*)v,v) +$$

$$\Delta^{-1}(y, x^*)(k(f'_{(r)}(x^*))^T U_r^* f'_{(r)}(x^*)v,v) - ((f'_{(r)}(x^*))^T u_{(r)}^{*T} u_{(r)}^* f'_{(r)}(x^*)v,v))]$$

$$= \Delta^{-1}(y, x^*)[(L''_{xx}(x^*, u^*)v,v) + \Delta^{-1}(y, x^*)(k - \Sigma u_i^*)((f'_{(r)}(x^*))^T U_r^* f'_{(r)}(x^*)v,v)$$

$$+ \Delta^{-1}(y, x^*)((\Sigma u_i^*)(\Sigma u_i^*(f'_i(x^*), v)^2) - (\Sigma u_i^*(f'_i(x^*), v))^2)]$$

Taking into account identity

$$(\sum_{i=1}^m u_i^*)(\sum_{i=1}^m u_i^* (f'_i(x^*), v)^2) - (\sum_{i=1}^m u_i^*(f'_i(x^*), v))^2 \qquad (4.1)$$

$$= \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m u_i^* u_j^* (f'_i(x^*) - f'_j(x^*), v)^2 \geq 0$$

we obtain

$$(F''_{xx}(x^*, y, u^*, k)v,v) \geq \Delta^{-1}(y, x^*)[(L''_{xx}(x^*, u^*) \qquad (4.2)$$

$$+ \Delta^{-1}(y, x^*)(k - \Sigma u_i^*)((f'_{(r)}(x^*))^T U_r^* f'_{(r)}(x^*)v,v)]$$

So for a convex programming problem (1.1) the function $F(x, y, u^*, k)$ is convex in x for any

$y \in$ int Ω and $k \geq \Sigma u_i^*$.

15

If one of $f_0(x)$, $-f_i(x)$, $i = 1,...,r$ are strongly convex then due to $u_i^* > 0$, $i = 1,...,r$ the

Classical Lagrangean $L(x, u^*)$ is strongly convex in the neighborhood of x^* while the matrix

$(k - \Sigma u_i^*) \Delta^{-1}(y,x^*) (f'_{(r)}(x^*))^T U_r^* f'_{(r)}(x^*)$ is non negative defined for any $y \in \text{int } \Omega$ and

$k \geq \Sigma u_i^*$, therefore $F(x, y, u^*, k)$ is strongly convex in the neighborhood of x^*. If $f_0(x)$ and all

$-f_i(x)$ are convex then $L(x, u^*)$ is convex in x, if addition (1.4) is satisfied and $r = n$, then for any

$y \in \text{int } \Omega$ and $k \geq \Sigma u_i^*$ the matrix $(k - \Sigma u_i^*) \Delta^{-1}(y,x^*)(f'_{(r)}(x^*))^T U_r^* f'_{(r)}(x^*)$ is positive defined

and again $F(x, y, u^*, k)$ is strongly convex.

Note, due to $f_i(x) \in C^2$, $i = 0, 1,..., m$ the MIDF $F(x, y, u^*, k)$ will remain strongly convex

in x for any $u \in \mathbb{R}_+^m$ close enough to u^*.

2) Now let's consider the case when none of $f_0(x)$ and $-f_i(x)$, $i = 1,..., r$ are strongly

convex and $r < n$. If $k > \Delta(y,x^*) k_0 + \Sigma u_i^*$, then due to (4.2) we obtain

$$(F''_{xx}(x^*,y,u^*,k)v,v) \geq \Delta^{-1}(y,x^*)([L''_{xx}(x^*,u^*)$$
$$+ k_0(f'_{(r)}(x^*))^T U_r^* f'_{(r)}(x^*)]v,v), \qquad \forall v \in \mathbb{R}^n \tag{4.3}$$

Therefore if the second order optimality condition (1.4) - (1.5) are satisfied, then due to the Assertion

1 with $A = L''_{xx}(x^*,u^*)$ and $B = f'_{(r)}(x^*)$ for $k_0 > 0$ large enough, any "center" $y \in \text{int } \Omega$ and

any $k > \Delta(y,x^*) k_0 + \Sigma u_i^*$ there exists $\mu > 0$:

$$(F''_{xx}(x^*,y,u^*,k)v,v) \geq \Delta^{-1}(y,x^*) \mu (v,v), \qquad \forall v \in \mathbb{R}^n \tag{4.4}$$

It is also clear that for a fixed $y \in \text{int } \Omega$ and fixed $k > \Delta(y, x^*) k_0 + \Sigma u_i^*$ there exists $M < \infty$:

$$(F''_{xx}(x^*,y,u^*,k)v,v) \leq \Delta^{-1}(y,x^*) M (v,v), \forall v \in \mathbb{R}^n \tag{4.5}$$

16

remark table

So the condition number of the Hessian $F''_{xx}(x^*, y, u^*, k)$ is fixed at the K-K-T's pair (x^*, u^*) and due to $f_i(x) \in C^2$ it remains to be true in the neighborhood of (x^*, y^*) for any fixed "center" $y \in \text{int}\,\Omega$ and any fixed barrier parameter $k > \Delta(y, x^*) k_0 + \Sigma u_i^*$.

<u>Remark 1</u>.   The second part of the theorem 1 remain true even for nonconvex problem if the second order optimality conditions are satisfied. In other words the barrier parameter k not only allows to retain the convexity in x of the MIDF $F(x, y, u, k)$ but also provide convexification of the $F(x, y, u, k)$ in x in case when the Classical Lagrangean $L(x, u)$ for the initial problem is not convex in $x \in \mathbf{R}_+^n$.

<u>Remark 2</u>.   Theorem 1 holds true for the MIDF $H(x, y, u, k)$. For any $y \in \text{int}\,\Omega$ and any fixed $k \geq 0.5\, k_0\, \Delta(y, x^*) + \Sigma u_i^*$, there exists $\mu > 0$ and $M < \infty$ that for $\forall v \in \mathbf{R}^n$ the following is true:

a) $H''_{xx}(x^*, y, u^*, k)v, v) \geq \Delta^{-2}(y, x^*)([L''_{xx}(x^*, y^*)$

$+ k_0 (f'_{(r)}(x^*))^T U_r^* f'_{(r)}(x^*)]v, v) \geq \Delta^{-2}(y, x^*)\, \mu(v, v)$,

b) $H''_{xx}(x^*, y, u^*, k)v, v) \leq \Delta^{-2}(y, x^*)\, M(v, v)$

5. Modified Center Method It follows from the Theorem 1 that to solve a constrained optimization problem for which the second order optimality conditions are fulfilled, it is enough to find a minimizer for a strongly convex and smooth in x function $F(x, y, u^*, k)$ with any fixed $y \in \text{int}\,\Omega$ as a "center" and any fixed $k > \Delta(y, x^*) k_0 + \Sigma u_i^*$. Due to the strong convexity of $F(x, y, u, k)$ in x to find an approximation to x^* it is enough to find a minimizer

$$\hat{x} = \hat{x}(y, u, k) = \text{argmin}\{F(x, y, u, k) / x \in \mathbf{R}^n\} \tag{5.1}$$

for a given Lagrange multipliers vector $u \in R_{+}^{m}$ close enough to u^{*}, when both y and k are fixed. Moreover, as it turns out, having the minimizer $\hat{x}$ one can find a better approximation $\hat{u}$ for the vector u^{*} without changing both $y \in \text{int } \Omega$ and $k > 0$.

Let's consider it with more details. Assuming that the minimizer $\hat{x}$ exists, we obtain

$$F_x'(\hat{x},y,u,k) = (1 - k^{-1} \Sigma u_i + k^{-1} \Sigma \hat{u})f_0'(\hat{x}) - \Sigma \hat{u}_i f_i'(\hat{x}) = 0^n \qquad (5.2)$$

where the components of the new vector of Lagrangean multipliers $\hat{u} = \hat{u}(y,u,k)$ are defined by formulas:

$$\hat{u}_i(y,u,k) = u_i \Delta (y,\hat{x}) (kf_i(\hat{x}) + \Delta (y,\hat{x}))^{-1}, \quad i=1,...,m$$

Let $d_i(x, y, k) = kf_i(x) + \Delta(y, x)$ then for the Lagrange multipliers update we have the following formulas

$$\hat{u}_i(y,u,k) = u_i \Delta (y,\hat{x}) d_i^{-1}(\hat{x},y,k), \quad i=1,...,m \qquad (5.3)$$

Formulas (5.3) are critical for our further considerations.

First, we have $\hat{u}(u^{*},y,k) = u^{*}$ for any fixed $y \in \text{int } \Omega$ and properly chosen $k > 0$, i.e. u^{*} is a fixed point of the map $u \rightarrow \hat{u}(u,y,k)$.

Second, we will show later that for the new vector $\hat{u}$ the following estimation:

$$| \hat{u} - u^{*}| \leq c k^{-1} \Delta (y,x^{*}) | u - u^{*} | \qquad (5.4)$$

holds, and $c > 0$ is independent on $y \in \text{int } \Omega$ and $k > 0$, where $| x | = | x |_{\infty} = \max_{1 \leq i \leq n} |x_i|$.

Third, it turns out that the estimation (5.4) is taking place not only for $\hat{u}$ but for the minimizer

18

$\hat{x}$ as well, i.e.

$$| \hat{x} - x^* | \le c\, k^{-1} \Delta (y,x^*) \, | u - u^* | \qquad (5.5)$$

In other words, finding a minimizer $\hat{x}$ and updating the vector $u \in \mathbf{R}_+^m$ is equivalent to applying to $u \in \mathbf{R}_+^m$ an operator :

$$C_{y,k} : C_{y,k} u = \hat{u}(u,y,k) = \hat{u}$$

Note that $C_{y,k}\, u^* = u^*$. The operator $C_{y,k}$ is a contractive one if

$$| C_{y,k} u - u^* | = | C_{y,k}(u - u^*) | < | u - u^* |$$

The contractibility of $C_{y,k}$ is defined by

$$Contr\, C_{y,k} = \gamma_{y,k} = c\, k^{-1} \Delta (y,x^*)$$

The constant $c > 0$ depends on the input data and the size of a given problem and independent on y and k. We will characterize the constant $c > 0$ in the course of proving the basic theorem.

So, for a given problem, the contractibility $0 < \gamma_{y,k} < 1$ depends on the "center" $y \in \text{int } \Omega$ and the barrier parameter $k > 0$.

The independence c on y and k makes possible to reduce $\gamma_{y,k} > 0$ to any apriori given level by increasing $k > 0$ under the fixed y, or reducing $\Delta (y, x^*)$ under the fixed k or by changing both the "center" $y \in \text{int } \Omega$ and the barrier parameter $k > 0$ in the process of solution.

In particular, for any "center" $y \in \text{int } \Omega$ and any given $0 < \gamma < 1$, one can find such a barrier parameter $k > 0$ that the operator $C_{y,k}$ will shrink the distance between current approximation (x,u)

and the primal dual solution (x^*, u^*) by a factor $0 < \gamma < 1$. Now we will describe the basic version of the Modified Interior Center Method. The convergence and rate of convergence will be considered later.

We start with $y \in$ int Ω, $u^0 = e_m = (1,...,1) \in \mathbf{R}^m$ and $k > m$. Let's assume that the couple (x^s, u^s) has been found already. Take $k > \Sigma u_i^s$, then the next approximation (x^{s+1}, u^{s+1}) we find by formulas:

$$x^{s+1} = \text{argmin} \; \{ F(x, y, u^s, k) / x \in \mathbf{R}^n \} \qquad (5.6)$$

$$u^{s+1} : \; u_i^{s+1} = u_i^s \, \Delta \, (y, x^{s+1}) \, d_i^{-1}(x^{s+1}, y, k) \quad i=1,...,m \qquad (5.7)$$

First, let us consider conditions for the problem (1.1), under which the method (5.6) - (5.7) is executable.

To simplify our consideration, we assume

A1. $\qquad \inf_{x \in \mathbf{R}^n} f_0(x) > -\infty$

We also assume that the set of optimal solutions for the problem (1.1) is not empty and bounded, i.e.

A2. $\qquad X^* = \text{Argmin} \{ f_0(x) / x \in \Omega \} \neq 0$ is bounded.

Taking into account the Corollary 20 (see [Fiac M68] p 94) and assumptions A1 - A2, we conclude that the set $\Omega_k(y) = \{ x : k f_i(x) + \Delta(y, x) \geq 0, \; i=1,...,m ; \; \Delta(y, x) > 0 \}$ is bounded for any $y \in$ int Ω and $k > 0$. Also $x \to \partial \Omega_k(y) \Rightarrow F(x, y, u^s, k) \to \infty$, therefore for any $u^s \in \mathbf{R}_+^m$, $y \in$ int Ω and $k > \Sigma u_i^s$ the function $F(x, y, u^s, k)$ is convex in $x \in \Omega_k(y)$, and the minimizer

$$x^{s+1} = \text{argmin} \; \{ F(x, y, u^s, k) / x \in \mathbf{R}^n \}$$

20

exists. Therefore

$$F_x'(x^{t+1}, y, u^t, k) = (1 - k^{-1}\sum u_i^t + k^{-1}\sum u_i^{t+1})f_0'(x^{t+1}) - \sum u_i^{t+1}f_i'(x^{t+1}) = 0 \quad (5.8)$$

and the vector $u^{t+1} \in \mathbf{R}_{++}^m$ if $u^t \in \mathbf{R}_{++}^m$

Hence, starting with a vector $u^0 \in \mathbf{R}_{++}^m$ one can guarantee that the Lagrange multipliers will remain positive up to the end of the process without any particular care about it.

Before discussing the convergence results we would like to describe briefly the dual interpretation of the MCM (5.6) - (5.7).

Let's consider the dual function

$$h(u) = \inf \{L(x, u) / x \in \mathbf{R}^n\}$$

and the dual problem

$$h(u^*) = \max \{h(u) / u \in \mathbf{R}_+^m\} \quad (5.9)$$

Along with the Classical Lagrangean $L(x, u)$ for the initial problem, we will consider an approximation for it

$$L(x, u, u^t, k) = (1 - k^{-1}\sum u_i^t + k^{-1}\sum u_i)f_0(x) - \sum u_i f_i(x)$$

Note that $L(x, u^t, u^t, k) = L(x, u^t)$. We also consider an associated with $L(x, u, u^t, k)$ approximation for the dual function

$$\hbar(u, u^t, k) = \inf \{L(x, u, u^t, k) / x \in \mathbf{R}^n\},$$

which is equal to the dual function $h(u)$ when $u = u^t$, i.e., $\hbar(u^t, u^t, k) = h(u^t)$ for any $k > 0$.

Along with the dual problem (5.9) we consider the following convex programming problem

$$\max \{\hbar(u, u^t, k) / u \in \mathbf{R}_+^m\}$$

The function $\hbar(u, u^t, k)$ is concave in $u \in \mathbf{R}_+^m$ and due to (5.8) $L_x'(x^{t+1}, u^{t+1}, u^t, k) = 0$.

21

Therefore $L(x^{r+1}, u^{r+1}, u', k) = h(u^{r+1}, u', k)$. For the subgradient of $h(u, u', k)$ at u^{r+1} we have

$$\partial h(u^{r+1}, u', k) = k^{-1} f_0(x^{r+1}) e_m - f(x^{r+1}) \qquad (5.10)$$

$f(x) = (f_1(x),\dots,f_m(x))$. Using formulas (5.7) for the Lagrange multipliers update, we obtain

$$f(x^{r+1}) = -k^{-1} \Delta(y, x^{r+1}) u'(u^{r+1})^{-1} + k^{-1} f_0(x^{r+1}) e_m - k^{-1} f_0(y) e_m \qquad (5.11)$$

where $u'(u^{r+1})^{-1} = (u_1'(u_1^{r+1})^{-1},\dots, u_m'(u_m^{r+1})^{-1})$.

So in view of (5.10) and (5.11) we obtain

$$\partial h(u^{r+1}, u', k) = -k^{-1} \Delta(y, x^{r+1}) u'(u^{r+1})^{-1} - k^{-1} f_0(y) e_m$$

or

$$\partial h(u^{r+1}, u', k) + k^{-1} \Delta(y, x^{r+1}) u'(u^{r+1})^{-1} + k^{-1} f_0(y) e_m = 0$$
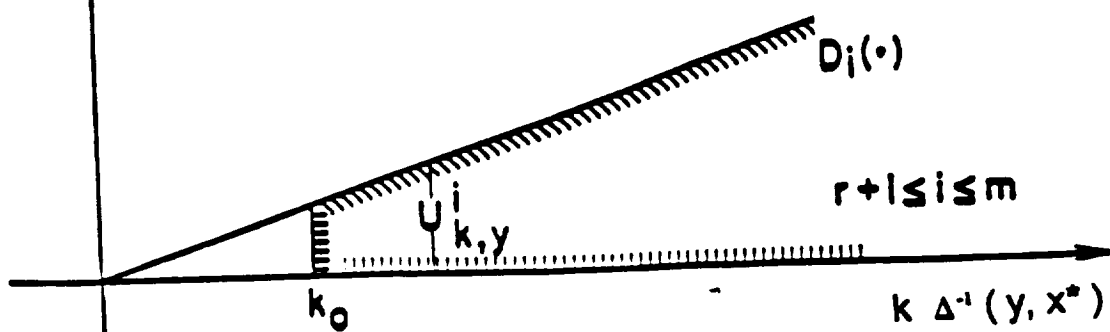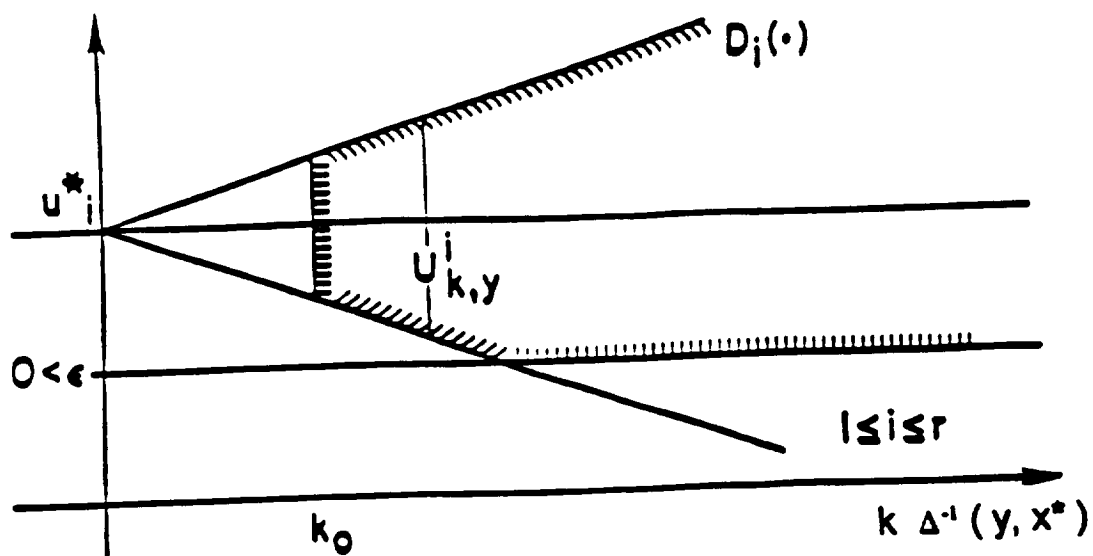
Therefore

$$u^{r+1} = \mathrm{argmax}\, \{\, h(u, u', k) + k^{-1} \Sigma u_i'\, [\, \Delta(y, x^{r+1}) \ln u_i (u_i')^{-1}$$
$$+ f_0(y) u_i (u_i')^{-1} - f_0(y)\,]\, /\, u \in \mathbb{R}^m_{++} \,\} \qquad (5.12)$$

The method (5.12) has some similarities with the prox-method with eutropy-like kernel (see [PolTeb95]), which corresponds to MBF (see [Pol92]), however, there is a fundamental difference between them as well. In contrast to the prox-method, which corresponds to MBF, the dual to MCM is dealing not with the dual objective function $h(u)$, but with an approximation $h(u, u', k)$ to $h(u)$. Therefore the convergence results for MBF method cannot be applied to (5.12).

We will obtain the convergence results for the method (5.12) as byproduct of the correspondent results for MCM (5.6) - (5.7).

These results will follow from the Basic Theorem, which we are going to prove in the next section.

22

$D(\cdot) = D_1(\cdot) \circledast \ldots \circledast D_r(\cdot) \circledast \ldots \circledast D_m(\cdot)$

$U_{k,y} = U_{k,y}^1 \circledast \ldots \circledast U_{k,y}^r \circledast \ldots \circledast U_{k,y}^m$

6. Basic Theorem . The Basic Theorem establishes the contractibility properties of the operator $C_{y,k}$. We will start by characterizing the domain, where the operator $C_{y,k}$ is defined and possesses these properties. Let's consider a small enough number $\tau > 0$, a fixed $y_0 \in \text{int } \Omega$ and a subset

$\Omega_\tau = \{x : f(x) \geq \tau\, e\} \cap \{x : \Delta\, (y_0,x) > 0\}$ of the RFS $\Omega\, (y)$. Note that due to A1 - A2 and the Corollary 20 (see [FiacM68]) the set Ω_τ is bounded. We will choose the "center" y from Ω_τ. For any $y \in \Omega_\tau$ we have $\Delta\, (y, x^*) > 0$. Along with $\tau > 0$ we consider a couple of small numbers $\varepsilon > 0$ and $\delta > 0$ and a large enough $k_0 > 0$. In the course of proving the Basic Theorem it will become clear what "small" and "large" mean.

To characterize the domain, where the operator $C_{y,k}$ is defined, we will consider two types of sets.

The first type

$$D_i(\bullet) = \{(y,u,k) : u_i \geq \varepsilon,\, |u_i - u_i^*| \leq \delta\, \Delta^{-1}(y,x^*)k,\, k \geq k_0\, \Delta\, (y,x^*) + \Sigma u_i^*\},\, i = 1,\ldots,r$$

is related to the active constraints.

The second type is associated with the passive constraints.

$$D_i(\bullet) = \{(y,u,k) : 0 \leq u_i \leq \delta\, \Delta^{-1}(y,x^*)k,\, k \geq k_0\, \Delta\, (y,x^*) + \Sigma u_i^*\},\, i = r+1,\ldots,m$$

The set $D(\bullet) = D_1(\bullet) \times \ldots \times D_r(\bullet) \times \ldots \times D_m(\bullet)$ is the domain, where the operator $C_{y,k}$ is defined.

We will prove later that for any fixed $y \in \Omega_\tau$ there exists $k_0 > 0$ that for any $k \geq k_0 \Delta(y,x^*) + \Sigma u_i^*$ the operator $C_{y,k}$ is a contractive on $D(\bullet)$.

For a fixed $y \in \Omega_\tau$ and a fixed $k \geq k_0 \Delta(y,x^*) + \Sigma u_i^*$, the domain $D(\bullet)$ shrinks (see Fig.1) to

$$U_{y,k} = U_{y,k}^1 \times \ldots \times U_{y,k}^r \times \ldots \times U_{y,k}^m.$$

We are particularly interested in the set $U_{y,k}$, because as soon as both the "center" y and the

23

barrier parameter k are fixed, the set $U_{y,k}$ is the only feasible set for the Lagrange multipliers, moreover if $C_{y,k}$ is a contractive operator then $u \in U_{y,k} \to \hat{u} \in U_{y,k}$.

Before we turn to the Basic Theorem, let's briefly describe the main idea of the proof. In view of A1-A2 for any $u \in R_+^m$, $y \in \Omega_\tau$ and $k > \Sigma u_i$, there exists the MIDF's minimizer $\hat{x} = \hat{x}(y,u,k)$ and

$$F_x'(\hat{x},y,u,k) = f_0'(\hat{x}) - \sum_{i=1}^{r} \hat{u}_i f_i'(\hat{x}) - h(\hat{x},y,u,k) + g(\hat{x},y,u,k) = 0 \tag{6.1}$$

where

$$\hat{u}_i = u_i \, \Delta(y,\hat{x}) \, d_i^{-1}(\hat{x},y,k), \quad i = 1,...,r, \tag{6.2}$$

$$h(\hat{x},y,u,k) = \sum_{i=r+1}^{m} u_i \, \Delta(y,\hat{x}) \, d_i^{-1}(\hat{x},y,k) f_i'(\hat{x})$$

and

$$g(\hat{x},y,u,k) = k^{-1} \sum_{i=1}^{m} u_i \left(-1 + \Delta(y,\hat{x}) \, d_i^{-1}(\hat{x},y,k) \right) f_0'(\hat{x})$$

Considering (6.1) and (6.2) as a system of equations for $\hat{x}$ and $\hat{u}_{(r)}$, it is easy to verify that $\hat{x} = x^*$ and $\hat{u}_{(r)} = u_{(r)}^*$ satisfy the system for any $y \in \Omega_\tau$, $k > \Sigma u_i^*$, and $u = u^*$.

Moreover for any triple $(y, u, k) \in D(\cdot)$, the system (6.1) - (6.2) can be solved for $\hat{x}$ and $\hat{u}_{(r)}$.

Having the solution $\hat{x} = \hat{x}(y, u, k)$ and $= \hat{u}_{(r)} = \hat{u}_{(r)}(y,u,k)$ one can find the Jacobians

$\hat{x}_u'(\cdot) = J_u(\hat{x}(y,u,k))$ and $\hat{u}_{(r)u}'(\cdot) = J_u(\hat{u}_{(r)}(y,u,k))$ and estimate

$|\hat{x}_u'(y,u^*,k)|$ and $|\hat{u}_{(r)u}'(y,u^*,k)|$

It turns out that under second order optimality condition, there is such $k_0 > 0$ that for any $y \in \Omega_\tau$ and $k \geq k_0 \Delta(y,x^*) + \Sigma u_i^*$ the following estimation

$$\max\{ |\hat{x}_u'(y,u^*,k)|, |\hat{u}_{(r)u}'(y,u^*,k)| \} \leq c \tag{6.3}$$

24

takes place and $c > 0$ is independent on y and k.

Due to the continuity $\hat{x}'_u(\cdot)$ and $\hat{u}'_{(r)u}(\cdot)$ in u the estimation (6.3) is taking place in the neighborhood of u^*.

In view of $x^* = \hat{x}(y, u^*, k)$ and $u^* = \hat{u}(y, u^*, k)$ and using (6.3) one can estimate $|\hat{x} - x^*|$ and $|\hat{u} - u^*|$ through $|u - u^*|$.

The independence $c > 0$ on y and k makes possible to prove that for any fixed $y \in \Omega_\tau$ there exists $k_0 > 0$ such that for any $k \geq k_0 \Delta(y, x^*) + \Sigma u_i^*$ the operator $C_{y,k}$ is a contractive one, i.e. $0 < \gamma_{y,k} < 1$, therefore $u \in U_{y,k} \rightarrow C_{y,k} \, u = \hat{u} \in U_{y,k}$.

In the course of proving the Basic Theorem we will assume $\min \{ f_i(x^*) \, / \, i = r+1, \ldots, m \} = \sigma > 0$, $\min \{ \Delta (y, x^*) \, / \, y \in \Omega_\tau \} = \tau_0 > 0$, $O^{p,q}$ be the $p \times q$ zero matrix; I^r be the $r \times r$ identity matrix, $S(a, \varepsilon) = \{ x \in \mathbb{R}^n : |x - a| \leq \varepsilon \}$. We remind that $d_i(x, y, k) = (k f_i(x) + \Delta (y, x))$ and introduce three diagonal matrices $d(x, y, k) = [\text{diag } d_i(x, y, k)]_{i=1}^m$, $d_{(r)}(x, y, k) = [\text{diag } d_i(x, y, k)]_{i=1}^r$, $d_{(m-r)}(x, y, k) = [\text{ diag } d_i(x, y, k)]_{i=r+1}^m$.

Theorem 2.   1. If A1 - A2 are taking place, then for any $y \in \Omega_\tau$, $u \in \mathbb{R}_{++}^m$ and $k > \Sigma u_i^*$ there exists $\hat{x} = \hat{x}(y, u, k) = \text{argmin} \{ F(x, y, u, k) \, / \, x \in \mathbb{R}^n \} : F'_x(\hat{x}, y, u, k) = 0^n$.

2. If $f_i(x) \in C^2$, $i = 0, \ldots, m$ and standard second order optimality conditions (1.4)-(1.5) are taking place then:

a) for any triple $(y, u, k) \in D(\cdot)$ the minimizer $\hat{x} = \hat{x}(y, u, k)$ exists, $F'_x(\hat{x}, y, u, k) = 0^n$ and for the pair $\hat{x}$ and $\hat{u} = \hat{u}(y, u, k)$ the following estimate

$$\max \{ |\hat{x} - x^*| \, , \, |\hat{u} - u^*| \} \leq c \, k^{-1} \Delta (y, x^*) |u - u^*| \qquad (6.4)$$

holds and $c > 0$ is independent on y and k.

b) for any fixed $y \in \Omega_\tau$ and $k > k_0 \Delta(y,x^*) + \Sigma u_i^*$, the MIDF $F(x, y, u, k)$ is strongly convex in the neighborhood of $\hat{x}$ and for any $u \in U_{y,k}$ there exists $\hat{\mu} > 0$ and $\hat{M} < \infty$:

$$\text{mineigval} \quad F''_{xx}(\hat{x},y,u,k) \geq \Delta^{-1}(y,x^*)\,\hat{\mu} \tag{6.5}$$

$$\text{maxeigval} \quad F''_{xx}(\hat{x},y,u,k) \leq \Delta^{-1}(y,x^*)\,\hat{M} \tag{6.6}$$

<u>Proof</u> 1) In view of the assumptions A1 - A2 and the Corollary 20 (see [Fiac M68] p94) the set

$\Omega_k(y) = \{x : kf_i(x) + \Delta(y,x) \geq 0, i = 1,...,m; \Delta(y,x) > 0\}$ is bounded for any $y \in \Omega_\tau$ and $k > 0$. Also $x \in \partial\,\Omega_k(y) \rightarrow F(x, y, u^*, k) \rightarrow \infty$, therefore for any $y \in \Omega_\tau$, $u \in R_{++}^m$ and $k > \Sigma u_i$ the function $F(x, y, u, k)$ is convex in $x \in \Omega_k(y)$ and $\hat{x} = \hat{x}(y,u,k)$ is an unconstrained minimizer of $F(x, y, u, k)$, i.e. $F'_x(\hat{x},y,u,k) = 0^n$.

2) For technical reasons, we introduce a vector $t = (t_1,...,t_m)$, $t_i = k^{-1}\Delta(y,x^*)(u_i - u_i^*)$ instead of the vector of Lagrange multipliers u, then $u = u^* \rightarrow t = O^m$. Such transformation translates the neighborhood of u^* into the neighborhood $S(o,\delta) = \{t : |t_i| \leq \delta, i = 1,...,m\}$ of the origin of dual space.

We will split the vector $\hat{u}$ on two parts, which correspond to the active and passive constraints.

Let $\hat{u}_{(r)} = (\hat{u}_i, i = 1,...,r)$ is a vector of Lagrange multiplier, which corresponds to the active constraints, while $\hat{u}_{(m-r)} = \hat{u}_{(m-r)}(x,y,t,k) = (\hat{u}_i(x,y,t,k), i = r+1,...,m)$ is the vector of Lagrange multipliers, which corresponds to the passive contraints.

We have $\hat{u}_i(x,y,t,k) = k\Delta^{-1}(y,x^*)t_i\,\Delta(y,x)d_i^{-1}(x,y,k)$ $i = r+1,...,m$, $\hat{u} = (\hat{u}_{(r)}, \hat{u}_{(m-r)})$ and for the vector function $h(x, y, u, k)$, $g(x, y, u, k)$ we will have the following replacement.

26

$$h(x,y,t,k) = \sum_{i=r+1}^{m} \hat{u}_i(x,y,t,k)\,(f_i'(x))^T = (\hat{u}_{(m-r)}(x,y,t,k)f_{(m-r)}'(x))^T,$$

$$g(x,y,t,k) = k^{-1}\left\{ \sum_{i=1}^{m} (kt_i\Delta^{-1}(y,x^*) + u_i^*)\,[\,-1 + \Delta(y,x)\,d_i^{-1}(x,y,k)]\right\}(f_0'(x))^T$$

So for any $k > 0$, small enough $\varepsilon_0 > 0$ and $y \in \Omega_\varepsilon$, the vector functions $h(x, y, t, k)$ and $g(x, y, t, k)$ are smooth in $x \in S(x^*, \varepsilon_0)$ and $t \in S(0,\delta)$. Then we have $h(x^*,y,0,k) = 0^n$, $g(x^*,y,0,k) = 0^n$,

$h_x'(x^*,y,0,k) = 0^{n,n}$, $g_x'(x^*,y,0,k) = -\Delta^{-1}(y,x^*)f_0'^{T}(x^*)f_0'(x^*)$ also $h'_{\hat{u}_{(r)}}(x^*,y,0,k) = 0^{n,r}$

and $g'_{\hat{u}_{(r)}}(x^*,y,0,k) = 0^{n,r}$. On $S(x^*,\varepsilon_0) \times S(u_{(r)}^*,\varepsilon_0) \times \Omega_\varepsilon \times S(0,\delta) \times (0,+\infty)$ we consider the

map $\Phi(x,\hat{u}_{(r)},y,t,k): \mathbf{R}^{2n+r+m+1} \to \mathbf{R}^{n+r}$ defined as follows:

$$\Phi(x,\hat{u}_r,y,t,k) = \{f_0'^{T}(x) - \sum_{i=1}^{r} \hat{u}_i f_i'^{T}(x) - h(x,y,t,k) + g(x,y,t,k)\,;$$

$$k^{-1}\Delta(y,x^*)\,[\,(k\Delta^{-1}(y,x^*)t_i + u_i^*)\,\Delta(y,x)\,d_i^{-1}(x,y,k) - \hat{u}_i\,]\,,\, i = 1,\dots,r\}$$

Taking into account (1.3) and $h(x^*, y, 0, k) = g(x^*, y, 0, k) = 0^n$ we obtain

$$\Phi(x^*,u_{(r)}^*,y,0,k) = 0^{n+r} \text{ for any } k > 0 \text{ and } y \in \Omega_\varepsilon.$$

Let $\Phi'_{x\hat{u}_{(r)}} = \Phi'_{x\hat{u}_{(r)}}(x^*,u_{(r)}^*,y,0,k)$, $L_{xx}'' = L_{xx}''(x^*,u^*)$, $f' = f'(x^*)$, $f_{(r)}' = f_{(r)}'(x^*)$,

$U_r^* = [\operatorname{diag} u_i^*]_{i=1}^r$, $u_{(r)}^* = (u_i^*, i = 1,\dots,r)$.

In view of $h_x'(x^*,y,0,k) = 0^{n,n}$, $h'_{\hat{u}_{(r)}}(x^*,y,0,k) = 0^{n,r}$, $g_x'(x^*,y,0,k)$

$= -\Delta^{-1}(y,x^*)f_0'^{T}(x^*)f_0'(x^*)$, $g'_{\hat{u}_{(r)}}(x^*,y,0,k) = 0^{n,r}$ we obtain

$$\Phi_{(y,k)} = \Phi'_{x\hat{u}_{(r)}}(x^*,u_{(r)}^*,0,y,k) = \begin{bmatrix} L_{xx}'' - \Delta^{-1}(y,x^*)f_0'^{T}f_0' & -f_{(r)}'^{T} \\ -U_r^* f_{(r)}' & -k^{-1}\Delta(y,x^*)I^r \end{bmatrix}$$

Now we will prove the nondegeneracy of the matrix $\Phi_{(y,k)}$ for any $y \in \operatorname{int}\Omega_\varepsilon$ and any

$k \geq k_0\Delta(y,x^*) + \sum u_i^*$. Let us consider $w = (z, v) \in R^{n+r}$, then the system $\Phi_{(y,k)}w = 0^{n+r}$ can be

rewritten as follows:

$$L_{xx}'' z - \Delta^{-1}(y,x^*) f_0'^T f_0' z - f_{(r)}'^T v = 0^n \tag{6.7}$$

$$U_r^* f_{(r)}' z - k^{-1} \Delta(y,x^*) v = 0^r \tag{6.8}$$

We find v from (6.8) and substitute in (6.7). Taking into account the K-K-T's condition

$f'_0 = u^*_{(r)} f'_{(r)}$ we obtain

$$L_{xx}'' z + \Delta^{-1}(y,x^*) [k f_{(r)}'^T U_r^* f_{(r)}' - (f_{(r)}'^T u_{(r)}^{*T} u_{(r)}^* f_{(r)}')] z = 0^n$$

i.e.

$$(L_{xx}'' z, z) + \Delta^{-1}(y,x^*) [k(U_r^* f_{(r)}' z, f_{(r)}' z) - (u_{(r)}^* f_{(r)}' z)^2] = 0$$

The inequality

$$k \geq k_0 \Delta(y,x^*) + \Sigma u_i^*$$

implies

$$(L_{xx}'' z, z) + \Delta^{-1}(y,x^*) [k(U_r^* f_{(r)}' z , f_{(r)}' z) - (u_{(r)}^* f_{(r)}' z)^2]$$

$$\geq (L_{xx}'' z, z) + \Delta^{-1}(y,x^*) [k_0 \Delta(y,x^*) (U_r^* f_{(r)}' z , f_{(r)}' z)$$

$$+ (\Sigma u_i^*) (U_r^* f_{(r)}' z , f_{(r)}' z) - (u_{(r)}^* f_{(r)}' z)^2]$$

$$= (L_{xx}'' z, z) + k_0 (f_{(r)}'^T U_{(r)}^* f_{(r)}' z, z) + \Delta^{-1}(y,x^*) [(\Sigma u_i^*)(\Sigma u_i^* (f_i', z)^2) - (\Sigma u_i^* (f_i', z))^2]$$

Due to identity (4.1) we obtain $(\Sigma u_i^*)(\Sigma u_i^* (f_i', z)^2) - (\Sigma u_i^* (f_i', z))^2 \geq 0$. Therefore taking into

account Assertion 1, we have $0 = ((L_{xx}'' + k_0 f_{(r)}'^T U_{(r)}^* f_{(r)}') z, z) \geq \mu(z,z), \mu > 0,$ i.e. $z = 0^n,$ hence

from (6.7) we obtain $f_{(r)}'^T v = 0,$ so due to (1.4) we have $v = 0,$ i.e. $\Phi_{(y,k)} w = 0^{n+r} \rightarrow w = 0^{n+r},$ i.e.

$\Phi_{(y,k)}$ is a nonsingular matrix.

Let k_1 be large enough. We consider a compact

28

$K = \{0^n\} \times \{(y,k) : y \in \Omega_{\tau}, k_1 \geq k \geq k_0 \, \Delta\,(y,x^{\cdot}) + \Sigma\, u_i^{\cdot}\}$

Since $\Phi\,(x^{\cdot}, \hat{u}_{(r)}, y, 0, k) = 0^{n+r}$, the matrix $\Phi_{y,k} = \Phi'_{x\hat{u}_{(r)}}\,(x^{\cdot}, u_{(r)}^{\cdot}, y, 0, k)$ is nonsingular,

$f_i(x) \in C^2, i = 1, ..., m$ and K is compact, it follows from the second implicit function theorem

(see [Ber82]p.12) that there is a small enough $\delta > 0$ that in the neighborhood

$S(K, \delta) = \{(y,t,k) : |t_i| \leq \delta,$

$i = 1, ..., m, y \in \Omega_{\tau}, k \in [k_0\,\Delta\,(y,x^{\cdot}) + \Sigma\,u_i^{\cdot}, k_1]\}$ of the compact K there exist unique continuously

differentiable vector - functions $x(\cdot) = x(y,t,k) = (x_1(y,t,k), ..., x_n(y,t,k))$ and

$\hat{u}_{(r)}(\cdot) = \hat{u}_{(r)}(y,t,k) = (\hat{u}_1(y,t,k), ..., \hat{u}_r(y,t,k))$ such that $x(y,0,k) = x^{\cdot}, \hat{u}_{(r)}(y,0,k) = u_{(r)}^{\cdot}$

and for any triple $(y, t, k) \in S(K, \delta)$ there is $\varepsilon_0 > 0$ that

$$\max\, \{\,|x(y,t,k) - x^{\cdot}|\,,\ |\hat{u}_{(r)}(y,t,k) - u_{(r)}^{\cdot}|\,\} \leq \varepsilon_0 \tag{6.9}$$

The identity

$$\Phi\,(x(y,t,k), \hat{u}_{(r)}(y,t,k), y,t,k) = \Phi\,(x(\cdot), \hat{u}_{(r)}(\cdot), \cdot) = 0^{n+r} \tag{6.10}$$

holds true for all $(y, t, k) \in S(K, \delta)$.

So we obtain

$$f_0'^T(x(\cdot)) - \sum_{i=1}^{r} \hat{u}_i(\cdot) f_i'^T(x(\cdot)) - h(x(\cdot),\cdot) + g(x(\cdot),\cdot) \tag{6.11}$$

$$= \Delta\,(y,x(\cdot)) F_x'(x(\cdot),y,u,k) = 0^n$$

which is the necessary optimality condition for the vector $x(\cdot)$ to be a minimizer of the function

$F(x, y, u, k)$ in x under the fixed (y, u, k). Also from (6.10) we obtained the identities

$\hat{u}_i(\cdot) = (k\Delta^{-1}(y,x^{\cdot})t_i + u_i^{\cdot})\, \Delta\,(y,x(\cdot)) d_i^{-1}(x(\cdot),y,k)\ i = 1,...,r$ for the Lagrange multipliers

that corresponds to the active contraints.

After multiplying both sides by $k^{-1}\Delta(y, x^{\cdot})$, it can be rewritten as follows:

$$(t_i + k^{-1}\Delta\,(y,x^{\cdot})u_i^{\cdot})\, \Delta\,(y,x(\cdot))\, d_i^{-1}(x(\cdot),y,k) - k^{-1}\Delta\,(y,x^{\cdot})\hat{u}_i(\cdot) = 0\ \ r = 1,...,r \tag{6.12}$$

The Lagrange multipliers that correspond to the passive constraints , we can rewrite in the following way:

$$\hat{u}_i(x(\bullet),\bullet) \equiv \hat{u}_i(\bullet) = k\Delta^{-1}(y,x^\ast)t_i\,\Delta\,(y,x(\bullet))\,d^{-1}(x(\bullet),y,k)\quad i=r+1,...,m \qquad (6.13)$$

let $\hat{u}_{(m-r)}(\bullet) = (\hat{u}_i(\bullet),\ i=r+1,...,m$ and $\hat{u}(\bullet) = (\hat{u}_{(r)}(\bullet),\hat{u}_{(m-r)}(\bullet))$.

To prove the sufficient optimality condition for the vector $x(\bullet)$ to be a minimizer of the function $F(x,\,y,\,u,\,k)$ in x under fixed $(y,\,u,\,k)$ we will show later that the function $F(x,\,y,\,u,\,k)$ is strongly convex in the neighborhood of $x(\bullet)$ for any $(y,\,u,\,k)\in D(\bullet)$. But first of all we will ascertain the estimation (6.4).

To this end let us first prove that for small enough $\delta > 0$ and large enough k_0 there exists $\rho > 0$ such that the inequality

$$\mid (\Phi'_{x,\hat{u}_{(r)}}(x(\bullet),\hat{u}_{(r)}(\bullet),\bullet))^{-1}\mid\ \leq \rho \qquad (6.14)$$

holds true for all $(t,\,y,\,k)\in S(K,\,\delta)$.

We consider the matrix

$$\Phi_{(y,\bullet)} = \Phi'_{x,\hat{u}_{(r)}}(x^\ast,u^\ast_{(r)},0,y,\infty) = \begin{bmatrix} L''_{xx} - \Delta^{-1}(y,x^\ast)f'^T_0 f'_0 & -f'^T \\ -U^\ast_r f'_{(r)} & 0^{r,r} \end{bmatrix}$$

The matrix $\Phi_{(y,\bullet)}$ is nonsingular for any $y\in\Omega_\tau$. In fact, for a vector $w = (z,\,v)\in\mathbf{R}^{n+r}$ the system

$$\Phi_{(y,\bullet)}\,w = 0^{n+r}$$

can be rewritten in the following way

$$L''_{xx}z - \Delta^{-1}(y,x^\ast)f'^T_0 f'_0 z - f'^T_{(r)}v = 0^n \qquad (6.15)$$

$$-U^\ast_r f'_{(r)}z = 0^r \qquad (6.16)$$

30

Because of $u^*_{(r)} > 0^r$ from (6.16) we obtain $f'_{(r)} z = 0^r$, i.e. $(f'_i , z) = 0 , i = 1,...,r$, therefore

$(\sum_{i=1}^{r} u_i^* f'_i , z) = (f'_0, z) = 0$. Multiplying (6.15) by z we obtain

$$(L''_{xx} z , z) - \Delta^{-1}(y, x^*) (f'_0, z)^2 - (v, f'_{(r)} z) = 0 \qquad (6.17)$$

i.e.

$$(L''_{xx} z , z) = 0 , \quad \forall z : f'_{(r)} z = 0^r$$

so due to (1.5) we have $z = 0^n$ then from (6.14) one obtains $f'^T_{(r)} v = 0^n$, which due to (1.4) implies

$v = 0^r$.

Therefore, $\Phi_{(y, \infty)} w = 0^{n+r}$ implies $w = 0^{n+r}$ for any $y \in \Omega$, i.e. the matrix $\Phi_{(y, \infty)}$ is nonsingular, so

there exists a constant $\rho > 0$ independent of k and $y \in \Omega$, such that

$$| \Phi^{-1}_{(y, \infty)} | \le \rho_0 .$$

Hence, for the Gram matrix $G_{(y, \infty)} = \Phi^T_{(y, \infty)} \Phi_{(y, \infty)}$ we have mineigval $G_{(y, \infty)} = \mu_0 > 0$. Then there

exists a large enough $k_0 > 0$ such that for any $y \in \Omega$, and $k \ge k_0 \Delta(y, x^*) + \sum u_i^*$ i.e. $k^{-1} \Delta(y, x^*) <$

k_0^{-1} we obtain for the Gram matrix $G_{(y, k)} = \Phi^T_{(y, k)} \Phi_{(y, k)}$ the inequality

$$\text{mineigval } G_{(y, k)} \ge \tfrac{1}{2} \mu_0$$

and $\mu_0 > 0$ is independent of $y \in \Omega$, and $k \in [k_0 \Delta(y, x^*) + \sum u_i^*, k_1]$. Therefore $\Phi_{(y, k)}$ is not only

nonsingular, but there exists a constant $\rho > 0$ independent of $y \in \Omega$, and $k \ge k_0 \Delta(y, x^*) + \sum u_i^*$ such

that

$$| \Phi^{-1}_{(y, k)} | = | \Phi'^{-1}_{x, \hat{u}_{(r)}} (x(y, 0, k) ; \hat{u}_r(y, 0, k)) | \le \rho \qquad (6.18)$$

The last inequality implies (6.14) if $\delta > 0$ is small enough. Now we will prove estimation (6.4). First

let us estimate the norm $| \hat{u}_{(m-r)}(\cdot) |$

Due to (6.9) for any small enough $\delta > 0$ there exists such small enough $\epsilon_0 > 0$ that for

$\forall (y, t, k) \in S(K, \delta)$

$$\max \{ |f_0(x(y,t,k)) - f_0(x(y,0,k))| , |f_i(x(y,t,k)) - f_i(x(y,0,k))| \mid i = r+1,...,m \} \leq e_0.$$

Therefore, in view of $f_i(x^*) \geq \sigma > 0$ for the passive constraints, we obtain

$f_i(x(y,t,k)) \geq \dfrac{\sigma}{2} > 0$, $i = r+1,...,m$ and for the Lagrange multipliers, that correspond to the passive

constraints we have

$$u_i(y,t,k) = u_i(f_0(y) - f_0(x(y,t,k))) \cdot (kf_i(x(y,t,k)) + f_0(y) - f_0(x(y,t,k)))^{-1}$$

$$= u_i(f_0(y) - f_0(x^*) + f_0(x^*) - f_0(x(y,t,k)))[kf_i(x^*) - k(f_i(x^*) - f_i(x(y,t,k)))$$

$$+ f_0(y) - f_0(x^*) - (f_0(x(y,t,k) - f_0(x^*))]^{-1}$$

$$\leq u_i(f_0(y) - f_0(x^*) + e_0)(kf_i(x^*) - (k+1)e_0 + f_0(y) - f_0(x^*))^{-1}$$

$$\leq \dfrac{2u_i}{k}(f_0(y) - f_0(x^*) + e_0)(f_i(x^*) - k^{-1}(k+1)e_0 + k^{-1}(f_0(y) - f_0(x^*)))^{-1}.$$

Hence, for small enough $e_0 < \dfrac{\sigma}{2} k(k+1)^{-1}$ we obtain

$$u_i(\bullet) = u_i(y,t,k) \leq \dfrac{2u_i(f_0(y) - f_0(x^*))}{k[\sigma - \dfrac{\sigma}{2} + (f_0(y) - f_0(x^*))k^{-1}]} \leq \dfrac{4u_i(f_0(y) - f_0(x^*))}{k\sigma} , \quad i = r+1,...,m$$

So we have

$$|u_{(m-r)}(y,t,k) - u^*_{(m-r)}| \leq \dfrac{4}{\sigma} k^{-1}(f_0(y) - f_0(x^*)) |u_{(m-r)} - u^*_{(m-r)}|$$

Now we will show that the estimation (6.4) holds for $\hat{x}(y,t,k)$ and $\hat{u}_{(r)}(y,t,k) =$

$(\hat{u}_i(y,t,k), i = 1,..., r)$. To this end we differentiate the identities (6.11) and (6.12) with respect

to t. Let $x'_t(\bullet) = J_t(x(\bullet)) = (x'_{j,t}(\bullet), j = 1,...,n)$, $\hat{u}'_{(r),t}(\bullet) = J_t(u_{(r)}(\bullet)) = (\hat{u}'_{i,t}(\bullet), i = 1,...,r)$ are

the Jacobians of the vector functions $x(\bullet)$ and $\hat{u}_{(r)}(\bullet)$. Also $\overline{L}(x(\bullet), \hat{u}_r(\bullet)) = f_0(x(\bullet)) -$

$\sum\limits_{i=1}^{r} \hat{u}_i(\bullet) f_i(x(\bullet))$ and let $J_t(h(x(\bullet), \bullet))$ and $J_t(g(x(\bullet), \bullet))$ are the Jacobians of the vector

functions $h(x(\bullet),\bullet)$ and $g(x(\bullet),\bullet)$. Then differentiating (6.11) with respect to t we obtain:

$$\overline{L_{xx}''}(x(\bullet),\hat{u}_{(r)}(\bullet))x_t'(\bullet) - (f_{(r)}'(x(\bullet))^T \hat{u}_{(r),t}'(\bullet)$$

$$- J_t(h(x(\bullet),\bullet)) + J_t(g(x(\bullet),\bullet) = 0 \qquad (6.19)$$

Let $F_{(r)}(x(\bullet)) = r\begin{bmatrix} f_0'(x(\bullet)) \\ \ldots \\ f_0'(x(\bullet)) \end{bmatrix}^{n}$, $\Psi(x(\bullet),y,t,k) = [\, \mathrm{diag}(t_i + k^{-1} u_i^* \Delta(y,x'))]_1^r \times$

$$[\, - d_{(r)}^{-1}(x(\bullet),y,k)F_{(r)}(x(\bullet)) - k\Delta(y,x(\bullet))d_{(r)}^{-2}(x(\bullet),y,k)f_{(r)}'(x(\bullet))$$

$$+ \Delta(y,x(\bullet)) \bullet d_{(r)}^{-2}(x(\bullet),y,k)F_{(r)}(x(\bullet))]\,,$$

then differentiating (6.12) with respect to t we obtain

$$\Psi(x(\bullet),y,t,k)x_t'(\bullet) - k^{-1}\Delta(y,x')\hat{u}_{(r),t}'(\bullet) =$$

$$- [\Delta(y,x(\bullet)) \bullet d_{(r)}^{-1}(x(\bullet),y,k)\,;\, 0^{r,m-r}] = S(x(\bullet),\bullet) \qquad (6.20)$$

Now we consider Jacobians $J_t(h(x(\bullet),\bullet))$ and $J_t(g(x(\bullet),\bullet))$ in more detail.

Recall that

$$h(x(y,t,k),y,t,k) = h(x(\bullet),\bullet) = \sum_{i=r+1}^{m} \hat{u}_i(x(\bullet),\bullet)(f_i'(x(\bullet)))^T = (u_{(m-r)}(x(\bullet),\bullet)f_{(m-r)}'(x(\bullet)))^T$$

Therefore

$$J_t(h(x(\bullet),\bullet)) = h_x'(x(\bullet),\bullet)x_t'(\bullet) + h_t'(x(\bullet),\bullet)$$

$$= [\sum_{i=r+1}^{m} \hat{u}_i(x(\bullet),\bullet)(f_i''(x(\bullet)) + f_{(m-r)}'^T(x(\bullet)))\hat{u}_{(m-r),x}'(x(\bullet),\bullet)]x_t'(\bullet) +$$

$$f_{(m-r)}'^T(x(\bullet))[0^{m-r,r}; k\Delta^{-1}(y,x^*)\Delta(y,x(\bullet))d_{(m-r)}^{-1}(x(\bullet),y,k)] =$$

$$N(x(\bullet),\bullet)x_t'(\bullet) + q(x(\bullet),\bullet)$$

Taking into account $x(y,0,k) = x^*$, $\hat{u}_i(x^*,y,0,k) = u_i^* = 0$, $i = r+1,...,m$,

$\hat{u}_{(m-r),x}'(x(y,0,k),y,0,k) = 0^{(m-r),n}$, we obtain $N(x^*,y,0,k) = 0^{n,n}$ and $q(x^*,y,0,k) =$

$f_{(m-r)}'^T(x^*)[0^{m-r,n}; kd_{(m-r)}^{-1}(x^*,y,k)]$.

Now let us consider the Jacobian

$$J_t(g(x(\bullet),\bullet)) = g_x'(x(\bullet),\bullet)x_t'(\bullet) + g_t'(x(\bullet),\bullet)$$

$$= g_x'(x(\bullet),\bullet)x_t'(\bullet) + \Delta^{-1}(y,x^*) F_m'^T(x(\bullet))[\operatorname{diag} kf_i'(x(\bullet))(kf_i'(\bullet)) + \Delta^{-1}(y,x(\bullet))]_1^m$$

$$= g_x'(x(\bullet),\bullet)x_t'(\bullet) + p(x(\bullet),\bullet)$$

For the Jacobian $g_x'(x(\bullet),\bullet)$ we obtain

$$g_x'(x(\bullet),\bullet) = k^{-1}[\sum_{i=1}^{m}(kt_i\Delta^{-1}(y,x(\bullet)) + u_i^*)(-1 + \Delta(y,x(\bullet))d_i^{-1}(x(\bullet),y,k))]f_0''(x(\bullet))$$

$$+ k^{-1}f_0'^T(x(\bullet))[\sum_{i=1}^{m}(kt_i\Delta^{-1}(y,x^*) + u_i^*) \times (-f_0'(x(\bullet))d_i^{-1}(x(\bullet),y,k)$$

$$- \Delta(y,x(\bullet))d_i^{-2}(x(\bullet),y,k)(kf_i'(x(\bullet)) - f_0'(x(\bullet)))] = G(x(\bullet),\bullet).$$

Taking into account (1.3) and $x(y,0,k) = x^*$, $[\operatorname{diag}(kt_i\Delta^{-1}(y,x^*) + u_i^*)]_{1/t=0}^r = U_r^*$,

34

$$F_{m} = F_{m}(x(0,y,k)) = m \begin{bmatrix} \overset{n}{f_0'(x^*)} \\ \dots \\ f_0'(x^*) \end{bmatrix} \text{ we obtain}$$

$$G(x^*,y,0,k) = -\Delta^{-1}(y,x^*)f_0'^T(x^*)f_0'(x^*),$$

$$p(x^*,0,y,k) = \Delta^{-1}(y,x^*)[0^{n,r}, F_{(m-r)}^T [\text{ diag } kf_i(x^*) d_i^{-1}(x^*,y,k)]_{r+1}^m]$$

Therefore,

$$J_t(g(x^*,y,0,k)) - \Delta^{-1}(y,x^*)f_0'^T(x^*)f_0'(x^*) x_t'(y,0,k) + p(x^*,0,y,k)$$

$$\Psi(x^*,y,0,k) = -U_r^* f_{(r)}'(x^*), S(x^*,y,0,k) = [I^r ; 0^{r,m-r}]$$

$$q(x^*,y,0,k) = k^{-1}[0^{n,r} ; f_{(m-r)}'^T(x^*) d_{(m-r)}^{-1}(x^*,y,k)]$$

We recall that

$$\Phi_{x,\hat{u}_{(r)}}'(x(\bullet),\hat{u}_{(r)}(\bullet),\bullet) = \Phi_{x,\hat{u}_{(r)}}'(\bullet) =$$

$$\begin{array}{cc} n & r \\ \begin{matrix} n \\ r \end{matrix} \begin{bmatrix} L_{xx}''(x(\bullet),\hat{u}_{(r)}(\bullet)) + G(x(\bullet),\bullet) + N(x(\bullet),\bullet) & -f_{(r)}'^T(x(\bullet)) \\ \Psi(x(\bullet),\bullet) & -k^{-1} \Delta(y,x^*) I^r \end{bmatrix} \end{array}$$

Then combining (6.19) and (6.20) we obtain

$$\begin{matrix} n \\ r \end{matrix} \begin{bmatrix} \overset{m}{x_t'(\bullet)} \\ \hat{u}_{(r),t}'(\bullet) \end{bmatrix} = (\Phi_{x,\hat{u}_{(r)}}'(\bullet))^{-1} \begin{matrix} n \\ r \end{matrix} \begin{bmatrix} \overset{m}{q(x(\bullet),\bullet) - p(x(\bullet),\bullet)} \\ S(x(\bullet),\bullet) \end{bmatrix} = (\Phi_{x,\hat{u}_{(r)}}'(\bullet))^{-1} R(x(\bullet),\bullet). \quad (6.21)$$

35

Now we consider the system (6.21) for $t = 0^-$. Taking into account

$$\overline{L''_{xx}}(x(y,0,k),\hat{u}_{(r)}(y,0,k) = L''_{xx}(x^*,u^*),\ G(x(y,0,k),y,0,k)$$

$$= -(f_0(y) - f_0(x^*))^{-1}f_0'^T(x^*)f_0'(x^*),\ N(x(y,0,k),y,0,k) = 0^{n,n},$$

$$\Psi(x(y,0,k);y,0,k) = -U_r^* f_{(r)}'(x^*),\ q(x(y,0,k);y,0,k) = q(x^*,y,0,k)$$

$$= k^{-1}[0^{n,r};f_{(m-r)}'^T(x^*)d_{(m-r)}^{-1}(x^*,y,k)],p(x(y,0,k),y,0,k) = p(x^*,y,0,k) =$$

$$[0^{n,r};F_{(m-r)}^T(x^*)[\operatorname{diag}f_i(x^*)d_i^{-1}(x^*,y,k)]_{r+1}^m],$$

$$S(x(y,0,k),y,0,k) = S(x^*,y,0,k) = (-I^r;0^{r,m-r})$$

we obtain the following system

$$\begin{bmatrix} x_t'(y,0,k) \\ \hat{u}_{(r),t}(y,0,k) \end{bmatrix} = \Phi_{(y,k)}^{-1} \begin{bmatrix} q(x^*,y,0,k) - p(x^*,y,0,k) \\ S(x^*,y,0,k) \end{bmatrix} = \Phi_{(y,k)}^{-1} \cdot \mathbf{R}(x^*,y,0,k). \quad (6.22)$$

Therefore

$$\max\{|x_t'(y,0,k)|, |\hat{u}_{(r),t}'(y,0,k)|\} \leq |\Phi_{(y,k)}^{-1}| |R(x^*,y,0,k)|. \quad (6.23)$$

Taking into account $\min\{\Delta(y,x^*)\forall y \in \Omega_t\} = \tau_0 > 0$,

$$|[\operatorname{diag}(f_i(x^*) + k^{-1}\Delta(y,x^*))^{-1}]_{r+1}^m| \leq \sigma^{-1},$$

$$|[\operatorname{diag}(f_i(x^*)(f_i(x^*) + k^{-1}\Delta(y,x^*))^{-1}]_{r+1}^m| \leq 1,$$

one obtains

$$| q(x^*, t, 0, k) | \leq \sigma^{-1} | f'^T_{(m-r)}(x^*) |, \; | p(x^*, y, 0, k) | \leq \tau_0^{-1} | F^T_{(m-r)}(x^*) |,$$

$$| S(x^*, 0, y, k) | \leq 1 \; \text{and} \; | R(x^*, y, 0, k) | \leq \sigma^{-1} | f'^T_{(m-r)}(x^*) | + \tau_0^{-1} | F^T_{(m-r)}(x^*) | + 1.$$

In view of (6.18) and (6.23) we have

$$\max\{ | x_t'(y, 0, k) |, | \hat{u}'_{(r), t}(y, 0, k) | \} \leq \rho (1 + \sigma^{-1} | f'^T_{(m-r)}(x^*) | + \tau_0^{-1} | F^T_{(m-r)}(x^*) |) = c_0$$

So there exists a small enough $\delta > 0$ such that for any $(y, t, k) \in S(K, \delta)$ the inequality

$$| (\Phi'_{x, \hat{u}_{(r)}}(x(y, \alpha t, k), \hat{u}_{(r)}(y, \alpha t, k); y, \alpha t, k))^{-1} R(x(y, \alpha t, k); y, \alpha t, k) |$$

$$\leq 2\rho (1 + \sigma^{-1} | f'^T_{(m-r)}(x^*) | + \tau_0^{-1} | F^T_{(m-r)}(x^*) |) = c_0 \tag{6.24}$$

holds true for any $0 \leq \alpha \leq 1$. Also we have

$$\begin{bmatrix} x(y, t, k) - x(y, 0, k) \\ \hat{u}_r(y, t, k) - \hat{u}_r(y, 0, k) \end{bmatrix} =$$

$$\int_1^0 \Phi'^{-1}_{x, \hat{u}_{(r)}}(x(y, \alpha t, k), \hat{u}_{(r)}(y, \alpha t, k), y, \alpha t, k) \cdot R(x(y, \alpha t, k); y, \alpha t, k) [t] d\alpha.$$

From (6.24) and (6.25) we obtain

$$\max\{ | x'(y, t, k) - x^* |, | \hat{u}'_{(r)}(y, t, k) - u_{(r)}^* | \} \leq c_0 | t | = c_0 k^{-1} \Delta(y, x^*) | u - u^* |.$$

37

Let $\hat{x}(y,u,k) = x(y,k^{-1}\Delta(y,x^{\cdot})(u-u^{\cdot}),k), \hat{u}(y,u,k) = (u_{(r)}(y,k^{-1}\Delta(y,x^{\cdot})(u-u^{\cdot}),k),$

$\hat{u}_{(m-r)}(y,k^{-1}\Delta(y,x^{\cdot})(u-u^{\cdot}),k))$ and $c = \max\{c_0, 4\sigma^{-1}\}$, then

$$\max\{|\hat{x}(y,u,k)-x^{\cdot}|, |\hat{u}(y,u,k)-u^{\cdot}|\} \leq ck^{-1}\Delta(y,x^{\cdot})|u-u^{\cdot}|.$$

So we ascertained the estimation (6.4). Also $\hat{x}(y,u^{\cdot},k) = x^{\cdot}$ and $\hat{u}(y,u^{\cdot},k) = u^{\cdot}$ follows from

(6.4) for any triple $(y, u^{\cdot}, k) \in D(\cdot)$ i.e. $u^{\cdot}$ is the fixed point of the mapping $u \to \hat{u}(y,u,k)$.

3) Now we will prove that $F(x, y, u, k)$ is strongly convex in a neighborhood of

$\hat{x} = \hat{x}(y,u,k)$ for any $(y,u,k) \in D(\cdot)$.

Using the formula for $F''_{xx}(\hat{x},y,u,k)$ (see Appendix A2) and taking into account the estimation (6.4)

we obtain for a small enough δ and for any triple $(y, u, k) \in D(\cdot)$ that

$$F''_{xx}(\hat{x},y,u,k) \approx \Delta^{-1}(y,x^{\cdot})[L''_{xx}(x^{\cdot},u^{\cdot}) + \Delta^{-1}(y,x^{\cdot})(kf'^{T}_{(r)}(x^{\cdot})U_r^{\cdot}f'_{(r)}(x^{\cdot}) - f'^{T}_0(x^{\cdot})f'_0(x^{\cdot}))$$

$$+ (k\Delta(y,x^{\cdot}))^{-1}(\Sigma(u_i^{\cdot}-u_i))f'^{T}_0(x^{\cdot})f'_0(x^{\cdot}) + k^{-1}(\Sigma(u_i^{\cdot}-u_i))f''_0(x^{\cdot})].$$

For any triple $(y, u, k) \in D(\cdot)$ we have $k \geq k_0 \Delta(y,x^{\cdot}) + \Sigma u_i^{\cdot}$, and

$k^{-1}|u_i - u_i^{\cdot}| \leq \delta\Delta^{-1}(y,x^{\cdot}), i = 1,...,m$. Keeping in mind $\min\{\Delta(y,x^{\cdot})| y \in \Omega_{\cdot}\} = \tau_0 > 0$ for

any $v \in \mathbb{R}^n$ we obtain

$$(F''_{xx}(\hat{x},y,u,k)v,v) \geq \Delta^{-1}(y,x^{\cdot})[((L''_{xx}(x^{\cdot},u^{\cdot}) + k_0 f'^{T}_{(r)}U_r^{\cdot}f'_{(r)}(x^{\cdot}))v,v)$$

$$+ \Delta^{-1}(y,x^{\cdot})(((\Sigma u_i^{\cdot})f'^{T}_{(r)}U_r^{\cdot}f'_{(r)}(x^{\cdot}) - f'^{T}_{(r)}(x^{\cdot})u_{(r)}^{\cdot T}u_{(r)}^{\cdot}f'_{(r)}(x^{\cdot}))v,v)$$

$$-(k\Delta(y,x'))^{-1}\sum|u_i^* - u_i|(f_0'(x'),v)^2 - k^{-1}\sum|u_i^* - u_i|(f_0''(x')v,v)$$

$$\geq \Delta^{-1}(y,x')[((L_{xx}''(x^*,u^*) + k_0 f_{(r)}'^T(x')U_r^* f_{(r)}'(x'))v,v)$$

$$- \delta m\tau_0^{-1}(\tau_0^{-1}(f_0'(x'),v)^2 + (f_0''(x')v,v))].$$

So due to Assertion 1, there exists a k_0 large enough such that

$$(F_{xx}''(\hat{x},y,u,k)v,v) \geq \Delta^{-1}(y,x')[\mu(v,v) - \delta\tau_0^{-1}m(\tau_0^{-1}(f_0'(x'),v)^2 + (f_0''(x')v,v))].$$

So for small enough $\delta > 0$ and any triple $(y, u, k) \in D(\bullet)$ there exists $0 < \hat{\mu} < \mu$:

$$F_{xx}''(\hat{x},y,u,k)v,v) \geq \Delta^{-1}(y,x')\hat{\mu}(v,v), \quad \forall v \in \mathbb{R}^n$$

i.e. for $\forall(y, u, k) \in D(\bullet)$, We have

$$\text{mineigval } F_{xx}''(\hat{x},y,u,k) \geq \Delta^{-1}(y,x')\hat{\mu}$$

To complete the proof we note that for any triple $(y, u, k) \in D(\bullet)$ we have

$$k^{-1}\sum u_i \leq (\sum u_i^*)(k_0\Delta(y,x') + \sum u_i^*)^{-1} + \delta m\Delta^{-1}(y,x') \leq (\sum u_i^*)(k_0\tau_0 + \sum u_i^*)^{-1} + \delta m\tau_0^{-1}.$$

Therefore if $0 < \delta < \dfrac{\tau_0}{m(1 + (\sum u_i^*)(k_0\tau_0)^{-1})}$ then $k^{-1}\sum u_i < 1$. So for small enough $\delta > 0$ the

function $F(x, y, u, k)$ is convex in $x \in \Omega_k(y)$ for any $(y, u, k) \in D(\bullet)$. Hence the vector

$\hat{x} = \hat{x}(y,u,k)$ is a unique minimum of the function $F(x, y, u, k)$ in $\Omega_k(y)$ and $F_x'(\hat{x},y,u,k) = 0$.

Due to the definition of $F(x, y, u, k)$ we obtain $\hat{x} = \text{argmin}\{F(x,y,u,k) \mid x \in \mathbb{R}^n\}$.

Using the formula for $F_{xx}''(\hat{x},y,u,k)$ one can find $\hat{M}$ such that for any triple $(y, u, k) \in D(\bullet)$

the estimate (6.6) is taking place.

We completed the proof of the basic theorem.

<u>Remark 3.</u>   All statements of Theorem 2 remain true for the MIDF $H(x, y, u, k)$. To prove it

we consider instead of $\Phi(x,\hat{u}_{(r)},y,t,k)$ the mapping $\Phi_H(x,\hat{u}_{(r)},y,t,k): \mathbb{R}^{2n+r+m+1} \to \mathbb{R}^{n+r}$

defined by

$$\Phi_H(x,\hat{u}_r,y,t,k) = (f'^T_0(x) - \sum_{i=1}^{r} \hat{u}_i f'^T_i(x) - h(x,y,t,k) + g(x,y,t,k);$$

$$k^{-1}\Delta(y,x') [(k\Delta^{-1}(y,x')t_i + u_i') \Delta^2(y,x) d_i^{-2}(x,y,k) - \hat{u}_i], i = 1,...,r)$$

where

$$h(x,y,t,k) = \sum_{i=r+1}^{m} \hat{u}_i(x,y,t,k)(f'_i(x))^T,$$

$$g(x,y,t,k) = k^{-1}\{ \sum_{i=1}^{m} (kt_i\Delta^{-1}(y,x') + u_i') [-1 + \Delta^2(y,x) d_i^{-2}(x,y,k)]\} (f'_0(x))^T$$

and

$$\hat{u}_i(x,y,t,k) = k\Delta^{-2}(x,y')t_i \Delta^2(y,x) d_i^{-2}(x,y,k), i = r+1,...,m$$

The MIDFs and MBFs [see(Pol 92)] have some common features, however, there are essential differences between them as well.

We will consider few small examples to illustrate some of the differences.

7. **Examples**. Let us consider a convex programming problem

$$x^* = \text{argmin} \{f_0(x) = x \mid f(x) = -x^2 + x \geq 0\} = 0 \tag{7.1}$$

The corresponding Classical Lagrangian is $L(x,u) = x - u(-x^2 + x)$, then $L'_x(x^*,u^*) = 1 + 2u^*x^* - u^* = 0$, i.e. $u^* = 1$. The feasible set $\Omega = \{x: -x^2 + x \geq 0\} = [0,1]$. Now

we fixed $0 < y < 1$, then the following problem

$$x^* = \text{argmin}\{-\ln(y-x) \mid k^{-1}[\ln(k(-x^2+x)+y-x) - \ln(y-x)] \geq 0, y > x\}$$

is equivalent to (7.1) and the corresponding MIDF is:

$$F(x,y,u,k) = -\ln(y-x) - k^{-1}u[\ln(k(-x^2+x)+y-x) - \ln(y-x)]$$

$$= (-1+uk^{-1})\ln(y-x) - k^{-1}u\ln(k(-x^2+x)+(y-x))$$

So $F(x,y,u^*,k) = (-1+k^{-1})\ln(y-x) - k^{-1}\ln(-kx^2+(k-1)x+y)$ and

$F_x'(x,y,u^*,k) = (1-k^{-1})(y-x)^{-1} + (2x-1+k^{-1})(-kx^2+kx+y-x)^{-1}$, therefore

$F_x'(x^*, y, u^*, k) = 0$. Let us consider $F_{xx}''(x, y, u^*, k)$. We obtain $F_{xx}''(x, y, u^*, k) =$

$(1-k^{-1})(y-x)^{-2} + (2kx^2+2(y-x)+4x-2kx+(k-1)^2k^{-1})(-kx^2+(k-1)x+y)^{-2}$.

Therefore $F_{xx}''(x^*,y,u^*,k) = (2y+k-1)y^{-2}$ and taking $y = \frac{1}{8}$ and $k = \frac{1}{2}$ we obtain

$F_{xx}''(x^* ; \frac{1}{8} ; u^*, \frac{1}{2}) = -16$, i.e. the Classical Lagrangian for the equivalent problem is strongly

concave at the solution, while problem (1.7) is convex and the Classical Lagrangian $L(x, u^*) = x^2$ for

problem (1.7) is strongly convex at the solution $x^* = 0 = \text{argmin } L(x, u^*)$.

Moreover, $F(x,\frac{1}{8},u^*,\frac{1}{2}) = \ln(\frac{1}{8}-x) - 2\ln(-x^2-x+\frac{1}{4}) + 2\ln 2$, so for $x = \frac{1}{8} - e$

we obtain $F(\frac{1}{8}-e, \frac{1}{8}, u^*, \frac{1}{2}) = \ln e - 2\ln(-(\frac{1}{8}-e)^2 - (\frac{1}{8}-e) + \frac{1}{4}) + 2\ln 2 =$

$\ln e - 2\ln(\frac{5}{4}e - e^2 + \frac{7}{64}) + 2\ln 2 = F(e)$. Therefore $\inf_{0 \leq x \leq 1} F(x, \frac{1}{8}, u^*, \frac{1}{2}) = \inf_{e \to 0} F(e) = -\infty$.

This example shows that without the condition $k > \Sigma u_i^*$ Theorem 1 and Proposition 3 are invalid

even for the convex programming problem.

In case of Modified Barrier Function the situation is different. For the MBF which corresponds to

(7.1) we obtain $F(x,u,k) = x - k^{-1}u\ln(k(-x^2+x)+1)$ then $F(x^*,u^*,k) = f_0(x^*) = x^* = 0$

$F_x'(x^*,u^*,k) = 0$ and $F_{xx}''(x^*,u^*,k) > 0, x^* = \text{argmin }\{F(x,u^*,k) \mid x \in R^1\}$ for any $k > 0$ i.e.

41

the Proposition and the first part of Theorem 1 remain true in the case of the MBF for any $k > 0$ if the problem is convex.

As for the MIDF $F(x, y, u, k)$ then the Proposition 3 and the first part of Theorem 1 are true only if $k > \sum u_i^*$.

As far as the Theorem 2 is concerned then the results for the MBF remain true for MIDF only if $k \geq k_0 \Delta(y,x^*) + \sum u_i^*$, however instead of the estimation

$$\max\{\, |\hat{x} - x^*\,|, |\hat{u} - u^*\,|\,\} \leq ck^{-1}\,|u - u^*\,| \tag{7.2}$$

(see [Pol92] p185) for the MBF, we obtain the estimation (6.4) for MIDF, i.e.

$$\max\{\, |\hat{x} - x^*\,|, |\hat{u} - u^*\,|\,\} \leq ck^{-1}\,\Delta(y,x^*)\,|u - u^*\,|\,.$$

Therefore one can improve the convergence as compared to MBF method by choosing a fixed $y \in \Omega$, that $\Delta(y,x^*) < 1$. Now we will show that it is possible for all statements of Theorem 1 and Theorem 2 to remain true, even when the functions $f_i(x)$, $i = 1,\ldots, m$ are non-concave.

To show this we consider the following problem:

$$x^* = \operatorname{argmin}\{f_0(x) = x \,|f(x) = e^x - 1 \geq 0\} = \operatorname{argmin}\{x\,|\, -e^x + 1 \leq 0\} = 0 \tag{7.3}$$

The function $f(x) = -e^x + 1$ is strongly concave, therefore, the Classical Lagrangian $L(x, u) = x + u(-e^x + 1)$ is strongly concave for any $u > 0$. Then $L_x'(x^*, u) = 0 \Rightarrow 1 - u = 0$ i.e. $u^* = 1$, so $L(x,u^*) = x - e^x + 1$ and $\inf L(x,u^*) = -\infty$, moreover $\inf L(x,u) = -\infty$ for any $u > 0$.

Now let us consider the MIDF, which corresponds to problem (7.3). We obtain $F(x, y, u, k) =$

$$-\ln(y-x) - k^{-1}u(\ln(k(e^x - 1) + y - x) - \ln(y-x)) = (-1 + k^{-1}u)\ln(y-x) - k^{-1}u$$

$\ln(k(e^x - 1) + y - x)$. So $F(x,y,u^*,k) = (-1 + k^{-1})\ln(y-x) - k^{-1}\ln(k(e^x - 1) + y - x)$,

$F_x'(x,y,u^*,k) = (1 - k^{-1})(y-x)^{-1} - (e^x - k^{-1})(k(e^x - 1) + y - x)^{-1}$ and $F_x'(x^*,y,u^*,k) = 0$

for any $y > 0$ and $k > 0$. Then $F_{xx}''(x,y,u^*,k) = (1 - k^{-1})(y-x)^{-2} + ((k-2)e^x - (y-x)e^x + k^{-1})$

42

$(k(e^{x}-1)+y-x)^{2})^{-1}$ and $F_{xx}''(x^{*},y,u^{*},k)=(k-y-1)y^{-2}$. Therefore for any $y\geq\tau>0$ and

any $k>y+1$ the MIDF $F(x,y,u^{*},k)$ is strongly convex in a neighborhood of the solution

$x^{*}=\text{argmin}\{F(x,y,u^{*},k)\mid y>x\}=0$ while the Classical Lagrangian $L(x,u^{*})=x-e^{x}+1$ is

strongly concave in the same neighborhood and $\min_{x}L(x,u^{*})=-\infty$.

Now we would like to make a few comments about the estimate (6.4) and its relation with the

corresponding estimate (7.2) for the MBF. Let us consider the matrix $\Phi_{(y,k)}$ for problem (7.3). We

have

$$\Phi_{(y,k)}=\begin{bmatrix}-1-y^{-1} & -1 \\ -1 & -yk^{-1}\end{bmatrix},\text{ so }\Phi_{(y,k)}^{-1}=\begin{bmatrix}\dfrac{y}{k-y-1} & \dfrac{-k}{k-y-1} \\ \dfrac{-k}{k-y-1} & \dfrac{k(y+1)}{y(k-y-1)}\end{bmatrix}$$

It is easy to see that $q(x^{*},y,0,k)=p(x^{*},y,0,k)=0$ and

$$R(x^{*},y,0,k)=\begin{bmatrix}q(x^{*},y,0,k)-p(x^{*},y,0,k) \\ S(x^{*},y,0,k)\end{bmatrix}=\begin{bmatrix}0 \\ -1\end{bmatrix}.$$

Therefore, taking into account (6.22) we obtain

$$\begin{bmatrix}\hat{x}'(y,0,k) \\ \hat{u}_{t}(y,0,k)\end{bmatrix}=\Phi_{(y,k)}^{-1}R(x^{*},y,0,k)=\begin{bmatrix}\dfrac{k}{k-y-1} \\ -\dfrac{k(y+1)}{y(k-y-1)}\end{bmatrix}.$$

For all $(y,u,k)\in D(\cdot)$ there exists

$\hat{x}=\hat{x}(y,u,k)=\text{argmin}\{-\ln(y-x)-k^{-1}u(\ln(k(e^{x}-1)+y-x)-\ln(y-x))\mid y>x\}$

43

Also for $\hat{x}$ and $\hat{u} = \hat{u}(y,u,k) = (y - \hat{x})(k(e^{\hat{x}} - 1) + y - \hat{x})^{-1}$ and $t = y|u - u^*|k^{-1}$ small enough

we have $|\hat{x} - x^*| \le 2k(k - y - 1)^{-1}t$, $|\hat{u} - u^*| \le 2k(y + 1)(y(k - y - 1))^{-1}t$, i.e. the following

estimations

$$|\hat{x} - x^*| \le 2y(k - y - 1)^{-1}|u - u^*| \tag{7.4}$$

$$|\hat{u} - u^*| \le 2(y + 1)(k - y - 1)^{-1}|u - u^*| \tag{7.5}$$

holds true.

Now we apply the MBF to the same problem (7.3). We obtain $F(x,u,k) = x - k^{-1}u\ln(k(e^x - 1) + 1)$.

So $F(x^*,u^*,k) = f_0(x^*) = 0$, $F_x'(x,u^*,k) = 1 - e^x(k(e^x - 1) + 1)^{-1}$, $F_x'(x^*,u^*,k) = 0$,

$F_{xx}''(x,u,k) = (k - 1)ue^x(ke^x - k + 1)^{-2}$ and $F_{xx}''(x^*,u^*,k) = k - 1$. Note that for any $k > 1$ the

MBF $F(x, u^*, k)$ is strongly convex at the solution while the Classical Lagrangian $L(x, u^*) =$

$x - e^x + 1$ is strongly concave. Then (see [Pol92] p185)

$$\Phi_k = \begin{bmatrix} L_{xx}'' & -f' \\ -u^* f' & -k^{-1} \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ -1 & -k^{-1} \end{bmatrix}, \ \Phi_k^{-1} = \begin{bmatrix} \dfrac{1}{k-1} & -\dfrac{k}{k-1} \\ -\dfrac{k}{k-1} & \dfrac{k}{k-1} \end{bmatrix}$$

$R(0,k) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ therefore

$$\begin{bmatrix} \hat{x}_t'(0,k) \\ \hat{u}_t(0,k) \end{bmatrix} = \Phi_k^{-1} R(0,k) = \begin{bmatrix} \dfrac{k}{k-1} \\ -\dfrac{k}{k-1} \end{bmatrix}$$

Due to Theorem 1 (see [Pol92]) for all $(u, k) \in D(\bullet)$ there exists $\hat{x} = \hat{x}(u,k) =$

argmin$\{F(x,u,k) \mid x \in R^1\}$ and for $\hat{x}$ and $\hat{u} = u(k(e^{\hat{x}} - 1) + 1)^{-1}$ the following estimations

44

$$|\hat{x} - x^*| \le 2(k-1)^{-1}|u - u^*|,\qquad\qquad (7.6)$$

$$|\hat{u} - u^*| \le 2(k-1)^{-1}|u - u^*|\qquad\qquad (7.7)$$

holds true. Comparing the estimation (7.4) and (7.6) we find that the extra tool in the MIDF has an essential influence on the rate of convergence for the primal sequence. By changing the "center" y it is possible to speed up the convergence of the primal sequence even with fixed parameter $k > 0$.

8. Shifted Interior Distance Function (SIDF)

The SIDF one obtains from the MIDF by setting $u = e = (1,...,1) \in \mathbf{R}_+^m$, i.e. the SIDF, which corresponds to $F(x, y, u, k)$ is defined by formula:

$$\varphi(x,y,k) = F(x,y,e,k) = (-1 + k^{-1}m)\ln(f_0(y) - f_0(x)) - k^{-1}\sum_{i=1}^{m}\ln(kf_i(x) + f_0(y) - f_0(x))$$

$$\qquad\qquad (8.1)$$

$$= (-1 + k^{-1}m)\ln\Delta(y,x) - k^{-1}\sum_{i=1}^{m}\ln d_i(x,y,k)$$

If A1 and A2 are taking place then the set $\Omega_k(y)$ is bounded. For any $k > m$ and $y \in \Omega$, SIDF $\varphi(x, y, k)$ is convex.

Therefore there exists

$$x(\bullet) = x(y,k) = \mathrm{argmin}\{\varphi(x,y,k)\,/\,x \in \mathbf{R}^n\}$$

and

$$\varphi_x'(x(\bullet),\bullet) = (1 - mk^{-1})\Delta^{-1}(y,x(\bullet)) - \sum_{i=1}^{m} d_i^{-1}(x(\bullet),\bullet)(f_i'(x(\bullet)) - k^{-1}f_0'(x(\bullet))) = 0 \quad (8.2)$$

By setting

$$u_i(\bullet) = u_i(y,k) = \Delta(y,x(\bullet))d_i^{-1}(x(\bullet),\bullet)\,,\, i = 1,...,m \qquad\qquad (8.3)$$

and $u(\bullet) = (u_i(\bullet)\,,\, i = 1,...,m)$ we can rewrite (8.2) as follows:

$$f_0'(x(\bullet)) - \sum_{i=1}^{m} u_i(\bullet)f_i'(x(\bullet)) + k^{-1}[\sum_{i=1}^{m} u_i(\bullet) - m]f_0'(x(\bullet)) = 0^n \qquad (8.4)$$

The following proposition is taking place.

<u>Proposition 5</u> . If conditions A1 and A2 are satisfied, then for any monotone increasing sequence

$\{ k_s \} : k_1 > m$, $\lim k_s = \infty$ and $y \in \Omega_r$:

1) the sequence $\{w\,(y,k_s) = (\,x\,(y,\,k_s\,),\,u\,(y,\,k_s\,))\}$ is bounded and any limit point $(\bar{x},\,\bar{u})$ of

$\{w\,(y,k_s)\}$ is a K-K-T's pair

$$L_x'(\bar{x},\bar{u}) = 0 \,, \ \bar{u}_i f_i(\bar{x}) = 0 \,, \ \bar{u}_i \geq 0 \,, \ i = 1,...,m \,,$$

i.e. $\bar{x} = x^* , \bar{u} = u^*$

2) $\lim \varphi\,(x(y,k_s),y,k_s) = \varphi^* = \ln(f_0(y) - f_0(x^*))^{-1}$ i.e. $f_0(x^*) = f_0(y) - \exp(-\varphi^*)$

<u>Proof</u>   Due to the Corollary 20 (see [FiacM68]p94) it follows from assumptions A1-A2 that the set

$\Omega_{k_1}(y)$ is bounded, therefore in view of the inclusions $\Omega_{k_1}(y) \supset \Omega_{k_2}(y) \supset ... \supset \Omega_{k_s}(y) \supset \Omega_{k_{s+1}}(y)...$

the sequence $\{x\,(\,y,\,k_s\,)\}$ is bounded for any $y \in \Omega_r$. So it contains a converging subsequence. We

can assume without loosing generality that

$$\lim_{s \to \infty} x(y,k_s) = \bar{x}$$

It is clear that $\bar{x}$ depends on $y \in$ int Ω , we will omit this indication to simplify notations.

Then $\lim \varphi\,(x(\cdot),\cdot) = \lim_{s \to \infty} \varphi\,(x(y,k_s),y,k_s) = \lim_{s \to \infty} [(-1 + mk_s^{-1})\ln(f_0(y) - f_0(x(\cdot)))$

$$- k_s^{-1} \sum_{i=1}^{m} \ln\,(f_i(x(\cdot)) + (f_0(y) - f_0(x(\cdot)))k_s^{-1}) + mk_s^{-1}\ln k_s\,]$$

$$= \lim_{s \to \infty} [\,-\ln(f_0(y) - f(x(\cdot))) - k_s^{-1} \sum_{i=1}^{m} \ln f_i(x(\cdot))] \qquad\qquad (8.5)$$

Therefore $\bar{x} \in \Omega$, moreover $x \in \partial\,\Omega$, i.e. there is at least one index $i : f_i(\bar{x}) = 0$, otherwise we would

be able to find $\lambda > 0$:

$$f_i(x(y,k_s)) \geq \lambda \,, \ \forall_i = 1,...,m$$

for any large enough k_s.

46

Therefore using formula (8.3) for the Lagrange multipliers we obtain

$$\lim_{k_s \to \infty} u_i(y, k_s) = 0, \, i = 1, \dots, m$$

So, from (8.4) we have $f'_0(\bar{x}) = 0$ $\bar{x} \in \Omega$ which is impossible, because (1.1) is a constrained optimization problem.

Hence, $\bar{x} \in \partial \Omega$ and we will assume that $\bar{I} = \{ i : f_i(\bar{x}) = 0 \}$ is the active set.

Now we want to prove that $\{ u(\bullet) = u(y, k_s) \}$ is a bounded sequence. Assuming the opposite we can find $i : u_i(\bullet) = u_i(y, k_s) \to \infty$. Dividing both sides of (8.4) by $\sum_{i=1}^{m} u_i(\bullet)$ we obtain

$$(\sum_{i=1}^{m} u_i(\bullet))^{-1} f'_0(x(\bullet)) - \sum_{i=1}^{m} \bar{u}_i(\bullet) f'_i(x(\bullet)) + k_s^{-1} (1 - m(\sum u_i(\bullet))^{-1}) f'_0(x(\bullet)) = 0^n \qquad (8.6)$$

where $\bar{u}_i(\bullet) = u_i(\bullet) (\sum u_i(\bullet))^{-1}, \, i = 1, \dots, m$

Taking a limit in (8.6) we obtain

$$\sum_{i \in \bar{I}} \bar{u}_i f'_i(\bar{x}) = 0 \text{ and } \bar{u}_i \geq 0, \, i \in \bar{I} \qquad (8.7)$$

and not all $\bar{u}_i = 0$.

However, (8.7) is impossible because it contradicts Slater condition (1.2), so $\{ u(y, k_s) \}$ is bounded.

Without losing generality we can assume

$$\bar{u} = \lim_{k_s \to \infty} u(y, k_s)$$

Taking the limit in (8.4) we obtain

$$f'_0(\bar{x}) - \sum \bar{u}_i f'_i(\bar{x}) = 0 \text{ and } \bar{u}_i f_i(\bar{x}) = 0, \, i = 1, \dots, m$$

In view of $\bar{u} \in \mathbb{R}^m_+$ and $\bar{x} \in \Omega$ we have $\bar{x} = x^*, \bar{u} = u^*$ i.e. $(\bar{x}, \bar{u})$ is a K-K-T's pair.

To find $\lim_{k_s \to \infty} \varphi(x(y, k_s), y, k_s)$ we consider first the passive contraints $i : f_i(\bar{x}) > 0$. It is clear that for $y \in \Omega$,

$$\lim k_s^{-1} \ln f_i(x(y, k_s)) = 0, \, i \in \bar{I} \qquad (8.8)$$

47

Now we are going to consider the active constraints keeping in mind the formulas for the Lagrange multipliers (5.7). We have

$$0 \le \overline{u}_i = \lim_{s \to \infty} \frac{f_0(y) - f_0(x(\cdot))}{k_s f_i(x(\cdot)) + f_0(y) - f_0(x(\cdot))} , \quad i \in \overline{I}$$

For $\overline{u}_i > 0$ we obtain $f_i(x(\cdot)) = 0$ (k_s^{-1}), for $\overline{u}_i = 0$ we have $\lim_{s \to \infty} k_s f_i(x(\cdot)) = \infty$, therefore for any $y \in \Omega_\tau$

$$\lim_{s \to \infty} k_s^{-1} \ln f_i(x(\cdot)) = 0 , i \in \overline{I} \tag{8.9}$$

Using (8.6) and taking into account (8.8) and (8.9) for any $y \in \Omega_\tau$ we obtain

$$\varphi^* = \lim_{s \to \infty} \varphi(x(y,k_s),y,k_s) = \lim_{s \to \infty} [-\ln (f_0(y) - f_0(x(y,k_s)) - k_s^{-1} \sum_{i \in I} \ln f_i(x(y,k_s)))]$$

$$= \lim_{s \to \infty} [-\ln (f_0(y) - f_0(x(y,k_s)))] = \ln [f_0(y) - f_0(x^*)]^{-1} . \text{ Therefore } f_0(x^*) = f_0(y) - e^{-\varphi^*}$$

If for the problem (1.1) the standard second order optimality conditions are satisfied then the following statement, which is a corollary of the Basic Theorem, is taking place.

<u>Assertion 2</u>   If the second order optimality conditions (1.4) - (1.5) are satisfied, then there exists $k_0 > 0$ such that for any $y \in \Omega_\tau$ and $k \ge k_0 \Delta(y,x^*) + \sum u_i^* > m$ the following statements are taking place.

1) there exists a vector

$$x(y,k) = \text{argmin} \{ \varphi(x,y,k)/x \in \mathbb{R}^n \}$$

such that $\varphi_x'(x(y,k),y,k) = 0^n$.

2) for the pair of vectors $x(y,k)$, $u(y,k) = (u_i(y,k) = \Delta(y,x(y,k)) d_i^{-1}(x(y,k),y,k) , i = 1,...,m)$ the following estimation

$$\max \{ |x(y,k) - x^*|, |u(y,k) - u^*| \} \le ck^{-1} \Delta(y,x^*) \tag{8.10}$$

48

holds and $c > 0$ is independent on k and y.

3) the shifted interior distance function φ (x, y, k) is strongly convex in a neighborhood of x (y, k).

9. Numerical Realization of the MCM   The numerical realization of the MCM requires:

1) to find a "center" $y \in \text{int } \Omega_\tau$;

2) to find a triple $(y, u, k) \in D(\bullet)$;

3) to replace the infinite procedure of finding $\hat{x} = \hat{x}(y, u, k)$ for a finite procedure, which retains the contractibility properties of the operator $C_{y,k}$.

To find $y \in \text{int } \Omega_\tau$ we consider the following convex programming problem

$$\max\{ x_{n+1} / f_i(x) - x_{n+1} \geq 0 , \ i = 1,...,m \} \tag{9.1}$$

Starting with a "warm" start (x^0, x_{n+1}^0) : $x_{n+1}^0 < \min \{ f_i(x^0)/1 \leq i \leq m \}$ we apply any IPM for solving (9.1) up to the point $(\bar{x}, \bar{x}_{n+1})$: $\bar{x}_{n+1} \geq \tau$, then we set $y = \bar{x}$.

To find a triple $(y, u, k) \in D(\bullet)$ one can again apply an IPM, starting with $x = y$, $u = e_m$ and $k > m$ and increasing the barrier parameter followed by a Newton step for solving the system $\varphi'_x (x, y, k) = 0^n$ in x.

The other option is to find an approximation for a vector x (y, k) : $\varphi'_x (x(y, k), y, k) = 0^n$, using smooth unconstrained optimization techniques. In particular, one can use Newton method with step length, which we will describe later.

Due to Proposition 5 or Assertion 2 to find $(y, u, k) \in D(\bullet)$ it is enough to find an approximation for $x(y, k)$ and $u(y, k)$ when $k > m$ is large enough.

Also the parameter $k > m$ has to be large enough to guarantee that $C_{y,k}$ is a contractive operator for a given $y \in \Omega_\tau$, i.e. $0 < \gamma_{y,k} = c k^{-1} \Delta(y, x^*) < 1$. The constant $c > 0$ is associated with

the input data and the size of the problem and apriori unknown as well as the value $\Delta(y,x^*)$ for a chosen $y \in \Omega_\tau$.

So there is no explicit way to find for a given $0 < \gamma < 1$ a "center" $y \in \Omega_\tau$ and $k > 0$ such that $\gamma < c\,k^{-1}\,\Delta(y,x^*) < 1$.

Therefore we have to find an implicit way to adjust $y \in \Omega_\tau$ and $k > 0$ to an apriori chosen $0 < \gamma < 1$ to guarantee the inequality $\gamma < c\,k^{-1}\,\Delta(y,x^*) < 1$.

First, let's consider for a fixed $y \in \Omega_\tau$ and $k > \Sigma u_i$, a non negative function $v(w, y, k) = v(x, y, u, k) : W_{y,k} = \Omega_k(y) \times U_{y,k} \to \mathbf{R}_+^1$, which is defined by formula

$$v(w,y,k) = \max\{\ |F_x'(x,y,u,k)|, \ \max_{1 \le i \le m}\{-f_i(x)\}, \ \sum_{i=1}^{m} u_i\,|f_i(x)|\}$$

The following proposition is a consequence of the K-K-T's optimality conditions.

<u>Proposition 6</u>. For a convex programming problem (1.1), any fixed $y \in \Omega_\tau$, $u \in \mathbf{R}_+^m$ and $k > \Sigma u_i$ the following statement is true

$$v(x,y,u,k) = 0 \ \leftrightarrow \ (x,u) = (x^*,u^*) = w^* \tag{9.2}$$

<u>Proof</u>. First, $v(x,y,u,k) = 0 \to \max\{-f_i(x)\,/\,1 \le i \le m\} \le 0$, i.e. $-f_i(x) \le 0$ or $f_i(x) \ge 0$, $i = 1,\dots, m$, so $x \in \Omega$.

Second, $v(x, y, u, k) = 0 \to \sum_{i=1}^{m} u_i\,|f_i(x)| = 0$, i.e. $u_i\,|f_i(x)| = 0$, $i = 1,\dots,m$,

Therefore for any $1 \le i \le m$

$$u_i > 0 \ \to \ f_i(x) = 0 \quad \text{and} \quad f_i(x) > 0 \ \to \ u_i = 0 \tag{9.3}$$

Third, $v(x,y,u,k) = 0 \to |F_x'(x,y,u,k)| = 0$

Therefore in view of (9.3) we obtain

50

$$F_x'(x,y,u,k) = \frac{1-k^{-1}\sum u_i}{f_0(y)-f_0(x)}f_0'(x) - \sum_{i=1}^{m}\frac{u_i\,[f_i'(x)-k^{-1}f_0'(x)]}{kf_i(x)+f_0(y)-f_0(x)}$$

$$= \Delta^{-1}(y,x)(f_0'(x)-\sum u_i f_i'(x)) = 0^n \text{ and } \Delta(y,x)>0$$

Hence, $v(x,y,u,k)=0$ leads to $x \in \Omega$ and $L_x'(x,u)=0^n$, $\sum_{i=1}^{m} u_i f_i(x)=0$, $u \in \mathbf{R}_+^m$, so

$w=(x,u)=(x^*,u^*)=w^*$.

Let's consider the second part of the statement. We have to prove that $v(x^*,y,u^*,k)=0$.

First, $x^* \in \Omega$, therefore $\max\{-f_i(x^*)/1 \le i \le m\} \le 0$, then in view of (x^*,u^*) is a K-K-T's

pair, we have $\sum u_i^*|f_i(x^*)|=\sum u_i^* f_i(x^*)=0$ and due to the Proposition 2 $F_x'(x^*,y,u^*,k)=$

$\Delta^{-1}(y,x^*)L_x'(x^*,u^*)=0^n$. Therefore $v(x^*,y,u^*,k)=0$.

For a fixed "center" $y \in \Omega_r$ and a fixed barrier parameter $k > \sum_{i=1}^{m} u_i^*$ one can consider the

non-negative function $v(w,y,k): W_{y,k}=\Omega_k(y)\times U_{y,k} \to \mathbf{R}_+^1$ as a merit function, which measures

the proximity from $w=(x,u)$ to $w^*=(x^*,u^*)$.

Due to the smoothness of $f_i(x)$, $i=0,1,\ldots,m$ and boundness of $W_{y,k}$ there is a constant L that

$$v(w,y,k)=v(w,y,k)-v(w^*,y,k)\le L\,|w-w^*|, \quad \forall w \in W_{y,k} \tag{9.4}$$

Without loosing the generality we can assume that $L \le 1$.

Further, if for a given $0<\gamma<1$ the "center" $y \in \Omega_r$ and the barrier parameter $k > 0$ are

chosen appropriately, i.e. $\gamma_{y,k}=c\,k^{-1}\Delta(y,x^*)<\gamma<1$ then, due to (6.4) for any $u \in U_{y,k}$ we have

$$\max\{\,|\hat{x}-x^*|,\ |\hat{u}-u^*|\,\}\le \gamma\,|u-u^*|$$

and again $\hat{u} \in U_{y,k}$.

Let's assume that for $u^0 \in U_{y,k}$ we have $|u^0-u^*| \le 1$, then MCM (5.6) - (5.7) produces

51

a sequence $\{ w^s = (x^s, u^s) \}$:

$$\max \{ | x^s - x^* |, \quad | u^s - u^* | \} \le \gamma^s \tag{9.5}$$

Then in view of (9.4) we obtain

$$v(w^s, y, k) = v(w^s, y, k) - v(w^*, y, k) \le \gamma^s \tag{9.6}$$

In other words if for a given $0 < \gamma < 1$ the "center" and the barrier parameter are such, that

$\gamma_{y,k} \le \gamma$, then $\{ \gamma^s \}_{s=0}^{\infty}$ is a majorant for the sequence $\{ v(w^s, y, k) \}_{s=1}^{\infty}$.

We will say that the pair (y, k) is <u>consistent</u> with a given ratio $0 < \gamma < 1$ if (9.6) is taking

place for all $s \ge 1$.

To check the consistency of a chosen pair (y, k) for a given ratio $0 < \gamma < 1$ one has to solve

infinite number of unconstrained optimization problems. Moreover, each problem requires infinite

number of arithmetic operations.

So, first we will show how to keep the estimation (6.4) without solving an unconstrained opti-

mization problem at every step.

Then we will show that the numerical realization of MCM does not require the consistency

of chosen couple (y, k) with a given ratio $0 < \gamma < 1$ from the beginning of the process. We will

achieve the consistency by adopting the barrier parameter in the process of solution.

Let's consider a number $\theta > 0$, $k > \Sigma u_i$ and a pair $(\bar{x}, \bar{y})$:

$$| F_x'(\bar{x}, y, u, k) | \le \theta k^{-1} | \Delta (y, \bar{x}) d^{-1}(\bar{x}, y, k) u - u | \tag{9.7}$$

We update the Lagrange multipliers by the following formula

$$\bar{u} = \Delta (y, \bar{x}) d^{-1}(\bar{x}, y, k) u \tag{9.8}$$

The following assertion is taking place.

<u>Assertion 2</u> If the second order optimality conditions (1.4) - (1-5) are satisfied then for any $\mu > 0$

and $y \in \Omega$, there exists such $k_0 > 0$ that for any $(y, u, k) \in D(\bullet)$ the following estimate holds

$$\max\{ \, |\bar{x} - x^*| \, , \, |\bar{u} - u^*| \, \} \leq ck^{-1}\Delta(y,x^*)(1+\theta) \, |u - u^*| \qquad (9.9)$$

The estimate (9.9) can be proved by a slight modification of the considerations, which have been use

to prove Lemma 2 (see[Pol92]).

So the consistency of a couple (y, k) with a given ratio $0 < \gamma < 1$ one can check using instead

of (9.6) the following inequality

$$v(\bar{w}^s, y, k) \leq \gamma^s \, , \quad s \geq 1 \qquad (9.10)$$

where $\bar{w}^s = (\bar{x}^s, \bar{u}^s)$:

$$| \, F_x'(\bar{x}^{s+1}, y, \bar{u}^s, k) \, | \leq \theta k^{-1} | \, \Delta(y, \bar{x}^{s+1}) \, d(\bar{x}^{s+1}, y, k) \, \bar{u}^s - \bar{u}^s \, | \qquad (9.11)$$

$$\bar{u}^{s+1} = \Delta(y, \bar{x}^{s+1}) \, d^{-1}(\bar{x}^{s+1}, y, k) \, \bar{u}^s \qquad (9.12)$$

Due to the Assertion 2 for any fixed "center" and a given ratio $0 < \gamma < 1$ there exists a threshold

$k_{y,\gamma} \geq k_0$ that for any $k \geq k_{y,\gamma}$ we have $ck^{-1}\Delta(y,x^*)(1+\theta) \leq \gamma < 1$ and $u \in U_{y,k} \rightarrow \bar{u} \in U_{y,k}$.

In other words, for $k \geq k_{y,\gamma} > k_0$ and $u \in U_{y,k}$:

1) the approximation $\bar{x}$ that satisfies (9.7) can be found for a finite number of operations;

2) after every Lagrange multipliers update, the distance from the approximation $\bar{x}$ and $\bar{u}$ to

x^* and u^* shrinks by a factor $0 < \gamma < 1$;

3) the new vector Lagrange multipliers $\bar{u}$ belongs to $U_{y,k}$ again.

So, the inequality (9.10) holds for any $s \geq 1$, i.e. the fixed couple (y, k) is consistent with a fixed

given ratio $0 < \gamma < 1$.

We would like to emphasize that the convergence of the method (9.11) - (9.12) is not due to

the "center" or barrier parameter update, but rather due to the Lagrange multipliers update.

53

Therefore when it comes to numerical realization of the MCM (5.6) - (5.7) the main problem is to find for a chosen fixed "center" $y \in \Omega_\tau$ and a given ratio $0 < \gamma < 1$ the threshold $k_{y,\tau} > k_0$ that for any $k \geq k_{y,\tau}$ the couple (y, k) will be consistent with $0 < \gamma < 1$.

In the procedure describe below the "merit" function $v(w, y, k)$ is the key element, which we are going to use to adjust the barrier parameter to the level, which will make (y, k) consistent with $0 < \gamma < 1$ while the "center" $y \in \Omega_\tau$ and the ratio $0 < \gamma < 1$ are fixed.

To describe the MCM we introduce the Relaxation Operator $R : \Omega_k(y) \to \Omega_k(y)$ by formula

$$R x = \bar{x}$$

where $\bar{x}$ is defined by (9.7).

We start with a "center" $y \in \Omega_\tau$, we set $\bar{x}^0 = y$ and $\bar{u}^0 = e_m$. We choose the ratio $0 < \gamma < 1$ and a monotone increasing sequence $\{ k_s \} : k_0 > m$, $k_s < k_{s+1}$, $\lim k_s = \infty$. Let $\varepsilon > 0$ as the required accuracy.

Let's assume that the pair $(\bar{x}^s, \bar{u}^s)$ and the barrier parameter $k = \bar{k}_s$ have been found already. The next approximation $(\bar{x}^{s+1}, \bar{u}^{s+1})$ and $k = \bar{k}_{s+1}$ we find using the following operations.

1) $x : = \bar{x}^s$, $u : = \bar{u}^s$;

2) $\bar{x} : = R x$, $\bar{u} = \Delta(y, \bar{x}) d^{-1}(\bar{x}, y, k) u$;

3)   if $v(\bar{w}, y, k) \leq \varepsilon$, then $x^* = \bar{x}$, $u^* = \bar{u}$ and stop;

if $v(\bar{w}, y, k) > \varepsilon$

4)   a) if $v(\bar{w}, y, k) \leq \gamma^{s+1}$ and $k > \Sigma \bar{u}_i$, then $\bar{k}_{s+1} = k$, $\bar{x}^{s+1} = \bar{x}$, $\bar{u}^{s+1} = \bar{u}$,

   $s + 1 := s$ and go to 1.

   b) if $v(\bar{w}, y, k) \leq \gamma^{s+1}$, but $k \leq \Sigma \bar{u}_i$, then $k := \bar{k}_{s+1} = \Sigma \bar{u}_i + 1$,

   $\bar{x}^{s+1} = \bar{t}_{s+1} x + (1 - \bar{t}_{s+1}) y$, $\bar{t}_{s+1} = \max\{1 \geq t > 0 : y + t(\bar{x} - y) \in \Omega_k(y)\}$,

54

$\overline{u}^{s+1} = \overline{u}$, $s+1 := s$ and go to 1.

c) if $v(\overline{w},y,k) > \gamma^{s+1}$, $k = \max\{k_{s+1}, \Sigma\overline{u}_t + 1\}$,

$x(s) = \operatorname{argmin}\{f_0(x^i) / 1 \leq i \leq s\}$, $\overline{t}_{s+1} = \max\{1 \geq t > 0:$

$y + t(x(s) - y) \in \Omega_k(y)\}$, $x := \overline{t}_{s+1}x(s) + (1 - \overline{t}_{s+1})y$, $u = e_m$ and go to 2.

It comes to the point when y and k became consistent with $0 < \gamma < 1$, then every application of the relaxation operator leads to $\overline{x} \in \Omega_k(y)$ that $\overline{u} = \Delta(y,\overline{x})d^{-1}(\overline{x},y,u) \in U_{y,k}$.

From this point on the "center" $y \in \Omega_t$ and the barrier parameter $k = \overline{k}_t$ is fixed, every Lagrange multipliers update shrinks the distance from $\overline{x}$ and $\overline{u}$ to x^* and u^* by a factor $0 < \gamma < 1$ and

$u \in U_{y,k} \to \overline{u} \in U_{y,k}$.

The numerical realization of the operator R is based on smooth unconstrained minimization technique.

We will describe it based on Newton Method with a step length for finding an approximation $\overline{x}$ for the $\hat{x} = \hat{x}(y,u,k) = \operatorname{argmin}\{f(x,y,u,k) / x \in \mathbf{R}^n\}$.

The Newton direction ζ we find from the system

$$F''_{xx}(x,y,u,k)\zeta = -F'_x(x,y,u,k) \tag{9.13}$$

The step length t we find using Armijo rule. We check the following inequality for $t = 1$.

$$F(x + t\zeta,y,u,k) - F(x,y,u,k) \leq \tfrac{1}{3}t(F'_x(x,y,u,k),\zeta) \tag{9.14}$$

If (9.14) is satisfied we set

$$x := x + t\zeta \tag{9.15}$$

if not, we set $t := \tfrac{1}{2}t$ and check (9.14) again up to the point when (9.14) is true, then update x by formula (9.15) and go to the first phase of the Newton Method - find the Newton's direction ζ.

We continue the process (9.13), (9.15) up to the point $\bar{x}$, which satisfies (9.7).

Note that if x is within the Newton area of $\hat{x}(y, u, k)$, then it takes only $0 (\ln \ln \varepsilon^{-1})$ Newton steps to get $\bar{x}$.

In view that from some point on both the "center" $y \in \Omega$, and the barrier parameters $k > 0$ are fixed and $u \in U_{y,k} \to \bar{u} \in U_{y,k}$ the primal approximation will remain in the Newton area for the system $F'_x(x, y, u, k) = 0$ in x after every Lagrange multipliers update.

Such a point we call "hot" start. From this point on at most $0 (\ln \ln \varepsilon^{-1})$ Newton steps requires for every Lagrange multipliers update. Each update shrinks the distance from the current primal - dual pair $(\bar{x}, \bar{u})$ to the primal - dual solution (x^*, u^*) by a factor $0 < \gamma < 1$. So from this point on it takes $0 (\ln \varepsilon^{-1}) \cdot 0 (\ln \ln \varepsilon^{-1})$ to get the approximation to (x^*, u^*) with accuracy $\varepsilon > 0$.

To reach the "hot" start sometimes one has to increase significantly the barrier parameter.

For a very large barrier parameter, the MIDF became close to IDF, so in such case we might lose the the potential advantages of MIDF.

Also, as it was shown in Section 5, MCM is close to prox method with entropy-like kernel. It is well known that prox methods converge with linear rate and again to decrease the ratio one has to increase the parameter.

On the other hand being a Classical Lagrangean for the equivalent problem the MIDF possesses properties which make them fundamentally different from the Classical Lagrangean $L(x, u)$ for the initial problem.

Therefore we can expect that the dual function and the dual problem, which is based on MIDF, might have some interesting properties, which will enable us to improve the convergence without increasing the barrier parameter.

56

In the next section we are going to consider some aspect of the duality theory, which is based on MIDF.

10. Dual Problems Based on MIDF.

The MIDF, as we mentioned above, is a Classical Lagrangian for a problem, which is equivalent to (1.1), but not every result of the duality theory can be translated automatically when instead of the Classical Lagrangians for the initial problem, one considers the MIDF. The situation at this point is different from the situation with the Modified Barrier Functions (see [Pol92]), which are also Classical Lagrangians for a problem which is equivalent to the initial problem (1.1).

On the other hand, using MIDF, it is possible to obtain some new important characteristics for the dual functions and dual problems, which are impossible to obtain by using the Classical Lagrangians for the initial problem.

We are going to start with the basic optimality criteria for the convex programming problem. Theorem 3. Let for the convex programming problem (1.1) the Slater condition (1.2) is satisfied , then

1)   if for $y \in \text{int}\Omega$ and $k > 0$ there exists a vector $u^* = (u^*_1 ,..., u^*_m) \geq 0^n$ such that

$$u^*_i f_i(x^*) = 0 , \ i = 1,...,m \ \text{ and } \ F(x,y,u^*,k) \geq F(x^*,y,u^*,k) , \ \forall x \in \mathbb{R}^n : \Delta (y,x) > 0 \ (10.1)$$

then the vector x^* is a solution of problem (1.1).

2)   if $f_i(x) \in C^1$, $i = 0, 1,..., m$, and if x^* is a solution of problem (1.1), then there exists $u^* \geq 0$ such that (10.1) holds true for any $y \in \text{int } \Omega$ and $k > \Sigma u_i^*$,

3)   if (x^*, u^*) is a saddle point of the MIDF $F (x, y, u, k)$:

$$F(x,y,u^*,k) \geq F(x^*,y,u^*,k) \geq F(x^*,y,u,k) , \ \forall x \in \mathbb{R}^n : \Delta (y,x) > 0 , u \in \mathbb{R}^n_+ \quad (10.2)$$

57

then (x^*, u^*) is the K-K-T's pair for any $y \in$ int Ω and $k > 0$.

4)   if (x^*, y^*) is a K-K-T's pair then the pair (x^*, u^*) is a saddle point of the MIDF $F(x, y, u, k)$

for any $y \in$ int Ω and $k > \Sigma u_i^*$.

<u>Proof</u>:

1)   Let (10.1) holds true, then

$$-\ln \Delta(y,x) - k^{-1} \Sigma u_i^* [\ln(kf_i(x) + \Delta(y,x)) - \ln \Delta(y,x)]$$

$$\geq -\ln \Delta(y,x^*) - k^{-1} \Sigma u_i^* \ln[k\Delta^{-1}(y,x^*)f_i(x^*) + 1)]$$

$$= -\ln \Delta(y,x^*),$$

i.e.

$$-\ln \Delta(y,x) \geq -\ln \Delta(y,x^*) + k^{-1} \sum_{i=1}^{m} u_i^* \cdot \ln[k\Delta^{-1}(y,x)f_i(x) + 1]$$

In view of $\Delta(y, x) > 0$ the last term at the right hand side is non-negative for $\forall x \in \Omega$, therefore

$$\ln \Delta(y,x) \leq \ln \Delta(y,x^*) \rightarrow f_0(x) \geq f_0(x^*), \ \forall x \in \Omega : \Delta(y,x) > 0$$

and x^* is the solution of problem (1.1).

2)   if x^* is the solution of problem (1.1), then there exists $u^* \geq 0^m$ such that $u_i^* f_i(x^*) = 0$, $r = 1,...,m$

and $L'_x(x^*, u^*) = 0^n$, therefore due to the Proposition 2 we have

$$F'_x(x^*,y,u^*,k) = \Delta^{-1}(y,x^*)L'_x(x^*,u^*) = 0$$

The MIDF $F(x,y,u^*,k) = (-1 + k^{-1} \Sigma u_i^*)\ln \Delta(y,x) - k^{-1}\Sigma u_i^* \ln(kf_i(x) + \Delta(y,x))$ is convex

function in $x \in \mathbf{R}^n$: $\Delta(y, x) > 0$ for any $y \in$ int Ω and $k > \Sigma u_i^*$. Therefore $F(x, y, u^*, k) \geq$

$F(x^*, y, u^*, k)$ for any $x \in \mathbf{R}^n$: $\Delta(y, x) > 0$.

58

3) if (x^*, u^*) is a saddle point of the $F(x, y, u, k)$, i.e.

$$-\ln\Delta(y,x) - k^{-1}\sum_{i=1}^{m} u_i^* [\ln(kf_i(x) + \Delta(y,x)) - \ln\Delta(y,x)]$$

$$\geq -\ln\Delta(y,x^*) - k^{-1}\sum_{i=1}^{m} u_i^* [\ln(kf_i(x^*) + \Delta(y,x^*)) - \ln\Delta(y,x^*)]$$

$$\geq -\ln\Delta(y,x^*) - k^{-1}\sum_{i=1}^{m} u_i [\ln(kf_i(x^*) + \Delta(y,x^*)) - \ln\Delta(y,x^*)]$$

then $x^* \in \Omega$. In fact, suppose the contrary, i.e. there exists $i_0 : f_{i_0}(x^*) < 0$, therefore

$\ln(kf_{i_0}(x^*) + \Delta(y,x^*)) < 0$ for $y \in \text{int } \Omega$ and $k > 0$. So by fixing $u_i = 0$, $i \neq i_0$ and increasing

u_{i_0} we obtain a contradiction with the right side of the inequality (10.2), so $f_i(x^*) \geq 0$ $i = 1,...,m$.

Now we will prove that if (10.2) holds true, then $u_i f_i(x^*) = 0$, $i = 1,..., m$. In fact for $u = 0$

from the right side of (10.2), we obtain $\sum_{i=1}^{m} u_i^* \ln(k\Delta^{-1}(y,x^*)f_i(x^*) + 1) \leq 0$ but we just proved

that $f_i(x^*) \geq 0$, $i = 1,...,$ m, therefore $\ln(k\Delta^{-1}(y,x^*)f_i(x^*) + 1) \geq 0$, $i = 1,...,m$ and

$\sum_{i=1}^{m} u_i^* \ln(k\Delta^{-1}(y,x^*)f_i(x^*) + 1) \geq 0$, so for any $y \in \text{int } \Omega$ and $k > 0$ we have

$u_i^* \ln(k\Delta^{-1}(y,x^*)f_i(x^*) + 1) = 0$, $i = 1,...,m$. Hence $u_i f_i(x^*) = 0$, $i = 1,..., m$ and from the left side

of the inequality (10.2) we obtain

$$-\ln\Delta(y,x) \geq -\ln\Delta(y,x^*) + k^{-1}\sum u_i^* \ln(k\Delta^{-1}(y,x)f_i(x) + 1).$$

In view that for any $x \in \Omega$: $\Delta(y, x) > 0$ we have $\ln(k\Delta^{-1}(y,x)f_i(x) + 1) \geq 0$ we obtain that

$-\ln\Delta(y,x) \geq -\ln\Delta(y,x^*)$ or

$$\ln(f_0(y) - f_0(x)) \leq \ln(f_0(y) - f_0(x^*)) \rightarrow f_0(x) \geq f_0(x^*), \quad \forall x \in \Omega : \Delta(y,x) > 0$$

i.e. x^* is the solution of problem (1.1).

4) First of all, note that if (x^*, u^*) is the K-K-T's pair then x^* is the solution of problem (1.1). Using

the same consideration as in 2) we obtain $F'_x(x^*, y, u^*, k) = 0$ and $F(x^*, y, u^*, k) = \ln\Delta(y, x^*)$

for any $y \in \text{int } \Omega$. Taking into account the convexity $F(x, y, u^*, k)$ in x for any $k > \sum u_i^*$ we obtain

$F(x, y, u^*, k) \geq F(x^*, y, u^*, k)$, $\forall x \in \Omega_k(y)$, and due to the definition of the function

59

$F(x, y, u, k)$ we obtain

$$F(x, y, u^*, k) \geq F(x^*, y, u^*, k), \quad \forall \; x \in \mathbf{R}^n : \Delta(y, x) > 0$$

Also $k^{-1} \sum u_i \ln(k\Delta^{-1}(y, x^*) f_i(x^*) + 1) \geq 0$, $i = 1,...,m$. Therefore

$$F(x^*, y, u^*, k) = -\ln(f_0(y) - f_0(x^*))$$

$$\geq -\ln(f_0(y) - f_0(x^*)) - k^{-1}\sum u_i [\ln(kf_i(x^*) + \Delta(y, x^*) - \ln\Delta(y, x^*)]$$

$$= F(x^*, y, u, k), \quad \forall \; u \in \mathbf{R}^n_+ .$$

So (x^*, u^*) is a saddle point of the MIDF $F(x, y, u, k)$ for any $y \in \text{int } \Omega$ and $k > \sum u_i^*$.

We have completed the proof of Theorem 3.

Remark 4 . We want to emphasize again the difference between MIDF $F(x, y, u, k)$ and MBF $F(x, u, k)$. The statements 2) and 4) are not true (see Section 7, example 1) without the condition $k > \sum u_i^*$ while for the MBF (see Theorem 4 [Pol 92]) the statements 1) - 4) are true for any $k > 0$.

Now we are going to consider the dual pair constrained optimization problems, which are based on MIDF. Let $y \in \text{int } \Omega$ and $k > \sum u_i$ be fixed and $\psi_{y,k}(x) = \sup_{u \geq 0} F(x, y, u, k)$. Then

$$\psi_{y,k}(x) = \begin{cases} -\ln(f_0(y) - f_0(x)), & \text{if } f_0(y) - f_0(x) > 0 , f_i(x) \geq 0 , i = 1,...,m \\ \infty , & \text{otherwise} \end{cases}$$

and the initial problem (1.1) reduces to finding

$$x^* = \text{argmin} \{ \psi_{y,k}(x) \mid x \in \mathbf{R}^n \} \tag{10.3}$$

We define the dual function by the formula $\varphi_{y,k}(u) = \inf_{x \in \mathbf{R}^n} F(x, y, u, k)$. Then the dual to (1.1)

problem consists of finding

$$u^* = \text{argmax}\{\,\varphi_{y,k}(u)\mid u\geq 0\,\} \qquad (10.4)$$

Due to the definition of $\psi_{y,k}(x)$ and $\varphi_{y,k}(u)$ for any $y\in\text{int }\Omega$ and $k>0$ we have

$$-\ln(f_0(y)-f_0(x)) = \psi_{y,k}(x) \geq \varphi_{y,k}(u), \quad \forall x\in\Omega:\Delta(y,x)>0,\; u\in\mathbf{R}^n_+$$

Therefore if $\bar{x}$ and $\bar{u}$ are feasible solutions of the primal and dual problems respectively and

$\psi_{y,k}(\bar{x}) = \psi_{y,k}(\bar{u})$ then $\bar{x} = x^*$ and $\bar{u} = u^*$.

The smoothness of the dual function $\varphi_{y,k}(u)$ depends on the smoothness of $f_i(x)$, $i=0$, $1,\dots m$ and convexity properties of the MIDF $F(x,y,u,k)$ in $x\in\mathbf{R}^n$.

The next theorem describes the smoothness properties of the dual function $\varphi_{y,k}(u)$.

__Theorem 4__. If (1.1) is a convex programming problem, $f_i(x)\in C^2$, $i=0,1,\dots,m$ and conditions (1.3) - (1.5) are satisfied, then there exists k_0 such that for any fixed $y\in\text{int }\Omega$ and any fixed $k\geq$

$k_0\Delta(y,x^*)+\Sigma\, u_i^*$:

1) the concave function $\varphi_{y,k}(u)$ is twice continuously differentiable in $U_{y,k}$.

2) the gradient of the dual function is defined by formula

$$\varphi'_{(y,k)_u}(u) = F'_u(\hat{x}(y,u,k),y,u,k) = F'_u(\hat{x}(\bullet),\bullet)$$

$$= -k^{-1}(\ln(kf_1'(\hat{x}(\bullet))\Delta^{-1}(y,\hat{x}(\bullet))+1),\dots,\ln(kf_m'(\hat{x}(\bullet)\Delta^{-1}(y,\hat{x}(\bullet))+1)) \qquad (10.5)$$

3) the Hessian of $\varphi_{y,k}(u)$ is defined by formula

$$\varphi''_{(y,k)_{uu}}(u) = -F''_{ux}(\hat{x}(\bullet),\bullet)(F''_{xx}(\hat{x}(\bullet),\bullet))^{-1}F''_{xu}(\hat{x}(\bullet),\bullet) \qquad (10.6)$$

where $F''_{ux}(\hat{x}(\bullet),\bullet) = F''^T_{xu}(\hat{x}(\bullet),\bullet)$.

__Proof__: First of all, note that $\varphi_{y,k}(u)$: $\mathbf{R}^m_+ \to R$ is a concave function whether or not the functions $f_i(x)$, $i=1,\dots,m$ are convex. If the standard second order optimality conditions (1.3) - (1.5) are

satisfied then due to Theorem 2 the function $F(x, y, u, k)$ is strongly convex in the neighborhood

of $\hat{x} = \hat{x}(y,u,k)$ for $\forall (y,u,k) \in D(\bullet)$.

Therefore $\hat{x}(y,u,k) = \hat{x}(\bullet)$ is a unique minimum of $F(x, y, u, k)$ in x while $\varphi_{y,k}(u) =$

$F(\hat{x}(y,u,k);y,u,k)$ is smooth in $u \in U_{y,k}$ i.e., there exists the gradient $\varphi'_{(y,k)u}(u) =$

$F'_x(\hat{x}(\bullet),\bullet) \cdot \hat{x}'_u(\bullet) + F'_u(\hat{x}(\bullet),\bullet)$. Taking into account $F'_x(\hat{x}(\bullet),\bullet) = 0$ we obtain

$$\varphi'_{(y,k)u}(u) = F'_u(\hat{x}(\bullet),\bullet) = \left(\frac{\partial \varphi_{(y,k)}(u)}{\partial u_1} ,..., \frac{\partial \varphi_{(y,k)}(u)}{\partial u_m} \right)$$

$$= -k^{-1}(\ln(kf_1(\hat{x}(\bullet))\Delta^{-1}(y,\hat{x}(\bullet)) + 1) ,..., \ln(kf_m((\hat{x}(\bullet))\Delta^{-1}(y,\hat{x}(\bullet)) + 1)).$$

Further $\hat{x}(y,u^*,k) = x^*$ and $f_i(\hat{x}(y,u^*,k)) = f_i(x^*) = 0, i = 1,...,r$, therefore

$$\varphi'_{(y,k)u}(u^*) = -k^{-1}(\ln(kf_1(x^*)\Delta^{-1}(y,x^*) + 1) ,..., \ln(kf_r(x^*)\Delta^{-1}(y,x^*) + 1),$$

$$\ln(kf_{r+1}(x^*)\Delta^{-1}(y,x^*) + 1) ,..., \ln(kf_m(x^*)\Delta^{-1}(y,x^*) + 1)$$

$$= (0, \overset{-r-}{...} ,0, -k^{-1}\ln(k\Delta^{-1}(y,x^*)f_{r+1}(x^*) + 1) ,..., -k^{-1}\ln(k\Delta^{-1}(y,x^*)f_m(x^*) + 1))$$

Since the matrix $F''_{xx}(\hat{x}(\bullet),\bullet)$ is a positive definite for any fixed $y \in \text{int } \Omega$ and $k \geq k_0\Delta(y,x^*) +$

Σu_i^* the system $F'_x(x, y, u, k) = 0$ yields a unique vector function $x(y, u, k)$ such that

$$F'_x(\hat{x}(y,u,k),y,u,k) = F'_x(x(\bullet),\bullet) = 0^n, \quad \forall u \in U_{y,k} \text{ and } \hat{x}(y,u^*,k) = x^*.$$

Differentiating the last identity by u we obtain

$$F''_{xx}(\hat{x}(y,u,k);y,u,k)\hat{x}'_u(y,u,k) + F''_{xu}(\hat{x}(y,u,k);y,u,k) = 0^{n,m}$$

therefore

$$\hat{x}'_u(y,u,k) = \hat{x}'_u(\bullet) = -(F''_{xx}(\hat{x}(\bullet),\bullet))^{-1}F''_{xu}(\hat{x}(\bullet),\bullet), \quad \forall u \in U_{y,k}.$$

Hence

$$\varphi''_{(y,k)uu}(u) = F''_{ux}(\hat{x}(\bullet),\bullet)\hat{x}'_u(\bullet) =$$

$$-F''_{ux}(\hat{x}(\bullet),\bullet)F''_{xx}(\hat{x}(\bullet),\bullet)^{-1}F''_{xu}(\hat{x}(\bullet),\bullet) \tag{10.7}$$

$$F''_{ux}(\hat{x}(\bullet),\bullet) = F''^{T}_{xu}(\hat{x}(\bullet),\bullet) =$$

$$-d^{-1}(\hat{x}(\bullet),y,k)f'(\hat{x}(\bullet)) - \Delta^{-1}(y,\hat{x}(\bullet))D((\hat{x}(\bullet))d^{-1}(\hat{x}(\bullet),y,k)F_m(\hat{x}(\bullet)).$$

To compute $\varphi''_{(y,k)uu}(u^*)$ we first consider

$$F''_{ux}(\hat{x}(y,u^*,k),y,u^*,k) = F''_{ux}(x^*,y,u^*,k) =$$

$$-d^{-1}(x^*,y,k)f'(x^*) - \Delta^{-1}(y,x^*)D(x^*)d^{-1}(x^*,y,k)F_m(x^*) =$$

$$-\begin{pmatrix} \Delta^{-1}(y,x^*)I^r, & 0^{r,m-r} \\ 0^{m-r,r}, & d^{-1}_{m-r}(x^*,y,k) \end{pmatrix} \bullet$$

$$\left[f'(x^*) - \Delta^{-1}(y,x^*)\begin{pmatrix} 0^{r,r}, & 0^{r,m-r} \\ 0^{m-r,r}, & [\operatorname{diag} f_i(x^*)]^m_{m-r+1} \end{pmatrix} F_m(x^*) \right] \tag{10.8}$$

then

$$\varphi''_{(y,k)uu}(u^*) = -F''_{ux}(x^*,y,u^*,k)F''_{xx}(x^*,y,u^*,k))^{-1}F''_{xu}(x^*,y,u^*,k) = -F''_{ux}(F''_{xx})^{-1}F''_{xu}$$

<u>Remark 5</u>. The dual function $\varphi(u) = \inf_{x \in \mathbb{R}^n} L(x,u)$, which is based on the Classical Lagrangian $L(x,u)$ for the initial problem is in general nonsmooth under the conditions of Theorem 4. It becomes

63

smooth, if, for example, $f_0(x)$ is strongly convex, while the dual function $\varphi_{(y,k)}(u)$, which is based on the MIDF $F(x,y,u,k)$ is smooth even for the nonconvex programming problem, when along with (1.3) - (1.5) the growth conditions (see [Pol 92] p. 181) are satisfied and $k \geq k_0 \Delta(y,x^*) + \Sigma u_i^*$.

The next theorem establishes the main result of the duality theory which is based on the MIDF.

Theorem 5. Let $f_0(x)$ and all $-f_i(x)$ be convex, then

1) if $f_i(x) \in C^1$, $i = 0,..., m$ and the Slater condition holds, then the existence of the solution x^* of the primal problem (1.1) implies the existence of the solution u^* of the dual problem (10.4) and $\psi_{y,k}(x^*) = \varphi_{y,k}(u^*)$ for any $y \in \text{int } \Omega$ and $k > \Sigma u_i^*$.

2) if $f_i(x) \in C^2$, $i = 0,..., m$ and the optimality conditions (1.3) - (1.5) are taking place, then the existence of the solution $u^* \in U_{y,k}$ for the dual problem implies the existence of the solution x^* for the primal problem and $\psi_{y,k}(x^*) = \varphi_{y,k}(u^*)$ for any $y \in \text{int } \Omega$ and $k \geq k_0 \Delta(y,x^*) + \Sigma u_i^*$.

3) if $f_i(x) \in C^2$, $i = 0,..., m$ and the second order optimality conditions (1.4) - (1.5) are satisfied for the primal problem, then the corresponding conditions are taking place for the dual problem (10.4) for any $y \in \text{int } \Omega$ and $k \geq k_0 \Delta(y,x^*) + \Sigma u_i^*$.

Proof:

1) Let x^* be a solution of problem (1.1), then there exists (see Theorem 3) a vector u^* such that (10.1) holds true for any $y \in \text{int } \Omega$ and $k > \Sigma u_i^*$. Hence for any $u \geq 0$ we obtain

$$\varphi_{y,k}(u^*) = \min_{x \in \mathbb{R}^n} F(x,y,u^*,k) = F(x^*,y,u^*,k) =$$

$$\psi_{y,k}(x^*) \geq F(x^*,y,u,k) \geq \min_{x \in \mathbb{R}^n} F(x,y,u,k) = \varphi_{y,k}(u), \quad \forall u \in \mathbb{R}_+^m$$

so u^* is a solution of the dual problem and $\psi_{y,k}(x^*) = \varphi_{y,k}(u^*)$ i.e.

64

$$f_0(x^*) = f_0(y) - \exp(-\varphi_{y,k}(u^*))\qquad\qquad(10.9)$$

for any $y \in$ int Ω and $k > \Sigma u_i^*$.

2) Conditions (1.3) - (1.5) imply that the function $F(x, y, u, k)$ is strongly convex (see Theorem 2) in the neighborhood of $\hat{x}(y,u,k) = \operatorname{argmin}\{F(x,y,u,k) \mid x \in \mathbb{R}^n\}$, therefore the vector $\hat{x}(y,u,k)$ is unique for any $(y, u, k) \in D(\bullet)$ and because of the smoothness of $f_i(x)$, $i = 0,..., m$ the gradient $\varphi'_{(y,k)u}(u)$ exists.

Let $\bar{u} \in U_{y,k}$ be a solution of the dual problem (10.4) and $\bar{x} = x(y,\bar{u},k)$. Then the optimality conditions for the dual problem (10.4) are fulfilled at $\bar{u}$, i.e.

$$\varphi'_{(y,k)u_i}(\bar{u}) = -k^{-1}(\ln(kf_i(\bar{x}) + \Delta(y,\bar{x}) - \ln\Delta(y,\bar{x}) \le 0, \text{ for } i : \bar{u}_i = 0$$

$$\varphi'_{(y,k)u_i}(\bar{u}) = -k^{-1}(\ln(kf_i(\bar{x}) + \Delta(y,\bar{x}) - \ln\Delta(y,\bar{x}) = 0, \text{ for } i : \bar{u}_i > 0$$

Then $\bar{u}_i = 0$ implies $\ln(kf_i(\bar{x}) + \Delta(y,\bar{x}) - \ln\Delta(y,\bar{x}) \ge 0 \rightarrow f_i(\bar{x}) \ge 0$ while $\bar{u}_i > 0$ implies $\ln(kf_i(\bar{x}) + \Delta(y,\bar{x})) = \ln\Delta(y,\bar{x}) \rightarrow f_i(\bar{x}) = 0$, i.e. $\bar{x} \in \Omega$ and for the pair $(\bar{x}, \bar{u})$ the complementarity conditions $f_i(\bar{x})\bar{u}_i = 0$, $i = 1,...,m$ hold. Therefore

$$\varphi_{(y,k)}(\bar{u}) = -\ln\Delta(y,\bar{x}) - k^{-1}\Sigma\bar{u}_i(\ln(kf_i(\bar{x}) + \Delta(y,\bar{x})) - \ln\Delta(y,\bar{x})) = -\ln\Delta(y,\bar{x}) = \psi_{y,k}(\bar{x})$$

i.e. for the primal and dual feasible pair $(\bar{x}, \bar{u})$ we have

$$\varphi_{y,k}(\bar{u}) = \psi_{y,k}(\bar{x}) = -\ln\Delta(y,\bar{x})$$

hence $\bar{x} = x^*$, $\bar{u} = u^*$.

3) We will now show that the second order optimality conditions hold for the dual problem (10.4) in the strict form .

First, we note that the gradients $e_i = (0,...,0,0,..., \overset{i}{1},..., 0)$, $i = r+1,...,m$ of the active constraints $u_i \ge 0$, $i = r + 1,..., m$ of the dual problem are linearly independent. Later we will prove that the Lagrange Multipliers, which correspond to the active constraints of the dual problem are

65

positive. Along with the linear independence of the e_i, $i = r + 1, ..., m$, this forms condition (1.4) for dual problem (10.4). Now we will prove that the condition type (1.5) is also satisfied for the dual problem.

Let us consider the Classical Lagrangian $L_{(y,k)}(u, \lambda) = \varphi_{(y,k)}(u) + \sum_{i=1}^{n} \lambda_i u_i$ for the dual problem (10.4). The Hessian of the Classical Lagrangian for the dual problem is $L''_{(y,k)uu}(u, \lambda) = \varphi''_{(y,k)uu}(u)$ and the second order optimality condition for the dual problem (10.4) is

$$(L''_{(y,k)uu}(u^*, \lambda^*)v, v) \le -\tilde{\mu}(v,v), \quad \forall v \in \mathbf{R}^m : (v, e_i) = 0, i = r+1, ..., m, \tilde{\mu} > 0 \quad (10.10)$$

To prove (10.10) we first consider the matrix $F''_{xx} = F''_{xx}(x^*, y, u^*, k)$. Taking into account conditions (1.4) - (1.5) for the primal problem, we obtain from Theorem 1 that

$$\text{mineigval } F_{xx} = \mu \Delta^{-1}(y, x^*) > 0 \ .$$

for any $y \in \text{int } \Omega$ and $k \ge k_0 \Delta(y, x^*) + \Sigma u_i^*$. Let

$$\text{maxeigval } \{ L''_{xx}(x^*, u^*) + k_0 f'^T_{(r)}(x^*) U^*_r f'_{(r)}(x^*) \} = M$$

Then maxeigval $F''_{xx} = M\Delta^{-1}(y, x^*)$, so for any $w \in \mathbf{R}^n$ we obtain

$$\mu^{-1}\Delta(y, x^*)(w, w) \ge (F''^{-1}_{xx} w, w) \ge M^{-1}\Delta(y, x^*)(w, w)$$

$$-\mu^{-1}\Delta(y, x^*)(w, w) \le (-F''^{-1}_{xx} w, w) \le -M^{-1}\Delta(y, x^*)(w, w)$$

Therefore due to (10.6)

$$(L''_{(y,k)uu}(u^*, \lambda^*)v, v) = (\varphi''_{(y,k)uu}(u^*)v, v) =$$

$$(F''_{ux}(-F''_{xx})^{-1}F''_{xu}v, v) = ((-F''_{xx})^{-1}F''_{xu}v, F''_{xu}v)$$

Let $v = (v_1, ..., v_m)$, then $(v, e_i) = 0 \rightarrow v_i = 0$, therefore any vector $v \in \mathbf{R}^m : (v, e_i) = 0, i = r + 1, ..., m$ has the form $v = (v_1, ..., v_r, 0, ..., 0) = (v_{(r)}, 0, ..., 0)$. Therefore, due to (10.8), $d^{-1}(x^*, y, k) v = v$ and $D(x^*) v = 0^m$ we have

$$F''_{xu}(x^*, y, u^*, k)v = -f'^T(x^*)d^{-1}(x^*, y, k)v - \Delta^{-1}(y, x^*)F''_u(x^*)d^{-1}(x^*, y, k)D(x^*)v$$

$$= -\Delta^{-1}(y, x^*) f'^T(x^*) v$$

So

$$(L''_{uu}(u^*, \lambda^*, y, k)v, v) = ((F''_{xx})^{-1} F''_{xu}v, F''_{xu}v)$$

$$\leq -\Delta^{-1}(y, x^*) M^{-1}(F''_{xu}v, F''_{xu}v) = -\Delta^{-1}(y, x^*) M^{-1}(f'^T_{(r)}(x^*)v_{(r)}, f'^T_{(r)}(x^*)v_{(r)})$$

$$= -\Delta^{-1}(y, x^*) M^{-1}(f'_{(r)}(x^*) f'^T_{(r)}(x^*)v_{(r)}, v_{(r)})$$

It follows from (1.4) that the Gram matrix $(f'_{(r)}(x^*) f'^T_{(r)}(x^*)$ is nonsingular. Therefore,

$$\text{mineigval } (f'_{(r)}(x^*) f'^T_{(r)}(x^*) = \mu_0 > 0$$

and

$$(f'_{(r)}(x^*) f'^T_{(r)}(x^*) v_{(r)}, v_{(r)}) \geq \mu_0 (v_{(r)}, v_{(r)}) = \mu_0 (v, v)$$

therefore, for $\tilde{\mu} = \Delta^{-1}(y, x^*) M^{-1} \mu_0 > 0$ we obtain

$$(L''_{(y,k)uu}(u^*, \lambda^*)v, v) \leq -\tilde{\mu}(v, v), \quad \forall v : (v_1 e_i) = 0, \ i = r+1, \dots, m$$

i.e. the condition type (1.5) for the dual problem (10.4) is satisfied.

Also in view of $f_i(x^*) > 0$, $i = r + 1, \dots, m$ we have

$$\lambda_i^* = -\varphi'_{(y,k)u_i}(u^*) = k^{-1} \ln(k\Delta^{-1}(y, x^*) \cdot f_i(x^*) + 1) > 0, i = r+1, \dots, m$$

which along with the linear independence of the gradients e_i, $i = r + 1, \dots, m$ of the active constraints

$u_i \geq 0$, $i = r + 1, \dots, m$, and (10.10) comprise the second order optimality conditions for the dual

problem. We completed the proof of Theorem 5.

Corrollary. The restriction $\overline{\varphi}_{(y,k)}(u) = \varphi_{(y,k)}(u) \mid u_{r+1} = \dots = u_m = 0$ of the dual function

67

$\varphi_{(y,k)}(u)$ to the manifold of the active constraints $u_i = 0$, $i = r + 1,..., m$ of the dual problem (10.4)

is strongly concave if the conditions of Theorem 5 are fulfilled.

<u>Remark 6</u> . Statements 2) and 3) of Theorem 5 are in general not true even for a convex programming problem if the dual function $\varphi(u) = \inf_{x \in \mathbb{R}^n} L(x,u)$ is based on the Classical Lagrangian

$L(x, u)$ for the initial problem. However, these results are valid for any $k \geq k_0 \Delta(y,x^*) + \Sigma u_i^*$ and

any $y \in \text{int}\Omega$ even for a nonconvex programming problem if the conditions (1.3) - (1.5) and growing

conditions (see [Pol 92]p.181) are fulfilled and the dual functions are based on the MIDF

$F(x, y, u, k)$.

<u>Remark 7</u> . All statements of Theorem 5 hold for the Modified Interior Distance Function $H(x, y, u, k)$

and the corresponding dual function

$$h_{y,k}(u) = \min \{ H(x,y,u,k) \mid x \in \mathbb{R}^n \}$$

with the only difference that instead of (10.9) we have

$$f_0(x^*) = f_0(y) - h_{y,k}^{-1}(u^*)$$

for any $y \in \text{int } \Omega$ and $k > \Sigma u_i^*$.

Let $y \in \text{int } \Omega$ and $k \geq k_0 \Delta(y,x^*) + \Sigma u_i^*$ also let $(y, u, k) \in D(\bullet)$. We consider the dual

problem (10.4) in the following form

$$u^* = \text{argmax} \{ F(x,y,u,k) \mid F_x'(x,y,u,k) = 0, u \geq 0 \} \tag{10.11}$$

(P. Wolfe's duality [Wol 61]), then the MCM method

$$x^{s+1} : F_x'(x^{s+1},y,u^s,k) = \Delta^{-1}(y,x^{s+1})L_x'(x^{s+1}, u^{s+1}) = 0 \tag{10.12}$$

$$u^{s+1} = \Delta(y,x^{s+1})d^{-1}(x^{s+1},y,k)u^s \tag{10.13}$$

is an Interior Point Method for the dual problem (10.11) and the estimation

$$\max \{ |x^{s+1} - x^*|, |u^{s+1} - u^*| \leq ck^{-1} \Delta(y,x^*) |u^s - u^*| \tag{10.14}$$

holds true under the conditions of Theorem 2 for any $(y, u^0, k) \in D(\bullet)$.

Finally note that the properties of the dual problem (10.4) (see Theorem 5) allows to improve significantly estimation (10.14) by using smooth optimization methods for solving the dual problem (10.4) in a way which is similar to (see [Pol 92], pp.206-208).

11 Concluding Remarks . The Lagrange multipliers, the specific role of the barrier parameter along with the extension of the feasible set give rise to properties P1 - P5, which makes MIDF $F(x, y, u, k)$ substantially different from IDF $F(x, \alpha) = F(x, f_0(y))$.

One can view both MIDF $F(x, y, u, k)$ and IDF $F(x, \alpha)$ as smooth approximations for a nonsmooth and convex in $x \in \Omega$: $\Delta(y, x) > 0$ function :

$$\lambda(x, \alpha) = \max\{-\ln(\alpha - f_0(x)), -\ln(\alpha - f_0(x) + f_i(x)) \ i = 1, ..., m\}.$$

Assuming again $\ln t = -\infty$ for $t \leq 0$ it is easy to see that solving the problem (1.1) is equivalent to solving the following unconstrained optimization problem

$$\lambda(x^*, \alpha) = \min\{\lambda(x, \alpha) / x \in \mathbb{R}^n\} = -\ln(\alpha - f_0(x^*))$$

Therefore the unconstrained minimizers

$$\hat{x}(y, u, k) = \operatorname{argmin}\{F(x, y, u, k) / x \in \mathbb{R}^n\}$$

or

$$x(\alpha) = \operatorname{argmin}\{F(x, \alpha) / x \in \mathbb{R}^n\}$$

one can view as an approximations for the solution x^* . The "quality" of approximations $\hat{x}(y, u, k)$ and $\hat{x}(\alpha)$ to x^* depends on the "quality" of the smooth approximations $F(x, y, u, k)$ and $F(x, \alpha)$ for the function $\lambda(x, \alpha)$. It is clear that these approximations are very different.

In case of MIDF $F(x, y, u, k)$ we have $\lim_{u \to u^*} \hat{x}(y, u, k) = x^*$ for any fixed $y \in \text{int } \Omega$ and

$k > \Sigma\, u_i$.

Moreover

$$\lim_{u \to u^*} [\, F(\hat{x}(y,u,k),y,u,k) - \lambda(x^*,\alpha)\,] = 0$$

In case of $F(x, \alpha)$ we have $\lim x(\alpha) = x^*$ only when $\alpha = f_0(y) \to \alpha^* = f_0(x^*)$. However

for any fixed "center" $y \in \mathrm{int}\,\Omega$ we have

$$\lim_{x \to x^*} F(x, \alpha) = \infty$$

So, while the MIDF $F(x, y, u^*, k)$ is an <u>exact</u> smooth approximation for $\lambda(x, \alpha)$ at the

solution x^* for any fixed "center" $y \in \mathrm{int}\,\Omega$ and any fixed $k > \Sigma\, u_i^*$, IDF $F(x, \alpha)$ does not exist at x^*

and

$$\lim_{x \to x^*} [F(x, \alpha) - \lambda(x^*,\alpha)] = +\infty\,.$$

for any fixed "center" $y \in \mathrm{int}\,\Omega$.

The "quality" of the approximation $F(x, y, u, k)$ has a substantial impact on the "quality"

of the MCM (5.6) - (5.7). Method (5.6) - (5.7) converges only due to the Lagrange multipliers

update. Therefore from some point on the function $F(x, y, u, k)$ as well as its gradient and Hessian

in x is not changing much after each Lagrange multipliers update, because $\lim_{s \to \infty} u^s = u^*$. Being

well conditioned at the primal - dual solution (x^*, u^*) the Hessian $F''_{xx}(x, y, u, k)$ remains

well conditioned in the neighborhood of (x^*, u^*), so from some point on if x^{s-1} is "well" defined in

terms of Newton method (see[Sm86]) for the system $F'_x(x, y, u^s, k) = 0$" the new approximation x^s

will be "well" defined for the system $F'_x(x, y, u^{s+1}, k) = 0$" and it will remain true up to the end of

the process. Such x^s we will call "hot" start (see[MeIP95]) .Therefore from the "hot" start on it takes

only $0 (\ln\ln \varepsilon^{-1})$ Newton steps to find a new primal approximation and update the Lagrange

multipliers.

Due to the estimation (6.4) every Lagrange multipliers update shrinks the distance from x' and u' to x^* and u^* by a factor $0 < \gamma < 1$. Therefore it takes only $0 (\ln \varepsilon^{-1}) \cdot 0 (\ln \ln \varepsilon^{-1})$ Newton's step from the "hot" start to the solution.

To reach the "hot" start one can use any of path-following methods (see [NesN94]).

The moment when the above mentioned acceleration begins depends on the triple (y, u, k). First of all the triple (y,u,k) has to be in $D (\cdot)$ and let 's assume that $\gamma_{y,k} = c\, k^{-1} \Delta(y, x^*) = \gamma = 0.5$, i. e.

$$k \geq \max \{ 2c\Delta (y,x^*), k_0 \Delta (y,x^*) + \sum_{i=1}^{m} u_i^* \} \qquad (11.1)$$

The constant $c > 0$ depends on the norm $| \Phi_{(y,k)}^{-1} R(x^*,y,0,k) |$ and $\sigma > 0$, the constant k_0 depends (see Assertion 1) on $\lambda > 0$, $\min \{ u_i^* \mid i = 1,..., r \}$, $\max \{ u_i^* \mid i = 1,..., r \}$ and mineigval $f'^T_{(r)}(x^*) f'_{(r)}(x^*)$. So the value of both $c > 0$ and $k_0 > 0$ depends on the "measure" of the nondegeneracy of the constrained optimization problem.

Recall that in unconstrained optimization the properties of the smooth optimization methods depend very much on the condition of the Hessian at the solution. One can consider this condition as the "measure" of nondegeneracy of an unconstrained optimization problem.

The analysis which was undertaken above, highlights the parameters which are responsible for the nondegeneracy of the constrained optimization problem and the way in which it influences the complexity of the NMCM. To complete the discussion we will make one more comment.

Due to (11.1) the "hot" start depends very much not only on the "measure" of nondegeneracy of the constrained optimization problem, but also on the value of Σu_i^*. The sum Σu_i^* is critical for the theory of MIDF (see Theorem 1-5). On the other hand, Σu_i^* depends on the condition number of the feasible set, i.e. on $\pi = r_0 R_0^{-1}$, where r_0 is the largest radius of the sphere inscribed in Ω and

71

R_0 is the smallest radius of the sphere circumscribed around Ω.

Recall the the value $\Sigma \, u_i^*$ is crucial in the theory of exact penalty methods (see [Ber82]) . However, in constrast to the exact penalty function, the MIDF is smooth and for some convex programming problems possesses the self concordant (see [NesN94] properties when $u = e_m$.

Therefore, the NMCM combines the self-concordant properties of the Shifted Interior Distance Function which guarantees the polynomial complexity of the Interior Points Methods, with important local P1 - P5 properties. It allows us to speed up the process in the final stage, to make the process numerically more stable, and has the potential for improving the complexity bounds at least for nondegenerate convex problems (see [MelP95]).

Due to (11.1) the moment when the process switches from the first phase to the second depends on the "measure" of nondegeneracy of the constrained optimization problem as well as on the condition number of the feasible set.

There are still a number of questions, which have to be answered.

First, we have to understand the behavior of the MCM in the absence of nondegeneracy assumptions.

Second, the convergence under the fixed $y \in$ int Ω and $k > k_0$ is in no way a suggestion that both the "center" and the barrier parameter have to be fixed. To find the optimal strategy for changing the "center", the barrier parameter or both is another important issue.

Third, to find conditions, which provide the convergence of the prox-type method (5.12).

Fourth, to specify the MIDF and MCM for LP and QP problems.

Finally, the complexity of the Newton MCM is one of the important and still unanswered questions.

72

Acknowledgement

I am very grateful to Professor Clovis Gonzaga for several thoughtful comments and suggestons.

I am thankful to anonymous referees for their comments.

I also would like to express my appreciation to Eileen Count and Angel Manzo for their excellent work in preparing the manuscript for publication.

Appendix

A1. <u>Proof of the Proposition 4</u>. We consider the Hessian $F''_{xx}(x^*, y, u^*, k)$ for a fixed $y \in \text{int } \Omega$ and $k > 0$.

$$F''_{xx}(x,y,u^*,k) = [(1 - k^{-1}\Sigma u_i^*)(f_0(y) - f_0(x))^{-1} f'_0(x) - k^{-1}\Sigma u^*_i(kf_i(x) + f_0(y)$$

$$- f_0(x))^{-1}(kf'_i(x) - f'_0(x))]'_{/x=x^*}$$

$$= [(1 - k^{-1}\Sigma u^*_i)(f_0(y) - f_0(x))^{-2}(f'_0(x))^T f'_0(x) + (1 - k^{-1}\Sigma u^*_i)(f_0(y) - f_0(x))^{-1} f''_0(x) +$$

$$k^{-1}\Sigma u^*_i(kf_i(x) + f_0(y) - f_0(x))^{-2}(kf'_i(x) - f_0'(x))^T(kf'_i(x) - f_0'(x)) - k^{-1}\Sigma u^*_i(kf_i(x) +$$

$$f_0(y) - f_0(x))^{-1}(kf''_i(x) - f''_0(x))]_{/x=x^*}$$

$$= [(1 - k^{-1}\Sigma u^*_i)(f_0(y) - f_0(x^*))^{-2}(f'_0(x^*))^T f'_0(x^*) + (1 - k^{-1}\Sigma u^*_i)(f_0(y) - f_0(x^*))^{-1} f''_0(x^*)$$

$$+ k^{-1}\Sigma u^*_i(f_0(y) - f_0(x^*))^{-2}(kf'_i(x^*) - f_0'(x^*))^T(kf'_i(x^*) - f'_0(x^*)) -$$

$$k^{-1}\Sigma u^*_i(f_0(y) - f_0(x^*))^{-1}(kf''_i(x^*) - f_0''(x^*))]$$

$$= (f_0(y) - f_0(x^*))^{-1}[(f_0(y) - f_0(x^*))^{-1}(f_0'(x^*))^T f_0'(x^*) - k^{-1}(\Sigma u^*_i)(f_0(y) - f_0(x^*))^{-1} \cdot$$

$$(f_0'(x^*))^T f_0'(x^*) + f_0''(x^*) - k^{-1}(\Sigma u^*_i)f_0''(x^*) + (f_0(y) - f_0(x^*))^{-1}(k\Sigma u^*_i(f_i'(x^*))^T f_i'(x^*)$$

$$- (f_0'(x^*))^T(\Sigma u^*_i f_i'(x^*)) - (\Sigma u^*_i f'_i(x^*))^T f_0'(x^*)) + k^{-1}(\Sigma u^*_i)(f_0'(x^*))^T f_0'(x^*))$$

$$- \Sigma u^*_i f_i''(x^*) + k^{-1}(\Sigma u^*_i)f_0''(x^*)]$$

Taking into account the K-K-Ts relation $f_0(x^*) = \Sigma u_i^* f_i(x^*)$ we obtain

$$F''_{xx}(x^*,y,u^*,k) = (f_0(y) - f_0(x^*))^{-1}[L''_{xx}(x^*,u^*)$$

$$+ k(f_0(y) - f_0(x^*))^{-1}(f'_{(r)}(x^*))^T U_r^* f'_{(r)}(x^*) - (f_0(y) - f_0(x^*))^{-1}(f_0'(x^*))^T f_0'(x^*)]$$

$$= (f_0(y) - f_0(x^*))^{-1}[L''_{xx}(x^*,u^*) +$$

$$(f_0(y) - f_0(x^*))^{-1}(kf'^T_r(x^*) U_r^* f'_{(r)}(x^*) - f'^T_{(r)}(x^*)u^{*T}_{(r)} u_r^* f'_{(r)}(x^*)]$$

Using the same considerations we obtain

$$H''_{xx}(x^*, y, u^*, k) = (f_0(y) - f_0(x^*))^{-2} [L''_{xx}(x^*, u^*))$$

$$+ 2(f_0(y) - f_0(x^*))^{-1}(k(f'_{(r)}(x^*))^T U^* f'_{(r)}(x^*) - (f'_{(r)}(x^*))^T u^{*T}_{(r)} u^*_{(r)} f'_{(r)}(x^*))].$$

A2. Formula for the MIDF Hessian $F''_{xx}(\hat{x}, y, u, k)$

$$F''_{xx}(x, y, u, k)_{/x=\hat{x}} = (1 - k^{-1}\Sigma u_i)(f_0(y) - f_0(\hat{x}))^{-2} f'^T_0(\hat{x}) f'_0(\hat{x}) + (1 - k^{-1}\Sigma u_i)$$

$$\bullet (f_0(y) - f_0(\hat{x}))^{-1} f''_0(\hat{x}) - k^{-1}\Sigma u_i (k f_i(\hat{x}) + f_0(y) - f_0(\hat{x}))^{-1}(k f''_i(\hat{x}) - f''_0(\hat{x})) +$$

$$k^{-1}(\Sigma u_i (k f_i(\hat{x}) + f_0(y) - f_0(\hat{x}))^{-2}(k f'_i(\hat{x}) - f'_0(\hat{x}))^T (k f'_i(\hat{x}) - f'_0(\hat{x}))$$

$$= (f_0(y) - f_0(\hat{x}))^{-1}[(f_0(y) - f_0(\hat{x}))^{-1}(f'^T_0(\hat{x}) f'_0(\hat{x}) - k^{-1}(\Sigma u_i) f'^T_0(\hat{x}) f'_0(\hat{x}))$$

$$+ f''_0(\hat{x}) - k^{-1}(\Sigma u_i) f''_0(\hat{x}) - \Sigma u_i (f_0(y) - f_0(\hat{x}))(k f_i(\hat{x}) + f_0(y) - f_0(\hat{x}))^{-1} f''_i(\hat{x}))$$

$$+ k^{-1}\Sigma u_i (f_0(y) - f_0(\hat{x}))(k f_i(\hat{x}) + f_0(y) - f_0(\hat{x}))^{-1} f''_0(\hat{x}) + k^{-1}\Sigma u_i (f_0(y) - f_0(\hat{x}))$$

$$\bullet (k f_i(\hat{x}) + f_0(y) - f_0(\hat{x}))^{-1}(k f_i(\hat{x}) + f_0(y) - f_0(\hat{x}))^{-1}(k^2 f'^T_i(\hat{x}) f'_i(\hat{x})$$

$$- k f'^T_0(\hat{x}) f'_i(\hat{x}) - k f'^T_i(\hat{x}) f'_0(\hat{x}) + f'^T_0(\hat{x}) f'_0(\hat{x}))]$$

$$= (f_0(y) - f_0(\hat{x}))^{-1}[(f_0(y) - f_0(\hat{x}))^{-1}(f'^T_0(\hat{x}) f'_0(\hat{x}) - k^{-1}(\Sigma u_i) f'^T_0(\hat{x}) f'_0(\hat{x}) +$$

$$f''_0(\hat{x}) - k^{-1}(\Sigma u_i) f''_0(\hat{x}) - \Sigma \hat{u}_i f''_i(\hat{x}) + k^{-1}(\Sigma \hat{u}_i) f''_0(\hat{x}) +$$

$$k^{-1}\Sigma \hat{u}_i (k f_i(\hat{x}) + f_0(y) - f_0(\hat{x}))^{-1}(k^2 f'^T_i(\hat{x}) f'_i(\hat{x}) - k f'^T_0(\hat{x}) f'_i(\hat{x}) - k f'^T_i(\hat{x}) f'_0(\hat{x}) + f'^T_0(\hat{x}) f'_0(\hat{x}))].$$

References

[Ad RVK89] Adler, I., Resende, M., Veiga, G., Karmarkar, N., An implementation of Karmarkar's algorithm for linear programming, *Mathematical Programming*, Vol. 44, (1989) pp.297-335.

[Bar 86] Barnes, E.R., A variation on Karmarkar's algorithm for solving linear programming problems, *Mathematical Programming*, Vol. 36, (1986) pp.174-182.

[Ber 82] Bertsekas, D., Constrained Optimization and Lagrange Multiplier Methods , Academic Press, New York (1982).

[Bui H 66] Bui-Trong Lieu, Huard, P., La methods des centres dans un espace topologique, *Num. Mat.* , Vol. 8, No. 1, (1966) pp. 56-67.

[Dik 67] Dikin, I.I., Iterative solution of problems of linear and quadratic programming, *Soviet Mathematics Doklady*, Vol. 8, (1967), pp. 674-675.

[Dik 74] Dikin, I.I., On the speed of an iterative process, *Upravlyaemye Sistemy* , 12, (1974) pp. 54-64.

[Fiac M 68] Fiacco, A.V., McCormick, G.P., Nonlinear Programming : Sequential Unconstrained Minimization Techniques, Wiley, New York (1968).

[Gol T 89] Gol'shtein, E.G., Tret'yakov, N.V., Modified Lagrangian functions. Theory and optimization methods, Nauka, Moscow (1989) (in Russian).

[Gon 88] Gonzaga, C., An algorithm for solving linear programming problems in $0\ (n^3\ L\ )$ operations, N. Megiddo, editor, Progress in Mathematical Programming - (Springer Verlag, Berlin, 1988) pp. 1 - 82.

[Gon 92] Gonzaga, C., Path-following methods for Linear Progamming, *SIAM Review*, 34, N2, (1992) pp. 167-224

[Gros K 81] Grossman, K., Kaplan, A., Nonlinear Programming based on unconstrained minimization, Novosibirsk Nauca, 1981, (in Russian).

[Hes 69] Hestenes, M.R., Multiplier and gradient methods, *JOTA* , Vol. 4, No. 5, (1969), pp. 303-320.

[Huar 67a]   Huard, P., Resolution of mathematical programming with nonlinear constraints by the method of centres. In: Nonlinear Programming, North Holland, Amsterdam (1967), pp. 207-219.

[Huar 67b]   Huard, P., A method of centres by upper-bounding function with applications. In: Nonlinear Programming, North Holland, Amsterdam (1967), pp. 1-30.

[Jar S S 88] Jarre, F., Sonnevend, G., Stoer, J., An implementation of the method of analytic centers, In: Benoussan, A., Lions, J.L., eds., Lecture Notes in Control and Information Sciences, No. 111, (Springer, Berlin 1988).

[Kar 84]   Karmarkar, N.A., A new polynomial-time algorithm for linear programming, *Combinatorica*, Vol. 4, (1984) pp. 373-395.

[Kh 79]   Khachian, L.G., Polynomial algorithm in linear programming *Soviet Math. Doklady*, Vol. 244, No. 5, (1979).

[Lev 65]   Levin, A.Y., About the algorithm for the minimization convex functions, *Dokl Acad. Nauk SSSR*, Vol. 160, No. 2, (1965) (in Russian).

[Man 75]   Mangasarian, O.L., Unconstrained Lagrangians in nonlinear programming, *SIAM J. Control*, Vol. 13, No. 4, (1975), pp. 772-791.

[Mel P 94]   Melman A., Polyak, R., Newton Modified Barrier Method for QP problems, Annals of Operations Research - (special volume on Interior Point Methods in Mathematical Programming) to appear, 1995

[Miff 76]   Mifflin, R., Rates of convergence for a method of centres algorithm, *JOTA*, Vol. 18, No. 2, (1976), pp. 199-288.

[Nes N 94]   Nesterov, Y., Nemirovsky, A., Interior Point Polynomial Methods in Convex Programming: Theory and Algorithms. SIAM, Philadelphia (1994).

[New 65]   Newman, D.J. Location of the maximum on unimodal surfaces, *J. ACM*, Vol. 12, No. 3, (1965), pp. 395-398.

[Pol E 71]   Polak, E., Computational Method in Optimization: A Unified Approach. Academic Press, New York, 1971.

[Pol T 73]   Polyak, B.T., Tret'yakov, N.V., The method of penalty estimates for conditional extremum problems, *Z. Vychisl. Matimat. Fiz.*, Vol. 13, (1973), pp. 34-46.

77

[Pol 86]   Polyak, R., Controlled processes in external and equilibrium problems. Deposited
              manuscript, *VINITI* (1986), 420p., Moscow, (in Russian).

[Pol 92]   Polyak, R., Modified Barrier Functions. (Theory and Methods), 54, Vol. 2, (1992),
              pp. 177-222.

[Pol Teb 95] Polyak, R., Teboulle, M., Nonlinear Rescaling and Proximal-like Methods in Convex
              Optimization, *Research Report*, Dept. of OR, George Mason University, (1995)
              pp. 1-23.

[Pow 69]   Powell, M.J.D., A method for nonlinear constraints in minimization problems. In:
              Optimization, Academic Press, New York (1969), pp. 283-298.

[Ren 88]   Renegar, J., A polynomial-time algorithm, based on Newton's method for linear
              programming, *Mathematical Programming*, Vol. 40, (1988) pp. 59-93.

[Rock 74]   Rockafellar, R.T., Augmented Lagrange multiplier functions and duality in nonconvex
              programming, *SIAM J. Control*, Vol. 12, No. 2, (1974), pp. 268-285.

[Sh 70]   Shor, N.Z., Utilization of the operation of space dilatation in the minimization of
              convex function, Kibernetica 1 (1970) pp. 6-12 , English translation : Cybernetics 6,
              pp. 7-15.

[Sm 86]   Smale, S., Newton's method estimates from data at one point. In: The Merging of
              Disciplines in Pure, Applied and Computational Mathematics (Ewing, R.E., et al)
              Springer-Verlag, New York-Berlin (1986) pp. 185-196.

[Son 85]   Sonnevend, G., An analytical centre for polyhedrons and new classes of global
              algorithms for linear (smooth convex) programming., In lecture Notes Control
              Inform. Sci. 84, Springer-Verlag, New York, NY, (1985), pp. 866-876.

[Vaid 87]   Vaidya, P.M., An algorithm for linear programming which requires $O\ (((m + n\ )\ n^2 +
              (m + n)^{1.5})L\ )$ arithmetic operations, Technical Report, AT & T Bell Laboratories,
              Murray Hill, NJ (1987).

[Van M F 86] Vanderbei, R.J., Meketon, M.S., Freedman, B.A., A modification of Karmarkar's
              linear programming algorithm, *Algorithmica*, Vol. 1, (1986) pp. 395-407.

[Wol 61]   Wolfe, P., A duality Theorem for Non-linear Programming, *Quart. Appl. Math.* 19, (1961), pp. 239-244.

[Yud N 76]   Yudin, D.B., Nemirovsky, A.S., Informational complexity and efficient methods for the solution of convex extremal problems, *Econ. Mat. Metody* 12 (1976), pp. 357-369 (in Russian) *Matecon* 13 (2), pp. 3-25 (in English).

[Zout 60]   Zoutendijk, G., Methods of Feasible Directions. Elsevier Amsterdam Publishing Co., New York (1960).

[ZPP 63]   Zuchovitsky, S.I., Polyak, R.A., Primak, M.E., An algorithm for solving a convex programming problem, *Dokl. Acad., Nauk. SSSR*, Vol. 153, No. 5, (1963).

[ZPP 69]   Zuchovitsky, S.I., Polyak, R.A., Primak, M.E., Two methods for finding an equilibrium point on n-person game, *Soviet Math. Dokl.*, Vol. 10, No. 2, (1969).