Mathematical Programming 77 (1997) 191–224

The long-step method of analytic centers for fractional problems

A. Nemirovski [1]

Faculty of Industrial Engineering and Management, Technion - The Israel Institute of Technology, 3200 Technion City, Haifa, Israel

Received 1 December 1994

Abstract

We develop a long-step surface-following version of the method of analytic centers for the fractional-linear problem $\min\{t_0 \mid t_0B(x) - A(x) \in H,\ B(x) \in K,\ x \in G\}$, where H is a closed convex domain, K is a convex cone contained in the recessive cone of H, G is a convex domain and $B(\cdot)$, $A(\cdot)$ are affine mappings. Tracing a two-dimensional surface of analytic centers rather than the usual path of centers allows to skip the initial "centering" phase of the path-following scheme. The proposed long-step policy of tracing the surface fits the best known overall polynomial-time complexity bounds for the method and, at the same time, seems to be more attractive computationally than the short-step policy, which was previously the only one giving good complexity bounds. © 1997 The Mathematical Programming Society, Inc. Published by Elsevier Science B.V.

Keywords: Fractional programming; Generalized eigenvalue problem; Interior point methods; Analytic centers

1. Introduction

In this paper we develop a long-step path-following method for the *linear-fractional* optimization problem

$$(\mathrm{P}) \qquad \min\{t_0 \mid t_0B(x) - A(x) \in H,\ B(x) \in K,\ x \in G\}, \qquad (1)$$

where

- $B(x) = \beta x + b$ and $A(x) = \alpha x + a$ are affine mappings from $\mathbb{R}^n$ to $\mathbb{R}^m$;

[1] E-mail: nemirovs@ie.technion.ac.il. The research was partly supported by the Israeli–American Binational Science Foundation (BSF).

PII S0025-5610(96)00083-4

- $H \subset \mathbb{R}^m$ is a closed convex domain which does not contain lines and K is a closed convex cone with a nonempty interior in $\mathbb{R}^m$ which is contained in the recessive cone of H:

$$H + K = H;$$

- G is a closed convex domain in $\mathbb{R}^n$.

Problem (1) covers a lot of applications, e.g., as follows:

Example 1 (*Convex problems*). Let $m = 1$, $K = H = \mathbb{R}_+$, $B(x) \equiv 1$, $A(x) = a^{\mathrm{T}}x$; under these assumptions (1) becomes a problem of minimizing a linear objective over a closed convex domain G, which is a universal, in the natural sense, form of a convex program.

Example 2 (*Simple linear-fractional problem*). Let, as above, $m = 1$, $K = H = \mathbb{R}_+$ and let $B(x)$ be a linear form which is positive on G; now (1) becomes the problem of minimizing the linear-fractional objective $A(x)/B(x)$ over G. This is the simplest problem of quasiconvex programming.

Example 3 (*von Neumann problem of economic growth*). Let B and A be $m \times n$ matrices with nonnegative entries, let $K = H = \mathbb{R}^m_+$ and let G be the standard simplex $\{x \geqslant 0 \mid \sum_i x_i = 1\}$ in $\mathbb{R}^{n+1}$ (G is regarded as a subset of its affine hull). Then (1) with $B(x) = Bx$, $A(x) = Ax$ is the well-known von Neumann problem of finding the largest rate of economic growth: *find the largest α such that for some nonzero nonnegative x one has $Bx \geqslant \alpha Ax$.*

Examples 2 and 3 are related to the case when $K = H$ is the nonnegative orthant in $\mathbb{R}^m$; a general problem (1) associated with this cone is as follows:

Example 4 (*Minimize the maximum of m linear-fractional functions over a given closed convex domain where all denominators are nonnegative*). This is a universal form of the *generalized concave fractional problem*

$$\max_{x \in S} \min_{j \leqslant m} \frac{g_j(x)}{h_j(x)}$$

($S \subset \mathbb{R}^n$ is convex, $g_j(\cdot)$ are concave and nonnegative on S, and $h_j(\cdot)$ are convex and positive on S). For applications of the latter problem in economics, see [3,7,19,20,22] and the references therein. The standard methods for solving the problem are Dinkelbach's algorithm [6] and its variants, see [20].

Now, nonpolyhedral cones K also lead to interesting problems, especially the cone of positive semidefinite matrices. If $K = H$ is the cone of positive semidefinite symmetric matrices of a given order, then (1) becomes:

Example 5 (*Generalized eigenvalue problem*). Given two symmetric matrices $B(x)$ and $A(x)$ of the same size with entries affinely depending on a vector x of design variables, minimize over $x \in G$, under additional restriction that $B(x)$ is positive semidefinite, the largest generalized eigenvalue of the pencil (B, A), i.e., the smallest $\lambda = \lambda(x)$ such that $A(x) \leqslant \lambda B(x)$ (the inequalities between symmetric matrices are always understood in the operator sense, i.e., as positive semidefiniteness of the corresponding difference). The problem of minimizing the largest generalized eigenvalue of a matrix pencil possesses a lot of applications in modern Control Theory (see [5]).

The development of polynomial-time interior-point methods for Linear and Convex Programming, started by the landmark paper of Karmarkar [12], initiated activity in Fractional Programming as well. To the best of our knowledge, the very first paper on an interior-point polynomial-time algorithm for fractional problems was the one of Anstreicher [1] (Example 2, G is a polytope). A Karmarkar-like algorithm for general fractional problems (including those in Examples 4 and 5) was recently developed by Nesterov and Nemirovski [17]. In what follows we deal with another interior-point method for (1) – the *method of analytic centers*. The method is as follows: we associate with G, H and K appropriate *barriers* – interior penalty functions $\Psi_G(x)$, $\Psi_H(y)$ and $\Psi_K(y)$, respectively – and trace the path

$$x^*(t_0) = \operatorname*{argmin}_x F_{t_0}(x),$$

$$F_{t_0}(x) \equiv \Psi_G(x) + \Psi_H(t_0 B(x) - A(x)) + \Psi_K(B(x)),$$

as t_0 approaches from above the optimal value of the problem. This is a quite traditional scheme; its potential in the context of interior-point methods for convex problems (cf. Example 1) was discussed by Sonnevend [21], although without any polynomial-time results. The results for this latter type of scheme were first established in the seminal paper of Renegar [18] for the case of Linear Programming (Example 1, G is a polytope); the polynomial-time results for the method were then extended by several authors to more general classes of convex problems. As far as quasiconvex problems (i.e., with nonconstant $B(\cdot)$) are concerned, this case seems to have been studied significantly less. Boyd and El Ghaoui [4] were the first to suggest using the method of analytic centers for the generalized eigenvalue problem; in their important paper, however, they do not establish an overall polynomial-time efficiency estimate. Polynomial-time complexity of the method was proved by Ye [23] (for the von Neumann economic growth problem), Freund and Jarre [8,9] ($K = H = \mathbb{R}^m_+$, cf. Example 4) and Nemirovski [13]; the complexity bound of the latter paper extends to the general case the bound established in [23] and seems to be the best known so far.

Locally quadratically convergent methods for the generalized eigenvalue problem were developed in [10, 11], although without any global convergence analysis.

From the practical viewpoint, the main disadvantage of the known polynomial-time results on the method of analytic centers for fractional problems is that they relate to a short-step version of the method, where the steps in the parameter t_0 are subject to certain a priori restrictions based on theoretical worst-case analysis. In practical computations, it is highly desirable to use "long-step" tactics, but the current theoretical understanding of the method, as far as we know, does not provide a practitioner with theoretically justified (i.e., consistent with known complexity bounds) tools for long steps.

In what follows we develop a long-step version of the method of analytic centers for fractional problems; our approach is mainly based on the recent long-step path-following schemes for convex optimization problems [14, 15]. In fact, we are developing a method which traces a *two-parameter surface of analytic centers* rather than a single-parameter path; this approach has its origin in [15] and avoids the need to come close to the usual path of analytic centers, which, in the traditional schemes, is the goal of a special initial phase of the method.

The paper is organized as follows. Section 2 contains prerequisites on self-concordant functions and barriers, the basic tools we use in our construction. In Section 3 we introduce the notion of the surface of analytic centers associated with problem (1), present the generic scheme of tracing the surface and motivate the advantages of tracing a *surface* rather than the usual *path* of analytic centers. In Section 4 we develop duality-based techniques which underlie the "long-step" tracing of the surface of analytic centers, and Section 5 contains the main results underlying the complexity analysis of the proposed method. These sections deal with "tactics" of tracing the surface of analytic centers: we explain *how* one can move around the surface, not *where* to move. The latter issue is discussed in the concluding Section 6, which contains also the overall polynomial time complexity results.

2. Self-concordant functions and barriers

In this section we present the basic facts from [16] which underlie all our further constructions. Let Q be a nonempty open convex domain in $\mathbb{R}^k$. A function $F : Q \to \mathbb{R}$ is called *strongly self-concordant* (s.s.-c.) on Q, if it is convex, C^3-smooth, Q is the natural domain of F (i.e., $F(x_i) \to \infty$ along any sequence of points $x_i \in Q$ converging to a boundary point of Q) and F satisfies the following differential inequality:

$$|\mathrm{D}^3F(x)[h,h,h]| \leqslant 2(\mathrm{D}^2F(x)[h,h])^{3/2}, \quad x \in Q, \ h \in \mathbb{R}^k$$

($\mathrm{D}^lF(x)[h_1,\dots,h_l]$ is lth differential of F taken at x along the directions $h_1,\dots,h_l$).

Let P be a closed convex subset in $\mathbb{R}^k$ with a nonempty interior Q, and let $\vartheta \geqslant 1$. A function F is called a *ϑ-self-concordant barrier* (ϑ-s.-c.b.) for P, if it is s.s.-c. on Q, and

$$|\mathrm{D}F(x)[h]| \leqslant \vartheta^{1/2}(\mathrm{D}^2F(x)[h,h])^{1/2}, \quad x \in Q, \ h \in \mathbb{R}^k.$$

For explicit self-concordant barriers for a wide variety of convex domains arising in convex optimization, see [16]; several important examples of these barriers will be given in Section 4. In what follows we heavily exploit the following results on self-concordant functions/barriers:

P.0 (see [16, Propositions 2.1.1 and 2.3.1]). *If $\alpha_i \geqslant 1$, F_i are s.s.-c. on convex domains $Q_i \subset \mathbb{R}^k$, $i = 1, \dots, q$, and the intersection Q of these domain is nonempty, then the function $F = \sum_i \alpha_i F_i$ is s.s.-c. on Q; if, in addition, F_i are ϑ_i-s.-c.b.'s for* $\mathrm{cl}\, Q_i$, *then F is a $(\sum_i \alpha_i \vartheta_i)$-s.-c.b. for* $\mathrm{cl}\, Q$.

If F is s.s.-c. on a convex domain $Q \subset \mathbb{R}^k$ and $\mathcal{A}$ is an affine mapping from $\mathbb{R}^q$ to $\mathbb{R}^k$ with the image intersecting Q, then $F^+(\cdot) = F(\mathcal{A}(\cdot))$ is s.s.-c. on the inverse image Q^+ of Q under the mapping $\mathcal{A}$; if, in addition, F is ϑ-s.-c.b. for $\mathrm{cl}\, Q$, *then F^+ is a ϑ-s.-c.b. for* $\mathrm{cl}\, Q^+$.

P.1 (see [16, Theorem 2.1.1 and Proposition 2.3.2]). *Let Q be a convex domain in $\mathbb{R}^k$ which does not contain lines and F be s.s.-c. on Q. Then the Hessian $F''(x)$ of F at any point $x \in Q$ is non-singular, the* Dikin ellipsoid *of F centered at $x \in Q$*

$$W_F(x) = \{y \in \mathbb{R}^k \mid |y - x|_x \equiv [(y-x)^{\mathrm{T}} F''(x)(y-x)]^{1/2} \leqslant 1\}$$

belongs to $\mathrm{cl}\, Q$, *and in the interior of this ellipsoid F'' is "almost proportional" to $F''(x)$, namely,*

$$r \equiv |y - x|_x < 1 \quad \Rightarrow \quad (1-r)^2 F''(x) \leqslant F''(y) \leqslant (1-r)^{-2} F''(x);$$

it follows, in particular, that

$$|y - x|_x < 1 \quad \Rightarrow \quad F(y) \leqslant F(x) + (y-x)^{\mathrm{T}} F'(x) + \rho(|y-x|_x),$$
$$\rho(s) = -\ln(1-s) - s.$$

If, in addition, F is ϑ-s.-c.b. for $\mathrm{cl}\, Q$, *then F attains its minimum on Q if and only if Q is bounded, and the minimizer of F is unique.*

P.2 (see [16, Section 2.2.3]). *Let Q be a bounded convex domain in $\mathbb{R}^k$ and let F be s.s.-c. on Q. Given $y^0 \in Q$, consider the damped Newton minimization of F starting at y^0, i.e., the process*

$$y^{i+1} = y^i - \frac{1}{1 + \lambda(F, y^i)} [F''(y^i)]^{-1} F'(y^i), \tag{2}$$

where the Newton decrement $\lambda(F, x)$ is given by

$$\lambda(F, x) = [(F'(x))^{\mathrm{T}} (F''(x))^{-1} F'(x)]^{1/2}.$$

The process (2) *is well-defined (i.e., keeps the iterates in Q) and, for any $\kappa \in (0, 0.2]$, generates a point y^i with $\lambda(F, y^i) \leqslant \kappa$ in no more than*

$$N = \mathrm{O}(1)\left(\left[F(y^0) - \min_Q F\right] + \ln\ln(1/\kappa)\right)$$

steps, $\mathrm{O}(1)$ *being an absolute constant.*

P.3 (see [16, Corollary 2.3.1]). *Let P be a closed convex domain in $\mathbb{R}^n$, F be a s.-c.b. for P, let $y \in \mathrm{int}\, P$ and h be a recessive direction for P: $P + \mathbb{R}_+ h = P$. Then*

$$-h^{\mathrm{T}} F'(y) \geqslant (h^{\mathrm{T}} F''(y) h)^{1/2}.$$

P.4 (see [16, Section 2.4.2]). *Let $Q \subset \mathbb{R}^n$ be an open convex domain, and let F be a s.s.-c. function on Q with nondegenerate Hessian. Then the domain* $\operatorname{Dom} F^*$ *of the Legendre transformation*

$$F^*(u) = \sup_{x \in Q}\{u^{\mathrm{T}}x - F(x)\}$$

of F (by definition, the domain is comprised of those u for which the right-hand side in the latter expression is finite) is an open convex set and F^* is s.s.-c. on $\operatorname{Dom} F^*$.

P.5 (see [16, Proposition 2.3.2]). *Let F be a ϑ-s.-c.b. for a bounded convex domain P, and let x^* be the minimizer of F on P. Then*

$$P \in \{y \mid |y - x^*|_x \leqslant 1 + 3\vartheta\}.$$

(It was proved by F. Jarre that $1 + 3\vartheta$ in the latter inclusion can be replaced with $\vartheta + 2\sqrt{\vartheta}$.)

3. Surface of analytic centers and basic updating scheme

3.1. Assumptions and notation

Given problem (1), we set

$$G_K = \operatorname{cl}\{x \in \mathbb{R}^n \mid B(x) \in \operatorname{int} K\};$$

from now on we assume that

A. The intersection D of G_K and G is a solid (closed and bounded convex set with a nonempty interior), and we are given in advance a starting point $x^\# \in \operatorname{int} D$.

B. We are given self-concordant barriers Φ_H for H, F_K for G_K and F_G for G, parameters of the barriers being ϑ_H, ϑ_K, ϑ, respectively. It is assumed that

$$\vartheta \geqslant \max\{\vartheta_H, \vartheta_K, 10\}.$$

(This latter assumption does not restrict generality, since a ϑ-s.-c.b. is also ϑ'-s.-c.b. for any $\vartheta' \geqslant \vartheta$. Note also that the only goal of the restriction $\vartheta \geqslant 10$ is to reduce absolute constants coming from terms with ϑ^{-1} in the forthcoming estimates.)

From now on we set

$$\Omega_H = \vartheta/\vartheta_H, \qquad \Omega_K = \vartheta/\vartheta_K.$$

3.2. Surface of analytic centers

Let c be a nonzero vector from $\mathbb{R}^n$, and let $t = (t_0, t_1)^{\mathrm{T}}$ be a 2-dimensional "parameter" vector. We denote by T the set of all values of t for which the domain

$$D_t = \mathrm{cl}\{x \in \operatorname{int} D \mid t_0 B(x) - A(x) \in \operatorname{int} H,\ c^{\mathrm{T}} x < t_1\}$$

is nonempty. If $x \in \operatorname{int} D$, then, by construction of D, $B(x) \in \operatorname{int} K$, so that $B(x)$ is a recessive direction of H; it follows that T is a monotone $(t \in T, t' \geqslant t \Rightarrow t' \in T)$ subset of $\mathbb{R}^2$, and $D_t \subset D_{t'}$ whenever $t \in T$ and $t' \geqslant t$; T is clearly open and nonempty. We denote by T^+ the set of all *primal feasible pairs* (t, x), i.e., those with $t \in T$ and $x \in \operatorname{int} D_t$. Now, for $t \in T$ let

$$F_t(x) = -\vartheta \ln(t_1 - c^{\mathrm{T}} x) + \Omega_H \Phi_H(t_0 B(x) - A(x)) \\ + \Omega_K F_K(x) + F_G(x) : \operatorname{int} D_t \to \mathbb{R}. \tag{3}$$

In view of P.0 F_t is a ϑ^*-s.-c.b. for the domain D_t, with

$$\vartheta^* = 4\vartheta;$$

since D (and, consequently, D_t) is bounded, this barrier attains its minimum at a unique point $x^*(t)$ of the domain D_t (see P.1). Thus, we come to the *surface of analytic centers*

$$\mathcal{S} = \mathcal{S}(c) = \left\{ (t, x^*(t)) \;\middle|\; t \in T,\ x^*(t) = \operatorname*{argmin}_{x \in D_t} F_t(x) \right\}.$$

Let

$$t_0^* = \inf\{t_0 \mid t \in T\};$$

this quantity is the greatest lower bound of those t_0 for which the system of *strict inclusions*

$$x \in \operatorname{int} G, \qquad B(x) \in \operatorname{int} K, \qquad t_0 B(x) - A(x) \in \operatorname{int} H$$

is solvable. We call t_0^* the *regularized optimal value in* (1), and we shall see that under reasonable regularity assumptions this regularized optimal value is the same as the actual optimal value in (1).

By origin of t_0^*, one can travel along the surface $\mathcal{S}(c)$ in a way which enforces the coordinate t_0 to tend to t_0^*, thus obtaining feasible solutions with the value of the objective converging to the regularized optimal value of the problem. This is exactly what we are going to do in order to solve the problem, except the fact that we shall generate strictly feasible pairs which are close, in a sense, to the surface rather than on the surface exactly. Note that the traditional method of analytic centers acts in the same way, but it traces a single-parameter *path* $\mathcal{S}_0$ given by

$$x(t_0) = \operatorname{argmin}\{\Omega_H \Phi_H(t_0 B(x) - A(x)) + \Omega_K F_K(x) + F_G(x)\},$$

not a two-parameter surface. Before coming to detailed description of the method, let us explain what are the advantages of tracing a surface instead of a single-parameter path.

In order to approximate t_0^*, it is, of course, sufficient to trace the path $\mathcal{S}_0$, but to trace the path, one should first come close to it. The standard way to do this is as follows. Given a starting point $x^\# \in \operatorname{int} D$, one chooses $t_0^\#$, $t_0^\# B(x^\#) - A(x^\#) \in \operatorname{int} H$, and then traces, as $t_1 \to \infty$, the auxiliary path $t_0 = t_0^{\#}$ at the surface $\mathcal{S}(c)$, starting with $t_1 = t_1^{\#}$; here $c, t_1^{\#}$ are readily given by the requirement $((t_0^{\#}, t_1^{\#}), x^{\#}) \in \mathcal{S}(c)$, e.g., as

$$c = -\vartheta^{-1}(\Omega_H \nabla \Phi_H(t_0^{\#} B(x^{\#}) - A(x^{\#})) + \Omega_K \nabla F_K(x^{\#}) + \nabla F_G(x^{\#})), \tag{4}$$

$$t^{\#} = (t_0^{\#}, c^{\mathrm{T}} x^{\#} + 1) \tag{5}$$

(from now on ∇ acts with respect to x). It is clearly seen that with this choice of $c, t^{\#}$ the starting pair $(t^{\#}, x^{\#})$ belongs to $\mathcal{S}(c)$.

As $t_1 \to \infty$, the auxiliary path $t_0 = t_0^{\#}$ on $\mathcal{S}(c)$ converges to the point $t_0 = t_0^{\#}$ on the "target" path $\mathcal{S}_0$; thus, tracing the auxiliary path, we eventually come close to the target one and can switch to tracing this latter path. Note that in this traditional two-phase path-following scheme we in fact all the time are traveling along the surface $\mathcal{S}(c)$ (the target path clearly belongs to the closure of the surface). After this is realized, we ask: why should we restrict ourselves to this particular route, where, in the first phase, we disregard the objective? We see that it is reasonable to investigate our abilities to trace surfaces of analytic centers; this is the issue we now examine.

3.3. Basic updating scheme

Let $\kappa \leqslant 0.2$ be a fixed positive tolerance. We say that a pair (t, x) is *close* to $\mathcal{S}$, if the pair satisfies the following predicate

$$\mathcal{P}_\kappa(t, x): \qquad (t, x) \in T^+ \ \& \ \lambda(F_t, x) \leqslant \kappa;$$

recall that the *Newton decrement* $\lambda(F, x)$ of a function F twice continuously differentiable at a point x and possessing a nonsingular Hessian at x is the quantity

$$\lambda(F, x) = ((\nabla F(x))^{\mathrm{T}} [\nabla^2 F(x)]^{-1} \nabla F(x))^{1/2}.$$

It is assumed that we are given a surface $\mathcal{S} \equiv \mathcal{S}(c)$ of analytic centers and a starting pair $(t^{\#}, x^{\#})$ which is close to $\mathcal{S}$, and our goal is to trace the surface, staying close to it, in order to approach a certain "target" point belonging to the closure of the surface. To this end we consider

Basic updating scheme. Given a pair (t, x) close to $\mathcal{S}$, replace it by a new pair (t^+, x^+), also close to $\mathcal{S}$, according to the following rules:

(1) Choose a direction δt in the plane of parameters, and form the associated *primal search ray*

$$X = \{X(r) \equiv (t(r), x(r)) = (t, x + \mathrm{d}_x(t, x)) + r(\delta t, \delta x) \mid r \geqslant 0\}, \tag{6}$$

where

$$\mathrm{d}_x(t, x) = -[\nabla^2 F_t(x)]^{-1} \nabla F_t(x) \tag{7}$$

and δx is given by the relation

$$(\delta t, \delta x) \in \Pi(t, x) = \left\{ (\mathrm{d}t, \mathrm{d}x) \;\middle|\; \left[\frac{\partial}{\partial t} \nabla F_t(x)\right] \mathrm{d}t + \nabla^2 F_t(x)\, \mathrm{d}x = 0 \right\}. \tag{8}$$

Note that

$$\delta x = -[\nabla^2 F_t(x)]^{-1}\{-\vartheta(t_1 - c^{\mathrm{T}}x)^{-2}\delta t_1 c + \Omega_H \delta t_0 \beta^{\mathrm{T}} \Phi'_H(t_0 B(x) - A(x)) + \Omega_H \delta t_0 (t_0\beta - \alpha)^{\mathrm{T}} \Phi''_H(t_0 B(x) - A(x)) B(x)\}. \tag{9}$$

(2) (*predictor step*) Choose a stepsize $\bar{r} > 0$ along the primal search ray and form the forecast

$$(t^+, \bar{x}) = X(\bar{r});$$

the forecast should belong to T^+ (this is a restriction on the stepsize).

(3) (*corrector step*) Apply the damped Newton minimization

$$y^{i+1} = y^i - \frac{1}{1 + \lambda(F_{t^+}, y^i)}[\nabla^2 F_{t^+}(y^i)]^{-1}\nabla F_{t^+}(y^i), \qquad y^0 = \bar{x}, \tag{10}$$

until the pair (t^+, y^i) satisfies $\mathcal{P}_\kappa$; then, set $x^+ = y^i$, thus forming the updated pair (t^+, x^+) which satisfies $\mathcal{P}_\kappa$.

This is the natural two-parameter analogy to the usual predictor-corrector scheme. The origin of relations from item (1) is clear: the surface $\{(\tau, x^*(\tau))\}$ of analytic centers is given by the equation $\nabla F_\tau(\cdot) = 0$; linearizing the equation at (t, x), we get $x^*(t + r\delta t) \approx x + \mathrm{d}_x(t, x) + r\delta x$, with $\mathrm{d}_x(t, x)$ and δx given by (7), (9) (these relations make sense, since $\nabla^2 F_t$ is nonsingular, see P.1 and (A)).

Note that from P.2 it follows that process (10) is well-defined, keeps the iterates y^i in the interior of D_{t^+} and terminates in no more than $\mathrm{O}(1)(V(t^+, x^+) + \ln\ln(1/\kappa))$ Newton iterations $y^i \mapsto y^{i+1}$; from now on $\mathrm{O}(1)$ are positive absolute constants and

$$V(\tau, y) = F_\tau(y) - \min_{u \in \operatorname{int} D_\tau} F_\tau(u) \equiv F_\tau(y) - f^*(\tau).$$

In order to bound from above the Newton complexity (number of Newton iterations) of a corrector step, we fix certain constant $\hat{\kappa} \geqslant 7$ and impose on rule (2)) the following restriction

$\mathcal{R}$: the stepsize $\bar{r}$ is such that $V(t^+, \bar{x}) \leqslant \hat{\kappa}$.

In order to choose the largest possible $\bar{r}$ satisfying $\mathcal{R}$, one could use a line search. The difficulty is, however, that the left-hand side in the latter inequality involves the implicitly defined quantity $f^*(t^+)$, so that we need certain "computationally cheap" technique for bounding $V(\cdot,\cdot)$ from above. To this end we intend to use *dual bounds*, which we now discuss.

4. Dual bounds

To get an upper bound on the quantity $V(\tau, y)$ is the same as to get a lower bound on the quantity $f^*(\tau) = \min_y F_\tau(y)$. This latter problem would be absolutely trivial if we knew the Legendre transformation of $F_\tau(\cdot)$ – the value of this transformation at 0 is exactly $-\min_y F_\tau(y)$. Of course, we have no hope of knowing explicitly the Legendre transformation of F_τ (since otherwise we would immediately know not only the optimal value, but also the minimizer of F_τ – it is the gradient of the Legendre transformation at 0). Nevertheless, in many cases we have certain partial information on the Legendre transformation of F_τ – namely, we can represent $F_\tau(y)$ as a superposition of an affine mapping $y \mapsto u = \pi_\tau y + p_\tau$ and a function $\mathcal{F}(u)$ with known Legendre transformation $\mathcal{F}_*$:

$$(*) \qquad F_\tau(y) = \mathcal{F}(\pi_\tau y + p_\tau).$$

It is possible "to see" $\min_y F_\tau(y)$ in $\mathcal{F}_*$: if $s \in \operatorname{Dom} \mathcal{F}_*$ is such that $\pi_\tau^{\mathrm{T}} s = 0$, then

$$\begin{aligned}
-\min_y F_\tau(y) &= \sup_y [s^{\mathrm{T}} \pi_\tau y - \mathcal{F}(\pi_\tau y + p_\tau)] \\
&= \sup_y [s^{\mathrm{T}}[\pi_\tau y + p_\tau] - \mathcal{F}(\pi_\tau y + p_\tau) - s^{\mathrm{T}} p_\tau] \\
&\leqslant \sup_u [s^{\mathrm{T}} u - \mathcal{F}(u) - s^{\mathrm{T}} p_\tau] \\
&= \mathcal{F}_*(s) - s^{\mathrm{T}} p_\tau,
\end{aligned}$$

so that every $s \in \operatorname{Dom} \mathcal{F}_*$ such that $\pi_\tau^{\mathrm{T}} s = 0$ generates a lower bound on $\min_y F_\tau(y)$; it is easily seen that for properly chosen s this bound is exact. This is the way we intend to use in order to generate lower bounds on $f^*(\cdot)$, and we start with the assumptions on the barriers Φ, F_K, F_G which ensure the possibility of representing F_τ in the form of $(*)$.

4.1. Structural assumptions on the barriers

From now on we make the following assumptions on the barriers Φ_H, F_K and F_G under consideration:

C.1. There exist closed convex domains $G_K^+ \subset \mathbb{R}^{m_K}$ and $G^+ \subset \mathbb{R}^{m_G}$ that do not contain lines, self-concordant barriers Φ_K and Φ_G for these domains, parameters of the barriers being ϑ_K and ϑ, respectively, and affine mappings

$$x \mapsto \pi_K x + p_K : \mathbb{R}^n \to \mathbb{R}^{m_K}, \qquad x \mapsto \pi_G x + p_G : \mathbb{R}^n \to \mathbb{R}^{m_G}$$

such that

$$G = \mathrm{cl}\{x \in \mathbb{R}^n \mid \pi_G x + p_G \in \operatorname{int} G^+\}, \qquad G_K = \mathrm{cl}\{x \in \mathbb{R}^n \mid \pi_K x + p_K \in \operatorname{int} G_K^+\}$$

and

$$F_K(x) = \Phi_K(\pi_K x + p_K), \qquad F_G(x) = \Phi_G(\pi_G x + p_G).$$

C.2. We know the Legendre transformations $\Phi_H^*(\cdot)$, $\Phi_K^*(\cdot)$, $\Phi_G^*(\cdot)$ of the functions Φ_H, Φ_K, Φ_G.

"We know" means that, given a vector s of the corresponding dimension, we may check whether the vector belongs to the domain of the Legendre transformation in question, and if it is the case, we can compute the value of the transformation at s.

Let us demonstrate that the aforementioned assumptions on the barriers are satisfied in a number of interesting and important particular cases.

First of all, our assumptions are "stable with respect to intersections": if, say, we can represent the domain G as an intersection $\bigcap_{i=1}^{k} G^i$ of finitely many domains in such a way that every G^i admits a s.-c.b. of the type $\Phi_i(\pi_i x + p_i)$, where Φ_i is a ϑ_i-s.-c.b. with known Legendre transformation, we may take as Φ_G the function

$$\Phi_G(u_1, \dots, u_k) = \Phi_1(u_1) + \cdots + \Phi_k(u_k)$$

(which results in $\vartheta = \sum_i \vartheta_i$; note that the Legendre transformation of Φ_G is the "direct sum" of those of Φ_i) and set

$$F_G(x) = \Phi_G(\pi_C x + p_C), \qquad \pi_C x + p_C = (\pi_1 x + p_1, \dots, \pi_k x + p_k);$$

of course, we can similarly handle F_K.

Further, our assumptions are "stable with respect to affine substitutions of variables": if, say, we can represent G (similarly for G_K) as an inverse image of a domain $\bar{G}$:

$$G = \mathrm{cl}\{x \mid \mathcal{A}(x) \in \mathrm{int}\, \bar{G}\}$$

under *affine* mapping $\mathcal{A}$ and $\bar{G}$ admits a barrier $\bar{F}_C(u) = \Phi_G(\bar{\pi}_C u + \bar{p}_C)$, Φ_G being a ϑ-s.-c.b. with known Legendre transformation, then G itself admits a desired barrier

$$F_G(x) = \Phi_G(\pi_C x + p_C),$$

the affine mapping $x \mapsto \pi_C x + p_C$ being the superposition of the mappings $x \mapsto \mathcal{A}(x)$ and $u \mapsto \bar{\pi}_C u + \bar{p}_C$.

Thus, the family of "good" domains – those possessing barriers of the required type – is closed with respect to the basic operations such us taking intersections and inverse images under affine mappings (and, as it is immediately seen, taking direct products).

Now let us indicate several "building blocks" which can be used, as G^+ and Φ, in the aforementioned combination rules (for justifications, see [16, Chapter 5]):

(1) *The nonnegative half-axis* $G^+ = \mathbb{R}_+$:

$$\Phi(u) = -\ln u, \qquad \Phi^*(s) = -\ln(-s) - 1 \quad (\vartheta = 1);$$

due to the combination rules, this example in fact covers all our needs in the case when $K = H = \mathbb{R}^m_+$ and G is a polytope;

(2) *The second-order cone* $G^+ = \{u \in \mathbb{R}^q \mid u_q \geqslant (\sum_{i=1}^{q-1} u_i^2)^{1/2}\}$:

$$\Phi(u) = -\ln\Big(u_q^2 - \sum_{i=1}^{q-1} u_i^2\Big),$$

$$\Phi^*(s) = -\ln\Big(s_q^2 - \sum_{i=1}^{q-1} s_i^2\Big) - 2 + 2\ln 2 \quad (\vartheta = 2);$$

due to the combination rules, this observation covers convex quadratic quadratically constrained problems (since the Lebesgue set $\{x \mid f(x) \leqslant 0\}$ of a convex quadratic form f can be represented as an inverse image of a second-order cone under affine mapping) and even more general family of convex programs (e.g., we may handle the hyperbolic domain of the type $\sum_{i=1}^{n-1} x_i^2 + 1 \leqslant x_n^2$, $x_n > 0$).

(3) *The epigraph of the exponent* $G^+ = \{(t,x) \in \mathbb{R}^2 \mid t \geqslant \exp\{x\}\}$:

$$\Phi(t,x) = -\ln(\ln t - x) - \ln t,$$

$$\Phi^*(s,\xi) = (\xi + 1)\ln\left(\frac{\xi+1}{-s}\right) - \xi - \ln \xi - 2 \quad (\vartheta = 2);$$

due to the combination rules, this observation covers Geometrical Programming in the exponential form.

(4) *The cone of positive semidefinite matrices* G^+ *in the space of* $m \times m$ *symmetric matrices*:

$$\Phi(u) = -\ln \operatorname{Det} u, \qquad \Phi^*(s) = -\ln \operatorname{Det}(-s) - m \quad (\vartheta = m);$$

due to the combination rules, this example allows to handle *Linear Matrix Inequality constraints* given by the requirement that a symmetric matrix $\mathcal{A}(x)$ affinely depending on the design vector is positive semidefinite (cf. Example 5).

Thus, our assumption on the structure of barriers Φ_H, F_K and F_G is compatible with a wide spectrum of important fractional problems.

4.2. Dual bounds

From now on we set

$$\mathcal{F}(u) \equiv \mathcal{F}(u_1; u_H; u_K; u_G) = -\vartheta \ln(u_1) + \Omega_H \Phi_H(u_H) + \Omega_K \Phi_K(u_K) + \Phi_G(u_G),$$

so that

$$F_t(x) = \mathcal{F}(U(t,x)), \qquad U(t,x) = \sigma t_1 + t_0[\xi x + p] + \pi x + q, \tag{11}$$

where

$$\sigma t_1 = (t_1; 0; 0; 0); \tag{12}$$

$$\xi x + p = (0; \beta x + b; 0; 0); \tag{13}$$

$$\pi x + q = (-c^{\mathrm{T}} x; -\alpha x - a; \pi_K x + p_K; \pi_G x + p_G). \tag{14}$$

Let $\mathcal{F}^*$ be the Legendre transformation of $\mathcal{F}$:

$$\begin{aligned}\mathcal{F}_*(s) &\equiv \mathcal{F}^*(s_1; s_H; s_K; s_G) \\ &= -\vartheta \ln(-s_1) + \Omega_H \Phi_H^*(s_H/\Omega_H) + \Omega_K \Phi_K^*(s_K/\Omega_K) \\ &\quad + \Phi_G^*(s_G) + \vartheta(\ln \vartheta - 1).\end{aligned}$$

Note that, in view of P.4, $\mathcal{F}_*$ is s.s.-c. on its domain.

Let us make the following simple observation:

Lemma 6. *Let s belong to* $\operatorname{Dom}\mathcal{F}_*$ *and satisfy, for some τ_0, the linear homogeneous equation*

$$\mathcal{E}_{\tau_0}[s] \equiv [\tau_0\xi + \pi]^{\mathrm{T}} s = 0.$$

Then, for all τ_1 such that $\tau = (\tau_0, \tau_1) \in T$, one has

$$f^*(\tau) \geqslant f_s(\tau) \equiv s^{\mathrm{T}}[\sigma\tau_1 + \tau_0 p + q] - \mathcal{F}_*(s).$$

Proof. Let y be such that $(\tau, y) \in T^+$. Since $\mathcal{F}$ is convex and closed, it is the Legendre transformation of $\mathcal{F}_*$, so that from (11) it follows that

$$\begin{aligned} F_\tau(y) \geqslant s^{\mathrm{T}}U(\tau, y) - \mathcal{F}_*(s) &= s^{\mathrm{T}}[\sigma\tau_1 + \tau_0[\xi y + p] + \pi y + q] - \mathcal{F}_*(s) \\ &= f_s(\tau) + s^{\mathrm{T}}[\tau_0\xi + \pi]y = f_s(\tau). \qquad \square \end{aligned}$$

Thus, any vector s satisfying, with respect to a given τ_0, the premise of the above lemma (let us call such s *dual feasible w.r.t.* τ_0) induces lower bounds on the quantities $f^*(\tau_0, \cdot) = \min_x F_{\tau_0, \cdot}(x)$, and these are the bounds we intend to use at the predictor step in order to ensure $\mathcal{R}$. Let us present a systematic way to form these dual feasible vectors.

4.3. Dual search parabola

Let $(t, x) \in T^+$ be close to the surface $\mathcal{S}$, δt be a direction in the parameter space and

$$X = \{X(r) \equiv (t(r) = t + r\delta t,\ x(r) = x + \Delta_x(r)),$$
$$\Delta_x(r) = \mathrm{d}_x(t, x) + r\delta x \mid r \geqslant 0\}$$

be the corresponding primal search ray (see (6)–(8)). The mapping $(\tau, y) \mapsto U(\tau, y)$ (see (11)) transforms this ray into the parabola

$$U(X(r)) = U(t, x) + \Delta_u(r) + \delta_u(r),$$

where

$$\Delta_u(r) = r\sigma\delta t_1 + r\delta t_0[\xi x + p] + [t_0\xi + \pi]\Delta_x(r), \tag{15}$$
$$\delta_u(r) = r\delta t_0 \xi \Delta_x(r). \tag{16}$$

Our goal is to associate with these "primal" entities a dual one – the *dual search parabola*

$$S = \{(t(r) = t + r\delta t, s(r)) \mid r \geqslant 0\}$$

which will provide us with dual feasible vectors. To define S, we first define the matrix

$$Q(x) = \Omega_K \pi_K^{\mathrm{T}} \Phi_K''(\pi_K x + p_K)\pi_K + \pi_G^{\mathrm{T}} \Phi_G''(\pi_G x + p_G)\pi_G$$

and then set

$$s = \mathcal{F}'(U(t,x)), \tag{17}$$

$$\Delta_s(r) = \mathcal{F}''(U(t,x))\Delta_u(r), \tag{18}$$

$$\widehat{s}(r) = s + \Delta_s(r), \tag{19}$$

$$\varepsilon(r) = -r\delta t_0 Q^{-1}(x)\xi^{\mathrm{T}}\Delta_s(r), \tag{20}$$

$$\delta_s(r) = (0; 0; \Omega_K \Phi_K'' \pi_K \varepsilon(r); \Phi_G''(\pi_G x + p_G)\pi_G \varepsilon(r)), \tag{21}$$

$$s(r) = \widehat{s}(r) + \delta_s(r) = s + \Delta_s(r) + \delta_s(r). \tag{22}$$

Note that the construction is well defined, since the matrix $Q(x)$ is positive definite (indeed, it is the Hessian at x of the s.-c.b. $\Omega_K F_K(\cdot) + F_G(\cdot)$ for *bounded* domain D).

In what follows, given a positive semidefinite symmetric matrix Q and a vector u of the corresponding dimension, we denote by $|u|_Q$ the Euclidean seminorm

$$|u|_Q = (u^{\mathrm{T}} Q u)^{1/2}.$$

Our local goal is to demonstrate that $s(r)$ is, for small r, dual feasible w.r.t. $t_0 + r\delta t_0$.

Lemma 7. *Let $(t,x) \in T^+$ be close to $\mathcal{S}$. Then*

$$\mathcal{E}_{t_0 + r\delta t_0}[s(r)] \equiv 0. \tag{23}$$

Furthermore,

$$|\Delta_s(r)|_{\mathcal{F}_*''(s)} = |\Delta_u(r)|_{\mathcal{F}''(U(t,x))}; \tag{24}$$

in particular,

$$|\Delta_s(0)|_{\mathcal{F}_*''(s)} = |\Delta_u(0)|_{\mathcal{F}''(U(t,x))} = |\mathrm{d}_x(t,x)|_{\nabla^2 F_t(x)} = \lambda(F_t, x) \leqslant \kappa, \tag{25}$$

and $s(0)$ is dual feasible w.r.t. t_0.

Proof. To simplify notation, let us omit explicit arguments in $\mathcal{F}$, $\mathcal{F}_*$ and F_t; in what follows, these arguments are, respectively, $U(t,x)$, $s = \mathcal{F}'(U(t,x))$ and x.

Let us first verify that

$$e(r) \equiv \mathcal{E}_{t_0 + r\delta t_0}[\widehat{s}(r)] = r\delta t_0 \xi^{\mathrm{T}} \Delta_s(r). \tag{26}$$

To this end, note that (8) can be rewritten as

$$[t_0\xi + \pi]^{\mathrm{T}} \mathcal{F}''[\sigma\delta t' + \delta t_0[\xi x + p] + [t_0\xi + \pi]\delta x] = -\delta t_0 \xi^{\mathrm{T}} \mathcal{F}', \tag{27}$$

while (7) means that

$$[t_0\xi+\pi]^{\mathrm{T}}\mathcal{F}''[t_0\xi+\pi]\mathrm{d}_x(t,x) = -[t_0\xi+\pi]^{\mathrm{T}}\mathcal{F}'. \tag{28}$$

Multiplying (27) by r and adding (28), we have

$$\begin{aligned}[t_0\xi+\pi]^{\mathrm{T}}\Delta_s(r) &\equiv [t_0\xi+\pi]^{\mathrm{T}}\mathcal{F}''\Delta_u(r) = -[(t_0+r\delta t_0)\xi+\pi]^{\mathrm{T}}\mathcal{F}' \\ &\equiv -[(t_0+r\delta t_0)\xi+\pi]^{\mathrm{T}}s,\end{aligned} \tag{29}$$

whence $\mathcal{E}_{t_0+r\delta t_0}[s+\Delta_s(r)] = r\delta t_0\xi^{\mathrm{T}}\Delta_s(r)$, as required in (26).

Now let us prove (23):

$$\begin{aligned}\mathcal{E}_{t_0+r\delta t_0}[s(r)] &= \mathcal{E}_{t_0+r\delta t_0}[\widehat{s}(r)] + \mathcal{E}_{t_0+r\delta t_0}[\delta_s(r)] \quad \text{(by (22))} \\ &= e(r) + \mathcal{E}_{t_0+r\delta t_0}[\delta_s(r)] \quad \text{(origin of } e(r)) \\ &= e(r) + [(t_0+r\delta t_0)\xi+\pi]^{\mathrm{T}}\delta_s(r) \quad \text{(origin of } \mathcal{E}) \\ &= e(r) + \Omega_K\pi_K^{\mathrm{T}}\Phi_K''(\pi_Kx+p_K)\pi_K\varepsilon(r) + \pi_G^{\mathrm{T}}\Phi_G''(\pi_Gx+p_G)\pi_G\varepsilon(r) \\ &\qquad \text{(by (13), (14) and (21))} \\ &= e(r) + Q(x)\varepsilon(r) \quad \text{(origin of } Q(x)) \\ &= 0 \quad \text{(by (20) and (26))}.\end{aligned}$$

Further, $\mathcal{F}_*'' = [\mathcal{F}'']^{-1}$, since the argument in the left-hand side is the gradient of $\mathcal{F}$ at the point involved into the right-hand one; this observation, in view of $\Delta_s(r) = \mathcal{F}''\Delta_u(r)$, immediately results in (24).

Last, we have $\Delta_u(0) = [t_0\xi+\pi]\mathrm{d}_x(t,x)$, and from (11) it follows that $\nabla^2 F_t = [t_0\xi+\pi]^{\mathrm{T}}\mathcal{F}''[t_0\xi+\pi]$; therefore $|\Delta_u(0)|^2_{\mathcal{F}''} = |\mathrm{d}_x(t,x)|^2_{\nabla^2 F_t} = |\nabla F_t|^2_{[\nabla^2 F_t]^{-1}} = \lambda^2(F_t,x)$ (the second equality is given by (7)), as required in (25). It remains to note that since (t,x) is close to $\mathcal{S}$, we have $\lambda(F_t,x) \leqslant \kappa < 1$, and (25) implies that $s(0)$ belongs to the open Dikin ellipsoid of the (as we know, s.s.-c.) function $\mathcal{F}_*$, the ellipsoid being centered at $s = \mathcal{F}' \in \operatorname{Dom}\mathcal{F}_*$; therefore $s(0) \in \operatorname{Dom}\mathcal{F}_*$, and since $\mathcal{E}_{t_0}[s(0)] = 0$ by (23), $s(0)$ is dual feasible w.r.t. t_0. □

4.3.1. Acceptability test

Lemma 7 states that, for small r, the point $s(r)$ is close to the point $s(0)$ from the domain of $\mathcal{F}_*$ and, consequently, itself belongs to this domain; besides this, $s(r)$ satisfies the equation $\mathcal{E}_{t_0+r\delta t_0}[\cdot] = 0$. Thus, at least for small r the point $s(r)$ is dual feasible w.r.t. $t_0(r) = t_0 + r\delta t_0$. This observation underlies the sufficient condition for $\mathcal{R}$ that we are about to present.

Acceptability test. Given a primal search ray $\{X(r) = (t, x+\mathrm{d}_x(t,x)) + r(\delta t,\delta x) \mid r \geqslant 0\}$ ((t,x) satisfies $\mathcal{P}_\kappa$, $(\delta t,\delta x) \in \Pi(t,x)$), and a candidate stepsize r, act as follows:

(a) Check whether $X(r) \in T^+$; if not, reject r.
(b) Compute $s(r)$ according to (17)-(22), and check whether $s(r) \in \operatorname{Dom} \mathcal{F}_*$; if not, reject r.
(c) Compute the quantity (see Lemma 6)

$$\begin{aligned} V(r) &= F_{t+r\delta t}(x + \mathrm{d}_x(t,x) + r\delta x) - f_{s(r)}(t + r\delta t) \\ &= F_{t+r\delta t}(x + \mathrm{d}_x(t,x) + r\delta x) + \mathcal{F}_*(s(r)) \\ &\quad - [s(r)]^{\mathrm{T}}[\sigma(t_1 + r\delta t_1) + (t_0 + r\delta t_0)p + q]. \end{aligned}$$

If $V(r) > \hat{\kappa}$, reject r.
If r was not rejected, claim that r satisfies $\mathcal{R}$.

An immediate consequence of Lemmas 6 and 7 is the following proposition:

Proposition 8. *The Acceptability test is valid: if a stepsize r passes the test, then $(t + r\delta t, x + \mathrm{d}_x(t,x) + r\delta x) \in T^+$ and*

$$F_{t+r\delta t}(x + \mathrm{d}_x(t,x) + r\delta x) - \min_{u \in \operatorname{int} D_{t+r\delta t}} F_{t+r\delta t}(u) \leqslant \hat{\kappa}.$$

5. Main propositions

In this section we formulate the main results underlying our policy of tracing the surface of analytic centers and the complexity analysis of the resulting method. The corresponding proofs are given in the Appendix.

Acceptable steps. The above constructions give a possibility of implementing the basic scheme for tracing the surface $\mathcal{S}$ in a way which ensures a fixed Newton complexity of the corrector step; in view of P.3 and Proposition 8, to this end it suffices to choose as the stepsize $\bar{r}$ a quantity passing the Acceptability test. To derive polynomial time complexity bounds, we should, of course, know that the test is "reasonable", i.e., that it for sure accepts stepsizes of certain "not too small" length. The corresponding statement is:

Theorem 9. *Let $(t,x) \in T^+$ satisfy $\mathcal{P}_\kappa$, and let $\omega \leqslant 0.05$ be a positive real. Assume that a direction δt satisfies the assumptions*

$$\vartheta \delta t_i^2 \leqslant (t_1 - c^{\mathrm{T}}x)^2 \omega^2, \tag{30}$$

$$\mu(t,x)|\delta t_0| \leqslant \omega \sqrt{\frac{\Omega_K}{\Omega_H^2 + \Omega_H \Omega_K}}, \quad \mu(t,x) = -B^{\mathrm{T}}(x)\Phi_H'(t_0 B(x) - A(x)), \tag{31}$$

and let δx be defined by δt in such a way that $(\delta t, \delta x) \in \Pi(t,x)$. Then any stepsize $r \in [0,1]$ passes the Acceptability test.

Growth of the potential. The next statement is the key to complexity analysis of the method:

Theorem 10. *The function*

$$f^*(\tau) = \min_{y \in \mathrm{int}\, D_\tau} F_\tau(y) : T \to \mathbb{R}$$

is nonincreasing in τ_0, and if (t, x) satisfies $\mathcal{P}_\kappa$ and

$$\Delta t_0 = -\omega \sqrt{\frac{\Omega_K}{\Omega_H^2 + \Omega_H \Omega_K}} \mu^{-1}(t, x), \quad 0 < \omega \leqslant \frac{1}{20}, \tag{32}$$

(see (31)), then $t^+ = (t_0 + \Delta t_0, t_1) \in T$ and

$$f^*(t^+) \geqslant f(t) + \frac{\omega}{16} \sqrt{\frac{\Omega_H \Omega_K}{\Omega_H + \Omega_K}}. \tag{33}$$

6. Tracing the surface

We have developed a technique which allows, given a pair (t, x) close to the surface and a direction δt in the plane of parameters, to perform something like the largest predictor step compatible with the predicate $\mathcal{R}$ (and thus ensuring an a priori bound on the Newton complexity of the subsequent corrector step). Thus, we know *how* to travel along the surface, but we did not discuss *where* to travel. In the usual path-following scheme the latter question does not arise at all – the only reasonable strategy is to decrease t_0, the only parameter of interest. This is not the case with a two-parameter surface, since here there are many candidate strategies for traveling from the starting point to the desired optimal solution.

Recall that we have associated with problem (P) equipped with a starting point $x^\# \in \mathrm{int}\, D$ $(D = G_K \cap G,\ G_K = \mathrm{cl}\{x \mid B(x) \in \mathrm{int}\, K\})$ and with a starting value $t_0^\#$ of the parameter t_0, $t_0^\# B(x^\#) - A(x^\#) \in \mathrm{int}\, H$, two-parameter surface $\mathcal{S} = \mathcal{S}(c)$ (see (4)). This surface which passes through the point $x^\#$, the corresponding value of the parameter vector $t = (t_0, t_1)$ being $t^\# = (t_0^\#, t_1^\#)$, see (5). In order to solve the problem, it suffices to travel along $\mathcal{S}$, starting from $(t^\#, x^\#)$, in a way which ensures that the "parameter of interest" t_0 approaches the regularized optimal value t_0^* of (P). As for the "centering" parameter t_1, we should make it large enough, since with too small values of the parameter the artificial constraint $c^{\mathrm{T}} x \leqslant t_1$ may change the optimal value of t_0 we actually are interested in. In fact all we need is to make the artificial constraint *redundant*, i.e., to ensure that

$$t_1 > \max\{c^{\mathrm{T}} x \mid x \in D_t\};$$

then, of course, the quantity

$$t_0^*(t_1) = \inf\{t_0 \mid (t_0, t_1) \in T\}$$

coincides with the regularized optimal value

$$t_0^* = \inf\{t_0 \mid (t_0, \tau) \in T\}$$

of (P). Consequently, after redundancy is detected, we may fix the value of the centering parameter and completely focus on decreasing the parameter of interest. Thus, at the *initial phase*, before the redundancy is detected, we decrease t_0 and increase t_1, and at the *main phase*, after the redundancy is detected, we decrease t_0 and keep t_1 constant.

To implement this idea, we need, first, a test for detecting redundancy, and, second, a "safe" strategy for the initial phase, to ensure a reasonable duration of the phase. These are the issues we now discuss.

6.1. Detecting redundancy

To detect redundancy, one can use the following

Redundancy test. Given an iterate (t,x) satisfying $\mathcal{P}_\kappa$, compute the quantity

$$\psi = \left(c^{\mathrm{T}}[\nabla^2 F_t(x)]^{-1}c\right)^{1/2}$$

and check whether

$$7\vartheta\psi \leqslant t_1 - c^{\mathrm{T}}x; \tag{34}$$

if it is the case, claim that redundancy is achieved.

Lemma 11. *Let $\kappa \leqslant 0.2$. Then the aforementioned test is correct: if (t,x) satisfies $\mathcal{P}_\kappa$, then (34) implies that*

$$D_{(\tau,t_1)} = \{x \in D \mid \tau B(x) - A(x) \in H\} \quad \forall \tau \in (t_0^*, t_0], \tag{35}$$

t_0^* being the regularized optimal value in (P).

Proof. For the sake of brevity, let us write F instead of F_t. Let x^* be the minimizer of F; we have

$$\lambda(F,x) \leqslant \kappa \leqslant \frac{1}{5}.$$

This relation, in view of [16, Theorem 2.2.2(iii)], implies that

$$\sqrt{(x-x^*)^{\mathrm{T}}F''(x)(x-x^*)} \leqslant 1-(1-3\kappa)^{1/3} \leqslant 0.27. \tag{36}$$

In view of this latter fact and P.1 we have

$$F''(x^*) \geqslant (0.73)^2 F''(x). \tag{37}$$

Furthermore, since F is a (4ϑ)-s.-c.b. for D_t, $\vartheta \geqslant 10$ and x^* is the minimizer of the barrier, P.5 implies that

$$\sqrt{(y-x^*)^{\mathrm{T}}F''(x^*)(y-x^*)} \leqslant 4.9\vartheta, \quad y \in D_t.$$

Thus, (36)–(37) imply that

$$\sqrt{(y-x)^{\mathrm{T}}F''(x)(y-x)} \leqslant 0.27 + (4.9/0.73)\vartheta < 7\vartheta, \quad y \in D_t, \tag{38}$$

which combined with the definition of ψ and (34) results in

$$\max_{y \in D_t} c^{\mathrm{T}} y < c^{\mathrm{T}} x + 7\psi\vartheta \leqslant t_1.$$

Thus, (34) does imply the equality in (35) for $\tau = t_0$ and, consequently, for all $\tau \in (t_0^*, t_0]$, as claimed. □

6.2. *The method*

Now we are able to summarize the description of the method for solving (P). Our strategy will be as follows: given a starting point $x^\# \in \operatorname{int} D$, we first define a surface of analytic centers $\mathcal{S}(c)$ which passes through $x^\#$ for some explicitly defined value $t^\#$ of the parameter vector. Then we use the basic updating scheme equipped with the Acceptability test in order to generate a sequence of (close to $\mathcal{S}$) pairs (t^i, x^i) $((t^1, x^1) = (t^\#, x^\#))$ with t_0^i converging to the regularized optimal value of the problem. In this process, we all the time decrease the "parameter of interest" t_0 and never decrease the "centering parameter" t_1. The centering parameter is increased at the *initial phase*, until the constraint $c^{\mathrm{T}} x \leqslant t_1^i$ becomes redundant in the description of the domain D_{t^i}, and is kept constant at the subsequent *main phase*. At the main phase, the parameter of interest is decreased as fast as it is allowed by the Acceptability test; this is not the case at the initial phase, where our policy is aimed to ensure reasonable duration of the phase; to this end we choose the directions δt^i in a way which guarantees "fast decrease" of the quantities $f^*(t^i) = \min_x F_{t^i}(x)$ with i, which is a convenient implicit way to make t_1 redundant.

The implementation of the outlined strategy is as follows.

Initialization. Given $x^\# \in \operatorname{int} D$, choose $t_0^\#$ such that

$$t_0^\# B(x^\#) - A(x^\#) \in \operatorname{int} H \tag{39}$$

and

$$\Omega_H \nabla^2 \Phi_H (t_0^\# B(x^\#) - A(x^\#)) \leqslant \nabla^2 (\Omega_K \Phi_K(\pi_K x + p_K) + \Phi_G(\pi_G x + p_G)) \tag{40}$$

and define c and $t_1^\#$ according to (4)–(5), thus defining the 2-parameter surface $\mathcal{S} \equiv \mathcal{S}(c)$ of analytic centers and a pair $(t^\#, x^\#)$, $t^\# = (t_0^\#, t_1^\#)$, belonging to the surface.

Remark 12. We shall demonstrate that (39) is satisfied for all large enough values of $t_0^\#$ and that $\nabla^2 \Phi_H(t_0 B(x^\#) - A(x^\#)) \to 0$, $t_0 \to \infty$, so that it is easy to ensure (39)-(40); it suffices to start with an arbitrary trial value τ of $t_0^\#$ and to test the values, say, $\tau + 2^k$, $k = 0, 1, \ldots$ until the required relations are satisfied.

ith iteration, $i \geqslant 1$. Given a pair (t^i, x^i) satisfying $\mathcal{P}_\kappa$ $((t^1, x^1) = (t^\#, x^\#))$, act as follows:

(0) Compute the quantities $\mu(t^i, x^i)$, $\psi^*(t^i, x^i)$, where

$$\mu(t,x) = -B^{\mathrm{T}}(x)\Phi'_H(t_0 B(x) - A(x));$$
$$\psi_i^*(t,x) = \max\{z \mid (t_0 - z)B(x) - A(x) \in H\};$$

(1) Choose a direction $\delta t^i = (\delta t_0^i, \delta t_1^i)$ in the plane of parameters as

$$\delta t^i = 0.05 \begin{cases} \left(\dfrac{t_1^i - c^{\mathrm{T}} x^i}{\sqrt{\vartheta}}, -\min\left[\mu^{-1}(t^i, x^i)\sqrt{\dfrac{\Omega_K}{\Omega_H^2 + \Omega_H \Omega_K}}; 0.5\dfrac{\psi^*(t^i, x^i)}{\sqrt{\vartheta}}\right]\right) \\ \text{initial phase,} \\ \left(0, -\mu^{-1}(t^i, x^i)\sqrt{\dfrac{\Omega_K}{\Omega_H^2 + \Omega_H \Omega_K}}\right) \\ \text{main phase.} \end{cases} \tag{41}$$

(2) Apply to the pair (t^i, x^i) and the direction δt^i the basic updating scheme from Section 3.3 equipped with the Acceptability test in order to ensure $\mathcal{R}$, the corresponding stepsize r_i being subject to the restrictions

$$1 \leqslant r_i \leqslant R_i,$$

where

$$R_i = \begin{cases} 0.5\sqrt{\vartheta} & \text{initial phase,} \\ +\infty & \text{main phase.} \end{cases}$$

The new iterate (t^{i+1}, x^{i+1}) is the result given by the basic updating scheme.

Remark 13. From (41) and Theorem 9 it follows that the stepsize $r_i = 1$ passes the Acceptability test, so that (2) is consistent; moreover, the " short-step" version of the method ($r_i \equiv 1$) does not require any line search and dual bounding.

To get a "practical" algorithm, it is, of course, reasonable to use a line search to get the largest possible stepsize $r_i \in [1, R_i]$ accepted by the Acceptability test; note that this line search is computationally inexpensive compared to our natural "complexity unit" – the arithmetic cost of a Newton step.

(3) At the initial phase, subject (t^{i+1}, x^{i+1}) to the Redundancy test; if the pair passes the test, switch to the main phase. Loop.

6.3. Complexity analysis

To present complexity analysis of the method, we need an additional regularity assumption as follows:

D. (P) is solvable and there exists an optimal solution ($t_0 = \tau^*, x = x^*$) to the problem such that

$$B(x^*) \in \operatorname{int} K.$$

Assumption (D) can be interpreted as "well-posedness" of (P); when it is violated, it seems to be impossible to bound the complexity of solving (P). Indeed, consider the following example: G is a solid in $\mathbb{R}^n$, $H = K = \mathbb{R}^2_+$, $B(x) = (b^{\mathrm{T}}x, 1)$, $A(x) = (b^{\mathrm{T}}x, 0)$, where the vector b is such that $\min\{b^{\mathrm{T}}x \mid x \in G\} = 0$. In this example, the optimal value in (P) is 0 and is achieved at the set of minimizers of $b^{\mathrm{T}}x$ on G; outside this set the inclusion $t_0 B(x) - A(x) \in H$ implies that $t_0 \geqslant 1$. Thus, to solve the indicated fractional program within accuracy, say, $1/2$, i.e., to find a feasible pair (t_0, x) with $t_0 \leqslant 1/2$, is the same as to solve the program *exactly*; this can be done with finite computational effort only for a very restricted family of solids G.

Theorem 14. *Let problem* (P) *satisfying assumptions* (A), (B), (C) *be solved by the method presented in Section* 6.2. *Then*

(i) *The duration of the initial phase does not exceed*

$$\mathcal{N}_{\mathrm{ini}} = \mathrm{O}(1)\sqrt{\vartheta}\ln(\vartheta[D : x^{\#}]) \tag{42}$$

iterations; *here and in what follows all* $\mathrm{O}(1)$*s are absolute constants, and*

$$[D : x^{\#}] = \max\{\alpha \mid \exists h : x^{\#} - h \notin D,\ x^{\#} + \alpha h \in D\}$$

is the asymmetry coefficient of D *with respect to* $x^{\#}$.

(ii) *Let* $(t^{\mathrm{m}}, x^{\mathrm{m}})$ *be the pair which starts the main phase of the method, and let* $\varepsilon \in (0, 1)$. *The number of iterations of the main phase which results in a pair* (t^i, x^i) *such that*

$$t^i \leqslant \tau^* + \varepsilon(t^{\mathrm{m}} - \tau^*)$$

does not exceed the quantity

$$\mathcal{N}(\varepsilon) = \mathrm{O}(1)\sqrt{\vartheta(\vartheta_H + \vartheta_K)}\ln\left(1 + \frac{\Theta}{\varepsilon}\right), \tag{43}$$

where

$$\Theta = \min\{s \geqslant 0 \mid sB(x^*) - B(x^*(t^{\mathrm{m}})) \in K\}$$

and $x^*(t) = \operatorname{argmin}_{x \in \operatorname{int} D_t} F_t(x)$.

(iii) *The Newton complexity* (*number of Newton iterations*) *of any corrector step of the method does not exceed*

$$\mathrm{O}(1)(\hat{\kappa} + \ln\ln(1/\kappa)).$$

Proof. Let us start with proving correctness of the initialization rule for t_0; as it was explained in Remark 12, to this end it suffices to verify that $t_0 B(x^{\#}) - A(x^{\#}) \in \operatorname{int} H$

for all large enough t_0 (this is evident, since $B(x^{\#}) \in \operatorname{int} K$ and K is contained in the recessive cone of H) and that

$$R(t_0) \equiv \nabla^2 \Phi_H(t_0 B(x^{\#} - A(x^{\#})) \to 0, \quad t_0 \to \infty. \tag{44}$$

To prove the latter relation, let us choose t_0^* in such a way that $y = t_0^* B(x^{\#}) - A(x^{\#}) \in \operatorname{int} H$, and let U be a convex symmetric neighbourhood of the origin in $\mathbb{R}^m$ such that $B(x^{\#}) + U \subset K$. Since K is contained in the recessive cone of H, for $z \geqslant 0$ we have

$$y(z) + zU \subset H, \qquad y(z) = (t_0^* + z)B(x^{\#}) - A(x^{\#}).$$

It remains to note that (44) is an immediate consequence of the latter relation due to the following general fact:

(*) *Let P be a closed convex domain, F be a γ-self-concordant barrier for P, y be an interior point of P and V be a convex symmetric neighbourhood of the origin such that $y + V \subset P$. Then*

$$z^{\mathrm{T}} \nabla^2 F(y) z \leqslant (1 + 5\gamma)^2 \quad \forall z \in V.$$

The proof is immediate: to simplify notation, let $y = 0$. The function $\Psi(x) = F(x) + F(-x)$ is 2γ-s.-c.b. for the convex domain $P' = P \cap (-P)$, and 0 is the minimizer of the Ψ on P'. From P.5 it follows that

$$z^{\mathrm{T}}(2\nabla^2 F(0))z \leqslant (1 + 6\gamma)^2 \quad \forall z \in P' \supset V,$$

and (*) follows.

Let us prove (i).

(i.1) Let us first verify that the quantities $f_i = F_{t^i}(x^i)$ "quickly decrease" during the initial phase: if step i belongs to the phase, then

$$f_i - f_{i+1} \geqslant \mathrm{O}(1)\sqrt{\vartheta}. \tag{45}$$

Indeed, we have

$$\begin{aligned} f_i - f_{i+1} = &\{F_{t^i}(x^i) - F_{t^{i+1}}(x^i)\}_{\mathrm{I}} \\ &+ \{F_{t^{i+1}}(x^i) - \min_x F_{t^{i+1}}(x)\}_{\mathrm{II}} \\ &+ \{\min_x F_{t^{i+1}}(x) - f_{i+1}\}_{\mathrm{III}}. \end{aligned} \tag{46}$$

The quantity $\{\cdot\}_{\mathrm{II}}$ is nonnegative. Further, since $\lambda \equiv \lambda(F_{t^{i+1}}, x^{i+1}) \leqslant \kappa \leqslant 0.2$, we have

$$\{\cdot\}_{\mathrm{III}} \geqslant -\rho(\lambda) \geqslant -0.024 \tag{47}$$

(see [16, Theorem 2.1.1]). It remains to evaluate $\{\cdot\}_{\mathrm{I}}$. Taking into account the structure of F_t, dropping index i ($r = r_i$, $\delta t = \delta t^i$, and so on), setting $y = t_0 B(x) - A(x)$, $\mathrm{d}y = \delta t_0 B(x)$, we have

$$\{\cdot\}_{\rm I} = \vartheta \ln\left(1 + \frac{r\delta t_1}{t_1 - c^{\rm T}x}\right) + \Omega_H[\Phi_H(y) - \Phi_H(y + r\,{\rm d}y)]$$
$$= \vartheta \ln(1 + 0.05r\vartheta^{-1/2}) + \Omega_H[\Phi_H(y) - \Phi_H(y + r\,{\rm d}y)]. \tag{48}$$

Now, since $0 > \delta t_0 \geqslant -0.025\vartheta^{-1/2}\psi^*(t,x)$ by (41), we have $y + 40\vartheta^{1/2}\,{\rm d}y \in H$ (definition of ψ^*). Now let us use the following fact [16, Proposition 2.3.2]:

(**) Let F be a ϑ-s.c.b. for a domain P, let $y \in \operatorname{int} P$, and let $y + R\,{\rm d}y \in P$. Then

$$F(y + t\,{\rm d}y) \leqslant F(y) - \vartheta \ln(1 - r/R), \quad 0 \leqslant r < R.$$

Applying (**) to the barrier $\Omega_H\Phi_H$ for $P = H$ and our y and ${\rm d}y$ ($R = 40\vartheta^{1/2}$), we get

$$\Omega_H[\Phi_H(y) - \Phi_H(y + r\,{\rm d}y)] \geqslant \vartheta \ln(1 - 0.025\vartheta^{-1/2}r),$$

so that (48) results in

$$\{\cdot\}_{\rm I} \geqslant \vartheta \ln(1 + g - 2g^2), \quad g = 0.025r\vartheta^{-1/2}.$$

Since at the initial phase $1 \leqslant r \leqslant 0.5\vartheta^{1/2}$, we have $2g^2 \leqslant 0.025g$, so that $\{\cdot\}_{\rm I} \geqslant \vartheta \ln(1 + 0.975g)$. Summarizing our observations, we conclude from (46) that

$$f_i - f_{i+1} \geqslant \vartheta \ln(1 + 0.024\vartheta^{-1/2}r) - 0.024 \geqslant 0.5\vartheta \ln(1 + 0.024\vartheta^{-1/2}r)$$

(we have taken into account that $\vartheta \geqslant 10$ and $r \geqslant 1$), and (45) follows.

(i.2) Now we are able to bound from above the duration of the initial phase. To simplify notation, in the below reasoning we assume that $x^\# = 0$ (which, of course, does not restrict generality). Assume that step i belongs to the phase and is not the final step of it. Setting

$$\Psi_\tau(x) = \Omega_H\Phi_H(\tau B(x) - A(x)) + F(x),$$
$$F(x) = \Omega_K\Phi_K(\pi_K x + p_K) + \Phi_G(\pi_G x + p_G),$$

we conclude from (45) that

$$\vartheta \ln(t^{i+1} - c^{\rm T}x^{i+1}) \geqslant {\rm O}(1)\vartheta^{1/2}i + \Psi_{t_0^{i+1}}(x^{i+1}) - \Psi_{t_0^\#}(0)$$

(we have taken into account that $t_1^\# - c^{\rm T}x^\# = 1$ by (5)). Since $t_0^{i+1} \leqslant t_0^\#$ and $\Psi_\tau(x)$ clearly is nonincreasing in τ whenever $x \in \operatorname{int} D$ is such that $\tau B(x) - A(x) \in \operatorname{int} H$ (see P.3 and take into account that $B(x)$ is a recessive direction for H when $x \in \operatorname{int} D$), we have $\Psi_{t_0^{i+1}}(x^{i+1}) \geqslant \Psi_{t_0^\#}(x^{i+1})$, and we come to

$$\vartheta \ln(t^{i+1} - c^{\rm T}x^{i+1}) \geqslant {\rm O}(1)\vartheta^{1/2}i + \Psi(x^{i+1}) - \Psi(0), \quad \Psi(\cdot) \equiv \Psi_{t_0^\#}(\cdot). \tag{49}$$

Now let

$$W = \{y \mid y^{\rm T}\nabla^2 F(0) y \leqslant 1\}, \qquad W' = \{y \mid y^{\rm T}\nabla^2 \Psi(0) y \leqslant 1\}$$

be the Dikin ellipsoids, centered at $x^{\#} = 0$, of the barriers F and Ψ, respectively, and let $D' = D \cap (-D)$ be the symmetrization of D. By definition of the asymmetry coefficient $\alpha \equiv [D : x^{\#}]$, we have $D \subset \alpha D'$, while by $(*)$ we have $D' \subset (1 + 15\vartheta)W$; thus,

$$D \subset 16\alpha\vartheta W \subset 32\alpha\vartheta W' \tag{50}$$

(the concluding inclusion follows from $W \subset \sqrt{2}W'$, see (40)). Now, the function $\Psi(x)$ is 3ϑ-s.-c.b for the domain

$$P = \{x \in D \mid t_0^{\#}B(x) - A(x) \in H\},$$

and both x^{i+1} and W' are contained in this domain; setting $y = x^{i+1}$, $\mathrm{d}y = x^{\#} - x^{i+1} = -x^{i+1}$, we get $y \in \operatorname{int} P$, $y + [1 + (32\alpha\vartheta)^{-1}]\,\mathrm{d}y \in P$ (the latter inclusion follows from (50) and $W' \subset P$). Applying $(**)$, we get

$$\Psi(0) \leqslant \Psi(x^{i+1}) + 3\vartheta \ln(1 + 32\alpha\vartheta),$$

so that (49) implies that

$$\ln(t^{i+1} - c^{\mathrm{T}}x^{i+1}) \geqslant \mathrm{O}(1)\vartheta^{-1/2}i - 3\ln(1 + 32\alpha\vartheta). \tag{51}$$

On the other hand, we know that redundancy was not detected at the step i, i.e., that

$$7\vartheta(c^{\mathrm{T}}[\nabla^2 F_{t^{i+1}}(x^{i+1})]^{-1}c)^{1/2} > t_i^{i+1} - c^{\mathrm{T}}x^{i+1}. \tag{52}$$

The quantity $2(c^{\mathrm{T}}[\nabla^2 F_{t^{i+1}}(x^{i+1})]^{-1}c)^{1/2}$ is the variation (i.e., maximum minus minimum) of the linear form $c^{\mathrm{T}}x$ on the Dikin ellipsoid, centered at x^{i+1}, of the barrier $F_{t^{i+1}}$; this ellipsoid, say V, is contained in $D_{t^{i+1}}$ and, consequently, in D. Taking into account (50), we conclude that the variation of $c^{\mathrm{T}}x$ on V is at most $32\alpha\vartheta$ times the variation $2(c^{\mathrm{T}}[\nabla^2 F_{t^{\#}}(0)]^{-1}c)^{1/2}$, the latter quantity clearly being $\leqslant 2(c^{\mathrm{T}}[\nabla^2\Psi(0)]^{-1}c)^{1/2}$. From (4) and from the fact that Ψ is 3ϑ-s.-c.b. it follows that $(c^{\mathrm{T}}[\nabla^2\Psi(0)]^{-1}c)^{1/2} \leqslant 2\vartheta^{-1/2}$. Combining our observations, we come to

$$(c^{\mathrm{T}}[\nabla^2 F_{t^{i+1}}(x^{i+1})]^{-1}c)^{1/2} \leqslant 64\alpha\sqrt{\vartheta},$$

so that (52) results in

$$\mathrm{O}(1)\alpha\vartheta^{3/2} \geqslant t_1^{i+1} - c^{\mathrm{T}}x^{i+1}.$$

This inequality, combined with (51), immediately implies upper bound on i required in (i).

Now let us prove (ii). From Theorem 10 it follows that if $N(i)$ is the number of iterations of the main phase preceding an iteration i of the phase, then

$$\gamma_i \equiv \min_{x \in \operatorname{int} D_{t^i}} F_{t^i}(x) \geqslant \gamma_{i^*} + \mathrm{O}(1)N(i)\sqrt{\frac{\Omega_H\Omega_K}{\Omega_H + \Omega_K}}. \tag{53}$$

On the other hand, let $\tilde{x} = x^*(t^{\mathrm{m}})$ be the analytic center of the domain $D_{t^{\mathrm{m}}}$. Given $\varepsilon \in (0,1)$, set

$$q = \frac{\varepsilon}{\varepsilon + \Theta}, \qquad x_q = (1 - q)x^* + q\tilde{x}.$$

Then $x_q \in \operatorname{int} D_{t^{\mathrm{m}}}$. For $\tau = \tau^* + \varepsilon(t_0^{\mathrm{m}} - \tau^*)$ we have

$$\begin{aligned} u_q &\equiv \tau B(x_q) - A(x_q) = y_q + \eta_q, \\ y_q &= (1-q)(\tau^* B(x^*) - A(x^*)) + q(t_0^{\mathrm{m}} B(\tilde{x}) - A(\tilde{x})), \\ \eta_q &= (1-q)(\tau - \tau^*) B(x^*) - q(t_0^{\mathrm{m}} - \tau) B(\tilde{x}). \end{aligned} \tag{54}$$

The vector y_q clearly belongs to H. By definition of Θ one has $\delta \equiv \Theta B(x^*) - B(\tilde{x}) \in K$, whence

$$\begin{aligned} \eta_q &= (1-q)(\tau - \tau^*) B(x^*) - q(t_0^{\mathrm{m}} - \tau) B(\tilde{x}) \\ &= [(1-q)(\tau - \tau^*) - q(t_0^{\mathrm{m}} - \tau)\Theta] B(x^*) + q(t_0^{\mathrm{m}} - \tau)\delta \\ &= pB(x^*) + p'\delta, \quad p > 0,\ p' \geqslant 0. \end{aligned}$$

Thus,

$$\eta_q \in \operatorname{int} K, \tag{55}$$

and since we already know that $y_q \in H$, (54) implies that $\tau B(x_q) - A(x_q) \in \operatorname{int} H$, so that $x_q \in D_{(\tau, t_1^{\mathrm{m}})}$.

Now let

$$F(x,y) = -\vartheta \ln(t_1^{\mathrm{m}} - c^{\mathrm{T}} x) + \Omega_H \Phi_H(y) + \Omega_K \Phi_K(\pi_K x + p_K) + \Phi_G(\pi_G x + p_G);$$

then F is a 4ϑ-s.-c.b. for the domain $\bar{D} = \{(x,y) \mid x \in D, y \in H, t_1^{\mathrm{m}} \geqslant c^{\mathrm{T}} x\}$. We have

$$\tilde{z} \equiv (\tilde{x}, t_0^{\mathrm{m}} B(\tilde{x}) - A(\tilde{x})) \in \operatorname{int} \bar{D}, \qquad z^* \equiv (x^*, \tau^* B(x^*) - A(x^*)) \in \bar{D}.$$

Let

$$z_q = (1-q) z^* + q\tilde{z} = (x_q, y_q).$$

From (55) it follows that $u_q - y_q \in K$, whence, in view of P.3, $\Phi_H(u_q) \leqslant \Phi_H(y_q)$ and, consequently,

$$F(x_q, u_q) \leqslant F(x_q, y_q) = F(z_q).$$

One clearly has $F_{(\tau, t_1^{\mathrm{m}})}(x_q) = F(x_q, u_q)$, whence

$$f^*(\tau, t_1^{\mathrm{m}}) \leqslant F(z_q);$$

on the other hand, z_q is a convex combination of a pair of points from $\bar{D}$ with the coefficients q and $1-q$, and from [16, Proposition 2.3.2(ii)], it follows that

$$F(z_q) \leqslant F(\tilde{z}) + 4\vartheta \ln(1/q) = F_{t^{\mathrm{m}}}(\tilde{x}) + 4\vartheta \ln(1/q) = f^*(t^{\mathrm{m}}) + 4\vartheta \ln(1/q).$$

Thus, we come to

$$f^*(\tau, t_1^{\mathrm{m}}) - f^*(t^{\mathrm{m}}) \leqslant 4\vartheta \ln(1/q).$$

Now, if i is such that at the iteration i belonging to the main phase one has $t_0^i \geqslant \tau = \tau^* + \varepsilon(t_0^{\mathrm{m}} - \tau^*)$, then, in view of the monotonicity of $f^*(\cdot, t_1)$ (see Theorem 10) and (53) the left-hand side of the latter inequality is $\geqslant O(1)N(i)\sqrt{\Omega_H\Omega_K(\Omega_H+\Omega_K)^{-1}}$, and we come to the inequality

$$N(i) \leqslant O(1)\vartheta\sqrt{\frac{\Omega_H+\Omega_K}{\Omega_H\Omega_K}}\ln(1/q) = O(1)\sqrt{\vartheta(\vartheta_H+\vartheta_K)}\ln(1/q);$$

since $1/q = 1+\Theta/\varepsilon$, we come to (43).

(iii) is an immediate consequence of $\mathcal{R}$, which, as we know, is ensured by our construction, and P.2. □

Appendix

Proof of Theorem 9. To simplify notation, in what follows we omit explicit arguments to the functions involved in the calculation; all quantities related to $\mathcal{F}$, Φ_H, Φ_K, Φ_G are taken at the points $U(t,x)$, $t_0B(x)-A(x)$, π_Kx+p_K, π_Gx+p_G, respectively; we also write B instead of $B(x)$ and u instead of $U(t,x)$. Similarly, all quantities related to $\mathcal{F}_*$ are taken at the point $s=\mathcal{F}'$.

(1) Let us start with the following observation:

Lemma A.1. *If (t,x) satisfies the premise of Theorem 9, then $\mu(t,x)>0$, and for all $z\in\mathbb{R}^n$, $v\in\mathbb{R}^m$ one has*

$$|z^{\mathrm{T}}\beta^{\mathrm{T}}\Phi'_H| \leqslant \mu(t,x)|z|_{\nabla^2F_K(x)},\quad F_K(x)=\Phi_K(\pi_Kx+p_K), \tag{A.1}$$

$$B|_{\Phi''_H} \leqslant \mu(t,x), \tag{A.2}$$

$$|\beta z|_{\Phi''_H} \leqslant 3\mu(t,x)|z|_{\nabla^2F_K(x)}, \tag{A.3}$$

$$|\beta^{\mathrm{T}}v|_{Q^{-1}(x)} \leqslant 3\frac{\mu(t,x)}{\sqrt{\Omega_K}}|v|_{[\Phi''_H]^{-1}}, \tag{A.4}$$

$$|\beta^{\mathrm{T}}\Phi'_H|_{Q^{-1}(x)} \leqslant \frac{\mu(t,x)}{\sqrt{\Omega_K}}. \tag{A.5}$$

Proof. Let us start with (A.1) and (A.2). Let z be such that $|z|_{\nabla^2F_K(x)}\leqslant 1$; then, due to the origin of F_K and P.1, $B(x+z)\in K$, so that the direction $B+\beta z$ is a recessive direction for the domain H; in view of P.3 it follows that

$$-(B+\beta z)^{\mathrm{T}}\Phi'_H \geqslant |B+\beta z|_{\Phi''_H}. \tag{A.6}$$

Thus, the affine form $-(B+\beta z)^{\mathrm{T}}\Phi'_H$ is nonnegative at the (centered at the origin) ellipsoid given by $|z|_{\nabla^2F_K(x)}\leqslant 1$, and the value of the form at the origin is nothing but $\mu=\mu(t,x)$, which immediately implies (A.1). From (A.6) it follows that $|B|_{\Phi''_H}\leqslant\mu$, as required in (A.2).

Since $x\in\operatorname{int}D\subset\operatorname{int}G_K$, we have $B\in\operatorname{int}K$, and, consequently, $B\neq 0$; since H does not contain lines, Φ''_H is positive definite (see P.1), and we conclude that $\mu>0$.

Now, under assumption $|z|_{\nabla^2 F_K(x)} \leqslant 1$, the left-hand side of (A.6), in view of (A.1), does not exceed 2μ, and from (A.6), (A.2) and the triangle inequality it follows that, for the indicated z, one has $|\beta z|_{\Phi''_H} \leqslant 3\mu$, which immediately results in (A.3).

It remains to prove (A.4) and (A.5). We have

$$|\beta^{\mathrm{T}} v|_{Q^{-1}(x)} = \max\{w^{\mathrm{T}}\beta^{\mathrm{T}} v \mid w^{\mathrm{T}} Q(x) w \leqslant 1\}; \tag{A.7}$$

since one clearly has $Q(x) \geqslant \Omega_K \nabla^2 F_K(x)$, from $w^{\mathrm{T}} Q(x) w \leqslant 1$ it follows that $|w|_{\nabla^2 F_K(x)} \leqslant \Omega_K^{-1/2}$, and consequently, in view of (A.3), $|\beta w|_{\Phi''_H} \leqslant 3\mu\Omega_K^{-1/2}$; thus,

$$w^{\mathrm{T}} Q(x) w \leqslant 1 \quad \Rightarrow \quad w^{\mathrm{T}}\beta^{\mathrm{T}} v = [\beta w]^{\mathrm{T}} v \leqslant |\beta w|_{\Phi''_H} |v|_{[\Phi''_H]^{-1}} \leqslant 3\mu\Omega_K^{-1/2} |v|_{[\Phi''_H]^{-1}},$$

and (A.4) follows. To prove (A.5), let us set in (A.7) $v = \Phi'_H$; according to (A.1), $|w^{\mathrm{T}}\beta^{\mathrm{T}}\Phi'_H| \leqslant \mu|w|_{\nabla^2 F_K(x)} \leqslant \mu\Omega_K^{-1/2}|w|_{Q(x)}$, and (A.5) follows from (A.7). Lemma A.1 is proved. □

(2) Now let us formulate the main estimates we need. In what follows δt is fixed and satisfies (30)–(31), and δx is such that $(\delta t, \delta x) \in \Pi(t, x)$.

Lemma A.2. *For (t, x) satisfying the premise of Theorem 9 one has*

$$|[t_0\xi + \pi]\delta x|_{\mathcal{F}''} = |\delta x|_{\nabla^2 F_t(x)} \leqslant 3\omega, \tag{A.8}$$

$$|[t_0\xi + \pi]\mathrm{d}_x(t,x)|_{\mathcal{F}''} = |\mathrm{d}_x(t,x)|_{\nabla^2 F_t(x)} \leqslant \kappa; \tag{A.9}$$

$$|\delta t_0 \beta \delta x|_{\Phi''_H} \leqslant 9\Omega_H^{-1/2}(\Omega_H + \Omega_K)^{-1/2}\omega^2, \tag{A.10}$$

$$|\delta t_0 \beta \mathrm{d}_x(t,x)|_{\Phi''_H} \leqslant 9\Omega_H^{-1/2}(\Omega_H + \Omega_K)^{-1/2}\omega\kappa, \tag{A.11}$$

$$|\Delta_s(r)|_{\mathcal{F}''_*} = |\Delta_u(r)|_{\mathcal{F}''} \leqslant 2\omega + 3(\kappa + \omega), \quad 0 \leqslant r \leqslant 1; \tag{A.12}$$

$$|\delta_u(r)|_{\mathcal{F}''} \leqslant 9(\Omega_H + \Omega_K)^{-1/2}\omega(\omega + \kappa), \quad 0 \leqslant r \leqslant 1; \tag{A.13}$$

$$|\delta_s(r)|_{\mathcal{F}''_*} \leqslant 3(\Omega_H + \Omega_K)^{-1/2}\omega(4\omega + \kappa), \quad 0 \leqslant r \leqslant 1. \tag{A.14}$$

Proof. To simplify notation, in what follows we write F instead of F_t and d instead of $(t_1 - c^{\mathrm{T}} x)^{-1}$.

(a) The equality in (A.8) is evident (see (11)). We have

$$\delta x = -[\nabla^2 F]^{-1}[-\vartheta d^2 \delta t_1 c + \Omega_H \beta^{\mathrm{T}} \Phi'_H \delta t_0 + \Omega_H (t_0\beta - \alpha)^{\mathrm{T}} \Phi''_H B \delta t_0],$$

whence

$$\begin{aligned}
|\delta x|^2_{\nabla^2 F} &= |[\nabla^2 F]\delta x|^2_{[\nabla^2 F]^{-1}} \\
&\leqslant 3\{\vartheta^2 d^4 \delta t_1^2 c^{\mathrm{T}} [\nabla^2 F]^{-1} c\}_{\mathrm{I}} \\
&\quad + 3\{\Omega_H^2 \delta t_0^2 (\Phi'_H)^{\mathrm{T}} \beta [\nabla^2 F]^{-1} \beta^{\mathrm{T}} \Phi'_H\}_{\mathrm{II}} \\
&\quad + 3\{\delta t_0^2 \Omega_H^2 B^{\mathrm{T}} \Phi''_H [t_0\beta - \alpha][\nabla^2 F]^{-1}[t_0\beta - \alpha]^{\mathrm{T}} \Phi''_H B\}_{\mathrm{III}}.
\end{aligned} \tag{A.15}$$

We clearly have $\nabla^2 F \geqslant \vartheta d^2 cc^{\mathrm{T}}$, which immediately implies that

$$\{\cdot\}_{\mathrm{I}} \leqslant \vartheta d^2 \delta t_1^2 \leqslant \omega^2. \tag{A.16}$$

(the concluding inequality is nothing but (30)).

We also evidently have $\nabla^2 F \geqslant Q(x)$, so that from (A.5) it follows that

$$\{\cdot\}_{\mathrm{II}} \leqslant \delta t_0^2 \Omega_H^2 \mu^2(t,x) \Omega_K^{-1} \leqslant \omega^2 \tag{A.17}$$

(the concluding inequality is an immediate consequence of (31)).

Finally, $\nabla^2 F \geqslant \Omega_H [t_0\beta - \alpha]^{\mathrm{T}} \Phi_H'' [t_0\beta - \alpha$, so that

$$\{\cdot\}_{\mathrm{III}} \leqslant \delta t_0^2 \Omega_H |B|^2_{\Phi_H''} \leqslant \delta t_0^2 \Omega_H \mu^2(t,x) \leqslant \omega^2 \tag{A.18}$$

(the second inequality follows from (A.2), the third from (31)).

Combining (A.15)–(A.18), we come to the inequality required in (A.8).

(b) The equality in (A.9) is evident (see (11)), and the inequality is given by (25).

(c) In view of (A.3) and $\nabla^2 F \geqslant \Omega_K \nabla^2 F_K$ we have

$$|\beta\delta x|_{\Phi_H''} \leqslant 3\mu(t,x)|\delta x|_{\nabla^2 F_K} \leqslant 3\mu(t,x)\Omega_K^{-1/2}|\delta x|_{\nabla^2 F} \leqslant 9\mu(t,x)\Omega_K^{-1/2}\omega$$

(the concluding inequality follows from (A.8)), which combined with (31) implies (A.10). (A.11) is given by similar reasoning, with (A.9) playing the role of (A.8).

(d) Equality in (A.12) is given by (24). To prove the inequality, note that in view of (15) one has

$$\begin{aligned}
|\Delta_u(r)|_{\mathcal{F}''} &\leqslant |r\sigma\delta t'|_{\mathcal{F}''} + |r\delta t_0[\xi x + p]|_{\mathcal{F}''} + |[t_0\xi + \pi]\Delta_x(r)|_{\mathcal{F}''} \\
&= \sqrt{\vartheta} d|r\delta t_1| + |r||\delta t_0|\Omega_H^{1/2}|B|_{\Phi_H''} + |[t_0\xi + \pi](\mathrm{d}_x(t,x) + r\delta x)|_{\mathcal{F}''} \\
&\qquad \text{(structure of } \mathcal{F}) \\
&\leqslant |r|\omega + |r||\delta t_0|\Omega_H^{1/2}\mu(t,x) + 3(\kappa + |r|\omega) \\
&\qquad \text{(by (30), (A.2), (A.8) and (A.9))} \\
&\leqslant 2\omega + 3(\kappa + \omega), \quad 0 \leqslant r \leqslant 1, \quad \text{(by (31))}
\end{aligned}$$

as claimed in (A.12).

(e) To prove (A.13), note that in view of (16) and (13) one has

$$\begin{aligned}
|\delta u(r)|_{\mathcal{F}''} &= \Omega_H^{1/2}|r\delta t_0\beta\Delta_x(r)|_{\Phi_H''} \\
&= \Omega_H^{1/2}|r\delta t_0\beta[\mathrm{d}_x(t,x) + r\delta x]|_{\Phi_H''} \\
&\leqslant 9(\Omega_H + \Omega_K)^{-1/2}\omega(\omega + \kappa), \quad 0 \leqslant r \leqslant 1, \quad \text{(by (A.10) and (A.11))}
\end{aligned}$$

as required in (A.13).

(f) It remains to prove (A.14). From (17)–(22) it follows that

$$\delta_s(r) = \mathcal{F}''\zeta, \qquad \zeta = (0; 0; \pi_K\varepsilon(r); \pi_G\varepsilon(r)),$$

and since $\mathcal{F}_*'' = [\mathcal{F}'']^{-1}$, we come to

$$\begin{aligned}
|\delta_s(r)|^2_{\mathcal{F}''_*} &= |\zeta|^2_{\mathcal{F}''} \\
&= \Omega_K[\pi_K\varepsilon(r)]^{\mathrm{T}}\Phi''_K\pi_K\varepsilon(r) + [\pi_G\varepsilon(r)]^{\mathrm{T}}\Phi''_G\pi_G\varepsilon(r) \\
&= \varepsilon^{\mathrm{T}}(r)[\Omega_K\pi_K^{\mathrm{T}}\Phi''_K\pi_K + \pi_G^{\mathrm{T}}\Phi''_G\pi_G]\varepsilon(r) \\
&= \varepsilon^{\mathrm{T}}(r)Q(x)\varepsilon(r) \quad \text{(origin of } Q(x)) \\
&= [Q(x)\varepsilon(r)]^{\mathrm{T}}Q^{-1}(x)[Q(x)\varepsilon(r)] \\
&= |r\delta t_0|^2|\xi^{\mathrm{T}}\Delta_s(r)|^2_{Q^{-1}(x)} \quad \text{(by (20))} \\
&= |r\delta t_0|^2|\beta^{\mathrm{T}}[\Omega_H\Phi''_H][r\delta t_0 B + [t_0\beta-\alpha](\mathrm{d}_x(t,x)+r\delta x)]|^2_{Q^{-1}(x)} \\
&\qquad \text{(by (17)–(22) and (13))} \\
&\leqslant 9\Omega_K^{-1}\mu^2(t,x)|r\delta t_0|^2 \\
&\qquad \times |\Omega_H\Phi''_H[r\delta t_0 B + [t_0\beta-\alpha](\mathrm{d}_x(t,x)+r\delta x)]|^2_{[\Phi''_H]^{-1}} \\
&\qquad \text{(by (A.4))} \\
&= 9\Omega_H^2\Omega_K^{-1}\mu^2(t,x)|r\delta t_0|^2|r\delta t_0 B + [t_0\beta-\alpha](\mathrm{d}_x(t,x)+r\delta x)|^2_{\Phi''_H},
\end{aligned}$$

whence

$$\begin{aligned}
|\delta_s(r)|_{\mathcal{F}''_*} &\leqslant 3\Omega_H\Omega_K^{-1/2}\mu(t,x)|r\delta t_0|[|r\delta t_0 B|_{\Phi''_H} + |[t_0\beta-\alpha](\mathrm{d}_x(t,x)+r\delta x)|_{\Phi''_H}] \\
&\leqslant 3\Omega_H\Omega_K^{-1/2}\mu(t,x)|r\delta t_0|[|r\delta t_0|\mu(t,x) + \Omega_H^{-1/2}|\mathrm{d}_x(t,x)+r\delta x|_{\nabla^2F_t}] \\
&\qquad \text{(by (A.2) and the evident relation} \\
&\qquad\quad \Omega_H^{1/2}|[t_0\beta-\alpha]h|_{\Phi''_H} \leqslant |h|_{\nabla^2F_t(x)}) \\
&\leqslant 3(\Omega_H+\Omega_K)^{-1/2}\omega(4\omega+\kappa), \\
&\qquad \text{(by (31), (A.8) and (A.9); note that } 0\leqslant r\leqslant 1)
\end{aligned}$$

as required in (A.14). Lemma A.2 is proved. □

(3) Let us fix $r\in[0,1]$ and set

$$u = U(t,x), \qquad u^+ = U(X(r)) = U(t+r\delta t, x+\mathrm{d}_x(t,x)+r\delta x),$$

$$s^+ = s(r), \qquad t^+ = t+r\delta t, \qquad x^+ = x+\mathrm{d}_x(t,x)+r\delta x.$$

From (A.12), (A.13) and (A.14) it follows that

$$\begin{aligned}
|u^+-u|_{\mathcal{F}''}, |s^+-s|_{\mathcal{F}''_*} \leqslant \zeta \equiv 2\omega + 3\max\{&(1+3\omega)(\kappa+\omega), \\
&\kappa+\omega+\omega(4\omega+\kappa)\} < 0.9 \qquad \text{(A.19)}
\end{aligned}$$

(the concluding relation follows from $\omega\leqslant 0.05$, $\kappa\leqslant 0.2$). In view of P.1 as applied to s.s.-c. functions $\mathcal{F}$ and $\mathcal{F}_*$, we conclude that $u^+\in\operatorname{Dom}\mathcal{F}$, and, consequently, $(t+r\delta t, x+\mathrm{d}_x(t,x)+r\delta x)\in T^+$, $s^+\in\operatorname{Dom}\mathcal{F}_*$, and, besides this,

$$\mathcal{F}(u^+) \leqslant \mathcal{F}(u) + s^{\mathrm{T}}[u^+-u] + \rho(\zeta),$$

$$\mathcal{F}_*(s^+) \leqslant \mathcal{F}_*(s) + [s^+-s]^{\mathrm{T}}u + \rho(\zeta), \quad \rho(\zeta) = -\ln(1-\zeta)-\zeta$$

(we have taken into account that $s = \mathcal{F}'(u)$, whence also $u = \mathcal{F}'_*(s)$).

Since $s = \mathcal{F}'(u)$, we have

$$\begin{aligned}\mathcal{F}(u^+) + \mathcal{F}_*(s^+) &\leqslant \mathcal{F}(u) + \mathcal{F}_*(s) + s^{\mathrm{T}}[u^+ - u] + [s^+ - s]^{\mathrm{T}}u + 2\rho(\zeta)\\ &= s^{\mathrm{T}}u + s^{\mathrm{T}}[u^+ - u] + [s^+ - s]^{\mathrm{T}}u + 2\rho(\zeta)\\ &= [s^+]^{\mathrm{T}}u^+ + 2\rho(\zeta) - [s^+ - s]^{\mathrm{T}}[u^+ - u]. \end{aligned}\tag{A.20}$$

Now, $u^+ = \sigma t_1^+ + t_0^+[\xi x^+ + p] + \pi x^+ + q$ and, as we know, $[t_0^+\xi + \pi]^{\mathrm{T}}s^+ = 0$, whence

$$[s^+]^{\mathrm{T}}u^+ = [s^+]^{\mathrm{T}}[\sigma t_1^+ + t_0^+ p + q],$$

and in view of (A.20) we come to

$$\begin{aligned}V(r) &\equiv F_{t^+}(x^+) + \mathcal{F}_*(s^+) - [s^+]^{\mathrm{T}}[\sigma(t^+)' + t_0^+ p + q]\\ &= \mathcal{F}(u^+) + \mathcal{F}_*(s^+) - [s^+]^{\mathrm{T}}u^+\\ &\leqslant 2\rho(\zeta) - [s^+ - s]^{\mathrm{T}}[u^+ - u].\end{aligned}$$

Thus, to prove that r passes the Acceptability test it suffices to demonstrate that the right-hand side of the latter inequality is $\leqslant \hat{\kappa}$. This is immediate:

$$2\rho(\zeta) - [s^+ - s]^{\mathrm{T}}[u^+ - u] \leqslant 2\rho(\zeta) + |s^+ - s|_{\mathcal{F}''_*}|u^+ - u|_{\mathcal{F}''}$$

(since $\mathcal{F}''_* = [\mathcal{F}'']^{-1}$), and the concluding quantity, in view of (A.19), is $\leqslant 2\rho(\zeta) + \zeta^2 \leqslant 6.02 \leqslant \hat{\kappa}$. □

Proof of Theorem 10. If $(\tau, y) \in T^+$, and $\tau^+ \geqslant \tau$ differs from τ only in the zero coordinate, then, as we know, $D_\tau \subset D_{\tau^+}$, and $B(y)$ is a recessive direction for H; in view of P.3 we have $B(y)^{\mathrm{T}}\Phi'_H(\tau_0 B(y) - A(y)) \leqslant 0$, whence $(\partial/\partial\tau_0)F_\tau(y) \leqslant 0$ and $F_{\tau^+}(y) \leqslant F_\tau(y)$. Since the latter inequality holds for all $y \in D_\tau \subset D_{\tau^+}$, we conclude that $f^*(\tau^+) \leqslant f(\tau)$, so that $f^*(\tau)$ is nonincreasing in τ_0.

Now let (t, x) satisfy $\mathcal{P}_\kappa$, let $x^* = x^*(t) \equiv \operatorname{argmin} F_t(\cdot)$ and let $\mathrm{d}t$ be the direction defined by $\mathrm{d}t_1 = 0$ and $\mathrm{d}t_0$ given by (32) *as applied to the pair* (t, x^*):

$$\mathrm{d}t_0 = -\omega\sqrt{\frac{\Omega_K}{\Omega_H^2 + \Omega_H\Omega_K}}\mu^{-1}(t, x^*).$$

Let

$$s^+(l) = (s_1(l); s_H(l); s_K(l); s_G(l)), \quad 0 \leqslant l \leqslant 1,$$

be the vectors associated with the pair (t, x^*), the direction $\mathrm{d}t$ and stepsize l by our Acceptability test. From Theorem 9 it follows that $t + l\,\mathrm{d}t \in T$ and that $s^+(l)$ is dual feasible with respect to $t_0 + l\,\mathrm{d}t_0$. In view of Lemma 6 we have (note that $\mathrm{d}t_1 = 0$)

$$f^*(t + l\,\mathrm{d}t) \geqslant [s^+(l)]^{\mathrm{T}}[\sigma t_1 + (t_0 + l\,\mathrm{d}t_0)p + q] - \mathcal{F}_*(s^+(l)). \tag{A.21}$$

As in the proof of Theorem 9 we have (we use the same short notation as in the indicated proof, but x, κ are now replaced by $x^*, 0$; below $s = \mathcal{F}'(U(t, x^*))$)

$$
\begin{aligned}
\mathcal{F}_*(s^+(l)) &\leqslant \mathcal{F}_*(s) + [s^+(l) - s]^{\mathrm{T}} U(t, x^*) + \rho(|s^+(l) - s|_{\mathcal{F}_*''}) \\
&= -F_t(x^*) + s^{\mathrm{T}} U(t, x^*) + [s^+(l) - s]^{\mathrm{T}} U(t, x^*) + \rho(|s^+(l) - s|_{\mathcal{F}_*''}) \\
&\leqslant -F_t(x^*) + [s^+(l)]^{\mathrm{T}} [\sigma t_1 + t_0(\xi x^* + p) + \pi x^* + q] \\
&\quad + \rho(5l\omega + 12 l^2 \omega^2) \\
&\qquad \text{(by (A.12) and (A.14) as applied with } \kappa = 0,\ r = 1,\ \delta t = l\,\mathrm{d}t, \\
&\qquad \text{which allows replacing } \omega \text{ by } l\omega) \\
&= -F_t(x^*) + [s^+(l)]^{\mathrm{T}} [\sigma t_1 + (t_0 + l\,\mathrm{d}t_0)(\xi x^* + p) + \pi x^* + q] \\
&\quad - l\,\mathrm{d}t_0 [s^+(l)]^{\mathrm{T}} (\xi x^* + p) + \rho(6l\omega) \\
&\qquad \text{(since } 5l\omega + 12 l^2 \omega^2 \leqslant 6l\omega \text{ due to } 0 \leqslant l \leqslant 1 \text{ and } 0 < \omega \leqslant 0.05) \\
&= -F_t(x^*) + [s^+(l)]^{\mathrm{T}} [\sigma t_1 + (t_0 + l\,\mathrm{d}t_0) p + q] \\
&\quad - l\,\mathrm{d}t_0 [s^+(l)]^{\mathrm{T}} (\xi x^* + p) + \rho(6l\omega). \\
&\qquad \text{(since } s^+(l) \text{ is dual feasible w.r.t. } t_0 + l\,\mathrm{d}t_0).
\end{aligned}
$$

Combining this inequality, (A.21) and taking into account that $F_t(x^*) = f^*(t)$, we come to

$$
\begin{aligned}
f^*(t + l\,\mathrm{d}t) &\geqslant f^*(t) - \rho(6l\omega) + l\,\mathrm{d}t_0 [s^+(l)]^{\mathrm{T}} (\xi x^* + p) \\
&= f^*(t) - \rho(6l\omega) + l\,\mathrm{d}t_0 s_H^{\mathrm{T}}(l) B(x^*).
\end{aligned} \tag{A.22}
$$

Furthermore, from relations defining $s^+(l)$ (see (15), (17)–(22)) it follows that

$$
s_H(l) = \Omega_H \Phi_H' + (\Delta_s(l))_H \tag{A.23}
$$

and, as we know from (A.12) (applied to $r = 1$, $\kappa = 0$, $\delta t = l\,\mathrm{d}t$, which allows to replace ω by $l\omega$),

$$
|(\Delta_s(l))_H|_{[\Omega_H \Phi_H'']^{-1}} \leqslant |\Delta_s(l)|_{\mathcal{F}_*''} \leqslant 5l\omega. \tag{A.24}
$$

Combining (A.22) and (A.23), we come to

$$
\begin{aligned}
f^*(t + l\,\mathrm{d}t) - f^*(t) &\geqslant -\rho(6l\omega) + l\,\mathrm{d}t_0 \Omega_H B^{\mathrm{T}}(x^*) \Phi_H' \\
&\qquad - l|\mathrm{d}t_0| |(\Delta_s(l))_H|_{[\Omega_H \Phi_H'']^{-1}} |B(x^*)|_{\Omega_H \Phi_H''} \\
&\geqslant -\rho(6l\omega) + l|\mathrm{d}t_0| [\Omega_H \mu(t, x^*) - 5l\omega \Omega_H^{1/2} \mu(t, x^*)] \\
&\qquad \text{(origin of } \mu(t, x^*),\ \text{(A.24) and (A.2))} \\
&= -\rho(6l\omega) + l|\mathrm{d}t_0| \mu(t, x^*) \Omega_H (1 - 5l\omega) \\
&\geqslant -\rho(6l\omega) + \frac{3}{4} l|\mathrm{d}t_0| \mu(t, x^*) \Omega_H \\
&= -\rho(6l\omega) + \frac{3}{4} l\omega \sqrt{\frac{\Omega_H \Omega_K}{\Omega_H + \Omega_K}} \quad \text{(by (32))}.
\end{aligned}
$$

Thus, we have

$$f^*(t+l\,dt)-f^*(t)\geqslant \frac{3}{4}l\omega\sqrt{\frac{\Omega_H\Omega_K}{\Omega_H+\Omega_K}}-\rho(6l\omega),\quad 0\leqslant l\leqslant 1. \tag{A.25}$$

Since $\rho(u)=-\ln(1-u)-u$ and $\omega\leqslant 1/20$, for $0\leqslant l\leqslant 1$ we have

$$\rho(6l\omega)\leqslant\frac{2}{3}(6l\omega)^2=24l^2\omega^2,$$

and (A.25) implies that

$$f^*(t+l\,dt)-f^*(t)\geqslant \frac{1}{4}l\omega\sqrt{\frac{\Omega_H\Omega_K}{\Omega_H+\Omega_K}},\quad 0\leqslant l\leqslant \frac{1}{4}. \tag{A.26}$$

Since (t,x) satisfies $\mathcal{P}_K$, we have (see (36))

$$|x^*-x|_{\nabla^2F_t(x)}\leqslant 0.27; \tag{A.27}$$

consequently,

$$|x^*-x|_{\nabla^2F_K(x)}\leqslant \Omega_K^{-1/2}\theta,\qquad F_K(y)=\Phi_K(\pi_Ky+p_K), \tag{A.28}$$

whence, in view of the origin of F_K and P.1,

$$B(x\pm\Omega_K^{1/2}\theta^{-1}[x^*-x])\in K,$$

and, consequently,

$$B(x^*)-\frac{\sqrt{\Omega_K}-\theta}{\sqrt{\Omega_K}}B(x)\equiv B(x^*)-\nu B(x)\in K. \tag{A.29}$$

Since $[-\Phi_H'(t_0B(x^*)-A(x^*))]^{\mathrm T}u$ is nonnegative for all recessive directions u of H (see P.3) and, in particular, for all directions from the cone K, we conclude from (A.29) that

$$\begin{aligned}\mu(t,x^*)&=-B^{\mathrm T}(x^*)\Phi_H'(t_0B(x^*)-A(x^*))\\&\geqslant-\nu B^{\mathrm T}(x)\Phi_H'(t_0B(x^*)-A(x^*)).\end{aligned} \tag{A.30}$$

Now let $u=t_0B(x)-A(x)$, $u^*=t_0B(x^*)-A(x^*)$ and $d=u^*-u$; from (A.27) it immediately follows that

$$|d|_{\Phi_H''(u)}\leqslant\Omega_H^{-1/2}\theta\leqslant 0.27\Omega_H^{-1/2},$$

whence, in view of P.1,

$$\Phi_H''(u+\tau d)\leqslant(0.73)^{-2}\Phi_H''(u),\quad 0\leqslant\tau\leqslant 1,$$

so that

$$|\Phi_H'(u^*)-\Phi_H'(u)|_{[\Phi_H''(u)]^{-1}}\leqslant 2|d|_{\Phi_H''(u)}\leqslant 2\Omega_H^{-1/2}\theta.$$

Consequently,

$$|B^{\mathrm{T}}(x)(\Phi'_H(u) - \Phi'_H(u^*))| \leqslant |B(x)|_{\Phi''_H(u)}|\Phi'_H(u) - \Phi'_H(u^*)|_{[\Phi''_H(u)]^{-1}}$$
$$\leqslant 2\mu(t,x)\Omega_H^{-1/2}\theta$$

(we have used (A.2)), so that

$$-B^{\mathrm{T}}(x)\Phi'_H(u^*) \geqslant \mu(t,x)(1 - 2\Omega_H^{-1/2}\theta) \geqslant 0.46\mu(t,x).$$

From the latter inequality and (A.30) we conclude that

$$\mu(t,x^*) \geqslant 0.46\nu\mu(t,x) \geqslant 0.3\mu(t,x).$$

It follows that $|\mathrm{d}t_0| \leqslant |4\Delta t_0|$, so that $t + l\,\mathrm{d}t \geqslant t^+$ whenever $0 \leqslant l \leqslant 1/4$; thus, (A.26), in view of already proved monotonicity of f^*, implies the desired relation

$$f^*(t^+) - f(t) \geqslant f^*(t + \frac{1}{4}\mathrm{d}t) - f^*(t) \geqslant \frac{\omega}{16}\sqrt{\frac{\Omega_H\Omega_K}{\Omega_H + \Omega_K}}. \qquad \square$$

References

[1] K.M. Anstreicher, A monotone projective algorithm for fractional linear programming, *Algorithmica* 1 (1986) 483–498.

[2] C.R. Bector and A. Cambini, Fractional programming – Some recent results, in: *Generalized Convexity and Fractional Programming with Economic Applications*, Pisa, 1988, Lecture Notes in Economics and Mathematical Systems, Vol. 345 (Springer, Berlin, 1989) 493–504.

[3] Y. Benadada, J.P. Crouzeix and J. Ferland, An interval-type algorithm for generalized fractional programming, in: *Generalized Convexity and Fractional Programming with Economic Applications*, Pisa, 1988, Lecture Notes in Economics and Mathematical Systems, Vol. 345 (Springer, Berlin, 1989) 106–120.

[4] S. Boyd and L. El Ghaoui, Method of centers for minimizing generalized eigenvalue, *Linear Algebra and Applications* (Special Issue on Numerical Linear Algebra Methods in Control, Signals and Systems) 188 (1993) 63–111.

[5] S. Boyd, L. El Ghaoui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory* (SIAM, Philadelphia, 1994).

[6] W. Dinkelbach, On nonlinear fractional programming, *Management Science* 13 (1967) 492–498

[7] J.A. Ferland and J.Y. Potvin, Generalized fractional programming: Algorithms and numerical experimentation, *European Journal of Operational Research* 20 (1985) 92–101.

[8] R.W. Freund and F. Jarre, An interior-point method for convex fractional programming, AT&T Numerical Analysis Manuscript No. 93-03, Bell Laboratories, Murray Hill, NJ, 1993.

[9] R.W. Freund and F. Jarre, An interior-point method for multi-fractional programs with convex constraints, *Mathematical Programming* 67 (1994) 407–440.

[10] J.-P.A. Haeberly and M.L. Overton, A hybrid algorithm for optimizing eigenvalues of symmetric definite pencils, *SIAM J. Matrix Anal. Appl.* 15 (1994) 1141–1156.

[11] J.-P.A. Haeberly and M.L. Overton, Optimizing eigenvalues of symmetric definite pencils, in: *Proceedings of the 1994 American Control Conference*, Baltimore, MD, 1994.

[12] N. Karmarkar, A new polynomial-time algorithm for linear programming, *Combinatorica* 4 (1984) 373–395.

[13] A. Nemirovski, On polynomiality of the method of analytic centers for fractional problems, *Mathematical Programming* 73 (1996) 175–198.

[14] Yu. Nesterov, Long-step strategies in interior point potential reduction methods, Preprint, Departement d'Economie Commerciale et Industrielle, Université de Geneve, Geneve, 1993.

[15] Yu. Nesterov and A. Nemirovski, Multi-parameter surfaces of analytic centers and long-step path-following interior point methods, Research Report 2/94, Technion – Israel Institute of Technology, Haifa, 1993; also: *Mathematics of Operations Research*, to appear.

[16] Yu. Nesterov and A. Nemirovskii, *Interior Point Polynomial Algorithms in Convex Programming: Theory and Applications* (SIAM, Philadelphia, 1994)

[17] Yu. Nesterov and A. Nemirovskii, An interior point method for generalized linear-fractional problems, *Mathematical Programming* 69 (1995) 117–204.

[18] J. Renegar, A polynomial time algorithm, based on Newton's method, for linear programming, *Mathematical Programming* 40 (1988) 59–93.

[19] S. Schaible, Bibliography in fractional programming, *Zeitschrift für Operations Research* 27 (1983) 211–241.

[20] S. Schaible, Multi-ratio fractional programming - A survey, in: *Optimization, Parallel Processing and Applications*, Lecture Notes in Economics and Mathematical Systems, Vol. 304 (Springer, Berlin, 1988) 57–66.

[21] G. Sonnevend, An "analytical centre" for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming, in: *System Modeling and Optimization*, Budapest, Lecture Notes in Control and Information Sciences, Vol. 84 (Springer, Berlin, 1985) 866–875.

[22] J. von Neumann, A model of general economic equilibrium, *Review of Economic Studies* 13 (1945) 1–9.

[23] Y. Ye, On the von Neumann economic growth problem, Working Paper Series No. 92-9, March 1992, Department of Management Sciences, The University of Iowa, Iowa City, IA.