МАТЕРИАЛЫ ПО МАТЕМАТИЧЕСКОМУ ОБЕСПЕЧЕНИЮ ЭВМ

ЦЕНТРАЛЬНЫЙ
ЭКОНОМИКО-
МАТЕМАТИЧЕСКИЙ
ИНСТИТУТ
АКАДЕМИИ НАУК
СССР

САМОСОГЛАСОВАННЫЕ ФУНКЦИИ И ПОЛИНОМИАЛЬНЫЕ АЛГОРИТМЫ В ВЫПУКЛОМ ПРОГРАММИРОВАНИИ

Москва
1989

USSR ACADEMY OF SCIENCES
Central Economic & Mathematic Institute

SELF-CONCORDANT FUNCTIONS AND POLYNOMIAL-TIME METHODS IN CONVEX PROGRAMMING

Ju. E. Nesterov, A.S. Nemirovsky

Moscow 1989

In the paper there is developed a general approach to the design of polynomial-time interior point algorithms for Convex Programming problems. This approach results in a number of algorithms for such problems as Linear and Quadratic (including quadratically constrained QP) Programming, Geometric Programming, approximation in L_p - norms, minimization over matrices, finding of extremal ellipsoids.

Authors: Ju.E.Nesterov, A.S.Nemirovsky

Section 0. Introduction

Recently a number of new polynomial-time algorithms for linear and quadratic programming problems were suggested which require $O(m^{1/2} \ln(1/\varepsilon))$ iterations to obtain an approximate solution to the accuracy ε (m is the number of constraints ; the accuracy is estimated in some natural relative scale). The first method of that type was developed by Renegar [Re. 1988]. Now (July, 1988) the authors also know the methods of Gonzaga [Go.1987], Kojima, Mizuno, Yoshise [KMY. 1987], Monteiro, Adler [MA.1987 1], Vaidya [Va. 1987], Todd, Ye [TY. 1987], Nesterov [Ne.1988 1], concerning LP problems, and of Goldfarb, Liu [GL. 1988], Mehrotra, Sun [MS. 1987, 1988], Nesterov [Ne. 1988 2,3], concerning convex quadratic programming; see also [So. 1985], [Ja. 1987], [Fr. 1988]. Being quite different from the analytical viewpoint, these methods are close to each other in their background. Namely, for each of the methods one can find a family of smooth convex functions $F_t(x)$ defined on some convex regions $Q_t \subset R^n$ and depending on the parameter $t \in \Delta$ ($\Delta \subset \mathbb{R}$ is an open interval), such that the trajectory $x_t^* = \text{argmin}\{ F_t(x) \mid x \in Q_t\}$ converges to the solution of the problem under consideration for $t \to t^*$, where t^* is an appropriate endpoint of Δ. The methods construct approximations x_t for the points x_t^* along a sequence $t = t_i \in \Delta$ which converges to t^*. The transformation $x_{t_i} \to x_{t_{i+1}}$ is based on Newton's minimization method: $x_{t_{i+1}}$ is obtained by one step of the method as applied to $F_{t_{i+1}}$ with x_{t_i} chosen as the starting point.

Of course, this scheme of *path-following methods*, is quite traditional; classical results on it are described in [FMcC 1968].But the well-known general results on this scheme do not ensure the polynomial-time convergence even for LP problems. This is the reason for the complicated and ver special analysis one can find in the papers mentioned above. Therefore it seems important to explain these methods from a general point of view and to understand which elements of the constructions are the key ones and what is the widest class of problems the methods can be applied to. For instance, it is interesting to find out, whether or not these methods can be extended to non-linear and non-quadratic problems.

In what follows we try to develop an approach of that kind. We guess we have found out a key property of the family $\{Q_t, F_t\}_{t\in\Delta}$, which underlies the polynomial-time results (this property originates from [Ne. 1988 2] and we shall call it *self-concordance*).

Notice that below the polynomiality is understood in a manner slightly different from the traditional one. Usually a polynomial-time method is defined as a method which produces the exact solution of any problem from the class under consideration and such that the total number of arithmetic operations is bounded by a polynomial function of the problem size, L (the length of the input data). These operations must be performed with $O(p(L))$-bits numbers and the accuracy of $O(p(L))$ bits, where $p(\ )$ is a polynomial. Such a definition is oriented to relatively simple problems (LP, QP); really, in more complicated situations the solution itself need not be rational even for the rational data. Moreover, the above definition, being convenient from the theoretical viewpoint, contradicts the practice of computation, because of the following reasons. (1) Usually numbers are represented not with fixed, but with floating point, while the size of a number in the fixed point form is not bounded by the polynomial of the size of the number in the floating point form. (2) In practice arithmetic operations are not performed with the desired, but with certain fixed accuracy. (3) All known polynomial-time algorithms in numerical (excluding combinatorial) optimization in their nature are "infinite" - i.e. they are based on conceptual converging (but not finitely converging) iterative procedures using precise real arithmetic. The polynomiality in the traditional sense is obtained by "external" (and standard) termination and rounding rules which, as far as we know, are not used in practical computation.

By the described reasons *polynomial-time method* below means an iterative procedure for a class of problems such that the total number of operations of precise real arithmetic, which is required to produce an approximate solution up to the accuracy ε, is bounded by a polynomial of $\ln(1/\varepsilon)$ and of the problem's dimension. Above ε is the *relative* accuracy, measured in some natural scale, and the dimension of a problem usually can be defined as the sum of the numbers of variables and constraints.

In what follows we give summary of the results developed by the authors and taking their origin in [Ne. 1988 1,2,3,4]; some of these results were announced in [NN. 1988].

Below E (possibly, with sub- or superscripts) means a finite-dimensional real vector space, $C(E)$ $(C_B(E))$ means the family of all convex subsets of E with a nonempty interior (bounded convex subsets of E with a nonempty interior, respectively). If F is a smooth function from E into $\mathbb{R}$ then $D^kF(x)[h_1,\dots,h_k]$ means the value of its k-th differential (taken at $x \in E$) at the set of vectors $h_1,\dots,h_k \in E$.

Other often used notations are as follows:

$$\lambda_* = 2 - 3^{1/2},$$

$$\omega(\lambda) = 1 - (1 - 3\lambda)^{1/3},$$

$$\lambda^+ = \lambda^2 (1 - \lambda)^{-2}, \; 0 \le \lambda \le \lambda_*.$$

To simplify cross-references, they are abbreviated: T.1.2 means Theorem 1.2 (prefix 1. denotes the section number), P.1.1 means Proposition 1.1, C.2.1 - Corollary 2.1, and so on. The proofs of the statements are given at the latter subsections of the corresponding sections; each of the proofs is supplied by its own numeration of the expressions.

When necessary, the beginnings of statements, sets of formulae and so on are marked by "□" and the ends - by "■".

Section 1. Self-concordant functions and Newton's method

1.1. Self-concordance. To motivate this notion, let us start with analysing the following traditional situation: given some convex smooth function $F: R^n \to \mathbb{R}$, one desires to minimize it by Newton's method. What are the sufficient conditions for the quadratic convergence? The usual answer is that these are the Lipshitz continuity and the non-degeneracy of the Hessian matrix of F. Notice that the answer requires a Euclidean structure on R^n in which the Hessian matrix condition number and Lipschitz constant are to be evaluated. Of course, the Hessian non-degeneracy and the Lipschitz continuity are independent of the manner in which the Euclidean structure is chosen, but the "numeric characterization" of these properties and hence - the explicit description of the "quadratic convergence region" do depend on this choice and not on F only. Now notice, that the second order differential of F induces an infinitesimal Euclidean metric in R^n, intrinsically connected with F. It turns out that the Lipshitz continuity of D^2F with respect to this metric leads to very interesting consequences concerned with Newton's minimization of F. The property described will be called *self-concordance*. The precise definition is as follows.

Definition 1.1. Let E be a finite-dimensional real vector space, Q be an open nonempty convex subset of E, $F : Q \to \mathbb{R}$ be a function, $a > 0$. F is called *self-concordant* on Q with the parameter value a (notation: $F \in S_a(Q,E)$), if F is C^3 - smooth and convex function on Q, and for each $x \in Q$ and $h \in E$ the following inequality holds:

$$|D^3F(x)[h,h,h]| \leq 2\, a^{-1/2}(D^2F(x)[h,h])^{3/2} \qquad (1.1)$$

A function $F \in S_a(Q,E)$ is called *strongly self-concordant* ($F \in S_a^+(Q,E)$), if the sets $\{x \in Q \mid F(x) \leq t\}$ are closed in E for each $t \in \mathbb{R}$.

Self-concordance is an affine-invariant property:

Proposition 1.1. (i) Let $F \in S_a(Q,E)$ ($F \in S_a^+(Q,E)$) and let $x = \mathcal{A}(y) = Ay + b$ be an affine transformation from a finite-dimensional real vector space E^+ into E, such that $Q^+ = \{y \mid$

$\mathcal{A}(y) \in Q\} \neq \emptyset$, and let $F^+(y) \equiv F(\mathcal{A}(y)): Q^+ \to \mathbb{R}$. Then $F^+ \in S_a(Q^+,E^+)$, $F^+ \in S_a^+(Q^+,E^+)$, respectively.

(ii) Let $F_i \in S_{a_i}(Q_i,E)$, $p_i > 0$, $i = 1,2$, $Q \equiv Q_1 \cap Q_2 \neq \emptyset$, $F(x) = p_1F_1(x) + p_2F_2(x): Q \to \mathbb{R}$ and let $a = \min\{ p_1a_1, p_2a_2 \}$. Then $F \in S_a(Q,E)$. If under the above assumptions eihter $F_i \in S_{a_i}^+(Q_i,E)$, $i=1,2$, or $F_1 \in S_{a_1}^+(Q_1,E)$ and $Q_1 \subseteq Q_2$, then $F \in S_a^+(Q,E)$. ■

The following statement is a finite-difference version of infinitesimal relation (1.1); this statement is the main technical tool in what follows.

Theorem 1.1. Let $F \in S_a(Q,E)$, $x \in Q$ and $e \in E$. Denote $p_x(e) = (D^2F(x)[e,e]/a)^{1/2}$, $\Delta_x(e) = \{ s \geq 0 \mid s\, p_x(e) < 1 \}$, and let $x(s) = x + s\, e$. Then

(i) For each $s \in \Delta_x(e)$ such that $x(s) \in Q$ we have

$$\forall h \in E : (1 - s\, p_x(e))^2 D^2F(x)[h,h] \leq D^2F(x(s))[h,h] \leq (1 - s\, p_x(e))^{-2} D^2F(x)[h,h] \qquad (1.2)$$

(ii) If $F \in S_a^+(Q,E)$ then $x(s) \in Q$ for each $s \in \Delta_x(e)$. ■

Corollary 1.1. Let $F \in S_a(Q,E)$. Then the subspace $E_F = \{h \in E \mid D^2F(x)[h,h] = 0 \}$ does not depend on the choice of $x \in Q$.

Corollary 1.2. Let Q be a convex region in E and let $F \in S_a^+(Q,E)$. For $x \in Q$ and $r > 0$ let $W_r(x) = \{y \in E \mid D^2F(x)[y - x, y - x] < a\, r^2\}$. Then $W_1(x) \subset Q$ for each $x \in Q$. ■

1.2. Newton's method and self-concordant functions. In this item we describe the behaviour of Newton's method on self-concordant functions.

Let $F \in S_a(Q,E)$. For $x \in Q$ denote

$$\lambda(F,x) = \inf\{ \lambda \mid |DF(x)[h]| \leq \lambda\, a^{1/2}(D^2F(x)[h,h])^{1/2} \ \forall h \in E \} \qquad (1.3)$$

(if the set on the right is empty, then $\lambda(F,x) = \infty$ by definition). The quantity $\lambda(F,x)$ can be interpreted as follows. Let us consider the quadratic approximation

$$\Phi_x(y) = F(x) + DF(x)[y - x] + D^2F(x)[y - x, y - x]/2$$

of the function F at the point x. Then, obviously,

$$a\, \lambda^2(F,x)/2 = \Phi_x(x) - \inf\{ \Phi_x(y) \mid y \in E \} \qquad (1.4)$$

or, which is the same,

$$\alpha\,\lambda^2(F,x) = 2\sup\{\, DF(x)[h] - D^2F(x)[h,h]/2 \mid h \in E \,\}. \quad (1.5)$$

It is worthy to notice that

$$\lambda(F,y) \equiv \min\{\, \lambda \geq 0 \mid \forall\, h \in R^n: |DF(y)[h]| \leq \alpha^{1/2}\lambda\,(D^2F(y)[h,h])^{1/2} \,\} \quad (1.6)$$

(i.e. $\lambda(F,x)$ is, within a constant factor, the norm of the linear form $DF(x)$ in the metric defined by $D^2F(x)$).

The quantity $\lambda(F,x)$ will be called *Newton's decrement* of F at x.

Proposition 1.2. For $F \in S_\alpha(Q,E)$ either $\lambda(F,x) \equiv \infty$ for all $x \in Q$, or $\lambda(F,x)$ is a finite continuous function on Q. ■

Let $F \in S_\alpha(Q,E)$, $x \in Q$ and $\lambda(F,x) < \infty$. Then the form $\Phi_x(y)$ is bounded from the below in $y \in E$ and thus attains its minimum over y. Let $x^*(F,x)$ be some minimizer of this form, and let $e(F,x) = x^*(F,x) - x$. Obviously,

$$DF(x)[h] = -D^2F(x)[e(F,x),h] \qquad \forall\, h \in E, \quad (1.7)$$

$$D^2F(x)[e(F,x),e(F,x)] = \alpha\,\lambda^2(F,x). \quad (1.8)$$

Notice that $x^*(F,x)$ is Newton's iterate of x.

The following result shows how the quantities $\lambda(F,\ )$ and $F(\ )$ vary on the segment $[x,x^*(F,x)]$.

Theorem 1.2. Let $F \in S_\alpha(Q,E)$, $x \in Q$ and $\lambda(F,x) < \infty$. Let σ, $0 < \sigma \leq \min\{\, 1, \lambda^{-1}(F,x) \,\}$, be such that the points $x(s) = x + s\,e(F,x)$ belong to Q for $s \in \Delta \equiv [0,\sigma)$. Then for $s \in \Delta$ we have

$$\lambda(F,x(s)) \leq \frac{1 - s - s\lambda + 2s^2\lambda}{(1 - s\lambda)^2}, \quad (1.9)$$

$$F(x) - F(x(s)) \geq \alpha\,\lambda^2\left(s\,\frac{1+\lambda}{\lambda} + \frac{1}{\lambda^2}\ln(1 - s\lambda) \right), \quad (1.10)$$

where $\lambda = \lambda(F,x)$. ■

The following version of the above theorem is some more convenient.

Theorem 1.3. Let $F \in S_\alpha(Q,E)$, $x \in Q$ and $\lambda(F,x) < 1$. Let one of the sets $X_x = \{\, y \in Q \mid \lambda(F,y) \leq \lambda(F,x) \,\}$, $Y_x = \{\, y \in Q \mid F(y) \leq F(x) \,\}$ be closed in E. Then

(i) F attains its minimum over Q: $X_* \equiv \mathrm{Argmin}_Q\, F \neq \emptyset$;

(ii) If $\lambda(F,x) \leq \lambda_* \equiv 2 - 3^{1/2} = 0.2679\ldots$, then $x^*(F,x) \in Q$ and

$$\lambda(F,x^*(F,x)) \leq \frac{\lambda^2(F,x)}{(1 - \lambda(F,x))^2} \leq \lambda(F,x)/2\ ; \quad (1.11)$$

(iii) For each $y \in Q$ such that $\lambda(F,y) < 1/3$ and for each $x^* \in X_*$ we have

$$a^{-1}(F(y) - \min_Q F) \leq \frac{1}{2}\,\omega^2(\lambda(F,y))\,\frac{1 + \omega(\lambda(F,y))}{1 - \omega(\lambda(F,y))}, \quad (1.12)$$

$$a^{-1}\, D^2F(x^*)[y-x^*, y-x^*] \leq (1 - \omega(\lambda(F,y)))^{-2}\,\omega^2(\lambda(F,y)), \quad (1.13)$$

$$a^{-1}\, D^2F(y)[x^*-y, x^*-y] \leq \omega^2(\lambda(F,y)),$$

where $\omega(\lambda) = 1 - (1 - 3\lambda)^{1/3}$.

(iv) For each $y \in Q$ such that $\delta^2(F,y) \equiv 2\,a^{-1}\,(F(y) - \min_Q F) < 4/9$, we have

$$\lambda(F,y) \leq \frac{24\,\delta(F,y)}{(3+(9-12\delta(F,y))^{1/2})((9-12\delta(F,y))^{1/2}-1)^2}. \quad ■ \quad (1.14)$$

The following theorem summaries the above statements.

Theorem 1.4. Let $F \in S_a(Q,E)$ $x \in Q$ and let the set $X = \{y \in Q \mid F(y) \leq F(x)\}$ be closed in E. Then

(i) F is bounded from below on Q iff it attains its minimum over Q. If $\lambda(F,x) < 1$, then F attains its minimum over Q;

(ii) Let $\lambda(F,x) < \infty$, $\lambda_* = 2 - 3^{1/2} = 0.2679...$ and $\lambda' \in [\lambda_*, 1)$. Consider the *Newton iteration* starting at x:

$$x_0 = x; \quad x_{i+1} = x_i + \sigma'(\lambda(F,x_i))\, e(F,x_i), \; i \geq 0, \quad (1.15)$$

where

$$e(F,x_i) \in \mathrm{Argmin}\{DF(x_i)[h] + \tfrac{1}{2}\, D^2F(x_i)[h,h] \mid h \in E\}, \quad (1.16)$$

$$\sigma'(\lambda) = \begin{cases} (1+\lambda)^{-1}, & \lambda > \lambda' \\ (1-\lambda)\,\lambda^{-1}(3-\lambda)^{-1}, & \lambda' \geq \lambda \geq \lambda_* \\ 1, & \lambda < \lambda_* . \end{cases} \quad (1.17)$$

The iterations are well-defined (i.e. for all i we have $x_i \in X$, $\lambda_i \equiv \lambda(F,x_i) < \infty$ and $e(F,x_i)$ is well-defined), and the following relations hold:

$$\lambda_i > \lambda' \Rightarrow F(x_{i+1}) \leq F(x_i) - a\,(\lambda_i - \ln(1 + \lambda_i)) \leq F(x_i) - a\,(\lambda' - \ln(1 + \lambda')); \quad (1.18)$$

$$\lambda' \geq \lambda_i \geq \lambda_* \Rightarrow$$

$$\lambda_{i+1} \leq \frac{6\,\lambda_i - \lambda_i^2 - 1}{4} < \lambda_i ,$$

$$1 - \lambda_{i+1} \geq (1 - \lambda_i)\,\frac{5 - \lambda_i}{4} \geq (1 - \lambda_i)\,\frac{5 - \lambda'}{4}; \quad (1.19)$$

$$\lambda_i < \lambda_* \Rightarrow \lambda_{i+1} \leq \lambda_i^2\,(1 - \lambda_i)^{-2} < \tfrac{1}{2}\,\lambda_i . \quad (1.20)$$

Moreover,

$$\lambda_i < 1/3 \Rightarrow F(x_i) - \min_Q F \leq \tfrac{1}{2}\, a\, \omega^2(\lambda_i)\,\frac{1 + \omega(\lambda_i)}{1 - \omega(\lambda_i)}. \quad (1.21)$$

Comments. Let $F \in S_a(Q,E)$ be bounded from below on Q and let $x \in Q$. Assume that the set $\{y \in Q \mid F(y) \leq F(x)\}$ is closed in E. By T.1.4 F attains its minimum over Q and $\lambda_0 \equiv \lambda(F,x) < \infty$, while the above described Newton's iterations converge (in objective's values) to the minimizer of F over Q. Moreover, $\lambda_t \to 0$, $t \to \infty$. The theorem shows that Newton's process can be divided into three sequential stages with the values of iteration number t as follows:

$$t < t_*(1) \equiv \min\{ t \mid \lambda_t \leq \lambda' \};$$

$$t_*(1) \leq t < t_*(2) \equiv \min\{ t \mid \lambda_t < \lambda_* \};$$

$$t \geq t_*(2) .$$

At the first stage F decreases at each iteration by a quantity which is not smaller than $a\,(\lambda' - \ln(1 + \lambda')) \equiv a\,\lambda^*$; the iterations number $t_*(1)$ of this stage is not greater than

$$t(1) \equiv \,](a\,\lambda^*)^{-1}\,(F(x) - \min_Q F)[.$$

At the second stage the quantities λ_t decrease, and the quantities $(1 - \lambda_t)$ increase as a geometric progression with the ratio $\varkappa = (5 - \lambda')/4$; the iterations number $t_*(2) - t_*(1)$ of this stage is not greater than

$$t(2) \equiv 1 + \,]\ln^{-1}(\varkappa)\,\ln\left(\frac{1 - \lambda_*}{1 - \lambda'}\right)[.$$

At the third stage the quantities λ_t quadratically decrease; it is important that the behaviour of λ_t at the second stage depends on λ' only, and at the third stage it does not depend on any parameter of the objects involved.

The inequality $\lambda(F,x) < 1$, which under the theorem conditions ensures the boundness of F from below, can not be weakened. This is demonstrated by the example

$$F(x) = \ln 1/x \in S_1^+((0,\infty),\mathbb{R}),$$

where $\lambda(F,x) \equiv 1$.

1.3. Self-concordant functions and duality. It turns out that the Legendre transformation of a strongly self-concordant function is strongly self-concordant with the same parameter value. The corresponding definitions and results are as follows.

Let E be a finite-dimensional real vector space and E^* be the space conjugate to E. The value of a form $\xi \in E^*$ at the vector $x \in E$ will be denoted by $\langle\xi,x\rangle$. Let $a > 0$. An *(a,E)-pair* is, by definition, an arbitrary pair (Q,F), where Q is nonempty convex open set in E and $F \in S_a^+(Q,E)$ is such that $E_F = \{0\}$ (i.e. $D^2F(x)$ is non-degenerate for each $x \in Q$). The Legendre transformation of an (a,E)-pair (Q,F) is defined as the pair $(Q,F)^* \equiv (Q^*,F^*)$, where

$$Q^* = \Phi(Q) \quad (\Phi(x) = DF(x)[\ ]: Q \to E^*),$$
$$F^*(\xi) = \sup\{\langle\xi,x\rangle - F(x) \mid x \in Q\}.$$

The following statement is true.

Proposition 1.3. Let (Q,F) be an (a,E)-pair and (Q^*,F^*) be its Legendre transformation. Then (Q^*,E^*) is an (a,E^*)-pair and

$$Q^* = \{\xi \in E^* \mid \text{the function } F_\xi(x) = F(x) - \langle\xi,x\rangle \text{ is bounded from below on } Q\}.$$

Moreover, $(Q^*,F^*)^* = (Q,F)$ (we use the standard equivalence between $(E^*)^*$ and E). ∎

1.4. Proofs

1.4.1. Proposition 1.1 can be proved by a straightforward verification of the corresponding definitions. ∎

1.4.2. Theorem 1.1.

(1). Let $s \in \Delta_x(e)$ be such that $x(s) \in Q$, and let $h \in E$. Let for $\Delta = [0,s]$ and $\rho \in \Delta$ the function $\psi(\rho)$ be defined as

$\psi(\rho) = D^2F(x(\rho))[e,e]$ and let $\phi(\rho) = D^2F(x(\rho))[h,h]$. By virtue of (1.1) for each triple of vectors $h_i \in E$, $i=1,2,3$, we have $D^3F(u)[h_1,h_2,h_3] \leq 2\,a^{-1/2} \prod_{i=1}^{3} (D^2F(u)[h_i,h_i])^{1/2}$, $u \in Q$, which implies

$|\psi'(\rho)| \leq 2\,a^{-1/2}(\psi(\rho))^{3/2}$, $|\phi'(\rho)| \leq 2\,a^{-1/2}\psi^{1/2}(\rho)\phi(\rho)$. (1)

By the first relation in (1) either $\psi(\rho)\equiv 0$, $\rho \in \Delta$, and then, by virtue of the second relation in (1), $\phi(s) = \phi(0)$, or $\psi(\rho)$ is positive over Δ, and then $|(\psi^{-1/2}(\rho))'_\rho| \leq a^{-1/2}$, $\rho \in \Delta$, which implies

$$\psi^{-1/2}(\rho) \geq \psi^{-1/2}(0) - \rho\, a^{-1/2}. \qquad (2)$$

In the latter case, by $\psi(0) = p_x^2(e)\, a$, we have $\psi^{-1/2}(0) > \rho\, a^{-1/2}$ for $\rho \in \Delta \subseteq \Delta_x(e)$, and (2) implies

$$\psi^{1/2}(\rho) \leq a^{1/2}p_x(e)/(1 - \rho\, p_x(e)).$$

So the second relation in (1) can be rewritten as

$$|\phi'(\rho)| \leq (2\, p_x(e)/(1 - \rho\, p_x(e)))\, \phi(\rho),\ \rho \in \Delta.$$

Thus, either $\phi\equiv 0$ on Δ, or ϕ is positive over Δ, and in the latter case $|\ln(\phi(s)/\phi(0))| \leq 2\ln(1/(1 - s\, p_x(e)))$, which, by definition of ϕ, leads to (1.2). Obviously, (1.2) holds in the above situations $\psi(\rho) \equiv 0$, $\rho \in \Delta$, and $\phi(\rho) = \phi(0)$, $\rho \in \Delta$. (i) is proved.

(ii). Let $\sigma = \sup\{ s \in \Delta_x(e) \mid x(s) \in Q \}$. We desire to prove that $\sigma = 1/p_x(e)$ ($1/0 = \infty$). Assume the latter does not hold, so $\sigma\, p_x(e) < 1$. Then, by (1.2), the function $g(s) = F(x(s))$ has bounded second derivative for $0 \leq s < \sigma$ and hence is bounded for these s. Since $F \in S_a^+(Q,E)$ this leads to $x(\sigma) = \lim_{s\to\sigma} x(s) \in Q$; since Q is open, we have $x(s) \in Q$ for certain $s > \sigma$. The latter under the condition $\sigma\, p_x(e) < 1$ contradicts the definition of σ. (ii) is proved. ■

1.4.3. Corollary 1.1. The set $\{ x \in Q \mid D^2F(x)[h,h] = 0 \} \equiv X_h$ is closed in Q by virtue of the continuity of $D^2F(x)[h,h]$ in x and is open in Q by (1.2); hence this set is either empty, or coincides with Q. ■

1.4.4. Corollary 1.2. This Corollary is reformulation of T.1.1.(i). ■

1.4.5. Proposition 1.2. Let $x \in Q$. It is clear that

$$(\lambda(F,x) < \infty) \Leftrightarrow (DF(x)[h] = 0 \ \forall \ h \in E_F).$$

Assume that $\lambda(F,x) < \infty$ for given x, and let $h \in E_F$. For $\psi(y) = DF(y)[h]$ we have $D\psi(y)[e] = D^2F(y)[h,e] = 0$, so ψ is a constant in Q; since $\psi(x) = 0$, we have $\psi \equiv 0$. Thus, if $\lambda(F,x) < \infty$, then $DF(y)[h] = 0$ for each $h \in E_F$ and each $y \in E$, and in this situation

$$a^{1/2}\ \lambda(F,y) = \min\{\ |DF(y)[h]|(D^2F(x)[h,\ h])^{-1/2}\ |\ h \in E^F,\ h \neq 0\ \},$$

where E^F is a complement to E_F in E. The form $D^2F(y)[h,h]$ is positive defined on the subspace E^F, hence the continuity of the first and the second derivatives of F implies the statement. ∎

1.4.6. Theorem 1.2. Let $e = e(F,x)$. Notice that, by (1.8), $\lambda = p_x(e)$. By T.1.1, since $\sigma \leq 1/\lambda$, we have for all $s \in \Delta$ and $h \in E$:

$$(1 - s\ \lambda)^2\ D^2F(x)[h,h] \leq D^2F(x(s))[h,h] \leq \leq (1 - s\ \lambda)^{-2}\ D^2F(x)[h,h], \qquad (1)$$

so (below $s \in \Delta$)

$$|D^2F(x)[h,h] - D^2F(x(s))[h,h]| \leq \leq \left(\frac{1}{(1 - s\ \lambda)^2} - 1\right) D^2F(x)[h,h]. \qquad (2)$$

We see that

$$\left|\ \frac{d}{ds} DF(x(s))[h] - D^2F(x)[e,h]\ \right| \leq$$

$$\leq \left(\frac{1}{(1 - s\lambda)^2} - 1\right) (D^2F(x)[h,h])^{1/2}\ (D^2F(x)[e,e])^{1/2} \leq$$

$$\leq \left(\frac{1}{(1 - s\lambda)^2} - 1\right)(D^2F(x)[h,h]\ )^{1/2} a^{1/2}\ \lambda, \qquad (3)$$

or, by (1.7),

$$|DF(x(s))[h] - (1-s)\ DF(x)[h]| \leq$$

$$\leq a^{1/2}\ \frac{(s\lambda)^2}{(1 - s\lambda)}\ (D^2F(x)[h,h])^{1/2}. \qquad (4)$$

Now, by definition of $\lambda(F,x)$, (4) and (1), we have:

$$2\ \sup\{DF(x(s))[h] - D^2F(x(s))[h,h]/2\ |\ h \in E\ \} \leq$$

$$\leqslant 2 \sup(DF(x)[h]\ (1 - s) + a^{1/2} \frac{(s\lambda)^2}{(1 - s\lambda)} (D^2F(x)[h,h])^{1/2} -$$

$$- \frac{(1-s\lambda)^2}{2} D^2F(x)[h,h] \mid h \in E \} \leqslant$$

$$\leqslant 2 \sup(a^{1/2} \lambda\ (1 - s)\ (D^2F(x)[h,h])^{1/2} + a^{1/2} \frac{(s\lambda)^2}{(1 - s\lambda)}.$$

$$(D^2F(x)[h,h])^{1/2} - \frac{(1 - s\lambda)^2}{2} D^2F(x)[h,h] \mid h \in E \} \leqslant$$

$$\leqslant a\ \lambda^2 \left(\frac{1 - s - s\lambda + 2\ s^2\lambda}{(1- s\lambda)^2} \right)^2,$$

which together with (1.6) implies (1.9) (notice that $(1 - s) \geqslant 0$).

Let $f(s) = F(x) - F(x(s))$. Relation (4) with $h = e$ and relation (1.7) lead to

$$f'(s) = DF(x(s))[e] \leqslant$$

$$\leqslant (1 - s)\ \ DF(x)[e] + a^{1/2} \frac{(s\lambda)^2}{(1 - s\lambda)}(D^2F(x)[e,e])^{1/2} =$$

$$= -(1 - s)D^2F(x)[e,e] + a^{1/2} \frac{(s\lambda)^2}{(1 - s\lambda)} (D^2F(x)[e,e])^{1/2} =$$

$$= - (1 - s)\ a\ \lambda^2 + a\ \lambda \frac{(s\lambda)^2}{(1 - s\lambda)},$$

hence

$$f(s) \leqslant f(0) - a\ \lambda^2 \int_0^s (1 - \rho - \frac{\rho^2\ \lambda}{1 - \rho\ \lambda}\ )\ d\rho,$$

which implies (1.10) ∎

1.4.7. Theorem 1.3.

1^0. Let $\sigma(\lambda) = \min(1, \frac{1 - \lambda}{\lambda\ (3 - \lambda)})$, $\Delta_\lambda = [0,\sigma(\lambda)]$ and

$$\phi_\lambda(s) = \frac{1 - s - s\lambda + 2s^2\lambda}{(1 - s\lambda)^2},$$

$$\psi_\lambda(s) = s\ \frac{1 + \lambda}{\lambda} + \lambda^{-2} \ln(1 - s\lambda),\ s \in \Delta_\lambda,$$

for $0 \leqslant \lambda < 1$. It is easy to show that ϕ_λ decreases on Δ_λ, and ψ_λ is nonnegative on Δ_λ. Let $X = X_x \cap Y_x$ and $u \in X$; then $\lambda \equiv \lambda(F,u) \leqslant \lambda(F,x) < 1$. Let

$$s' = \sup(s \in \Delta_\lambda \mid u + s\ e(F,u) \in Q \}.$$

By (1.9), (1.10), for $0 \leqslant s < s'$ we have

$$\lambda(F,u+se(F,u)) \leqslant \lambda\ \phi_\lambda(s) \leqslant \lambda,$$

$$F(u+se(F,u)) \leqslant F(u) - a\,\lambda^2\,\psi_\lambda(s) \leqslant F(u). \qquad (1)$$

The sets X_x and Y_x are closed in Q (since $\lambda(F,\ )$ and $F(\ )$ are continuous over Q), and one of the sets is closed in E; hence X is closed in E. By (1), $u+se(F,u) \in X$, $0 \leqslant s < s'$, and by virtue of the closedness of X in E we have $u+s'e(F,u) \in X$, and (1), by continuity arguments, holds for $s = s'$. Thus, $u+s'e(F,u) \in X \subset Q$; since Q is open, the inclusion $u+s'e(F,u) \in Q$ is possible only if $s' = \sigma(\lambda)$ (the definition of s'). Thus, we get

$$u \in X \quad \Rightarrow \quad u^*(u) \equiv u + \sigma(\lambda(F,u))e(F,u) \in X$$

and

$$\lambda(F,u^*(u)) \leqslant \lambda(F,u)\,\psi_{\lambda(F,u)}(\sigma(\lambda(F,u)), \qquad (2)$$

$$F(u^*(u)) \leqslant F(u) - a\,\lambda^2(F,u)\,\psi_{\lambda(F,u)}(\sigma(\lambda(F,u))). \qquad (3)$$

2^0. Assume that $u \in X$ is such that $\lambda(F,u) \leqslant 2 - 3^{1/2}$. Then $\sigma(\lambda(F,u)) \geqslant 1$, hence $u^*(u) = x^*(F,u) \in Q$, and, by (2),

$$\lambda(F,x^*(F,u))\,\phi_{\lambda(F,u)}(1) = \lambda^2(F,u)/(1 - \lambda(F,u))^2 \leqslant$$

$$\leqslant ((2 - 3^{1/2})/(3^{1/2} - 1)^2)\,\lambda(F,u),$$

which leads to (1.11).

3^0. Consider the following process:

$$x_0 = x; \quad x_{t+1} = u^*(x_t) \equiv x_t + \sigma(\lambda(F,x_t))e(F,x_t),\ t \geqslant 0 \qquad (4)$$

(the process is well defined by the arguments of sect. 1^0, namely, $x_t \in X \subset Q\ \forall\ t \geqslant 0$). This is the Newton minimization of F starting at x with certain step length choice; we shall see that for all large enough t it turns out that $\sigma(\lambda(F,x_t)) = 1$, hence for these t (4) is the usual Newton minimization.

Let $\lambda_t = \lambda(F,x_t)$. Then, by (2),(3):

$$\lambda_{t+1} \leqslant \lambda_t\,\phi_{\lambda_t}(\sigma(\lambda_t)), \quad F(x_{t+1}) < F(x_t), \qquad (5)$$

which implies that $\lambda_t \to 0$, $t \to \infty$; in particular, for all large enough t we have $\lambda_t < \lambda_*$, or $\sigma(\lambda_t) = 1$, as was promised above. Let $t_* = \min\{ t \mid \lambda_t < \lambda_* \}$. Then for $t \geqslant t_*$ we have, by (1.11),

$$\lambda_{i+1} \leq \lambda_i^2/(1 - \lambda_i)^2 < \lambda_i/2, \quad \lambda_{i_*} < 2 - 3^{1/2}. \tag{6}$$

Notice, that the behaviour of λ_i depends on $\lambda_0 \equiv \lambda(F,x)$ only (this quantity must be < 1), and λ_i quadratically converge to 0 by virtue of (6).

4^0. Let us prove (1). We can assume that $\lambda_i > 0$, $i > 0$ – otherwise (1) is obvious. Let E^F be a complement to E_F in E. Let

$$V_i = \{ y \in x_i + E^F \mid D^2F(x_i)[y-x_i, y-x_i] \leq 100\, a\, \lambda_i^2 \};$$

then V_i is a compact (because $D^2F(x)[\ ,\ ]$ is positively defined on E^F).

Let

$$\omega_{i,\varepsilon}(s) = a\,\{ s^2/2 + \varepsilon\,(\lambda_i s + \int_0^s \rho^2(1-\rho)^{-1}d\rho)\}$$

for $\varepsilon = \pm 1$. Assume that $i > 2$ is such that for $s_i = 10\,\lambda_i$ one has

$$s_i < 1;$$

$$\omega_{i,-1}(s_i) > 0; \quad \omega_{i,1}(s) \leq F(x_0) - F(x_1),\ 0 \leq s < s_i;$$

$$(1 - s)^{-2}(s + \lambda_i - \lambda_i s) < \lambda_0,\ 0 \leq s \leq s_i \tag{7}$$

(since $\lambda_i > 0$ and $\lambda_i \to 0$, (7) holds for all large enough i). Let us verify that for chosen value of i the following inclusion holds:

$$V_i \subset X. \tag{8}$$

Indeed, let $e \in E^F$ be such that $D^2F(x_i)[e,e] = a$, and let

$$\sigma_e = \sup\{ s \in [0,1] \mid x(s,e) \equiv x_i + se \in Q\}.$$

By T.1.1 for all $s \in [0,\sigma_e)$ and $h \in E$ we have:

$$|D^2F(x(s,e))[h,h] - D^2F(x_i)[h,h]| \leq \{(1-s)^{-2}-1\}\, D^2F(x_i)[h,h],$$

$$D^2F(x(s,e))[h,h] \geq (1 - s)^2\, D^2F(x_i)[h,h],$$

which leads to

$$\left|\frac{d}{ds}\,(DF(x(s,e))[h]) - D^2F(x_i)[e,h]\right| \leq$$

$$\leq a^{1/2}\,\{(1 - s)^{-2} - 1\}\,(D^2F(x_i)[h,h])^{1/2},$$

or

$$|DF(x(s,e))[h] - s\,D^2F(x_i)[e,h] - DF(x_i)[h]| \leq$$

$$\leq a^{1/2}\ s^2\ (1 - s)^{-1}\,(D^2F(x_i)[h,h])^{1/2}.$$

Hence

$2\ \sup(DF(x(s,e))[h] - \frac{1}{2}\ D^2F(\ x(s,e)\ )[h,h]\ |\ h \in E) \leq$
$\leq 2\ \sup(DF(x_i)[h] + s|D^2F(x_i)[e,h]| +$
$+ a^{1/2}\ s^2\ (1-s)^{-1}\ (D^2F(x_i)[h,h])^{1/2} - \frac{1}{2}\ (1-s)^2\ D^2F(x_i)[h,h]\ |$
$h \in E\ ) \leq 2\ \sup(DF(x_i)[h] + a^{1/2}\ (\ s + s^2\ (1 -$
$s)^{-1})(D^2F(x_i)[h,\ h])^{1/2} - \frac{1}{2}\ (1 - s)^2\ D^2F(x_i)[h,h]\ |\ h \in E\ ) \leq$
$2\ \sup(\ a^{1/2}\ (\lambda_i + s\ (1 - s)^{-1})(D^2F(x_i)[h,h])^{1/2} - \frac{1}{2}\ (1 - s)^2$
$D^2F(x_i)[h,h]\ |\ h \in E) \leq a\ (s + \lambda_i - \lambda_i s)^2\ (1 - s)^{-4},$

which, by virtue of (I.6) and (7), leads to

$$\lambda(x(s,e)) < \lambda_0,\quad 0 \leq s < \sigma_e. \tag{9}$$

Let $f(s) = F(x(s,e)) - F(x_i)$; then for $0 \leq s < \sigma_e$: $f''(s) = D^2F(x(s,e))[e,e]$, so

$$|f''(s) - D^2F(x_i)[e,e]| \leq a\ ((1 - s)^{-2} - 1),$$

which, by virtue of $f'(0) = DF(x_i)[e]$, leads to s
$DF(x_i)[e] + a\ (\ \frac{1}{2}\ s^2 - \int_0^s \rho^2\ (1 - \rho)^{-1}\ d\rho\ ) \leq f(s) \leq$
$\leq s\ DF(x_i)[e] + a\ (\ \frac{1}{2}\ s^2 + \int_0^s \rho^2\ (1 - \rho)^{-1}\ d\rho\ ),$

so by virtue of

$$|DF(x_i)[e]| \leq \lambda_i\ a^{1/2}\ (D^2F(x_i)[e,\ e])^{1/2} = a\ \lambda_i$$

we have ($\sigma_e^* = \min(\ s_i,\ \sigma_e\ )$):

$$f(s) \leq \omega_{i,1}(s),\quad f(s) \geq \omega_{i,-1}(s),\ 0 \leq s < \sigma_e \tag{10}$$

By (7) and (5) the relations (9), (10) imply the inclusion $x(s,e) \in X$ for $0 \leq s < \sigma_e^*$. Since X is closed in E, we have $x(\sigma_e^*,e) \in X$; since Q is open, the latter, by definition of σ_e^*, is possible only when $\sigma_e = s_i$; this implies (8), since e is an arbitrary vector from E^F such that $D^2F(x_i)[e,e] = a$.

Notice that the points belonging to the (relative) boundary ∂V_i of V_i are of the form $x_i + s_i\ e,\ e \in E^F$, $D^2F(x_i)[e,e] = a$; so, taking into account (10) and (7), we get $F(u) > F(x_i),\ u \in \partial V_i$. Hence there exist a point $x_* \in V_i$, such that $DF(x_*)[h] = 0,\ h \in E^F$. By virtue of P.1.2 under the conditions of the theorem under consideration one has $DF(u)[h] = 0$ for each $u \in Q$ and $h \in E_p$; hence $DF(x_*) = 0$. (1) is proved.

5^0. Let us prove (iii). Let $e = y - x_*$, $\lambda = \lambda(F,y)$, $\omega = p_y(e)$ ($\equiv (D^2F(y)[e,e]/a)^{1/2}$), $x(s) = x_* + se$, $y(s) = y - se = x(1-s)$, $f(s) = F(y(s)) - F(x_*)$, $0 \leq s \leq 1$. We have

$$DF(y(s))[e] = -f'(s) \geq 0,\ 0 \leq s \leq 1. \tag{11}$$

Let $\sigma = \min\{1, \omega^{-1}\}$. Since $\frac{d}{ds}\{-DF(y(s)[e]\} = D^2F(y(s)[e,e]$, we, by virtue of T.1.1, get

$$\frac{d}{ds}\{-DF(y(s))[e]\} \geq (1 - s\omega)^2 D^2F(y)[e,e] = a\,\omega^2 (1 - s\omega)^2,$$

$0 \leq s \leq \sigma$, thus

$$DF(y(s))[e] \leq DF(y)[e] - a\,\omega^2 \int_0^s (1 - \rho\omega)^2\, d\rho \leq$$

$$\leq a\,\omega\,\{\lambda - s\,\omega\,(3 - 3s\omega + s^2\omega^2)/3\},\quad 0 \leq s \leq \sigma.$$

This, together with (11), implies

$$3\lambda \geq a\,\omega\,(3 - 3s\omega + s^2\omega^2),\ 0 \leq s \leq \sigma. \tag{12}$$

If $\omega \geq 1$, then $\sigma = \omega^{-1}$, and (12) holds for $s = \omega^{-1}$, which implies $\lambda \geq 1/3$; this contradicts to the conditions in (iii). Hence $\omega < 1$, so $\sigma = 1$, and (12) implies $\omega(3 - 3\omega + \omega^2) \leq 3\lambda$. In the latter inequality the left hand side is monotone in $\omega > 0$, so

$$\omega \leq \omega(\lambda) \equiv 1 - (1 - 3\lambda)^{1/3} \tag{13}$$

($\omega(\lambda)$ is the unique root of the equation $\omega(3 - 3\omega + \omega^2) = 3\lambda$); (13) is the second relation in (1.13).

Now let $g(s) = F(x(s)) - F(x(0))$ ($= F(x(s)) - F(x_*)$). Then $g(0) = g'(0) = 0$, and for $0 \leq s \leq 1$ we have

$$g''(s) = D^2F(x(s))[e,e] \leq (1 - (1 - s)\,\omega)^{-2} D^2F(y)[e,e]$$

(we have taken into account T.1.1), so

$$g(1) \leq a\,\omega^2 \int_0^1 \{\int_0^s (1 - (1 - \rho)\,\omega)^{-2}\, d\rho\}\, ds =$$

$$= a\,\{\omega\,(1 - \omega)^{-1} + \ln(1 - \omega)\} = a\,\{(\omega + \omega^2 + \omega^3 + \dots) -$$

$$- (\omega + \tfrac{1}{2}\omega^2 + \tfrac{1}{3}\omega^3 + \dots)\} \leq a\,\{\tfrac{1}{2}\omega^2 + \omega^3/(1 - \omega)\} =$$

$$= \tfrac{1}{2}\, a\,\omega^2\, \frac{1 + \omega}{1 - \omega};$$

this together with (13) implies (1.12). Further, by (1.13) and T.1.1 one has

$$D^2F(x_*)[e,e] \leq (1-\omega)^{-2}\, D^2F(y)[e,e],$$

which, together with (13), implies (1.13). (iii) is proved.

6^0. Let us prove (iv). Let $x_* \in \mathrm{Argmin}_D\, F$ and let $y \in Q$ be such that $\delta^2 = 2\,a^{-1}\,(F(y) - F(x_*)) < 4/9$. Let

$$e = y - x_*,\quad \omega = p_{x_*}(e),\quad x(s) = x_* + se,\ 0 \leq s \leq 1,$$

$$\sigma = \min\{1, \omega^{-1}\},\quad f(s) = F(x(s)) - F(x_*).$$

By T.1.1 for $0 \leq s \leq \sigma$ we have $f''(s) > a\,\omega^2\,(1-\omega s)^2$, so, by $f'(0) = 0$,

$$f(s) \geq a\,\omega^2 \int_0^s \left(\int_0^t (1 - z\,\omega)^2\, dz\right) dt =$$

$$= \frac{1}{12}\, a\,\omega^2\, s^2\, (6 - 4\omega s + \omega^2 s^2),\ 0 \leq s \leq \sigma.$$

If $\omega \geq 1$, then $\sigma = \omega^{-1}$, and we get $\delta^2 = 2\,a^{-1}\ f(1) \geq 1/2 > 4/9$, which is impossible.

Hence $\omega < 1$, so $\sigma = 1$, and our inequality implies

$$\frac{1}{12}\,\omega^2\,(6 - 4\omega + \omega^2) \leq \delta^2 \quad\text{and}\quad \omega < 1, \tag{14}$$

and hence

$$\frac{1}{18}\,\omega^2\,(3-\omega)^2 \leq \delta^2 \quad\text{and}\quad \omega < 1. \tag{15}$$

For $0 \leq s \leq 1$ and $h \in E$ we, by T.1.1, have

$$\left|\frac{d}{ds}\ F(x(s))[h]\right| = |D^2F(x(s))[h,e]| \leq$$

$$\leq (D^2F(x(s))[h,h])^{1/2}\ (D^2F(x(s))[e,e])^{1/2} \leq$$

$$\leq a^{1/2}\,\omega\,(1 - s\omega)^{-2}\ (D^2F(x_*)[h,h])^{1/2},$$

so by virtue of $DF(x_*)[h] = 0$ we get

$$|DF(y)[h]| \leq a^{1/2}\,\omega\,(1-\omega)^{-1}\ (D^2F(x_*)[h,h])^{1/2}.$$

By T.1.1 we also have $D^2F(y)[h,h] \geq (1-\omega)^2\, D^2F(x_*)[h,h]$; the inequalities obtained lead to $\lambda(F,y) \leq \omega\ (1-\omega)^{-2}$, which, together with (15), implies (1.14). ∎

1.4.8. Theorem 1.4.

1^0. Let $\lambda(F,x) < \infty$. Let J be the set of all integers $j \geq 0$ satisfying the conditions as follows:

(1_j) process (1.15) is well defined for $0 \leq i \leq j$, i.e. for the above i one has $x_i \in X$, $\lambda_i < \infty$, $e(F,x_i)$ are well defined;

(2_j) for $0 \leq i < j$ the implications (1.18)-(1.20) hold.

Let us verify that $J = \{j \geq 0\}$. First of all, $0 \in J$. Indeed,

(2_0) is obvious, and to prove (1_0) we need to verify that $e(F,x)$ is well defined; the latter fact follows from the above assumption that $\lambda(F,u)$ is finite for $u = x$ and P.1.2.

It remains to verify the implication $j \in J \Rightarrow j+1 \in J$. Let $j \in J$, so $x_j \in X$ and $e_j = e(F,x_j)$ is well defined.

Assume that $\lambda_j > \lambda'$. Let

$\sigma = \sup\{ s \in [0, \sigma'(\lambda_j)] \mid x(s) = x_j + s\, e_j \in Q\}$.

Then $\sigma \leqslant \min(1, \lambda_j^{-1})$. By T.1.2 (see (1.10)) in the case under consideration for $0 \leqslant s \leqslant \sigma$ one has

$$F(x(s)) - F(x_j) \leqslant - a\,\lambda_j^2 \left(s \frac{1 + \lambda_j}{\lambda_j} + \lambda_j^{-2} \ln(1 - s\,\lambda_j)\right) \leqslant 0$$

(we have taken into account that $\sigma \leqslant (1 + \lambda_j)^{-1} = \sigma'(\lambda_j)$). Hence $x(s) \in X, 0 \leqslant s < \sigma$, and, by the fact that X is closed, we have $x(\sigma) \in X$. By definition of σ the latter is possible only if $\sigma = \sigma'(\lambda_j)$, which implies $x_{j+1} \in X$; this, together with (1_j) and the above remarks on $\lambda(F,u)$ and $e(F,u)$, leads to (1_{j+1}). Since $x_{j+1} \in X$, the above inequality for $F(x(s)) - F(x_j)$ holds, by the continuity arguments, for $s = \sigma = \sigma'(\lambda_j)$ as well, which together with (2_j) implies (2_{j+1}). Thus, in the case under consideration we have $j+1 \in J$.

Now let $\lambda_j < \lambda'$. By the arguments from the subsection 1.4.6.1^0 (where one must set $u = x = x_j$; the theorem is appliable since X_{x_j} is closed in E together with X, because $x_j \in X$) and by virtue of $\sigma'(\lambda) = \sigma(\lambda)$ for $\lambda < \lambda'$ we have $x_{j+1} \in X$, which, together with (1_j) implies (1_{j+1}). Relations (1.13), as applied to the above u, prove the implications (1.19)-(1.21) for $i = j$, which together with (2_j) leads to (2_{j+1}). Thus $j+1 \in J$.

2^0. Now we can prove (ii). All the statements in (ii), excluding (1.21), immediately follows from $\{(1_j), (2_j) \mid j \geqslant 0\}$. Let us verify (1.21). Let $\lambda_i < 1/3$. The set

$\{y \in Q \mid F(y) \leqslant F(x_i)\}$

is closed in E together with X by virtue of $x_i \in X$ so (1.21) follows from T.1.3.(iii), where one must set $x = x_i$, $y = x_i$.

$\in X$) and by virtue of $\sigma'(\lambda) = \sigma(\lambda)$ for $\lambda < \lambda'$ we have $x_{j+1} \in X$, which, together with (1_j) implies (1_{j+1}). Relations (1.13), as applied to the above u, prove the implications

is closed in E together with X by virtue of $x_i \in X$ so (1.21) follows from T.1.3.(iii), where one must set $x = x_i$, $y = x_i$.

3^o. It remains to prove (i). Under the conditions of the theorem the set $\{y \in Q \mid F(y) \leq F(x')\}$ is closed in E for each $x' \in X$; so by T.1.3 the implication

$$(\exists\, x' \in X : \lambda(F, x') < 1) \Rightarrow F \textit{ attains its minimum over } Q \quad (1)$$

holds. So to prove (i) it suffices to establish that if F is bounded from the below then $\lambda(F,x) < \infty$ and the premise in (1) is true.

The first statement immediately follows from the fact that in the case of $\lambda(F,x) = \infty$ there exists $h \in E$ such that $D^2F(u)[h,h] = 0$ for all $u \in Q$, while $DF(x)[h] < 0$; so on the intersection of the ray $\{x + t\,h \mid t \geq 0\} \cap Q$ F linearly decreases; since X is closed and Q is open, the above ray is contained in Q, and F is not bounded from the below over Q, which contradicts our condition.

So in the case of F bounded from below we have $\lambda(F,x) < \infty$. Consider the process (1.15). By virtue of (ii) and the comment to the theorem the first stage of this process does terminate, so $x_j \in X$ and $\lambda(F,x_j) \leq \lambda' < 1$ for some j; thus, the premise in (1) holds. ■

1.4.9. Proposition 1.3.

Let $Q' = \{\xi \in E^* \mid$ *the function* $F_\xi(x) = F(x) - \langle \xi, x \rangle$ *is bounded from below over* $Q\}$. Let us verify that $Q' = Q^*$. The inclusion $Q^* \subset Q'$ is obvious; let us establish the inverse inclusion. Let $\xi \in Q'$, so $F_\xi(x)$ is bounded from the below over Q. It is obvious that a linear form is α' - self-concordant on E for each $\alpha' > 0$, so $F_\xi \in S^+_\alpha(Q,E)$ (see P.1.1.(ii)). Since this function is bounded from the below over Q, it attains its minimum at a point of this set (T.1.4), which means that $\xi \in \Phi(Q)$.

Since $D^2F(x)$ is non-degenerate for $x \in Q$, Q^* is open, while Q' is obviously a convex set; so Q^* is nonempty, open and convex. By virtue of the standard properties of the Legendre transformation, the C^3-smoothness and the convexity of F together with non-degeneracy of D^2F imply that F^* has the same properties with respect to Q^*.

Let us verify that F^* is self-concordant with the parameter value a. Let us fix $x \in Q$ and notice that for all e, $h \in E$ one has

$$DF(x)[h] = \langle\Phi(x),h\rangle,$$
$$DF^*(\Phi(x))[\phi] = \langle\phi,x\rangle,$$
$$\langle\Phi'(x)h,e\rangle = D^2F(x)[h,e],$$
$$\langle(\Phi'(x)h)'h,h\rangle = D^3F(x)[h,h,h]$$

and

$$F^*(\Phi(x)) = \langle\Phi(x),x\rangle - F(x).$$

Taking the derivatives of these identities in the directions h and e, we have

$$DF^*(\Phi(x))[\Phi'(x)h] = \langle\Phi'(x)h,x\rangle = D^2F(x)[h,x],$$
$$D^2F^*(\Phi(x))[\Phi'(x)h,\Phi'(x)e] = (DF^*(\Phi(x))[\Phi'(x)h])'e -$$
$$- DF^*(\Phi(x))[(\Phi'(x)h)'e] = \langle(\Phi'(x)h)'e,x\rangle + \langle\Phi'(x)h,e\rangle -$$
$$- \langle(\Phi'(x)\ h)'e,x\rangle = \langle\Phi'(x)h,e\rangle = D^2F(x)[h,e],$$
$$D^3F^*(\Phi(x))[\Phi'(x)h,\Phi'(x)h,\Phi'(x)h] =$$
$$= (D^2F^*(\Phi(x))[\Phi'(x)h,\ \Phi'(x)h])'h - 2\ D^2F^*(\Phi(x))[\Phi'(x)h,$$
$$(\Phi'(x)h)'h] = (\langle\Phi'(x)h,h\rangle)'h-2\ D^2F^*(\Phi(x))[\Phi'(x)h,(\Phi'(x)h)'h] =$$
$$= D^3F(x)[h,h,h] - 2\ D^2F^*(\Phi(x))[\Phi'(x)h,(\Phi'(x)h)'h] =$$
$$= D^3F(x)[h,h,h] - 2\ D^2F(x)[(\Phi'(x))^{-1}((\Phi'(x)h)'h],h] =$$
$$= D^3F(x)[h,h,h] - 2\ \langle\Phi'(x)(\Phi'(x))^{-1}((\Phi'(x)h)'h],h\rangle =$$
$$= D^3F(x)[h,h,h] - 2\ \langle(\Phi'(x)h)'h,h\rangle = - D^3F(x)[h,h,h].$$

So for all $x \in Q$ and $h \in E$:

$$|D^3F^*(\Phi(x))[\Phi'(x)h,\Phi'(x)h,\Phi'(x)h]| = |D^3F(x)[h,h,h]| \le$$
$$\le 2\ a^{-1/2}(D^2F(x)[h,h])^{3/2} =$$
$$= 2\ a^{-1/2}(D^2F^*(\Phi(x))[\Phi'(x)h,\Phi'(x)h])^{3/2} \qquad (1)$$

While (x,h) passes through $Q \times E$, $(\Phi(x),\Phi'(x)h)$ passes through $Q^* \times E^*$, so (1) means that $F^* \in S_a(Q^*,E^*)$. It remains to verify that $F^*(\xi_i) \to \infty$ for each sequence $\{\xi_i \in \text{int } Q^*\}$

converging to a point $\xi \in \partial Q^*$. Indeed, assume that $(F^*(\xi_i))$ is bounded from above; then the functions $F_{\xi_i}(x)$ are uniformly in i bounded from below, so the same is true for F_ξ; the latter, by virtue of $Q' = Q^*$, leads to $\xi \in Q^*$, which is impossible (since Q^* is open and $\xi \in \partial Q^*$). Thus, $F^* \in S_a^+(Q^*,E^*)$, which together with the above remarks demonstrates that (Q^*,F^*) is an (a,E^*) - pair. The equality $(Q^*,F^*)^* = (Q,F)$ is an immediate corollary of the above established facts and the standard properties of the Legendre transformation. ∎

Section 2. Self-concordant families

Assume we desire to solve a problem

$f(x) \to \min \mid x \in G \subset E.$

One of the most traditional approaches to the numerical solution puts into correspondence with the problem a parametrized family of problems

$F_t(x) \to \min \mid x \in G_t \subset E,$

such that the trajectory $x^*(t)$ of the minimizers of F_t converges to the solution of the problem as, for example, $t \to \infty$; the trajectory $x^*(t)$ is approximated in an appropriate way along a sequence of parameter's values converging to ∞, which gives approximate solutions. The approximation of the trajectory usually is realized as follows: having produced a good enough approximation, $x(t)$, to $x^*(t)$ for some t, we replace t by a close value t' of the parameter, regard $x(t)$ as an approximation to a new minimizer, $x^*(t')$, and then improve this approximation by some numerical optimization method, producing $x(t')$.

In this section we shall study the above scheme under the assumption that the family considered consists of self-concordant functions and that the improvement of the previous approximation is performed by Newton's method. We desire to find out the conditions on the family which allows for the polynomiality of the method resulted.

2.1. Self-concordant families.

Definition 2.1. Let E be a finite-dimensional real vector space, $\mathcal{F} = \{Q_t, F_t, E\}_{t\in\Delta}$ be a family of functions defined on nonempty open convex subsets $Q_t \subset E$, Δ be an open nonempty interval on the real axis and Q_* be the set $\{(t,x) \in E_* \equiv \mathbb{R} \times R^n \mid t \in \Delta,\ x \in Q_t\}$. Let $\alpha(t)$, $\gamma(t)$, $\mu(t)$, $\xi(t)$, $\eta(t)$ be continuous positive scalar functions defined on Δ, where α, γ, μ are assumed to be continuously differentiable, and let $æ \in (0,\lambda_*)$. The family $\mathcal{F}$ is called *self-concordant* with the parameters α, γ, μ, ξ, η, $æ$, (notation: $\mathcal{F} \in \Sigma_\Delta(\alpha,\gamma,\mu,\xi,\eta,æ)$), if

(i) Q_* is an open subset of E_*; $F_t(x)$ is convex in x, continuous in $(t,x) \in Q_*$ and has three derivatives in x, $D^i F_t(x)$, continuous in $(t,x) \in Q_*$ for $i = 1,2,3$ and continuously differentiable in t for $i = 1,2$;

(ii) $(\forall\ t \in \Delta)$ the function $F_t : Q_t \to \mathbb{R}$ is self-concordant with the parameter value $\alpha(t)$;

(iii) the set $\{(t,x) \in Q_* \mid \lambda(F_t,x) \le æ\}$ is closed in $\Delta \times E$, and there exists some neighborhood (in $\Delta \times E$) of this set, X, such that for each $(t,x) \in X$ and $h \in E$ the following inequalities hold:

$$|\{DF_t(x)[h]\}'_t - \{\ln \mu(t)\}'_t\, DF_t(x)[h]| \le \le \xi(t)\, \alpha^{1/2}(t)\, \{D^2F_t(x)[h,h]\}^{1/2}, \tag{2.1}$$

$$|\{D^2F_t(x)[h,h]\}'_t - \{\ln\ \gamma(t)\}'_t \cdot D^2F_t(x)[h,h]| \le \le 2\,\eta(t)\, D^2F_t(x)[h,h] \tag{2.2}$$

(henceforth, D and $\{\ \}'_t$ mean the derivatives in x and in t, respectively).

The family $\mathcal{F}$ is called *strongly self-concordant* with the parameters α, γ, μ, ξ, η (notation: $\mathcal{F} \in \Sigma^+_\Delta(\alpha,\gamma,\mu,\xi,\eta)$), if it satisfies the conditions (i), (ii) and (iv), where

(iv) inequalities (2.1), (2.2) hold for each $(t,x) \in Q_*$ and $h \in E$, and the set $X(a) = \{(t,x) \in Q_* \mid F_t(x) \le a\}$ is closed in $\Delta \times E$ for each $a \in \mathbb{R}$.

2.2. "Categorial" properties of self-concordant families.

Proposition 2.1. Let $\mathcal{F} = (Q_t, F_t, E)_{t\in\Delta}$ be a family. Then

(i) The following implication holds:

$$\mathcal{F} \in \Sigma^+_\Delta(\alpha,\gamma,\mu,\xi,\eta) \Rightarrow \mathcal{F} \in \Sigma_\Delta(\alpha,\gamma,\mu,\xi,\eta,\varkappa) \quad \forall\, \varkappa \in (0,\lambda_*);$$

(ii) Let $x = \mathcal{A}(y) = Ay + b$ be an affine transformation of a finite-dimensional real vector space E^+ into E, $Q^+_t = \{y \in E^+ \mid Ay + b \in Q_t\}$ and $F^+_t(y) = F_t(Ay + b)\colon Q^+_t \to \mathbb{R}$. Then the following implications hold:

(ii.1) $\mathcal{F} \in \Sigma_\Delta(\alpha,\gamma,\mu,\xi,\eta,\varkappa)$, $\mathcal{A}(E^+) = E \Rightarrow$

$$\mathcal{F}^+ = (Q^+_t, F^+_t, E^+)_{t\in\Delta} \in \Sigma_\Delta(\alpha,\gamma,\mu,\xi,\eta,\varkappa)$$

(ii.2) $\mathcal{F} \in \Sigma^+_\Delta(\alpha,\gamma,\mu,\xi,\eta)$, $(Q_t \neq \emptyset \ \forall\, t \in \Delta) \Rightarrow$

$$\mathcal{F}^+ = (Q^+_t, F^+_t, E^+)_{t\in\Delta} \in \Sigma^+_\Delta(\alpha,\gamma,\mu,\xi,\eta)$$

(iii) Let $\mathcal{F} = (Q_t, F_t, E)_{t\in\Delta} \in \Sigma^+_\Delta(\alpha,\gamma,\mu,\xi,\eta)$, $\mathcal{F}^* = (Q^*_t, F^*_t, E)_{t\in\Delta} \in \Sigma^+_\Delta(\alpha^*,\gamma,\mu,\xi^*,\eta^*)$, p, $p^* > 0$ and let $Q^+_t = Q_t \cap Q^*_t \neq \emptyset$ for each $t \in \Delta$. Let α^+ be a positive continuously differentiable function on Δ, such that

$$\alpha^+(t) \leq \min\{\, p\,\alpha(t),\ p^*\,\alpha^*(t)\}, \quad t \in \Delta,$$

and let

$$\eta^+(t) = \max\{\eta(t), \eta^*(t)\},$$

$$\xi^+(t) = 2(\alpha^+(t))^{-1/2} \max\{(p\alpha(t))^{1/2}\, \xi(t), (p^*\alpha^*(t))^{1/2}\, \xi^*(t)\,\}.$$

Then the family $\mathcal{F}^+ = (Q^+_t, F^+_t = p\,F_t + p^*\,F^*_t, E)_{t\in\Delta}$ belongs to $\Sigma^+_\Delta(\alpha^+,\gamma,\mu,\xi^+,\eta^+)$.

2.3. Metric corresponding to self-concordant family.

Assume that $\mathcal{F} = (Q_t, F_t, E)_{t\in\Delta} \in \Sigma_\Delta(\alpha,\gamma,\mu,\xi,\eta,\varkappa)$. Let

$$\psi(\mathcal{F},t) = (\gamma(t)\ \alpha(t))^{1/2}\ \mu^{-1}(t) \tag{2.3}$$

and introduce metrics on Δ parametrized by $\nu > 0$:

$$\rho_\nu(\mathcal{F};t,\tau) = \max\,\{|\ln(\psi(\mathcal{F},u)/\psi(\mathcal{F},v))| \mid u,v \in [t,\tau]\} +$$

$$+ \nu^{-1}\left|\int_t^\tau \xi(s)ds\right| + \left|\int_t^\tau \eta(s)ds\right|. \qquad (2.4)$$

The following result shows that the property of self-concordance and the metrics corresponding to a family are invariant under rescalings and parameter's replacements.

Proposition 2.2. Let $\mathcal{F} = (Q_t, F_t, E)_{t\in\Delta} \in \Sigma_\Delta(\alpha,\gamma,\mu,\xi,\eta,\varkappa)$, let Δ^+ be an open interval on the real axis, $p(t)$ be a continuously differentiable positive function on Δ and $\pi(\tau)$ be a continuously differentiable one-to-one mapping from Δ^+ onto Δ. Denote $\mathcal{F}^+ = (Q_{\pi(\tau)}, p(\pi(\tau))\, F_{\pi(\tau)}, E)_{\tau\in\Delta^+}$.

Then $\mathcal{F}^+ \in \Sigma_{\Delta^+}(\alpha^+,\gamma^+,\mu^+,\xi^+,\eta^+,\varkappa)$, where

$\alpha^+(\tau) = \alpha(\pi(\tau))p(\pi(\tau))$, $\mu^+(\tau) = \mu(\pi(\tau))p(\pi(\tau))$,

$\gamma^+(\tau) = \gamma(\pi(\tau))p(\pi(\tau))$,

$\xi^+(\tau) = \xi(\pi(\tau))|\pi'(\tau)|$, $\eta^+(\tau) = \eta(\pi(\tau))\,|\pi'(\tau)|$,

and for all $\nu > 0$, τ, $\tau' \in \Delta^+$ one has

$$\rho_\nu(\mathcal{F}^+;\tau,\tau') = \rho_\nu(\mathcal{F};\pi(\tau),\pi(\tau')).$$

2.4. Main result on self-concordant families is as follows.

Theorem 2.1. Let $\mathcal{F} = (Q_t, F_t, E)_{t\in\Delta} \in \Sigma_\Delta(\alpha,\gamma,\mu,\xi,\eta,\varkappa)$. Assume that $(t,x) \in Q_*$ satisfies the inequality $\lambda(F_t,x) < \varkappa$ and that $t' \in \Delta$ is such that

$$\rho_\varkappa(\mathcal{F};t,t') \leq \varkappa^{-1}(\varkappa - \lambda(F_t,x)). \qquad (2.5)$$

Then $(t',x) \in Q_*$ and

$$\lambda(F_{t'},x) \leq \varkappa. \qquad (2.6)$$

Combining this theorem with the above results on Newton's method, we obtain the following

Corollary 2.1. Let $\mathcal{F} = (Q_t, F_t, E)_{t\in\Delta} \in \Sigma_\Delta(\alpha,\gamma,\mu,\xi,\eta,\varkappa)$, let $(t_0, x_{-1}) \in Q_*$ be a point such that

$$\lambda(F_{t_0},x_{-1}) \leq \varkappa, \qquad (2.7)$$

and let $\varkappa^+ = \varkappa^2/(1 - \varkappa)^2$. Assume $(t_i \in \Delta)_{i \geq 0}$ to be such that

$$\rho_{\varkappa}(\mathcal{F}; t_i, t_{i+1}) \leq (\varkappa - \varkappa^+)/\varkappa, \quad i \geq 0. \tag{2.8}$$

Let

$$x_i = x^*(F_{t_i}, x_{i-1}). \tag{2.9}$$

Then x_i are well-defined, belong to Q_{t_i} and

$$\lambda(F_{t_i}, x_i) \leq \varkappa \tag{2.10}$$

for all $i \geq 0$. ■

Thus, being given a sufficiently close approximation, x_{-1}, to the F_{t_0} - *center* of Q_{t_0}, i.e. to the minimizer of F_{t_0}, we can follow the path $x^*(t)$ formed by the minimizers of F_t, using a fixed step-length in the parameter t (the step-length is measured in the metric corresponding to the family).

In the next Section we describe techniques which allows for constructing a spectrum of self-concordant families and the corresponding polynomial-time algorithms.

2.5. Proofs of the results

2.5.1. Proposition 2.1.

(i). It suffices to verify that if $\varkappa \in (0, \lambda_*)$ then $X_*(\varkappa)$ is closed in E_Δ. For $t \in \Delta$ the function $F_t(\ )$, considered as a function of $x \in Q_t$, obviously belongs to $S^+_{\alpha(t)}(Q_t, E)$, so, by T.1.3, one has $t \in \Delta$, $\lambda(F_t, x) \leq \varkappa \Rightarrow F_t(x) - \phi(t) \leq \alpha(t)\ \ g(\varkappa)$, where $\phi(t) = \min\{ F_t(y) \mid y \in Q_t\}$. The function ϕ is upper semicontinuous by virtue of $(\Sigma.1)$ and thus is bounded from above on each compact set $\Delta^+ \subset \Delta$; so $\{(t, x) \in X_*(\varkappa) \mid t \in \Delta^+\}$ $\equiv X(\Delta^+, \varkappa) \subset X^*(a)$ for some $a \in \mathbb{R}$. So there exists a set $Y(\Delta^+, \varkappa)$, which is contained in Q_* and is closed in E_* such that $X(\Delta^+,\varkappa) \subset Y(\Delta^+,\varkappa)$. By $(\Sigma.1)$ the set $X_*(\varkappa)$ is closed in Q_*, thus $X(\Delta^+,\varkappa)$ is closed in Q_*. The latter fact is valid for each compact set Δ^+ which is contained in the interval Δ, so $X_*(\varkappa)$ is closed in E_Δ, Q.E.D.

(ii). Under the conditions of (ii.1), as well as of (ii.2), $\mathcal{F}^+$ obviously satisfies $(\Sigma.1)$ and $(\Sigma.2)$. To verify $(\Sigma.3)$, respectively, $(\Sigma^+.3)$, let us consider the mapping $\pi(t,y) = (t,\mathcal{A}(y))$: $E^+_\Delta \to E_\Delta$: this mapping obviously is continuous. We have $X^*_+(a) = \{ (t,y) \mid F^+_t(y) \leq a \} = \pi^{-1}(\{(t,x) \mid F_t(x) \leq a\})$, so under the assumptions of (ii.2) the sets $X^*_+(a)$ are closed in E^+_Δ for each $a \in \mathbb{R}$. It is clear that if $\mathcal{F}$ satisfies (2.2) and (2.3) for some $(t,x) \in Q_*$, then the corresponding inequalities hold for $\mathcal{F}^+$ for all (t,y) such that $\pi(t,y) = (t,x)$. (ii.2) is proved. To prove (ii.1) by the same arguments it remains to verify that for each $\varkappa$ the equality $X^+_*(\varkappa) = \{(t,y) \in Q^+_* \mid \lambda(F^+_t,y) \leq \varkappa\} = \pi^{-1}(X_*(\varkappa))$ holds. The inclusion of the second set in the first one is obvious; to prove the inverse inclusion, let us notice, that if $(t,y) \in X^+_*(\varkappa)$, then

$$|DF_t(\mathcal{A}(y))[Ah]| \leq \alpha^{1/2}(t)\, \varkappa\, (D^2F_t(\mathcal{A}(y))[Ah,Ah])^{1/2};$$

when h passes through E^+, Ah passes through E (since $\mathcal{A}$ is an onto mapping), so we have $\pi(t,y) \in X_*(\varkappa)$. (ii) is proved.

(iii). All the relations which must be satisfied by $\mathcal{F}^+$, a^+, γ, μ, ξ^+, η^+ by virtue of $(\Sigma.1)$, $(\Sigma.2)$, $(\Sigma^+.3)$, are obviously true, excluding the closedness of the sets

$$X^*_+(a) = \{(t,x) \in E_* \mid t \in \Delta,\ x \in Q^+_t,\ F^+_t(x) \leq a\},\quad a \in \mathbb{R},$$

in E_Λ. Let us verify that $X^*_+(a)$ is closed in E_Δ. Assume that $(t_i, x_i) \in X^*_+(a)$ and $(t_i,x_i) \to (t,x) \in E_\Delta \setminus X^*_+(a)$. Then (t,x) does not belong to one of the sets Q_*, Q^*_* (otherwise (t,x)

belongs to Q_*^+, and F^+ is continuous on this set). If $(t,x) \notin Q_*$, then $F_{t_i}(x_i) \to \infty$ for $i \to \infty$ because of the closedness of the sets $\{(\tau,u) \in Q_* \mid F_\tau(u) \leqslant \text{const}\}$ in E_Δ, and if $(t,x) \in Q_*$, then $\{F_{t_i}(x_i)\}$ is bounded from below by virtue of the continuity of F on Q_*. By the same reasons $\{F^*_{t_i}(x_i)\}$ is either bounded, or tends to $+\infty$. Since one of the sequences $\{F_{t_i}(x_i)\}$, $\{F^*_{t_i}(x_i)\}$ tends to $+\infty$ and both of them are bounded from below, we have $F^+_{t_i}(x_i) \to \infty$, which contradicts the inclusion $(t_i, x_i) \in X^+_*(a)$, $i \geqslant 1$ ■

2.5.2. Theorem 2.1.

1^0. Let $\delta = \{\tau \mid \tau \in \Delta,\ (\tau,x) \in X^+(x)\}$. Then δ is open in Δ and contains t. Let us denote by δ^* the connectedness component of t in δ.

2^0. Let us fix $h \in E$ and consider two functions of $\tau \in \delta^*$:

$$a(\tau) = DF_\tau(x)[h],\quad b(\tau) = D^2F_\tau(x)[h,h] \qquad (1)$$

By $(\Sigma.3)$ we have ($(\ )'$ means the derivative in τ):

$$|a'(\tau) - (\ln(\mu(\tau)))'\, a(\tau)| \leqslant \alpha^{1/2}(\tau)\ \xi(\tau)\ b^{1/2}(\tau), \qquad (2)$$

$$|b'(\tau) - (\ln(\gamma(\tau)))'\, b(\tau)| \leqslant 2\ \eta(\tau)\ b(\tau). \qquad (3)$$

By (3) either (*the case* I_h) $b(\tau) \equiv 0$, $\tau \in \delta^*$, or (*the case* II_h) $b(\tau)$ does not take the zero value over δ^*. In the case I_h, by (2) and by virtue of $|a(t)| \leqslant \lambda(F_t,x)\ \alpha^{1/2}(t)\ b^{1/2}(t) = 0$, we have $a(\tau) \equiv 0$, $\tau \in \delta^*$.

3^0. Now assume that the case II_h takes place. Let

$$\phi(\tau) = (a^2(\tau)\ \alpha^{-1}(\tau)\ b^{-1}(\tau))^{1/2},\ \tau \in \delta^*.$$

Let $t'' \in \delta^*$ be such that

$$\rho_{æ}(\mathcal{F};t,t'') < \rho_{æ}(\mathcal{F};t,t') \tag{4}$$

Denote by t^* the nearest to t'' point of the segment $[t,t'']$, in which ϕ equals zero, if such exists; otherwise let $t^* = t$. Let also δ^+ be the segment with the endpoints t^* and t''. We have

$$\rho_{æ}(\mathcal{F};t^*,t'') < \rho_{æ}(\mathcal{F};t,t');$$

$$\phi(t^*) \leqslant \lambda = \lambda(F_t,x); \quad \phi(\tau) \neq 0, \quad \tau \in (t^*, t''] = \delta_0^+. \tag{5}$$

For $\tau \in \delta^+$ the function $\phi(\tau)$ is continious, and for $\tau \in \delta_0^+$ it is continiously differentiable and differs from zero.

For $\tau \in \delta_0^+$ we have

$$2\phi'(\tau)\phi(\tau) = 2a^2(\tau)(\ln(\mu(\tau)))'(a(\tau)b(\tau))^{-1} - a^2(\tau)(\ln(\gamma(\tau)))'$$

$$(b(\tau)a(\tau))^{-1} - a^2(\tau)(\ln(a(\tau)))'(a(\tau)b(\tau))^{-1} + \omega(\tau),$$

where

$$\omega(\tau) = 2\,(a'(\tau) - (\ln(\mu(\tau)))'a(\tau))\,a(\tau)\,(a(\tau)b(\tau))^{-1} -$$

$$- a^2(\tau)\,(b'(\tau) - (\ln(\gamma(\tau)))'\,b(\tau))\,(a(\tau)b^2(\tau))^{-1}.$$

Since $\tau \in \delta_0^*$, we have $(\tau,x) \in X^+(æ)$, and by (2.2),(2.3) we get

$$|\omega(\tau)| \leqslant 2\,a^{1/2}(\tau)\,\xi(\tau)\,a(\tau)\,b^{1/2}(\tau)\,(a(\tau)b(\tau))^{-1} +$$

$$+ 2\,a^2(\tau)\,\eta(\tau)\,b(\tau)\,(a(\tau)b^2(\tau))^{-1} = 2\,\phi(\tau)\xi(\tau) + 2\,\eta(\tau)\,\phi^2(\tau).$$

Thus, for $\tau \in \delta_0^+$ we have

$$|\phi'(\tau) + (\ln(\phi(\mathcal{F};\tau))_\tau'\,\phi(\tau)| \leqslant \xi(\tau) + \phi(\tau)\eta(\tau). \tag{6}$$

Let $\phi^* = \max\{\,\phi(\tau)\mid t \in \delta^+\}$. By (6)

$$|\phi'(\tau) + (\ln(\phi(\mathcal{F};\tau))_\tau'\,\phi(\tau)| \leqslant \xi(\tau) + \phi^*\,\eta(\tau),\ \tau \in \delta_0^+. \tag{7}$$

Let $\psi_- = \min\{\psi(\tau) \mid \tau \in \delta^+\}$, $\psi_+ = \max\{\psi(\tau) \mid \tau \in \delta^+\}$; then by (7) and the continuity of ϕ on δ^+

$$\tau \in \delta^+ \;\Rightarrow\; \phi(\tau) \leqslant \exp\{\rho_1\}\,(\phi(t^*) + æ\,\rho_2 + \phi^*\,\rho_3), \tag{8}$$

where

$$\rho_1 = \ln(\psi_+/\psi_-),\quad \rho_2 = æ^{-1}\left|\int_{t^*}^{t''}\xi(s)ds\right|,\quad \rho_3 = \left|\int_{t^*}^{t''}\eta(s)ds\right|.$$

So

$$\rho_1 + \rho_2 + \rho_3 = \rho_{æ}(\mathcal{F};t^*,t'') < æ^{-1}(æ - \lambda). \tag{9}$$

By (9) $\rho_1 + \rho_3 < 1$, so $\exp(\rho_1)\,\rho_3 < 1$, which, by (8), leads to

$$\phi^* \leq (1 - \exp(\rho_1)\,\rho_3)^{-1}(\phi(t) + \rho_2).$$

This relation, by virtue of (9) and the second relation in (5), implies $\phi^* \leq æ$.

The latter inequality together with the definition of t^* means that in the case II_h the implication

$$(t''\in\delta^*,\ \rho_{æ}(\mathcal{F};t,t'') < \rho_{æ}(\mathcal{F};t,t')) \Rightarrow \max\{\ \phi(\tau)\ |\ \tau \in [t,t'']\} \leq æ$$

holds. By the continuity arguments this proves the implication

$$(t''\in\delta^*,\ \rho_{æ}(\mathcal{F};t,t'') \leq \rho_{æ}(\mathcal{F};t,t')) \Rightarrow$$
$$\max\{\ \phi(\tau)\ |\ \tau \in [t,t'']\} \leq æ. \qquad (10)$$

Taking into account the definition of ϕ, we obtain from (10) that

$$(t''\in\delta^*,\ \rho_{æ}(\mathcal{F};t,t'') \leq \rho_{æ}(\mathcal{F};t,t')) \Rightarrow$$
$$|D\overset{\circ}{F}_{t''}(x)[h]| \leq \alpha^{1/2}(t'')\ æ\ (D^2F_{t''}(x)[h,h])^{1/2} \qquad (11)$$

The relation (11) has been proved in the case II_h; in the case I_h (where, as we have seen, $DF_{t''}(x)[h] = 0,\ t'' \in \delta^*$) it is obvious. Thus we have

$$(t''\in\delta^*,\ \rho_{æ}(\mathcal{F};t,t'') \leq \rho_{æ}(\mathcal{F};t,t')) \Rightarrow \lambda(F_{t''},x) \leq æ \qquad (12)$$

4^o. To complete the proof it suffices to show that $t' \in \delta^*$ - it will allow us to take $t'' = t'$ in (12). If $t' \notin \delta^*$ then there exist t^+ which lies between t and t' and is a boundary point of the interval δ^*. Assume that t_i lie in δ^* between t and t^+ and tend to t^+ as $i \to \infty$. Each t_i satisfies the premise in (12) (since t_i lies between t and t' and belongs to δ^*); hence by virtue of (12) the inclusions $(t_i,x) \in X_*(æ)$ hold. For $i \to \infty$ the points (t_i,x) converge to (t^+,x). The latter point belongs to $E_*(\Delta)$, since t^+ lies between $t \in \Delta$ and $t' \in \Delta$ and hence itself belongs to Δ. Since $X_*(æ)$ is closed in $E_*(\Delta)$, we have $(t^+,x) \in X_*(æ)$. Hence for all $\tau \in \Delta$ close enough to t^+ the points (τ,x) belong to $X^+(æ)$, so these τ belong to δ; the latter fact contradicts the assumption that t^+ is a boundary point of δ^*. ■

Section 3. Barrier-generated families and barrier method

In this Section we develop a barrier method for the solution of the problem

$$f(x) \to \min \mid x \in G \subset E. \tag{3.1}$$

Barrier methods are path-following methods which correspond to families of the form

$$F_t(x) = t\, f(x) + F(x),$$

where F is some barrier (interior point cost function) for the feasible region G.

Below we implement this scheme using the results on self-concordant families. To ensure the self-concordance of the above families we need some special barriers. So we begin with the definitions and results on the barriers required.

3.1. Self-concordant barriers and barrier-generated families.

Definition 3.1. Let $G \in C(E)$, $\vartheta \geq 1$, $\beta \geq 0$.

(i) A function F: int $G \to \mathbb{R}$ is called a *ϑ-self-concordant barrier* for G (notation: $F \in \mathcal{B}(G,\vartheta)$), if $F \in S_1^+(\text{int } G, E)$ and

$$\lambda(F) \equiv \sup\{\lambda(F,x) \mid x \in \text{int } G\} \leq \vartheta^{1/2}.$$

(ii) A function f: $G \to \mathbb{R} \cup \{+\infty\}$ is called *β-compatible* with $F \in \mathcal{B}(G,\vartheta)$ (notation: $f \in \mathcal{A}(F,\beta)$), if f is lower semicontinuous convex function on G, finite and C^3-smooth on int G and such that for all $x \in$ int G and $h \in E$ the following inequality holds:

$$|D^3f(x)[h,h,h]| \leq \beta\, \{3\, D^2f(x)[h,h]\}\, \{3\, D^2F(x)[h,h]\}^{1/2}. \tag{3.2}$$

The following fact underlies our further developments:

Proposition 3.1. Let $G \in C(E)$, $\vartheta \geq 1$, $\beta \geq 0$, $F \in \mathcal{B}(G,\vartheta)$ and $f \in \mathcal{A}(F,\beta)$. Denote $\Delta = (0,\infty)$ and consider a family

$$\mathcal{F} = \mathcal{F}(F,f) = \{D_t \equiv \text{int } G,\ F_t(x) = t\, f(x) + F(x),\ E\}_{t\in\Delta}.$$

This family $\mathcal{F}$ is strongly self-concordant with the parameters

$$\alpha(t) = (1+\beta)^{-2}, \quad \mu(t) = \gamma(t) = t,$$
$$\xi(t) = \vartheta^{1/2}(1+\beta)/t, \quad \eta(t) = 1/(2t). \tag{3.3}$$

In particular, $\phi(\mathcal{F},t) = (1+\beta)^{-1} t^{-1/2}$ and

$$\rho_\nu(\mathcal{F};t,\tau) = (1 + (1+\beta)\vartheta^{1/2}\nu^{-1})\,|\ln(t/\tau)|. \tag{3.4}$$

3.2. Barriers' properties.

To proceed, let us state some useful properties of self-concordant barriers.

Proposition 3.2. Let $G \in C(E)$, $F \in \mathfrak{s}(G,\vartheta)$. Then

(i) Let $x = \mathcal{A}(y) \equiv Ay + b$ be an affine transformation from a space E^+ into E, such that $\mathcal{A}(E^+) \cap \operatorname{int} G \neq \emptyset$, let $G^+ = \mathcal{A}^{-1}(G)$, $f \in \mathcal{A}(F,\beta)$, $F^+(y) = F(\mathcal{A}(y))$: $\operatorname{int} G^+ \to \mathbb{R}$, $f^+(y) = f(\mathcal{A}(y))$: $\operatorname{int} G^+ \to \mathbb{R}$.

Then $F^+ \in \mathfrak{s}(G^+,\vartheta)$ and $f^+ \in \mathcal{A}(F^+,\beta)$.

(ii) If $f_i \in \mathcal{A}(F,\beta_i)$, $p_i \geq 0$, $i = 1,2$, then $p_1 f_1 + p_2 f_2 \in \mathcal{A}(F, \max(\beta_1,\beta_2))$. If f is a convex quadratic form on E, then $f \in \mathcal{A}(F,0)$. Moreover, $F \in \mathcal{A}(F,1)$ (F is extended to ∂G by the value $+\infty$).

(iii) Let $G_i \in C(E)$, $F_i \in \mathfrak{s}(G_i,\vartheta_i)$, $1 \leq i \leq m$, be such that $G^+ = \bigcap_{i=1}^m G_i \in C(E)$. Let $F^+ = \sum_{i=1}^m F_i$: $\operatorname{int} G^+ \to \mathbb{R}$. Then $F^+ \in \mathfrak{s}(G^+, \sum_{i=1}^m \vartheta_i)$ and $\mathcal{A}(F_i,\beta) \subset \mathcal{A}(F^+,\beta)$ $\forall i$.

(iv) Let $x,y \in \operatorname{int} G$ and let for $w \in G$

$$\pi_w(z) = \inf\{\, t > 0 \mid w + t^{-1}(z-w) \in G \,\}$$

be the Minkovsky function of G with the pole at w. Denote

$$W_r(x) = \{\, z \in E \mid D^2F(x)[z-x,z-x] < r^2 \,\}.$$

Then

(iv.1) $W_1(x) \subset \operatorname{int} G$;

(iv.2) The following inequalities hold

$$DF(x)[x-y] \leq \vartheta\, \pi_y(x)/(1-\pi_y(x)); \tag{3.5}$$
$$DF(x)[y-x] \leq \vartheta \tag{3.6}$$
$$F(x) \leq F(y) + \vartheta \ln(1/(1-\pi_y(x))); \tag{3.7}$$

$$F(x) \geq F(y) + DF(y)[y-x] + \ln(1/(1-\pi_y(x))) - \pi_y(x); \quad (3.8)$$

$$|DF(x)[h]| \leq \vartheta\ (1 - \pi_y(x))^{-1}\ (D^2F(y)[h,h])^{1/2},\ h\in E; \quad (3.9)$$

$$D^2F(x)[h,h] \leq (1 + 3\vartheta)^2\ (1 - \pi_y(x))^{-2}\ D^2F(y)[h,h]. \quad (3.10)$$

Moreover, if $z \in \partial G$ and $\pi_z(x) \leq (\vartheta^{1/2} + 1)^{-2}$, then

$$DF(x)[z-x] \geq 1 - \pi_z(x)(\vartheta^{1/2} + 1)^2. \quad (3.11)$$

(v) $G = G + E_F$ (cf. C.1.1) and F does not vary along the directions parallel to E_F.

F is bounded from below on int G iff the image of G in the factor-space E/E_F is bounded.

If F is bounded from below, it attains its minimum over int G at the set X_F of the form $x(F) + E_F$, and the following inclusions hold:

$$\{\ x \in E \mid D^2F(x(F))[x-x(F),x-x(F)] < 1\ \} \subset \text{int}\ G \subset$$
$$\subset \{x \in E \mid D^2F(x(F))[x-x(F),x-x(F)] \leq (1 + 3\vartheta)^2\ \}. \quad (3.12)$$

(vi) Let $x \in$ int G , $h \in E$ and $q_x(h) = \sup\ \{\ t \mid x \pm t\ h \in G\}$. Then

$$(D^2F(x)[h,h])^{-1/2} \leq q_x(h) \leq (1 + 3\vartheta)\ (D^2F(x)[h,h])^{-1/2}. \quad (3.13)$$

3.3. Barrier method.

Let us fix the objects $G \in C_B(E)$, $F \in \mathfrak{s}(G,\vartheta)$ and $\beta \geq 0$. Our purpose is to describe a method for the solution of (3.1) under the assumption that the objective f is β-compatible with F.

Denote by λ^+ the value of the function $\lambda \to \lambda^2(1 - \lambda)^{-2}$ at the point λ, and let $\zeta(\lambda) = \omega^2(\lambda)(1 + \omega(\lambda))(1 - \omega(\lambda))^{-1}$. If G is bounded, then the form $D^2F(x)[h,e]$ defines a scalar product on E (P.3.2.(v)); this product will be denoted $\langle h,e\rangle_{x,F}$, and the corresponding norm - $\|\ \|_{x,F}$. We omit the subscript x, if $x = x(F)$ is the minimizer of F over int G; notice that this minimizer does exist and is unique, see P.3.2.(v).

The barrier method is defined by the parameters λ_1' , λ_1,

λ_2, λ_3', λ_3, such that

$$0 < \lambda_1^+ \leq \lambda_1' < \lambda_2 < \lambda_3 < \lambda_* ;$$

$$\lambda_1' < \lambda_1 < \lambda_* ; \quad \lambda_3^+ \leq \lambda_3' < \lambda_3 ; \tag{3.14}$$

$$\zeta(\lambda_1') \leq 1/9, \ (1 + \beta)\lambda_2 < \lambda_3 ;$$

$$(1 - \omega(\lambda_3'))^{-2} w^2(\lambda_3') < 1,$$

$$\omega^2(\lambda_2)\ (1 - \omega(\lambda_2))^{-2} \leq 1/9, \tag{3.15}$$

and by a *starting point*

$$w \in \text{int}\ G. \tag{3.16}$$

The method works in two stages, *the preliminary* and *the main* ones.

3.3.1. **The preliminary stage** produces an approximation, u, to $x(F)$ such that $\lambda(F,u) \leq \lambda_2$. To do this, we follow the minimizers trajectory of the family

$$\mathcal{F}^{(1)} = \mathcal{F}(F,g) = \{\text{int}\ G,\ F_t^{(1)}(x) = t\ g(x) + F(x),\ E\}_{t>0},\ t \to 0,$$

where

$$g(x) = -\ DF(w)[x - w]. \tag{3.17}$$

It is clear that $g \in \mathcal{A}(F,O)$, so the family $\mathcal{F}^{(1)}$ is strongly self-concordant (P.2.1); notice that for this family

$$\alpha(t) = 1; \quad \rho_\nu(\mathcal{F}^{(1)};t,t') = (1+\nu^{-1}\vartheta^{1/2})\ |\ln(t/t')|. \tag{3.18}$$

The approximation under consideration is constructed as follows: let

$$t_i = æ_1^{-i},\ i \geq 0,\ æ_1 = \exp\left(\frac{\lambda_1 - \lambda_1'}{\lambda_1(1+\lambda_1^{-1}\vartheta^{1/2})} \right), \tag{3.19}$$

thus

$$\rho_{\lambda_1}(\mathcal{F}^{(1)};t_i,t_{i+1}) \leq \lambda_1^{-1}(\lambda_1 - \lambda_1'),\ \ i \geq 0, \tag{3.20}$$

and let us produce the points x_i:

$$x_{-1} = w;\ \ x_i = x^*(F_{t_i}^{(1)},x_{i-1}),\ \ i \geq 0 \tag{3.21}$$

($x^*(\ \ ,\ \ )$ is defined in Sect. 1.2).

Process (3.21) is interrupted at the first moment i^* when the relation

$$\lambda(F, x_{i-1}) \leq \lambda_2 \quad (3.2$$

holds; the result of the preliminary stage is

$$u = x_{i^*-1} \quad (3.2$$

Proposition 3.3. (i) The preliminary stage is well-defined: ... are well-defined and belong to int G, $-1 \leq i \leq i^*$, $i^* < \infty$, a... the following relations hold

$$\lambda(F^{(1)}_{t_i}, x_{i-1}) \leq \lambda_1, \quad (3.24$$

$$\lambda(F^{(1)}_{t_i}, x_i) \leq \lambda'_1, \quad (3.25$$

(ii) The result of the preliminary stage satisfies th... relations

$$\lambda(F, u) \leq \lambda_2, \quad (3.26$$

$$u \in W_{1/3}(x(F)); \quad (3.27$$

(iii) The number i^* of the preliminary stage iteratio... satisfies the inequality

$$i^* \leq 1 + \frac{\lambda_1 + \vartheta^{1/2}}{\lambda_1 - \lambda'_1}\left(\ln \frac{31}{\lambda_2 - \lambda'_1} + \ln \frac{\vartheta}{1 - \pi_{x(F)}(w)} \right). \quad (3.2$$

3.3.2. **The main stage** minimizes f: at this stage th... minimizers trajectory of the family

$$\mathcal{F}^{(2)} = \mathcal{F}(F, f) = (\text{int } G,\ F^{(2)}_t(x) = t f(x) + F(x),\ E)_{t>0}$$

is approximated along a sequence $t_i \to \infty$. Notice that thi... family is strongly self-concordant with

$$\alpha(t) = (1 + \beta)^{-2},$$

$$\rho_\nu(\mathcal{F}^{(2)}; t, t') = (1 + \nu^{-1}(1+\beta)\vartheta^{1/2})|\ln(t/t')|. \quad (3.29$$

Let

$$t_0 = \frac{\lambda_3(1 + \beta)^{-1} - \lambda(F, u)}{|f'(u)|_{u,F}}, \quad (3.30$$

where f' is the gradient of f with respect to the Euclidea... structure $\langle\ ,\ \rangle_{u,F}$; we assume that $f'(u) \neq 0$ (otherwise u is ...

solution to (3.1)). Let

$$t_i = \varkappa_2^i t_0, \ i \geq 0, \ \varkappa_2 = \exp\left(\frac{\lambda_3 - \lambda_3'}{\lambda_3(1 + \lambda_3^{-1}(1+\beta)\vartheta^{1/2})}\right) \qquad (3.31)$$

thus

$$\rho_{\lambda_3}(\mathcal{F}^{(2)}; t_i, t_{i+1}) = \frac{\lambda_3 - \lambda_3'}{\lambda_3}, \quad i \geq 0, \qquad (3.32)$$

and let us produce the points x_i:

$$x_{-1} = u; \ x_i = x^*(F^{(2)}_{t_i}, x_{i-1}), \ i \geq 0. \qquad (3.33)$$

The points x_i are regarded as the approximate solutions produced by the barrier method.

Proposition 3.4. (i) The main stage is well-defined: x_i, $i \geq -1$, are well-defined and belong to int G, and the following inequalities hold:

$$\lambda(F^{(2)}_{t_i}, x_{i-1}) \leq \lambda_3 \, , \qquad (3.34_i)$$

$$\lambda(F^{(2)}_{t_i}, x_i) \leq \lambda_3' \, . \qquad (3.35_i)$$

(ii) For each $i \geq 0$ we have

$$f(x_i) - f(x^*) \leq \frac{9\,(1 + \beta)}{\lambda_3 - (1 + \beta)\lambda_2}\left(2\vartheta + \frac{\zeta(\lambda_3')}{2(1 + \beta)^2}\right) \cdot$$

$$\exp\left(-\frac{\lambda_3 - \lambda_3'}{\lambda_3 + (1 + \beta)\vartheta^{1/2}}\, i\right) V_F(f). \qquad (3.36)$$

From now on x^* denotes the minimizer of f over G, and

$$V_F(f) = \sup\{ f(x) \mid x \in W_{1/2}(\, x(F)\,)\,\} -$$

$$- \inf\{ f(x) \mid x \in W_{1/2}(\, x(F)\,)\,\}. \qquad (3.37)$$

3.3.3. Below we use the following statement, which summarizes the results of P.3.3 and P.3.4:

Theorem 3.1. Let $G \in C_B(E)$, $F \in \mathcal{S}(G, \vartheta)$, $f \in \mathcal{A}(F, 0)$ (i.e. f is a quadratic form), $w \in$ int G and let λ_1, λ_1', λ_2, λ_3, λ_3' satisfy (3.14), (3.15) for $\beta = 0$. Consider the application of the barrier method to problem (3.1) generated by f (the method is defined by the parameters $\beta = 0$, λ_1, λ_1', λ_2, λ_3, λ_3' and by the starting point w). Then for each $\varepsilon \in (0,1)$ the total number of the preliminary and the main stage iterations, $N(\varepsilon)$, which is required to produce an approximate solution, $x_\varepsilon \in$ int G, such that

$$f(x_\varepsilon) - \min_G f \le \varepsilon V_F(f),$$

satisfies the inequality

$$N(\varepsilon) \le O(\vartheta^{1/2} \ln(\frac{2\vartheta}{\varepsilon(1 - \pi_{x(F)}(w))})) \qquad (3.38)$$

(the constant factors in $O(\ )$ depend on λ_1, λ_1', λ_2, λ_3, λ_3' only).

Each of the above iterations can be reduced to a step of Newton's method as applied to a convex combination of f and F (or of F and a linear form).

Good (approximately optimal for large ϑ) choice of the parameters λ_1, λ_1', λ_2, λ_3, λ_3' in the case of $\beta = 0$ is:

$$\lambda_1 = \lambda_3 = 0.193; \quad \lambda_1' = \lambda_3' = \lambda_1^+ \cong 0.057; \quad \lambda_2 = 0.150.$$

Under this choice of the parameters and for large enough ϑ the principal term of the asymptotics (for $\varepsilon \to 0$) of the right hand side of (3.38) is $c\,\vartheta^{1/2} \ln(1/\varepsilon)$, where

$$c \cong 7.36.$$

3.4. Examples of barriers.

The main question arising now is how to obtain a self-concordant barrier for a given convex set. In what follows we describe some techniques which enables one to find such barriers.

Let us start with the following useful statement.

Proposition 3.5. (i) Let G be a closed convex set in R^n with int $G \ne \emptyset$, and let $\mathcal{A}(y): R^k \to R^n$ be an affine transformation such that $\mathcal{A}(R^k)$ intersects int G. If F is a ϑ-self-concordant barrier for G, then the function $F(\mathcal{A}(y))$ is ϑ-self-concordant barrier for $\mathcal{A}^{-1}(G)$.

(ii) Let G_i be closed convex sets in R^n and F_i be ϑ_i-self-concordant barriers for G_i, $1 \le i \le k$. Assume that the set $G = \bigcap_{i=1}^{k} G_i$ has a nonempty interior. Then the function $F = \sum_{i=1}^{k} F_i$ is $\sum_{i=1}^{k} \vartheta_i$ -self-concordant barrier for G.

The following theorem gives a spectrum of concrete self-concordant barriers.

Theorem 3.2. For appropriately chosen absolute constants taken as the constant factors in the below $O(\ )$, the following statements are true:

(i) *(Barriers for the intersection of regions bounded by first and second order surfaces)*

If the function $\Phi: R^n \to \mathbb{R}$ is a convex quadratic form such that the region $G' = \{x \in R^n \mid \Phi(x) < 0\}$ is nonempty, then the function $\ln(1/(-\Phi(x)))$ is a 1-self-concordant barrier for the set $G = \{x \in R^n \mid \Phi(x) \le 0\}$. Consequently, any set with a nonempty interior, which is an intersection of m convex sets bounded each by certain first or second order surface (for example, a convex polytope with m facets) admits a m-self-concordant barrier.

Notice that n-facet convex cone in R^n (as well as the intersection of such a cone with any convex set containing the vertex of the cone in its interior) admits no ϑ-self-concordant barrier with $\vartheta < n$.

(ii) *(Barriers for the epigraphs of functions of the Euclidean norm)*

A. The function $\ln(1/(t^2 - x^T x))$ is an 2-self-concordant barrier for the set $G = \{(t,x) \in \mathbb{R} \times R^n \mid t \ge |x|_2\}$.

B. Let $\zeta(t)$ be a nondecreasing continuous and concave function on $[0,\infty)$, C^3-smooth on $(0,\infty)$, satisfying $\zeta(0) = 0 < \zeta(t)$, $t > 0$, and such that one of the quantities

$\alpha_\zeta^{(1)} = \min\{ \alpha \ge 0 \mid |\zeta'''(t)|\ \zeta(t) \le \alpha\ \zeta'(t)\ |\zeta''(t)| \quad \forall\ t > 0\}$,

$$a_\zeta^{(2)} = \min\{ \alpha > 0 \mid |\zeta'''(t)|\, \zeta^{1/2}(t) \le \alpha\, |\zeta''(t)|^{3/2} \ \forall\, t > 0\},$$

$$a_\zeta^{(3)} = \min\{ \alpha > 0 \mid |\zeta'''(t)| \le \alpha\, |\zeta''(t)|/t \ \forall\, t > 0\}$$

is finite. Let

$$a_\zeta = \min\{a_\zeta^{(1)}, a_\zeta^{(2)}, a_\zeta^{(3)}\} + 1.$$

Then the function

$$O(a_\zeta^2)\ \ln(1/(\zeta^2(t) - x^T x)) - \ln t$$

is a $O(a_\zeta^2)$-self-concordant barrier for the set

$$G = \{(t,x) \in \mathbb{R} \times R^n \mid t \ge 0,\ \zeta(t) \ge (x^T x)^{1/2}\}.$$

Thus, the sets

$$G_n^p = \{(t,x) \in \mathbb{R} \times R^n \mid t \ge (|x|_2)^p\},\ 1 \le p < \infty,$$

admit $O(1)$-self-concordant barriers, $O(1)$ does not depend on p.

(iii) *(Barriers for the epigraphs of functions of one variable)*

A. Let $\zeta(t)$ be a nondecreasing C^3-smooth concave function on $(0,\infty)$ such that the quantity

$$a_\zeta = \min\{ \alpha > 0 \mid |\zeta'''(t)| \le \alpha\, |\zeta''(t)|/t \ \forall\, t > 0\} + 1$$

is finite. Then the function

$$O(a_\zeta^2)\ (\ln(1/t) + \ln(1/(\zeta(t) - x)))$$

is an $O(a_\zeta^2)$-self-concordant barrier for the set

$$G = \mathrm{Cl}\ \{(t,x) \in R^2 \mid t > 0,\ \zeta(t) \ge x\}.$$

Thus, the sets

$$G^p = \{(t,x) \in R^2 \mid t \ge (x_+)^p\},\ 1 \le p < \infty,$$

admit $O(1)$-self-concordant barriers, $O(1)$ does not depend on p.

B. Let $f(x) \ne \mathrm{const}$ be a C^3 function on $\mathbb{R}$, such that f', f'', $f''' \ge 0$. Assume that

$$f'(x)\ f'''(x) \in [\tfrac{1}{2}(f''(x))^2,\ \lambda\,(f''(x))^2]$$

for some $\lambda \in [3/2,2)$ and for all x. Let $\Delta = \{x \mid f'(x) > 0\}$ and let $\zeta(x): f(\Delta) \to \Delta$ be the inverse to $f:\Delta \to f(\Delta)$. Then the function

$$O((2-\lambda)^{-2})\ (\ln(1/(t - f(x)) + \ln(1/(\zeta(t) - x)))$$

is an $O((2-\lambda)^{-2})$ -self-concordant barrier for the set

$$G = \{(t,x) \in R^2 \mid t \geq f(x)\}.$$

Thus, the epigraph $\{(t,x) \in R^2 \mid t \geq \exp(x)\}$ of e^x admits an $O(1)$ -self-concordant barrier.

(iv) *(A barrier for the cone of symmetric positively semidefinite matrices)*

Let S_n be the space of symmetric $n \times n$ -matrices with real entries and let S_n^+ be the cone of positively semidefinite matrices from S_n. The function

$$\ln(1/\mathrm{Det}(x))$$

is a n-self-concordant barrier for S_n^+.

(v) *(A barrier for the epigraph of the matrix norm)*

Let $L_{m,n}$ be the space of $m \times n$ - matrices with real entries. The function

$$O(1)\ \ln(\ 1/\mathrm{Det}(t^2\ I_n - x^T x))$$

is an $O(n)$-, and the function $O(1)$

$$\ln(1/\mathrm{Det}(t^2\ I_m - x\,x^T))$$

is an $O(m)$ - self- concordant barrier for the set

$$G = \{(t,x) \in \mathbb{R} \times L_{m,n} \mid t \geq \|x\|\}.$$

Herein I_k means the $k \times k$ unit matrix and $\|\ \|$ is the standard matrix norm (the spectral radius of $(A\,A^T)^{1/2}$).

(vi) *(A barrier for the epigraph of "fractional-quadratic" function)*

Let S_n be the space of $n \times n$ - symmetric matrices. Then the function

$$F(t,X,x) = -\ O(1)\ (\ln \mathrm{Det}\ X + \ln(t - x^T\ X^{-1}\ x))$$

is an $O(n)$ - self-concordant barrier for the set

$$G = \mathrm{Cl}\ \{(t,X,x) \in \mathbb{R} \times S_n \times R^n \mid X \text{ is positive definite}, \\ t > x^T X^{-1} x\}.$$

3.5. Coverings and barriers calculus.

So far we have considered the barrier method under the assumption that the objective f in (3.1) is quadratic (or β-compatible with the barrier for the feasible region G). Of course, this is not a severe restriction. Indeed, replacing G by the epigraph of $f|_G$, one can reduce (3.1) to a problem of the same type with a linear f. We see that an appropriate choice of extra variables may simplify the situation. This idea can be implemented as follows.

Definition 3.2. 1) Let $G \in C(E)$ and let $\Gamma = (E',G',\pi, F)$ be a collection consisting of:

a finite-dimensional real vector space E', $\dim E' = \dim E + l$;

a set $G' \in C(E')$;

an affine transformation $\pi: E' \to E$, such that $\pi(G') = G$ and each compact $K \subset G$ is π-image of some compact $K' \subset G'$;

a ϑ-self-concordant barrier F for G'.

In this situation we call Γ a *(ϑ,l)-covering* for G, and G itself is called *(ϑ,l)-regular*.

2) Let $G \in C(E)$ and let $\phi: G \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous on G and finite on int G convex function. The pair (G,ϕ) is called a *functional element* (f.e.) on E. A (ϑ,l)-covering, Γ, for the *epigraph* $\mathfrak{G}(G,\phi) = \{(t,x) \in \mathbb{R} \times E \equiv E_+ \mid x \in G,\ t \geq f(x)\}$ of the f.e. (G,ϕ) is called a *(ϑ,l)-covering for the f.e. (G,ϕ)*. A f.e. (G,ϕ) is called *(ϑ,l)-regular*, if it admits a (ϑ,l)-covering.

We do not distinguish between a continuous convex function $\phi: E \to \mathbb{R}$ and the f.e. (E,ϕ); thus, we can speak about (ϑ,l)-regular functions.

Our purposes now are as follows. In this subsection we describe some kind of calculus for regular sets and functional elements. In the next subsection we show that a convex programming problem with regular components can be reduced to a problem of the form (3.1) with *linear* f and G possessing a self-concordant barrier; the latter problem (and hence - the original one) can be solved by the above barrier method.

We start with some calculus of coverings.

Theorem 3.3. (i) Let $\Gamma = (E',G',\pi,F)$ be a (ϑ,l) covering for $G \in C(E)$ and let $\sigma : E \to E_1$ be an affine transformation such that $G_1 \equiv \sigma(G) \in C(E_1)$ and such that each compact contained in G_1 is σ-image of some compact contained in G. Then Γ induces a $(\vartheta, l + (\dim E - \dim E_1))$-covering Γ_1 for G_1.

(ii) Let $\Gamma = (E',G',\pi,F)$ be a (ϑ,l) - covering for $G \in C(E)$ and let $\sigma : E_1 \to E$ be an affine transformation such that $\sigma(E_1) \cap \operatorname{int} G \neq \emptyset$. Let $G_1 = \sigma^{-1}(G)$; then $G_1 \in C(E_1)$, and Γ induces a (ϑ,l)-covering Γ_1 for G_1.

(iii) Let $\Gamma_i = (E'_i,G'_i,\pi_i,F_i)$ be (ϑ_i,l_i)-coverings for $G_i \in C(E)$, $1 \leq i \leq k$, and let $G \equiv \bigcap_{i=1}^{k} G_i \in C(E)$. Then the coverings Γ_i induce a $(\sum_{i=1}^{k} \vartheta_i, \sum_{i=1}^{k} l_i)$-covering Γ for G.

The above reductions are "explicit" - i.e. they are straightforward and require only the application of "rational" linear algebra techniques to the initial coverings.

Now we state the following superposition theorem for regular functional elements:

Theorem 3.4. Let (G_i,ϕ_i) be (ϑ_i,l_i)-regular functional elements on E, $1 \leq i \leq k$, and let (G,ϕ) be a (ϑ,l)-regular functional element on R^k. Assume that

the set $H = \bigcap_{i=1}^{k} G_i$ has a nonempty interior;

each of the functions ϕ_i is bounded on bounded subsets of G_i, $1 \leq i \leq k$, and the function ϕ is bounded on bounded subsets of G;

the image of H under the mapping $f = (\phi_1,\dots,\phi_k)$ is contained in G, thus the function $g(x) = (\phi \circ f)(x) : H \to \mathbb{R}$ is well-defined.

Moreover, assume that for each $x \in H$ the set $f(x) + (R^k)_+$ is contained in G, and on this set $\phi(u) \geq \phi(f(x))$ (herein $((R^k)_+$ is the nonnegative ortant in R^k).

Then (H,g) is a $(\sum_{i=1}^{k} \vartheta_i + \vartheta, \sum_{i=1}^{k} l_i + l + k)$-regular functional element. The covering for this element is induced in explicit form by the coverings of the initial f.e.

The following corollary of the above theorem is more convenient:

Corollary 3.1. Let $f = (f_1,\dots,f_k) : E \to R^k$ be a vector-function which components f_i are (ϑ_i, l_i) regular, and let $\phi: R^k \to \mathbb{R}$ be a monotone (with respect to the usual partial ordering on R^k) and (ϑ, l)-regular function. Then the superposition $g(x) = \phi(f(x)) : E \to \mathbb{R}$ is $(\sum_{i=1}^{k} \vartheta_i + \vartheta, \sum_{i=1}^{k} l_i + l + k)$-regular, and the corresponding covering for this superposition is induced in explicit form by the initial coverings.

The above statement holds, if ϕ is monotone on $(R^k)_+$ only and f is nonnegative on E.

The following simple statement is also useful:

Proposition 3.6. Let (G,ϕ) be a (ϑ,l)-regular f.e. on E and let the set $H' = \{ x \in \text{int } G \mid \phi(x) < 0 \}$ be nonempty. Then the set $H = \text{Cl } H'$ is $(\vartheta + 1, l + 1)$-regular, and the corresponding covering is induced in explicit form by the covering for the initial f.e.

We summarize the above results on regularity in the following statements.

A. Composition rules.

(i) The product of a (ϑ,l) - regular function by a positive constant, the sum of such a function and an affine form and the superposition of the function with an affine transformation of the argument are (ϑ,l) - regular;

(ii) If functions $f_i : R^n \to \mathbb{R}$ are (ϑ_i, l_i) - regular, $1 \leq i \leq k$, then the maximum of these functions over i is $(\sum_{i=1}^{k} \vartheta_i, \sum_{i=1}^{k} l_i)$ - regular, and their sum is $(1 + \sum_{i=1}^{k} \vartheta_i, k + \sum_{i=1}^{k} l_i)$ - regular;

(iii) The superposition of a (ϑ,l) - regular and monotone on R^k (or on the nonnegative ortant in R^k) function ϕ and a k-dimensional (k-dimensional nonnegative, respectively) vector- function f possessing (ϑ_i,l_i) -regular components, $1 \leq i \leq k$, is $(\vartheta + \sum_{i=1}^{k} \vartheta_i,\ k + l + \sum_{i=1}^{k} l_i)$ - regular.

B. The regularity of certain functions of one variable. *For appropriately chosen absolute constants taken as the constant factors in the below O( ):*

(i) the function $f(x) \equiv x$ is $(1,0)$-, and the function $f(x) \equiv (x)_+$ is $(2,0)$ - regular;

(ii) the functions $|x|^p$, $(x_+)^p$, $1 \leq p < \infty$, are $(O(1),0)$ - regular (where $O(1)$ does not depend on p);

(iii) the function $\exp(x)$ is $(O(1),0)$ - regular.

C. The regularity of certain functions of many variables. *For appropriately chosen absolute constants taken as the constant factors in the below O( ):*

(i) a convex quadratic form on R^n is $(1,0)$ - regular;

(ii) the function $|x|_2$ is $(2,0)$ - regular;

(iii) the matrix norm $|x|$ on the space of $m \times n$ - matrices is $(O(\min(m,n)),0)$ - regular;

(iv) the functions $(|x|_p)^p : R^n \to \mathbb{R}$, $1 \leq p < \infty$, are $(O(n),n)$ - regular; the functions $(|x|_p)^{p/2}: R^n \to \mathbb{R}$, $2 \leq p < \infty$, are $(O(n),n)$ - regular (the constant factor in $O(\ )$ does not depend on p).

3.6. Barrier method for problems with regular components.

Consider the convex programming problem

$$f_0(x) \to \min \mid x \in G_{m+1} \subset R^n,\ f_i(x) \leq 0,\ 1 \leq i \leq m. \quad (3.39)$$

3.6.1. Assume that the objects involved into (3.39) are as follows:

the set G belongs to $C(E)$, and a (ϑ,l)-covering, $\Gamma = (E', G', \pi, F)$, for this set is given;

the functions f_i are represented by (ϑ_i, l_i)-regular f.e. (G_i, f_i), such that $G \subset G_i$, and (ϑ_i, l_i)-coverings, $\Gamma_i = (E_i', G_i', \pi_i, F_i)$, for these elements are given;

the Slater condition holds: the set $H' = \{x \in \text{int } G \mid f_i(x) < 0,\ 1 \leqslant i \leqslant m\}$ is nonempty;

the feasible region $H = \{x \in G \mid f_i(x) \leqslant 0,\ 1 \leqslant i \leqslant m\}$ of the problem is bounded.

Under these assumptions problem (3.39) can be solved as follows. Let $E^+ = \mathbb{R} \times E$, $G^+ = \mathbb{R} \times G$, $G_i^+ = \mathbb{R} \times G_i$, and let $\psi_i(t,x): G_i^+ \to \mathbb{R}$ be defined as $f_i(x)$ for $i > 0$ and as $f_0(x) - t$ for $i = 0$. The set G^+ and the epigraphs of the f.e. (G_i^+, ψ_i), $1 \leqslant i \leqslant m$, are the inverse images of G and the epigraphs of the f.e. (G_i, f_i), respectively, under appropriate linear epimorphisms. By T.3.3.(ii) the coverings Γ, Γ_i, $i \geqslant 1$, induce coverings Γ^+, Γ_i^+ for G^+ and the f.e. (G_i^+, f_i). The epigraph

$$\mathfrak{G}(G_0^+, \psi_0) = \{(\tau, t, x) \in \mathbb{R}\times(\mathbb{R} \times E) \mid (t,x) \in G_0^+,\ \tau \geqslant \psi_0(x)\}$$

of the f.e. (G_0^+, ψ_0) is the inverse image of $\mathfrak{G}(G_0, f_0)$ under the linear epimorphism $(\tau, t, x) \to (\tau + t,\ x)$, thus by the same theorem Γ_0 induces a covering Γ_0^+ for the f.e. (G_0^+, ψ_0). Notice that the parameters of the initial coverings coincide with these ones for the induced coverings.

The coverings Γ_i^+, $0 \leqslant i \leqslant m$, induce a $(\vartheta^+ = \sum_{i=0}^{m} \vartheta_i,\ l^+ = \sum_{i=0}^{m} l_i)$-covering Γ^{++} for the f.e. $(Q = \bigcap_{i=0}^{m} G_i^+,\ \psi = \max_{0\leqslant i\leqslant m} \psi_i)$ (T.3.3.(iii); this theorem is appliable, because $\bigcap_{i=0}^{m} \text{int } G_i^+$ under the above assumptions about the problem is obviously nonempty). Since (3.39) satisfies the Slater condition, the covering Γ^{++}, by P.3.6, induces a $(1 + \vartheta^+,\ 1 + l^+)$-covering, Γ', for the set $Q' = \{(t,x) \in Q \mid \psi(t, x) \leqslant 0\} = \text{Cl } \{(t,x) \in Q \mid \psi(t,x) < 0\}$ (the latter equality holds because ψ is lower semicontinuous on Q). The Slater condition also implies that $\text{int } G^+ \cap \text{int } Q' \neq \emptyset$, so Γ' and Γ^+ induce a $(\vartheta^* = 1 + \vartheta + \vartheta^+, l^* = 1 + l + l^+)$-covering $\Gamma^* = (E^*, G^*, \pi^*, F^*)$ for the set

$$G^\# = G^+ \cap Q' =$$
$$\{(t,x) \in E^+ = \mathbb{R} \times E \mid x \in G,\ f_i(x) \leqslant 0,\ 1 \leqslant i \leqslant m,\ t \geqslant f_0(x)\}.$$

It is clear, that the problem

$$\phi(u) \equiv t(\pi^*(u)) \to \min \mid u \in G^* \qquad (3.40)$$

(where $t(z) = t$ for $z = (t,x) \in E'$) is equivalent to the problem

$$t \to \min \mid (t,x) \in G^{\#},$$

and the latter is equivalent to (3.39). Now notice, that problem (3.40) "almost satisfies" the conditions under which it can be solved by the barrier method: this problem is of the form (3.1), the objective is linear, and F^* is a ϑ^*-self-concordant barrier for G^*. Notice that this barrier is induced in explicit form by the initial barriers. The only obstacle for application of the barrier method to (3.40) is the possibility for G^* to be unbounded. This obstacle can be removed as follows.

The feasible set H was assumed to be bounded; assume that we are given some constants $t_* < t^*$ such that $f_0(x) \in (t_*,t^*)$ for $x \in H$. Let $G^{\#\#} = \{(t,x) \in G^{\#} \mid t_* \leq t \leq t^*\}$; then $G^{\#\#}$ is a bounded subset of $G^{\#}$; since Γ^* is a covering for $G^{\#}$, then $G^{\#\#}$ is contained in $\pi^*(G^{**})$ for certain *bounded* $G^{**} \subset G^*$. Without loss of generality we can assume that

$$G^{**} = \{u \in G^* \mid \ |u|_2 \leq R\}$$

for an appropriate R and that int $G^{**} \neq \emptyset$. Now let

$$G_R^* = \{u \in G^* \mid |u|_2 \leq R,\ t(\pi^*(u)) \in [t_*,t^*]\}.$$

Obviously, $G_R^* \in C(E^*)$ and $\pi^*(G_R^*) = G^{\#\#}$. Moreover, G_R^* is the part of G^* singled out by one quadratic and two linear constraints, thus F^* induces a $(\vartheta^* + 3,\ l^*)$-self- concordant barrier, F_R^*, for G_R^*. By the above arguments, $\Gamma_R^* := (E^*,G_R^*,\pi^*,F_R^*)$ is a $(\vartheta^* + 3,\ l^*)$-covering for $G^{\#\#}$, so the problem

$$t(\pi^*(u)) \to \min \mid u \in G_R^* \qquad (3.41)$$

is equivalent to the problem $t \to \min \mid (t,x) \in G^{\#\#}$, and the latter is equivalent to (3.40) by definition of t_*, t^*. Problem (3.41) can be solved by the barrier method, because G_R^* is bounded.

3.6.2. Under some more restrictions on the objects involved into (3.39) this problem can be solved by the barrier method in slightly different manner. Let us assume that G is bounded, the problem is consistent and that we are given the following data:

a ϑ-self-concordant barrier F for G;

a point $z \in \operatorname{int} G$ and $\sigma \geqslant 1$, such that $G \subset z + \sigma\ ((G - z) \cap (z - G))$ (it means that z is a "symmetry center of G within the factor σ");

a constant V, such that $|f_0(x)| \leqslant V$, $f_i(x) \leqslant V$, $1 \leqslant i \leqslant m$, for all $x \in G$.

Suppose that, being given $\varepsilon \in (0,V)$, we desire to find an *ε-solution* to (3.39), i.e. a point $x_\varepsilon \in G$, such that

$$f_0(x_\varepsilon) \leqslant f_0(x^*) + \varepsilon,\ f_i(x_\varepsilon) \leqslant \varepsilon,\ 1 \leqslant i \leqslant m,$$

where x^* is a solution of (3.39). Obviously, we can restrict ourselves to the case of $\varepsilon < V$.

Let

$$\Omega(\varepsilon) = 4\ V/\varepsilon; \quad \delta(\varepsilon) = \varepsilon^2/(4\ V^2), \tag{3.42}$$

and

$$G^* = \{(t,x) \in \mathbb{R} \times E \mid x \in G,\ t \leqslant 3\ V\ \Omega(\varepsilon),\ t \geqslant f_0(x) + V,\ t \geqslant \Omega(\varepsilon)\ f_i(x),\ 1 \leqslant i \leqslant m\}.$$

By definition of V the point $w = (3V\Omega(\varepsilon)/2,\ z)$ obviously belongs to the interior of the convex compact G^*, hence $G^* \in C_B(\mathbb{R} \times E)$. Moreover, let G' be the intersection of G with the image of G under the symmetry with the center at z. Then the convex set

$$Q = \{(t,x) \mid x \in G',\ |t - 3V\Omega(\varepsilon)/2| \leqslant V\Omega(\varepsilon)/2\}$$

is symmetric with respect to w and is contained in G^*, while the image of Q under the enlargement with the center at w and the ratio $\sigma_* = \max\{\sigma, 3\}$ contains G^*. So

$$1/(1-\pi_v(w)) \leqslant \sigma_* \tag{3.43}$$

for each $v \in G^*$.

It is clear that the function

$$F^*(t,x) = F(x) + F_0(t - V,x) + \sum_{i=1}^{m} F_i(t/\Omega(\varepsilon),x) + \ln(1/(3\Omega(\varepsilon)V - t))$$

is a ϑ_*-self-concordant barrier for G^*, where

$$\vartheta_* = 1 + \vartheta + \sum_{i=0}^{m} \vartheta_i .$$

Now consider the problem

$$t \to \min \mid (t,x) \in G^* . \qquad (3.44)$$

Let us solve it by the barrier method, generated by the barrier F^* and the starting point w (the method corresponds to $\beta = 0$; the parameters λ_1, λ_1', λ_2, λ_3, λ_3' are assumed to be some fixed absolute constants satisfying (3.14), (3.15)). By (3.43) and T.3.1 it is clear, that after

$$N(\varepsilon) = O(\vartheta_*^{1/2} \ln(\vartheta_*^2 \sigma_*/\delta(\varepsilon))) = O(\vartheta_*^{1/2} \ln(\vartheta_* \sigma_*^{1/2} V/\varepsilon)) \qquad (3.45)$$

iterations of the preliminary and the main stages an approximate solution, $(t',x') \in \operatorname{int} G^*$, to the problem (3.44), will be produced, such that

$$t' - \min\{t \mid (t,x) \in G^*\} \leqslant \delta(\varepsilon)\,(3\,\Omega(\varepsilon)\,V) = \varepsilon \qquad (3.46)$$

(notice that $f_0(x) + V \geqslant 0$ and hence $0 \leqslant t \leqslant 3\ \Omega(\varepsilon)\ V$ for $(t,x) \in G^*$).

Let us verify that x' is an ε-solution to problem (3.39). Indeed, let

$$\phi(x) = \max(f_0(x) + V,\ \Omega(\varepsilon)\ f_1(x), \ldots, \Omega(\varepsilon)\ f_m(x)),$$

and let ϕ^* be the minimum value of ϕ over G. Then, by definition of G^* and by virtue of (3.46), we have

$$\phi(x') < t' \leqslant \phi^* + \varepsilon .$$

Moreover, $\phi^* \leqslant \phi(x^*) = f_0(x^*) + V$ (notice that $f_0(x^*) + V \geqslant 0 \geqslant f_i(x^*)$, $1 \leqslant i \leqslant m$). We get

$$f_0(x') + V \leqslant \phi(x') \leqslant \phi^* + \varepsilon \leqslant f_0(x) + V + \varepsilon,$$

or $f_0(x') \leqslant f_0(x^*) + \varepsilon$, and, by the same arguments,

$$\Omega(\varepsilon)\ f_i(x') \leqslant f_0(x) + V + \varepsilon \leqslant 2\,V + \varepsilon \leqslant 3\,V$$

for $1 \leqslant i \leqslant m$. This leads to

$$f_i(x') \leqslant 3\ V/\Omega(\varepsilon) \leqslant \varepsilon$$

for $1 \leqslant i \leqslant m$. So x' is the desired ε-solution and its generation requires no more than

$$N(\varepsilon) \leqslant O((m + k)^{1/2}\ (m + n)\ n^2)\ \ln(2(m + k)V/\varepsilon))$$

iterations of the barrier method as applied to (3.44).

3.7. Application examples.

Let us describe the application of the barrier method to a spectrum of convex programming problems of the form (3.39) (the method is applied in the manner described in sect. 3.6.2). In each of the below examples we give expressions for two efficiency estimates: $N(\varepsilon)$, the upper bound for the number of iterations required by the above described barrier method to obtain an ε-solution, and $M(\varepsilon)$, the upper bound for the total number of the arithmetic operations performed at these iterations.

The constant factors in the below $O(\ )$ are absolute constants.

For simplicity sake, G in the examples **A-D** is assumed to be an Euclidean ball of the radius R centered at O.

A. Linear and quadratic programming. Assume that f_i, $0 \leq i \leq m$, are convex quadratic forms (possibly, degenerate or linear). The problem can be reduced to the form (3.44) with

$$F^*(t,x) = -\sum_{i=1}^{m} \ln(t - \Omega(\varepsilon)f_i(x)) - \ln(t - V - f_0(x)) - \ln(R^2 - \|x\|_2^2) - \ln(3\Omega(\varepsilon)V - t),$$

$\vartheta_* = m + 3,$

which implies (we assume that $m > 0$)

$$N(\varepsilon) \leq O(m^{1/2} \ln(2mV/\varepsilon)),$$

$$M(\varepsilon) \leq O(m^{1/2}(m\, n^2 + n^3) \ln(2mV/\varepsilon)).$$

The barrier method of the above type was independently developed in [Go. 1987] for LP and in [Ne. 1988 2,3] - for LP and linearly constrained QP.

Notice, that in the case of linear f_i, $1 \leq i \leq m$, and quadratic f_0 the total number of operations can be reduced in order (see Sect. 6 below).

Notice also, that the Mehrotra and Sun method for quadratically constrained quadratic programming [MS. 1988] has $O(m^{3/2})$ times worse efficiency than the above barrier method (given a good initial point, their method converges at the same rate as the latter one, but the initialization scheme of [MS. 1988] is worse than our preliminary stage).

B. Geometrical programming (in the exponential form). Assume that

$$f_i(x) = \sum_{j=1}^{r_i} c_{i,j} \exp\{ a_{i,j}^T x \} + d_i,\ c_{i,j} \geq 0,\ 0 \leq i \leq m.$$

Let K be the number of different elements, $a_1, \dots, a_K$, in the array $\{a_{i,j} \mid 0 \le i \le m,\ 1 \le j \le r_i\}$, and let s_p, $1 \le p \le K$ be the largest (over the constraints and the objective) among the coefficients $c_{\cdot\cdot\cdot}$ at the term $\exp\{a_p^T x\}$.

Let us introduce the extra variables vector $\tau = (\tau_1, \dots, \tau_K)^T$ and a function $l(i,j)$ taking values in $\overline{1,K}$, such that $l(i,j) = l(i',j')$ iff $a_{i,j} = a_{i',j'}$. Then the problem under consideration is equivalent to the problem

$$g_0(\tau,x) \equiv \sum_{j=1}^{r_0} c^*_{0,j}\, \tau_{l(0,j)} + d_0 \to \min \mid$$

$$g_p(\tau,x) \equiv s_p \exp\{a^T x\} - \tau_p \le 0,\ 1 \le p \le K,$$

$$g_{K+i}(\tau,x) \equiv \sum_{j=1}^{r_i} c^*_{i,j}\, \tau_{l(i,j)} + d_i \le 0,\ 1 \le i \le m,$$

$$(\tau,x) \in G' \equiv \{(\tau,x) \mid \|x\|_2^2 \le R^2,\ 0 \le \tau_p \le V',\ 1 \le p \le K\},$$

where

$$c^*_{i,j} = c_{i,j} / s_{l(i,j)} \le 1,\ V' = \max\{V + |d_i| \mid 0 \le i \le m\}.$$

The set G' admits a $(K+1)$-self-concordant barrier

$$F(\tau,x) = -\ln(R^2 - \|x\|_2^2) - \sum_{p=1}^{K} \ln(\tau_p(V' - \tau_p));$$

(it is easy to verify that the barrier parameter is $K+1$ instead of $2K+1$, the value implied by our general theory). The latter problem satisfies the conditions from the beginning of Sect. 3.6.2; in particular, the corresponding V can be taken equal to $V^* = (K+1)V'$. Obviously, G' is symmetric with respect to $z = (v'/2, 0)$, so $\sigma = 1$.

By T.3.2.((i),(ii),(iv).B) under appropriate choice of absolute constants in the below $O(\ )$ one can take as F^* the function

$$F^*(t,\tau,x) = O(1) \sum_{p=1}^{K} \Big\{ -\ln\big[\Omega^{-1}(\varepsilon)\, s_p^{-1} t + s_p^{-1} \tau_p - \exp\{a_p^T x\}\big] -$$

$$- \ln\big[\ln\big[\Omega^{-1}(\varepsilon)\, s_p^{-1} t + s_p^{-1}\tau_p\big] - a_p^T x\big]\Big\} - \sum_{i=1}^{m} \ln\big[\Omega^{-1}(\varepsilon)\, t -$$

$$- \sum_{j=1}^{r_i} c^*_{i,j}\,\tau_{l(i,j)} - d_i\Big] - \ln\Big[t - V^* - g_C(\tau,x)\Big] +$$
$$+ F(\tau,x) - \ln(3V^*\Omega(\varepsilon) - t),$$

with $\Omega(\varepsilon) = 4\,V^*/\varepsilon$, which results in

$$\vartheta_* = O(m + K).$$

So,

$$N(\varepsilon) \leq O((m + K)^{1/2} \ln(2(m + K)V^*/\varepsilon)),$$
$$M(\varepsilon) \leq O((m + K)^{1/2} K\,(n + m)\,(n + K) \ln(2(m + K)V^*/\varepsilon))$$

(the estimate for $M(\varepsilon)$ corresponds to the computation of Newton's direction by the conjugate gradient method).

C. L_p -approximation. Assume that

$$f_0(x) = \sum_{j=1}^{k} |a_j^T x - b_j|^p,\ x,\ a_j \in R^n \quad (p \in [1,\infty)).$$

For simplicity sake, $f_i(x)$, $1 \leq i \leq m$, are assumed to be convex quadratic forms. Introducing an extra variables vector $\tau = (\tau_1,\dots,\tau_k)^T$, we can rewrite the problem as

$$g_0(\tau,x) \equiv \sum_{j=1}^{k} \tau_j \to \min \mid g_j(\tau,x) \equiv |a_j^T x - b_j|^p - \tau_j \leq 0,$$
$$1 \leq j \leq k,$$
$$g_{k+i}(\tau,x) \equiv f_i(x) \leq 0,\ 1 \leq i \leq m,$$
$$(\tau,x) \in G' = \{(\tau,x) \mid \|x\|_2 \leq R,\ 0 \leq \tau_j \leq V,\ 1 \leq j \leq k\}.$$

The set G' admits a $(k + 1)$-self-concordant barrier

$$F(\tau,x) = - \ln(R^2 - \|x\|_2^2) - \sum_{j=1}^{k} \ln(\tau_j(V - \tau_j)),$$

and this set is symmetric with respect to its F-center

$$(x=0,\ \tau_j = V/2,\ 1 \leq j \leq k),$$

i.e. $\sigma = 1$. The parameter V for the transformed problem is the same as for the initial one.

By T.3.2.(iii), under appropriate choice of an absolute constant $O(1)$ (which does not depend on p) the function

$$\psi_{(p)}(t,u) = O(1)(2\ln(1/t) + \ln(1/(t^{2/p} - u^2))):$$
$$H \equiv \{(t,u) \in R^2 \mid t > |u|^p\} \to \mathbb{R}$$

is an $O(1)$-self-concordant barrier for the set Cl H. Hence, we can take as F^* the function

$$F^*(t,\tau,x) = \sum_{j=1}^{k} \psi_{(p)}(\Omega^{-1}(\varepsilon)t + \tau_j, a_j^T x - b_j) - \sum_{i=1}^{m} \ln(\Omega^{-1}(\varepsilon)\ t -$$
$$- f_i(x)) - \ln(t - V - \sum_{j=1}^{k} \tau_j) + F(\tau,x) - \ln(3V\Omega(\varepsilon) - t),$$

which results in

$\vartheta_* = O(m + k)$.

So,

$N(\varepsilon) = O((m + k)^{1/2} \ln(2(m + k) V/\varepsilon))$,

$M(\varepsilon) \leqslant O((m + k)^{1/2} n^2 (m + n + k)) \ln(2(m + k)V/\varepsilon))$.

D. Matrix norm minimization. Let $n = kl$ and let the elements of R^{kl} be regarded as $k \times l$ - matrices. Assume that $f_0(x) = \|x\|$ is the standard matrix norm (corresponding to the Euclidean norms in R^k and R^l); f_i, $1 \leqslant i \leqslant m$, as above, are convex quadratic forms. Without loss of generality, assume that $k \leqslant l$.

The problem can be reduced to the form (3.44) with

$$F^*(t,x) = - O(1) \ln \operatorname{Det}((t- V)^2 I_k - x x^T) - \sum_{i=1}^{m} \ln(\Omega^{-1}(\varepsilon)\ t - f_i(x)) - \ln(R^2 - \|x\|_2^2) - \ln(3V\Omega(\varepsilon) - t)$$

(**T**.3.2.(vi)) and with

$\vartheta_* = O(m + k)$.

So,

$N(\varepsilon) \leqslant O((m + k)^{1/2} \ln(2(m + k)V/\varepsilon))$,

$M(\varepsilon) \leqslant O((m + k)^{1/2}(m + n)\ n^2) \ln(2(m + k)V/\varepsilon))$, $n = k\ l$.

E. Optimization over positive-defined symmetric matrices.

Let $n = (k^2 + k)/2$ and let the elements of R^n be regarded as symmetric $k \times k$ -matrices. Suppose that the constraints defining G include the positively semidefiniteness condition. For the sake of simplicity, assume that

$G = \{ x \mid 0 \leqslant x \leqslant I_k \}$,

(the inequalities are understood in the operator sense). The functions f_i are assumed to be convex quadratic forms, $0 \leqslant i \leqslant m$.

One can take the function

$$F(x) = - \ln \operatorname{Det}(x) - \ln \operatorname{Det}(I_k - x)$$

as $2k$-self-concordant barrier for G (**T.3.2.**(v)). Notice that G is symmetric with respect to its F-center $z = \frac{1}{2} I_k$ $(\sigma = 1)$. So,

$$F^*(t,x) = - \ln(t - V - f_0(x)) - \sum_{i=1}^{m} \ln(\Omega^{-1}(\varepsilon)\ t - f_i(x)) + F(x) - \ln(3V\Omega(\varepsilon) - t),$$

$\vartheta_* = m + 2\ k + 1$,

$N(\varepsilon) \leqslant O((m + k)^{1/2} \ln(2(m + k)V/\varepsilon)),$

$M(\varepsilon) \leqslant O((m + k)^{1/2} (m + n) n^2) \ln(2(m + k)V/\varepsilon)).$

F. Inscribing the maximal volume ellipsoid into a convex polytope. The problem is as follows. Given a convex compact polytope K of the form

$$\{x \in R^n \mid a_i^T x \leqslant b_i, \ 1 \leqslant i \leqslant m\}.$$

We desire to find an ellipsoid

$$W(B,u) = \{u + B v \mid v^T v \leqslant 1\}$$

contained in Q and having maximum possible volume. This problem is considered in Sect. 7.

3.8. Universal barrier

3.8.1. We have noticed that the main question which arises in the connection with the barrier method is the problem of the choice of a self-concordant barrier for a given convex region. First of all, we desire to know if such a barrier does exist. The answer is positive.

Theorem 3.5. There exists an absolute constant C such that for each integer $n > 0$ each set $G \in C(R^n)$ admits a $C\,n$ - self-concordant barrier. If G does not contain any straight line, then one can take as the above barrier the function

$$F(x) = O(1) \ln |G^*(x)| : \operatorname{int} G \to \mathbb{R}, \tag{3.52}$$

where $O(1)$ is an appropriately chosen absolute constant,

$$G^*(x) = \{\phi \in R^n \mid \phi^T (y - x) \leqslant 1 \quad \forall\, y \in G\}$$

is the polar of G with respect to the point x and $|\ |$ means the Lebesque n-dimensional measure.

In particular, if (G,g) is a functional element on R^n and its epigraph does not contain any straight line, then under an appropriate choice of an absolute constant $O(1)$ the function

$$F(t,w) = O(1) \ln\{ \int_{R^n} (t - w^T x + g^*(w))^{-n-1} dw\} \tag{3.53}$$

is an $O(1)\ n$ - self-concordant barrier for the epigraph $\mathfrak{G}(G,g)$ of the functional element (G,g); herein

$$g^*(w) = \sup\{w^T x - g(x) \mid x \in G\}$$

is the Legendre transformation of the functional element (G,g).

Notice that the above statement can not be strengthen:

Proposition 3.7. Let G be a convex polytope in R^n, such that certain boundary point of G belongs exactly to k $(n-1)$-dimensional facets of G, with the normals to these facets being linearly independent. Then the parameter value ϑ of any ϑ - self-concordant barrier F for G is $\geq k$.

Comments. Of course, the result of T.3.5 is more theoretical than practical: in general case the barrier given by this theorem is not "practicable". Nevertheless the result seems to be of great importance. First of all, in the case of $n = 2$ formula (3.52) seems to be "computable"; at all events, it is not difficult to use it for polygones. Hence we can construct an $O(1)$ - self- concordant barrier for the epigraph of a given convex function of one variable (may be it would be necessary to approximate the function by a piecewise linear one). So we obtain a "regular" method to construct self-concordant barriers for sums of convex one-dimensional functions, and hence can apply the above techniques to the separable convex programming.

Moreover, even in multidimensional case the above result sometimes help in construction of "computable" barriers, as is demonstrated by two examples which follow.

3.8.1.1. A self-concordant barrier for the cone S_n^+ of $n \times n$ - symmetric positive semidefinite matrices. Such a barrier ($-\ln \mathrm{Det}\, x$, the parameter value is equal to n) has been described above (T.3.2.(iv)). It turns out that (3.52) leads to "almost" this barrier. Indeed, we have

$$G^*(x) = \{-\psi \mid \psi \in S_n^+,\ \mathrm{Tr}\{\psi\, x\} \leq 1\}$$

(notice that $\mathrm{Tr}\{\phi\, x\}$ is the usual scalar product on the space S_n of symmetric $n \times n$ - matrices). Hence

$$f(x) \equiv |G^*(x)| = \int_{\Omega(x)} d\psi,$$

where $x \in \mathrm{int}\ S_n^+$ and

$$\Omega(x) = \{\psi \in S_n \mid \mathrm{Tr}\{\psi\, x\} \leq 1\}.$$

Under the change of variables $\psi = x^{-1/2}\ \xi\ x^{-1/2}$ (the corresponding Jacobian is equal to $\mathrm{Det}^{-(n+1)/2}\, x$) we get

$$f(x) = \int_{\Omega(I_n)} (\mathrm{Det}^{-(n+1)/2}\, x)\, d\xi.$$

So (3.52) gives $F(x) = -O(n) \ln \mathrm{Det}\, x + \mathrm{const}$. By our theorem this is an $O(\dim S_n) = O(n^2)$ - self-concordant barrier for S_n^+. Of course, the result is too rough, but the barrier obtained can be easily improved: we can try to find a better barrier of the form $\lambda F(x)$, choosing λ as small as it is possible under the restriction that $\lambda F(x)$ must be a 1 - self-concordant function. This leads to the above mentioned n - self-concordant barrier - $\ln \operatorname{Det} x$; the latter has the best possible value, n, of the parameter (notice that an appropriate n-dimensional cross-section of S_n^+ is the usual positive orthant in R^n, so by virtue of P.3.7 the parameter value of any self-concordant barrier for S_n^+ is $\geq n$).

3.8.1.2. Now let us construct a barrier for the epigraph of the function

$$g(u) = x^T X^{-1} x, \quad u = (X,x) \in \operatorname{int} G,$$

$$G = \{ (X,x) \in S_n \times R^n \mid X \in S_n^+ \},$$

i.e. a barrier for the set

$$G^* = \operatorname{Cl} \{(t,X,x) \in \mathbb{R}\times S_n \times R^n \mid X \in \operatorname{int} S_n^+, \ t \geq x^T X^{-1} x\}.$$

Assume that g is extended from $\operatorname{int} G$ onto G as a lower semicontinuous convex function taking values in $\mathbb{R} \cup \{+\infty\}$; the extended function is denoted by g, too. Now (G,g) is a functional element, and we can use (3.53) to obtain the desired barrier. A straightforward computation (which is omitted here) leads to

$$F(t,X,x) = O(1) \frac{n+1}{2} \{-\ln \operatorname{Det} X - \ln(t - x^T X^{-1} x)\},$$

with the parameter value $O(n^2)$. As in the above example, the barrier can be improved; the resulting barrier is

$$F^*(t,X,x) = O(1) \{-\ln \operatorname{Det} X - \ln(t - x^T X^{-1} x)\}$$

with the parameter value $O(n)$ (see T.3.2.(vi)).

Remark. Notice that the fractional-quadratic function g is connected with the approach to the combinatorial optimization suggested in [Sh. 1987]. Namely, let us consider a quadratic programming problem of the form

$(\mathcal{K})$: $\quad K_0(x) \to \min \mid x \in R^n, \ K_i(x) = 0, \ 1 \leq i \leq q,$

where K_i, $0 \leq i \leq q$ are (nonconvex) polynomials of the second degree. For example, we can take $K_i(v) = v_i - v_i^2$, $1 \leq i \leq q = n$, which means Boolean restrictions on the variables. The application of branch and bounds methods to such problems requires a lower esimate of the objective's optimal value, $æ^*$, for $(\varkappa)$. In [Sh. 1987] such an estimate is taken in the form as follows. Let

$$-h(\lambda) = \min\{ K(x,\lambda) \equiv K_0(x) + \sum_{i=1}^{q} \lambda_i K_i(x) \mid x \in R^n \}.$$

The function $K(x,\lambda)$ considered as a function of x is a quadratic form which coefficients are linear in λ. Let

$$\Lambda = \{\lambda \mid K(x,\lambda) \text{ is positive definite form of } x\};$$

then Λ is an open convex set, and for $\lambda \in \Lambda$ we have

$$h(\lambda) = b^T(\lambda)\, A^{-1}(\lambda)\, b(\lambda),$$

where $b(\lambda)$, $A(\lambda)$ are some linear in λ vector and symmetric matrix, respectively. Assume that Λ is nonempty, and let Λ^* be the closure of Λ; let h be extended from Λ onto Λ^* as a lower semicontinuous convex function taking values in $\mathbb{R} \cup \{+\infty\}$ (the extended function also is denoted by h). The quantity

$$\phi^* = -\inf\{h(\lambda) \mid \lambda \in \Lambda^*\}$$

is a natural lower bound for $æ^*$. So we can produce a lower bound for $æ^*$ by solving the problem

$$h(\lambda) \to \min \mid \lambda \in \Lambda^*.$$

The latter problem can be solved by the barrier method, which requires a self-concordant barrier for the epigraph, Φ, of the functional element (Λ^*,h). Φ is the inverse image of the epigraph of the above introduced functional element (G^*,g). The latter set posesses a self-concordant barrier with the parameter value $O(n)$ (see T.3.2.(vi)). An iteration of the corresponding barrier method costs no more than $O(\max^3(n,q))$.

3.8.2. Above our barrier method was extended from the regions which possess self-concordant barriers to the regions which possess coverings. The following statement shows, that this generalization, being considered from the theoretical

viewpoint, does not extend the family of regions.

Proposition 3.8. Let $G \in C(E)$ and $\Gamma = (E', G', \pi, F)$ be a (ϑ, l)-covering for G, such that G' does not contain any straight line and such that $\pi^{-1}(x) \cap G'$ is bounded for each $x \in G$. Then the function

$$\Phi(x) = \inf \{ F(y) \mid y \in \text{int } G', \pi(y) = x \} : \text{int } G \to \mathbb{R}$$

is a ϑ-self-concordant barrier for G. ■

Of course, this proposition does not depreciate the above considerations connected with coverings. Indeed, we need the computation of the barrier and its derivatives, and these operations can be easy for the covering set and complicated for the original one.

3.9. Proofs of the results

3.9.1. Proposition 3.1. Let us verify that under the parameters choice described in (3.3) the relations $(\Sigma.1)$, $(\Sigma.2)$, $(\Sigma^{+}.3)$ hold. $(\Sigma.1)$ is obviously fulfilled. To prove $(\Sigma.2)$, let $\omega = (1 + \beta)^{-1}$ and notice that, by virtue of $F \in S_1^{+}(\text{int } G, E)$, for $x \in \text{int } G$, $h \in E$ we have

$$|D^3F(x)[h,h,h]| \leq 2\ (D^2F_t(x)[h,h])^{3/2}. \tag{1}$$

Let us fix x and h and let

$$p = (D^2f(x)[h,h])^{1/2}, \quad q = (D^2F(x)\ [h,h])^{1/2};$$

then, by $f \in \mathcal{A}(F,\beta)$, we have:

$$|D^3f(x)[h,h,h]| \leq 3^{3/2}\ \beta\ p^2\ q = 2\ \beta\ 3^{3/2}\ 2^{-1}\ p^2\ q =$$

$$= 2\ \beta\ \{(3^{2/3}\ 2^{-2/3}\ p^{4/3}\ \ t^{2/9})\ (3^{1/3}\ q^{2/3}\ t^{-2/9})\}^{3/2} \leq$$

$$\leq 2\ \beta\ \{ \tfrac{2}{3}\ (3^{2/3}\ 2^{-2/3}\ p^{4/3}\ t^{2/9})^{3/2} +$$

$$+ \tfrac{1}{3}\ (3^{1/3}\ q^{2/3}\ t^{-2/9})^3\}^{3/2} = 2\ \beta\ \ \{ p^2\ t^{1/3} + q^2\ t^{-2/3}\ \}^{3/2} =$$

$$= 2\ \beta\ t^{-1}\ \{ p^2\ t + q^2\}^{3/2} = 2\ \beta\ t^{-1}\ (D^2F_t(x)[h,h])^{3/2},$$

which by (i) implies $t\ |D^3f(x)[h,h,h]| + |D^3F(x)[h,h,h]| \leqslant \leqslant 2\ (1 + \beta)\ (D^2F_t(x)[h,h])^{1/2}$, and the latter relation together with (3.3) leads to the inequality required in $(\Sigma.2)$.

It remains to verify $(\Sigma^+.3)$. The closedness in $E_*(\Delta)$ of the sets $\{(t,x) \mid t \in \Delta,\ F_t(x) \leqslant a\}$ is an immediate corollary of the inclusion $F \in S_1^+(\text{int } G,E)$ and the continuity and boundness from below of f over each bounded subset of int G. Let us prove that for $x \in \text{int } G$, $h \in E$ the relations (2.2), (2.3) hold. By $F \in \mathfrak{B}(G,\vartheta) \subset S_1^+(\text{int } G,E)$ we have:

$$|(DF_t(x)[h])'_t - t^{-1}DF_t(x)[h]| = t^{-1}|DF(x)[h]| \leqslant$$

$$\leqslant \lambda(F,x)\ t^{-1}(D^2F(x)[h,h])^{1/2} \leqslant \vartheta^{1/2}\ t^{-1}\ (\ D^2F_t(x)[h,h])^{1/2} =$$

$$= \vartheta^{1/2}\ (1 + \beta)\ t^{-1}\ \alpha^{1/2}(t)\ (D^2F_t(x)[h,h])^{1/2},$$

which is required in (2.2). Furthermore, $|(D^2F_t(x)[h,h])'_t - t^{-1}D^2F_t(x)[h,h]| = t^{-1}\ D^2F(\ x)[h,h] \leqslant\ t^{-1}\ D^2F_t(x)[h,h]$, which leads to (2.3) ■

3.9.2. Proposition 3.2. (i), (ii) and (iii) admit a straightforward verification (sf. P.1.1). Let us prove (iv). (iv.1) is contained in C.1.2. Let us verify (iv.2). Denote the left hand side of (3.5) by γ, and let

$$\Delta = \{\ t \in \mathbb{R} \mid y + t\ (x - y) \in \text{int } G\ \} = (-T',\ T),$$

$T',\ T > 0$. Let $\phi(t) = F(y + t(x - y)): \Delta \to \mathbb{R}$; by (i) $\phi \in \mathfrak{B}(\text{Cl}\ \Delta,\ \vartheta)$. It is possible that ϕ is a constant; then (3.5), (3.6), (3.7) for x and y under consideration are obvious. Moreover, in this case $x - y \in E_F$, thus either $x = y$, or the whole straight line (x,y) is contained in G; in both of the cases

$\pi_y(x) = 0$, so (3.8) holds.

Now assume that ϕ is not a constant. Since ϕ is a barrier for Cl Δ, we have $\phi''(t) > 0$ (C.1.1) and $(\phi'(t))^2/\phi''(t) \leqslant \vartheta$, $t \in \Delta$, or $\phi''(t) \geqslant \vartheta^{-1} (\phi'(t))^2$. Let $\psi(t) = \phi'(t)$ and $\psi(t_0) > 0$ for some $t_0 \in \Delta$. By the comparison theorem for $\eta(t) = \vartheta\psi(t_0)(\vartheta - (t - t_0)\psi(t_0))^{-1}$ (notice that $\eta' = \vartheta^{-1}\eta^2$, $\eta(t_0) = \psi(t_0)$) we have $\psi(t) \geqslant \eta(t)$ for each $t \geqslant t_0$, such that ϕ and η are well defined at t; thus, $T - t_0 \leqslant \vartheta/\psi(t_0)$.

Let us verify (3.5). This relation is obvious for $DF(x)[x-y] \leqslant 0$; now assume that $DF(x)[x-y] > 0$. Since $DF(x)[x-y] = \phi'(1) = \psi(1)$, we have $T - 1 \leqslant \vartheta / \psi(1)$, so (3.5) holds. It is clear that (3.5) holds for $y \in \partial G$ as well.

Now let us prove (3.7). The application of (3.5) to the barrier ϕ for Cl Δ gives for $0 \leqslant t < T < \infty$: $\phi'(t) \leqslant \vartheta/(T - t)$, so $F(x) = \phi(1) \leqslant \phi(0) + \int_0^1 \vartheta(T-t)^{-1}dt = F(y) + \vartheta \ln(T/(T-1))$, which implies (3.7). If $T = \infty$, or, that is the same, $\pi_y(x) = 0$, we have by (3.5) (the latter relation is applied to ϕ): $\phi'(t) \leqslant 0$, $t \in \Delta$, so (3.7) is obvious.

Now let us prove (3.6). Since ϕ is convex, then $\phi'(t) \geqslant \phi'(0)$, $0 < t < T$, and by (3.6) for $T < \infty$ we have $\phi'(t) \leqslant \vartheta/(T - t)$. So $\phi'(0) \leqslant \vartheta/T = \vartheta\pi_y(x) \leqslant \vartheta$, or $DF(y)[x-y] \leqslant \vartheta$. If $T = \infty$, then, by (3.5), $\phi'(t) \leqslant 0$, $t > 0$, so in this case again $DF(y)[x-y] \leqslant \vartheta$; under necessary renotations the inequality obtained is (3.6).

Let us prove (3.8). ϕ is a barrier for Cl Δ, so in the case of $T < \infty$ the relation $0 \leqslant t < T$ implies, by (1v.1), the inequality $t + (\Phi''(t))^{-1/2} \leq T$, or $\Phi''(t) \geq (T - t)^{-2}$. So

$$F(x) = \Phi(1) = \Phi(0) + \Phi'(0) + \int_0^1 \Phi''(t)(1 - t)\,dt \geq F(y) +$$

$$+ DF(y)[x - y] + \int_0^1 (1 - t)(T - t)^{-2}\,dt = F(y) + DF(y)[x-y] +$$

$$+ \ln(1/(1 - \pi_y(x))) - \pi_y(x),$$

which is required in (3.8). In the case of $T = \infty$ we have $\pi_y(x) = 0$, and (3.8) is an immediate corollary of the convexity of F.

Let us prove (3.9). The situation in an obvious manner can be reduced to the case of $E_F = \{0\}$. Let us provide E by the scalar product of the form $\langle h, g\rangle = D^2F(y)[h, g]$, let $\| \;\|$ be the corresponding norm and let us identify the first and second order differentials with the gradients and Hessians. We have $F''(y) = I$, and the open unit ball V centered at y is contained in int G ((iv.1)). Let y' - be the point of the ray $[y, x)$, such that x lies between y and y', and let V' be the image of V under the homothety with the center at y' and the coefficient $\alpha = \|x - y'\|/\|y - y'\|$; $V' \subset$ int G is an opened ball with the radius α centered at x. Let $0 < \alpha' < \alpha$, let h be the unit normalization of $F'(x)$ and let $z = x - \alpha' h$. Then $z \in$ int G, $\pi_z(x) \leq 1/2$, which, by (3.5), gives $\langle F'(x), x - z\rangle \leq \vartheta$, or $\|F'(x)\| \leq \vartheta/\alpha'$. Under an appropriate choice of y' and α' the quantity ϑ/α' can be done a number arbitrary close to $\vartheta/(1 - \pi_y(x))$; the inequality $\|F'(x)\| \leq \vartheta/(1 - \pi_y(x))$ is, by the choice of the scalar product, the desired (3. 9).

Let us prove (3.10). Since $W_1(y) \subset$ int G, the set

$$V = \{ z \in E \mid D^2F(y)[z-x, z-x] < (1 - \pi_y(x))^2 \},$$

which is a union of the images of $W_1(y)$ under homotheties with the centers in int G, is contained in int G. It suffices to prove (3.10) under the assumption that, что $D^2F(x)[h,h] = 1$; moreover, it is possible to assume that $DF(x)[h] \geq 0$ (otherwise we can replace h by $-h$). Under the notations

$$x(t) = x + th, \quad \Phi(t) = DF(x(t))[h],$$

$0 \leqslant t < T \equiv \sup(t \mid x(t) \in \text{int } G)$

we have $T \geqslant 1$ ((iv.1)) and $\phi'(t) \geqslant (1 - t)^2$, $0 \leqslant t < 1$ (the latter - by T. 1.1 and by $D^2F(x)[h,h] = 1$). These relations together with the inequality $\phi(0) \geqslant 0$ for $0 < t < 1$ lead to $\phi(t) \geqslant t(3 - 3t + t^2)/3$, or to

$$DF(x(t))[x(t) - x] = t\,\phi(t) \geqslant t^2(3 - 3t + t^2)/3 \equiv \alpha(t).$$

By (3.5) this means that $\pi_x(x(t)) \geqslant \alpha(t)/(\vartheta + \alpha(t))$. Taking t being close to 1, we find out that $\pi_x(x + h) \geqslant (1 + 3\vartheta)^{-1}$, so the point $x + (1 + 3\vartheta)h$ does not belong to int G and hence belongs to V. The latter fact means that

$$(1 + 3\vartheta)^2 D^2F(y)[h,h] \geqslant (1 - \pi_y(x))^2 =$$

$$= (1 - \pi_y(x))^2 D^2F(x)[h,h],$$

which is required in (3.10).

Let us prove (3.11). Let $\Delta = \{ t \mid z + t(x - z) \in \text{int } G\}$. Then $\Delta = (0,T)$, $T = \pi_z^{-1}(x) \geqslant (1 + \vartheta^{1/2})^2$. Let $\phi(t) = F(z + t(x - z))$; then $\phi(t)$ is a barrier for Cl Δ, so $\phi''(t) \geqslant t^{-2}$, $t \in \Delta$ ((iv. 1)). When $1 < t < T$, we have by (3.5):

$$t\,\phi'(t) \leqslant \vartheta\, t/(T - t)$$

(when $T = \infty$, we set $1/(T - t) = 0$); so

$$\phi'(1) + \int_1^t \tau^{-2}\, d\tau \leqslant \vartheta/(T - t), \quad 1 \leqslant t < T,$$

or $\phi'(1) \leqslant (\vartheta/(T - t) - 1 + t^{-1})$, $1 \leqslant t < T$. If $T < \infty$, then in the above inequality one can set $t = (1 + \vartheta^{1/2})^{-1} \pi_z^{-1}(x)$ (this quantity, by the assumption, is $\geqslant 1$), which leads to

$$\phi'(1) \leqslant -1 + (1 + \vartheta^{1/2})^2 \pi_z(x).$$

If $T = \infty$, then the same relation follows from the above inequality when $t \to \infty$. So

$$DF(x)[z - x] = -\phi'(1) \geqslant 1 - (1 + \vartheta^{1/2})^2 \pi_z(x),$$

Q.E.D. (iv) is proved.

(v): The fact that F is a constant along the intersections of int G with $E_F + x$, $x \in$ int G, is obvious because $\lambda(F, \ ) < \infty$; since F tends to ∞ as the argument approaches to a boundary point of G, the sets $x + E_F$, $x \in$ int G, are contained in int G. When proving the remaining statements of (v) we can assume that $E_F = \{0\}$ (otherwise consider the reduction of F onto an intersection of G and some subspace which complements E_F with respect to E). It is clear that in the case of bounded G F attains its minimum over int G at an unique point (since in the case of $E_F = \{0\}$ F is strongly convex. Now assume that G is unbounded; let us prove that then F is unbounded from below. Indeed, int G contains a ray $L = [y,x)$, y, $x \in$ int G. By (3.5) $DF(z)[x-y] \leqslant 0$, $z \in L$; if $\pi_x(y) = 0$, then int G contains the ray $[x,y)$, hence $DF(z)[y-x] \leqslant 0$, $z \in L$, or $D^2F(z)[y-x] = 0$, which contradicts the assumption $E_F = \{0\}$. Thus, $\pi_x(y) > 0$, and hence for $z_t = y + t(x - y)$ we have $\lim_{t\to\infty} \pi_{z_t}(y) = 1$. By (3.8) we have

$$F(y) \geqslant F(z_t) + DF(z_t)[y-z_t] + \ln(1/(1 - \pi_{z_t}(y))) - \pi_{z_t}(y) \geqslant$$
$$\geqslant F(z_t) + \ln(1/(1 - \pi_{z_t}(y))) - 1$$

(since $DF(z_t)[y - z_t] \geqslant 0$ by the above arguments), which implies

$$F(z_t) \leqslant F(y) + \ln(1 - \pi_{z_t}(y)) + 1 \to -\infty, \ t \to \infty.$$

It remains to verify (3.12). The left inclusion follows from (iv.1). To prove the right inclusion it suffices to show that if $x(F)$ is the minimizer of F over int G and $h \in E$ is such that $D^2F(x(F))[h,h] = 1$, then the point $x(F) + \rho h$, $\rho = (1 + 3\vartheta)$, does not belong to int G. Let $x(t) = x(F) + t h$ and $\phi(t) = DF(x(t))[h]$; then $\phi(0) = 0$. So, by the choice of h and T.1.1, we have $\phi'(t) \geqslant (1 - t)^2$, so

$$\phi(t) \geqslant t(3 - 3t + t^2)/3, \ 0 \leqslant t < 1.$$

On the other hand, by (3.5)

$$t\,\phi(t) \leqslant \vartheta\, \pi_{x(F)}(x(t))/(1 - \pi_{x(F)}(x(t))).$$

These results imply

$$\pi_{x(F)}(x(1)) \geqslant (1 + 3\vartheta)^{-1},$$

which is the right inclusion. (v) is proved.

Let us prove (vi). The left inequality in (vi) follows from (iv.1). To prove the right inequality it suffices to consider the case of $D^2F(x)[h,h] = 1$, $DF(x)[h] \geq 0$ and to verify that the point $x + (1 + 3\theta) h$ does not belong to int G; the latter can be done in the same manner as in the proof of the right inclusion in (3.12). ■

3.9.3. Proposition 3.3.

1^0. Let us prove (3.24_i), (3.25_i) inductively. By definition of g and t_0 we have $\lambda(F^{(1)}_{t_0}, x_{-1}) = 0$, so (3.24_0) holds. Assume that x_{i-1} are well defined, belong to int $\dot{G}$ and (3.24_i), $0 \leq i \leq k$, (3.25_i), $0 \leq i < k$, hold. Relation (3.24_k), by $\lambda_1 \in (0, \lambda_*)$ and T.1.3.(ii) (the theorem is applied with $F = F^{(1)}_{t_k}$, $x = x_{k-1}$), implies that x_k is well defined and (3.25_k) holds. Furthermore, by (3.25_k), (3.20), P.2.1 and T.2.1 (the theorem is applied with $\varkappa = \lambda_1$, $x = x_k$, $t = t_k$, $t' = t_{k+1}$, $F_{\cdot} = F^{(1)}_{\cdot}$), (3.24_{k+1}) holds. The induction is over.

2^0. Let us prove that $t^* < \infty$ and that (iii) holds. Let us fix $i < t^* - 1$ and denote $F^{(1)}_{t_i}$ by Φ. Then $\Phi \in S^+_1(\text{int } G, E)$ and $\lambda(\Phi, x_i) \leq \lambda'_1 < \lambda_* < 1/3$ by virtue of (3.24_i). Henceby (1.13) we have

$$\Phi(x_i) - \Phi(x(F)) \leq \Phi(x_i) - \inf\{\Phi(x) \mid x \in \text{int } G\} \leq$$
$$\leq \zeta(\lambda'_1)/2 \leq 1/18 \tag{1}$$

(we have taken into account (3.14), (3.15)). By T.1.1 and P.3.2.(iv.1) we have $\|e\|_F = 1$, $t \in [0,1) \Rightarrow x(F) + te \in$ int G,

$$\frac{d^2}{dt^2} F(x(F) + te) \geq (1 - t)^2,$$

or, by $\frac{d}{dt} F(x(F) + te)|_{t=0} = 0$,

$$\|e\|_F = 1,\ t \in [0,1) \Rightarrow x(F) + te \in \text{int } G,$$
$$F(x(F) + te) - F(x(F)) \geq t^2(6 - 4t + t^2)/12, \tag{2}$$

and hence

$$\|e\|_F = 1,\ t \in [0,1) \Rightarrow \Phi(x(F) + te) - \Phi(x(F)) \geq$$
$$t^2(6 - 4t + t^2)/12 - t_i\, t\, \|g'\|_F, \tag{3}$$

where g' is the gradient of g with respect to the Euclidean structure defined by the scalar product $\langle\ ,\ \rangle_F$. By virtue of (3.9) we have $|g'|_F \le \vartheta\,(1 - \pi_{x(F)}(w))^{-1} = \Omega$, and we get

$$|e|_F = 1,\ t \in [0,1) \Rightarrow \Phi(x(F) + te) - \Phi(x(F)) \geq t^2(6 - 4t + t^2)/12 - t\, t_i\, \Omega. \qquad (4)$$

Let us verify that

$$t_i \geq \min\{\frac{19}{288\,\Omega},\ \frac{\lambda_2 - \lambda_1'}{2\,\Omega}\}.$$

Indeed, otherwise for $x \in \partial W_{1/2}(x(F))$, by virtue of (4), we have

$$\Phi(x) - \Phi(x(F)) > 17/192 - 19/576 = 1/18\ ,$$

and (1) leads to $x_i \in W_{1/2}(x(F))$. Hence by T.1.1 we have

$$|g'|_{x_i,F} \le 2\ |g'|_F \le 2\,\Omega,\ \text{or}\ \lambda(F,x_i) \le \lambda(\Phi,x_i) + t_i\ |g'|_{x_i,F} \le \lambda_1' + 2t_i\Omega < \lambda_2,$$

which contradicts the assumption $i < i^* - 1$.

It remains to notice that (3.26) is equivalent to (3.22); relation (3.26) by (1.13), (3.15) implies (3.27). ■

3.9.4. Proposition 3.4.

1^o. By (3.14) we have $\lambda_3(1+\beta)^{-1} > \lambda_2 \geq \lambda(F,u)$ (the latter is a corollary of (3.26)), so $t_0 > 0$. Let us verify that

$$t_0 \geq \frac{\lambda_3 - (1 + \beta)\lambda_2}{9(1+\beta)V_F(f)}. \qquad (1)$$

Indeed, by (3.27) we have $u \in W_{1/3}(x(F))$, so by T.1.1

$$|e|_{u,F} \le |e|_F(1 - 1/3)^{-1} = \tfrac{3}{2}\ |e|_F,\ e \in E.$$

Hence the ellipsoid $W_{1/9}(u)$ is contained in $W_{1/2}(x(F))$, which leads to $|f'(u)|_{u,F} \le 9\ V_F(f)$, and (1) follows.

2^o. Let us prove (3.38_i), (3.39_i) inductively. We have

$$\lambda(F^{(2)}_{t_0},u) = \alpha^{-1/2}(t_0)\sup\{\,|DF_{t_0}(u)[h]|\,(D^2F_{t_0}(u)[h,h])^{-1/2}\ |\ h \ne 0\ \} \le (1+\beta)\ \sup\{\ |DF_{t_0}(u)[h]|\,(D^2F(u)[h,h])^{-1/2}\ |\ h \ne 0\ \} \le$$

$\leqslant (1 + \beta) \sup\{ \{|DF(u)[h]| + t_0|Df(u)[h]|\}\langle h,h\rangle_{u,F}^{-1/2} \mid h \neq 0\} \leqslant$

$\leqslant (1 + \beta) (\lambda(F,u) + t_0\|f'(u)\|_{u,F}) = \lambda_3$

(the latter - by no (3.34)), so (3.38_0) holds. Assume that $k \geqslant 0$ is such that (3.38_i) holds for $0 \leqslant i \leqslant k$, and ($3.39_i$) holds for $0 \leqslant i < k$. Relation (3.38_k) by T.1.3.(11) implies that x_k is well defined, belongs to int G and that (3.39_k) holds. Furthermore, by T.2.1 and (3.36) relation (3.39_k) leads to (3.38_{k+1}). (1) is proved.

3^o. Let us prove (ii). Let us fix i and denote $t_i = t$, $F_{t_i}^{(2)} = \Phi$, $x_i = z$. By (3.39_i) we have $\lambda(\Phi,z) \leqslant \lambda_3' < 1$; moreover, $\Phi \in S_a^+(\text{int } G,E)$, $a = (1 + \beta)^{-2}$. By T.1.3 the function Φ attains its minimum over int G in some point v, and (1.12), (1.13) and (3.15) imply:

$$\Phi(z) - \Phi(v) \leqslant \tfrac{1}{2}a\,\zeta(\lambda_3') \equiv \nu, \quad D^2\Phi(v)[z-v,z-v] < a. \qquad (2)$$

Let us verify that

$$DF(v)[z-v] \geqslant -\vartheta. \qquad (3)$$

Indeed, by virtue of the second relation in (3.42) and C.1.2, the point $z' = v + (v - z)$ belongs to int G; (3.6), as applied to $x = v$, $y = z'$, implies (3.43).

Let x^* be the minimizer of f over G (the point does exist since G is bounded and f is lower semicontinuous on G). We have

$$f(x^*) \geqslant f(v) + Df(v)[x^* - v];$$

further, by definition of v we have

$$Df(v)[h] = - t^{-1}DF(v)[h],$$

so

$$f(x^*) \geqslant f(v) - t^{-1}DF(v)[x^* - v] \geqslant f(v) - t^{-1}\vartheta$$

(the latter - by (3.6)). At the same time by (3.42) we have

$$f(z) \leqslant f(v) + (F(v) - F(z) + \nu)\, t^{-1} \leqslant f(v) +$$
$$+ (DF(v)[v-z] + \nu)\, t^{-1} \leqslant f(v) + (\vartheta + \nu)\, t^{-1}$$

(the latter - by (3.43)). The above inequalities imply

$$f(z) \leqslant f(x^*) + t^{-1}(\,2\vartheta + \nu\,),$$

which inequality together with (3.41) proves (3.40). ■

3.9.5. **Theorem 3.2.**

In what follows the quantities denoted by ρ (with sub- or superscripts) are nonnegative.

3.9.5.1. (i):

By virtue of P.3.2.(i) and P.3.2.(iii) it suffices to prove that if $\Phi(x)$ is a convex quadratic form on E and the set $\{x \in E \mid \Phi(x) < 0\}$ is nonempty, then the function $F(x) = \ln(1/(-\Phi(x)))$: $\operatorname{int} G \to \mathbb{R}$ belongs to $\mathcal{S}(G,1)$, where $G = \{x \in E \mid \Phi(x) \leq 0\}$. For $x \in \operatorname{int} G$ and $h \in E$ we have:

$DF(x)[h] = -\Phi^{-1}(x)\, D\Phi(x)[h]$;

$D^2F(x)[h,h] = -\Phi^{-1}(x)\, D^2\Phi(x)[h,h] + \Phi^{-2}(x)\, (D\Phi(x)[h])^2 = |\Phi^{-1}(x)|\, D^2\Phi(x)[h,h] + \Phi^{-2}(x)\, (D\Phi(x)[h])^2$;

$|D^3F(x)[h,h,h]| = |\, 3\, \Phi^{-2}(x)\, D^2\Phi(x)[h,h]\, D\Phi(x)[h] - 2\, \Phi^{-3}(x)\, (D\Phi(x)[h])^3|$.

Hence

$|D^3F(x)[h,h,h]| \leq 3\, \lambda\, \rho + 2\, \rho^3$,

where λ, $\rho \geq 0$ are such that

$\lambda + \rho^2 = D^2F(x)[h,h]$,

and we get

$|D^3F(x)[h,h,h]| \leq 2\, (D^2F(x)[h,h])^{3/2}$,

so $F \in S_1^+(\operatorname{int} G, R^n)$.

The inequality

$|DF(x)[h]| \leq (D^2F(x)[h,h])^{1/2}$,

or, which is the same, $\lambda(F,x) \leq 1$, is obvious by virtue of the above expressions for the derivatives of F. ∎

3.9.5.2. (ii):

A). Let $\Psi(z) = -\ln(\Phi(z))$, $\Phi(z) = x^Tx - t^2$, $z = (t,x) \in \mathbb{R} \times R^n$ (Ψ and Φ are defined on $H = \operatorname{int}\{z \in \mathbb{R} \times R^n \mid t \geq |x|\} \equiv \operatorname{int} G$).

1^0. Let $h = (s,u) \in \mathbb{R} \times R^n$. For $z = (t,x) \in H$ we have:

$$D\Psi(z)[h] = 2\,(x^T u - t\, s)/\gamma, \tag{1}$$

$$D^2\Psi(z)[h,h] = 2\,\nu\,\gamma^{-2}\,(s - 2\,t\,\nu^{-1}\,x^T u)^2 + 4\,\gamma^{-1}\nu^{-1}\,\{(x^Tx)\,(u^Tu) - (x^Tu)^2\} + 2\,\nu^{-1}u^T u, \tag{2}$$

$D^3\Psi(z)[h,h,h] = 12\,\gamma^{-2}(x^Tu - ts)(u^Tu - s^2) +$

$+ 16\ \gamma^{-3}\ (x^T u - ts)^3,$ (3)

where

$$\gamma = t^2 - x^T x,\ \nu = t^2 + x^T x; \quad (4)$$

relations (1)-(4) can be obtained by a straightforward computation.

By (2) Φ is convex (and, of course, C^∞) on H.

2^0. Let us verify that

$$\xi \equiv 2\ \sup\{-D\Phi(z)[h] - \tfrac{1}{2}\ D^2\Phi(z)[h,h]\ |\ h \in \mathbb{R}\times R^n\} \leq 2. \quad (5)$$

Let us fix $h = (t,u)$ and let $\eta = x^T u$. Assume that $x \neq 0$; by (2) and by virtue of $u^T u \geq \eta^2/(x^T x)$ the following inequality holds:

$$D^2\Phi(z)[h,h] \geq 2\ \gamma^{-2}\ \nu^{-1}\ (\nu\ s - 2\ t\ \eta)^2 + 2\ \ \nu^{-1}\ \eta^2/(x^T x). \quad (6)$$

Hence for $x \neq 0$ we have

$$\xi \leq 2\ \sup\{-2\ \gamma^{-1}(\eta - ts) - \gamma^{-2}\ \nu^{-1}\ (\nu s - 2t\eta)^2 - - \nu^{-1}\eta^2/(x^T x)\ |\ s,\ \eta \in \mathbb{R}\}.$$

Introducing new variables instead of (s,η) - namely, $(\sigma = \nu s - 2t\eta,\ \eta)$, we get

$$\xi \leq 2\ \sup\{-\ 2\ \gamma^{-1}\ \eta\ (1 - 2\ t^2/\nu) + 2\ t\ \gamma^{-1}\nu^{-1}\ \sigma - - \gamma^{-2}\ \nu^{-1}\ \sigma^2 - \nu^{-1}\ \eta^2/(x^T x)\ |\ \sigma,\ \eta \in \mathbb{R}\}.$$

But $1 - 2t^2/\nu = -\ \gamma/\nu$, so

$$\xi \leq 2\ \sup\{[2\ \eta - \eta^2/(x^T x)]\ \nu^{-1} + (2\ t\ \gamma\sigma - \sigma^2)\ \gamma^{-2}\ \nu^{-1}\ |\ \sigma,\ \eta \in \mathbb{R}\} = 2\ (\nu^{-1}\ x^T x + \nu^{-1}\ t^2) = 2.$$

Thus, in the case of $x \neq 0$ (5) is proved.

Now let $x = 0$. Then, by (1) - (4),

$$\xi \leq 2\ \sup\{2\ s/t - s^2/t^2 - u^T u/t^2\ |\ (s,u) \in \mathbb{R}\times R^n\} = 2,$$

and the proof of (5) is completed.

3^0. Assume that $h \neq 0$. Let us evaluate the quantity

$$\zeta = |D^3\Phi(z)[h,h,h]|/(D^2F(z)[h,h])^{3/2}.$$

By (1) - (4) this quantity is invariant with respect to the change of z by $\lambda\ z$ ($\lambda > 0$).

$3^0.1$. Assume, first, that $x \neq 0$; as we have noticed, we can restrict ourselves to the case of $x^T x = 1$. In this situation (1) - (4) lead to

$$\rho^2 \equiv D^2\Phi(z)[h,h] = 2\ \nu\ \gamma^{-2}\ (s - 2\ t\ \gamma^{-1}\ x^T u)^2 +$$

$$+ 4\gamma^{-1}\nu^{-1}(u^Tu - (x^Tu)^2) + 2\nu^{-1}u^Tu = \rho_1^2 + \rho_2^2 + \rho_3^2, \qquad (6)$$

where

$$\rho_1^2 = 2\nu\gamma^{-2}(s - 2t\nu^{-1}x^Tu)^2,$$

$$\rho_2^2 = 4\gamma^{-1}\nu^{-1}(u^Tu - (x^Tu)^2), \quad \rho_3^2 = 2\nu^{-1}u^Tu, \qquad (7)$$

and

$$\xi \equiv |D^3\Phi(z)[h,h,h]| = |\,12\gamma^{-2}(x^Tu - ts)(u^Tu - s^2) +$$

$$+ 16\gamma^{-3}(x^Tu - ts)^3|,$$

while

$$\gamma = t^2 - 1 > 0, \quad \nu = t^2 + 1. \qquad (8)$$

Let us first verify that

$$\mu_1 \equiv |x^Tu - ts|\,\gamma^{-1} \le 2^{-1/2}(\rho_1^2 + \rho_3^2)^{1/2}. \qquad (9)$$

Indeed, let $\sigma = (x^T u - t\,s)/\gamma$, $\eta = x^T u$; then, by $u^Tu \ge \eta^2$ (since $x^T x = 1$), we have:

$$\rho_1^2 + \rho_3^2 \ge 2\gamma^{-2}\nu^{-1}(\gamma(\eta - \sigma\nu)/t)^2 + 2\eta^2/\nu =$$

$$= 2\nu^{-1}((\eta - \sigma\nu)^2 t^{-2} + \eta^2) \ge 2\nu^{-1}(\nu^2\sigma^2/(1 + t^2)) = 2\sigma^2.$$

Now let us evaluate the quantity

$$\mu_2 = |u^Tu - s^2|/\gamma.$$

Let $u = \eta x + v$, $v^Tx = 0$; under the notation $\tau^2 = v^Tv$ we have

$$u^Tu = \eta^2 + \tau^2, \quad x^T u = \eta. \qquad (10)$$

So

$$\rho_1^2 = 2\nu\gamma^{-2}(s - 2t\eta/\nu)^2, \quad \rho_2^2 = 4\gamma^{-1}\nu^{-1}\tau^2,$$

$$\rho_3^2 = 2\nu^{-1}(\eta^2 + \tau^2), \quad \mu_2 = |\eta^2 + \tau^2 - s^2|/\gamma.$$

We have

$$\rho_2^2 + \rho_3^2 - 2\eta^2/\nu = (4\gamma^{-1}\nu^{-1} + 2\nu^{-1})\tau^2 =$$

$$= 2\nu^{-1}\gamma^{-1}(2 + t^2 - 1)\tau^2 = 2\gamma^{-1}\tau^2,$$

so

$$|\mu_2| \le \tfrac{1}{2}(\rho_2^2 + \rho_3^2 - 2\eta^2\nu^{-1}) + |\eta^2 - s^2|/\gamma \equiv$$

$$\equiv \tfrac{1}{2}(\rho_2^2 + \rho_3^2 - 2\eta^2\nu^{-1}) + \mu_2^*. \qquad (11)$$

Let $\sigma = \nu\gamma^{-1}(s - 2t\eta/\nu)$; then

$$s^2 - \eta^2 = \gamma^2\nu^{-2}\sigma^2 + 4t\gamma\nu^{-2}\sigma\eta + 4t^2\nu^{-2}\eta^2 - \eta^2 =$$

$$= \gamma^2\nu^{-2}(\sigma^2 - \eta^2) + 4t\gamma\nu^{-2}\sigma\eta$$

and $\eta^2 + \sigma^2 = \nu(\rho_1^2 + 2\eta^2/\nu)/2$, which leads to

$$|s^2 - \eta^2|/\gamma \le \gamma\nu^{-2}\max\{|\sigma^2 - \eta^2 + 4t\sigma\eta\gamma^{-1}| \mid \sigma^2 + \eta^2 \le$$

$\leq \nu(\rho_1^2 + 2\eta^2/\nu)/2\} = \gamma\nu^{-2} \max\{ -r^2 \cos2\phi + 2t\gamma^{-1}r^2\sin2\phi \mid r^2 \leq$
$\leq \nu\ (\rho_1^2 + 2\ \eta^2/\nu)/2\ \} = (\rho_1^2 + 2\ \eta^2/\nu)/2$
(we have taken into account that $(1 + 4\ t^2\ \ \gamma^{-2}) = (\nu/\gamma)^2)$),
which, by (11) and (6), implies

$$|\mu_2| \leq \rho^2/2. \tag{12}$$

(9), (12) and (3) lead to

$$|D^3\Psi(z)[h,h,h]| \leq 6\ \rho^2\ 2^{-1/2}\ (\rho_1^2 + \rho_3^2)^{1/2} + 16(2^{-1/2}[\rho_1^2 + \rho_3^2]^{1/2})^3 \leq 9\ (D^2\Psi(z)[h,h])^{3/2}. \tag{13}$$

3^0.2. Now assume that $x = 0$. By (1) - (4) in the case under consideration for $h = (s,u)$ we have

$$D^2\Psi(z)[h,h] = 2\ s^2\ t^{-2} + 2\ u^Tu\ t^{-2},$$
$$|D^3\Psi(z)[h,h,\ h]| = |12\ t^{-3}\ s\ (s^2 - u^Tu) + 16\ t^{-3}\ s^3| =$$
$$= |4\ t^{-3}\ s^3 + 12\ t^{-3}\ s\ u^Tu| \leq 9\ (D^2\Psi(z)[h,h])^{3/2},$$

and the resulting inequality in (13) holds. Thus, (13) is proved.

4^0. (13) means that for $c = (9/2)^2$ and

$$F(t,x) = c\ \Psi(t,x):\ H \to \mathbb{R}$$

the inclusion $F \in S_1(H,\mathbb{R} \times R^n)$ holds; in fact, obviously, $F \in S_1^+(H,\mathbb{R}\ \dot{\times}\ R^n)$. In view of (5) we have $\vartheta(F) \leq 2\ c$. Thus, F is a $2c$ - self-concordant barrier for G. ■

B). Let

$$G^* = \{(\tau,x) \in \mathbb{R} \times R^n \mid \tau \geq (x^Tx)^{1/2}\},$$
$$\Psi^*(\tau,x) = -\ \ln(\tau^2 - x^Tx):\ H^* \equiv \text{int}\ G^* \to \mathbb{R},$$
$$\Phi(t,x) = \Psi^*(\zeta(t),x) - \ln t:\ H \equiv \text{int}\ G \to \mathbb{R}.$$

It is clear that G is closed and convex, and

$$H = \{(t,x) \in \mathbb{R} \times R^n \mid \zeta(t) > \|x\|\}.$$

Moreover, the relations $\zeta(0) = 0 < \zeta(t)$, $t > 0$, imply that if $z_i \in H$ and $z_i \to z \in \partial H$ as $i \to \infty$, then $\Phi(z_i) \to \infty$.

1^0. Let us fix $z = (t,x) \in H$ and set $\tau_* = \zeta(t)$ (> 0); then $z^* \equiv (\tau_*,x) \in H^*$. For $h = (s,u) \in \mathbb{R} \times R^n$ let $h^* = (s_*,u)$, $s_* = \zeta'(t)\ s$. We have

$$D\Phi(z)[h] = D\Psi^*(z^*)[h^*] - s/t, \tag{14}$$
$$D^2\Phi(z)[h,h] = D^2\Psi^*(z^*)[h^*,h^*] + D\Psi^*(z^*)[(\zeta''(t)s^2,0)] + (s/t)^2, \tag{15}$$

$$D^3\Psi(z)[h,h,h] = D^3\Psi^*(z^*)[h^*,h^*,h^*] +$$
$$+3D^2\Psi^*(z^*)[h^*,(\zeta''(t)s^2,0)]+D\Psi^*(z^*)[(\zeta'''(t)s^3,0)] -2(s/t)^3. \quad (16)$$

One has $D\Psi^*(z^*)[(\zeta''(t)s^2,0)] = 2\,\tau_*\,(\tau_*^2-x^Tx)^{-1}\cdot|\zeta''(t)|\,s^2$ (we have taken into account the concavity of ζ), whence

$$D^2\Psi(z)[h,h] = D^2\Psi^*(z^*)[h^*,h^*] +2\,\tau_*(\tau_*^2-x^Tx)^{-1}|\zeta''(t)|\,s^2 +$$
$$+ (s/t)^2 \geq D^2\Psi^*(z^*)[h^*,h^*] + (s/t)^2. \quad (17)$$

In particular, Ψ is convex (and, of course, C^3) on H.

2^0. By (17), (15) and (5) we have

$$2\,\sup\{-D\Psi(z)[h] - \tfrac{1}{2}\,D^2\Psi(z)[h,h] \mid h \in \mathbb{R}\times R^n\} \leq 3. \quad (18)$$

3^0. Let

$$\xi = |D^3\Psi(z)[h,h,h]|,\ \rho^2 = D^2\Psi(z)[h,h] \equiv \rho_1^2 + \rho_2^2 + \rho_3^2,$$

where (see (17))

$$\rho_1^2 = D^2\Psi^*(z^*)[h^*,h^*],\quad \rho_2^2 = 2\,\tau_*\,(\tau_*^2 - x^Tx)^{-1}\,|\zeta''(t)|\,s^2,$$
$$\rho_3^2 = (s/t)^2, \quad (20)$$

and let us prove that

$$\xi \leq (9 + \alpha_\zeta^*)\,\rho^3,\ \alpha_\zeta^* = \min(\alpha_\zeta^{(1)},\alpha_\zeta^{(2)},\alpha_\zeta^{(3)}). \quad (19)$$

From (16) and the resulting inequality in (13) we get, with the help of Cauchy's inequality:

$$\xi \leq 2\,\tau_*\,(\tau_*^2 - x^Tx)^{-1}\cdot|\zeta'''(t)\,s^3| + 9\,\rho_1^3 +$$
$$+ 3\,|D^2\Psi^*(z^*)\,[(\zeta''(t)s^2,0),h^*]| + 2\,\rho_3^3 \leq$$
$$\leq 2\,\tau_*(\tau_*^2 - x^Tx)^{-1}\,|\zeta'''(t)\,s^3| + 9\,\rho_1^3 + 3(D^2\Psi^*(z^*)[h^*,h^*])^{1/2}$$
$$D^2\Psi^*(z^*)[(\zeta''(t)s^2,0),(\zeta''(t)s^2,0)])^{1/2} + 2\,\rho_3^3; \quad (20)$$

but

$$D^2\Psi^*(z^*)[(\zeta''(t)s^2,0),(\zeta''(t)s^2,0)] = \frac{\partial^2\Psi^*(z^*)}{\partial\tau^2}(\zeta''(t)s^2)^2 =$$
$$= 2\,(\tau_*^2 + x^Tx)(\tau_*^2 - x^Tx)^{-2}\,(\zeta''(t)s^2)^2 \leq$$
$$\leq 4\,\tau_*^2\,(\tau_*^2 - x^Tx)^{-2}\,(\zeta''(t)s^2)^2 \leq \rho_2^4$$

(we have taken into account that $\tau_*^2 \geq x^Tx$). Thus,

$$\xi \leq 2\,\tau_*\,(\tau_*^2 - x^Tx)^{-1}\,|\zeta'''(t)\,s^3| + 9\,\rho_1^3 + 3\,\rho_1\,\rho_2^2 + 2\,\rho_3^3. \quad (21)$$

Let us evaluate the quantity

$$\xi^* = 2\,\tau_*\,(\tau_*^2 - x^Tx)^{-1}\,|\zeta'''(t)\,s^3|.$$

Under the assumption that

$$\alpha_\zeta^{(1)} < \infty \tag{22}$$

we have $|\zeta'''(t)| \leq \alpha_\zeta^{(1)} \zeta'(t)|\zeta''(t)|\ \zeta^{-1}(t)$, whence, in view of $\tau_* = \zeta(t)$, one has

$$\xi^* \leq 2\ (\tau_*^2 - x^T x)^{-1}\ \zeta'(t)|\zeta''(t)|\ |s^3|\ \alpha_\zeta^{(1)} =$$
$$= \alpha_\zeta^{(1)}\ (\zeta'(t)s/\tau_*)\ (2\ \tau_*\ (\tau_*^2 - x^T x)^{-1}|\zeta''(t)|\ \ s^2)\ =$$
$$= \alpha_\zeta^{(1)}\ \ \rho_2^2\ |s_*|/\tau_*.$$

Thus, in the case of (22)

$$\xi^* \leq \alpha_\zeta^{(1)}\ \rho_2^2\ |s_*|/\tau_*. \tag{23}$$

By definition of h^* and by virtue of (2)

$$\rho_1^2 = 2\ (\tau_*^2 + x^T x)(\tau_*^2 - x^T x)^{-2}\ (s_* - 2\ \tau_*\ x^T u\ (\tau_*^2 + x^T x)^{-1})^2 +$$
$$+ 4\ (\tau_*^2 + x^T x)^{-1}\ (\tau_*^2 - x^T x)^{-1}\ (\ (x^T x)(u^T u) - (x^T u)^2) + 2\ (\tau_*^2 +$$
$$+ x^T x)^{-1}\ u^T u \geq 2\ \nu\ \gamma^{-1}\ (s_* - 2\ \tau_*\ \nu^{-1}\ x^T u)^2 + 2\nu^{-1}\ u^T u \equiv$$
$$\equiv \rho_{1,1}^2 + \rho_{1,2}^2, \tag{24}$$

where $\rho_{1,1} = (2\ \nu\ \gamma^{-2})^{1/2}\ |s_* - 2\ \tau_*\ \nu^{-1}\ x^T u|$,

$\rho_{1,2} = (2\ \nu^{-1}\ u^T u)^{1/2}$, $\gamma = \tau_*^2 - x^T x$, $\nu = \tau_*^2 + x^T x$.

(24) implies

$$|s_*| \leq 2\ \tau_*\ \nu^{-1}\ |x^T u| + \gamma\ (2\ \nu)^{-1/2}\ \rho_{1,1} \tag{25}$$

and, besides this,

$$|x^T u| \leq (x^T x)^{1/2}\ (u^T u)^{1/2} \leq (x^T x)^{1/2}\ (\nu/2)^{1/2}\ \rho_{1,2}. \tag{26}$$

Hence

$$|s_*|/\tau_* \leq 2\ \nu^{-1}\ (x^T x)^{1/2}\ (\nu/2)^{1/2}\ \rho_{1,2} +$$
$$+ \gamma\ \tau_*^{-1}\ (2\ \nu)^{-1/2}\ \ \rho_{1,1} \equiv \beta_1\ \rho_{1,1} + \beta_2\ \rho_{1,2}, \tag{27}$$

where $\beta_1 = \gamma\ \tau_*^{-1}\ (2\ \nu)^{-1/2}$, $\beta_2 = \ 2\ \ (x^T x)^{1/2}\ \ (2\ \nu)^{-1/2}$.

We have $\beta_1^2 + \beta_2^2 = (2\ \nu)^{-1}\ (\gamma^2\ \tau_*^{-2} + 4\ x^T x) = (2\ \nu)^{-1}$ $(\tau_*^4 - 2\ \tau_*^2\ x^T x + (x^T x)^2 + 4\ x^T x\ \tau_*^2)\ \tau_*^{-2} = (2\ \nu\ \ \tau_*^2)^{-1}\ \ (\tau_*^2 +$ $x^T x)^2 = (\tau_*^2 + x^T x)/(2\ \tau_*^2) \leq 1$,

so (26) and (24) lead to $|s_*|/\tau_* \leq \rho_1\ \rho_2^2\ \alpha_\zeta^{(1)}$.

Thus, in the case of (22) (see (23) and (21))

$$\xi \leqslant 9\,\rho_1^3 + 3\,\rho_1\,\rho_2^2 + \alpha_\zeta^{(1)}\,\rho_1\,\rho_2^2 + 2\,\rho_3^3. \tag{28}$$

Now assume that $\alpha_\zeta^{(2)} < \infty$. Then

$\xi^* \leqslant \alpha_\zeta^{(2)}\ |\zeta''(t)|^{3/2}\,\zeta^{-1/2}(t)\,2\,\tau_*\,(\tau_*^2 - x^Tx)^{-1}\,|s|^3 =$
$= \alpha_\zeta^{(2)}\,(|\zeta''(t)|\,s^2\,2\,\tau_*\,(\tau_*^2 - x^Tx)^{-1})^{3/2}((\tau_*^2 -$
$- x^Tx)\,\tau_*^{-1}\,\zeta^{-1}(t))^{1/2} \leqslant \alpha_\zeta^{(2)}\,\rho_2^3,$

since $\zeta(t) = \tau_*$, so $(\tau_*^2 - x^Tx)\,\tau_*^{-1}\,\zeta^{-1}(t) \leqslant 1$. Thus, in the case under consideration

$$\xi \leqslant 9\,\rho_1^3 + 3\,\rho_1\,\rho_2^2 + \alpha_\zeta^{(2)}\,\rho_2^3 + 2\,\rho_3^3. \tag{29}$$

Relations (28), (29) in view of $\rho_1^2 + \rho_2^2 + \rho_3^3 = \rho^2$ prove (19).

Now assume that

$$\alpha_\zeta^{(3)} < \infty. \tag{30}$$

In this case we have $\xi^* = 2\,\tau_*\,(\tau_*^2 - x^Tx)^{-1}\,|\zeta'''(t)\,s^3| \leqslant$
$\leqslant 2\,\tau_*\,(\tau_*^2 - x^Tx)^{-1}\,|\zeta''(t)|\,s^2\,|s/t|\,\alpha_\zeta^{(3)} = \alpha_\zeta^{(3)}\,\rho_2^2\,\rho_3,$
which results in $\xi \leqslant 9\,\rho_1^3 + 3\,\rho_1\,\rho_2^2 + 2\,\rho_3^3 + \alpha_\zeta^{(3)}\ \rho_2^2\ \rho_3$ and proves (19).

4^0. Let $c_\zeta = \max(9, 3 + \alpha_\zeta^*)/2$, $F(z) = c_\zeta^2\ \Phi(z)$. Then in view of (19):

$$|D^3F(z)[h,h,h]| \leqslant 2\,(D^2F(z)[h,h])^{3/2},\ z \in H,\ h \in \mathbb{R} \times R^n,$$

which together with the remarks from the beginning of the proof means that $F \in S_1^+(H, \mathbb{R} \times R^n)$. By virtue of (18) we have $\vartheta(F) \leqslant 2\,c_\zeta^2$. ■

3.9.5.3. (iii):

A) Let $G = \{(t,x) \in R^2 \mid t \geqslant 0,\ \zeta(t) \geqslant x\}$; then $H = \operatorname{int} G = \{(t,x) \in R^2 \mid t > 0,\ \zeta(t) > x\}$. Let

$$\Phi(t,x) = -\ln(t) - \ln(\zeta(t) - x) : H \to \mathbb{R}.$$

It is clear that $G \in C(R^2)$ and that Φ tends to ∞ as the argument approaches a boundary point of G. Moreover, it is clear that Φ is a C^3 - smooth on H.

1^0. First of all, let us verify that Φ is convex and satisfy the inequality ($\forall\ z = (t,x) \in H$):

$$2\sup\{-D\Phi(z)[h] - \tfrac{1}{2}\,D^2\Phi(z)[h,h] \mid h \in R^2\} \leqslant 2. \tag{1}$$

Indeed, the functions $\Phi(t,x) = t$ and $H(t,x) = \zeta(t) - x$ are concave, C^3-smooth and positive on H, so (1) is an immediate corollary of the following general statement.

Lemma 3.1. Let Φ be a C^3-smooth concave positive function on a convex open subset of $H \subset R^n$ and let $F(x) = -\ln(\Phi(x)): H \to \mathbb{R}$. Then F is a C^3-smooth convex function such that for each $x \in H$ we have

$$2\sup\{-DF(x)[h] - \frac{1}{2}D^2F(x)[h,h] \mid h \in R^n\} \leq 1.$$

Proof. We have $DF(x)[h] = -\Phi^{-1}(x)\, D\Phi(x)[h]$. Hence

$$D^2F(x)[h,h] = \Phi^{-2}(x)\,(D\Phi(x)[h])^2 - \Phi^{-1}(x)\,D^2\Phi(x)[h,h] \geq$$
$$\geq (DF(x)[h])^2,$$

Q.E.D. ■

2^0. Let $z = (t,x) \in H$, $h = (s,u) \in R^2$. We have

$$D\Psi(z)[h] = -s/t - (\zeta'(t)\,s - u)/(\zeta(t) - x), \tag{2}$$

$$D^2\Psi(z)[h,h] = s^2/t^2 + |\zeta''((t)|\,s^2/(\zeta(t) - x) +$$
$$+ (\zeta'(t)\,s - u)^2/(\zeta(t) - x)^2, \tag{3}$$

$$D^3\Psi(z)[h,h,h] = -2\,s^3/t^3 - \zeta'''(t)\,s^3/(\zeta(t) - x) +$$
$$+ 3\,\zeta''(t)\,s^2(\zeta'(t)\,s - u)/(\zeta(t) - x)^2 -$$
$$- 2\,(\zeta'(t)\,s - u)^3/(\zeta(t) - x)^3. \tag{4}$$

Let

$$\xi = |D^3\Psi(z)[h,h,h]|,\ \rho^2 = D^2\Psi(z)[h,h] = \rho_1^2 + \rho_2^2 + \rho_3^2,$$
$$\rho_1 = |s|/t,\ \ \rho_2^2 = |\zeta''(t)|\,s^2/(\zeta(t) - x),$$
$$\rho_3^2 = (\zeta'(t)\,s - u)^2/(\zeta(t) - x)^2. \tag{5}$$

Then

$$\xi \leq 2\,\rho_1^3 + 3\,\rho_2^2\,\rho_3 + 2\,\rho_3^3 + |\zeta'''(t)\,s^3|/(\zeta(t) - x),$$

which in view of $|\zeta'''(t)| \leq \alpha_\zeta^*\,|\zeta''(t)|/t$, $\alpha_\zeta^* = \alpha_\zeta - 1$, implies

$$\xi \leq 2\rho_1^3 + 3\rho_2^2\,\rho_3 + 2\rho_3^3 + (\alpha_\zeta^*\,|\zeta''(t)|\,s^2(\zeta(t) - x)^{-1})\,(|s|/t) \leq$$
$$\leq 2\,\rho_1^3 + 3\,\rho_2^2\,\rho_3 + 2\,\rho_3^3 + \alpha_\zeta^*\,\rho_2^2\,\rho_1,$$

hence, in view of (5),

$$|D^3\Psi(z)[h,h,h]| \leq (3 + \alpha_\zeta^*)\,(D^2\Psi(z)[h,h])^{3/2}. \tag{6}$$

Relations (6) and (1) in our standard manner (see subsect. 4^0 of 3.9.4.2) imply (111).A. ■

B). 1^0. We have $f'' \geq 0$, so $\Delta = (x^*,\infty)$, and, since $f''' \geq$

0, we have $f(\Delta) \equiv \Delta^* = (t^*, \infty)$ (we do not exclude that $x^* = -\infty$ and/or $t^* = -\infty$). Obviously, the inverse to $f|_\Delta$ function ζ: $\Delta^* \to \Delta$ is C^3-smooth, increases and is concave. If

$$G = \{(t,x) \in R^2 \mid t \geq f(x)\},$$

then $H \equiv \text{int}\, G = \{(t,x) \in R^2 \mid t > f(x)\}$. The function

$$\Phi(t,x) = -\ln(t - f(x)) - \ln(\zeta(t) - x): H \to \mathbb{R}$$

obviously is C^3 - smooth and tends to ∞ as the argument approaches a boundary point of G. L.3.1 implies that Φ is convex on H and satisfies the condition

$$2 \sup\{-D\Phi(z)[h] - \frac{1}{2} D^2\Phi(z)[h,h] \mid h \in R^2\} \leq 2. \qquad (7)$$

To complete in our standard manner the proof of (111).B, it is necessary to establish an inequality of the form

$$|D^3\Phi(z)[h,h,h]| \leq O((2 - \lambda)^{-1}) \, (D^2\Phi(z)[h,h])^{3/2} \qquad (8)$$

for all $z \in H$ and $h \in R^2$. This will be done in what follows.

2^0. Let us fix $z = (t,x) \in H$ and $h = (s,u) \in R^2$. We have

$$\rho^2 \equiv D^2\Phi(z)[h,h] = f''(x)\, u^2/(t - f(x)) + (f'(x)\, u - s)^2/(t - f(x))^2 + |\zeta''(t)|\, s^2/(\zeta(t) - x) + (\zeta'(t)\, s - u)^2/(\zeta(t) - x)^2 = \rho_1^2 + \rho_2^2 + \rho_3^2 + \rho_4^2,\ \rho_1^2 = f''(x)\, u^2/(t - f(x)),$$

$$\rho_2^2 = (f'(x)\, u - s)^2/(t-f(x))^2, \quad \rho_3^2 = |\zeta''(t)| s^2/(\zeta(t)-x),$$

$$\rho_4^2 = (\zeta'(t)s-u)^2/(\zeta(t) - x)^2, \qquad (9)$$

$$D^3\Phi(z)[h,h,h] = f'''(x)\, u^3/(t-f(x)) - \zeta'''(t)\, s^3/(\zeta(t)-x) + 3 f''(x)\, u^2\, (f'(x)\, u - s)/(t - f(x))^2 + 2\, (f'(x)\, u - s)^3/(t - f(x))^3 + 3\, |\zeta''(t)|\, s^2(\zeta'(t)\, s - u)/(\zeta(t) - x)^2 - 2\, (\zeta'(t)s - u)^3/(\zeta(t) - x)^3. \qquad (10)$$

(9) and (10) imply that

$$\xi \equiv |D^3\Phi(z)[h,h,h]| \leq |f'''(x)u^3|/(t - f(x)) + |\zeta'''(t)s^3|/(\zeta(t) - x) + 3\, \rho_1^2\, \rho_2 + 2\, \rho_2^3 + 3\, \rho_3^2\, \rho_4 + 2\rho_4^3. \qquad (11)$$

3^0. Let $y = \zeta(t)$, thus $t = f(y)$, $y \in \Delta$, $f'(y) > 0$. Let also $\sigma = s/f'(y)$. Let

$$\delta_1 = |f'''(x)\, u^3|/(t - f(x)) = |f'''(x)\, u^3|/(f(y) - f(x)),$$

$$\delta_2 = |\zeta'''(t)\, s^3|/(\zeta(t)-x) = |\zeta'''(t)\, (f'(y))^3\, \sigma^3|/(y - x). \qquad (12)$$

Notice that

$$x < y;\ f(y) > f(x);\ f'(y) > 0. \tag{13}$$

4^0. In view of the correspondence between f and ζ we have:

$$\zeta'(t) = 1/f'(y);\quad \zeta''(t) = -f''(y)/(f'(y))^3;$$
$$\zeta'''(t) = -f'''(y)/(f'(y))^4 + 3\,(f''(y))^2/(f'(y))^5.$$

Hence (see (9) - (13))

$$\rho_1^2 = f''(x)\,u^2/(f(y) - f(x));$$
$$\rho_2^2 = (f'(x)\,u - f'(y)\,\sigma)^2/(f(y) - f(x))^2;$$
$$\rho_3^2 = f''(y)\,\sigma^2(f'(y))^{-1}/(y - x);$$
$$\rho_4^2 = (\sigma - u)^2/(y - x)^2; \tag{14}$$
$$\delta_1 = |f'''(x)\,u^3|/(f(y) - f(x)); \tag{15}$$
$$\delta_2 = |f'''(y)\,\sigma^3|\,(f'(y))^{-1}\,(y - x)^{-1} +$$
$$+ 3\,|(f''(y))^2\,\sigma^3|\,(f'(y))^{-2}\,(y - x)^{-1} = \delta_{2,1} + 3\,\delta_{2,2},$$
$$\delta_{2,1} = |f'''(y)\ \sigma^3|\,(f'(y))^{-1}\,(y - x)^{-1},$$
$$\delta_{2,2} = |(f''(y))^2\,\sigma^3|\,(f'(y))^{-2}\,(y - x)^{-1}. \tag{16}$$

By the assumption of the statement under consideration

$$0 \leqslant f'''(y) \leqslant \lambda\,(f''(y))^2\,(f'(y))^{-1},$$

hence

$$\delta_2 \leqslant (3 + \lambda)\,\delta_{2,2}. \tag{17}$$

5^0. Assume first that

$$f''(x) > 0;$$

then $f'(x) > 0$.

By virtue of (14) we have:

$$|u| \leqslant \rho_1((f(y) - f(x))/f''(x))^{1/2},$$
$$|\sigma| \leqslant \rho_2(f(y) - f(x))/f'(y) + f'(x)\,(f'(y))^{-1}\,\rho_1((f(y) -$$
$$- f(x))/f''(x))^{1/2} = \rho_2\,(f(y) - f(x))/f'(y) +$$
$$+ \rho_1\,((f'(x))^2\,(f(y) - f(x))\,(f'(x))^{-2}\,(f''(x))^{-1})^{1/2},$$

whence

$$f''(y)\,|\sigma|\,(f'(y))^{-1} \leqslant \rho_2\,f''(y)\,(f(y) - f(x))\,(f'(y))^{-2} +$$
$$+ \rho_1\,f''(y)\,(f'(y))^{-1}\,((f'(x))^2\,(f(y) -$$
$$- f(x))\,(f'(x))^{-2}\,(f''(x))^{-1})^{1/2}. \tag{18}$$

7^0. For $\omega \in \Delta$, under the assumption that $f''(y) \neq 0$ and independently on the positiveness or equality to 0 of the quantity $f''(x)$, we have by definition of λ:

$$\frac{d}{d\omega}\,(f''(\omega)\,(f'(\omega))^{-\lambda}) \leqslant 0.$$

Hence for $A = f''(y)\,(f'(y))^{-\lambda}$ and $x \leqslant \omega \leqslant y$ the inequality

$$f''(\omega)\quad(f'(\omega))^{-\lambda} \geqslant A,$$

or

$$-(\lambda - 1)^{-1}\,\frac{d}{d\omega}\,((f'(\omega))^{-(\lambda-1)}) \geqslant A,$$

holds. Thus

$$(f'(\omega))^{1-\lambda} \geqslant (f'(y))^{1-\lambda} + (\lambda - 1)\,A\,(y - \omega),$$

whence

$$f'(\omega) \leqslant f'(y)\,(1 + (\lambda - 1)\,A\,(y - \omega)\,(f'(y))^{\lambda - 1})^{1/(\lambda-1)}$$

for $x \leqslant \omega \leqslant y$. After integration over $\omega \in [x,y]$, we establish the implication

$$f''(y) \neq 0 \Rightarrow$$

$$f(y) - f(x) \leqslant (f'(y))^2\,(f''(y))^{-1}\,c(\lambda),\ c(\lambda) = \int_0^\infty (1 + (\lambda - 1)\,\omega)^{-1/(\lambda-1)}\,d\omega. \tag{19}$$

(18) and (19) under the assumption that $f''(y) \neq 0$ imply

$$f''(y)\,|\sigma|/f'(y) \leqslant \rho_2\,c(\lambda) + \rho_1\,c^{1/2}(\lambda)\,(f''(y)/f''(x))^{1/2}\,(f'(x)/f'(y)). \tag{20}$$

Moreover, $f''(x) \geqslant A\,(f'(x))^{\lambda} = f''(y)\quad(f'(x)/f'(y))^{\lambda}$, which together with (20) leads to

$$f''(y)\,|\sigma|/f'(y) \leqslant \rho_2 c(\lambda) + \rho_1 c^{1/2}(\lambda)\,(f'(x)/f'(y))^{1-\lambda/2} \leqslant \leqslant \rho_2\,c(\lambda) + \rho_1\,c^{1/2}(\lambda) \tag{22}$$

(we have taken into account that $1 - \lambda/2 \geqslant 0$ и $f'(x) \leqslant f'(y)$).

In view of (22), (14), (16) and (17) we get in the case of $f''(x) \neq 0$:

$$\delta_2 \leqslant (3 + \lambda)\,(\rho_2\,c(\lambda) + \rho_1\,c^{1/2}(\lambda))\,\rho_3^2. \tag{23}$$

8^0. Now assume that $f''(x) = 0$. Let us prove that (23) is true in this situation, too.

a). It is possible that $f'(x) = 0$. Then, by (14),

$$|\sigma| \leqslant \rho_2 \,(f(y) - f(x))/f'(y).$$

If $f''(y) \neq 0$, then (19) holds, and we have

$$f''(y)\,|\sigma|\,(f'(y))^{-1} \leqslant \rho_2\, c(\lambda),$$

which together with (14) and (16) implies $\delta_{2,2} \leqslant \rho_3^2 \;\; \rho_2 \;\; c(\lambda)$; the latter together with (17) immediately leads to (23): If $f''(y) = 0$, then (23) is obvious since in this situation $f'''(y) = 0$.

b). Now assume that $f''(x) = 0$, $f'(x) \neq 0$. Then, in view of the condition of the statement under consideration,

$$0 \leqslant f'''(\omega) \leqslant 2\,(f''(\omega))^2/f'(\omega),\ \omega \geqslant x,$$

or, for $g(\omega) = f''(\omega)$:

$$0 \leqslant g'(\omega) \leqslant 2\, g^2(\omega)/f'(x),\ g(x) = 0$$

(we have taken into account that $f'(\omega) \geqslant f'(x)$ for $\omega \geqslant x$). The resulting inequality implies $g(\omega) \equiv 0$, $\omega \geqslant x$, so $f''(y) = 0$, $\delta_2 = 0$ and (23) holds.

9^0. Now let us evaluate δ_1. One can assume that $f'(x) > 0$, $f''(x) > 0$ – otherwise $f'''(x) = 0$, and hence $\delta_1 = 0$. Thus, we assume that

$$f'(x) > 0,\quad f''(x) > 0. \tag{24}$$

By the condition $f'''(x) \leqslant 2\,(f''(x))^2/f'(x)$, which implies

$$\delta_1 \leqslant 2\,(f''(x))^2\;|u|^3\;(f'(x))^{-1}\;(f(y) - f(x))^{-1}. \tag{25}$$

By virtue of the conditions on f we have:

$$x \leqslant \omega \leqslant y \;\Rightarrow\; \frac{d}{d\omega}\,(f'(\omega)/f''(\omega)) \leqslant 1/2. \tag{26}$$

Assume first that

$$f'(y)/f''(y) \geqslant y - x. \tag{27$_a$}$$

Then (26) implies

$$y - x \leqslant f'(y)/f''(y) \leqslant (f'(x)/f''(x)) + (y - x)/2,$$

so $y - x \leqslant 2\, f'(x)/f''(x)$ and $f'(y)/f''(y) \leqslant 2\; f'(x)/f''(x)$. Hence, in view of (14),

$$|\sigma| \leqslant \rho_3\,(f'(y)\,(y - x)/f''(y))^{1/2} \leqslant 2\,\rho_3\;\, f'(x)/f''(x),$$

which by the same argument gives us

$$|u| \leqslant 2\,(\rho_4 + \rho_3)\, f'(x)/f''(x).$$

Therefore (25) and (14) imply

$$\delta_1 \leqslant 2\,(f''(x)\,u^2/(f(y) - f(x)))\,(f''(x)\,|u|/f'(x)) \leqslant$$
$$\leqslant 4\,\rho_1^2\,(\rho_4 + \rho_3). \qquad (28_a)$$

Now assume that

$$f'(y)/f''(y) < y - x. \qquad (27_b)$$

Then, by (14), $|\sigma| \leqslant \rho_3\,(f'(y)\,(y - x)/f''(y))^{1/2}$, so (14) and (27_b) lead to

$$|u| \leqslant \rho_4(y - x) + \rho_3(f'(y)\,(y - x)/f''(y))^{1/2} \leqslant (\rho_3+\rho_4)(y - x).$$

Thus, (25) implies

$$\delta_1 \leqslant 2\,(f''(x))^2\,(f'(x))^{-1}\,(f(y) - f(x))^{-1}(\rho_3 + \rho_4)^3\,(y - x)^3 =$$
$$= 2\,(\rho_3 + \rho_4)^3\,S,$$
$$S = (f''(x))^2\,(f'(x))^{-1}\,(f(y) - f(x))^{-1}\,(y - x)^3. \qquad (29_b)$$

For $x \leqslant \omega \leqslant y$ we, in view of the conditions on f, have:

$$(f'''(\omega)/f''(\omega)) \geqslant (f''(\omega)/f'(\omega))/2,$$

or $(\ln(f''(\omega)))' \geqslant \frac{1}{2}\,(\ln(f'(\omega)))'$, whence

$$f''(\omega)/f''(x) \geqslant (f'(\omega)/f'(x))^{1/2}, \quad \omega \geqslant x. \qquad (30_b)$$

Let $\pi = f''(x)\ (f'(x))^{-1/2}$, $g(\omega) = f'(\omega)$. We have

$$g'(\omega)\,(g(\omega))^{-1/2} \geqslant \pi, \text{ or } g^{1/2}(\omega) \geqslant g^{1/2}(x) + \pi\,(\omega - x)/2,$$

whence

$$f(y) - f(x) \geqslant \int_0^{y-x} ((f'(x))^{1/2} + \pi\,t/2)^2\,dt =$$
$$= f'(x)\,(y - x) + \pi\,(f'(x))^{1/2}\,(y - x)^2/4 + \pi^2\,(y - x)^3/12 \geqslant$$
$$= \pi^2\,(y - x)^3/12. \qquad (31_b)$$

Therefore

$$S \leqslant (f''(x))^2\,(y - x)^3\,(f'(x))^{-1}\,(12\,(y - x)^3/\pi^2) = 12$$

(we have taken into account the definition of π). Thus, (29_b) implies

$$\delta_1 \leqslant 24\,(\rho_3 + \rho_4)^3. \qquad (32_b)$$

Implications $(27_a) \to (28_a)$, $(27_b) \to (32_b)$ together with (23) imply

$$\delta_1 + \delta_2 \leqslant O(1 + c(\lambda))\,\rho^3 \qquad (33)$$

with an absolute constant in $O(\ )$. (33), (11) and (12) lead to inequality $\xi \leqslant O(1 + c(\lambda))\ \rho^3$. The latter relation by definition of $c(\lambda)$ coincides with (8). ■

3.9.5.4. (iv):

It is known that the function $(\mathrm{Det}\ x)^{1/n}$: $S_n^+ \to \mathbb{R}$ is concave. Therefore the function $F(x) = -\ln(\mathrm{Det}\ x)$: $\mathrm{int}\ S_n^+ \to \mathbb{R}$ is convex, C^3 - smooth and satisfies the relation

$$2 \sup\{- DF(x)[h] - \frac{1}{2} D^2F(x)[h,h] \mid h \in S_n\} \leqslant n \quad (1)$$

(L.3.1). Moreover, this function tends to ∞ as the argument approaches a point from ∂S_n^+. Therefore it suffices to verify the inclusion

$$F \in S_1(\mathrm{int}\ S_n^+, S_n). \quad (2)$$

For $x \in \mathrm{int}\ S_n^+$, $h \in S_n$ we have: $DF(x)[h] = - \frac{d}{dt}\big|_{t=0} \ln(\mathrm{Det}\ (x + t\ h)) = - \frac{d}{dt}\big|_{t=0} (\ln(\mathrm{Det}\ x) + \ln(\mathrm{Det}\ (I + t\ x^{-1}\ h))) = - \mathrm{Tr}\{x^{-1}h\}$, $D^2F(x)[h,h] = - \frac{d}{dt}\big|_{t=0} \mathrm{Tr}\{(x + th)^{-1}h\} = \mathrm{Tr}\{x^{-1}\ h\ x^{-1}\ h\}$, $D^3F(x)[h,h,h] = \frac{d}{dt}\big|_{t=0} \mathrm{Tr}\{(x + th)^{-1}\ h\ (x + th)^{-1}\ h\} = - 2\ \mathrm{Tr}(x^{-1}\ h\ x^{-1}\ h\ x^{-1}\ h)$.

Let $h^* = x^{-1/2}\ h\ x^{-1/2}$; then

$D^2F(x)[h,h] = \mathrm{Tr}\{(h^*)^2\}$, $|D^3F(x)[h,h,h]| = 2\ |\mathrm{Tr}\{(h^*)^3\}|$.

In other words, if $\{\nu_i\}$ denote the eigenvalues of the (symmetric) matrix h^*, then

$$|D^3F(x)[h,h,h]| \leqslant 2 \sum_{i=1}^{n} |\nu_i|^3 \leqslant 2(\sum_{i=1}^{n} \nu_i^2)^{3/2} = 2\ (D^2F(x)[h,h])^{3/2},$$

which leads to (2). ■

3.9.4.5. (v):

Let

$F_{m,n}(t,x) = - \ln(\mathrm{Det}\ (t^2\ I_n - x^T\ x))$:

$H = \{(t,x) \in \mathbb{R} \times L_{m,n} \mid t > |x|\} \to \mathbb{R}$.

We shall prove that in the case of $m \geqslant n$ under an appropriate choice of absolute constant factors in the $O(\ )$ which follows the function $O(1)\ F_{m,n}(x)$ is an $O(n)$ - self-concordant barrier for

$G = \mathrm{Cl}\, H = \{(t,x) \in \mathbb{R} \times L_{m,n} \mid t \geq |x|\}.$

First of all, let us verify that this fact implies (v) as a whole. Indeed, if $m < n$, then $L_{m,n}$ is in a natural sense imbedded into $L_{n,n}$ (we add to matrices $(n-m)$ zero rows); this imbedding preserves the matrix norm and therefore can be embedded to a linear imbedding of the epigraph of the matrix norm on the first space into similar epigraph on the second space. The latter imbedding induces an "inverse" transformation of barriers, which transforms $\lambda\ F_{n,n}$ into $\lambda\ F_{m,n}$, so the second function is a self-concordant barrier, if the first is. In other words, our hypothesis implies that $O(1)$ $F_{m,n}$ is an $O(n)$ - self-concordant barrier for the epigraph of the matrix norm on $L_{m,n}$ for all m and n. In view of the natural isometric isomorphism between $L_{m,n}$ and $L_{n,m}$ the latter statement is equivalent to (v).

Thus, from now on we assume that $m \geq n$ and consider the function

$F(t,x) = -\ln(\mathrm{Det}\ (t^2 I_n - x^T x))$:

$H = \{(t,x) \in \mathbb{R} \times L_{m,n} \mid t > |x|\} \to \mathbb{R}.$

Let also

$G = \mathrm{Cl}\, H = \{(t,x) \in \mathbb{R} \times L_{m,n} \mid t \geq |x|\}.$

For the sake of brevity we write E instead of $L_{m,n}$ and E^* instead of $\mathbb{R} \times L_{m,n}$. Below Greek capitals denote $n \times n$ -matrices.

1^0. It is clear that F tends to ∞ as the argument belonging to H approaches a boundary point of G.

2^0. Let us fix $z^* = (t^*,x^*) \in H$ and $h = (s,u) \in E^*$, such that

$$\mathrm{Det}\ (x^*)^T x^* > 0. \tag{1}$$

Let us derive the expressions for $DF(z^*)[h]$, $D^2F(z^*)[h,h]$, $D^3F(z^*)[h,h,h]$. Let

$$\Omega = ((t^*)^2 I_n - (x^*)^T x^*)^{-1/2} \tag{2}$$

and

$$J(t,x) = -\ln(\mathrm{Det}\ \Omega\ (t^2 I_n - x^T x)\ \Omega) = F(t,x) + \mathrm{const}. \tag{3}$$

Let for $z = (t,x) \in H$

$$Q(z) = \Omega\ (t^2 I_n - x^T x)\ \Omega; \tag{4}$$

then

$$Q(z^*) = I_n, \quad J(z) = -\ln(\mathrm{Det}\ Q(z)). \tag{5}$$

We have

$$DF(z^*)[h] = DJ(z^*)[h] = -\mathrm{Tr}\{Q^{-1}(z^*)\ DQ(z^*)[h]\}, \tag{6}$$

$$D^2F(z^*)[h,h] = D^2J(z^*)[h,h] =$$
$$= \mathrm{Tr}\{Q^{-1}(z^*)\ DQ(z^*)[h]\ Q^{-1}(z^*)\ DQ(z^*)[h]\} -$$
$$- \mathrm{Tr}\{Q^{-1}(z^*)\ D^2Q(z^*)[h,h]\}, \tag{7}$$

$$D^3F(z^*)[h,h,h] = -2\ \mathrm{Tr}\{(Q^{-1}(z^*)\ DQ(z^*)[h])^3\} +$$
$$+ 3\ \mathrm{Tr}\{Q^{-1}(z^*)\ D^2Q(z^*)[h,h]\ Q^{-1}(z^*)\ DQ(z^*)[h]\} \tag{8}$$

(we have taken into account that $D^3Q(z^*) = 0$).

In view of (5) we get from (6) - (8)

$$DF(z^*)[h] = -\mathrm{Tr}\{\Omega(2\ t^*\ s\ I_n - (x^*)^T u - u^T x^*)\Omega\} =$$
$$= -2\ \mathrm{Tr}\{t^*\ s\ \Omega^2 - (u\Omega)^T\ (x^*\Omega)\}, \tag{9}$$

$$D^2F(z^*)[h,h] = \mathrm{Tr}\{(2\ t^*\ s\ \Omega^2 - (u\Omega)^T(x^*\Omega) - (x^*\Omega)^T(u\Omega))^2\}, \tag{10}$$

$$D^3F(z^*)[h,h,h] = -2\ \mathrm{Tr}\{(2\ t^*\ s\ \Omega^2 - (u\Omega)^T\ (x^*\Omega) -$$
$$- (x^*\Omega)^T\ (u\Omega))^2\} + 3\ \mathrm{Tr}\{(2\ s^2\ \Omega^2 - 2\ (u\Omega)^T(u\Omega))\ (2\ t^*\ s\ \Omega^2 -$$
$$- (u\Omega)^T(x^*\Omega) - (x^*\Omega)^T\ (u\Omega))\}. \tag{11}$$

3^0. Let

$$\sigma = s/t^*; \quad \nu = u\,\Omega; \quad t^*\,\Omega = \theta; \quad x^*\,\Omega = \xi, \tag{12}$$

which, in view of (4), (5), implies

$$\theta^2 = I_n + \xi^T\,\xi = I_n + \mathrm{P}. \tag{13}$$

Furthermore, let

$$\nu^T\,\xi = \Pi + \Phi, \quad \Pi = \Pi^T, \quad \Phi = -\Phi^T; \tag{14}$$

and

$$P_m = \xi\ (\xi^T\xi)^{-1}\ \xi^T \in L_{m,m}; \tag{15}$$

then P_m is an orthoprojector of rank n, such that

$$(\Pi + \Phi)\ \mathrm{P}^{-1}\ (\Pi + \Phi)^T = \nu^T\ P_m\ \nu, \tag{16}$$

thus

$$\nu^T\,\nu = (\Pi + \Phi)\ \mathrm{P}^{-1}\ (\Pi + \Phi)^T + \nu^T\ (I_m - P_m)\ \nu. \tag{17}$$

4^0. Now we can rewrite (9) - (11) as

$$DF(z^*)[h] = -2\ \mathrm{Tr}\{\sigma\ (I_n + \mathrm{P}) - \Pi\}, \tag{18}$$

$$D^2F(z^*)[h,h] = 2\ \mathrm{Tr}\{\sigma^2\ (I_n + \mathrm{P})\ (I_n + 2\ \mathrm{P}) - 4\ \sigma\ \Pi\ (I_n + \mathrm{P}) +$$

$+ 2\,\Pi^2 + (\Pi + \Phi)\, P^{-1}\, (\Pi + \Phi)^T) + 2\, Tr(v^T\, (I_m - P_m)\, v).$ (19)

5^0. Expressions (9) - (11) for the differentials of F does not vary under the substitution $x \to x\,U$, $u \to u\,U$, where U is an orthogonal $n \times n$ - matrix. Therefore below we, without loss of generality, can assume that

$$P = Diag(\lambda_1^{-1}, \dots, \lambda_n^{-1}), \quad \lambda_i > 0, \quad 1 \leqslant i \leqslant n. \qquad (20)$$

Consider the function

$$K(E) = Tr((\Pi + E)\, P^{-1}\, (\Pi + E)^T) \qquad (21)$$

on the space C skew-symmetric $n \times n$ - matrices. This is a bounded from below (and hence attaining its minimum over C) quadratic form; let E^* be the minimizer of this form. Then

$$(\forall\, E \in C):\ Tr((\Pi + E + E^*)\, P^{-1}\, (\Pi + E + E^*)^T) \geqslant$$
$$\geqslant Tr((\Pi + E^*)\, P^{-1}\, (\Pi + E^*)^T),$$

or, taking the derivative in E:

$$Tr(\, E\, P^{-1}\, (\Pi + E^*)^T) = 0\ \forall\, E \in C,$$

whence $(P^{-1}(\Pi + E^*)^T) = (P^{-1}(\Pi + E^*)^T)^T$, or $P^{-1}(\Pi - E^*) = (\Pi + E^*)\, P^{-1}$, thus $P^{-1}\,\Pi - \Pi\, P^{-1} = P^{-1}\, E^* + E^*\, P^{-1}$, or, which is the same,

$$(\lambda_i - \lambda_j)\, \Pi_{ij} = (\lambda_i + \lambda_j)\, E^*_{ij}. \qquad (22)$$

Let Π^* be the symmetric matrix which is produced from Π when the diagonal entries are replaced by zeros. Then (22) means that

$$E^*_{ij} = (\lambda_i - \lambda_j)(\lambda_i + \lambda_j)^{-1}\, \Pi^*_{ij}. \qquad (23)$$

Let Δ be a diagonal matrix coinciding with the diagonal of Π. Then

$$(\Pi + E^*)_{ij} = 2\,\lambda_i\, (\lambda_i + \lambda_j)^{-1}\, \Pi_{ij} =$$
$$= \Delta_{ii} + 2\,\lambda_i\, (\lambda_i + \lambda_j)^{-1}\, \Pi^*_{ij}.\ (\Pi^* + E^*)_{ij} =$$
$$= 2\,\lambda_i\, (\lambda_i + \lambda_j)^{-1}\, \Pi^*_{ij}. \qquad (24)$$

6^0. We have

$$K(\Phi) = K(E^*) + Tr((\Phi - E^*)\, P^{-1}\, (\Phi - E^*)^T), \qquad (25)$$

so (19) can be rewritten as

$$\frac{1}{2}\, D^2F(z^*)[h,h] = d + d_5 + d_6. \qquad (26)$$

where

$$d = \mathrm{Tr}(\sigma^2 (I_n + P)(I_n + 2P) - 4\sigma \Pi (I_n + P) + 2\Pi^2 + (\Pi + E^*) P^{-1} (\Pi + E^*)^T), \tag{27}$$

$$d_5 = \mathrm{Tr}((\Phi - E^*) P^{-1} (\Phi - E^*)^T) \geq 0, \tag{28}$$

$$d_6 = \mathrm{Tr}(\nu^T (I_m - P_m) \nu) \geq 0. \tag{29}$$

Let q_i be defined by the relation

$$\Pi_{ii} = (2 + \lambda_i)^{-1} (2 (1 + \lambda_i^{-1}) \sigma + q_i); \tag{30}$$

then the first sequence of equalities in (24) implies that

$$d = \sum_{i=1}^{n} \Big\{ \sigma^2 (1 + \lambda_i^{-1})(1 + 2\lambda_i^{-1}) - 4\sigma (1 + \lambda_i^{-1}) (2 + \lambda_i)^{-1} (2 (1 + \lambda_i^{-1}) \sigma + q_i) + 2 (2 + \lambda_i)^{-2} (2 (1 + \lambda_i^{-1}) \sigma + q_i)^2 + 2 \sum_{j=1}^{n} (\Pi^*_{ij})^2 + \lambda_i \Pi^2_{ii} + 4 \sum_{j=1}^{n} \lambda_i^2 \lambda_j (\lambda_i + \lambda_j)^{-2} (\Pi^*_{ij})^2 \Big\} = \sum_{i=1}^{n} \Big\{ \sigma^2 (1 + \lambda_i)(2 + \lambda_i)^{-1} + q_i^2 (2 + \lambda_i)^{-1} + 2 \sum_{j=1}^{n} (\Pi^*_{ij})^2 + \lambda_i \Pi^2_{ii} + 4 \sum_{j=1}^{n} \lambda_i^2 \lambda_j (\lambda_i + \lambda_j)^{-2} (\Pi^*_{ij})^2 \Big\}. \tag{31}$$

In particular, $d \geq 0$; hence, in view of (26), (28), (29), it follows that $D^2F(z^*)[h,h] \geq 0$ for each $h \in E^*$ and each $z^* \in H$ satisfying (1). By continuity arguments we have $D^2F(z^*)[h,h] \geq 0$ for each $z^* \in H$ and $h \in E^*$, thus F is convex (and, of course, C^3 - smooth) on H.

7^0. Let

$$d_1 = \sum_{i=1}^{n} \sigma^2 (1 + \lambda_i)(2 + \lambda_i)^{-1}; \quad d_2 = \sum_{i=1}^{n} q_i^2 (2 + \lambda_i)^{-1};$$

$$d_3 = \sum_{i=1}^{n} \sum_{j=1}^{n} (\Pi^*_{ij})^2; \quad d_4 = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i^2 \lambda_j (\lambda_i + \lambda_j)^{-2} (\Pi^*_{ij})^2, \tag{32}$$

so in view of (26) - (29)

$$\frac{1}{2} D^2F(z^*)[h,h] = d_1 + d_2 + 2 d_3 + 4 d_4 + d_5 + d_6 \tag{33}$$

and each d_i is nonnegative.

8^0. Let us verify that

$$\gamma \equiv - D^2F(z^*)[h,h] - 2 DF(z^*)[h] \leq 2n. \tag{34}$$

Indeed, in view of (18) and (28), (29), (32), (33) we have

$$\gamma \leq 2 \sup \Big\{ \sum_{i=1}^{n} \Big\{ 2\sigma (1 + \lambda_i) \lambda_i^{-1} - 2 (2\sigma (1 + \lambda_i^{-1}) + q_i) (2 +$$

$+ \lambda_i)^{-1} - \sigma^2 (1 + \lambda_i) (2 + \lambda_i)^{-1} - q_i^2 (2 + \lambda_i)^{-1} \mid \sigma, q_1, \ldots, q_n$

$\in \mathbb{R} \}\} \leq 2 (\sum_{i=1}^{n} ((1 + \lambda_i)(2 + \lambda_i)^{-1} + (2 + \lambda_i)^{-1})) = 2n,$

which implies (34).

Thus, (34) holds for all $h \in E^*$ and all $z^* \in H$ satisfying (1); by continuity arguments (34) holds for all $h \in E^*$ and $z^* \in H$.

9^0. Now let us prove that under an appropriate choice of an absolute constant $O(1)$ identically in $z \in H$ and $\eta \in E^*$ one has

$$\xi \equiv |D^3F(z)[\eta,\eta,\eta]| \leq O(1)\, \rho^3, \quad \rho \equiv DF(z)[\eta,\eta]; \tag{35}$$

this relation together with (34) and the remarks from $1.^0$ proves in our standard manner the required statement.

By continuity arguments it suffices to prove that, under an appropriate choice of $O(1)$ relation (35) holds identically in $\eta \in E^*$ and in $z = (t,x) \in H$, such that $\mathrm{Det}\, x^T x > 0$. So we can deal with (35) for $z = z^*$, $\eta = h$; we shall use our previous notations.

By (11) we have

$$\xi \leq 16\, |\mathrm{Tr}\{\Xi^3\}| + 12\, |\mathrm{Tr}\{\Xi\, (\sigma^2 (I_n + P) - v^T v)\}|,$$
$$\Xi = \sigma(I_n + P) - \Pi \ (= \Xi^T). \tag{36}$$

Let for a matrix A $|A|$ means the operator norm, and $|A|_2$ the Hilbert-Schmidt norm. Recall that if A', A are such matrices that $A'\, A''$ is well defined then

$$|A'\, A''|_2 \leq |A'|\, |A''|_2, \quad |(A')^T|_2 = |A'|_2, \tag{37}$$

and if A', A'' are $k \times l$ - matrices, then

$$|\mathrm{Tr}(A'\, (A'')^T)| \leq |A'|_2\, |A''|_2; \tag{38}$$

notice that for each matrix A one has

$$|A| \leq |A|_2, \tag{39}$$

and for positive semidefinite matrix A one has

$$|A|_2 \leq \mathrm{Tr}(A). \tag{40}$$

10^0. Let us verify that

$$|\mathrm{Tr}\{\Xi^3\}| \leq (d_1 + d_2 + d_3)^{3/2}. \tag{41}$$

Indeed, S is a symmetric $n \times n$ - matrix; therefore

$$|\mathrm{Tr}(S^3)| \leqslant (\mathrm{Tr}(S^2))^{3/2},$$

and to prove (41) it suffices to establish that

$$\mathrm{Tr}(S^2) = |S|_2^2 \leqslant d_1 + d_2 + d_3. \tag{42}$$

In view of (30), (32) we have

$$\mathrm{Tr}(S^2) = \mathrm{Tr}(\sigma^2 (I_n + P)^2 - \sigma (I_n + P)\Pi - \sigma \Pi (I_n + P) + \Pi^2) =$$
$$= \sum_{i=1}^{n} \Big\{\sigma^2 (1 + \lambda_i)^2 \lambda_i^{-2} - 2\sigma (1 + \lambda_i) \lambda_i^{-1} \Pi_{ii} + \Pi_{ii}^2 +$$
$$+ \sum_{j=1}^{n} (\Pi_{ij}^*)^2\Big\} = \sum_{i=1}^{n} \Big\{\sigma^2 (1 + \lambda_i)^2 \lambda_i^{-2} - 4\sigma^2 (1 + \lambda_i)^2 \lambda_i^{-2} (2 +$$
$$+ \lambda_i)^{-1} - 2\sigma (1 + \lambda_i) \lambda_i^{-1} (2 + \lambda_i)^{-1} q_i + 4\sigma^2 (1 + \lambda_i)^2 \lambda_i^{-2}$$
$$(2 + \lambda_i)^{-2} + 4\sigma (1 + \lambda_i) \lambda_i^{-1} (2 + \lambda_i)^{-2} q_i + q_i^2 (2 + \lambda_i)^{-2} +$$
$$+ \sum_{j=1}^{n} (\Pi_{ij}^*)^2\Big\} = \sum_{i=1}^{n} \Big\{\sigma^2 (1 + \lambda_i)^2 (2 + \lambda_i)^{-2} - 2\sigma (1 + \lambda_i) (2 +$$
$$+ \lambda_i)^{-2} q_i + q_i^2 (2 + \lambda_i)^{-2}\Big\} + d_3 \leqslant \sum_{i=1}^{n} \Big\{\sigma^2 (1 + \lambda_i)^2 + q_i^2 + (1 +$$
$$+ \lambda_i) q_i^2 + (1 + \lambda_i) \sigma^2\Big\} (2 + \lambda_i)^{-2} + d_3 \leqslant$$
$$\leqslant \sum_{i=1}^{n} \Big\{\sigma^2 (1 + \lambda_i) (2 + \lambda_i) + q_i^2 (2 + \lambda_i)\Big\} (2 + \lambda_i)^{-2} + d_3 =$$
$$= \sum_{i=1}^{n} \sigma^2 (1 + \lambda_i) (2 + \lambda_i)^{-1} + \sum_{i=1}^{n} q_i^2 (2 + \lambda_i)^{-1} + d_3 =$$
$$= d_1 + d_2 + d_3,$$

which is required in (42).

11^0. Let us denote

$$\mathrm{A} = \sigma^2 (I_n + P) - \nu^T \nu \tag{43}$$

and verify that

$$|\mathrm{A}|_2 \leqslant 5\, d_1 + 5\, d_2 + d_3 + 12\, d_4 + 10\, d_5 + d_6. \tag{44}$$

We have

$$\mathrm{A} = - \sum_{i=1}^{9} \mathrm{A}_i, \tag{45}$$

where (see (17) and the definition of Δ)

$$\mathrm{A}_1 = - \sigma^2 (I_n + P) + \Delta P^{-1} \Delta, \tag{46}$$

$$\mathrm{A}_2 = \Delta P^{-1} (\Pi^* + E^*)^T = \mathrm{A}_3^T, \tag{47}$$

$$A_4 = \Delta P^{-1} (\Phi - E^*)^T + (\Phi - E^*) P^{-1} \Delta, \qquad (48)$$

$$A_5 = (\Pi^* + E^*) P^{-1} (\Pi^* + E^*)^T, \qquad (49)$$

$$A_6 = (\Phi - E^*) P^{-1} (E^* + \Pi^*)^T = A_7^T, \qquad (50)$$

$$A_8 = (\Phi - E^*) P^{-1} (\Phi - E^*)^T, \qquad (51)$$

$$A_9 = \nu^T (I_m - P_m) \nu. \qquad (52)$$

Let us evaluate $|A_i|_2$, $1 \leq i \leq 9$.

Since $(I_m - P_m)$ is an orthoprojector, by (40) one has

$$|A_9|_2 \leq \mathrm{Tr}\{\nu^T (I_m - P_m) \nu\} = d_6. \qquad (53.9)$$

By the same arguments

$$|A_8|_2 \leq \mathrm{Tr}\{(\Phi - E^*) P^{-1} (\Phi - E^*)^T\} = d_5 \qquad (53.8)$$

and

$$|A_5|_2 \leq \mathrm{Tr}\{(\Pi^* + E^*) P^{-1} (\Pi^* + E^*)^T\}. \qquad (54)$$

Furthermore,

$$|A_6|_2 = |A_7|_2 = |(\Phi - E^*) P^{-1} (E^* + \Pi^*)^T|_2 =$$

$$= |(P^{-1/2} (\Phi - E^*)^T) (P^{-1/2} (E^* + \Pi^*)^T)|_2 \leq$$

$$\leq |P^{-1/2}(\Phi - E^*)^T|_2 \, |P^{-1/2} (E^* + \Pi^*)^T|_2 =$$

$$= \mathrm{Tr}^{1/2}\{(\Phi - E^*) P^{-1}(\Phi - E^*)^T\} \, \mathrm{Tr}^{1/2}\{(E^* + \Pi^*) P^{-1}(E^* + \Pi^*)^T\},$$

or

$$|A_6|_2 = |A_7|_2 \leq d_5^{1/2} \, \mathrm{Tr}^{1/2}\{(E^* + \Pi^*) P^{-1} (E^* + \Pi^*)^T\}. \qquad (55)$$

In view of (32) and the resulting equality in (24) we have

$$\mathrm{Tr}^{1/2}\{(E^* + \Pi^*) P^{-1}(E^* + \Pi^*)^T\} = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_j (\Pi^*_{ij} + E^*_{ij})^2 =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} 4 \lambda_i^2 \lambda_j (\lambda_i + \lambda_j)^{-2} (\Pi^*_{ij})^2 = 4 d_4,$$

so (54) and (55) imply

$$|A_6|_2, \ |A_7|_2 \leq 4 \, d_5^{1/2} \, d_4^{1/2}, \qquad (53.6,7)$$

$$|A_5|_2 \leq 4 \, d_4. \qquad (53.5)$$

Now let us evaluate $|A_4|_2$. We have in view of (30):

$$P^{-1}\Delta = \Delta P^{-1} = \sigma I_n + P^{-1}\Delta^*, \quad \Delta^* = \mathrm{Diag}\{(\sigma + q_i)/(2 + \lambda_i)\}; \qquad (56)$$

since $\Phi - E^*$ is skew-symmetric, we have

$$A_4 = - \Delta P^{-1} (\Phi - E^*) + (\Phi - E^*) P^{-1} \Delta =$$

$$= (\Phi - E^*) (\sigma I_n + P^{-1} \Delta^*) - (\sigma I_n + P^{-1} \Delta^*) (\Phi - E^*) =$$

$= (\Phi - E^*) P^{-1} \Delta^* - P^{-1} \Delta^* (\Phi - E^*) = ((\Phi - E^*) \ P^{-1} \Delta^*) +$
$+ ((\Phi - E^*) P^{-1} \Delta^*)^T,$
whence

$$|A_4|_2 \leq 2 \ |P^{-1} \Delta^* (\Phi - E^*)^T|_2 =$$
$$= 2 \ |(\Delta^* P^{-1/2}) \ ((\Phi - E^*) \ P^{-1/2})^T|_2 \leq$$
$$\leq 2 \ |\Delta^* P^{-1/2}| \ |(\Phi - E^*) \ P^{-1/2}|_2 =$$
$$= 2 \ |\Delta^* P^{-1/2}| \ \mathrm{Tr}^{1/2}((\Phi - E^*) \ P^{-1}(\Phi - E^*)^T) = 2 \ |\Delta^* P^{-1/2}| \ d_5^{1/2}.$$

Thus,

$$|A_4|_2 \leq 2 \ d_5^{1/2} \ |\Delta^* \ P^{-1/2}|. \tag{57}$$

Furthermore, by (56) one has

$$|\Delta^* \ P^{-1/2}| = \max_i \ (|\sigma + q_i| \ \ \lambda_i^{1/2} \ (2 + \lambda_i)^{-1}) \leq$$
$$\leq (\max_i \ (\lambda_i \ (\sigma^2 + 2 \ \sigma \ q_i + q_i^2) \ (2 + \ \lambda_i)^{-2}))^{1/2} \ \leq$$
$$\leq \left\{\max_i \ \lambda_i \ (2 + \lambda_i)^{-2}\left[[\sigma^2 \ ((1 + \lambda_i) + 1)] + q_i^2 \ [(1 + \lambda_i)^{-1} + \right.\right.$$
$$\left.\left. + 1]\right]\right\}^{1/2} = \left\{\max_i \ \left\{\lambda_i \ \sigma^2 \ (2 + \lambda_i)^{-1} + q_i^2 \ \lambda_i \ (1 + \lambda_i)^{-1} \ (2 + \right.\right.$$
$$\left.\left. + \lambda_i)^{-1}\right\}\right\}^{1/2} \leq (d_1 + d_2)^{1/2}$$

(we have taken into account (32)), which together with (57) leads to

$$|A_4|_2 \leq 2 \ d_5^{1/2} \ (d_1 + d_2)^{1/2}. \tag{53.4}$$

Now let us evaluate $|A_2|_2$. Let $Z = \Pi^* + E^*$; then in view of (30) and the resulting equality in (24) we have

$$|A_2|_2^2 = \sum_{i=1}^{n} \Pi_{ii}^2 \ \lambda_i^2 \ \sum_{j=1}^{n} Z_{ij}^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \ \left\{(2 + \lambda_i)^{-1} \ (2 \ \sigma \ (1 + \lambda_i) + \right.$$
$$\left. + \lambda_i \ q_i)\right\}^2 4 \ \lambda_j^2 \ (\lambda_i + \lambda_j)^{-2} \ (\Pi_{ij}^*)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} (2 + \lambda_i)^{-2}\left\{4 \ \sigma^2 \ (1 + \right.$$
$$\left. + \lambda_i)^2 + 4 \ \sigma \ \lambda_i \ q_i \ (1 + \lambda_i) + q_i^2 \ \lambda_i^2\right\} 4 \ \lambda_j^2 \ (\lambda_i + \lambda_j)^{-2} \ (\Pi_{ij}^*)^2 \leq$$
$$\leq \sum_{i=1}^{n}\sum_{j=1}^{n} \ (2 + \lambda_i)^{-2}\left\{ \ 4 \ \sigma^2 \ (1 + \lambda_i)^2 + \lambda_i \ (1 + \lambda_i) \ (\tfrac{1}{8} \ (4 \ \sigma)^2 + \right.$$
$$\left. + 8 \ q_i^2) + \lambda_i^2 \ q_i^2\right\} 4 \ \lambda_j^2 \ (\lambda_i + \lambda_j)^{-2} \ (\Pi_{ij}^*)^2 \leq \sum_{i=1}^{n}\sum_{j=1}^{n} \ (2 + \lambda_i)^{-2}$$

$$\Big\{4\,\sigma^2\,(1+\lambda_i)\,(1+\tfrac{9}{8}\lambda_i)+\lambda_i\,(8+9\,\lambda_i)\,q_i^2\Big\}\,4\,\lambda_j^2\,(\lambda_i+\lambda_j)^{-2}\,(\Pi^*_{ij})^2 \leq \sum_{i=1}^{n}\sum_{j=1}^{n}\Big\{(\sigma^2\,(1+\lambda_i)\,2^{-1}\,(2+\lambda_i)^{-1})\,((8+9\,\lambda_i)\,(2+\lambda_i)^{-1})+\lambda_i\,((8+9\,\lambda_i)\,(2+\lambda_i)^{-1})\,(q_i^2\,(2+\lambda_i)^{-1})\Big\}\,4\,\lambda_j^2\,(\lambda_i+\lambda_j)^{-2}\,(\Pi^*_{ij})^2 \leq$$

$$\leq \sum_{i=1}^{n}\sum_{j=1}^{n}\Big\{\tfrac{9}{2}\,(\sigma^2(1+\lambda_i)\,2^{-1}(2+\lambda_i)^{-1})+9\,\lambda_i\,(q_i^2\,(2+\lambda_i)^{-1})\Big\}\,4\,\lambda_j^2\,(\lambda_i+\lambda_j)^{-2}\,(\Pi^*_{ij})^2 \leq \sum_{i=1}^{n}\sum_{j=1}^{n}\Big\{\tfrac{9}{2}\,d_1+9\,\lambda_i\,d_2\Big\}\,4\,\lambda_j^2\,(\lambda_i+\lambda_j)^{-2}\,(\Pi^*_{ij})^2 \leq \tfrac{9}{2}\,d_1\sum_{i=1}^{n}\sum_{j=1}^{n}4\,(\Pi^*_{ij})^2+9\,d_2\sum_{i=1}^{n}\sum_{j=1}^{n}4\,\lambda_j^2\,\lambda_i(\lambda_i+\lambda_j)^{-2}\,(\Pi^*_{ij})^2 \leq 18\,d_1\,d_3+36\,d_2\,d_4$$

(we have used (32)). Thus,

$$\|A_2\|_2,\ \|A_3\|_2 \leq 3\,(2\,d_1\,d_3+4\,d_2\,d_4)^{1/2}. \qquad (53.2.3)$$

It remains to evaluate $\|A_1\|_2$. In view of (30) we have

$$\|A_1\|_2^2=\sum_{i=1}^{n}(\lambda_i\,\Pi_{ii}^2-\sigma^2\,\lambda_i^{-1}\,(1+\lambda_i))^2=\sum_{i=1}^{n}\Big\{\lambda_i\,(2+\lambda_i)^{-2}\,(4\,(1+\lambda_i^{-1})^2\,\sigma^2+4\,(1+\lambda_i^{-1})\,\sigma\,q_i+q_i^2)-\sigma^2\,(1+\lambda_i^{-1})\Big\}^2=$$

$$=\sum_{i=1}^{n}\Big\{(1+\lambda_i^{-1})\,(2+\lambda_i)^{-2}\,(4\,(1+\lambda_i)\,\sigma^2+4\,\lambda_i\,\sigma\,q_i+q_i^2\,\lambda_i\,(1+\lambda_i^{-1})^{-1}-4\,\sigma^2-4\,\sigma^2\,\lambda_i-\sigma^2\,\lambda_i^2)\Big\}^2=$$

$$=\sum_{i=1}^{n}\Big\{(1+\lambda_i^{-1})\,(2+\lambda_i)^{-2}\,(4\,\lambda_i\,\sigma\,q_i+q_i^2\,\lambda_i\,(1+\lambda_i^{-1})^{-1}-\sigma^2\,\lambda_i^2)\Big\}^2=\sum_{i=1}^{n}\Big\{(2+\lambda_i)^{-2}\,(4\,(1+\lambda_i)\,\sigma\,q_i+q_i^2\,\lambda_i+\sigma^2\,\lambda_i\,(1+\lambda_i))\Big\}^2 \leq \sum_{i=1}^{n}\Big\{(2+\lambda_i)^{-2}\,(4\,(1+\lambda_i)\,|\sigma|\,|q_i|+q_i^2\,\lambda_i+\sigma^2\,\lambda_i\,(1+\lambda_i))\Big\}^2 \leq$$

$$\leq \sum_{i=1}^{n}\Big\{(2+\lambda_i)^{-2}\,(1+\lambda_i)\,\lambda_i\,\sigma^2+2\,(2+\lambda_i)^{-2}\,(1+\lambda_i)\,\Big\{\sigma^2\,(1+\lambda_i)^{1/2}+q_i^2\,(1+\lambda_i)^{-1/2}\Big\}+\lambda_i\,q_i^2\,(2+\lambda_i)^{-2}\Big\}^2 \leq \sum_{i=1}^{n}\Big\{\Big\{((2+\lambda_i)^{-1}\,(1+\lambda_i)\,\sigma^2)\,((2+\lambda_i)^{-1}\,(\lambda_i+2\,(1+\lambda_i)^{1/2})\Big\}+\Big\{(q_i^2\,(2+\lambda_i)^{-1})\,((2+\lambda_i)^{-1}\,(\lambda_i+2\,(1+\lambda_i)^{1/2})\Big\}\Big\}^2 \leq 4\sum_{i=1}^{n}\Big\{((2+\lambda_i)^{-1}\,(1+\lambda_i)\,\sigma^2)+$$

$$+ (q_i^2 (2 + \lambda_i)^{-1}) \}^2 \leq 4 (d_1 + d_2)^2$$

(we have taken into account (32)). Thus,

$$|A_1|_2 \leq 2 (d_1 + d_2). \qquad (53.1)$$

Combining relations (45) and (53.i), $i = 1,\ldots,9$, we get (44).

12°. Combining (36), (38), (41), (42), (33), we get an inequality of the form (35), which, by the arguments from 9°, completes the proof. ■

3.9.4.6. (vi):

Let

$$H = \{z \equiv (t,X,x) \in E = \mathbb{R} \times S_n \times R^n \mid X \text{ is positive definite, } t > x^T X^{-1} x\},$$

$$\Phi(z) = - \ln \operatorname{Det} X - \ln(t - x^T X^{-1} x): H \to \mathbb{R},$$

$$h = (\tau,R,u) \in E.$$

It is clear that Φ is C^∞ - smooth on H and tends to ∞ as the argument approaches a boundary point of H.

We have

$$D\Phi(z)[h] = -\operatorname{Tr}\{X^{-1}R\} - (t - x^T X^{-1} x)^{-1} (\tau - 2 x^T X^{-1} u + x^T X^{-1} R X^{-1} x), \qquad (1)$$

$$D^2\Phi(z)[h,h] = \operatorname{Tr}\{X^{-1} R X^{-1} R\} + (t - x^T X^{-1} x)^{-1} (2 u^T X^{-1} u - 4 x^T X^{-1} R X^{-1} u + 2 x^T X^{-1} R X^{-1} R X^{-1} x) + (t - x^T X^{-1} x)^{-2} (\tau - 2 x^T X^{-1} u + x^T X^{-1} R X^{-1} x)^2 \qquad (2)$$

$$D^3\Phi(z)[h,h,h] = - 2 \operatorname{Tr}\{X^{-1} R X^{-1} R X^{-1} R\} + (t - x^T X^{-1} x)^{-1} (- 6 u^T X^{-1} R X^{-1} u + 12 x^T X^{-1} R X^{-1} R X^{-1} u - 6 x^T X^{-1} R X^{-1} R X^{-1} R X^{-1} x) - 3 (t - x^T X^{-1} x)^{-2} (\tau - 2 x^T X^{-1} u + x^T X^{-1} R X^{-1} x) (2 u^T X^{-1} u - 4 x^T X^{-1} R X^{-1} u + 2 x^T X^{-1} R X^{-1} R X^{-1} x) - 2 (t - x^T X^{-1} x)^{-3} (\tau - 2 x^T X^{-1} u + x^T X^{-1} R X^{-1} x)^3 \qquad (3)$$

Let

$$X^{-1/2} R X^{-1/2} = \Omega, \quad t - x^T X^{-1} x = \sigma > 0,$$

$X^{-1/2} x = \xi, \quad X^{-1/2} u = v,$

$\gamma = (\tau - 2 x^T X^{-1} u + x^T X^{-1} R X^{-1} x);$

then the above expressions can be rewritten as

$$D\Phi(z)[h] = - \mathrm{Tr}(\Omega) - \sigma^{-1} \gamma, \tag{4}$$

$$D^2\Phi(z)[h,h] = \mathrm{Tr}(\Omega^2) + \sigma^{-2}\gamma^2 + 2\sigma^{-1} |v - \Omega \xi|_2^2 = \rho_1^2 + \rho_2^2 + 2\rho_3^2,$$

$$\rho_1 = \mathrm{Tr}^{1/2}(\Omega^2), \quad \rho_2 = \sigma^{-1} \gamma, \quad \rho_3 = \sigma^{-1/2} |v - \Omega \xi|_2 \tag{5}$$

$$D^3\Phi(z)[h,h,h] = - 2 \mathrm{Tr}(\Omega^3) - 6 \sigma^{-2} \gamma |v - \Omega \xi|_2^2 - $$

$$- 2 \sigma^{-3} \gamma^3 - 6 \sigma^{-1} (v - \Omega \xi)^T \Omega (v - \Omega \xi). \tag{6}$$

We see that Φ is a convex function and that

$$|D\Phi(z)[h]| \leq n^{1/2} \rho_1 + \rho_2, \tag{7}$$

$$|D^3\Phi(z)[h,h,h]| \leq 2 \rho_1^3 + 6 \rho_2 \rho_3^2 + 2 \rho_2^3 + 6 \rho_3^2 |\Omega|. \tag{8}$$

Since $|\Omega| \leq (D^2\Phi(z)[h,h])^{1/2}$ (see (5)), (8) implies

$$|D^3\Phi(z)[h,h, h]| \leq O(1) (D^2\Phi(z)[h,h])^{3/2}$$

with an absolute constant $O(1)$.

The above remarks immediately lead to (vi). ■

3.9.6. Theorem 3.3.

(i) is obvious: one can set $\Gamma_1 = (E',G', \sigma\circ\pi,F)$.

Let us prove (ii). Without loss from generality we can assume that $E' = E \times R^l$, π is a projector from E' onto E with $\mathrm{Ker}\, \pi = R^l$, $E_1 = H \times R^k$ and that σ is a projector from E_1 onto H with $\mathrm{Ker}\, \sigma = R^k$. Let $G_1' = G \cap H$, so $G_1 \in C(H)$.

It is clear that

$$\Gamma' = (H \times R^l,\ G'' = G'\cap(H \times R^l),\ \pi|_{H\times R^l},\ F' = F|_{\mathrm{int}G''})$$

is a (ϑ,l) - covering for G_1' (see P.3.2.(i)). Furthermore, we have $G_1 = G_1' \times R^k$. Let $G^+ = G'' \times R^k \in C(H \times R^l \times R^k)$, let π^+ be the projector from $H \times R^l \times R^k$ onto $H \times R^k$ with $\mathrm{Ker}\, \pi^+ = R^l$, and let $F^+(u) = F'(\pi^+(u))$: $\mathrm{int}\, G^+ \to \mathbb{R}$. It is clear that

$$\Gamma^+ = (H \times R^l \times R^k,\ G^+,\ \pi^+,\ F^+)$$

is the desired (ϑ,l) - covering for G_1.

Let us prove (iii). Without loss of generality we can assume that $E'_i = E \times H_i$, $\dim H_i = l_i$, and that π_i are projectors from E'_i onto E with $\mathrm{Ker}\,\pi_i = H_i$. Let

$$E^+ = E \times H, \quad H = H_1 \times \ldots \times H_k,$$

and let π be the natural projector from E^+ onto E and

$$G_i^+ = \{(x,u_1,\ldots,u_k) \in E' \mid (x,u_i) \in G'_i\}$$

(herein $x \in E$, $u_j \in H_j$). Let

$$F'_i(x,u_1,\ldots,u_k) = F_i(x,u_i) : \mathrm{int}\, G_i \to \mathbb{R}.$$

Then F'_i are ϑ_i - self-concordant barriers for G_i^+ (P.3.2.(i)). If $x_0 \in \bigcap_{i=1}^{k} \mathrm{int}\, G_i$ and $(x_0,u_i^0) \in G'_i$ (such u_i^0 obviously exist), then $(x_0,u_1^0,\ldots,u_k^0) \in \bigcap_{i=1}^{k} \mathrm{int}\, G_i^+$, thus the above intersection is nonempty; by P.3.2.(iii) the function

$$F^+(x,u_1,\ldots,u_k) = \sum_{i=1}^{k} F'_i(x,u_1,\ldots,u_k) : \mathrm{int}\, G^+ \to \mathbb{R},$$

$$G^+ = \bigcap_{i=1}^{k} G_i^+,$$

is a $\sum_{i=1}^{k} \vartheta_i$ - self-concordant barrier for G^+. Let us verify that $\Gamma^+ = (E^+, G^+, \pi, F^+)$ is a $(\sum_{i=1}^{k} \vartheta_i, \sum_{i=1}^{k} l_i)$ - covering for G.

It is necessary to prove that $\pi(G^+) = G$ and that each bounded subset of G is a π-image of some bounded subset of G^+. If $(x,u_1,\ldots u_k) \in G^+$, then $(x,u_i) \in G'_i$ for each i, hence $x \in G_i$ for each i, i.e. $x \in G$. At the same time if $x \in G$, then for each i there exists u_i such that $(x,u_i) \in G'_i$, hence $(x,u_1,\ldots,u_k) \in G^+$. The above arguments mean that $\pi(G^+) = G$. If $Q \subset G$ is bounded, then there exists $R < \infty$ such that

$$Q \subset \pi_i(\{(x,u_i) \in G'_i \mid \|(x,u_i)\|_2 \leq R\})$$

for each i; it is clear that $Q \subset \pi(Q^+)$, where

$$Q^+ = \{(x,u_1,\ldots,u_k) \in G^+ \mid \|(x,u_i)\| \leq R,\ 1 \leq i \leq k\}$$

is a bounded subset of G^+. ■

3.9.7. Theorem 3.4.

Under the assumptions of the theorem g obviously is a finite convex lower semicontinuous function on H, thus (H,g) is a functional element.

Let $s = \dim E$; consider $(s + k + 1)$ - dimensional space E' with the coordinates $(t,x_1,\dots,x_s,\tau_1,\dots,\tau_k)$, where $x_1,\dots,x_s$ are coordinates in E. Let for $1 \leqslant i \leqslant k$

$$Q_i = \Big\{(t,x_1,\dots,x_s,\tau_1,\dots,\tau_k) \in E' \mid (x_1,\dots,x_s) \in G_i, \quad \tau_i \geqslant \phi_i(x_1,\dots,x_s)\Big\},$$

$$Q = \Big\{(t,x_1,\dots,x_s,\tau_1,\dots,\tau_k) \in E' \mid (\tau_1,\dots,\tau_k) \in G, \; t \geqslant \phi(\tau_1,\dots,\tau_k)\Big\}.$$

It is clear that the sets Q_i, $1 \leqslant i \leqslant k$, and Q admit (ϑ_i,l_i) - coverings Γ_i and a (ϑ,l) - covering Γ, respectively, and that these coverings are induced in a straightforward manner by the given coverings for the initial functional elements (indeed, each of the above sets is of the form $\pi^{-1}(\mathfrak{G})$, where $\mathfrak{G}$ is the epigraph of the corresponding initial functional element and π is a projector from E' onto an appropriate coordinate subspace of E').

Let $x \in \operatorname{int} H$ and $\tau = (\tau_1,\dots,\tau_k) > f(x)$; then, in view of the conditions of our theorem, $\tau \in \operatorname{int} G$. Let $t > \phi(\tau)$. It is clear that the point $w = (t,x_1,\dots,x_s,\tau_1,\dots,\tau_k)$ belongs to $\operatorname{int} Q \cap \operatorname{int} Q_1 \cap \dots \cap \operatorname{int} Q_k$, hence the above intersection is nonempty. By P.3.2.(iii) the coverings Γ_i, Γ induces in a straightforward manner a $(\vartheta^* \equiv \sum_{i=1}^{k} \vartheta_i + \vartheta,\; l^* \equiv \sum_{i=1}^{k} l_i + l)$ - covering Γ^* for the set $Q^* \equiv Q \cap (\bigcap_{i=1}^{k} Q_i)$.

Now let us verify that if $\pi: E' \to \mathbb{R} \times E$ is defined as $\pi(t,x_1,\dots,x_s,\tau_1,\dots,\tau_k) = (t,x_1,\dots,x_s)$, then π maps Q^* onto $\mathfrak{G}(H,g)$, and each compact in $\mathfrak{G}(H,g)$ is an image of a compact belonging to Q^* (this together with T.3.3.(i) will complete the proof).

First of all, let us establish the equality $\pi(Q^*) = \mathfrak{S}(H,g)$. For $v = (t,x_1,\ldots,x_s) \in \mathfrak{S}(H,g)$ let $\tau_i = \phi_i(x)$; then $w = (t,x_1,\ldots,x_s,\tau_1,\ldots,\tau_k) \in Q^*$ (obviously) and $\pi(w) = v$, thus $\mathfrak{S}(H,g) \subset \pi(Q^*)$. To prove the inverse inclusion it suffices to verify that if $w = (t,x_1,\ldots,x_s,\tau_1,\ldots,\tau_k) \in Q^*$, then $\pi(w) \in \mathfrak{S}(H,g)$. Indeed, let $x = (x_1,\ldots,x_s)$; then $x \in G_i$, $\phi_i(x) \leqslant \tau_i$, $1 \leqslant i \leqslant k$, $\tau = (\tau_1,\ldots,\tau_k) \in G$ and $t \geqslant \phi(\tau)$ (by definition of Q^*), hence $x \in H$ and $\tau \in f(x) + (R^k)_+$. By the conditions of the theorem this implies $\phi(\tau) \geqslant \phi(f(x))$, so $t \geqslant \phi(\tau) \geqslant g(x)$, or $\pi(w) = (t,x) \in \mathfrak{S}(H,g)$; the inverse inclusion is proved.

It remains to verify that if X is a compact belonging to $\mathfrak{S}(H,g)$, then $X \subset \pi(Y)$ for some compact Y belonging to Q^*. Let $Z = f(X)$; then Z is a bounded subset of G and $\phi(\tau)$ is bounded on Z (by the conditions of the theorem); hence there exists a bounded subset $Y' \subset E'$ containing all the points w of the form $(g(x),x,f(x))$, $x \in X$; it is clear that $X \subset \pi(Y' \cap Q^*)$. Hence for $Y = \mathrm{Cl}\,(Y' \cap Q^*)$ we have $X \subset \pi(Y)$ while Y obviously is a compact ■

3.9.8. Proposition 3.7.

After an appropriate choice of the coordinates we can assume that the point involved into the statement is O, while in the neighbourhood of this point G is described by the inequalities $x_i \leqslant 0$, $1 \leqslant i \leqslant k$. Let

$$x(t) = -t\sum_{j=1}^{k} e_j\,,\quad x^{(i)}(t) = -t\sum_{\substack{1\leqslant j\leqslant k\\ j\neq i}} e_j,\quad 1 \leqslant i \leqslant k\,.$$

(e_j are the orts of our coordinate axes); then for all small enough $t > 0$ the points $x(t)$ belong to int G, the points $x^{(i)}(t)$ belong to ∂G and

$$\pi_{x^{(i)}(t)}(x(t)) \to 0,\quad t \to 0.$$

The latter fact, by virtue of (3.11), implies

$$\varliminf_{t\to 0} DF(x(t))[x^{(i)}(t) - x(t)] \geqslant 1,$$

hence

$$k \leq \lim_{t\to 0} \sum_{i=1}^{k} DF(x(t))[x^{(i)}(t) - \dot{x}(t)] =$$

$$= \lim_{t\to 0} DF(x(t))[0 - x(t)] \leq \vartheta$$

(the latter inequality by (3.6)), Q.E.D. ■

3.9.9. Theorem 3.5.

(3.53) is a simple corollary of (3.52), so we must prove the first statement of the theorem only. Without loss of generality we can restrict ourselves to the case when G does not contain any straight line.

1^0. For $x \in \text{int}\, G$ one obviously has $G^*(x) \in C_B(R^n)$ (int $G^*(x) \neq \emptyset$ since G does not contain any straight line, and the boundness of $G^*(x)$ follows from $x \in \text{int}\, G$). Hence the function

$$f(x) = |G^*(x)|$$

is well defined and positive on int G. If $x_i \in \text{int}\, G$ and $x_i \to x \in \partial G$, then all of the sets $G^*(x_i)$ contain certain fixed open nonempty set (since the sequence $\{x_i\}$ is bounded) and at the same time are not uniformly bounded (since $\lim x_i \in \partial G$). Since $G^*(\ )$ is a convex set, we have $f(x_i) \to \infty$. Thus, the function

$$\Phi(x) = \ln f(x)$$

is well defined on int G and tends to ∞ as the argument belonging to int G approaches a point from ∂G.

2^0. Let

$$p(\phi) = \sup\{ \phi^T y \mid y \in G\}: R^n \to \mathbb{R} \cup \{+\infty\}$$

be the support function for G. For $x \in \text{int}\, G$ we have

$$G^*(x) = \{\tau \phi \mid \phi \in S,\ 0 \leq \tau \leq r_x(\phi) \equiv (p(\phi) - \phi^T x)^{-1}\},$$

whence

$$f(x) = n^{-1} \int (p(\phi) - \phi^T x)^{-n}\, dS(\phi)$$

(the integral is taken with respect to the Lebesque area of the unit sphere in R^n; of course, $(+\infty)^{-n} = 0$). It is clear that f (and hence $-\Phi$) is C^∞ - smooth on int G; moreover,

$$D^l f(x)[h,\ldots,h] = (-1)^l\ (n+l-1)!\ (n!)^{-1} \int\ (\phi^T h)^l$$

$$(p(\phi) - \phi^T x)^{-l-n}\, dS(\phi) = (-1)^l\ (n+l)!\ (n!)^{-1} \int_{G(x^*)} (y^T h)^l\ dy \equiv$$

$$\equiv (-1)^l\ (n+l)!\ (n!)^{-1}\ I_l(h)$$

(we have used the description of $G^*(x)$ by means of $r_x(\Phi)$). A straightforward computation gives us ($x \in \operatorname{int} G$; I_0 does not depend on h):

$$D\Phi(x)[h] = -(n+1)\, I_1(h) I_0^{-1};$$

$$D^2\Phi(x)[h,h] = (n+1)(n+2)\, I_2(h)\, I_0^{-1} - (n+1)^2\, (I_1(h) I_0^{-1})^2,$$

$$D^3\Phi(x)[h,h,h] = -(n+3)(n+2)(n+1)\, I_3(h)\, I_0^{-1} + $$
$$+ 3(n+2)(n+1)^2\, I_2(h)\, I_1(h)\, I_0^{-2} - 2(n+1)^3\, I_1^3(h)\, I_0^{-3}.$$

Let us fix $x \in \operatorname{int} G$ and $h \in R^n$ such that $\|h\|_2 = 1$, and let

$$\Delta = \{t \in R \mid \exists\, y \in G^*(x):\ y^T h = t\},$$

$$\psi(t) = (\operatorname{mes}_{n-1} \{y \in G^*(x) \mid y^T h = t\})^{1/(n-1)},$$

mes_{n-1} is the $(n-1)$ - dimensional Lebesque measure. Then, obviously,

$$I_l(h)\, I_0^{-1} = \int_\Delta t^l\, \eta^{n-1}(t)\, dt,$$

$$\eta(t) = \psi(t)\, (\int_\Delta \psi^{n-1}(\tau)\, d\tau)^{-1/(n-1)},$$

hence $\eta(t) \geq 0$, $t \in \Delta$, and $\int_\Delta \eta^{n-1}(t)\, dt = 1$. Notice that *the function $\eta(t)$ is concave on the segment* Δ (the latter is the Brunn-Minkowsky's theorem).

We see that the quantities $I_l(h)\, I_0^{-1}$, by means of which the differentials of Φ are expressed, can be interpreted as the moments of some random variable ξ (which takes its values in Δ with the density of the probability distribution of the form $\eta^{n-1}(t)$). Let us express the initial moments by means of the central ones, i.e. let us denote ($\mathbf{E}$ means the averaging operator)

$$\mu = I_1(h)\, I_0^{-1} \equiv \mathbf{E}\,\xi;$$

$$\sigma^2 = I_2(h)\, I_0^{-1} - \mu^2 \equiv \mathbf{E}\,(\xi - \mathbf{E}\xi)^2;$$

$$\theta = \mathbf{E}\,(\xi - \mathbf{E}\xi)^3.$$

A straightforward computation leads to

$$D\Phi(x)[h] = -(n+1)\,\mu;$$

$$D^2\Phi(x)[h,h] = (n+2)(n+1)\,\sigma^2 + (n+1)\,\mu^2;$$

$$D^3\Phi(x)[h,h,h] = -(n+3)(n+2)(n+1)\,\theta - 6(n+2)(n+1)\,\sigma^2\,\mu - 2(n+1)\,\mu^3.$$

Thus Φ is convex and

$$|D\Phi(x)[h]| \leq (n+1)^{1/2}\,(D^2\Phi(x)[h,h])^{1/2}.$$

Taking into account the results of 1^0 we see that to prove the theorem it suffices to verify that Φ is self-concordant with an appropriate absolute constant being chosen as the parameter value, i.e. it suffices to establish the inequality

$$|(n+3)(n+2)(n+1)\,\theta + 6(n+2)(n+1)\,\sigma^2\,\mu + 2(n+1)\,\mu^3| \leq O(1)\,((n+2)(n+1)\,\sigma^2 + (n+1)\,\mu^2)^{3/2}.$$

The latter inequality, in turn, follows from the inequality

$$|\theta| \leq O(1)\,\sigma^3.$$

3^0. Thus, we have reduced our problem to that one as follows. We are given a segment $\delta = [-a,b] \subset \mathbb{R}$, $a,\ b > 0$, and a continuous nonnegative function $\psi(t)$ on δ, such that

$$\int_{a}^{b} t\,\psi^{n-1}(t)\,dt = 0, \tag{1}$$

$$\int_{-a}^{b} \psi^{n-1}(t)\,dt = 1. \tag{2}$$

Let

$$\theta = \int_{-a}^{b} t^3\,\psi^{n-1}(t)\,dt \quad \text{and} \quad \sigma = \left(\int_{-a}^{b} t^2\,\psi^{n-1}(t)\,dt\right)^{1/2};$$

it is necessary to prove that under an appropriate choice of an absolute constant $O(1)$ we have $\theta \leq O(1)\,\sigma^3$.

First of all, let $\lambda = 1/\sigma$ and $a = \lambda\,\tilde{a}$, $b = \lambda\,\tilde{b}$,

$$\psi(t) = \lambda^{1/(n-1)}\,\tilde{\psi}(\lambda^{-1}t),$$

$$\tilde{\theta} = \int_{-\tilde{a}}^{\tilde{b}} t^3\,\tilde{\psi}^{n-1}(t)\,dt = \lambda^3\,\theta = \theta/\sigma^3,$$

$$\tilde{\sigma} = \left(\int_{-\tilde{a}}^{\tilde{b}} t^2\,\tilde{\psi}^{n-1}(t)\,dt\right)^{1/2} = \lambda\,\sigma = 1,$$

and also, as above, $\int_{-\tilde{a}}^{\tilde{b}} \tilde{\psi}^{n-1}(t)\,dt = 1$ and $\int_{-\tilde{a}}^{\tilde{b}} t\,\tilde{\psi}^{n-1}(t)\,dt = 0$.

So the situation can be reduced to the case when the objects a, b, ψ, satisfy (1),(2) and the condition

$$\int_{-a}^{b} t^2\,\psi^{n-1}(t)\,dt = 1; \tag{3}$$

under these assumptions we desire to prove that

$$|\theta| \leqslant O(1). \tag{4}$$

It is convenient to introduce the body

$$G^* = \{(t,u) \in \mathbb{R} \times R^n \mid t \in \delta,\ \|u\|_2 \leqslant \psi(t)\}.$$

Under appropriate choice of the volume unit we have:

G^* is a convex compact body of unit volume (the latter - by (2));

the center of gravity of this body is at the origin (1).

Without loss of generality we can assume that

$$|\theta| \leqslant \int_0^b t^3\,\psi^{n-1}(t)\,dt \equiv \theta^*,$$

so it suffices to evaluate from above the quantity θ^*.

Each hyperplane passing through the gravity center of a convex compact body of unit volume divides this body into parts with the volume of each part being $\geqslant \exp(-1)$ [Gr. 1960]. In particular,

$$\int_0^b \psi^{n-1}(t)\,dt \equiv V \geqslant \exp(-1). \tag{5}$$

Let τ be such that

$$V' \equiv \int_\tau^b \psi^{n-1}(t)\,dt = (n/(n+1))^n\,V; \tag{6}$$

in view of (3) and (5) - (6) we have ("the Chebyshev's inequality")

$$\tau \leqslant O(1) \tag{7}$$

(from now on all the constant factors in $O(\ )$ are absolute constants). Therefore (3) implies

$$\int_0^\tau t^3\,\psi^{n-1}(t)\,dt \leqslant O(1). \tag{8}$$

Now let us introduce a *linear* function $\phi(t)$ and positive h satisfying the relations

$$\phi(h) = 0; \ \phi(\tau) = \psi(\tau); \ \int_\tau^h \phi^{n-1}(t)\,dt = V', \tag{9}$$

i.e. let us replace the part of G^* situated to the right of the hyperplane $t = \tau$ by the cone of the same volume and the same intersection with this hyperplane. It is clear that the graph of ϕ is a secant for the graph of ψ, and the t-coordinates of the intersection points of these graphs are τ and $\tau' > \tau$. Besides this, $h \geqslant b$. Notice that

$$\int_\tau^b t^3 \psi^{n-1}(t)\,dt - \int_\tau^h t^3 \phi^{n-1}(t)\,dt = \int_\tau^h t^3 \gamma(t)\,dt,$$

where γ is a function with the zero value of the integral over the segment $[\tau, h]$, such that γ is nonnegative on $[\tau, s]$ and nonpositive on $[s, h]$ for an appropriate s (we have taken into account the convexity of ψ and the linearity of ϕ). In view of these properties we have

$$\int_\tau^h t^3 \gamma(t)\,dt \leqslant 0,$$

thus

$$\theta^* \leqslant \int_\tau^h t^3 \phi^{n-1}(t)\,dt + \int_0^\tau t^3 \psi^{n-1}(t)\,dt = \theta^{**} + O(1),$$

$$\theta^{**} = \int_\tau^h t^3 \phi^{n-1}(t)\,dt$$

(see (8)). Thus, it suffices to prove that

$$\theta^{**} \leqslant O(1).$$

Let us verify that $h \leqslant n\tau$. Indeed, consider the cone

$$K = \{(t,u) \mid 0 \leqslant t \leqslant h, \ |u|_2 \leqslant \phi(t)\}.$$

The part of this cone situated between the hyperplanes $t = 0$ and $t = \tau$ contains similarly defined part of G^*, and the part, K', of the cone K, which is situated to the right of hyperplane $t = \tau$, has the same volume, V', as the corresponding part of G^*. Therefore

$$V \leqslant |K| = ((h+\tau)/h)^n \ |K'| = ((h+\tau)/h)^n \ V' =$$
$$= ((h+\tau)/h)^n \ (n/(n+1))^n \ V$$

(the latter - by the definition of V'), which implies $h \leqslant n\tau$.

Thus, we have

$$h = \eta\, n\, \tau\,, \quad \eta \leq 1, \quad \tau \leq O(1) \tag{10}$$

and

$$\theta^{**} = S \int_{\tau}^{h} t^3 \,((h - t)/(h - \tau))^{n-1}\, dt, \quad V' = n^{-1} S\,(h - \tau), \tag{11}$$

wherein $S = \psi^{n-1}(\tau)$. We have

$$\theta^{**} = S\,\tau^4 \int_{1}^{n\eta} l^3 \,((\eta n - l)/(\eta n - 1))^{n-1}\, dl =$$

$$= S\,\tau^4\,(n\eta - 1)^{-n+1} \int_{0}^{n\eta - 1} (n^3\eta^3 - 3\,n^2\eta^2\, s + 3\,n\eta s^2 - s^3)\, s^{n-1}\, ds =$$

$$= S\,\tau^4\,(n\,\eta - 1)\,\Big\{n^2\,\eta^3 - 3\,n^2\,\eta^2\,(n+1)^{-1}\,(n\,\eta - 1) +$$

$$+ 3\,n\,\eta\,(n+2)^{-1}\,(n\,\eta - 1)^2 - (n+3)^{-1}\,(n\,\eta - 1)^3\Big\} =$$

$$= V'\,\tau^3\,\Big\{n\,\Big\{(n^2\,\eta^3 - (n+3)^{-1}\,(n\,\eta - 1)^3)_1 -$$

$$- (3\,n^2\,\eta^2\,(n+1)^{-1}\,(n\,\eta - 1) - 3\,n\,\eta\,(n+2)^{-1}\,(n\,\eta - 1)^2)_2\Big\}\Big\}.$$

Since $V' \leq 1$ and $\tau \leq O(1)$, it suffices to verify that the expression denoted by $(\ )_1 - (\ )_2$ does not exceed $O(n^{-1})$ (this will lead to the desired estimate $\theta^{**} \leq O(1)$). A straightforward computation, using the relation $0 \leq \eta \leq 1$, leads to

$$(\ )_1 = (n^2 + 3\,n)^{-1}\,(3\,n^3\,\eta^2\,(\eta+1) - 3\,\eta\,n^2) + O(n^{-1}),$$

$$(\ )_2 = (n^2 + 3\,n + 2)^{-1}\,(3\,n^3\,\eta^2\,(\eta+1) - 3\,n^2\,\eta) + O(n^{-1});$$

hence $(\ )_1 - (\ )_2$ is of the desired order. ■

Section 4. Another self-concordant families and polynomial- time methods

We now give two more examples of self-concordant families and three more examples of polynomial-time barrier-generated methods.

4.1. Method of centers and Renegar's type family.

Let F be a ϑ-self-concordant barrier for $G \in C_B(E)$ and let $f(x)$ be a convex quadratic form on E. Let us fix a constant $\zeta \geq 1$.

Let

$$t^* = \min \{f(x) \mid x \in G\}, \quad \Delta = (t^*, +\infty),$$

and let

$$Q_t = \{x \in \text{int } G \mid f(x) < t\},$$
$$F_t(x) = \zeta \ln(1/(t - f(x))) + F(x) : Q_t \to \mathbb{R}$$

for $t \in \Delta$. Thus, a family

$$\mathcal{F}^* = \mathcal{F}^*(F,f) = \{Q_t, F_t, E\}_{t \in \Delta}$$

is defined.

Theorem 4.1. For each $\lambda \in (0,\lambda_*)$ and $\text{æ}' \in (\lambda,\lambda_*)$, under the parameters α, γ, μ, ξ, η, æ choice in accordance with the relations

$$\alpha(t) \equiv 1; \ \gamma(t) \equiv 1; \ \mu(t) \equiv 1;$$
$$\xi(t) = \zeta^{1/2}\, \Omega/(t - t^*);$$
$$\eta(t) = \Omega/(t - t^*),$$
$$\text{æ} = \lambda, \tag{4.1}$$

where

$$\Omega = 1 + (\delta + \vartheta)/\zeta,$$
$$\delta \equiv (\text{æ}' \ (1 - \beta)^{-1}) \ (\beta \ (1 - \beta)^{-1} + 1 + 3 \vartheta^*),$$
$$\vartheta^* = \vartheta + \zeta,$$
$$\beta = \omega(\text{æ}'), \tag{4.2}$$

the family $\mathcal{F}^*(F,f)$ is self-concordant with these parameters.

In particular, for this family we have

$$\psi(t) \equiv 1, \ t \in \Delta, \tag{4.3}$$

and

$$\rho_\lambda(\mathcal{F}^*;t,t') = \lambda^{-1}\, \Omega \ (\zeta^{1/2} + \lambda) \ \left|\ln\left(\frac{t - t^*}{t' - t^*}\right)\right|. \tag{4.4}$$

Moreover, the following implication holds:

$$t \in \Delta, \ x \in Q_t, \ \lambda(F_t,x) \leq \lambda \ \Rightarrow$$
$$\Rightarrow (t - f(x))^{-1} \leq \Omega \ (t - t^*)^{-1}. \tag{4.5}$$

Now we can describe the *method of centers* for the solution of the problem

$$f(x) \to \min \mid x \in G. \tag{4.6}$$

Let us fix constants λ, λ' such that

$$\lambda^+ \leq \lambda' < \lambda < \lambda_*, \tag{4.7}$$

and assume that we are given $t_0 \in \Delta$ and $x_{-1} \in Q_{t_0}$ satisfying the relation

$$\lambda(F_{t_0}, x_{-1}) \leq \lambda. \tag{4.8}$$

We produce a sequence of points x_i and numbers $t_i \in \Delta$ as follows:

being given t_i and x_{i-1}, such that

$$t_i \in \Delta;\ x_{i-1} \in Q_{t_i};\ \lambda(F_{t_i}, x_{i-1}) \leq \lambda, \tag{4.9$_i$}$$

we find a point x_i satisfying the relations

$$x_i \in Q_{t_i}, \quad \lambda(F_{t_i}, x_i) \leq \lambda'. \tag{4.10$_i$}$$

Notice that under condition (4.9$_i$) the point $x_i = x^*(F_{t_i}, x_{i-1})$ satisfy (4.10$_i$) (T.1.3.(ii)).

after x_i is produced, we define t_{i+1} in accordance with

$$\ln\left[\frac{t_{i+1} - f(x_i)}{t_i - f(x_i)}\right] = -\frac{\lambda - \lambda'}{\Omega\ (\zeta^{1/2} + \lambda)} \equiv -\chi. \tag{4.11}$$

(since $x_i \in Q_{t_i}$, it is clear that $t_i > f(x_i)$; thus t_{i+1} is well-defined).

Let us verify, that (4.10$_i$) implies the inclusion $t_{i+1} \in \Delta$ and relations (4.9$_{i+1}$), as well as the inequality

$$f(x_i) - t^* < t_{i+1} - t^* \leq \exp(-\nu)\ (t_i - t^*), \tag{4.12}$$

$$\nu = \ln(\Omega/(\Omega - 1 + \exp(-\chi))).$$

We have $t_i > f(x_i) \geq t^*$, so $t_{i+1} \in \Delta$. We also have

$$1 > (t_{i+1} - t^*)/(t_i - t^*) > (t_{i+1} - f(x_i))/(t_i - f(x_i)),$$

which, by (4.11) and T.4.1, leads to

$$\rho_\lambda(\mathcal{F}^*; t_i, t_{i+1}) \leq (\lambda - \lambda')/\lambda.$$

The latter relation, by T.2.1, implies (4.9$_{i+1}$). To verify (4.12), notice, that (4.11) and (4.5) imply

$$1 - \exp(-\chi) = (t_i - t_{i+1})/(t_i - f(x_i)) \leq$$
$$\leq \Omega\,(t_i - t_{i+1})/(t_i - t^*),$$

so

$$(t_i - t_{i+1})/(t_i - t^*) \geq \Omega^{-1}\,(1 - \exp(-\chi)),$$

which leads to the second inequality in (4.12); the first one follows from the inclusion in (4.9_{i+1}).

(4.12) leads to the estimate

$$f(x_i) - \min_G f < \exp(-(i+1)\,\nu)\,(t_0 - t^*). \qquad (4.13)$$

The value of ν depends on λ, λ' and ζ only. Assume that λ and λ' are absolute constants satisfying (4.7). Then, maximizing ν over ζ, we obtain

$$\zeta = O(\vartheta) \quad \text{and} \quad \nu = O(\vartheta^{-1/2}).$$

Thus, the rate of convergence of the method under consideration is the same as of the F-generated barrier method.

The rational choice of the parameters for the method is

$$\lambda = 0.136, \quad \lambda' = \lambda^+ \cong 0.025, \quad \zeta = 3\,\vartheta;$$

for large ϑ this choice results in

$$\nu \cong 0.011\,\vartheta^{-1/2}.$$

To initialize the method of centers, one needs a pair (t_0, x_{-1}) such that $x_{-1} \in Q_{t_0}$ and $\lambda(F_{t_0}, x_{-1}) \leq \lambda$. To produce such a pair, we can first approximate the F-center of G using, for example, the preliminary stage of the barrier method. The stage is terminated when a point x, $\lambda(F,x) \leq \lambda/2$, is produced. This point can be taken as x_{-1}. Then, obviously, $\lambda(F_t, x_{-1}) \leq \lambda$ for all sufficiently large t, which allows us to choose an appropriate t_0.

Notice that, being considered as geometrical objects, the minimizers trajectories for the barrier and the centers methods corresponding to (f,F) coincide (although the parametrization of the curve depends on the method). The approximations to this curve, generating by the methods, of course differ from each other.

The polynomial-time method of centers originates from [Re.1988] (where LP problems are considered and $G = \{x \in R^n \mid$

$a_i^T x > b_i,\ 1 \leq i \leq m\}$, $F(x) = -\sum_{i=1}^{m} \ln(a_i^T x - b_i)$, f is linear). This method is used in most of the papers mentioned in Sect. 0 (excluding [Go. 1987] and [Ne. 1988 1,2,3,4]).

4.2. Dual parallel trajectories method and homogeneous self-concordant families.

The next self-concordant family (we call it *homogeneous*) is defined as follows.

Let E^* be the space conjugate to E. Let for $\vartheta \geq 1$ $\mathcal{F}(E^*,\vartheta)$ be the set of all functions $F^* \in S_1(E^*,E^*)$ satisfying the relation

$$\vartheta^*(F^*) = \sup\{D^2F^*(\phi)[\phi,\phi] \mid \phi \in E^*\} \leq \vartheta. \tag{4.14}$$

Let E_0^* be a hyperplane in E^*, kodim $E^* = 1$, and let $b \in E^* \setminus E_0^*$, $\Delta = (0,\infty)$ and $F^* \in \mathcal{F}(E^*,\vartheta)$. These objects generate a family of functions defined on E_0^*:

$$\mathcal{F}^{**} = \mathcal{F}^{**}(F^*,E_0^*,b) = \{Q_t = E_0^*,\ F_t(\phi) = F^*(t\,\phi + t\,b),\ E_0^*\}_{t\in\Delta}. \tag{4.15}$$

Proposition 4.1. For each $\vartheta \geq 1$, $F^* \in \mathcal{F}(E^*,\vartheta)$, E_0^*, b, and for each $\varkappa \in (0,\lambda_*)$, the family $\mathcal{F}^{**}(F^*,E_0^*,b)$ is self-concordant with the parameters

$$\alpha(t) = 1;\ \mu(t) = t;\ \gamma(t) = t^2;\ \xi(t) = \eta(t) = \vartheta^{1/2}/t;\ \varkappa. \tag{4.16}$$

In particular, $\psi(\mathcal{F}^{**},t) = 1$ and

$$\rho_\nu(\mathcal{F}^{**};t,\tau) = (\nu^{-1} + 1)\,(\vartheta)^{1/2}\,|\ln(t/\tau)|. \tag{4.17}$$

The functions F^* of the kind mentioned above arise as the Legendre transformations of self-concordant barriers. More precisely, let $G \in C_B(E)$ and let

$\mathfrak{R}(G;\phi) = \max\{ \langle\phi,x\rangle \mid x \in G \}$

be the support function for G (from now on $\langle\phi,x\rangle$ means the value of a functional $\phi \in E^*$ at $x \in E$). Then the following proposition holds:

Proposition 4.2. A. Let $G \in C_B(E)$, $F \in \mathcal{B}(G,\vartheta)$, and let

$F^*(\phi) = \max(\langle\phi, x\rangle - F(x) \mid x \in \text{int } G) : E^* \to \mathbb{R}$

be the Legendre transformation of F. Then

(i) $F^* \in \mathcal{F}(E^*, \vartheta)$ and D^2F^* is non-degenerate on E^*;

(ii) $\mathfrak{K}^*(F^*;\phi) = \lim_{t\to\infty} DF^*(t\ \phi)[\phi] = \mathfrak{K}(G;\phi)$.

B. If $F^*: E^* \to \mathbb{R}$ satisfies (i) and $\mathfrak{K}^*(F^*;\ )$ is finite, then the domain of the Legendre transformation, F, of the function F^* is the interior of a set $G \in C_B(E)$, and $F \in \mathfrak{s}(G,\vartheta)$.

The following fact is obvious:

Remark 4.1. Let $F_i^* \in \mathcal{F}(E^*,\vartheta_i)$, $p_i \geq 1$, $i = 1,2$, $x(\phi)$ be an affine form on E^* and let $\phi = A\ \psi$ be a *homogeneous* affine transformation from H^* into E^*. Then

(i) $p_1\ F_1^*(\phi) + p_2\ F_2^*(\phi) \in \mathcal{F}(E^*,\ p_1\vartheta_1 + p_2\vartheta_2)$;

(ii) $F_1^*(\phi) + x(\phi) \in \mathcal{F}(E^*,\vartheta_1)$;

(iii) $F_1^*(A\ \psi) \in \mathcal{F}(H^*,\vartheta_1)$.

Corollary 4.1. If $G_i \in C_B(E)$ and $F_i \in \mathfrak{s}(G_i,\vartheta_i)$, $1 \leq i \leq k$, then the arithmetic sum $G = G_1 + \ldots + G_m$ of the sets G_i admits a $(\sum_{i=1}^{k} \vartheta_i)$-self-concordant barrier.

Indeed, the desired barrier for G is the Legendre transformation of the sum of F_i^*, the Legendre transformations of the given barriers.

Now we describe *the dual parallel trajectories method* for LP problems (the method is close to that one described in [Ne. 1988 1,4]).

Let $G \in C_B(R^m)$, F be a ϑ-self-concordant barrier for G; assume that we know the F-center of G (to simplify the description, let the center be O). Let A, rank $A = n$, be a $n \times m$ - matrix and $b \in R^n$. The dual parallel trajectories method solves problems of the form

$$\tau \to \max \mid A\,x = \tau\, b,\ x \in G. \tag{4.18}$$

If G is a polytope, then (4.18) is a LP problem. Notice that the assumption $F'(0) = 0$ is not a severe restriction, which is demonstrated by the following example:

$$\tau \to \max \mid x \in R^m,\ A x = \tau b,\ |x|_\infty \leq 1, \tag{4.19}$$

where $G = \{x \in R^m \mid |x|_\infty \leq 1\}$, $\vartheta = m$ and $F(x) = \sum_{i=1}^{m} \ln(1/(1 - x_i^2)) + 2\ln(m)$ (it is easy to show that the parameter value for the barrier F equals m). Notice that (4.19) is an "universal", in some natural sense, format for LP problems.

Without loss of generality, we can assume that

$$A\,(F''(0))^{-1}A^T = I_n, \tag{4.20}$$

(because the system $A x = \tau b$ can be replaced by an equivalent system such that the rows of its matrix are orthonormal with respect to the scalar product $e^T(F''(0))^{-1}h$).

Define a function on $R^n \times \operatorname{int} G$:

$$L(\phi, x) = -F(x) + \phi^T A x, \tag{4.21}$$

and let

$$F^+(\phi) = \max\{L(\phi,x) \mid x \in \operatorname{int} G\}. \tag{4.22}$$

F^+ is obtained from the Legendre transformation of the barrier F by a homogeneous affine transformation of argument, so

$$F^+ \in \mathcal{F}(R^n, m) \tag{4.23}$$

and D^2F^+ is non-degenerate. Notice, that in the case of problem (4.19) F^+ has an explicit representation:

$$F^+(\phi) = \sum_{i=1}^{m} \left(\frac{(a_i^T\phi)^2}{1 + [1 + (a_i^T\phi)^2]^{1/2}} - \ln(1 + [1 + (a_i^T\phi)^2]) \right), \tag{4.24}$$

where a_i, $1 \leq i \leq m$, are the columns of A.

Denote the minimizer of $L(\phi, x)$ in $x \in \operatorname{int} G$ by $X(\phi)$ (this point is well-defined). For problem (4.15) one has

$$X_i(\phi) = \frac{(a_i^T\phi)}{1 + [1 + (a_i^T\phi)^2]^{1/2}}, \quad 1 \leq i \leq m. \tag{4.25}$$

Let $F^\tau(\phi) = F^+(\phi) - \tau b^T\phi$ for $\tau \geq 0$. The background of the method is formed by the following

Lemma 4.1. Let $E^+ = R^n$, $t > 0$,

$$\phi \in E_t^+ = \{\phi \in R^n \mid \phi^T b = t \ \ b^T b\},$$

and let F_t be the restriction of F^+ onto E_t^+. Let also λ_ϕ = $\lambda(F_t,\phi) < 1/3$ be such that for $\zeta_\phi = \omega(\lambda_\phi)\ (1 - \omega(\lambda_\phi))^{-1}$ and $\xi_\phi = \zeta_\phi/(1 - \zeta_\phi)^2$ one has $\xi_\phi < 1$. Then

(i) the solution, τ_ϕ, to the problem

$$\tau \to \max \mid \lambda(F^\tau,\phi) \leq 1$$

is well-defined and positive, and $\lambda(F^{\tau_\phi},\phi) = 1$;

(ii) the projection, $X^*(\phi)$, of the point $X(\phi)$ onto the plane $E' = \{x \in R^m \mid A\,x = \tau_\phi\, b\}$, orthogonal in the Euclidean structure on R^m, induced by the scalar product

$$\langle h,e\rangle_\phi \ = \ D^2F(X(\phi))[h,e],$$

belongs to G;

(iii) the inequality

$$t^* - \tau_\phi \leq \Phi/(t\ b^T b), \tag{4.26}$$

holds, where t is the optimal value in (4.14).

The above results lead to the following method for (4.18). Let us choose $\lambda > 0$, such that

$$0 < \lambda < \lambda_*,\ \omega(\lambda) < 1/2,$$
$$\omega(\lambda)\ (1 - \omega(\lambda))\ (1 - 2\,\omega(\lambda))^{-2} < 1, \tag{4.27}$$

and let t_0 be the solution of the equation

$$t\,\delta\ (1 - t\,\delta)^{-2} = \lambda, \quad \delta = (b^Tb/2)^{1/2}, \tag{4.28}$$

belonging to $(0,1/\delta)$.

Let $\phi_{-1} = t_0\, b \in E^+ = R^n$ and

$$t_i = \exp\left(\frac{\lambda - \lambda^+}{(1 + \lambda)\ m^{1/2}}\right) t_{i-1}, \quad i > 0. \tag{4.29}$$

Having produced $\phi_{i-1} \in E^+_{t_i}$, we find $\phi_i' \in E^+_{t_i}$, Newton's iterate of ϕ_{i-1} (Newton's method is applied to the restriction of F^+ onto $E^+_{t_i}$) and then define

$$\phi_i = (t_{i+1}/t_i)\ \phi_i' \in E^+_{t_{i+1}}.$$

Then the next iteration is performed. The approximate solution to (4.18) produced at the i-th iteration is $x_i = X^*(\phi_i)$.

By virtue of the above stated properties of the family

$\mathcal{F}^{**}(F^+, E_0^+, b)$ (see P.4.1 and (4.17)), our standard arguments prove that the implication

$$\lambda(F_{t_0}, \phi_{-1}) \le \lambda \ \Rightarrow \ (\forall\ t\ ): \lambda(F_{t_t}, \phi'_t) \le \lambda^+, \quad \lambda(F_{t_{t+1}}, \phi_t) \le \lambda$$

holds, which, by L.4.1, proves the implication

$$\lambda(F_{t_0}, \phi_{-1}) \le \lambda \ \Rightarrow \ (\forall\ t): x_t \in G,\ A\, x_t = \tau_t\, b,$$

$$\varepsilon_t \equiv (t^* - \tau_t)/t^* \le \Omega \exp(- \frac{\lambda - \lambda^+}{(1 + \lambda)\, \vartheta^{1/2}}\ t\ ), \qquad (4.30)$$

$$\Omega = \vartheta/(t^*\ t_0\ b^T\ b). \qquad (4.31)$$

Let us verify that the premise in (4.30) is true. Indeed, obviously $\lambda(F_{t_0}, \phi_{-1}) \le \lambda(F^+, \phi_{-1})$. We have

$$D^2F^+(0)[\zeta,\zeta] = \frac{1}{2}\ \zeta^T A((F^*)''(0))A^T\zeta, = \frac{1}{2}\ \zeta^T A(F''(0))^{-1}A^T\zeta = \frac{1}{2}\ \zeta^T\ \zeta,$$

which implies $D^2F^+(0)[b,b] = \delta^2$; so, by T.1.1, we have for $0 < t\ \delta < 1$:

$$|D^2F^+(tb)[b,\zeta]| \le (1 - t\ \delta)^{-2}\ \delta\ (D^2F^+(0)[\zeta,\zeta])^{1/2},$$

which together with the relation $DF^+(0) = 0$ leads to

$$|DF^+(tb)[\zeta]| \le t\ \delta\ (1 - t\ \delta)^{-1}\ (D^2F^+(0)[\zeta,\zeta])^{1/2}$$

for each ζ, or, by virtue of T.1.1, to

$$|DF^+(tb)[\zeta]| \le t\ \delta\ (1 - t\ \ \delta)^{-2}\ (D^2F^+(tb)[\zeta,\zeta])^{1/2},$$

which means that $\lambda(F^+, tb) \le t\ \delta\ (1 - \ t\ \ \delta)^{-2}$. The resulting inequality, by virtue of the choice of t_0, leads to the desired relation $\lambda(F_{t_0}, \phi_{-1}) \le \lambda(F^+, \phi_{-1}) \le \lambda$.

To obtain the efficiency bound for the above method, it remains to evaluate Ω. Let us prove that $\Omega \le 2^{1/2}\vartheta/\lambda$. Indeed, since $A(F''(0))^{-1}A^T = I_n$, the point $w = (F''(0))^{-1}A^T\ \ b$ is the nearest to 0 (in the Euclidean metric on R^m, induced by the scalar product $(h,e) = h^T\ F''(0)\ e$) point of the plane $\{x \mid A\,x = b\}$. The ellipsoid $W = \{x \in R^m \mid x^T F''(0) x \le 1\}$ is contained in G (C.1.2), which implies $t^* \ge (w^T\ \ F''(0)\ \ w)^{-1/2} = |b|_2^{-1}$.

Besides this, obviously $t_0 \geq \lambda/(2\delta) = \lambda/(2^{1/2}|b|_2)$, hence

$$\Omega = \vartheta/(t^* t_0 |b|_2^2) \leq 2^{1/2}\vartheta/\lambda.$$

Q.E.D.

Now we obtain from (4.30) the following bound for the relative accuracy of x_i:

$$\varepsilon_i \leq 2^{1/2} \vartheta \lambda^{-1} \exp\left(- \frac{\lambda - \lambda^+}{(1+\lambda)\vartheta^{1/2}} i \right). \qquad (4.32)$$

The optimal choice of λ is

$\lambda = 0.206...;$

under this choice for each $\varepsilon \in (0,1)$ the inequality $\varepsilon_i \leq \varepsilon$ holds for all i such that

$$i \geq N(\varepsilon) \equiv 8.8\, \vartheta^{1/2} \ln(7 \vartheta \varepsilon^{-1}) + 1. \qquad (4.33)$$

Notice that the implementation of the dual parallel trajectories method needs an explicit representation of the Legendre transformation of F; this condition is satisfied for LP problems formatted as in (4.19).

The arithmetic cost per iteration for the above method as applied to (4.19) is $O(m\, n^2)$. A Karmarkar's type speed-up for this situation which reduces the cost to $O(m^{1/2}\, n^2)$ is described in [Ne. 1988 1,4].

4.3. Primal parallel trajectories method [Ne. 1988 2,3].

Consider a problem

$$c^T x \to \max \mid x \in G, \qquad (4.34)$$

where $G \in C_B(R^n)$. Assume that we are given a ϑ-self-concordant barrier F for G and we know the F-center of G; let this center be O:

$$O \in \operatorname{int} G,\ F'(O) = O \qquad (4.35)$$

(from now on F', F'' are the gradient and Hessian F with respect to the standard Euclidean structure on R^n). Without loss of generality assume $c^T c = 1$.

The primal parallel trajectories method for (4.34) is defined by parameters λ_1, λ_2, such that

$$0 < \lambda_1 < \lambda_*;\ 0 < \lambda_2 \leq 1/3; \tag{4.36}$$

$$\lambda_1^+ + \lambda_2 (1 - \lambda_2)^{-1} \leq \lambda_1 (1 - \lambda_2). \tag{4.37}$$

The method is as follows.

1. *Initialization.* Let

$$\tau_0 = \max(\tau \leq 1 \mid \tau (1 - \tau)^{-2} \leq \lambda_1),$$
$$e = (c^T [F''(0)]^{-1} c)^{-1/2} [F''(0)]^{-1} c,$$
$$x_{-1} = \tau_0 e. \tag{4.38}$$

2. *The i-th step.* Let $x_{i-1} \in \text{int } G$ be the previous approximate solution. Denote the set

$$\{y \in \text{int } G \mid c^T (y - x) = 0\}$$

by $E(x)$, and the restriction of F onto $E(x)$ - by $F_x(y)$. Let $x_i^+ \in E(x_{i-1})$ be Newton's iterate of x_{i-1} (Newton's method is applied to $F_{x_{i-1}}(\ )$; it will be shown that $x_i^+ \in \text{int } G$). Having produced x_i^+, we define x_i as

$$x_i = x_i^+ + \lambda_2 (c^T [F''(x_i^+)]^{-1} c)^{-1/2} [F''(x_i^+)]^{-1} c \tag{4.39}$$

(it will be shown that $x_i \in \text{int } G$). The i-th step is over.

Let $t^* = \max(c^T x \mid x \in G)$, $\Delta = (0, t^*)$ ($t^* > 0$ by (4.35)) and let $G^* = \{x \in G \mid c^T x \geq 0\}$. For each $t \in \Delta$ the set $G_t = \{x \in G \mid c^T x = t\}$ is defined. The restriction F_t of F onto the relative interior of G_t, by virtue of P.3.2.(1), is a ϑ-self-concordant barrier for G_t (the latter set is regarded as a full-dimensional subset of the corresponding hyperplane). Since G_t is bounded, F_t attains its minimum over the relative interior of G_t at the unique point $x^*(t)$ (P.3.2.(v)). By definition of $x^*(t)$, we have

$$F'(x^*(t)) = \delta(t)\, c \tag{4.40}$$

for certain $\delta(t)$ ($\delta(t) \geq 0$ by (4.35)). C^3-smoothness of F and the nondegeneracy of D^2F imply that $x^*(t)$ and $\delta(t)$ are C^2-smooth on Δ.

The main result on the primal parallel trajectories method is as follows.

Proposition 4.3. The primal parallel trajectories method is well-defined: for all i the points x_{i-1}, x_i^+ and x_i are well-defined and belong to int G. Moreover, for each $i \geq 0$ we have:

$$t_{i-1} \equiv c^T x_{i-1} > 0, \qquad (4.41_i)$$

$$\lambda(F_{t_{i-1}}, x_{i-1}) \leq \lambda_1, \qquad (4.42_i)$$

$$\delta(t_i) \geq (1 + \Omega\, \vartheta^{-1/2})\, \delta(t_{i-1}),$$

$$\Omega = \frac{5}{9}\,(1 - (1 - 3\lambda_2\omega(\lambda_1))^{5/3}), \qquad (4.43_i)$$

$$t^* - t_i \leq \vartheta\, \delta^{-1}(t_i). \qquad (4.44_i)$$

Moreover, the relative error of the i-th iterate satisfies the inequality

$$\varepsilon_i \equiv (c^T x^*)^{-1}\, (c^T x^* - c^T x_i) \leq$$
$$\leq \vartheta\, \gamma^{-1} \exp(-i \ln(1 + \Omega\, \vartheta^{-1/2})), \qquad (4.45)$$

where γ depends on λ_1, λ_2 only.

We see that the rate of convergence of the primal parallel trajectories method is the same as that one for our previous methods: it needs no more than $O(\vartheta^{1/2} \ln(2\vartheta/\varepsilon))$ iterations to produce an approximation x_i such that $\varepsilon_i \leq \varepsilon \in (0,1)$, the constant factor in $O(\ )$ depends on λ_1, λ_2 only.

Rational choice of the parameters is

$$\lambda_1 = 0.266, \quad \lambda_2 = 0.096;$$

under this choice (4.45) leads to

$$\varepsilon_i \leq 11.78\, \vartheta \exp(-i \ln(1 + 0.107\, \vartheta^{-1/2})).$$

4.4. Proofs of the results.

4.4.1. Theorem 4.1.

Let us verify that relations $(\Sigma.1)$, $(\Sigma.2)$, $(\Sigma.3)$ hold.
$(\Sigma.1)$ is obvious.

By T.3.2.(ii) the function

$f_t(x) = \ln(1/(t - f(x)))$

for each $t \in \Delta$ belongs, as a function of x, to $\mathfrak{s}(\{x|f(x)<t\},1)$. Since $\zeta \geq 1$, we have $\zeta f_t \in \mathfrak{s}(\{x|f(x)<t\},\zeta)$. Therefore by virtue of P.3.2.(iii) the inclusion

$$F_t \in \mathfrak{s}(\mathrm{Cl}\ Q_t,\vartheta^*),\quad \vartheta^* = \vartheta + \zeta$$

holds, so $F_t \in S_1^+(Q_t,E)$, which is required in $(\Sigma.2)$ when α is chosen in accordance with (4.1).

Let us verify $(\Sigma.3)$. Let

$$X^+(\varkappa) = \{(t,x) \in Q_* \mid \lambda(F_t,x) < \varkappa'\}$$

(from now on we use the notations from 2.1). It is clear that $X^+(\varkappa)$ is a neighbourhood of $X(\varkappa)$ in Q_*.

Let us verify that the set $X(\varkappa)$ is closed in E_Δ. Indeed, let $(t_i,x_i) \in X(\varkappa)$ and $(t_i,x_i) \underset{i\to\infty}{\to} (t,x)$, where $t \in \Delta$. By T.1.3.(iii) and in view of $F_\tau \in S_1^+(Q_\tau,E)$ we have

$$F_{t_i}(x_i) \leq \phi(t_i) + c$$

for certain constant c, where

$$\phi(\tau) = \min\{F_\tau(x) \mid x \in Q_\tau\},\ \tau \in \Delta.$$

The function ϕ is obviously bounded along the sequence (t_i) (because this sequence converges to a point from Δ), so $(F_{t_i}(x_i))$ is bounded. The latter in view of the definition of F_τ implies the inclusion $x \in Q_t$, or $(t,x) \in Q_*$. Thus, the closure of $X(\varkappa)$ in E_Δ is contained in Q_*; since $\lambda(\varkappa)$ is obviously closed in Q_*, $X(\varkappa)$ is closed in E_Δ, Q.E.D.

It remains to verify that under the parameters choice in accordance with (4.1) for our $X^+(\varkappa)$ the relations (2.2), (2.3) hold.

Let us fix $(t,x) \in X^+(\varkappa)$; then

$$\lambda(F_t,x) < \varkappa'. \tag{1}$$

Let

$$x^* = \mathrm{argmin}\{F_t(x) \mid x \in Q_t\}$$

(the existence and the uniqueness of x^* follows from P.3.2.(v) since Q_t is bounded and $F_t \in S_1^+(Q_t,E)$; notice that, by the same reasons, D^2F_t is non-degenerate on Q_t). Let us introduce an Euclidean structure on E with the help of scalar product

$$\langle h,s\rangle = D^2F_t(x^*)[h,s];$$

denote the corresponding norm by $\| \ \|$. Let W be the open unit ball centered at x^*. By P.3.2.(v) and in view of the inclusion $F_t \in \mathfrak{s}(\mathrm{Cl}\ Q_t, \vartheta^*)$ we have:

$$W \subset Q_t; \quad Q_t \subset \{y \mid \|y - x^*\| \leq (1 + 3\vartheta^*)\}. \qquad (2)$$

Furthermore, in view of (1) and T.1.3.(iii) we have ($\beta = \omega(\varkappa') < 1$):

$$D^2F_t(x)[x^* - x,\ x^* - x] < \beta^2, \quad \|x - x^*\| \leq \beta/(1 - \beta), \qquad (3)$$

whence by T.1.1

$$D^2F_t(x^*)[h,h] \geq (1 - \beta)^2\ D^2F_t(x)[h,h]. \qquad (4)$$

In view of (1) and (4) we have

$$\|\nabla F_t(x)\| \leq \varkappa'/(1 - \beta). \qquad (5)$$

Let u^* be the minimizer of f on $\mathrm{Cl}\ Q_t$ (or, which is the same, on G). Then, taking into account (2), we have

$$\delta \equiv (\varkappa'\ (1 - \beta)^{-1})\ (\beta\ (1 - \beta)^{-1} + 1 + 3\vartheta^*) \geq$$

$$\geq \langle \nabla F_t(x),\ x - u^*\rangle = (t - f(x))^{-1}\ \zeta\ \langle \nabla f(x),\ x - u^*\rangle +$$

$$+ \langle \nabla F(x),\ x - u^*\rangle \geq (t - f(x))^{-1}\ \ \zeta\ (f(x) - t^*) - \vartheta$$

(we have taken into account (3.6)). Thus,

$$(t - f(x))^{-1}\ (f(x) - t^*) \leq \zeta^{-1}(\delta + \vartheta) \equiv \Omega - 1, \qquad (6)$$

or

$$(t - f(x))^{-1} \leq \Omega\ (t - t^*)^{-1}; \qquad (7)$$

(4.5) is proved.

Now we have

$$|(DF_t(x)[h])'_t| = \zeta\ (t - f(x))^{-1}|Df_t(x)[h]| \leq$$

$$\leq \zeta\ (t - f(x))^{-1}\ (D^2f_t(x)[h,h])^{1/2} \leq$$

$$\leq \zeta^{1/2}\ (t - f(x))^{-1}\ (D^2F_t(x)[h,\ h])^{1/2},$$

which together with (4.1) and (7) implies (2.2) (we have taken into account that $f_t \in \mathfrak{s}(\{x|f(x)<t\},1)$).

By the same arguments

$$|(D^2F_t(x)[h,h])'_t| \leq 2\ \zeta\ (t - f(x))^{-1}\ |D^2f_t(x)[h,h]| \leq$$

$\leqslant 2\ (t - f(x))^{-1}\ D^2F_t(x)[h,h]$,

which in view of (4.1) and (7) implies (2.3). ■

4.4.2. Proposition 4.1.

$(\Sigma.1)$ obviously holds; $(\Sigma.2)$ immediately follows from the inclusion $F^* \in S_1(E^*,E^*)$ and from P.1.1.(1). Let us verify $(\Sigma.3)$; namely, let us prove that (2.2) and (2.3) hold for $X^+(\varkappa) = Q_* \equiv \Lambda \times E_0^*$. Indeed, let us fix $\psi \in E_0^*$, $t \in \Lambda$ and let $\Phi = t\ \psi + t\ b$. Then for $\zeta \in E_0^*$ we have:

$$DF_t(\psi)[\zeta] = t\ DF^*(\Phi)[\zeta],\quad D^2F_t(\psi)[\zeta,\zeta] = t^2\ D^2F^*(\Phi)[\zeta,\zeta],$$

$$|(DF_t(\psi)[\zeta])'_t - (\ln(t))'_t\ DF_t(\psi)[\zeta]| = |D^2F^*(\Phi)[\Phi,\zeta]| \leqslant$$

$$\leqslant t^{-1}\ (D^2F^*(\Phi)[\Phi,\Phi])^{1/2}\ (t^2\ D^2F^*(\Phi)[\zeta,\zeta])^{1/2} \leqslant$$

$$\leqslant t^{-1}\ \vartheta^{1/2}\ (D^2F_t(\psi)[\zeta,\zeta])^{1/2} = \xi(t)\ \alpha^{1/2}(t)\ (D^2F_t(\psi)[\zeta,\zeta])^{1/2}.$$

Furthermore,

$$|(D^2F_t(\psi)[\zeta,\zeta])'_t - (\ln(t^2))'_t\ D^2F_t(\psi)| = t\ |D^3F^*(\Phi)[\zeta,\zeta,\Phi]| \leqslant$$

$$\leqslant 2\ t\ D^2F^*(\Phi)[\zeta,\zeta]\ (D^2F^*(\Phi)[\Phi,\ \Phi])^{1/2} \leqslant$$

$$\leqslant 2\ t^{-1}\ D^2F_t(\psi)[\zeta,\zeta]\ \vartheta^{1/2} = 2\ \eta(t)\ D^2F_t(\psi)[\zeta,\zeta].$$

Inequalities (2.2), (2.3) are proved. ■

4.4.3. Proposition 4.2.

A. Since G is bounded, F is strongly convex on int G (P.3.2.(v)). Therefore (int G,F) is a $(1,E)$ - pair (see 1.6). By P.1.3 (E^*,F^*) is a $(1,E^*)$ - pair, so $F^* \in S_1^+(E^*,E^*)$ and D^2F^* is non-degenerate on E^*. A straightforward computation (see the proof of P.1.3) gives for $\Phi(x) = DF(x)[\ ]$: int $G \to E^*$:

$$D^2F^*(\Phi(x))[\Phi(x),\Phi(x)] = D^2F(x)[(\Phi'(x))^{-1}\Phi(x),(\Phi'(x))^{-1}\Phi(x)] =$$

$$= \langle\Phi(x),(\Phi'(x))^{-1}\Phi(x)\rangle = \lambda^2(F,x) \qquad (1)$$

(the latter - in view of $DF(x)[h] = \langle\Phi(x),h\rangle$, $D^2F(x)[h,h] = \langle\Phi'(x)h,\ h\rangle$, $h \in E$).

We have $\Phi(\text{int}\ G) = E^*$, thus (1) and the inclusion $F^* \in S_1^+(E^*,E^*)$ imply (i). (ii) follows from the standard properties of the Legendre transformation. A is proved. B can be proved by direct inversion of the above arguments. ■

4.4.4. Lemma 4.1.

Notice that $DF^+(O) = O$ (since $DF(O) = O$) and F^+ is strongly convex (P.4.2; recall that A is a matrix with full row rank). Hence $F^+(\phi)$ tends to ∞ as $|\phi| \to \infty$, and the minimizers ϕ_t^* of F_t are well defined. It is clear that $\nabla F^+(\phi_t^*) = \tau^*(t)\, b$ and that the function $\tau^*(t)$ increases on the positive ray.

Denote ϕ_t^* by ϕ^*, $\tau^*(t)$ by τ^*, and let $\Phi(\psi) = F^{\tau^*}(\psi)$.

(i): consider E^+ as being provided by a scalar product $\langle u,v\rangle = D^2\Phi(\phi^*)[u,v]$ and let $|\ \ |$ denotes the corresponding norm, $\Phi'(u)$, $\Phi''(u)$ denote the corresponding gradient and Hessian of Φ, respectively. By (4.23) and by virtue of the arguments from the beginning of the proof we have $\Phi \in S_1^+(E^+, E^+)$. Applying T.1.3.(iii) to the restriction, Ξ, of the function Φ onto E_t^+ and taking into account that $\lambda(\Xi,\phi) = \lambda_\phi < 1/3$, we get $|\phi - \phi^*| \leq \zeta_\phi$. Since $\zeta_\phi < 1$, we have

$$|\Phi''(\phi^* + s(\phi^* - \phi))| \leq (1 - s\,\zeta_\phi)^{-2},\ 0 \leq s \leq 1$$

(T.1.1). Moreover, $\Phi'(\phi^*) = O$; thus

$$|\Phi'(\phi)| \leq \zeta_\phi/(1 - \zeta_\phi).$$

Applying T.1.1, we get

$$\lambda(\Phi,\phi) \leq \zeta_\phi/(1 - \zeta_\phi)^2.$$

By the condition of Lemma the latter quantity is ≤ 1, thus τ_ϕ is well defined and positive. Moreover, we have

$$\tau_\phi \geq \tau^*(t). \qquad (1)$$

The latter equality in (1) is obvious. (1) is proved.

(ii): let $F^*(v)$ be the Legendre transformation of F, thus $F^+(\phi) = F^*(A^T\phi)$. Let

$$x = X(\phi), \quad \Xi(\psi) = F^+(\psi) - \tau_\phi\, b^T\psi\,, \quad \tau = \tau_\phi.$$

Replacing first and second order differentials by gradients and Hessians corresponding to the standard Euclidean structure, we get, in view of the standard properties of the Legendre transformation:

$$(1 = \lambda(\Xi,\phi)) \Rightarrow ((\Xi'(\phi))^T u \leq [(\Xi''(\phi)u)^T u]^{1/2},\ u \in R^n) \Rightarrow$$

$$\Rightarrow ((A\,x - \tau\, b)^T u \leq [(A\,H''(A^T\phi)\,A^T u)^T u]^{1/2},\ u \in R^n) \Rightarrow$$

$$\Leftrightarrow \{(A x - \tau b)^T u \leq \{[(A (F''(x))^{-1} A^T u)^T u]^{1/2}, u \in R^n\}. \quad (2)$$

Let x^* be the projection involved into (ii). Then

$$F''(x)(x - x^*) = A^T u^*$$

for certain $u^* \in R^n$, and

$$A x - \tau b = A (x - x^*);$$

the latter inequality in (2) as applied to $u = u^*$ leads to

$$(A (x - x^*))^T u^* \leq [(A (F''(x))^{-1} F''(x) (x - x^*))^T u^*]^{1/2};$$

in view of $F''(x)(x - x^*) = A^T u^*$ we get

$$(x - x^*)^T F''(x) (x - x^*) \leq [(x - x^*)^T F''(x) (x - x^*)]^{1/2},$$

whence $(x - x^*)^T F''(x) (x - x^*) \leq 1$. So the ellipsoid

$$\{y \in R^m \mid D^2F(x)[y-x,y-x] \leq 1\}$$

contains x^*. This ellipsoid is contained in G (P.3.2.(iv.1)), hence $x^* \in G$. (ii) is proved.

(iii): by the standard duality arguments, $X(\phi^*) \equiv x^*$ belongs to the set $G' = \{x \in \text{int } G \mid A x = \tau^*(t) b\}$, minimizes F over this set, and

$$(\forall w \in R^m): A w = 0 \Rightarrow w^T F'(x^*) = 0; \quad F'(x^*) = A^T \phi^*. \quad (3)$$

Let y^* be the solution to (4.18), and let

$$u^* = (t^*/\tau^*(t)) x^*.$$

Then the premise in (3) holds for $w = y^* - u^*$, which leads to

$$(u^* - x^*)^T F'(x^*) = (y^* - x^*)^T F'(x^*) \leq \vartheta \quad (4)$$

(the latter - by (3.6) and since F is a ϑ-s.c. barrier for G). The equality in (3), together with the obvious relation

$$A (u^* - x^*) = (t^* - \tau^*(t)) b$$

and (4), implies $(t^* - \tau^*(t)) b^T \phi^* \leq \vartheta$, whence, in view of $\phi^* \in E_t^+$, i.e. of $b^T \phi^* = t\, b^T b$),

$$t^* - \tau^*(t) \leq \vartheta/(t\, b^T b).$$

This inequality together with (1) proves (iii). ∎

4.4.5. Proposition 4.3.

1^O. Let us establish some properties of $x^*(t)$ and $\delta(t)$. Taking the derivative in t in (4.40) and in the identity

$c^T x^*(t) = t$, we get

$$\delta'(t) = c^T [F''(x^*(t))]^{-1} c,$$

$$(x^*(t))' = (c^T [F''(x^*(t)]^{-1} c)^{-1} [F''(x^*(t))]^{-1} c. \quad (1)$$

Let us fix $\tau \in \Delta$, and let $\|h\|_\tau = (h^T F''(x^*(\tau)) h)^{1/2}$. (1) implies

$$\|(x^*(\tau))'\|_\tau = (c^T [F''(x^*(\tau))]^{-1} c)^{-1/2} \equiv \phi(\tau). \quad (2)$$

Moreover, by T.1.1 and in view of (2) we have

$$\|x^*(\tau) - x^*(t)\|_t < 1 \Rightarrow$$

$$\Rightarrow \|(x^*(\tau))'\|_\tau \leqslant \phi(t) (1 - \|x^*(\tau) - x^*(t)\|_t)^{-1} \Rightarrow$$

$$\Rightarrow \|(x^*(\tau))'\|_t \leqslant \phi(t) (1 - \|x^*(\tau) - x^*(t)\|_t)^{-2}. \quad (3)$$

By C.1.2 the set $\{y \in R^n \mid \|y - x^*(t)\|_t < 1\}$ is contained in int G, which, together with (3), proves the implication

$$0 \leqslant \tau - t < (3\,\phi(t))^{-1} \Rightarrow$$

$$\Rightarrow \tau \in \Delta,\ \|x^*(\tau) - x^*(t)\|_t \leqslant 1 - (1 - 3(\tau - t)\,\phi(t))^{1/3},$$

$$\delta'(\tau) \geqslant \phi^2(t) (1 - 3(\tau - t)\,\phi(t))^{2/3} \quad (4)$$

(we have taken into account that

$\|x^*(\tau) - x^*(t)\|_t < 1 \Rightarrow$

$\Rightarrow c^T[F''(x^*(\tau)]^{-1}c \geqslant (1 - \|x^*(\tau) - x^*(t)\|_t)^2\, c^T[F''(x^*(t))]^{-1}c$

(see T.1.1) and have used (1)).

Now let us prove (4.41_i) - (4.44_i). Let

$$J = \{ i \geqslant 0 \mid (4.41_j),\ (4.42_j) \text{ hold for } 0 \leqslant j \leqslant i,$$
$$(4.43_j),\ (4.44_j) \text{ hold for } 0 \leqslant j < i,$$
$$x_{j-1} \in \text{int}\, G,\ 0 \leqslant j \leqslant i,\ x_j^+ \in \text{int}\, G,\ 0 \leqslant j < i\}.$$

We desire to prove that $J = \{i \geqslant 0\}$; it is sufficient to verify that $0 \in J$ and that

$$j \in J \Rightarrow (j + 1) \in J.$$

Let us first verify that $0 \in J$, i.e. that $x_{-1} \in$ int G, $t_{-1} > 0$ and (4.42_0) holds. By (4.38) we have $e^T F''(0)\, e = 1$, whence $\tau e \in$ int G for $0 \leqslant \tau < 1$, and, in view of T.1.1 and the relation $F'(0) = 0$, $\|F'(\tau e)\|_0 \leqslant \tau (1 - \tau)^{-1}$, or

$$|h^T F'(\tau e)| \leqslant \tau (1 - \tau)^{-1} (h^T F''(0)\, h)^{1/2} \leqslant$$

$$\leqslant \tau (1 - \tau)^{-2} (h^T F''(\tau e)\, h)^{1/2},$$

whence $\lambda(F, \tau_0 e) \leq \lambda_1$. It is clear that

$$\lambda(F_{t_{-1}}, \tau_0 e) \leq \lambda(F, \tau_0 e),$$

which implies (4.42_0). Furthermore,

$$t_{-1} = c^T \tau_0 e = \tau_0 (c^T [F''(0)]^{-1} c)^{1/2},$$

which, by virtue of the obvious unequality $\tau_0 \geq \lambda_1/2$, implies

$$t_{-1} \geq (c^T [F''(0)]^{-1} c)^{1/2} \lambda_1/2; \tag{5}$$

in particular, $t_{-1} > 0$. So $0 \in J$.

Now let $i \in J$; let us verify that then $i + 1 \in J$. First of all, in view of $\lambda_1 < \lambda_*$ and the fact that F_{t_i} is a barrier for the set $G_{t_{i-1}}$, (4.42_i) implies (see T.1.4) the relation

$$x_i^+ \in \mathrm{int}_0 G_{t_{i-1}} \subset \mathrm{int}\, G, \; \lambda(F_{t_{i-1}}, x_i^+) \leq \lambda_1^+ \tag{6}$$

(int_0 denotes the relative interior). Furthermore, let

$$e_i = (c^T [F''(x_i^+)]^{-1} c)^{-1/2} [F''(x_i^+)]^{-1} c;$$

then

$$x_i = x_i^+ + \lambda_2 e_i, \quad e_i^T F''(x_i^+) e_i = 1, \tag{7}$$

whence, in view of C.1.2 and the inclusion $\lambda_2 \in (0,1)$, $x_i \in$ int G. We have

$$t_i = c^T x_i = c^T x_i^+ + \lambda_2 c^T e_i = c^T x_i + $$
$$+ \lambda_2 (c^T [F''(x_i^+)]^{-1} c)^{1/2} = t_{i-1} + \lambda_2 (c^T [F''(x_i^+)]^{-1} c)^{1/2}. \tag{8}$$

In particular, $t_i > t_{i-1}$, and (4.41_{i+1}) holds.

By virtue of (6) and T.1.3.(iii) we have

$$(x^*(t_{i-1}) - x_i^+) F''(x_i^+) (x^*(t_{i-1}) - x_i^+) \leq \omega^2(\lambda_1^+),$$

whence, by T.1.1,

$$(1 - \omega(\lambda_1^+)) \phi^{-1}(t_{i-1}) \leq (c^T [F''(x_i^+)]^{-1} c)^{1/2} \leq$$
$$\leq (1 - \omega(\lambda_1^+))^{-1} \phi^{-1}(t_{i-1}).$$

Therefore (8) implies

$$t_i \geq \tau_i = t_{i-1} + \lambda_2 (1 - \omega(\lambda_1^+)) \phi^{-1}(t_{i-1}). \tag{9}$$

Since $\lambda_2 < 1/3$, the relations (9) and (4) imply the relations

$$\delta'(\tau) \geq \phi^2(t_{i-1})(1 - 3(\tau - t)\phi(t_{i-1}))^{2/3}, \quad t_{i-1} \leq \tau \leq \tau_{i-1}, \tag{10}$$

whence, since δ obviously is increasing,

$$\delta(t_i) \geq \delta(t_{i-1}) + \Omega\,\phi(t_{i-1}),$$
$$\Omega = \frac{5}{9}\,(1 - (1 - 3\lambda_2(1 - \omega(\lambda_1^+)))^{5/3}). \tag{11}$$

Since
$$\phi(t) = (c^T\,[F''(x^*(t))]^{-1}\,c)^{-1/2} =$$
$$= \delta(t)((F'(x^*(t)))^T\,[F''(x^*(t))]^{-1}\,F'(x^*(t)))^{-1/2} \geq \delta(t)\vartheta^{-1/2},$$
we get
$$\delta(t_i) \geq \delta(t_{i-1})\,(1 + \Omega\,\vartheta^{-1/2});$$
this is (4.43_i).

Furthermore, let x^* be the solution to (4.34). Then, by (3.6), we have
$$t^* - t_i = c^T\,(x^* - x_i) = c^T\,(x^* - x^*(t_i)) =$$
$$= \delta^{-1}(t)\,(F'(x^*(t_i)))^T\,(x^* - x^*(t_i)) \leq \vartheta\,\delta^{-1}(t_i),$$
which implies (4.44_i).

To prove the inclusion $i + 1 \in J$ it remains to verify that (4.42_{i+1}) holds. Let $c^T h = 0$ and let
$$x(s) = x_i^+ + s\,e_i,\ 0 \leq s \leq \lambda_2.$$
Then
$$\left|\frac{d}{ds}\,(F'(x(s))^T h\right| = |e_i^T\,F''(x(s))\,h| \leq$$
$$\leq (e_i^T\,F''(x(s))\,e_i)^{1/2}\,(h^T\,F''(x(s))\,h)^{1/2} \leq$$
$$\leq (1 - s)^{-2}\,(e_i^T\,F''(x(0)\,e_i)^{1/2}\,(h^T\,F''(x(0))\,h)^{1/2} =$$
$$(1 - s)^{-2}\,(h^T\,F''(x(0))\,h)^{1/2}$$
(we have taken into account that $e_i^T\,F''(x(0))\,e_i = 1$, and have used T.1.1). By (6) the relation $c^T h = 0$ implies
$$|h^T\,F'(x(0))| \leq \lambda_1^+\,(h^T\,F''(x(0))\,h)^{1/2},$$
so
$$|h^T\,F'(x(\lambda_2))| \leq (\lambda_1^+ + \lambda_2(1 - \lambda_2)^{-1})\,(h^T\,F''(x(0))\,h)^{1/2} \leq$$
$$\leq (1 - \lambda_2)^{-1}\,(\lambda_1^+ + \lambda_2\,(1 - \lambda_2)^{-1})\,(h^T\,F''(x(\lambda_2))\,h)^{1/2}$$
(we have used T.1.1). The resulting inequality means that
$$\lambda(F_{t_i}, x_i) \leq (1 - \lambda_2)^{-1}\,(\lambda_1^+ + \lambda_2(1 - \lambda_2)^{-1}),$$
which, by virtue of (4.37), leads to (4.42_{i+1}). The proposition is proved.

Section 5. Acceleration of the barrier method. I.

5.1. Introduction.

In this Section and in the next one we consider problems as follows:

$$(\mathcal{P}):\quad \psi(x) \equiv \frac{1}{2}\, x^T A\, x - a^T x \to \min \mid x \in R^n,$$
$$f_i(x) \equiv - a_i^T x + b_i \geqslant 0,\ 1 \leqslant i \leqslant m, \tag{5.1}$$

where A is a positive semidefinite symmetric $n \times n$ - matrix, $a,\ a_1,\dots,a_m \in R^n,\ b_1,\dots,b_m \in \mathbb{R}$. In other words, we deal with a linearly constrained convex quadratic programming problem.

From now on let

$$G = \{x \in R^n \mid f_i(x) \geqslant 0,\ 1 \leqslant i \leqslant m\}.$$

We assume that G is bounded set with a nonempty interior (hence $m > n$); without loss of generality we suppose that

$$a_i \neq 0,\ 1 \leqslant i \leqslant m. \tag{5.2}$$

Then

$$G' \equiv \operatorname{int} G = \{x \in R^n \mid f_i(x) > 0,\ 1 \leqslant i \leqslant m\}.$$

5.1.1. "Multistep" return to the trajectory. Recall that if we know a starting point $w \in G'$, then we can solve (5.1) by application of a path-following method, for example, the barrier method from Sect. 3, generated by the logarithmic self-concordant barrier

$$F(x) = - \sum_{i=1}^{m} \ln(f_i(x)) : G' \to \mathbb{R}. \tag{5.3}$$

The parameter value for this barrier equals m, so an ε-solution, x_ε, such that

$$x_\varepsilon \in G' \text{ and } \psi(x_\varepsilon) - \min_G \psi \leq \varepsilon\,(\max_G \psi - \min_G \psi) \tag{5.4}$$

will be produced by our method in no more than

$$N(\varepsilon) = O(m^{1/2} \ln(m\,\delta^{-1}\varepsilon^{-1})) \tag{5.5}$$

iterations with no more than $O(m\,n^2)$ arithmetic operations per iteration; herein $\delta = 1 - \pi_P(w)$, π_P is the Minkovsky's function of G with the pole at the minimizer, $x(F)$, of the barrier. Thus, the total arithmetic cost of an ε-solution does not exceed

$$M(\varepsilon) = O(m^{3/2}\,n^2 \ln(m\,\delta^{-1}\,\varepsilon^{-1})). \tag{5.6}$$

It is well-known that the cost given by (5.6) can be reduced. The idea of the acceleration originates from [Ka. 1984]; it is based on the use of approximations to the inversed Hessians instead of the exact inversed Hessians when computing the Newton directions. The compatibility (within a factor of order 1) of these approximations and the exact inversed Hessians is maintained by 1-rank corrections; it reduces the average (over the iterations) cost of an iteration by a factor $O(m^{1/2})$, so the total arithmetic cost of an ε-solution becomes

$$M^+(\varepsilon) = O(m\,n^2 \ln(m\,\delta^{-1}\,\varepsilon^{-1})). \tag{5.7}$$

This *Karmarkar's speed-up* is implemented in most of the papers mentioned in Sect. 0.

Notice that the above speed-up does not change the size of steps in the trajectory's parameter, t. But we can make a "large" step in t and then try to return onto the trajectory using an appropriate multistep procedure. In what follows we realize such an approach - it is the first purpose of this section. Our worst-case efficiency bound turns out to be the same as in (5.7), but now it is the worst-case bound and we can hope that on the real-world problems the behaviour of the method will be better; at the same time the usual acceleration with fixed size of steps in t gives no basis for such a hope.

5.1.2. "Advanced" linear algebra. The second purpose of Sect. 5, 6 is as follows. The improvement due to Karmarkar's speed-up strategy depends on which linear algebra techniques is used. The improvement mentioned corresponds to the traditional linear algebra where the inversion of a $k \times k$-matrix costs $O(k^3)$ arithmetic operations. It is well known that the inversion can be implemented with a lower cost of $O(k^{2+\gamma})$ operations for certain $\gamma < 1$ (the best known value of γ is $0.376...$ [CW. 1986]). Of course, such "advanced" linear algebra reduces the cost of the solution of $(\mathcal{P})$. Thus, the question arises: what is the upper bound $\mathbf{M}$ for the average (over iterations) cost of an ε-solution to $(\mathcal{P})$, if the efficiency of the inversion of a $k \times k$-matrix is $O(k^{2+\gamma})$ for some $\gamma \in (0,1]$.

It turns out that Karmarkar's speed-up in such a situation yields

$$\mathbf{M} = O(m^{s(\gamma)} \ln(m\, \delta^{-1}\, \varepsilon^{-1})),\ s(\gamma) = (5+\gamma)/2 \qquad (5.8)$$

(for simplicity sake we assume that $n = O(m)$). For the barriers method described above the authors have developed another speed-up strategy (it does not reduce to Karmarkar's one even in the traditional case of $\gamma = 1$). The new strategy yields

$$\mathbf{M} = O(m^{r(\gamma)} \ln(m\, \delta^{-1}\, \varepsilon^{-1})),\ r(\gamma) = 5/2 + 2\gamma^2/(2+3\gamma-\gamma^2). \qquad (5.9)$$

Notice that

$r(0) = s(0) = 5/2,\ r(1) = s(1) = 3,$

$r(\gamma) < s(\gamma),\ 0 < \gamma < 1.$

For example, $r(0.376...) = 2.594...$, $s(0.376...) = 2.688...$. The strategy mentioned (it is described in details in Sect. 6) is based on Karmarkar's one and on certain properties of the conjugate gradient method. This strategy differs from that one of Karmarkar even in the traditional case of $\gamma = 1$.

5.1.3. Preliminary results. From now on we fix $\gamma \in (0,1]$, such that for all $k \in \mathbb{N}$ the arithmetic cost of the inversion of a $k \times k$-matrix by certain method does not exceed $c_\gamma k^{2+\gamma}$. It is well known that under this assumption the multiplication of two $k \times k$ matrices can be performed in $O_\gamma(k^{2+\gamma})$ arithmetic operations (henceforth the constant factors in $O_\gamma(\ )$ depend on γ only). The following statement is a simple corollary of these assumptions:

Lemma 5.1. Let $\sigma(l,k,r) = l\ k\ r\ (\min(l,k,r))^{\gamma-1}$ for $l,k,r \in \mathbb{N}$. The product of a $l\times k$-matrix A and a $k\times r$-matrix B can be computed at the cost of $O_\gamma(\sigma(l,k,r))$ arithmetic operations. ■

Assume that the data in (5.1) are represented in natural way (by the list of the entries of the corresponding matrix and vectors), and let ϕ be a similarly represented convex quadratic form. Let for $x \in G'$, $t > 0$

$$\delta(t,x) = (t^{-1/2} f_1^{-1}(x),\dots,t^{-1/2} f_m^{-1}(x))^T \in R^m,$$
$$d(t,x) = (t^{-1} f_1^{-2}(x),\dots,t^{-1} f_m^{-2}(x))^T \in R^m,$$
$$D(t,x) = D_t(x) = \mathrm{diag}(d(t,x)) \in \mathcal{D},$$

where $\mathcal{D}$ is the set of diagonal $m \times m$ - matrices with positive diagonal entries. Let Z be a $n\times m$ - matrix with the rows $a_1,\dots,a_m$, and let $M(\phi,D) = \phi'' + Z\ D\ Z^T$, $D \in \mathcal{D}$. The $n\times n$ - matrix $M(\phi,D)$ is symmetric and positive defined (the latter - in view of the boundness of G'). We use the notation $M_t^\phi(x)$ for the matrix $M(\phi,D_t(x))$; notice that if

$$F_t^\phi(x) = t\ \phi(x) + F(x) \tag{5.10}$$

then

$$(F_t^\phi)''(x) = t\ M_t^\phi(x),\ x \in G'. \tag{5.11}$$

(from now on f' and f'' denote the gradient and Hessian of a function f: $G' \to \mathbb{R}$ with respect to the standard Euclidean structure on R^n).

For a couple h, s of positive m-dimensional vectors let

$$\nu(h,s) = \max(h_1/s_1, s_1/h_1,\dots,h_m/s_m, s_m/h_m) - 1.$$

The following lemma holds:

Lemma 5.2. (i) Being given $x \in G'$, $t > 0$ and $D \in \mathcal{D}$, we can:

compute $(F_t^\phi)'(x)$ at the cost of $O(m\ n)$ operations;

compute $D_t(x)$ at the cost of $O(m\ n)$ onerations;

compute the product of $M(\phi,D)$ and a given vector $h \in R^n$ at the cost of $O(m\ n)$ operations;

compute $M(\phi,D)$ at the cost of $O_\gamma(m\ n^{1+\gamma})$ operations.

(ii) Let us be given D, $D' \in \mathcal{D}$ and the matrix $L = [M(\phi,D)]^{-1}$, and let k be the number of diagonal positions in which the entries of D differ from that ones of D'. Then the matrix $[M(\phi,D')]^{-1}$ can be computed at the cost of $O_\gamma(m + l(n,k))$

operations, where

$$l(n,k) = \begin{cases} n^2 k^{\gamma}, & k \leq n, \\ k\, n^{1+\gamma}, & k > n. \end{cases} \blacksquare$$

5.2. The main inequality.

Let us fix a convex quadratic form ϕ on R^n and let for $t > 0$

$$x^*(t) = \operatorname{argmin}(F_t^{\phi}(x) \mid x \in G'), \tag{5.12}$$

$$\xi^*(t) = (t^{-1/2} f_1^{-1}(x^*(t)), \ldots, t^{-1/2} f_m^{-1}(x^*(t)))^T \tag{5.13}$$

(F_t^{ϕ} is defined by (5.10)).

Now we shall prove that the trajectory ξ^* can not, in a sense, vary too quick.

Lemma 5.3. Let $t_1, t_2 > 0$. Then

$$(t_1 t_2)^{1/2} (x^*(t_1) - x^*(t_2))^T \phi'' (x^*(t_1) - x^*(t_2)) + \sum_{i=1}^{m} (\xi_i^*(t_1) - \xi_i^*(t_2))^2 (\xi_i^*(t_1)\, \xi_i^*(t_2))^{-1} = m\, (t_1^{1/2} - t_2^{1/2})^2 (t_1 t_2)^{-1/2}. \quad \blacksquare \tag{5.14}$$

Corollary 5.1. Let $t_1, t_2 > 0$. Assume that $x(t_1), x(t_2) \in G'$ are such that

$$\lambda(F_{t_j}^{\phi}, x(t_j)) \leq \lambda \leq 0.1, \; j=1,2. \tag{5.15}$$

Then

$$(t_1 t_2)^{1/2} (x(t_1) - x(t_2))^T \phi'' (x(t_1) - x(t_2)) + \sum_{i=1}^{m} (\delta_i(t_1, x(t_1)) - \delta_i(t_2, x(t_2)))^2 (\delta_i(t_1, x(t_1))\delta_i(t_2, x(t_2)))^{-1} \leq \mu_0^2 \left\{ m\, (t_1^{1/2} - t_2^{1/2})^2 + \omega^2(\lambda)(t_1^{1/2} + t_2^{1/2})^2 \right\} (t_1 t_2)^{-1/2} \tag{5.16}$$

with an absolute constant $\mu_0 > 0$. $\blacksquare$

5.3. "Multistep" barrier methods: preliminary remarks.

Recall that the barrier method, as applied to (5.1), deals with trajectory (5.12), where ϕ is some linear form at the preliminary stage and $\phi = \psi$ at the main stage. The method produces approximations, $x(t)$, to $x^*(t)$ along a sequence $\{t_i \mid i \geq 0\}$ of t's values and maintains the inequality

$\lambda(F^{\phi}_{t_\iota}, x(t_\iota)) \leqslant \lambda,$

where λ is an appropriate absolute constant. This inequality is our only restriction on the quality of approximations: the manner in which these approximations are produced is not important. In this section we describe two strategies of approximation which differ from that of Sect. 3.

To simplify the descriptions, we consider a subproblem $\mathfrak{F}(\tau, y)$ as follows. Given some $\tau > 0$ and $y \in G'$ such that

$$\lambda(F^{\phi}_{\tau}, y) \leqslant \lambda \tag{5.17}$$

we desire to produce $y' \in G'$ and τ' satisfying the relation

$$\lambda(F^{\phi}_{\tau'}, y') \leqslant \lambda, \tag{5.17'}$$

with τ' either $\leqslant \tau/2$ (at the preliminary stage) or $\geqslant 2\ \tau$ (at the main stage). This subproblem will be called Notice that if $\mathcal{M}$ is a procedure which solves this subproblem (for each τ, y satisfying (5.17)) at the cost of $\leqslant M$ arithmetic operations, then the iterative application of $\mathcal{M}$ in the manner similar to that one of Sect. 3 produces an ε-solution to (5.1) in no more than

$$M^*(\varepsilon, \delta) = O(M \ln(m\, \varepsilon^{-1}\, \delta^{-1}))$$

arithmetic operations (that statement can be proved by the arguments used in the proofs of P.3.3, P.3.4).

In what follows we deal with the above subproblem and use the corresponding notations F^{ϕ}_t, $M^{\phi}_t(x)$, $x^*(t)$, $\xi^*(t)$ (see (5.10) - (5.13)). Let also

$$\Phi_t(x) = (F^{\phi}_t)'(x), \qquad \Psi_t(x) = (F^{\phi}_t)''(x).$$

5.4. Sets $K_\alpha(x)$.

For $x \in G'$ and $\alpha > 0$ let

$$K_\alpha(x) = \{y \in G' \mid f_\iota(x)/f_\iota(y),\ f_\iota(y)/f_\iota(x) < 1 + \alpha,\ 1 \leqslant \iota \leqslant m\}. \tag{5.18}$$

Notice that for cetrain *absolute constant* $\mu_1 > 0$ and for each $r, s > 0$ one has

$$|\ln(r/s)|^2 \leqslant \mu_1^2\, (r - s)^2\, (rs)^{-1}. \tag{5.19}$$

Lemma 5.4. (1) Let $x \in G'$, $\alpha > 0$, $t, s > 0$. Then for $y \in K_\alpha(x)$

one has

$$\|(\Psi_t(x))^{-1/2}\,\Psi_s(y)\,(\Psi_t(x))^{-1/2} - I_n\| \le$$
$$\le \max\{\,2\alpha + \alpha^2\,;\ |\,1 - s/t\,|\,\} \qquad (5.20)$$

(ii) There exists an *absolute constant* $\mu_2 > 0$ such that for each $t,\ t' > 0,\ x,\ x' \in G'$ and $\lambda \in (0,\ 0.1)$ satisfying the conditions

$$\lambda(F_t^\phi, x) \le \lambda,\ \lambda(F_{t'}^\phi, x') \le \lambda,\ 1/2 \le t/t' \le 2. \qquad (5.21)$$

the following implication holds:

$$\alpha \ge \mu_2\,(m\,\ln^2(t/t')| + \omega^2(\lambda)) \ \Rightarrow\ x' \in K_\alpha(x). \ \blacksquare \qquad (5.22)$$

Let for $\alpha > 0$ and $q \in (0,1)$ the function $g(t) \equiv g(\alpha,q,t)$ be defined as

$$g(t) = \begin{cases} -\ln(\alpha/q) - (q/\alpha)\,(t - n\,\alpha) + (q/\alpha)^2\,(t - \alpha/q)^2/2, \\ \qquad \text{if } t \ge \alpha/q; \\ -\ln(t), \text{ if } q\,\alpha \le t \le \alpha/q; \\ -\ln(q\,\alpha) - (q\,\alpha)^{-1}\,(t - q\,\alpha) + (q\,\alpha)^{-2}\,(t - q\,\alpha)^2/2, \\ \qquad \text{if } 0 < t \le q\,\alpha. \end{cases}$$

Let also

$$\alpha(q) = (1 - q)/q.$$

It is clear that the function $g(t)$ is a C^2-smooth extension of the function $-\ln t$ from the segment $[q\ \alpha,\ \alpha/q]$ onto $\mathbb{R}$ (the second derivative of the extended function is constant for $t > Q\,\alpha$ and for $t < q\,\alpha$).

For $u \in G'$ let

$$F_{u,q}(x) = \sum_{i=1}^{m} g(f_i(u), q, f_i(x)) : R^n \to \mathbb{R}.$$

The following statement is obvious.

Lemma 5.5. For each $u \in G'$ we have

$$x \in K_{\alpha(q)}(u) \ \Rightarrow\ F_{u,q}(x) = F(x); \qquad (5.23)$$

moreover, for each $x \in R^n$ we have

$$q^2\,F''(u) \le F''_{u,q}(x) \le q^{-2}\,F''(u). \ \blacksquare \qquad (5.24)$$

5.5. "Multistep" barrier method - I.

Let $2 \ge \rho > 1,\ 2 > \eta > 0,\ \lambda \in (0, 0.1)$. Let us describe a procedure $\mathcal{M}(\rho,\eta,\lambda)$ which solves the subproblem $\mathcal{P}(\tau, y)$. This procedure makes a "large" step in t (the size of the step depends on γ, m, n; in the traditional case of $\gamma = 1$ and $n = O(m)$ the optimal size is

$$t \to t' = (1 \pm O(m^{-3/7}))\, t$$

instead of the usual

$$t \to (1 \pm O(m^{-1/2}))\, t)$$

and then returns into the neighbourhood of the trajectory minimizing F_t^Φ, with the help of gradient descent method. To avoid some difficulties (for example, connected with the restrictions $x \in G'$) it is convenient to apply the gradient descent method not to $F_{t'}^\Phi$ itself, but to the function of the type $F_{u,q}$ which coincide with $F_{t'}^\Phi$ in a neighbourhood of $x^*(t')$; the latter property is provided by appropriate choice of u and q. Of course, the gradient descent method corresponds not to the initial Euclidean structure, but to the structure close to that one defined by the matrix $(F_{t'}^\Phi)''$.

The procedure is as follows.

Initialization. Let

β be the positive root of the equation $\beta^2/(1 - \beta) = \lambda^2$,

$q = (1 + \mu_2\, [m \ln^2 \eta + \omega^2(\lambda)])^{-1}$,

$h = (q^{-2} \rho\, \eta)^{-1}$,

$\omega = q^3 \beta\, (\rho\, \eta)^{-3/2}$,

$x_0 = y$, $t_0 = \tau$.

Compute $d^0 = d(t_0, x_0)$ and the matrices $M_{t_0}^\Phi(x_0)$ and

$$Q_0 = (M_{t_0}^\Phi)^{-1}.$$

The k-th step, $k \geq 0$. Assume that after k-1 step of the procedure we have produced a point $x_k \in G'$, a vector $d^k \in \mathcal{D}$, a number $t_k > 0$ and a matrix

$$Q_k = (M(\Phi, D_k))^{-1}, \quad D_k \equiv \mathrm{diag}(d^k), \qquad (5.25_k)$$

such that

$$\nu(d^k, d(t_k, x_k)) \leq \rho \qquad (5.26_k)$$

(notice that the initialization rules provide (5.25_0), (5.26_0)).

At the k-th step we:

a) set

$$t_{k+1} = \begin{cases} t_k \eta \text{ at the main stage} \\ t_k \eta^{-1} \text{ at the preliminary stage} \end{cases}$$

$$p_{k+1}(x) = t_{k+1}\, \phi(x) + F_{x_k, q}(x);$$

b) minimize the function p_{k+1} over $x \in R^n$ by a gradient method corresponding to the metric defined by the matrix Q_k^{-1}, i.e. set $x_{k,0} = x_k$ and compute

$$x_{k,l+1} = x_{k,l} - t_{k+1}^{-1}\, h\, Q_k\, p'_{k+1}(x_{k,l});$$

the process is terminated at the step, l, where the condition

$$(p'_{k+1}(x_{k,l}))^T\, t_{k+1}^{-1}\, Q_k\, p'_{k+1}(x_{k,l}) \leq \omega^2$$

holds;

c) set $x_{k+1} = x_{k,l}$;

d) compute $d(t_{k+1}, x_{k+1})$ and $d^{k+1} \in \mathcal{D}$:

$$d_i^{k+1} = \begin{cases} d_i^k, & (1+\rho)^{-1}\, d_i^k \leq d_i(t_{k+1}, x_{k+1}) \leq (1+\rho) d_i^k, \\ d_i(t_{k+1}, x_{k+1}), & \text{otherwise,} \end{cases}$$

(notice that this updating provides (5.26_{k+1})) and, using Q_k, compute Q_{k+1} in accordance with (5.25_{k+1}).

The k-th step of $\mathcal{M}(\rho,\eta,\lambda)$ is over.

If $t_{k+1}/t_0 \geq 2$ (at the main stage) or $t_{k+1}/t_0 \leq 1/2$ (at the preliminary stage), then set

$$k^* = k,\ \tau' = t_{k+1},\ y' = x_{k+1}$$

and terminate, otherwise perform the $(k+1)$-th step.

Theorem 5.1. All points x_k, $0 \leq k \leq k^* + 1$, produced by $\mathcal{M}(\rho,\eta,\lambda)$, belong to G' and satisfy the relation

$$\lambda(F_{t_k}^{\phi}, x_k) \leq \lambda \qquad (5.27_k)$$

and our procedure solves ($\mathcal{P}$). Moreover,

$$k^* \leq O((\eta - 1)^{-1}). \qquad (5.28)$$

Let η satisfy the condition

$$m^{1/2}\, (\eta - 1) \geq 1. \qquad (5.29)$$

Then the arithmetic cost of $\mathcal{M}(\rho,\eta,\lambda)$ does not exceed the quantity

$$\mathcal{N}^{(1)} = O_\gamma(q^{-4}\, (\eta - 1)^{-1}\, m\, n\, \ln(m\, q^{-1}\, \lambda^{-1}) + m\, n^{1+\gamma} + (\eta - 1)^{-1+\gamma}\, (\rho - 1)^{-\gamma}\, m^\gamma\, n^2). \qquad (5.30)$$

In particular, under the parameters choice

$$\eta = 1 + m^{-(5-\gamma)/(8-\gamma)}\, (n^4/\ln m)^{1/(8-\gamma)},\ \rho = 1.5,\ \lambda = 0.1, \qquad (5.31)$$

where

$$n^* = \max\{n, m^{(2-\gamma)/2} \ln m\} \qquad (5.32)$$

we have

$$N^{(1)} \leq O_\gamma(m^{(5+2\gamma)/(8-\gamma)} (n^*)^{(15-\gamma)/(8-\gamma)} (\ln m)^{(1-\gamma)/(8-\gamma)}) \leq$$
$$\leq O_\gamma(m^{(20+\gamma)/(8-\gamma)} (\ln m)^{(1-\gamma)/(8-\gamma)}). \qquad (5.33)$$

Hence under the parameters choice (5.31) we have

$$N^*(\varepsilon,\delta) \leq O_\gamma(N^{(1)} \ln(m\,\delta^{-1}\,\varepsilon^{-1})) \leq$$
$$\leq O_\gamma(m^{(20-\gamma)/(8-\gamma)} (\ln m)^{(1-\gamma)/(8-\gamma)} \ln(m\,\delta^{-1}\,\varepsilon^{-1})). \quad \blacksquare \quad (5.34)$$

Notice that in the case of traditional linear algebra ($\gamma = 1$) (5.33) - (5.34) transforms into

$$N^*(\varepsilon,\delta) \leq O(m\,(n^*)^2 \ln(m\,\delta^{-1}\varepsilon^{-1})), \quad n^* = \max\{n, m^{1/2} \ln m\}.$$

5.6. "Multistep" barrier method - II.

It is known that the rate of convergence of the gradient descent method as applied to strongly convex problems can be improved. The implementation of the "optimal" smooth convex optimization method ([Ne. 1983, 1988]) in the above scheme leads to an improvement of the results. Let us describe the corresponding procedure $\mathcal{M}^*(\rho,\eta,\lambda)$ (the parameters of the procedure are subjected to the same restrictions as in 5.4). he procedure is as follows.

Initialization. Let

β be the positive root of the equation $\beta^2/(1-\beta) = \lambda^2$,

$q = (1 + \mu_2\,[m \ln^2 \eta + \omega^2(\lambda)])^{-1}$,

$N = q^{-2}\,\rho\,\eta$,

$\omega = q^3\,\beta\,(\rho\,\eta)^{-3/2}$,

$x_0 = y$, $t_0 = \tau$.

Compute $d^0 = d(t_0, x_0)$ and the matrices $M^{\phi}_{t_0}(x_0)$ and

$Q_0 = (M^{\phi}_{t_0})^{-1}$.

The k-th step, $k > 0$. Assume that after $k-1$ step of the procedure we have produced a point $x_k \in G'$, a vector $d^k \in \mathcal{V}$, a number $t_k > 0$ and a matrix

$$Q_k = (M(\phi, D_k))^{-1}, \quad D_k = \mathrm{diag}(d^k), \qquad (5.35_k)$$

such that

$$\nu(d^k, d(t_k, x_k)) \le \rho \tag{5.36$_k$}$$

(notice that the initialization rules provide (5.35$_0$), (5.36$_0$)).

At the k-th step we:

a) set

$$t_{k+1} = \begin{cases} t_k\eta \text{ at the main stage} \\ t_k\eta^{-1} \text{ at the preliminary stage} \end{cases}$$

$$p_{k+1}(x) = t_{k+1}\,\phi(x) + F_{x_k, q}(x);$$

b) minimize the function p_{k+1} over $x \in R^n$ by the "optimal" method for smooth convex optimization corresponding to the metric defined by the matrix Q_k^{-1}, i.e.

set $x_{k,0} = x_k$, $A_{k,0} = M$, $v_{k,0} = x_k$;

at the l-th step of the minimization process ($l \ge 0$) we:

1. compute $a_{k,l} > 0$ as a root of the equation

$$a_{k,l}^2 = (1 - a_{k,l})\,A_{k,l}\,M^{-1};$$

2. set

$$y_{k,l} = a_{k,l}\,v_{k,l} + (1 - a_{k,l})\,x_{k,l},$$

$$A_{k,l+1} = a_{k,l}\,q^2 + (1 - a_{k,l})\,A_{k,l},$$

$$x_{k,l+1} = y_{k,l} - M^{-1}\,t_k^{-1}\,Q_k\,p'_{k+1}(y_{k,l}),$$

$$v_{k,l+1} = (1 - a_{k,l})A_{k,l}\,A_{k,l+1}^{-1}\,v_{k,l} + {}$$
$$+ a_{k,l}\,q^2 A_{k,l+1}^{-1}\,y_{k,l} - a_{k,l}\,A_{k,l+1}^{-1}\,t_k^{-1}Q_k p'_{k+1}(y_{k,l}).$$

This process is terminated at the first step, l, where the condition

$$(p'_{k+1}(x_{k,l}))^T\,t_{k+1}^{-1}\,Q_k\,p'_{k+1}(x_{k,l}) \le \omega^2$$

holds;

c) Set $x_{k+1} = x_{k,l}$.

d) compute $d(t_{k+1}, x_{k+1})$ and $d^{k+1} \in D$:

$$d_i^{k+1} = \begin{cases} d_i^k, & (1+\rho)^{-1}\,d_i^k \le d_i(t_{k+1}, x_{k+1}) \le (1+\rho)d_i^k, \\ d_i(t_{k+1}, x_{k+1}), & \text{otherwise,} \end{cases}$$

(notice that this updating provides (5.36$_{k+1}$)) and, using Q_k, compute Q_{k+1} in accordance with (5.35$_{k+1}$).

The k-th step of $\mathcal{M}^*(\rho,\eta,\lambda)$ is over. If $t_{k+1}/t_0 \geq 2$ (at the main stage) or $t_{k+1}/t_0 \leq 1/2$ (at the preliminary stage), then set

$$k^* = k, \quad \tau' = t_{k+1}, \quad y' = x_{k+1}$$

and terminate, otherwise perform the $(k+1)$-th step.

Theorem 5.2. All points x_k, $0 \leq k \leq k^* + 1$, produced by $\mathcal{M}^*(\rho,\eta,\lambda)$, belong to G' and satisfy the relation

$$\lambda(F^{\phi}_{t_k}, x_k) \leq \lambda \qquad (5.37_k)$$

and our procedure solves $(\mathcal{P})$.

Moreover,

$$k^* \leq O((\eta-1)^{-1}). \qquad (5.38)$$

Let η satisfy the condition

$$m^{1/2}\,(\eta - 1) \geq 1. \qquad (5.39)$$

Then the arithmetic cost of $\mathcal{M}^*(\rho,\eta,\lambda)$ does not exceed the quantity

$$M^{(2)} = c(\lambda)\, O_\gamma(q^{-2}\,(\eta-1)^{-1}\, m\, n\, \ln(m\, q^{-1}\, \lambda^{-1}) + m\, n^{1+\gamma} +$$
$$+ (\eta - 1)^{-1+\gamma}\,(\rho-1)^{-\gamma}\, m^\gamma\, n^2), \qquad (5.40)$$

where $c(\lambda)$ depends on λ only.

In particular, under the parameters choice

$$\eta = 1 + m^{-(3-\gamma)/(4-\gamma)}\,(n^*/\ln m)^{1/(4-\gamma)},\ \rho = 1.5,\ \lambda=0.1, \qquad (5.41)$$

where

$$n^* = \max\{n, m^{(2-\gamma)/2}\ln m\} \qquad (5.42)$$

we have

$$M^{(2)} \leq O_\gamma(m^{3/(4-\gamma)}\,(n^*)^{(7-\gamma)/(4-\gamma)}\,(\ln m)^{(1-\gamma)/(4-\gamma)}) \leq$$
$$\leq O_\gamma(m^{(10-\gamma)/(4-\gamma)}\,(\ln m)^{(1-\gamma)/(4-\gamma)}). \qquad (5.43)$$

Hence under the parameters choice (5.41) we have

$$M^*(\varepsilon,\delta) \leq O_\gamma(M^{(2)}\ln(m\,\delta^{-1}\,\varepsilon^{-1})) \leq$$
$$\leq O_\gamma(m^{(10-\gamma)/(4-\gamma)}\,(\ln m)^{(1-\gamma)/(4-\gamma)}\ln(m\,\delta^{-1}\,\varepsilon^{-1})).\ \blacksquare \qquad (5.44)$$

Notice that the optimal step size in t in the case of $n = O(m)$ of traditional linear algebra ($\gamma = 1$) is now

$$t \to (1 \pm O(m^{-1/3}))\, t$$

instead of $t \to (1 \pm O(m^{-3/7}))\, t$ in 5.4.

5.7. "Multistep" barrier method - III.

From now on we assume that (5.1) is a LP problem, i.e the *function* ψ *is linear*. This implies that the function ϕ involved into $(*)$ is linear too.

The procedures described in 5.4, 5.5 are as follows: at the k-th step they transform given approximation, x_k, to $x^*(t_k)$ into an approximation to $x^*(t_{k+1})$ for some prescribed t_{k+1}. Our new procedure, $\mathcal{A}^{**}(\rho,\alpha,\lambda)$, realizes another idea: t_{k+1} is, roughly speaking, as large as possible under the restriction that $x^*(t_{k+1}) \in K_\alpha(x_k)$, thus, we follows the trajectory until it does not leave the region in which $\Phi_t(x)$ is close to $\Phi_{t_k}(x_k)$.

The parameters of our procedure are subjected to the restrictions

$$1 < \rho \leq 1.1; \quad 0 < \alpha; \quad 0 < \lambda \leq 0.1;$$
$$\lambda/8 \geq \alpha > 4\,\mu_2\,\omega^2(\lambda); \quad \rho\,(1+\alpha)^2 \leq 1.25 \qquad (5.45)$$

(it is clear that (5.45) can be satisfied by an appropriate choice of *absolute constants* ρ, α, λ).

For $u \in G'$, $t > 0$ and $d \in \mathcal{D}$ let

$$Q^d = [\mathbf{K}, \phi, \mathrm{diag}(d))]^{-1};$$
$$\theta_{u,t}(x) = \nabla(t\,\phi(x) + F_{u,q(\alpha)}(x)): R^n \to R^n,$$
$$q(\alpha) = (1+\alpha)^{-1};$$
$$\Omega_{u,d,t}(x) = x - (t^{-1}\,Q^d)\,\theta_{u,t}(x): R^n \to R^n.$$

Lemma 5.6. Assume that ρ, α, λ satisfy (5.45). Let $u \in G'$, s, $t > 0$, $d \in \mathcal{D}$ be such that

$$\nu(d,d(t,x)) \leq \rho, \quad |\ln(s/t)| \leq 0.1 \qquad (5.46)$$

and

$$\lambda(F_t^\phi,u) \leq \lambda. \qquad (5.47)$$

Then

(i) for given Q^d, s, u, x the vector $\Omega_{u,d,s}(x)$ can be computed at the arithmetic cost $O(m\,n)$;

(ii) the relation

$$|S^{-1/2}\,\Omega'_{u,d,s}(x)\,S^{1/2}| \leq 0.25 \quad \forall\, x \in R^n, \qquad (5.48)$$

where

$$S = s^{-1}\,Q^d,$$

holds ($|\ |$ is the usual operator norm corresponding to the standard Euclidean structure on R^n);.

(iii) the implication

$$|\ln(s/t)| \le m^{-1/2} (\alpha/2\mu_2)^{1/2},\ 1/2 \le s/t \le 2 \Rightarrow$$
$$\Rightarrow x^*(s) \in K_{\alpha/2}(x),\quad \Omega_{u,d,s}(x^*(s)) = x^*(s) \qquad (5.49)$$

holds. ■

The procedure is as follows.

Initialization. Let

$\eta = \min(\exp(m^{-1/2} (\alpha/2\mu_2)^{1/2}), \exp(0.05))$;

$N =]\ln^{-1}(16) \ln(4 \rho^2 m \alpha^2 \lambda^{-2} (1 - \theta)^{-1})[$,

$\theta = \max((1 + \alpha)^2, \exp(0.1) - 1)$;

$L =]\ln m[$;

$x_0 = y,\ t_0 = \tau$.

Compute $d^0 = d(t_0, x_0)$ and the matrices $M^{\Phi}_{t_0}(x_0)$ and $Q_0 = (M^{\Phi}_{t_0})^{-1}$.

The k-th step, $k \ge 0$. Assume that after k-1 step of the procedure we have produced a point $x_k \in G'$, a vector $d^k \in D$, a number $t_k \in [\tau, \tau']$ and a matrix

$$Q_k = (M(\Phi, D_k))^{-1},\ D_k = \mathrm{diag}(d^k), \qquad (5.50_k)$$

such that

$$\nu(d^k, d(t_k, x_k)) \le \rho \qquad (5.51_k)$$

(notice that the initialization rules provide (5.50_0), (5.51_0)).

At the k-th step we:

a) set

$$\iota = 0,$$
$$\tau^0 = \tau_0 = \begin{cases} t_k\eta \text{ at the main stage} \\ t_k\eta^{-1} \text{ at the preliminary stage,} \end{cases}$$
$$\tau^* = \begin{cases} \tau_0\exp(0.05) \text{ at the main stage} \\ \tau_0\exp(-0.05) \text{ at the preliminary stage.} \end{cases}$$

b) set

$$y_{\iota,0} = x_k;\quad y_{\iota,j} = \Omega_{x_k, d^k, \tau_\iota}(y_{\iota,j-1}),\quad 1 \le j \le N.$$

Set

$$\tau^\iota = \begin{cases} \tau_\iota, \text{ if } y_{\iota,N} \in K_\alpha(x_k) \text{ or } \iota=0, \\ \tau^{\iota-1}, \text{ otherwise,} \end{cases}$$

$$y^{i} = \begin{cases} y_{i,N}, & \text{if } y_{i,N} \in K_{\alpha}(x_k) \text{ or } i=0, \\ y^{i-1}, & \text{otherwise,} \end{cases}$$

$$\tau_{i+1} = (\tau^{i}\,\tau')^{1/2}.$$

If $i < L$, then set $i = i + 1$ and go to b); otherwise set

$$x_{k+1} = y^{N}, \quad t_{k+1} = \tau^{N}$$

and go to c).

c) compute $d(t_{k+1}, x_{k+1})$ and $d^{k+1} \in \mathcal{D}$:

$$d_i^{k+1} = \begin{cases} d_i^{k}, & (1+\rho)^{-1}\, d_i^{k} \leq d_i(t_{k+1}, x_{k+1}) \leq (1+\rho) d_i^{k}, \\ d_i(t_{k+1}, x_{k+1}), & \text{otherwise,} \end{cases}$$

(notice that this updating provides (5.51_{k+1})) and, using Q_k, compute Q_{k+1} in accordance with (5.50_{k+1}).

The k-th step of $\mathcal{M}^{**}(\rho,\alpha,\lambda)$ is over. If $t_{k+1}/t_0 \geq 2$ (at the main stage) or $t_{k+1}/t_0 \leq 1/2$ (at the preliminary stage), then set

$$k^{*} = k, \quad \tau' = t_{k+1}, \quad y' = x_{k+1}$$

and terminate, otherwise perform the $(k+1)$-th step.

Comment: Let $y(x,d,s)$ be the N-th point of a sequence

$$y_0 = x; \quad y_j = \Omega_{x,d,s}(y_{j-1}), \ 1 \leq j \leq N.$$

Then a) - b) describe the usual L-step dichotomy as applied to the problem

(x_k): *being given* x_k, t_k, *find the greatest* $\zeta \in [\ln(\tau_0/t_k), \ln(\tau^{*}/t_k)]$ *such that*

$$y(x_k, d^{k}, s(\zeta)) \in K_{\alpha}(x_k), \quad s(\zeta) = t_k \exp(\zeta). \tag{5.52}$$

Indeed, i is the dichotomy's step number. First of all ($i = 0$) we verify if (5.52) holds for $\zeta = \ln(\tau_0/t_k)$, i.e. for $s(\zeta) = \tau_0$. Let us believe for a moment that the answer is positive (this assumption holds; the latter will be proved later). Then it is not difficult to verify that

$$y(x_k, d^{k}, t_{k+1}) \in K_{\alpha}(x_k). \tag{5.53}$$

Notice that the choice of η and τ^{*} leads to

$$\ln \eta \leq |\ln(\tau^*/t_k)|. \qquad (5.54)$$

It is not difficult to derive from a) - b) and (5.54) that either

$$|\ln(\tau^*/t_{k+1})| \leq 2^{-L},$$

or

$$|\ln(t/t_{k+1})| \leq 2^{-L}$$

for some t such that $y(x_k, d^k, t) \notin K_\alpha(x_k)$.

Notice also that (5.54) implies the relations

$$t_{k+1}/t_k \begin{cases} \geq \eta \text{ at the main stage} \\ \leq 1/\eta \text{ at the preliminary stage.} \end{cases} \qquad (5.55)$$

The following statement is true:

Theorem 5.3. All points x_k, $0 \leq k \leq k^* + 1$, produced by $\mathcal{M}^{**}(\rho,\alpha,\lambda)$, belong to G' and satisfy the relation

$$\lambda(F^{\Phi}_{t_k}, x_k) \leq \lambda \qquad (5.56_k)$$

and our procedure solves $(\mathcal{P})$. Moreover,

$$k^* \leq O(m^{1/2}). \qquad (5.57)$$

The arithmetic cost of $\mathcal{M}^{**}(\rho,\alpha,\lambda)$ does not exceed the quantity

$$c(\rho,\alpha,\lambda,\gamma)\ (m\, n^{1+\gamma} + n\, m^{3/2} \ln^2 m + m^{(1+\gamma)/2}\, n^2) \qquad (5.58)$$

where $c(\rho,\alpha,\lambda,\gamma)$ depends on r, α, λ, γ only. ■

5.8. Concluding remarks.

We have described a number of "multistep" strategies producing good enough approximations to the trajectory $x^*(t)$ of the "conceptual" barrier method. We hope that the "multistep" procedures with "large" steps in t in practice will be much more efficient than the barrier method with Karmarkar's speed-up and "small" steps in t of the type

$$t \to (1 \pm O(m^{-1/2}))\, t,$$

although all these methods posess similar worst-case efficiency's estimates.

5.9. Proofs of the results.

5.9.1. Lemma 5.1.

Let $s = \min(l,k,r)$. Without loss of generality we can assume that l,k,r are divisible by s. Dividing the matrices A and B into square $s\times s$ submatrices, we get $(l/s)\times(k/s)$ and $(k/s)\times(r/s)$ - matrices A', B' with elements from the ring $\varkappa$ of real $s\times s$-matrices. The multiplication of A' and B' in the traditional manner costs $O(lkrs^{-3})$ multiplications and additions of pairs of elements from $\varkappa$; each of these $\varkappa$-operations costs no more than $O_\gamma(s^{2+\gamma})$ arithmetic operations, which implies the statement of the lemma. ∎

5.9.2. Lemma 5.2.

(i). The first and the second statements are obvious; the third follows from the relation

$$M(\phi,D)\,h = \phi''h + (Z(D(Z^T h))).$$

The fourth statement follows from L.5.1, since the computation of the $n \times m$ -matrix DZ^T costs $O(m\ n)$ operations, the multiplication of Z and this matrix costs $O_\gamma(m\ n^{1+\gamma})$ operations and the addition of ϕ'' to the result costs $O(n^2)$ operations.

(ii). If $k = 0$, then the statement is obvious. Now let k be a positive integer. It is clear that

$$M'_{n,n} = M(\phi,D') = M(\phi,D) + V_{n,k}\,S_{k,n} = M_{n,n} + V_{n,k}\,S_{k,n}$$

(subscripts mean the numbers of rows and columns), where $V_{n,k}$ and $S_{k,n}$ can be computed at the cost $O(m + n\ k)$ operations.

Let $k \leq n$. By the well-known formula we have

$$[M'_{n,n}]^{-1} = M^{-1}_{n,n} - [M_{n,n}]^{-1}V_{n,k}(I_k + S_{k,n}\,V_{n,k})^{-1}S_{k,n}[M_{n,n}]^{-1},$$

where I_l means the $l \times l$ - identity matrix. By L.5.1 the matrix $(I_k + S_{k,n}\,V_{n,k})^{-1}$ can be computed at the cost $O_\gamma(n\ k^{1+\gamma})$; the resulting matrix can be inverted at the cost $O_\gamma(k^{2+\gamma})$; each of the subsequent matrix multiplications costs no more than $O_\gamma(n^2\ k^\gamma)$, thus $(M'_{n,n})^{-1}$ can be computed in no more than $O_\gamma(n^2\ k^\gamma)$ operations.

Now let $k > n$. We have

$$(M'_{n,n})^{-1} = (I_n + (M_{n,n})^{-1} V_{n,k} S_{k,n})^{-1} (M_{n,n})^{-1}.$$

The matrix $I_n + (M_{n,n})^{-1} V_{n,k} S_{k,n}$ can be computed at the cost $O(k\, n^{1+\gamma})$ (L.5.1), the resulting matrix can be inverted at the cost $O(n^{2+\gamma})$, and the result can be multiplied by $(M_{n,n})^{-1}$ at the cost $O(n^{2+\gamma})$. Thus, $(M'_{n,n})^{-1}$ can be computed at the cost of $O(k\, n^{1+\gamma})$ operations. ■

5.9.3. Lemma 5.3.

We have

$$\phi'(x^*(t)) - t^{-1} \sum_{i=1}^{m} f'_i f_i^{-1}(x^*(t)) = 0$$

(notice that f'_i does not depend on x). Substracting such an equality for $t = t_2$ from that one for $t = t_1$ and multiplying the resulting equality by $(x^*(t_1) - x^*(t_2))$, we get

$$(x^*(t_1) - x^*(t_2))^T \phi'' (x^*(t_1) - x^*(t_2)) =$$

$$= \sum_{i=1}^{m} (t_1^{-1} [f_i(x^*(t_1)) - f_i(x^*(t_2)] f_i^{-1}(x^*(t_1)) -$$

$$- t_2^{-1} [f_i(x^*(t_1)) - f_i(x^*(t_2))] f_i^{-1}(x^*(t_2))),$$

whence

$$\sum_{i=1}^{m} t_1^{-1/2} t_2^{-1/2} (\xi_i^*(t_1)/\xi_i^*(t_2) + \xi_i^*(t_2)/\xi_i^*(t_1)) +$$

$$+ (x^*(t_1) - x^*(t_2))^T \phi'' (x^*(t_1) - x^*(t_2)) = m (t_1^{-1} + t_2^{-1}),$$

which immediately leads to (5.14). ■

5.9.4. Corollary 5.1.

Let us define m-dimensional vectors and $m \times m$ - matrices as follows (below $t = t_1$ or $t = t_2$):

$$h(t) = ((f_i(x^*(t))/f_i(x(t)) \mid 1 \le i \le m)^T,$$

$$h\_(t) = (h_1^{-1}(t),\ldots,h_m^{-1}(t))^T,$$

$$H(t) = \mathrm{diag}(h(t)), \; e = (1,\ldots 1)^T$$

$$g = ((\xi_i^*(t_2)/\xi_i^*(t_1))^{1/2} \mid 1 \le i \le m)^T,$$

$$g_- = (g_1^{-1},\ldots, g_m^{-1})^T,$$

$$\eta(t) = ((h_1(t))^{1/2},\ldots,(h_m(t))^{1/2}).$$

We have

$$D^2F_t^{\phi}(x(t))[x^*(t)-x(t),x^*(t)-x(t)] = t\ (x^*(t)-x(t))^T\ \phi''\ (x^*(t)-x(t)) + \sum_{i=1}^{m} ((x^*(t)-x(t))^T f_i')^2 f_i^{-2}(x(t)) =$$

$$= t\ (x^*(t)-x(t))^T\ \phi''\ (x^*(t)-x(t)) + \sum_{i=1}^{m} (f_i^{-1}(x(t))\ (f_i(x^*(t)) - f_i(x(t))))^2 = t(x^*(t)-x(t))^T\ \phi''\ (x^*(t)-x(t)) + |h(t)-e|_2^2 \quad (1)$$

and similarly

$$D^2F_t^{\phi}(x^*(t))[x^*(t)-x(t),x^*(t)-x(t)] =$$

$$= t\ (x^*(t)-x(t))^T\ \phi''\ (x^*(t)-\ x(t)) + |h_-(t) - e|_2^2. \quad (2)$$

Thus, T.1.3.(iii) and (5.15) imply

$$t\ (x^*(t)-x(t))^T\ \phi''\ (x^*(t)-x(t)) + |h(t) - e|_2^2 \leq \omega^2(\lambda) \quad (3)$$

and

$$t\ (x^*(t)-x(t))^T\ \phi''\ (x^*(t)-x(t)) + |h_-(t) - e|_2^2 \leq$$
$$\leq \omega^2(\lambda)/(1 - \omega(\lambda))^2. \quad (4)$$

Furthermore, (5.14) implies

$$(t_1\ t_2)^{1/2}\ (x^*(t_1) - x^*(t_2))^T\ \phi''\ (x^*(t_1) - x^*(t_2)) + |g-g_-|_2^2 =$$
$$= m\ (t_1^{1/2} - t_2^{1/2})^2\ (t_1\ t_2)^{-1/2} \equiv \theta^2. \quad (5)$$

Since for positive s we have $(s - s^{-1})^2 \geq (s-1)^2 + (s^{-1}-1)^2$, (5) leads to

$$(t_1\ t_2)^{1/2}(x^*(t_1) - x^*(t_2))^T\ \phi''\ (x^*(t_1) - x^*(t_2)) +$$
$$+ |g - e|_2^2 + |g_- - e|_2^2 \leq \theta^2. \quad (6)$$

By (3), (4) we have

$$t\ (x^*(t)-x(t))^T\ \phi''\ (x^*(t)-x(t)) + |\eta(t) - e|^2 \leq \omega^2(\lambda), \quad (7)$$

$$t\ (x^*(t)-x(t))^T\ \phi''\ (x^*(t)-x(t)) + |\eta_-(t) - e|_2^2 \leq$$
$$\leq \omega^2(\lambda)/(1-\omega(\lambda))^2. \quad (8)$$

We have

$$\zeta \equiv (t_1\ t_2)^{1/2}\ (x(t_1) - x(t_2))^T\ \phi''\ (x(t_1) - x(t_2)) +$$
$$+ \sum_{i=1}^{m} (\delta_i(t_1,x(t_1))-\delta_i(t_2,x(t_2)))^2\ (\delta_i(t_1,x(t_1))\delta_i(t_2,x(t_2)))^{-1} =$$
$$= (t_1\ t_2)^{1/2}\ (x(t_1) - x(t_2))^T\ \phi''\ (x(t_1) - x(t_2)) +$$

$$+ |H^{1/2}(t_1)\, H^{-1/2}(t_2)\, g - H^{-1/2}(t_1)\, H^{1/2}(t_2)\, g_-|_2^2 \le$$
$$\le (t_1\, t_2)^{1/2}\, (x(t_1) - x(t_2))^T\, \phi''\, (x(t_1) - x(t_2)) +$$
$$+2|H^{1/2}(t_1)H^{-1/2}(t_2)g - e_-|_2^2 + 2|H^{1/2}(t_2)^{1/2}H^{-1/2}(t_1)g_- - e|_2^2. \quad (9)$$

Furthermore,

$$|H^{1/2}(t_1)H^{-1/2}(t_2)g - e|_2 \le |H^{1/2}(t_1)H^{-1/2}(t_2)(g - e)|_2 +$$
$$+ |H^{1/2}(t_1)((H^{-1/2}(t_2)\, e - e)|_2 + |H^{1/2}(t_1)\, e - e|_2 \le$$
$$\le |\eta(t_1)|_\infty |\eta_-(t_2)|_\infty |g-e|_2 + |\eta(t_1)|_\infty |\eta_-(t_2) - e|_2 + |\eta(t_1) - e|_2 \le$$
$$\le (1+\omega(\lambda))^{1/2}(1-\omega(\lambda))^{-1/2}\theta + (1+\omega(\lambda))^{1/2}\omega(\lambda)(1-\omega(\lambda))^{-1} + \omega(\lambda)$$

(we have taken into account (7), (8)). The same estimate holds for $|H^{1/2}(t_2)^{1/2}\, H^{-1/2}(t_1)\, g_- - e|_2^2$. Thus, (9) implies

$$\zeta \le (t_1\, t_2)^{1/2}\, (x(t_1) - x(t_2))^T\, \phi''\, (x(t_1) - x(t_2)) +$$
$$+ 4\Big\{(1 + \omega(\lambda))^{1/2}\, (1 - \omega(\lambda))^{-1/2}\, \theta +$$
$$+ ((1 + \omega(\lambda))^{1/2}\, (1 - \omega(\lambda))^{-1} + 1)\, \omega(\lambda)\Big\}^2 \quad (10)$$

We also have

$$(t_1\, t_2)^{1/2}\, (x(t_1) - x(t_2))^T\, \phi''\, (x(t_1) - x(t_2)) \le$$
$$\le 3\, (t_1\, t_2)^{1/2}\, \Big\{(x^*(t_1) - x^*(t_2))^T\, \phi''\, (x^*(t_1) - x^*(t_2)) +$$
$$+ (x^*(t_1) - x(t_1))^T\, \phi''\, (x^*(t_1) - x(t_1)) +$$
$$+ (x^*(t_2) - x(t_2))^T\, \phi''\, (x^*(t_2) - x(t_2))\Big\} \le$$
$$\le 3\, (t_1\, t_2)^{1/2}\, \Big\{(t_1\, t_2)^{-1/2}\, \theta^2 + t_1^{-1}\, \omega^2(\lambda) + t_2^{-1}\, \omega^2(\lambda)\Big\}$$

(the latter - in view of (6), (7)). The resulting inequality together with (10) proves (5.16). ■

5.9.5. Lemma 5.4.

(1). We have $\Phi_r(u) = r\,\phi'' + Z\ \mathrm{Diag}(f_i^{-2}(u))\ Z^T$, whence

$$\min(\, s/t,\, f_1^2(x)/f_1^2(y),\, \dots\, ,\, f_m^2(x)/f_m^2(y)\,)\, \Phi_t(x) \le$$
$$\le \Phi_s(y) \le \max(\, s/t,\, f_1^2(x)/f_1^2(y),\, \dots\, ,\, f_m^2(x)/f_m^2(y)\,)\, \Phi_t(x),$$

or

$$\min(\, s/t,\, f_1^2(x)/f_1^2(y),\, \dots\, ,\, f_m^2(x)/f_m^2(y)\,) \le$$
$$\le \Phi_t^{-1/2}(x)\, \Phi_s(y)\, \Phi_t^{-1/2}(x) \le$$
$$\le \max(\, s/t,\, f_1^2(x)/f_1^2(y),\, \dots\, ,\, f_m^2(x)/f_m^2(y)\,).$$

The latter relation immediately implies (5.20).

(ii). Let $\theta = (t/t')^{1/2}$, and let $v_t = \delta_t(t,x)/\delta_t(t',x')$. Then, by virtue of (5.16), we have

$(1 - v_t)^2/v_t \leqslant \mu_0 (m (1-\theta)^2 \theta^{-1} + \omega^2(\lambda) (1+\theta)^2 \theta^{-1})^{1/2}$;

moreover, $1/2 \leqslant \theta^2 \leqslant 2$. Therefore

$$\max(v_t, 1/v_t) \leqslant O(m \ln^2 \theta + \omega^2(\lambda)).$$

It remains to notice that $f_t(x')/f_t(x) = (t'/t)^{-1/2} v_t$. ■

5.9.6. Theorem 5.1.

1^0. (5.28) immediately follows from a).

2^0. Let us verify (5.27_k). For $k = 0$ this relation holds by virtue of the initialization rule and (5.17). Assume that (5.27_k) holds for some $k \leqslant k^*$ and let us prove that (5.27_{k+1}) also holds. Denote $C_k = f''_{k+1}(x_k)$. Then by virtue of L.5.5 we have

$$q^2 C_k \leqslant f''_{k+1}(x) \leqslant q^{-2} C_k \quad (1)$$

Furthermore, (5.25_k), (5.26_k) and the relation $|\ln(t_k/t_{k+1})| = \ln \eta$ imply

$$\rho^{-1} \eta^{-1} t_k Q_k^{-1} \leqslant C_k \leqslant \rho \eta t_k Q_k^{-1}. \quad (2)$$

Consider R^n as being provided by the scalar product

$$\langle u,v\rangle = u^T t_k Q_k^{-1} v,$$

and let $|\ |$ be the corresponding norm. (1) and (2) mean that f_{k+1} is a strongly convex function with respect to our Euclidean structure with the spectrum of the Hessian belonging to the segment

$$[r,R] = [q^2 \rho^{-1} \eta^{-1}, q^{-2} \rho \eta];$$

in view of h's definition process b) describes the usual gradient descent method as applied to f_{k+1}. Thus, by the standard arguments, we have for all i

$$(f'_{k+1}(x_{k,l}))^T\, t_k^{-1}\, Q_k\, f'_{k+1}(x_{k,l}) \le (1 - r/R)^l\, R^2\, |x_k - x_k^*|^2, \quad (3)$$

where x_k^* is the minimizer of f_{k+1}.

By definition of q and in view of (5.27_k) and L.5.4 we have $x^*(t_{k+1}) \in K_{a(q)}(x_k)$; hence $F^{\phi}_{t_{k+1}}$ coincides with f_{k+1} in a neighbourhood of $x^*(t_{k+1})$, which means that $x_k^* = x^*(t_{k+1})$. By virtue of C.5.1. relation (5.27_k) leads to

$$|x_k - x_k^*|^2 = (x_k - x_k^*)^T\, t_k\, Q_k^{-1}\, (x_k - x_k^*) \le$$
$$\le \rho\, (x_k - x_k^*)^T\, \Phi_{t_k}(x_k)(x_k - x_k^*) = \rho\, t_k\, (x_k - x_k^*)^T \phi''(x_k - x_k^*) +$$
$$+ \sum_{i=1}^{m} (1 - f_i^{-1}(x_k)\, f_i(x_k^*))^2 \le \rho\, (t_k/t_{k+1})^{1/2}\, \mu^2{}_0 \Big\{ m\, (t_k^{1/2} -$$
$$- t_{k+1}^{1/2})^2 + \omega^2(\lambda)\, (t_k^{1/2} + t_{k+1}^{1/2})^2 \Big\}\, (t_k\, t_{k+1})^{-1/2} +$$
$$+ \rho\, m\, (1 - q^{-1})^2 \le O(m\, q^{-2}). \quad (4)$$

Therefore (3) implies

$$(f'_{k+1}(x_{k,l}))^T\, t_k^{-1}\, Q_k\, f'_{k+1}(x_{k,l}) \le (1 - O(q^4))^l\, O(m\, q^{-6}), \quad (5)$$

whence

$$l \le l^* = O(q^{-4})\, (\ln(m/q) + \ln(1/\omega)). \quad (6)$$

Let $\nabla f_{k+1}(x) = t_k^{-1}\, Q_k\, f'_{k+1}(x)$ be the gradient of f_{k+1} with respect to our Euclidean structure. Then

$$|\nabla f_{k+1}(x_{k+1})|^2 \le \omega^2$$

(this is the termination rule in b)). Hence

$$|x_{k+1} - x^*(t_{k+1})| \le \rho\, \eta\, q^{-2}\, \omega. \quad (7)$$

Moreover, we have (see (1), (2))

$$\Phi_{t_{k+1}}(x^*(t_{k+1})) = f''_{k+1}(x^*(t_{k+1})) \le \rho\, \eta\, q^{-2}\, (t_k\, Q_k^{-1}),$$

and (7) implies

$$(x_{k+1} - x^*(t_{k+1}))^T\, \Phi_{t_{k+1}}(x^*(t_{k+1}))\, (x_{k+1} - x^*(t_{k+1})) \le$$
$$\le \rho^3\, \eta^3\, q^{-6}\, \omega^2, \quad (8)$$

which, by the choice of ω and by our standard arguments, leads to (5.27_{k+1}).

Notice that (5.27_{k^*+1}) implies (5.17'); the relation $\tau'/\tau \ge 2$ (at the main stage), $\tau'/\tau \le 1/2$ (at the preliminary stage) immediately follows from the termination rule (see c)). Thus, our procedure solves $(\mathcal{P})$.

3^0. It remains to evaluate the arithmetic cost of the procedure. It is easy to see that the total cost, M', of all computations *excluding the updating of the matrices* Q_k does not exceed $O(K^* l^* m n)$, $K^* = k^* + 1$. Now let us evaluate the total cost, M'', of the matrices updating. First of all, Q_0 can be produced at the cost, $\leq O_\gamma(m n^{1+\gamma})$ (see L.5.2). Now let

$$A_k = \{i \mid d_i^k \neq d_i^{k+1}\}, \quad r_k = |A_k|, \quad r = \sum_{k=0}^{k^*} r_k .$$

Then, by virtue of L.5.2,

$$M'' \leq O_\gamma(m n^{1+\gamma}) + \sum_{k=0}^{k^*} O_\gamma(l(n, r_k)) \leq$$

$$\leq O_\gamma(m n^{1+\gamma} + m K^*) + \sum_{k\in\mathcal{I}} O_\gamma(r_k n^{1+\gamma}) + \sum_{k\in\mathcal{J}} O_\gamma(n^2 r_k^\gamma), \tag{9}$$

where $\quad \mathcal{I} = \{k \mid 0 \leq k \leq k^*, \ r_k > n\},$

$\qquad\quad \mathcal{J} = \{k \mid 0 \leq k \leq k^*, \ r_k \leq n\}.$

Let

$h^0 = 0 \in R^m$, $(h^k)_i = |\ln(d_i(t_{k+1}, x_{k+1})/d_i(t_k, x_k))|$,

$1 \leq i \leq m$, $1 \leq k \leq k^*$.

It is clear from d) that

$$r = \sum_{k=0}^{k^*} r_k \leq O((\rho-1)^{-1}) \sum_{k=0}^{k^*} |h^k|_1 . \tag{10}$$

Furthermore,

$$\sum_{k=0}^{k^*} |h^k|_1 \leq m^{1/2} \sum_{k=0}^{k^*} |h^k|_2 .$$

We have (see (5.19) and C.5.1)

$$|h^k|_2^2 = \sum_{i=1}^{m} 4 \ln^2(\delta_i(t_{k+1}, x_{k+1})/\delta_i(t_k, x_k)) \leq$$

$$\leq 4 \mu_1^2 \sum_{i=1}^{m} (\delta_i(t_{k+1}, x_{k+1}) - \delta_i(t_k, x_k))^2 (\delta_i(t_{k+1}, x_{k+1}) \delta_i(t_k, x_k))^{-1} \leq$$

$$\leq O(m (\eta - 1)^2 + 1); \tag{11}$$

since $k^* \leq O((\eta - 1)^{-1})$, (11) implies

$$r \leq O((\rho-1)^{-1}) \, O(m + m^{1/2} (\eta-1)^{-1}) \leq O(m (\rho-1)^{-1}) \tag{12}$$

(the latter - by (5.29)). Hence (9) implies

$$\sum_{h\in\mathcal{J}} O_\gamma(r_h\, n^{1+\gamma}) \le O((\rho-1)^{-1})\, O_\gamma(m\, n^{1+\gamma}),$$

$$\sum_{h\in\mathcal{J}} O_\gamma(n^2\, r_h^\gamma) \le O_\gamma(|\mathcal{J}|\ (m/|\mathcal{J}|)^\gamma\, n^2) \le$$

$$\le O((\rho-1)^{-\gamma})\, O_\gamma((K^*)^{1-\gamma}\, m^\gamma n^2) \le O_\gamma((\eta-1)^{\gamma-1}\, m^\gamma n^2)\, O((\rho-1)^{-1}).$$

Thus, (9), (10), (11), (6) and (5.28) imply

$$N^{(1)} \le O_\gamma(q^{-4}\ (\eta-1)^{-1}\ m\ n\ \ln(m\ q^{-1}\ \lambda^{-1}) + m\ n^{1+\gamma} +$$
$$+ (\eta - 1)^{-1+\gamma}\ (\rho-1)^{-\gamma}\ \ m^\gamma\ n^2),$$

Q.E.D. ■

5.9.7. Theorem 5.2.

The proof of this theorem is quite similar to the proof of T.5.2; now we use the rate of convergence estimates for the "optimal" smooth convex minimization method (see [Ne. 1982]) instead of the estimates for the gradient descent method. ■

5.9.8. Lemma 5.5.

(i) is obvious.

(ii): by (5.24) we have

$$q^2(\alpha)\ F''(u) \le F''_{u,q(\alpha)}(x) \le q^{-2}(\alpha)\ F''(u),$$

thus for

$$f(x) = s\ \phi(x) + F_{u,q(\alpha)}(x)$$

we have

$$q^2(\alpha)\ f''(u) \le f''(x) \le q^{-2}(\alpha)\ f''(u).$$

Furthermore, (5.46) implies

$$\rho^{-1}\ s\ (Q^d)^{-1} \le f''(u) \le \rho\ s\ (Q^d)^{-1}.$$

Hence

$$\rho^{-1}\ q^2(\alpha)\ S^{-1} \le f''(x) \le \rho\ q^{-2}(\alpha)\ S^{-1},$$

or

$$\rho^{-1}\ q^2(\alpha)\ I_n \le S^{1/2}\ f''(x)\ S^{1/2} \le \rho\ \ q^{-2}(\alpha)\ \ I_n;$$

since

$$I_n - S^{1/2}\ f''(x)\ S^{1/2} = S^{-1/2}\ \ \Omega'_{u,d,s}(x)\ \ S^{1/2},$$

we have

$$|S^{-1/2}\ \Omega'_{u,d,s}(x)\ S^{1/2}| \le \rho\ q^{-2}(\alpha) - 1 = \rho\ (1 + \alpha)^2 - 1 \le 0.25$$

(the latter - by (5.45)). (ii) is proved.

(iii): Let s satisfy the premise in (5.49). Then, by (5.45), we have

$$a > 2\mu_2 (m \ln^2(s/t) + \omega^2(\lambda)), \quad 1/2 \leq s/t \leq 2,$$

whence, by L.5.4, applied with

$$x' = x^*(s), \quad t' = s, \quad x^*(s) \in K_{a/2}(x).$$

The latter inclusion means that $\theta_{u,s}(x^*(s)) = 0$, or that $\Omega_{u,d,s}(x^*(s)) = x^*(s)$. ■

5.9.9. Theorem 5.3.

1^0. Let us prove (5.56_k). (5.56_0) is true by the initialization rules and (5.17). Assume that (5.56_k) holds for some $k \leq k^*$ and prove that (5.56_{k+1}) also holds. Let us fix i, $0 \leq i \leq L$, and let

$$S = \tau_i^{-1} Q_k, \quad y_j = y_{i,j}, \quad v_j = S^{-1/2} y_j.$$

Then for $j \geq 1$

$$v_j = R(v_{j-1}), \quad R(v) = S^{-1/2} \Omega_{x_k, d^k, \tau_i}(S^{1/2} v).$$

We have

$$|R'(v)| = |S^{-1/2} \Omega'_{x_k, d^k, \tau_i}(S^{1/2} v) S^{1/2})| \leq 0.25$$

(L.5.6 as applied to $u = x_k$, $d = d^k$, $t = t_k$, $s = \tau_i$: the conditions of the lemma hold by (5.56_k), (5.51_k) and the relation $|\ln(\tau_i/t_k)| \leq |\ln(\tau_i/\tau_0)| + |\ln(\tau_0/t_k)| \leq |\ln(\tau^*/\tau_0)| + \ln \eta \leq 0.05 + 0.05 = 0.1$).

Thus,

$$|R(v) - R(v')| \leq 0.25 \, |v - v'|. \tag{1}$$

In particular, there exists an unique point, v^*, such that

$$R(v^*) = v^*,$$

and for each v one has

$$|v - v^*| \leq (4/3) \, |v - R(v)|. \tag{2}$$

We also have for $j > 0$: $|v_j - v^*| \leq 4^{-j} \, |v_0 - v^*|$, whence $(5/4) \, |v_0 - v_N| \geq |v_0 - v^*|$, or

$$|v_N - v^*| \leq (5/4) \, 4^{-j} \, |v_0 - v_N|. \tag{3}$$

Of course, we also have

$$|v_N - v^*| \leqslant 4^{-N} |v_0 - v^*|. \qquad (4)$$

2^0. Let us prove that

$$y_N \in K_\alpha(x_k) \Rightarrow \lambda(F^\Phi_{\tau_i}, y_N) \leqslant \lambda \qquad (5)$$

and

$$x^*(\tau_i) \in K_{\alpha/2}(x_k) \Rightarrow y_N \in K_\alpha(x_k). \qquad (6)$$

Assume that the premise in (5) holds. Then we have

$$|v_0 - v_N|^2 = |S^{-1/2}(y_0 - y_N)|^2 = (y_0 - y_N)^T S^{-1} (y_0 - y_N) =$$
$$= (y_0 - y_N)^T \tau_i Q_k^{-1}(y_0 - y_N) \leqslant \rho(y_0 - y_N)^T (F^\Phi_{\tau_i})''(x_k)(y_0 - y_N)$$

(we have taken into account (5.50_k), (5.51_k)). Hence

$$|v_0 - v_N|^2 \leqslant \rho \sum_{i=1}^{m} (f_i(x_k) - f_i(y_N))^2 f_i^{\,2}(x_k) \leqslant \rho\, m\, \alpha^2$$

(the latter – since $y_N \in K_\alpha(x_k)$). Hence (4) implies

$$|v_N - v^*| \leqslant (5/4)\, 4^{-N} \rho^{1/2} m^{1/2} \alpha. \qquad (7)$$

We have $R(v^*) = v^*$, or

$$|R(v_N) - v_N| \leqslant 1.25\, |v_N - v^*| \leqslant 2 \cdot 4^{-N} \rho^{1/2} m^{1/2} \alpha.$$

Since

$$\Omega_{x_k, d^k, \tau_i}(y) = y - S\, (F^\Phi_{\tau_i})'(y) \text{ for } y \in K_\alpha(x_k), \qquad (8)$$

our inequality implies

$$|S^{-1/2}(y_N - S\, (F^\Phi_{\tau_i})'(y_N)) - S^{-1/2} y_N| \leqslant 2 \cdot 4^{-N} \rho^{1/2} m^{1/2} \alpha,$$

or

$$[(F^\Phi_{\tau_i})'(y_N)]^T S\, [(F^\Phi_{\tau_i})'(y_N)] \leqslant 4^{-2N+1} \rho\, m\, \alpha^2. \qquad (9)$$

We also have for $y \in K_\alpha(x_k)$ (see (5.20)):

$$(F^\Phi_{t_k})''(x_k)(1 - \theta) \leqslant (F^\Phi_{\tau_i})''(y) \leqslant (F^\Phi_{t_k})''(x_k)(1 + \theta),$$
$$\theta = \max(\,(1 + \alpha)^2,\ |1 - \tau_i/t_k|\,) < 1,$$

or, since

$$\rho^{-1} S^{-1} \leqslant (F^\Phi_{t_k})''(x_k) \leqslant \rho\, S^{-1},$$

$$\rho^{-1} (1 - \theta)\, S^{-1} \leqslant (F^\Phi_{\tau_i})''(y) \leqslant \rho\, (1 + \theta)\, S^{-1}.$$

So

$$[(F^\Phi_{\tau_i})''(y)]^{-1} \leqslant \rho\, (1 - \theta)^{-1} S,$$

and (9) implies

$$\lambda^2(F^\Phi_{\tau_i}, y_N) = [(F^\Phi_{\tau_i})'(y_N)]^T [(F^\Phi_{\tau_i})''(y_N)]^{-1} [(F^\Phi_{\tau_i})'(y_N)] \leqslant$$

$$\leq 4^{-2N+1} \rho^2 (1-\theta)^{-1} m \alpha^2.$$

Hence, by the choice of N, the conclusion in (5) holds.

Now let us prove (6). Assume that

$$x^* = x^*(\tau_t) \in K_{\alpha/2}(x_h).$$

By (8) we have

$$\Omega_{x_h, d^h, \tau_t}(x^*) = x^*, \quad \text{or} \quad S^{-1/2} x^* = v^*.$$

Thus,

$$|v_0 - v^*|^2 = |S^{-1/2}(x_h - x^*)|^2 = (x_h - x^*)^T S^{-1} (x_h - x^*) \leq$$

$$\leq \rho (x_h - x^*)^T (F^{\Phi}_{t_h})''(x_h) (x_h - x^*) =$$

$$= \rho \sum_{i=1}^{m} (f_i(x_h) - f_i(x^*))^2 f_i^{-2}(x_h) \leq \rho m \alpha^2,$$

which together with (4) leads to

$$|v_N - v^*|^2 \leq 4^{-2N} \rho m \alpha^2, \tag{10}$$

or

$$(y_N - x^*)^T S^{-1} (y_N - x^*) \leq 4^{-2N} \rho m \alpha^2.$$

The latter inequality, as above, leads to

$$(y_N - x^*)^T (F^{\Phi}_{t_h})''(x_h) (y_N - x^*) \leq 4^{-2N} \rho^2 m \alpha^2,$$

or to

$$\sum_{i=1}^{m} (f_i(y_N) - f_i(x^*))^2 f_i^{-2}(x_h) \leq 4^{-2N} \rho^2 m \alpha^2. \tag{11}$$

Hence

$$|f_i(x^*)/f_i(x_h) - f_i(y_N)/f_i(x_h)| \leq 4^{-N} \rho m^{1/2} \alpha,$$

or, by virtue of $x^* \in K_{\alpha/2}(x_h)$,

$$f_i(y_N)/f_i(x_h) \leq f_i(x^*)/f_i(x_h) + 4^{-N} \rho m^{1/2} \alpha \leq$$

$$\leq 1 + \alpha/2 + 4^{-N} \rho m^{1/2} \alpha \tag{12}$$

and

$$f_i(x_h)/f_i(y_N) \leq (1 - \alpha/2 - 4^{-N} \rho m^{1/2} \alpha)^{-1}. \tag{13}$$

(12) and (13) together with (5.45) and the definition of N imply the conclusion of (6).

3^0. Notice that, by the choice of η, we have $x^*(\tau_0) \in K_{\alpha/2}(x_h)$ (see L.5.6.(iii)). So (6) means that $y_0 \in K_\alpha(x_h)$

(this was announced in Comment). By virtue of the Comment we have

$$x_{k+1} = y(x_k, d^k, t_{k+1}) \in K_\alpha(x_k),$$

which, by (5), implies (5.56_{k+1}).

Thus (5.56_k) holds for all k, $0 \leq k \leq k^* + 1$.

4^0. (5.56_{k^*+1}) together with (5.55) and the termonation rule proves that our procedure solves ($\mathcal{P}$).

5^0. It remains to evaluate the arithmetic cost of the procedure. By L.5.6.(iii) the total number of operations *excluding the updating of the matrices* Q_k does not exceed

$$M' = O(m\ n\ N\ L\ K^*),\ K^* = k^* + 1 \leq c_1(\rho,\alpha,\lambda)\ m^{1/2} \quad (14)$$

(the latter - in view of (5.55); from now on $c_i(\rho,\alpha,\lambda)$ depends on ρ, α, λ only).

Let us evaluate the total number, M'', of operations needed by the updating of the matrices. As in the situation of T.5.1, we have

$$M'' \leq c_2(\rho,\alpha,\lambda)\ O_\gamma(m\ n^{1+\gamma} + m\ K^* + (\rho-1)^{-\gamma}(K^*)^{1-\gamma}\ m^\gamma\ n^2 +$$

$$+ (\rho-1)^{-1} m\ n^{1+\gamma}) \leq c_3(\rho,\alpha,\lambda)\ O_\gamma(m\ n^{1+\gamma} + m^{3/2} + m^{(1+\gamma)/2}\ n^2)$$

(the latter - by (14)). The latter inequality together with (14) completes the proof. ■

Section 6. Acceleration of the barrier method, II.

In this Section we describe one more speed-up strategy for the barrier method as applied to the linearly constrained quadratic programming problem (5.1).

Below we preserve the notations from Sect. 5 and all the assumptions on (5.1) from the beginning f subsect. 5.1.

6.1. Description of the accelerated barrier method is as follows. The accelerated method, as well as the barrier method from Sect. 3, is defined by the parameters

$$\lambda_1',\ \lambda_1,\ \lambda_2,\ \lambda_3',\ \lambda_3,$$

satisfying the relations

□ $0 < \lambda_1^+ < \lambda_1' < \lambda_2 < \lambda_3 < \lambda_*$,

$\lambda_1' < \lambda_1 < \lambda_*$, $\lambda_3^+ < \lambda_3' < \lambda_3$; ■ (6.1)

□ $\zeta(\lambda_1') \leq 1/9$, $(1 - \omega(\lambda_2))^{-2} w^2(\lambda_2) < 1/9$,

$(1 - \omega(\lambda_3'))^{-2} w^2(\lambda_3') < 1$, ■ (6.2)

and by a starting point $w \in \text{int}\, G$.

In what follows we regard λ_1', λ_1, λ_2, λ_3', λ_3 as *absolute constants* satisfying (6.1) and (6.2).

Let

$$\left.\begin{aligned} n^* &= \max\{n, m^{(1+\gamma-\gamma^2)/(1+\gamma)}\} \ (\leq m), \\ \rho &= 10\,((n^*)^{(1+\gamma)}\, m^{-(1+\gamma-\gamma^2)})^{2/(2+3\gamma-\gamma^2)}, \\ M &= \rfloor (m^{\gamma/2}\, n^*)^{\gamma/(2+3\gamma-\gamma^2)} \lfloor, \\ K &= \rfloor m^{1/2} M^{-1} \lfloor,\ L = K M. \end{aligned}\right\} \quad (6.3)$$

Notice that $n^* \leq m$.

Denote

$$F(x) = \Sigma_{i=1}^{m} \ln(1/f_i(x)) : G' \to \mathbb{R} .$$

Then $F \in \mathfrak{B}(G,m)$ (T.3.1.(1)).

6.1.1. The accelerated method, as well as the basic one, consists of two stages - the preliminary and the main. Each of these stages corresponds to a set of objects as follows: a convex quadratic form ϕ on $R^n \equiv E$; a number $\varkappa > 0$; an initial value, t_0, of the iterations parameter t; an initial point $x_{-1} \in G'$; numbers λ, $\lambda' \in (0, \lambda_*)$.

These objects for the preliminary stage are as follows:

$\phi(x) = DF(w)(w - x)$;

$t_0 = 1$; $x_{-1} = w$;

$$\varkappa = \exp\left(- \frac{\lambda_1 - \lambda_1'}{\lambda_1(1+\lambda_1^{-1} m^{1/2})} \right); \quad \lambda = \lambda_1, \quad \lambda' = \lambda_1'.$$

For the main stage the above objects are:

$\phi(x) = f_0(x)$;

$x_{-1} = u$ is the result produced at the preliminary stage;
$t_0 = (\lambda_3 - \lambda(F,u))/\|\nabla f_0(u)\|_{u,F}$ (where $\| \ \|_{u,F}$ is the norm in R^n induced by the scalar product $D^2F(u)[\ ,\ ]$ and ∇ is the gradient in the corresponding Euclidean structure);

$$\varkappa = \exp\left\{ \frac{\lambda_3 - \lambda'_3}{\lambda_3(1 + \lambda_3^{-1} m^{1/2})} \right\};\ \lambda = \lambda_3,\ \lambda' = \lambda'_3.$$

6.1.2. A stage corresponds to a family

$$\mathcal{F}_\Phi = (G',\ F_t^\Phi(x) = t\,\phi(x) + F(x),\ R^n)_{t>0};$$

this family is strongly self-concordant with the metrics

$$\rho_\nu(F_\Phi; t, t') = (m^{1/2} + \nu)\,\nu^{-1}\,|\ln(t/t')| \qquad (6.4)$$

(P.3.1). Moreover, it is clear that for $x \in G'$, $t > 0$ one has

$$(F_t^\Phi)''(x) = t\,M_t^\Phi(x). \qquad (6.5)$$

At a stage of the accelerated method approximations x_i to the points

$$x_i^* = \operatorname{argmin}\{\ F_{t_i}^\Phi(x)\ |\ x \in G'\}$$

are produced, where $t_i = t_0\,\varkappa^i$, $i \geq 0$.

6.1.3. Let us start with some definitions. Let d be a m-dimensional vector with positive entries d_l, $l \in \overline{1,m}$. For $l \in \overline{1,m}$ let

$$\Gamma_j(d,l) = \{\ s > 0\ |\ \rho^{2j-1}\,d_l \leq s < \rho^{2j+1}\,d_l\ \},$$

$j \in Z$; the numbers $d_l\,\rho^{2j}$ will be called the *centers* of the *zones* $\Gamma_j(d,l)$. For a positive vector $h \in R^m$ the vector $q_{h,d}$ is defined as vector from R^m with the l-th coordinate being the center of that zone of the family $\{\Gamma_j(d,l)\ |\ j \in Z\}$, which contains the number h_l.

We need some classification of the iterations belonging to a stage. Let us regard the set $\{0,1,\ldots\}$ of values of the iteration number i as being divided into sequential L-element *segments*; each of the segments is regarded as being divided into K sequential M-element *groups* (see (6.3) in the connection with L, M, N).

6.1.4. A stage of the accelerated method is as follows. At its i-th iteration we are given positive m-dimensional vectors h^i, d^i, a $n{\times}n$ -matrix

$$Q_i = [M(\phi,\mathrm{diag}(d^i))]^{-1} \qquad (6.6_i)$$

and a point $x_i \in G'$.

From now on the subcripts at d's and h's mean the coordinate numbers of the corresponding vectors.

These objects are produced by the use of rules as follows (in the below description the numbers in angle brackets mean arithmetic cost of rule's implementation):

I. *Updating of* h^i, d^i, Q_i.

I.1. a) Compute

$$d^{*i} = d(t_i,x_{i-1}) \qquad (6.7)$$

$\langle O(m\ n)$, L.5.2.$\rangle$. If i is not an initial point of a segment, go to I.2.

b) If i is an initial point of a segment, set

$$h^i = d^i = d^{*i} \qquad (6.8)$$

and compute the matrix $M^{\phi}_{t_i}(x_{i-1})$ $\langle O_\gamma(m\ n^{1+\gamma})$, L.5.2$\rangle$. Then compute Q_i in accordance with (6.6_i) $\langle O_\gamma(n^{\gamma+2})$, by definition of $\gamma\rangle$.

At the preliminary stage also compute

$$\lambda(F,x_{i-1}) = t_i^{-1}\ (F'_{t_i}(x_{i-1}))^T\ Q_i\ (F'_{t_i}(x_{i-1}))$$

(the equality - by (6.7), (6.6_i) and (6.5)). If $\lambda(F,x_{i-1}) \le \lambda_2$, then the preliminary stage is terminated with the result $u = x_{i-1}$ $\langle O(m\ n)\rangle$. Otherwise go to II.

I.2. a) Compute

$$d^{+i} = q_{d^{*i},h^{i-1}} \qquad (6.9)$$

$\langle O(m)\rangle$ and, using Q_{i-1}, d^{+i}, d^{i-1}, compute

$$Q_t^+ = [M(\phi, \mathrm{diag}(d^{+t})]^{-1} \qquad (6.10)$$

$\langle O_\gamma(m + l(n, k(t)))$, $k(t) = |\{l \in \overline{1,m} \mid d_l^{+t} \neq d_l^{t-1}\}|$, by virtue of ($6.6_{t-1}$) and L.5.2$\rangle$. If t is not an initial element of a group, set

$$h^t = h^{t-1};\ d^t = d^{+t};\ Q_t = Q_t^+. \qquad (6.11)$$

(hence (6.6_t) holds) and go to II.

b) If t is an initial element of a group, set for $l \in \overline{1,m}$:

- *in the case of* $|\ln(d_l^{*t}/h_l^{t-1})| > 1$:

$$d_l^t = h_l^t = d_l^{*t}; \qquad (6.12)$$

- *in the case of* $|\ln(d_l^{*t}/h_l^{t-1})| \leq 1$:

$$h_l^t = h_l^{t-1};\ d_l^t = d_l^{+t}. \qquad (6.13)$$

$\langle O(m)\rangle$.

Then, using Q_t^+, d^{+t}, d^t, compute Q_t in accordance with (6.6_t) and go to II. $\langle O_\gamma(l(n,p(t)) + m)$, $p(t) = |\{l \mid d_l^t \neq d_l^{+t}\}|$, by virtue of (6.10) and L.5.2$\rangle$.

II. *Updating of* x_t.

II.1. Perform

$$N(\rho) =]\left(\tfrac{1}{2}\rho \ln(2\lambda(1-\lambda)^{-1}(\lambda' - \lambda^+)^{-1})\right)_+[+ 1 \qquad (6.14)$$

steps of the process

$$\square \quad u_0 = 0;\ s_1 = r_0 = -Q_t\, b(t_{t+1}, x_{t-1});$$
$$r_j = r_{j-1} + \alpha_j Q_t H(t_t, x_{t-1}) s_j;$$
$$s_{j+1} = r_j + \beta_j s_j;$$
$$u_j = u_{j-1} + \alpha_j s_j. \quad \blacksquare \qquad (6.15)$$

Herein:

$u_\cdot$, $r_\cdot$, $s_\cdot$ are vectors from R^n;

$$\alpha_j = -\frac{r_{j-1}^T S_t r_{j-1}}{r_{j-1}^T H(t_t, x_t)\, s_j};$$

$$\beta_j = \frac{r_j^T S_t r_j}{r_{j-1}^T S_t r_{j-1}^T};$$

$b(t,x)$ is the gradient of F_t^ϕ at a point x;

$H(t,x)$ is the Hessian of F_t^{ϕ} at a point x;

$S_t = M(\phi, \mathrm{diag}(d_t)) = Q_t^{-1}$.

II.2. Compute

$$x_t = x_{t-1} - u_{N(\rho)}. \tag{6.16}$$

The t-th iteration is over.

Comments to II. Process (6.15) is the conjugate gradient method solving the equation

$$H(t_t, x_{t-1})\, y = b(t_t, x_{t-1}), \tag{6.17}$$

and corresponding to the metric induced by the matrix S_t. It is easy to show that, under the notations (t is fixed)

$$\begin{aligned} &b = b(t_t, x_{t-1});\ H = H(t_t, x_{t-1});\ S = S_t;\ Q = Q_t; \\ &b_* = S^{-1/2}\, b;\ H_* = S^{-1/2}\, H\, S^{-1/2};\ z_j = S^{1/2}\, u_j, \end{aligned} \tag{6.18}$$

the sequence (z_j) is the trajectory of the standard conjugate gradients method minimizing the quadratic form

$$\Phi(z) = z^T H_* z - 2\, b_*^T z, \tag{6.19}$$

under the starting point choice $z_0 = 0$. Notice that we terminate method's implementation after $N(\rho)$ steps.

6.2. **The main result,** which was announced at the beginning of Sect. 5, is as follows:

Theorem 6.1. Assume that linearly constrained quadratic programming problem (5.1) satisfy the conditions from the beginning of subsection 5.1. Then the above accelerated barrier method as applied to this problem with starting point $w \in G'$ is such that:

(i) The amount of segments at the preliminary stage does not exceed the quantity

$$N_1 = O(\ln(2m/(1 - \dot{\pi}_{x(F)}(w))));$$

(ii) For each $\varepsilon \in (0,1)$ the number $N(\varepsilon)$ of the segments at the main stage which is required to produce an approximate solution $x^{\varepsilon} \in \mathrm{int}\, G$ such that

$$\psi(x^{\varepsilon}) - \min_G \psi \leq \varepsilon \ V_F(\psi),$$

satisfies the inequality

$$N(\varepsilon) \leq O(\ln(2m/\varepsilon)).$$

(iii) The arithmetic cost, $\mathfrak{m}$, of each segment of iterations (at the preliminary and at the main stage) satisfies the inequality

$$\mathfrak{m} \leq O_{\gamma}(m^{r_1(\gamma)} (n^*)^{r_2(\gamma)} + m\, n^{1+\gamma}) \leq O_{\gamma}(m^{r(\gamma)}), \qquad (6.20)$$

where

$$r_1(\gamma) = (2 + 5\gamma + \gamma^2)/(4 + 6\gamma - 2\gamma^2),$$
$$r_2(\gamma) = (4 + 5\gamma - \gamma^2)/(2 + 3\gamma - \gamma^2)$$

and

$$r(\gamma) = (10 + 15\gamma - \gamma^2)/(2 + 3\gamma - \gamma^2).$$

In particular, the total arithmetic cost $\mathfrak{m}(\varepsilon)$ of producing x^{ε} satisfies the relation

$$\mathfrak{m}(\varepsilon) \leq O_{\gamma}((m^{r_1(\gamma)} (n^*)^{r_2(\gamma)} + m\, n^{1+\gamma}) \ln(2\, m\, \varepsilon^{-1} \delta^{-1})) \leq$$
$$\leq O_{\gamma}(m^{r(\gamma)} \ln(2\, m\, \varepsilon^{-1} \delta^{-1})),$$
$$\delta = (1 - \pi_{x(F)}(w)). \qquad \blacksquare \qquad (6.21)$$

6.3. Proof of Theorem 6.1.

A. Lemma 6.1. For each iteration number i of the stage under consideration:

$$x_{i-1} \in \text{int}\, G; \qquad (1_i)$$

the matrices S_i and Q_i are symmetric positive definite and

$$t_i\, \rho^{-1}\, S_i \leq H(t_i, x_{i-1}) \leq t_i\, \rho\, S_i; \qquad (2_i)$$

the relations

$$\lambda(F^{\phi}_{t_i}, x_{i-1}) \leq \lambda, \qquad (3_i)$$

$$\lambda(F^{\phi}_{t_i}, x_i) \leq \lambda' \qquad (4_i)$$

hold.

Proof: induction on i.

(1_0) is obvious for the preliminary stage; this relation holds for the main stage by virtue of the fact that such a relation holds for the iterations of the preliminary stage (notice that the statements concerning the preliminary stage are included into the inductive premise when justifying the statements concerning the main stage). (3_0) is obvious for the preliminary stage by definition of the corresponding ϕ; this relation holds for the main stage by virtue of the termination rule used at the preliminary stage (see the proof of P.3.4).

To complete the induction we must prove the implications

$$(1_j),\ j \leq i \ \Rightarrow\ (2_i), \tag{5}$$

$$(2_i)\&(3_i) \ \Rightarrow\ (4_i)\&(1_{i+1}), \tag{6}$$

$$(4_i) \ \Rightarrow\ (3_{i+1}). \tag{7}$$

(7) is true by T.2.1, the definition of $\varkappa$ and (6.4) (see the proofs of P.3.3, P.3.4).

Let us verify (5). We have

$$S_i = K(\phi,\mathrm{diag}(d^i)),\quad H_i = H(t_i,x_{i-1}) = t_i\, M(\phi,\mathrm{diag}(d^{*i})).$$

In the case of (6.9) one has $d^{*i} = d^i$; in the cases of (6.11) or (6.12), (6.13) we have

$$\rho^{-1}\, d^{*i} \leq d^i \leq \rho\, d^{*i}.$$

These inequalities together with the definition of S_i immediately imply (2_i).

Now let us prove (6). Let i be fixed; let us also use the notations (6.18), as well as the abbreviations

$$t = t_i,\quad x = x_{i-1},\quad F_t = F^{\phi}_{t_i},$$

thus $F_t \in S_1^+(G',E)$ (by P.3.1; recall that ϕ is quadratic).

1°. Let $z_* = H_*^{-1}\, b_*$; by the usual properties of the congugate gradients for each j we have $(z_*-z_j)^T H_*(z_*-z_j) = 0$, thus

$$z_j^T\, H_*\, z_j = z_*^T H_*\, z_* - (z_* - z_j)^T\, H_*\, (z_* - z_j). \tag{8}$$

But $z_*^T\, H_*\, z_* = u_*\, H\, u_*$, where u_* is the solution to (6.17), or, that is the same,

$$z_*^T\, H_*\, z_* = \lambda^2(F_t,x) \leq \lambda^2 < 1/9$$

(we have taken into account that

$$\lambda^2(F_t,x) = (F_t'(x))^T [F_t''(x)]^{-1} F_t'(x) = b^T H^{-1} b = u_*^T H u_*,$$

and have used (3_t)). Thus (8) implies, in view of $z_j^T H_* z_j = u_j^T H u_j$:

$$u_j^T H u_j \leq z_*^T H_* z_* \leq \lambda^2 < 1/9, \tag{9}$$

so C.1.2 gives us

$$x - u_j \in G', \ j \geq 0, \tag{10}$$

and, in particular, $x_{t+1} \in G'$, which is required in (3_{t+1}).

2°. Let

$$\varepsilon_j = ((z_* - z_j)^T H_* (z_* - z_j))^{1/2}$$

$(= ((u_* - u_j)^T H (u_* - u_j))^{1/2})$ and let

$$g_j(\tau) = H^{-1/2} F_t'(x - \tau u_j)$$

for $0 \leq \tau \leq 1$. Then

$$g_j'(\tau) + H^{1/2} u_j = (H^{-1/2} (F_t''(x) - F_t''(x - t u_j))H^{-1/2})(H^{1/2} u_j),$$

whence, by T.1.1. and (9), in view of $F_t \in S_1^+(G',E)$, we have

$$|g_j'(\tau) + H^{1/2} u_j|_2 \leq \lambda((1 - \tau\lambda)^{-2} - 1), \ 0 \leq \tau \leq 1,$$

which leads to

$$|H^{-1/2} F_t'(x - u_j) - H^{-1/2} F_t'(x) + H^{1/2} u_j|_2 \leq \lambda^2/(1 - \lambda),$$

or, in view of $F_t'(x) = H u_*$, to

$$|H^{-1/2} F_t'(x-u_j)|_2 \leq \lambda^2/(1-\lambda) + |H^{1/2}(u_*-u_j)| = \lambda^2/(1-\lambda) + \varepsilon_j. \tag{11}$$

By T.1.1 and (9) we have

$$(1 - \lambda)^2 H \leq F_t''(x - u_j) \leq (1 - \lambda)^{-2} H,$$

which together with (11) implies

$$|(F_t''(x - u_j))^{-1/2} F_t'(x - u_j)|_2 \leq \lambda^2/(1-\lambda)^2 + \varepsilon_j/(1-\lambda),$$

or, that is the same,

$$\lambda(F_t, x - u_j) \leq \lambda^2/(1 - \lambda)^2 + \varepsilon_j/(1 - \lambda). \tag{12}$$

3°. Let $\mathfrak{P}_j$ be the space of real polynomials of real variable of degrees $< j$. Then

$$z_j = p_j(H_*) H_* z_*,$$

where

$$p_j \in \mathrm{Argmin}(\ (p(H_*)\ H_*\ z_*)^T\ H_*\ (p(H_*)\ H_*\ z_*) - \\ - 2\ z_*^T\ H_*\ p(H_*)\ H_*\ z_*\ |\ p \in \mathfrak{P}_j\ ),$$

or, that is the same,

$$p_j \in \mathrm{Argmin}\ (\ |H_*^{1/2}\ (\mathrm{Id} - H_* p(H_*))\ z_*|_2^2\ |\ p \in \mathfrak{P}_j\ ),$$

whence

$$\varepsilon_j^2 = (z_* - z_j)^T H_* (z_* - z_j) \leq |H_*^{1/2} (\mathrm{Id} - H_* p(H_*))\ z_*|_2^2\ \forall p \in \mathfrak{P}_j.$$

Thus, for each $p \in \mathfrak{P}_j$ we have

$$\varepsilon_j^2 \leq \max_{\tau \in \Sigma}{}^2\ |1 - \tau p(\tau)|\ |H_*^{1/2}\ z_*|_2^2,$$

where Σ is the spectrum of the matrix H_*; by (9) we have

$$\varepsilon_j \leq \lambda\ \max_{\tau \in \Sigma} |1 - \tau\ p(\tau)| \forall\ p \in \mathfrak{P}_j. \tag{13}$$

Let $q_j \in \mathfrak{P}_j$ be such that

$$1 - \tau\ q_j(\tau) =$$

$$= T_j((\rho - 2\ \tau\ t_t^{-1} + 1/\rho)/(\rho - 1/\rho))\ (T_j(\ (\rho + 1/\rho)\ /\ (\rho - 1/\rho)\ ))^{-1},$$

where $T_j(s) = \mathrm{ch}(j\ \mathrm{arcch}(s))$ is the Chebyshev polynomial of the degree j. (13) for $p = q_j$ gives us

$$\varepsilon_j \leq \lambda\ T_j^{-1}((\rho + 1/\rho)/(\rho - 1/\rho))$$

(we have taken into account (2_t)), which immediately implies the inequality

$$\varepsilon_j \leq 2\ \lambda\ \exp(\ -2j/\rho\ ),\ j \geq 1;$$

the latter together with (12) and the definition of $N(\rho)$ leads to the relation

$$\lambda(F_t, x - u_{N(\rho)}) \leq \lambda'$$

which is required in (4_t). ∎

B. It is not difficult to derive from (6.3) that if $M^* = (m^{\gamma/2}\ n^*)^{\gamma/(2+3\gamma-\gamma^2)}$, $K^* = m^{1/2}(M^*)^{-1}$, $L^* = K^*\ M^* = m^{1/2}$, then

$$M = O(M^*),\ K = O(K^*),\ L = O(L^*). \tag{14}$$

(14) immediately implies that if the numbers t, t' belong to a common segment of iterations of the stage then

$$|\ln(t_t/t_{t'})| \leq O(1), \tag{15}$$

and if these numbers belong to a common group of iterations then

$$|\ln(t_\iota/t_{\iota'})| \leqslant O(\Delta),\ \Delta = 1/K^* \in [O(m^{-1/2}),\ O(1)]. \qquad (16)$$

The results of L.6.1 and relations (15) in the same manner as in the proofs of P.3.3 and P.3.4 prove statements T.6.1.(i) and T.6.1.(ii).

C. It remains to prove T.6.1.(iii). Notice that (6.21) is an immediate corollary of the preceding statements of the theorem, so we must prove (6.20).

1^0. Let $T = \{t_\iota \mid \iota \geqslant 0\}$; let us write for $t = t_\iota$:

$x(t)$ instead of $x_{\iota-1}$,

F_t instead of F^{ϕ}_t.

Let also

$$\psi_l(t) = t^{1/2} f_l(x(t)),\ t \in T,$$

$$\psi_{*,l}(t) = t^{1/2} f_l(x_*(t)),$$

$$x_*(t) = \operatorname{argmin}\{ F_t(x) \mid x \in G'\},\quad t > 0.$$

By C.5.1 we have $(\forall\ t, t' \in T)$:

$$\sum_{l=1}^{m} (\psi_l(t) - \psi_l(t'))^2 (\psi_l(t)\ \psi_l(t'))^{-1} \leqslant$$
$$\leqslant O(m\ (t^{1/2} - (t')^{1/2})^2\ (t\ t')^{-1/2}). \qquad (17)$$

2^0. Let us fix a segment of iterations and let I be the corresponding set of values of the iteration number, ι, and denote by $J_1, \ldots, J_K$ the sets of iteration number values for the groups of the segment I. The remarks on the arithmetic cost rules involved into the method (see method's description) imply that

$$\mathfrak{m} \leqslant O_\gamma(m\ n^{1+\gamma} + m\ n\ N(\rho)\ L) + \sum_{j=1}^{K} \sum_{\iota\in J_j} O_\gamma(l(n,\ k(\iota))) +$$
$$+ \sum_{j=1}^{K} \sum_{\iota\in J_j} O_\gamma(l(n, p(\iota))). \qquad (18)$$

In view of the rules I.1 and I.2 we have

$$k(\iota) = \begin{cases} 0, & \iota \text{ isthe initialelement ofasegment} \\ |U(\iota)|, U(\iota) = \{l \in \overline{1,m} \mid d_l^{+\iota} \neq d_l^{\iota-1}\} & \text{otherwise} \end{cases} \qquad (19)$$

and

$$p(t) = \begin{cases} 0, & t \text{ isnot the initial element of a group or is ainitial element of the group } J_1 \\ |V(t)|, V(t) = \{l \in \overline{1,m} \mid d_l^t \neq d_l^{+t}\}, & \text{otherwise} \end{cases} \quad (20)$$

3^0. First of all let us evaluate the numbers $k(t)$, $t \in I$. Let $t(q)$ be the initial element of the group J_q, $1 \leq q \leq K$. It is clear from the description of the method that

$$h^{t(q)} = h^{t(q)+1} = \ldots = h^{t(q)+M-1}; \quad (21)$$

$$d^{+t} = q_{d^{*t}, h^{t(q)}}, \quad t(q) + 1 \leq t \leq t(q) + M, \ t \in I; \quad (22)$$

$$d^t = q_{d^{*t}, h^{t(q)}}, \quad t(q) \leq t < t(q) + M. \quad (23)$$

Let $I_q = \{t \in I \mid t(q) \leq t \leq t(q) + M\}$. By (6.7) and by definition of $\psi_l(t)$ we have

$$d^{*t} = d(t_t, x_{t-1}) = (\psi_1^{-2}(t_t), \ldots, \psi_m^{-2}(t_t))^T,$$

thus in view of (17), (14) for $1 \leq q \leq K$ and $t \in I_q$ one has

$$\sum_{l=1}^{m} ((d_l^{*t})^{1/2} - (d_l^{*t(q)})^{1/2})^2 (d_l^{*t} \, d_l^{*t(q)})^{-1/2} \leq O_\gamma(m \, \Delta^2). \quad (24_t)$$

Let q be fixed. Let us call a pair $(t, l) \in I_q \times \overline{1,m}$ an *event* if the number d_l^* does not belong to the zone $\Gamma_0(h^{t(q)}, l)$. Let us verify that if $t \in I_q \setminus \{t(q)\} = I_q^0$, and $l \in U(t)$, then either (t, l), or $(t-1, l)$ is an event. ndeed, if for some $t \in I_q^0$ and l neither (t, l) nor $(t-1, l)$ are events then $d_l^{*t-1} \in \Gamma_0(h^{t(q)}, l)$, which, by (23), implies $d_{t-1}(l) = h_{t(q)}(l)$; since $d_l^{*t} \in \Gamma_0(h^{t(q)}, l)$, then by (22) $d_l^{+t} = h_l^{t(q)}$. Thus $d_l^+ = d_l^{*t-1}$, and by (19) $l \in U(t)$, Q.E.D.

The above arguments mean that for $t \in I_q^0$ the quantity $k(t)$ does not exceed the total number of events of the form (t, l), $(t-1, l)$. If (t, l) is an event, then

$$|\ln(d_l^{*t} / h_l^{t(q)})| \geq \ln(\rho),$$

while (6.12), (6.13) imply

$$|\ln(d_l^{*t(q)} / h_l^{t(q)})| \leq 1.$$

Thus

$$|\ln(d_l^{*t} / d_l^{*t(q)})| \geq \ln(\rho/e) \geq 1$$

(we have taken into account that, by (6.3) $\rho \geq 10$, since $\omega_3(\gamma)$

< 1 and hence $n_* \leqslant m$). The latter inequality means that the item in (24_t) corresponding to l under consideration is not smaller than $O(\rho^{1/2})$; thus, the number of events of the form (t,l) does not exceed $O_\gamma(m\,\Delta^2\,\rho^{-1/2})$. The number of events of the form $(t-1,l)$ admits similar upper bound, thus, by the above arguments,

$$k(t) \leqslant O_\gamma(m\,\Delta^2\,\rho^{-1/2}),\quad t \in I^0_q,\quad 1 \leqslant q \leqslant K.$$

Since $\bigcup_{q=1}^{K} I^0_q = I \setminus (t^*)$ (t^* is the initial element of the segment I) and $k(t^*) = 0$ (see (19)), we obtain

$$\sum_{j=1}^{K} \sum_{t \in J_j} O_\gamma(l(n,\, k(t))) \leqslant O_\gamma((n^*)^2\, m^\gamma\, \Delta^{2\gamma}\, \rho^{-\gamma/2}\, L^*) \qquad (25)$$

(notice that $m\,\Delta^2\,\rho^{-1/2} \leqslant n^*$ by virtue of (6.3), so $l(n.k(t)) \leqslant (n^*)^2\, k^\gamma(t)$).

4^0. Now let us evaluate the latter sum in the right hand side of (18). This sum is of the form

$$S = \sum_{j=1}^{K} \sum_{t \in J_j} O_\gamma(l(n,p(t))).$$

Notice that, by (20),

$$S = \sum_{q=2}^{K} O_\gamma(l(n,p(t(q)))) \leqslant O_\gamma(n^2\, P^\gamma\, K^{1-\gamma}) + O_\gamma(n^{1+\gamma}\, P),$$

$$P = \sum_{q=2}^{K} p(t(q)). \qquad (26)$$

For $1 \leqslant q \leqslant K$ let s^q be the m-dimensional vectors with coordinates $\ln(h_l^{t(q)})$, r^q be the vectors with coordinates $\ln(d_l^{*t(q)})$, and for $2 \leqslant q \leqslant K$ let r^{+q} be the vectors with coordinates $\ln(d_l^{+t(q)})$, $1 \leqslant l \leqslant m$. From the description of the method it is clear that the evolution of these vectors is as follows:

$$s^1 = r^1; \qquad (27)$$

$$q > 1 \rightarrow s^q_l = \begin{cases} s^{q-1}_l, & |s^{q-1}_l - r^q_l| \leqslant 1; \\ r^q_l, & \text{otherwise} \end{cases} \qquad (28)$$

Let us verify that for $q \geqslant 2$ also

$$p(t(q)) \leqslant |V^*(q)|.$$

$$V^*(q) = \{l \in \overline{1,m} \mid |s_l^{q-1} - r_l^q| > 1\}. \tag{29}$$

Indeed, if $l \in V^*(q)$ then, in view of $h^{i(q)-1} = h^{i(q-1)}$ ($h^\cdot$ does not vary at the iterations of one group) and by I.2.b), we have

$$|\ln(h_l^{i(q)-1}/d_l^{*i(q)})| \leq 1,$$

thus, by (6.14) $d_l^{i(q)} = d_l^{*i(q)}$; therefore $l \in V(q)$; thus, $V(q) \subset V^*(q)$, which, by virtue of (20), proves (29).

It is clear that

$$|\ln(s/s')|^2 \leq O((s^{1/2} - (s')^{1/2})^2 (s\, s')^{-1/2}),\ s, s' > 0.$$

As in 3^0, we have

$$\sum_{l=1}^{m} ((d_l^{*i(q-1)})^{1/2} - (d_l^{*i(q)})^{1/2})^2 (d_l^{*i(q-1)}\, d_l^{*i(q)})^{-1/2} \leq$$
$$\leq O_\gamma(m\, \Delta^2),$$

which implies

$$\|r^q - r^{q-1}\|_2^2 \leq O_\gamma(m\, \Delta^2),$$

or $\|r^q - r^{q-1}\|_1 \leq O_\gamma(m\, \Delta)$. Since (27) - (29) imply

$$\sum_{q=2}^{K} p(i(q)) \leq \sum_{q=2}^{K} \|r^q - r^{q-1}\|_1,$$

we get

$$P \leq O_\gamma(m\, \Delta\, K) \leq O_\gamma(m). \tag{30}$$

Relations (30), (25), (26), (18) together with (14) and (6.3) prove (6.20). ■

Section 7. Extremal ellipsoids

7.1. Inscribed ellipsoid. Geometric formulation of the problem.

In this section we study a concrete geometric problem as follows. A polytope

$$K = \{ x \in R^n \mid a_i^T x \leq b_i,\ 1 \leq i \leq m\}$$

is given (by the list of the above inequalities); from now on we assume the polytope to be a compact with a nonempty interior ($K \in C_B(R^n)$). We also assume that $a_i \neq 0$, $1 \leqslant i \leqslant m$. The problem (it is denoted by $\mathcal{P}(K)$) is to find among the ellipsoids contained in K the one with maximum possible volume. We refer to the problem as to $\mathcal{P}(K)$.

This problem arises in connection with the IEM - *inscribed ellipsoid method* [TKE, 1988] for convex nondifferentiable optimization. The method minimizes a convex function f, for example, over n-dimensional cube up to relative accuracy ν in $O(n \ln(n/\nu))$ steps (i.e. evaluations of f and f'). Notice that this number of steps can not be reduced (for each $\nu < 1/2$) by more than an absolute multiplicative constant (for precise formulation of the latter remark see [NYu, 1978]). Each step of the IEM requires finding an ε-solution of the above geometrical problem (it is necessary to find an inscribed ellipsoid such that the ratio of its volume to the optimal one be $\geqslant \exp(-\varepsilon)$, where ε is an appropriate absolute constant). In [TKE, 1988] the latter problem is solved by use of the ellipsoid method, which requires about $O(m^6)$ arithmetic operations per step. It turns out that the above barrier method decreases this amount to $O(m^{4.5} \ln m)$. In this section we describe the corresponding implementation of the barrier method.

We study $\mathcal{P}(K)$ under the assumption as follows:

(I) K contains an unit Euclidean ball V centered at 0 and is contained in a concentric ball W with the radius r.

Herein r is a given parameter; notice that in the case of IEM without loss of generality one can take $r = 2n$ (and $m \leqslant O(n \ln n)$).

7.2. Algebraic formulation of the problem.

We can reformulate $\mathcal{P}(K)$ as follows. Let L_n be the space of real $n \times n$ - matrices and L_n^+ be the region in L_n formed by matrices with positive determinant. Each ellipsoid in R^n can be identified by its center $u \in R^n$ and by a matrix $B \in L_n^+$ in the following manner:

$$H(B,u) = \{x = By + u \mid |y|_2 \leqslant 1\}.$$

Notice that, under appropriate choice of the volume unit, the volume $|\ |$ of an ellipsoid $H(B,u)$ is

$$|H(B,u)| = \text{Det } B,$$

and the inclusion $H(B,u) \subset K$ is described by the inequalities system

$$|B^T a_i|_2 \leq b_i - a_i^T u, \ 1 \leq i \leq m;$$

this system will be referred to as $Q(a^m,b^m)$, where a^m denotes the collection of vectors a_i, $1 \leq i \leq m$, and b^m denotes the collection of numbers b_i, $1 \leq i \leq m$.

Let $\mathcal{V}(B) = \ln \text{Det } B : L_n^+ \to \mathbb{R}$. $\mathcal{P}(K)$ can be reformulate as follows:

$\mathcal{P}(K)$: *to find* $z = (B,u) \in L_n^+ \times R^n$ *satisfying* $Q(a^m,b^m)$ *and maximizing under this restriction the objective* $\mathcal{V}(z) \equiv -\ln$ Det B.

Notice that after the above reformulation the relative accuracy $(1 - \exp(-\varepsilon))$ in the volume value corresponds to absolute accuracy ε in the value of the objective involved into $\mathcal{P}(k)$. An ellipsoid $H(B,u)$ will be called ε-*optimal*, if it is contained in K, and its volume is $\geq (1 - \exp(-\varepsilon))\ V^*$, where V^* is the maximum volume of ellipsoids contained in K.

7.3. $\mathcal{P}(K)$ as a Convex Programming Problem.

A representation of a given ellipsoid as $H(B,u)$ is not unique; if U is an orthogonal $n \times n$ -matrix, then

$$H(B,u) = H(\ B\,U,\ u\ ).$$

Hence we can restrict the B-component of the variable $z = (B,u)$ involved into $\mathcal{P}(K)$ to be symmetric positive definite. This restriction leads to a convex programming problem. In fact there is a lot of convex programming problems being equivalent to $\mathcal{P}(K)$; now we describe these problems.

Let S_n be the space of symmetric real $n \times n$ -matrices and S_n^+ be the subset of S_n formed by positive definite matrices (this is an open convex cone in S_n; its closure we denote by S_n^O).

Let for $T \in L_n^+$ $TQ(a^m,b^m)$ denotes the inequalities system with respect to variable $(B,u) \in L_n \times R^N$:

$$|B^T T^T a_i|_2 \leqslant b_i - a_i^T u, \; 1 \leqslant i \leqslant m.$$

Consider the problem

$\mathcal{P}(T,K)$: *to find* $z = (B,u) \in S_n^+ \times R^n$ *satisfying* $T\mathcal{Q}(a^m,b^m)$ *and minimizing under this restriction the objective* $\mathcal{V}(z)$.

Problem $\mathcal{P}(K)$ and $\mathcal{P}(T,K)$, $T \in L_n^+$, obviously are consistent. Let the optimal values of their objectives be v^*, v_T^*, respectively, and let $\Delta(z) = \mathcal{V}(z) - v^*$ for a $\mathcal{P}(K)$-feasible point z, $\Delta_T(z) = \mathcal{V}(z) - v_T^*$ for a $\mathcal{P}(T,K)$-feasible point z.

Lemma 7.1. Let $T \in L_n^+$. If $z = (B,u)$ is a feasible point to $\mathcal{P}(T,K)$, then $Tz \equiv (TB,u)$ is a feasible point to $\mathcal{P}(K)$, and

$$\Delta(Tz) = \Delta_T(z). \tag{7.1}$$

The lemma shows that the solution of $\mathcal{P}(K)$ is equivalent to the solution of each $\mathcal{P}(T,K)$. Each of the latter problems is a convex programming problem.

7.4. Problems $\mathcal{P}(T,K)$ and the basic barrier method.

Let us discuss the basic barrier method application to a problem $\mathcal{P}(T,K)$. Let $E = S_n \times R^n$ and let this space be provided by the standard Euclidean structure with the scalar product

$$((B,u),(C,v)) = \mathrm{Tr}(B^T C) + u^T v.$$

Denote by $G(T)$ the closure of the feasible region of the problem $\mathcal{P}(T,K)$:

$$G(T) = \{z = (B,u) \in E \mid B \in S_n^O, \; |B^T T^T a_i|_2 \leqslant b_i - a_i^T u, \; 1 \leqslant i \leqslant m\}.$$

It is clear that $G(T) \in C_B(E)$ and that the function

$$F^T(z) = 2\,\mathcal{V}(z) - \sum_{i=1}^{m} \ln((b_i - a_i^T u)^2 - |B^T(T^T a_i)|_2^2) \equiv$$

$$\equiv 2\,\mathcal{V}(z) + \Phi^T(z)$$

is a ϑ-self-concordant barrier for $G(T)$, where

$$\vartheta = 2n + 2m < 4m$$

(T.3.2; since K is a compact, we have $n < m$). Notice that $\mathcal{V}(\;)$ is 1-compatible with this barrier (P.3.2). By the condition (I) the point $z = (\frac{1}{2} I_n, 0)$ is a good starting point for our problem, thus $\mathcal{P}(T,K)$ can be solved by the basic barrier method. It turns out that to find an ε-solution to this problem by the latter method it needs no more than

$O(m^{1/2} \ln(rm/\varepsilon))$

iterations of the preliminary and the main stages. Each of these iterations requires to form and to solve certain system of $\dim E \leq O(m^2)$ linear equations with $\dim E$ variables. It is easy to show that the standard implementation of these procedures costs no more than $O(m^6)$ arithmetic operations (even $O(m^5)$ operations, if the conjugate gradient method is applied, because the matrix of the system turns out to be sparse). Thus, the straightforward application of the barrier method to $\mathcal{P}(K)$ produces an ε-solution at the cost of $O(m^{5.5} \ln(rm/\varepsilon))$. The intrinsic symmetry of the problem allows us to improve the cost by the factor $O(m)$.

The idea of the speed-up can be easily described for the main stage, where we need to compute Newton's direction for functions of the form

$$F_t^T(z) = (2+t)\,\gamma(z) + \Phi^T(z).$$

This computation (at a point z of a general form) costs $O(m^5)$ operations (for simplicity sake we replace the powers of n by the same powers of m). But if z has a special form - namely, $z = (I_n, u)$ this computation costs only $O(m^4)$ operations. So let us try to perform the computations in such a manner that the only points in which Newton's direction is computed would be the above special points. It turns out to be possible because of the freedom in problem's formulation choice. Namely, assume we have performed i iterations of the main stage and an approximative solution (B_i, u_i), which is feasible to $\mathcal{P}(K)$, is produced, as well as the penalty parameter value t_i.

Consider the problem $\mathcal{P}(B_i, K)$. Since (B_i, u_i) is feasible to $\mathcal{P}(K)$, the point (I_n, u_i) is feasible to $\mathcal{P}(B_i, K)$. Let us compute Newton's iterate, (B_i', u_{i+1}), of (I_n, u_i) (the Newton method is applied to the function $F_{t_i}^{B_i}$). Now let us increase the penalty parameter value in the same manner as in the basic barrier method. Thus, we produce a new approximative solution $(B_{i+1} = B_i\, B_i', u_{i+1})$ to $\mathcal{P}(K)$ and a new penalty parameter value

t_{i+1}. Now we can perform the next iteration, and so on. Notice that the described procedure needs to be justified - now we have not any convex programming problem the barrier method is applied to. The main idea of the convergence and the rate of convergence proof is as follows. As in the case of the basic barrier method, our aim is to prove that the above method maintains the relation, say,

$$\lambda(F_{t_i}^{B_i},(I_n,u_i)) \le 0.1.$$

Assume that this relation holds for some i. The application of our usual arguments to $\mathcal{P}(K,B_i)$ leads to the relation

$$\lambda(F_{t_i}^{B_i},(B_i',u_{i+1})) \le (0.1)^2/(1-0.1)^2 \le 0.02,$$

and the latter inequality in the case of

$$t_{i+1} = (1 + O(\vartheta^{-1/2}))\, t_i$$

with an appropriate choice of the constant factor in $O(\ \ )$ implies

$$(*):\qquad \lambda(F_{t_{i+1}}^{B_i},(B_i',u_{i+1})) \le 0.05.$$

We wish to derive from the latter relation that

$$(**):\qquad \lambda(F_{t_{i+1}}^{B_{i+1}},(I_n,u_{i+1})) \le 0.1;$$

the difficulty lies in the fact that (*) and (**) involve *different* self-concordant functions. Let us use the following arguments. It is not difficult to show that there is a *nonlinear* one-to-one correspondence between the feasible regions $G(T)$ and $G(T')$ of the problems $\mathcal{P}(T,K)$ and $\mathcal{P}(T',K)$ such that the values of φ (as well as the values of Φ^T and $\Phi^{T'}$) at two corresponding to each other points coincide (up to an additive constant which depends on T, T' only).Hence F_t^T and $F_t^{T'}$ at the points in correspondence differ by a constant (which depends on t, T, T' only). It turns out that the under the above correspondence between $G(B_i)$ and $G(B_{i+1})$ the point (B_i',u_{i+1}) of the first set corresponds to the point (I_n,u_{i+1}) of the second set. The relation (*), by T.1.3.(iii,iv), means that $F_{t_{i+1}}^{B_i}$ at the point (B_i',u_{i+1}) differs from its minimum value over $G(B_i)$ by no more than $(0.06)^2/2$. Hence $F_{t_{i+1}}^{B_{i+1}}$ at

(I_n, u_{t+1}) differs from its minimum over $G(B_{t+1})$ by no more than the same amount. The latter relation, by T.1.3.(iv), implies (**).

The same trick at the preliminary stage needs some special effort. Indeed, at this stage we deal with the families of functions of the type

$$t \text{ (linear form) } + F^T(z).$$

We need this functions to be "almost invariant" under the above correspondence between $G(T)$ and $G(T')$. This condition holds if the *(linear form)* depends on u-component of z only, and we need some effort to provide the latter property of the linear perturbation which is dropped at the preliminary stage of the barrier method. This effort results in a *pre-preliminary stage* we insert in the method. This stage is as follows. All we need (see the description of the basic barrier method) is to find such a point $z^\#$ at which the partial derivative of the barrier in B-component is close to zero; having produced such a point, we can take the restriction of the first order differential of the barrier as the above *(linear form)*, and $z^\#$ - as the starting point for the preliminary stage. To obtain an appropriate $z^\#$, we set $u = 0$ and minimize the barrier in B-component only. This subproblem is relatively simple and is solved at the pre-preliminary stage by the use of the basic barrier method (at the cost of $O(m^{3.5} \ln(m\,r))$ operations).

In accordance with the above discussions we describe *three-stage* version of modified barrier method solving $\mathcal{P}(K)$.

Let us start with the description of the above mentioned correspondence between $G(T)$ and $G(T')$.

Lemma 7.2. Let $T, T' \in L_n^+$. Consider the mapping $Z_{T,T'}$, which transforms a pair $z = (B,u) \in S_n^0 \times R^n$ into the pair $z' = (B',u) \in S_n^0 \times R^n$ such that $T\,B^2\,T^T = (T')\,(B')^2\,(T')^T$ (it is clear that the latter relation do define a positively semidefinite symmetric B' as the function of positively semidefinite symmetric B). Then $Z_{T,T'}$ is a one-to-one mapping from $G(T)$ onto $G(T')$: $z \in G(T) \to z' \in G(T')$, such that $\nu(z) - \nu(z')$ does not depend on z, and $\Phi^T(z) = \Phi^{T'}(z')$. Moreover, $Z_{T',T}$ is the inverse to $Z_{T,T'}$.

The proof is quite straightforward and will be omitted.

7.5. Method's description.

To simplify our considerations, below we choose the parameters of the method as concrete numeric constants (our choice probably is not the best one).

7.5.1. Pre-preliminary stage. At this stage we deal with the problem

$\mathcal{P}_1(K)$: *to minimize* $R(C) = -2 \ln \operatorname{Det} C - 2 \sum_{i=1}^{m} \ln(b_i^2 - \langle C, A_i \rangle)$

under restrictions $C \in S_n^O$, $\langle C, A_i \rangle \leq b_i^2$, $1 \leq i \leq m$,

where $A_i = a_i a_i^T$ and $\langle Q, X \rangle = \operatorname{Tr}(Q^T X)$ is the natural scalar product in S_n.

Let G be the feasible region of $\mathcal{P}_1(X)$; it is clear that $G \in C_B(S_n)$ and R is a $\vartheta' \equiv 2(n + 2m) \leq 6m$ - self-concordant barrier for G. Let $C_0 = 0.5 I_n$. By (I), C_0 is an interior point of G. At the pre-preliminary stage we apply the preliminary stage of the basic barrier method to the barrier R, taking C_0 as the starting point and λ_1', λ_1, λ_2, λ_3', λ_3, such that (see (3.15) - (3.16))

$$0 < \lambda_1^+ \leq \lambda_1' < \lambda_2 < \lambda_3 < 0.01;$$

$$\lambda_1' < \lambda_1 < 0.01, \quad \lambda_3^+ \leq \lambda_3' < \lambda_3$$

$$0.01; \ \zeta(\lambda_1') \leq 0.01, \quad (1 - \omega(\lambda_3'))^{-2} \omega^2(\lambda_3') < 0.1,$$

$$\omega^2(\lambda_2)\,(1 - \omega(\lambda_2))^{-2} \leq 0.01,$$

as the parameters of the basic barrier method.

Let C^* be the result of this application; this point belongs to int G and satisfies the inequality

$$\lambda(R, C^*) \leq \lambda \equiv 0.01. \qquad (7.2)$$

Notice, that the number of iterations in the above procedure does not exceed

$$N_1 = O(m^{1/2} \ln(2m/(1 - \pi_R(C_0))) \qquad (7.3)$$

(П.3.3); herein π_R is the Minkovsky function of G with the pole at the minimizer of R over int G.

Proposition 7.1. The following relations hold:

$$N_1 \leqslant O(m^{1/2} \ln(\tau m)); \tag{7.4}$$

$$R(C^*) - \min_G R \leqslant 0.6\, \lambda^2. \tag{7.5}$$

Moreover, the arithmetic cost of the pre-preliminary stage does not exceed

$$M_1 \leqslant O(m^{3.5} \ln(\tau m) \tag{7.6}$$

operations.

7.5.2. Initialization of the preliminary stage. Having produced the positive-definite symmetric matrix C^*, we compute its factorization $C^* = B_* B_*^T$, where $B_* \in L_n^+$ (at the cost of $O(n^3)$ operations). Consider the problem

$\mathcal{P}_1(B_*, K)$: *to minimize* $P(B) = -2 \ln \operatorname{Det} B - \sum_{i=1}^{m} \ln(b_i^2 - |BB_*^T a_i|_2^2)$

by the choice of $B \in S_n^+$ under the restriction $(B,O) \in G(B_*)$.

It is not difficult to verify (cf. L.2.1) that there exists a one-to-one correspondence between the feasible sets of the problems $\mathcal{P}_1(K)$ and $\mathcal{P}_1(B_*,K)$ with the following property: if C is a feasible point to the first problem and B is the corresponding feasible point to the second problem, then $R(C) = 2\,P(B)$; notice that under our correspondence C^* transforms into I_n.

By the above arguments (7.5) implies the relation

$$P(I_n) - \min_{G_*} P \leqslant 0.3\, \lambda^2$$

($\lambda = 0.01$), where G_* is the feasible region of the problem $\mathcal{P}_1(B_*,K)$. Let $F^*(z) \equiv F^{B_*}(z)$ be the barrier for the feasible set, $G(B_*)$, of the problem $\mathcal{P}(B_*,K)$, and let $z^* = (I_n, O)$. As we have seen,

$$z^* \in \operatorname{int} G(B_*);\ F^*(z^*) - \min\{F^*(B,O) \mid (B,O) \in G(B_*)\} \leqslant 0.3\, \lambda^2. \tag{7.7}$$

Let $\langle\ ,\ \rangle_*$ denote the scalar product in E, defined by the form $D^2F^*(z^*)[\ ,\ ]$, and let $|\ |_*$ be the corresponding norm. By T.1.3.(iv), relation (7.7) implies the relation

$$\forall H \in S_n: \ |DF^*(z^*)[(H,0)]| \leq 0.07 \ |(H,0)|_*; \tag{7.8}$$

moreover, if $z^{**} = (X^{**},0)$ is the minimizer of F^* over the set $G^* = \{(X,u) \in G(B_*) \mid u = 0\}$, then

$$D^2F^*(z^{**})[z^*-z^{**},z^*-z^{**}] \leq 0.01 \tag{7.9}$$

(T.1.3.(iii)).

By (7.8), there exists a linear form $\psi(w) \equiv \langle\psi^*,w\rangle_* = (\psi^{**},w)$ on E ($(\ ,\ )$ is the standard scalar product on $E = S_n \times R^n$) such that $|\psi^*|_* \leq 0.07$ and the restriction of the form onto S_n coincides with the restriction of $DF^*(z^*)[w]$ onto this subspace.

Let us produce this form and consider the linear form

$$\phi(w) \equiv - DF^*(z^*)[w] + \psi(w) = \langle\phi^*,w\rangle = (\phi^{**},w).$$

For $T \in L^+_n$ and $t > 0$ let

$F^T_t(z) = t\phi(z) + 2\,\gamma(z) + \Phi^T(z) \equiv t\phi(z) + \Psi^T(z)$: int $G(T) \to \mathbb{R}$,

so F^T_t is a strongly self-concordant (with the parameter value 1) function defined on int $G(T)$.

Proposition 7.2. The following statements are true:

the linear form $\phi(w)$ depends only on u-component of $w \in E$; (7.10)

$\lambda(F^{B}_1{}^*,z^*) \leq 0.07$; (7.11)

$|\phi^*|_* \leq O(m^{1/2})$; (7.12)

let z^+ be the minimizer of $F^*(\ )$ over int $G(B_*)$, and let $\pi_+(z)$ be the Minkovsky function of $G(B_*)$ with the pole at z^+. Then

$\pi_+(z)) \leq 1 - O((\nu m)^{-2})$; (7.13)

the vector $\phi^{**} \in E$ can be produced at the cost of $O(m^2 n^2 + m^3)$ operations.

7.5.3. Preliminary stage. At this stage we produce matrices $B_i \in L^+_n$, vectors $u_i \in$ int K and numbers $t_i > 0$, $i \geq 0$, as follows:

$B_0 = B_*$, $u_0 = 0$, $t_0 = 1$;

$z_{t+1} = (B_{t+1}, u_{t+1}) = (B_t B^{(t+1)}, u_{t+1}),$

where $z^{(t+1)} = (B^{(t+1)}, u_{t+1})$ is the Newton iterate of $h^{(t)} = (I_n, u_t)$ (the Newton method is applied to the function $F_{t_t}^{B_t}(\ )$).

$t_{t+1} = t_t \exp(-\mu),$

where

$$\mu = \frac{0.05}{1 + \vartheta^{1/2}}, \quad \vartheta = 2\,(n + m). \tag{7.14}$$

The preliminary stage is terminated at the first iteration (its number is denoted by t^*) when the relation

$$\lambda(\Psi^{B_t}, (I_n, u_t)) \leq 0.1 \tag{7.15}$$

holds. The result of the stage is the point

$$z^\# = (B^\#, u^\#) \equiv (B_{t^*}, u_{t^*}).$$

Proposition 7.3. The following statements are true:

the preliminary stage is well-defined:

for all t, $0 \leq t \leq t^*$, we have $B_t \in L_n^+$, $(I_n, u_t) \in \text{int}\ G(b_t)$;

for all t, $0 \leq t \leq t^*$, the relations

$$\lambda(F_{t_t}^{B_t}, (I_n, u_t)) \leq 0.1, \tag{7.16$_t$}$$

$$\lambda(F_{t_t}^{B_t}, z^{(t+1)}) \leq 0.02 \tag{7.17$_t$}$$

hold;

the number t^* of the preliminary stage iterations satisfies the inequality

$$t^* \leq O(m^{1/2} \ln(m\mathbf{r})); \tag{7.18}$$

each of the preliminary stage iterations can be performed (including the verification of the termination condition) at the arithmetic cost of $O(m^2 n^2 + m^3)$ operations.

7.5.4. Main stage. At this stage we produce matrices $C_t \in L_n^+$, vectors $v_t \in$ int K and numbers $t_t > 0$, $t \geq 0$, as follows:

$(C_0, v_0) = (B^\#, u^\#), \quad t_0 = 1;$

$w_{t+1} \equiv (C_{t+1}, v_{t+1}) = (C_t C^{(t+1)}, v_{t+1}),$

where $w^{(t+1)} \equiv (C^{(t+1)}, v_{t+1})$ is the Newton iterate of the point $h^{(t)} \equiv (I_n, v_t)$ (the Newton method is applied to

$$\Xi_{t_t}^{C_t}(w) \equiv t_t\, \gamma(w) + \psi(w) + \Phi^{C_t}(w),$$

where, as above,

$$\Phi^T(w) \equiv - \sum_{t=1}^{m} \ln((b_t - a_t^T u)^2 - |B^T(T^T a_t)|_2^2): \text{int}\ G(T) \to \mathbb{R}.\ ),$$

$t_{t+1} = t_t \exp(\mu)$, where

$$\mu = \frac{0.05}{1 + 2\,\theta^{1/2}}, \quad \theta = n + 2\,m.$$

The properties of the stage are described by the following

Proposition 7.4. The following statements are true:

the main stage is well-defined: for all $t \geq 0$ we have

$$C_t \in L_n^+,\ (I_n, u_t) \in \text{int}\ G(C_t);$$

for all $t \geq 0$ the relations

$$\lambda(\Xi_{t_t}^{C_t}, (I_n, v_t)) \leq 0.1, \qquad (7.19_t)$$

$$\lambda(\Xi_{t_t}^{C_t}, w^{(t+1)}) \leq 0.02 \qquad (7.20_t)$$

hold;

each of the main stage iterations can be performed at the arithmetic cost of $O(m^2 n^2 + m^3)$ operations;

for each t the ellipsoid $H(C_t, v_t)$ is contained in K and

$$\ln |H(C_t, v_t)| \geq \ln |H(C^*, v^*)| - O(m/t_t). \qquad (7.21)$$

7.6. Main result.

The above propositions can be summarized in the following

Theorem 7.1. Assume that the condition (I) (see sect. 7.1) holds. Then the above described method produces an ε-optimal ($\forall\ \varepsilon \in (0,1)$) ellipsoid at the total (over all the stage) number of iterations $O(m^{1/2} \ln(m\tau/\varepsilon))$; the total arithmetic cost of these iterations does not exceed $O(m^{2.5}(n^2 + m) \ln(m\tau/\varepsilon))$ operations.

Notice that in the case of the IEM one can take
$r \leqslant 2n, \quad m \leqslant O(n \ln n)$.

7.7. A minimum volume ellipsoid which contains a given set.

A close to $\mathcal{P}(K)$ problem is to find a minimum volume ellipsoid containing a given finite set. The latter problem can be solved by the same techniques as above. Here we describe the corresponding results. Let Γ be a given m-element set in R^n, and K be the convex hull of Γ; we are required to produce an ellipsoid which contains Γ and has minimum possible volume. This problem will be referred to as $\mathcal{P}_n(\Gamma)$.

We use a traditional trick as follows. Let us regard R^n as an affine hyperplane A in R^{n+1}, defined by the equation $x_{n+1} = 1$; thus, $\Gamma \subset A \subset R^{n+1}$. Consider the problem

$\mathcal{P}^0_{n+1}(\Gamma)$: *to find a (n+1)-dimensional ellipsoid centereg at 0 and containing Γ with minimum possible volume.*

If $W \supset \Gamma$ is feasible to $\mathcal{P}^0_{n+1}(\Gamma)$, then W produces a n-dimensional ellipsoid $W \cap A$, which is feasible to $\mathcal{P}_n(\Gamma)$. It is not difficult to show that the solution of $\mathcal{P}^0_{n+1}(\Gamma)$ produces the solution of $\mathcal{P}_n(\Gamma)$. Moreover, if W is ε-optimal solution to $\mathcal{P}^0_{n+1}(\Gamma)$ (i.e. is feasible to this problem and its volume is $\leqslant \exp(\varepsilon)\, V^{**}$, V^{**} is the optimal objective's value for $\mathcal{P}^0_{n+1}(\Gamma)$), then $W' = W \cap A$ is $\frac{n-1}{n}$ - optimal solution to $\mathcal{P}_n(\Gamma)$. To transform our standard description of W into the standard description of W' it costs no more than $O(n^3)$ operations. So we can restrict ourselves to the problem $\mathcal{P}^0_{n+1}(\Gamma)$.

The algebraic reformulation of $\mathcal{P}^0_{n+1}(\Gamma)$ is as follows:

$\mathcal{P}^*$: *given a subset* $\Gamma = \{x_i \mid 1 \leqslant i \leqslant m\} \subset R^{n+1}$,
to minimize $\varphi(B) = -\ln \operatorname{Det} B$ *by the choice of* $B \in L^+_{n+1}$
under restrictions $|B x_i|_2 \leqslant 1,\ 1 \leqslant i \leqslant m$.

Let Q denote the feasible region of the latter problem. Each $B \in Q$ defines an ellipsoid $H(B^{-1}, 0)$ which is feasible to $\mathcal{P}^0_{n+1}(\Gamma)$; to produce an ε-solution to $\mathcal{P}^0_{n+1}(\Gamma)$ one has to find an ε-solution to $\mathcal{P}^*$, i.e. such $B \in Q$, that

$\mathcal{V}(B) - \inf_Q \mathcal{V} \leq \varepsilon.$

The optimal objective's value in $\mathcal{P}^*$ obviously is the same as in the problem $\mathcal{P}^{**}$, which is obtained from $\mathcal{P}^*$ when the restriction $B \in L^+_{n+1}$ is replaced by the restriction $B \in S^O_n$. The substitution $B^2 = C$ transforms $\mathcal{P}^{**}$ into the problem

$$\mathcal{P}^{***}: \mathcal{V}(C) \equiv - \ln \operatorname{Det} C \to \min \mid C \in S^O_n, \langle C, X_i \rangle \leq 1, 1 \leq i \leq m,$$

where $X_i = x_i x_i^T$.

If C is an ε-solution to $\mathcal{P}^{***}$ and $C = B^T B$ $(B \in L^+_{n+1}$; being given C, we can produce B in $O(n^3)$ operations), then B is an $(\varepsilon/2)$-solution to $\mathcal{P}^*$.

Assume that the following condition holds:

(II) The convex hull, K, of the set Γ contains the unit ball centered at O and is contained in a concentric ball with radius r (both of the balls - in R^n).

It is not difficult to show that under this condition one can insert into $\mathcal{P}^{***}$ (without loss of the optimal objective's value) $n+1$ extra constraints of the form $C_{jj} \leq (c\, r^2 m^4)$, $1 \leq j \leq n+1$, where $c \geq 1/8$ is an appropriate absolute constant. We obtain the problem

$$\mathcal{P}^{\#}: \quad \mathcal{V}(C) \to \min \mid C \in S^O_{n+1}, \langle C, X_i \rangle \leq a_i, 1 \leq i \leq m + n + 1.$$

(we have increased the list of matrices X_i to insert our extra constraints). All we need is to find an ε-solution to $\mathcal{P}^{\#}$.

The feasible set $G^{\#}$ of the latter problem admits an $O(m)$-self- concordant barrier

$$F(C) = \mathcal{V}(C) - \sum_{j=1}^{m+n+1} \ln(a_i - \langle C, X_i \rangle),$$

and the objective is 1-compatible with this barrier. The point $C_O = 0.25\, r^{-2} I_{n+1}$ belongs to int $G^{\#}$, and it is easy to show that

$$\ln(1/(1 - \pi_+(C_O)) \leq O(\ln(rm))$$

(see (II)), where π_+ is the Minkovsky function for $G^{\#}$ with the pole at the F-center of $G^{\#}$. Thus, problem $\mathcal{P}^{\#}$ can be solved by the basic barrier method with C_O being chosen as the starting point; the total number of iterations to produce an ε-solution to this (and hence - to the original) problem does not exceed

$$N(\varepsilon) = O(m^{1/2} \ln(r\, m/\varepsilon)).$$

The arithmetic cost of an iteration, as well as in the situation of P.7.1, does not exceed $O(m^3)$. Thus, we can produce an ε-solution to $\mathcal{P}$ at the total cost of

$$O(m^{3.5} \ln(r\, m/\varepsilon))$$

operations.

7.8. Proofs of the results.

7.8.1. Lemma 7.1.

Tz is obviously feasible for $\mathcal{P}(K)$. To verify (7.1), notice that

$$\Psi(Tz) - \Psi(z) = -\ln \operatorname{Det} T$$

does not depend on z. Thus, each feasible plan for $\mathcal{P}(T,K)$ corresponds to a feasible plan for $\mathcal{P}(K)$ with the same (within a constant term $-\ln \operatorname{Det} T$) objective's value. In particular,

$$v^* - v^*_T \leqslant -\ln \operatorname{Det} T. \tag{1}$$

To prove (7.1) it suffices to show that the latter inequality is an equality. Let (B^*,u^*) is the solution to $\mathcal{P}(K)$. With the help of the polar factorization of the matrix $(T^{-1}B^*)$, we can represent B^* as $B^* = TBU$ with orthogonal U and symmetric positive definite B. Since (B^*,u^*) satisfies the constraints $Q(a^m,b^m)$, the point $z^+ = (B,u^*)$ satisfies the constraints $TQ(a^m,b^m)$, so z^+ is feasible for $\mathcal{P}(T,K)$. Since U is orthogonal, we have:

$$v^* = \Psi(B^*,u^*) = \Psi(TB,u^*) = \Psi(B,u^*) - \ln \operatorname{Det} T =$$

$$= \Psi(z^+) - \ln \operatorname{Det} T \geqslant v^*_T - \ln \operatorname{Det} T,$$

thus $v^* - v^*_T \geqslant -\ln \operatorname{Det} T$. This inequality together with (1) proves the lemma. ■

7.8.2. Proposition 7.1.

To verify (7.4) we must prove, in view of (7.3), that

$$\alpha \equiv 1 - \pi_R(C_0) \geqslant O((rm)^{-s})$$

for certain *absolute constant* s.

Recall that K contains the unit ball centered at O and is contained on a ball with the radius r centered at O; moreover, C is feasible for $\mathcal{F}_1(K)$ C if and only if the ellipsoid $H(C^{1/2},O)$ is contained in K. The above arguments show that the ball (in S_n) with radius $1/4$ centered at C_0 is contained in G and that the diameter of G does not exceed $4\,n\,^{2}$ (the latter - since the semi-axes of the ellipsoid $H(C^{1/2},O)$ for $C \in G$, i.e. the eigenvalues of the matrix $C^{1/2}$, does not exceed r). Hence

$$\alpha \geqslant (1/4)/(4\,n\,r^2) \geqslant O((rm)^{-2}), \quad \text{Q.E.D.}$$

The relation (7.5) immediately follows from (7.2) and T.1.3.(iii).

To prove (7.6) it suffices, in view of (7.5), to verify that a Newton minimization step for a function of the form

$$(a \text{ linear function of } C) + R(C)$$

can be implemented at a cost $O(m^3)$. It is easy to see that the gradient, $2\,H$, of such a function at a given point $C \in \text{int } G$ can be computed at the above cost. A straightforward computation shows that the Hessian, $2\,\Psi$, of the function at C transforms $X \in S_n$ into the matrix

$$2\,\Psi\,X = 2\,C^{-1}\,X\,C^{-1} + \sum_{t=1}^{m} 2\,d_t\,\langle A_t,X\rangle\,A_t,$$

where the set of numbers

$$2\,d_t = (b_t^2 - \langle A_t,C\rangle)^{-2}, \quad 1 \leqslant t \leqslant m,$$

can be computad at a cost $O(m\,n^2)$ (recall that $A_t = a_t\ a_t^T$). The Newton displacement X is the solution to the system $\Psi\,X = H$; hence, it can be represented as

$$X = C\,(H + \sum_{t=1}^{m} x_t\,A_t)\,C \qquad (1)$$

where x_t, $1 \leqslant t \leqslant m$ are some scalars. Let us derive the system of linear equations for these scalars (the solution of the latter system, after substitution into (1) gives the desired X). To derive the system, let us substitute (1) into the equation $\mathcal{A} X = H$; after some simple transformations we get a matrix equation

$$\sum_{i=1}^{m} x_i A_i + \sum_{i=1}^{m} d_i \langle A_i, CHC\rangle A_i + \sum_{i=1}^{m} d_i \sum_{j=1}^{m} x_j \langle CA_jC, A_i\rangle A_i = 0. \quad (2)$$

This equation is eqvivalent to the system (*) of m scalar linear equations with m variables x_i produced when taking termwise scalar product of (2) and each of the matrices A_j, $1 \leq j \leq m$. The ij-th coefficient of the matrix of system (*) is

$$\langle A_j, A_i\rangle + \sum_{k=1}^{m} d_k \langle A_k, A_i\rangle\langle A_k, CA_jC\rangle,$$

and the i-th component of the right hand side vector is

$$- \sum_{k=1}^{m} d_k \langle A_k, A_i\rangle\langle A_k, CHC\rangle.$$

To produce the matrix of our system and the right hand side vector, it suffices to compute:

- all of the products $\langle A_i, A_j\rangle$ ($O(m^2 n)$ operations);

- m matrices CA_jC ($O(n^2 m)$ operations) and all of the scalar products of these matrices and matrices A_i ($O(m^2 n)$ operations more);

- the matrix $C H C$ ($O(n^3)$ operations) and its scalar products onto matrices A_k ($O(n^2 m)$ operations more)

(when evaluating the numoer of operations one must take into account that rank $A_i = 1$).

After the above quantities are produced, each of the coefficients of system (*) can be computed at the cost $O(m)$. Thus, system (*) can be formed at the cost $O(m^3)$; it can be then solved at the same cost. After the system is solved, the Newton displacement X can be computed at the cost $O(m n^2)$, see (1). ∎

7.8.3. Proposition 7.2.

(7.12) is obvious by definition of $\phi(z)$ (since the restriction of ψ onto S_n coincides with the restriction of the form $DF^*(z^*)[\ ]$, thus the restriction of ϕ onto S_n $\equiv 0$). Furthermore,

$$DF_1^{B*}(z^*)[w] = DF^*(z^*)[w] + \phi(w) = \psi(w),$$

thus (7.11) follows from the relation $|\psi^*|_* \leq 0.07$ (see the definition of ϕ). Moreover,

$$|\phi^*|_* \leq |(F^*)'(z^*)|_* + |\psi^*|_* \leq 0.07 + O(m^{1/2}),$$

since F^* is a self-concordant barrier with the parameter value $O(m)$; (7.12) is proved.

To verify (7.13), notice that the pair

$$z = (B,u) \in S_n^O \times R^n$$

is a feasible plan for $\mathcal{P}(B_*,K)$ if and only if the ellipsoid $H(B_*B,u)$ is contained in K. Let us introduce a norm

$$p(B,u) = |B_*B| + |u|_2$$

($|\ |$ is the usual operator norm) on E, and let $B_O = \frac{1}{2} C^{-1/2}$. Since $B_*B_*^T = C$, $B_*C^{-1/2}$ is an orthogonal matrix, so the ellipsoid $H(B_*B_O,O)$ is an Euclidean ball with radius $1/2$ centered at O. By (I) the $1/4$-neighbourhood of the point z_O = (B_O,O) (in the metric corresponding to the norm p) is contained in $G(B_*)$; at the same time (I) means that the diameter of $G(B_*)$ in the above metric does not exceed $O(m + r)$. Hence

$$\pi_+(z_O) \leq 1 - O((m + r)^{-1}). \qquad (1)$$

Furthermore, the restriction of F^* onto $G^* \equiv \{(B,O) \in G(B_*)\}$ is an $O(m)$-s.c. barrier for G^*, and the center of G^* with respect to this barrier is $z^{**} = (X^{**},O)$. Hence (P.3.2.(v)) G^* contains the ellipsoid

$$(\text{in } S_n)\ U = \{(Y,O) \mid D^2F^*(z^{**})[Y-X^{**},Y-X^{**}] \leq 1\}$$

and is contained in the ellipsoid

$$U' = \{(Y,O) \mid D^2F^*(z^{**})[Y-X^{**},Y-X^{**}] \leq O(m^2)\}.$$

This together with (7.9) implies that z^* can be represented as

$$z^* = \alpha z_O + (1 - \alpha) z$$

for certain $z \in G^*$ and $\alpha \in (0,1)$, $\alpha \geq O(1/m)$. By virtue of the convexity of π_+ and (1) we have

$$\pi_+(z^*) \leqslant \alpha\, \pi_+(z_0) + (1 - \alpha)\, \pi_+(z) \leqslant 1 - \alpha\, O(m^{-1}) \leqslant$$
$$\leqslant 1 - O(m^{-2}\, r^{-2}),$$

which is required in (7.13).

It remains to evaluate the cost at which ϕ^{**} can be computed. Let Q be the Hessian of F^* at the point z^*, q be the gradient of F^* at this point and Π be the orthoprojector of E onto S_n. Let $x = (X,0)$ be the solution of the system

$$\Pi(q - Q\,x) = 0;\ x \in S_n \tag{2}$$

(S_n is identified with the subspace $S_n \times \{0\}$ of E). It is not difficult to show that $\phi^{**} = Q\,x - q$. Indeed, for $w \in S_n$ we have

$$\langle x,w\rangle_* = (Q\,x,w) = (\Pi Qx,w) = (q,x);$$

if $w \in E$ is $\langle\ ,\ \rangle_*$ - orthogonal to S_n, then $\langle x,w\rangle_* = 0$ in view of $x \in S_n$. Hence x is $\langle\ ,\ \rangle_*$ - orthogonsl projection of the gradient of F^* at the point z^* (*the gradient is taken with respect to the Euclidean structure* $\langle\ ,\ \rangle_*$) onto S_n, or, which is the same, $x = \psi^*$. So $\psi^{**} = Q\,x$ and $\phi^{**} = Q\,x - q$.

Let us write the expressions for the first and second order differentials of F^T at the point (I_n,u):

$$DF^T(I_n,u)[(H,v)] = -\,2\,\langle I_n,H\rangle - \sum_{i=1}^{m} d_i\,(c_i a_i^T v - \langle H, T^T A_i T\rangle), \tag{3}$$

$$D^2F^T(I_n,u)[(H,v),(H,v)] =$$
$$= 2\,\langle H,H\rangle + \sum_{i=1}^{m} r_i\,(c_i a_i^T v - \langle H,T^T A_i T\rangle)^2 + \sum_{i=1}^{m} s_i\,\langle T^T A_i T\, H,H\rangle, \tag{4}$$

where $A_i = a_i\, a_i^T$, and the set of scalars d_i, c_i, r_i, s_i (which depend on u and T only) can be computed for given u, T at the cost of $O(m\, n^2)$ operations (when speaking about the costs of computations, we take into account that the matrices A_i are of rank 1). Notice that $s_i \geqslant 0$.

In particular, we see that the computation of q, as well as the multiplication of Q by a given vector, can be implemented at the cost $O(m\, n^2)$.

By the above arguments to prove that ϕ^{**} can be computed at the cost $O(m^2\, n^2 + m^3)$ it suffices to verify that one can priduce at this cost a symmetric solution, X, to the matrix equation

$$X + J X + X J + \sum_{i=1}^{m} (\alpha_i + \beta_i \langle T^T A_i T, X\rangle) T^T A_i T = H. \qquad (5)$$

In this equation $J = \frac{1}{2} \sum_{i=1}^{m} s_i T^T A_i T$ is a symmetric positive semidefinite matrix which can be computed at the cost $O(m n^2)$; at the same cost one can compute the symmetric matrix H and the set of scalars α_i, β_i.

1^0. To solve (5) we act as follows. Let us reduce the matrix J by an orthogonal transformation U to a three-diagonal form, i.e. let us find (at the cost $O(n^3)$) an orthogonal matrix U and a three-diagonal matrix matrix P such that $U J U^T = P$. The substitution $Y = U X U^T$ transforms (5) into the equation

$$Y + P Y + Y P + \sum_{i=1}^{m} (\alpha_i + \beta_i \langle S^T A_i S, Y\rangle) S^T A_i S = L, \qquad (6)$$

where $S = T U^T$, $L = U H U^T$ are matrices which can be computed at the cost $O(n^3)$. We desire to find a symmetric solution to (6); this solution at the cost $O(n^3)$ can be transformed into the desired solution to (5). Thus, we must verify tha' (6) can be solved at the cost $O(m^2 n^2 + m^3)$.

2^0. Let us find the solutions to $(m + 1)$ matrix equations

$$Y_i + P Y_i + Y_i P = S^T A_i S, \ 1 \leqslant i \leqslant m,$$

$$Y_{m+1} + P Y_{m+1} + Y_{m+1} P = L.$$

Since P is a three-diagonal matrix, each of these equations can be solved at the cost of $O(n^3)$. Indeed, the equation (with respect to a $n \times n$ - matrix Z)

$$\Phi(Z) \equiv Z + P Z + Z P = M \qquad (7)$$

has an unique solution, since the operator Φ (regarded as a linear operator in L_n) is symmetric and positive definite (since P is symmetric positive semidefin.te). The subspace of symmetric matrices is invariant for Φ, so if M is symmetric, then the solution, Z, to (7) is symmetric too. In particula , Y_i are symmetric, $1 \leqslant i \leqslant m + 1$.

3^0. Let us verify that (7) can be solved at the cost $O(n^3)$.

(7) regarded a system with n^2 variables (the entri s of

Z) cad be described as follows. The matrix P is symmetric, positive semidefinite and three-diagonal:

$$P e_i = \gamma_i e_i + \mu_i e_{i-1} + \mu_{i+1} e_{i+1},$$

where e_i, $1 \leq i \leq n$, are the standard orts in R^n, $e_0 = e_{n+1} = 0$, $\mu_1 = \mu_{n+1} = 0$, $\gamma_i \geq 0$. Let l_i be the columns of L and let $z_i = Z e_i$, $0 \leq i \leq n+1$. Then (7) can be rewritten as a system (u) of equations

$$(u_i): \quad \mu_i z_i + (\gamma_{i-1} + 1) z_{i-1} + P z_{i-1} + \mu_{i-1} z_{i-2} = l_{i-1}, \quad 2 \leq i \leq n+1,$$

with unknown vectors z_i, which are subjected to restrictions $z_0 = z_{n+1} = 0$. To solve the system, let us act as follows. The indexes of non-zero elements of the sequence $\mu^* = (\mu_1, \ldots, \mu_{n+1})$ can be divided into mutually disjoint sequential groups $I_r = \overline{s_r, t_r}$, $1 \leq r \leq k$, such that

$$\mu_{s_r - 1} = \mu_{t_r + 1} = 0$$

(notice th..t $\mu_0 = \mu_{n+1} = 0$) Let

$$I_r^- = I_r \cup (s_r - 1), \quad I_r^+ = I_r \cup (t_r + 1).$$

Now let $i_1, \ldots, i_f$ be the elements of $\overline{1,n}$, which does not belong to $\bigcup_r I_r^-$; $I_{r+j}^- = (i_j)$ and $I_{r+j}^+ = (i_j + 1)$ for $1 \leq j \leq f$. So we have defined the groups I_j^-, I_j^+, $1 \leq j \leq k + f$. Let $(u(r))$ denotes the subsystem of system (u), which consists of all the equations (u_i) of (u) with indexes i belonging to I_r^+; it is easy to verify that subsystem $(u(r))$ involves z_i with $i \in I_r^-$ only (so these subsystems have no common unknowns), and system (u) is a "direct product" of subsystems $(u(r))$, $r = 1, \ldots, k + f$. It suffices to prove that subsystem $(u(r))$ can be solved at the cost $O(n^2 a(r))$, where $a(r)$ is the number of elements in I_r^-.

To avoid cumbersome notations, assume that $(u(r))$ consists of the equations (u_t), $t = 2,\dots,p$; thus,

$$\mu_p = 0, \quad \mu_2,\dots,\mu_{p-1} \neq 0.$$

We desire to solve our subsystem at the cost of $O(n^2 p)$ operations. Let us act as follows. Let Z_0 be the zero and Z_1 be the identity $n \times n$ matrices and let for $2 \leqslant t < p$ the matrix Z_t be difined by the relation

$$Z_t = -\mu_t^{-1}(((\gamma_{t-1} + 1)I_n + P)\, Z_{t-1} + \mu_{t-1}\, Z_{t-2}).$$

It is clear that the general solution to he homogeneous system of equations

$$\mu_t\, z_t + (\gamma_{t-1} + 1)\, z_{t-1} + P\, z_{t-1} + \mu_{t-1}\, z_{t-2} = 0,$$

$$2 \leqslant t < p,$$

(where $z_0 = 0$) is of the form $z_t = Z_t\, \lambda$, $1 \leqslant t < p$, where $\lambda \in R^n$. A patricular solution $(z_t^*,\ 1 \leqslant t < p)$ to the system

$$\mu_t\, z_t + (\gamma_{t-1} + 1)\, z_{t-1} + P\, z_{t-1} + \mu_{t-1}\, z_{t-2} = l_{t-1},$$

$$2 \leqslant t < p,$$

can be find out recursively by formula

$$z_t^* = \mu_t^{-1}(l_{t-1} - (\gamma_{t-1} + 1)\, z_{t-1}^* - P\, z_{t-1}^* - \mu_{t-1}\, z_{t-2}^*),$$

$$2 \leqslant t < p,$$

where $z_0^* = z_1^* = 0$.

Since the matrix P is three-diagonal, the computation of all matrices Z_t and vectors z_t^*, $1 \leqslant t < p$, can be performed at the total cost $O(n^2\, p)$. Notice that the matrices Z_t are $O(p)$ - diagonal. It is clear that the solution to the subsystem under consideration is

$$z_t = Z_t\, \lambda^* + z_t^*, \quad 1 \leqslant t < p, \tag{8}$$

where λ^* is such that the equation (u_p) is satisfied by z_t given by (8). In other words, λ^* is the solution to the linear system

$$((\gamma_{p-1} + 1) + P)\,(Z_{p-1}\,\lambda + z^*_{p-1}) + \mu_{p-1}(Z_{p-2}\,\lambda + z^*_{p-2}) = l_{p-1}.$$

This sistem at the cost $O(n^2)$ can be reduced to the standard form, and the matrix of this system is $O(p)$-diagonal; so the cost at which the system can be solved by the conjugate gradient method does not exceed $O(n^2\,p)$. After λ^* is computed, it needs no more that $O(n\,p^2)$ operations to regenerate z_i in accordance with (8). Thus, the subsystem under consideration can be solved in $O(n^2\,p)$ operations, as was announced at the beginning of 3^0.

4^0. Let us return to equation (6). By 3^0 the total cost at which Y_i, $1 \leqslant i \leqslant m + 1$, can be computed does not exceed $O(m\,n^3)$. It is clear that the solution to (6) can be represented as

$$Y = \sum_{i=1}^{m+1} t_i\,Y_i \tag{9}$$

with appropriate scalars t_i. The substitution of this representation into (6) gives, by definition of Y_i, the equation for t_i of the form

$$\sum_{i=1}^{m} t_i\,S^T A_i S + t_{m+1}\,L + \sum_{i=1}^{m} (\alpha_i + \beta_i \sum_{j=1}^{m+1} t_j\,\langle S^T A_i S, Y_j\rangle)\,S^T A_i S = L. \tag{10}$$

Matrix equality (10) is equivalent to the system (*) of $m + 1$ scalar linear equations with unknowns t_i; these equations can be obtained by taking termwise scalar product of (10) and matrices $S^T A_i S$, $1 \leqslant i \leqslant m$, and L. Let us compute all quantities of the form

$$\langle S^T A_i S, Y_j\rangle, \quad \langle S^T A_i S, S^T A_j S\rangle, \quad \langle S^T A_i S, L\rangle, \quad \langle L, L\rangle\ .$$

Since rank $A_i = 1$, the total cost of this computation is $\leqslant O(m^2 n^2)$. After these quantities are computed it needs no more than $O(m)$ operations to compute each of the coefficients of (*). So (*) can be reduced to the standard form at the total cost $O(m^2\ n^2)$. Solving (*) ($O(m^3)$ operations) and then regenerating Y in accordance with (9) ($O(m\ n^2)$ operations more) we find out the desired solution to (6). The proof is over. ■

7.8.4. **Proposition 7.3.**

1^0. Assume that for some t the relations

$\mathcal{P}(t)$: *for each* j, $0 \leq j \leq t$, *one has* $B_j \in L_n^+$, $(I_n, u_j) \in \mathrm{int}\, G(B_j)$ *and* (7.16_j) *hold;* (7.17_j) *hold for* $0 \leq j < t$

hold.

Notice that $\mathcal{P}(0)$ is obviously true (see (7.11)). Let us verify that if $\mathcal{P}(t)$ holds, then $\mathcal{P}(t+1)$ holds. Indeed, the function $F_{t_t} = F_{t_t}^{B_t}$ is strongly self-concordant on $\mathrm{int}\ G(B_t)$, so by (7.16_t) and by T.1.3.(ii) we have $z^{(t+1)} \in G(B_t)$ and

$$\lambda(F_{t_t}, z^{(t+1)}) < (0.1)^2/(1-0.1)^2 < 0.013 \qquad (1)$$

(thus, (7.17_t) holds).

P.3.1 as applied to the strongly self-concordant family

$$\mathcal{F} \equiv (\mathrm{int}\ G(B_t),\ F_t(z) = t\ \phi(z) + \Phi^{B_t}(z),\ S_n \times R^n)_{t>0},$$

together with the fact that ϕ is 0-compatible with the corresponding barrier Φ^{B_t} (the parameter value for this barrier is $\vartheta = 2\ (n + m)$), implies:

$$\rho_{0.07}(\mathcal{F}; t_t, t_{t+1}) \leq (1 + \vartheta^{1/2} (0.07)^{-1})\ \mu \leq$$

$$\leq (0.07)^{-1}\ (1 + \vartheta^{1/2})\ \mu = 0.05\ (0.07)^{-1},$$

whence, by T.2.1 and by (1), one has

$$\lambda(F_{t_{t+1}}, z^{(t+1)}) \leq 0.05.$$

By T.1.3.(iii) the latter relation leads to

$$F_{t_{t+1}}(z^{(t+1)}) - \min_{\mathrm{int}G(B_t)} F_{t_{t+1}}(z) \leq 0.6\ (0.05)^2. \qquad (2)$$

By L.7.2 the left hand side in (2) is equal to

$$F_{t_{i+1}}^{B_{i+1}}(I_n, u_{i+1}) - \min \mathrm{int}_{G(B_{i+1})} F_{t_{i+1}}^{B_{i+1}},$$

which together with (2) and T.1.3.(iv) proves (7.16$_{i+1}$). Thus, the implication $\mathcal{P}(i) \Rightarrow \mathcal{P}(i+1)$ is proved.

2^0. Now let us prove (7.18). Let

$$\Psi(z) = \Psi^{B_*}(z) : \mathrm{int}\ G(B_*) \to \mathbb{R}, \quad F_t(z) = t\ \phi(z) + \Psi(z).$$

Let us put into correspondence to the points (I_n, u_i) the points $w_i \in G(B_*)$, using the transformation $G(B_*) \to G(B_i)$ described in L.7.2. In view of (2) and L.7.2 we have:

$$F_{t_i}(w_i) - \min_{\mathrm{int}G(B_*)} F_{t_i} \leqslant 0.6\ (0.05)^2, \tag{3}$$

whence by T.3.1.(iv)

$$\lambda(F_{t_i}, w_i) \leqslant 0.07. \tag{4}$$

Let w^* be the Ψ-center of $G(B_*)$ and let

$$W_{1/2} = \{w \in E = S_n \times R^n \mid D^2\Psi(w^*)[w - w^*, w - w^*] \leqslant 1/4\};$$

then $W_{1/2} \subset \mathrm{int}\ G(B_*)$ (C.1.2). Moreover,

$$D^2\Psi(w)[h,h] \geqslant 0.25\ D^2\Psi(w^*)[h,h]$$

for $w \in W_{1/2}$ (T.1.1), which implies

$$\Psi(w) - \Psi(w^*) \geqslant 1/32 \quad \text{for} \quad w \in \partial W_{1/2}. \tag{5}$$

Let us regard E as being provided by the scalar product $D^2\Psi(w^*)[\ ,\ ]$, and let ∇ and $\|\ \|$ denote the corresponding gradient and norm. In view of (7.12), (7.13) and P.3.2.(iv.2) we have

$$\|\nabla\phi\| \leqslant O((m\tau)^{2.5}),$$

which together with (5) implies

$$F_t(w) - F_t(w^*) \geqslant 1/32 - O(t\ (m\tau)^{2.5}).$$

Hence, under an appropriate choice of *absolute constants* as factors in $O(\ )$ which follow we have, by virtue of (3),

$$t_i \leqslant O((m\tau)^{-2.5}) \quad \Rightarrow \quad w_i \in W_{1/2}. \tag{6}$$

If the premise in (6) holds for given t, then

$$\lambda(\Psi,w_t) \leq \lambda(F_{t_t},w_t) + 2\ t_t\ |\nabla\Phi|$$

(since the norm induced by the form $D^2\Psi(w_t)[\ ,\ ]$ coinside, within a factor 2, with that one defined by the form $D^2\Psi(w^*)[\ ,\ ]$). In view of (4) we have

$$t_t \leq O((m\varkappa)^{-2.5}) \Rightarrow \lambda(\Psi,w_t) \leq 0.07 + O(t_t\ (m\varkappa)^{2.5}). \quad (7)$$

In particular, the implication

$$t_t \leq O((m\varkappa)^{-2.5}) \Rightarrow \lambda(\Psi,w_t) \leq 0.08$$

holds, whence, by T.1.3.(1)),

$$t_t \leq O((m\varkappa)^{-2.5}) \Rightarrow \Psi(w_t) - \min_{\mathrm{int}G(B_*)} \Psi \leq 0.6\ (0.08)^2,$$

and, by L.7.2

$$t_t \leq O((m\varkappa)^{-2.5}) \Rightarrow \Psi^{B_t}(I_n,u_t) - \min_{\mathrm{int}G(B_*)} \Psi^{B_t} \leq 0.6\ (0.08)^2,$$

so (T.1.3.(iv))

$$t_t \leq O((m\varkappa)^{-2.5}) \Rightarrow \lambda(\Psi^{B_t},(I_n,u_t)) \leq 0.1. \quad (8)$$

This implication together with the termination rule for the preliminary stage (see (7.15)) and the rules for t_t's updating leads to (7.18).

3^0. It remains to verify that an iteration of the preliminary stage can be performed in no more than $O(m^2\ n^2)$ operations. It is clear that the iteration's cost is

$$O(n^3) + \varkappa_1 + \varkappa_2,$$

where $\varkappa_1$ is the computation cost of $\lambda(\Psi^T,(I_n,u))$ for some given u and T (this computation is required in the termination rule) and $\varkappa^2$ is the computation cost of the Newton displacement for a function F_t^T at the point F_t^T. It is clear by (7.18), (7.19) that the gradients of Ψ^T and F_t^T at the point (I_n,u) can be computed at the cost $O(m\ n^2)$. After this gradients are produced, both of the above computations can be reduced to the solution of the equation (with unknowns $(X,v) \in E$))

$$D^2\Psi^T(I_n,u)[(X,v),(H,h)] = (s,(H,h)) \quad \forall\ (H,h) \in E \qquad (9)$$

for given $s \in E$. Thus, the cost of an iteration is

$$O(\ m\ n^2 + \varkappa),$$

where $\varkappa$ is the cost at which (9) can be solved. It suffices to prove that $\varkappa \leq O(m^2\ n^2)$.

In view of (7.19) relation (9) can be rewritten as a system of one matrix and one vector equations

$$X + \sum_{i=1}^{m} r_i\ (c_i a_i^T v - \langle X, T^T A_i T\rangle)\ T^T A_i T + $$
$$+ \sum_{i=1}^{m} s_i\ (T^T A_i T\ X + X\ T^T A_i T) \ = H, \qquad (10)$$

$$\sum_{i=1}^{m} d_i\ (c_i a_i^T v - \langle X, T^T A_i T\rangle)\ a_i = h, \qquad (11)$$

where $A_i = a_i\ a_i^T$ and the set of scalars $s_i > 0$, c_i, r_i, d_i, the vector $h \in R^n$ and the symmetric matrix H (these objects depend on eds u and T only) for given u, T can be computed at the cost $O(m\ n^2)$.

Let us act as in the proof of P.7.2.: first compute the symmetric matrix

$$W = \sum_{i=1}^{m} s_i\ T^T A_i T$$

($O(m\ \ n^2)$ operations), then reduce W by an orthogonal transformation to the three-diagonal form ($O(n^3)$ operations) and rewrite (10), (11) as

$$Y + R\ Y + Y\ R + \sum_{i=1}^{m} r_i\ (c_i a_i^T v - \langle Y, S^T A_i S\rangle)\ S^T A_i S \ = G, \qquad (12)$$

$$\sum_{i=1}^{m} d_i\ (c_i a_i^T v - \langle Y, S^T A_i S\rangle)\ a_i = h, \qquad (13)$$

where R, S and G are some known matrices (they can be computed at the cost $O(n^3)$; R is symmetric positive semidefinite and 3-diagonal). Notice that the solution (Y,v) to system (12), (13) at the cost $O(n^3)$ can be transformed to the solution (X,v) to (10), (11).

To solve (12), (13), we as in the proof of P.7.2, at the total cost $O(m^2 n^2 + m^3)$) produce the solutions Y_i, $1 \le i \le m + 1$, to the matrix equations

$$Y_i + R\,Y_i + Y_i\,R = S^T A_i S,\ 1 \le i \le m,$$
$$Y_{m+1} + R\,Y_{m+1} + Y_{m+1}\,R = G,$$

and represent the Y-component of the desired solution as

$$Y = \sum_{i=1}^{m+1} \tau_i\,Y_i$$

where scalars τ_i are our new unknowns. Substituting this representation into (12) and taking termwise scalar product of the resulting equation and each of the matrices $S^T A_i S$, $1 \le i \le m$, G, we obtain a system of scalar linear equations (which is equivalent to (12), (13)) of the form

$$\begin{cases} A\,\tau + B\,v = p & (*) \\ C\,\tau + D\,v = q, & (**) \end{cases}$$

where $\tau = (\tau_1, \dots, \tau_{m+1})^T$ and v are the unknowns, and the matrices A, B, C, D are of sizes $(m+1) \times (m+1)$, $(m+1) \times n$, $n \times (m+1)$, $n \times n$ ((*) corresponds to (12), (**) - to (13)). By the same arguments as in the proof of P.7.2 all the objects $A - D$, p, q can be computed at the cost $O(m^2\ n^2)$. Thus, to produce (*) - (**), to solve this system and to transform its solution into the solution to (12) - (13) it costs no more than $O(m^2 n^2 + m^3)$ operations. ■

7.8.5. Proposition 7.4.

The families

$$(\text{int}\ G(C),\ \Xi_t^C, E)_{t>0},\quad C \in L_n^+,$$

are strongly self-concordant families generated by $(n + 2m)$-self-concordant barriers $\gamma(w) + \Phi^C(w)$ for the sets $G(C)$ and by 1-compatible with these barriers functions $\gamma(w)$. By the termination rule for the preliminary stage (see (7.15)) we have

$$\lambda(\Phi^{B^{\#}}, (I_n, u^{\#})) \leq 0.01.$$

Obviously,

$$\Phi^{B^{\#}} = \Xi_1^{C_0},$$

thus, in view of $t_0 = 1$, (7.19_0) holds. By the same arguments as in the proof of P.7.3 this fact implies the well-definitness of the main stage iterations and validity of relations(7.19_t), (7.20_t) for each t. The cost of an iteration can be evaluated in the same manner as in the proof of P.7.3. It remains to verify (7.21). This inequality, by L.7.2, is equivalent to the inequality

$$\gamma(I_n, v_t) - \min_{\text{int } G(C_t)} \gamma \leq O(\theta/t_t);$$

the latter fact follows from (7.19_t) by virtue of arguments similar to these used in the proof of P.3.4. ■

REFERENCES

[CW. 1986]:
Coppersmith D., Winograd S. Matrix multiplication via arithmetic progression. Dept. of Math. Sci., IBM Thomas J. Watson Res. Centr., Preprint, Nov. 1986

[FMcC. 1968]:
Fiacco A.V., McCormick G.P. Nonlinear Programming, Wiley, N.Y., 1968

[Fr. 1988 1,2]:
1. Freund R.M. Projective transformations for interior point methods, Part I: Basic theory and Linear Programming. - M.I.T. Operations Research Center working paper OR 179-88, 1988
2. Freund R.M. Projective transformations for interior point methods, Part II: an algorithm for finding the weighted center of a polyhedral system. - M.I.T. Operations Research Center working paper OR 180-88, 1988

[GL. 1988]:
Goldfarb D., S. Liu An $O(n^3L)$ primal interior point algorithm for convex quadratic programming. Technical Report, Dept. of IEOR, Columbia University, New York, NY 10027

[Go. 1987]:
Gonzaga C.C. An algorithm for solving linear programming problems in $O(n^3L)$ operations. Tech. Report, Dept. of Electrical Engineering and Computer Sciences, Berkeley, California, 94720

[Ja. 1987]:
Jarre F. On the convergence of the method of analytic centers when applied to convex quadratic programs. - Report No. 35, Dec. 1987, Schwerpunktprogramm der Deutchen Forschungsgemeinschaft - Anwendungsbezogene Optimierung und Steuerung, revised Feb.1988.

[Ka. 1984]:
Karmarkar N. A new polynomial-time algorithm for linear programming. Combinatorica 1984, v. 4, 373 - 395

[KMY. 1987]:
Kojima M.S., S. Mizuno, A. Yoshise A primal-dual interior point algorithm for linear programming. Report No. B-193, Dept. of Information Sciences, Tokyo Institute of Technology, Tokyo, Japan

[MA. 1987 1,2]:
1. Monteiro R.C., I. Adler An $O(n^3)$ primal-dual interior point algorithm for linear programming. Report ORC 87-4, Operations Research Center, Dept. of Operations Research, University of California, Berkeley, CA
2. Monteiro R.C., I. Adler An $O(n^3)$ primal-dual interior point algorithm for convex quadratic programming. Report ORC 87-15, Operations Research Center, Dept. of Operations Research, University of California, Berkeley, CA

[MS. 1987, 1988]:
Mehrotra S., Sun J. (1987) An algorithm for convex quadratic programming that requires $O(n^{3.5} L)$ arithmetic operations. Technical report 87-24, Dept. of IE/MS, Northwestern University, Evanston, ILL. 60208
Mehrotra S., Sun J. (1988) A method of analytic centers for quadratically constrained convex quadratic programs. Technical report 88-01, Dept. of IE/MS, Northwestern University, Evanston, ILL. 60208

[Ne. 1983]:
Nesterov Ju.E. An $O(1/k^2)$-rate of convergence method for a convex programming problem. (in Russian) - *Doklady Akademii Nauk SSSR*, 1983, v.269, No. 3, 543-547

[Ne. 1988]:
Nesterov Ju.E. A general approach to design of optimal methods for smooth convex functions minimization. (in Russian) - *Ekonomika i matem. metody*, 1988, v.24, No. 3,

[Ne. 1988 1,2,3,4,5]:
1.Nesterov Ju. E. The method for linear programming which requires $O(n^3 L)$ operations. (in Russian) - *Ekonomika i matem. metody*, v. 24 (1988), No. 1, 174-176; translated as *Matekon*
2.Nesterov Ju. E. Polynomial-time methods in linear and quadratic programming. (in Russian) - *Izvestija AN SSSR, Technitcheskaya kibernetika*, 1988, No.3
3. Nesterov Ju. E. Polynomial-time iterative methods in linear and quadratic programming. (in Russian) - *Voprosy kibernetiki*, Moscow, 1988
4. Nesterov Ju. E. Dual polynomial-time algorithms for linear programming. (in Russian) - *Kibernetika*, 1989, No. 1
5. Nesterov Ju. E. New polynomial-time algorithms for linear and quadratic programming. - Report at the 13-th International Symposium on Mathematical Programming, 29 Aug. - 2 Sept. 1988, Tokyo, Japan

[NN. 1988 1,2]:
1. Nesterov Ju. E., Nemirovsky A.S. Polynomial-time barriers methods in convex programming. (in Russian) - *Ekonomika i matem. metody*, v. 24 (1988), No. 6; translated as *Matekon*
2. Nesterov Ju. E., Nemirovsky A.S. A general approach to polynomial-time algorithms design for convex programming. - Report at the 13-th International Symposium on Mathematical Programming, 29 Aug. - 2 Sept. 1988, Tokyo, Japan

[NYu. 1983]:
Nemirovsky A.S., Yudin D.B. Problem complexity and method efficiency in optimization. - J.Wiley&Sons, 1983

[Re. 1988]:
Renegar J. A polynomial-time algorithm, based on Newton's method, for linear programming. - Math. Progr. v. 40 (1988), 59 - 93

[Sh. 1987]:
Shor N.Z. Quadratic optimization problems. (in Russian) - *Izvestija AN SSSR, Technitcheskaya kibernetika*, 1987, No.1 128-139

[So. 1985]:
Sonnevend G. An analytic centre for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. - Proc. 12-th IFIP Conf. on System Modelling and Optimization, Lecture Notes in Control and Information Sciences, vol. 84

[TKE. 1988]
Tarasov S.V., Khachiyan L.G., Erlich A.I. The inscribed ellipsoid method. (in Russian) - *Doklady Akademii Nauk SSSR*, 1988, v. 298 No. 5, 1081-1085

[TY. 1987]:
Todd M.J., Y. Ye A centered projective algorithm for linear programming. Techn. Report No. 763, School of Operations Research and Industrial Engineering, College of Engineering, Cornell University, Ithaca, NY 14853

[Va. 1987]:
Vaidya P. An algorithm for linear programming which requires $O(((m+n)\,n^2 + (m+n)^{1.5}\,n)\,L)$ arithmetic operations. - AT&T Bell Laboratories, Murray Hill, NJ (1987)

[Ye 1987]:
Ye Y. Interior algorithms for linear, quadratic and linearly constrained convex programming. Ph. D. dissertation, Dept. of Engineering - Economic Systems, Stanford University, Stanford, CA

CONTENTS

Заказ № 30 А-09220 27/I-89г. Тираж I60экз.
Объем I0,0уч.изд.л. Цена Iруб.

ЦЭМИ АН СССР.