Found Comput Math (2008) 8: 533–560
DOI 10.1007/s10208-007-9019-4

FOUNDATIONS OF COMPUTATIONAL MATHEMATICS
The Journal of the Society for the Foundations of Computational Mathematics

Primal Central Paths and Riemannian Distances for Convex Sets

Y. Nesterov · A. Nemirovski

Received: 2 November 2005 / Revised: 29 May 2007 / Accepted: 5 September 2007 /
Published online: 15 December 2007

Abstract In this paper, we study the Riemannian length of the primal central path in a convex set computed with respect to the local metric defined by a self-concordant function. Despite some negative examples, in many important situations, the length of this path is quite close to the length of a geodesic curve. We show that in the case of a bounded convex set endowed with a ν-self-concordant barrier, the length of the central path is within a factor $O(\nu^{1/4})$ of the length of the shortest geodesic curve.

Keywords Riemannian geometry · Convex optimization · Structural optimization · Interior-point methods · Path-following methods · Self-concordant functions · Polynomial-time methods

Mathematics Subject Classification (2000) 52A41 · 53C22 · 68Q25 · 90C22 · 90C25 · 90C51 · 90C60

Communicated by Mike Todd.

This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming.

Y. Nesterov (✉)
CORE, Catholic University of Louvain, 34 voie du Roman Pays, 1348 Louvain-la-Neuve, Belgium
e-mail: nesterov@core.ucl.ac.be

A. Nemirovski
Technion-Israel Institute of Technology, Haifa, Israel
e-mail: nemirovs@ie.technion.ac.il

1 Introduction

Most modern interior-point schemes for convex optimization problems are based on tracing a path in the interior of a convex set. A generic case of this type is as follows. We are given a self-concordant function $f(x)$ with convex open domain $\operatorname{dom} f$ in a finite-dimensional space E (see Sect. 2 to recall the definitions). Let $\mathcal{A}$ be a linear operator from E to another linear space E_d and $d \in E_d$. We are interested in tracing the path

$$x(t) = \operatorname*{argmin}_{\substack{x \in \operatorname{dom} f: \\ \mathcal{A}x = d}} \left\{ f_t(x) \stackrel{\text{def}}{=} f(x) + t\langle e, x\rangle \right\}. \tag{1.1}$$

Note that this framework covers seemingly all standard short-step interior-point path-following schemes (e.g., [5, 7], and [10]):

- The *primal path-following method* solving the optimization program

  $$\min_x \left\{ \langle c, x\rangle : x \in \operatorname{cl} \operatorname{dom} f \right\}$$

  traces, as $t \to \infty$, the *primal central path* (1.1) given by $e = c$, $\mathcal{A} \equiv 0$, $b = 0$, and a ν-self-concordant barrier f for $\operatorname{cl} \operatorname{dom} f$.
- Let $\mathcal{K}$ be a closed pointed cone with nonempty interior in a finite-dimensional linear space G, A be a linear operator from G to R^m with conjugate operator A^*, G^* be a linear space dual to G, and $\mathcal{K}_* \subset G^*$ be a cone dual to $\mathcal{K}$. And let F be a ν-logarithmically homogeneous self-concordant barrier F for $\mathcal{K}$ with Legendre transform F_*. The *feasible-start primal-dual path-following method* for solving the primal-dual pair of conic problems

  $$(P): \quad \min_u \left\{ \langle c, u\rangle : Au = b,\ u \in \mathcal{K} \right\},$$

  $$(D): \quad \max_y \left\{ \langle b, y\rangle : c - A^* y \in \mathcal{K}_* \right\},$$

  traces, as $t \to \infty$, the *primal-dual central path* (1.1) given by

  $$x \equiv (u, y) \in E = G \times R^m, \qquad f(x) = F(u) + F_*(c - A^* y),$$

  $$\mathcal{A}(x) = Au, \qquad d = b, \qquad e = (c, -b).$$

- For the *infeasible-start primal-dual path-following method* as applied to (P), (D), we need to fix a reference points $u_0 \in \operatorname{int} \mathcal{K}$ and $s_0 \in \operatorname{int} \mathcal{K}_*$. Then we can trace, as $t \to 0$, the path (1.1) given by $E = (G \times R) \times R^{m+1}$,

  $$x \equiv (u, \tau, y, \kappa),$$

  $$\mathcal{A}x = \big(Au - \tau b, \langle c, u\rangle - \langle b, y\rangle + \kappa\big),$$

  $$d = \big(Ax_0 - b, \langle c, x_0\rangle + 1\big),$$

  $$f(x) = F(u) + F_*\big(s_0 + (\tau - 1)c - A^* y\big) - \ln \tau - \ln \kappa,$$

with e defined by the condition that starting point $x_0 = (u_0, 1, 0, 1)$ coincides with $x(1)$ (see [7] for details).

It can be shown that in all situations an *upper* complexity bound for tracing the corresponding trajectories by a short-step strategy is of the order $O(\sqrt{\nu} \ln \frac{\nu}{\epsilon})$ iterations, where ϵ is relative accuracy of the final approximate solution of the problem (see, for example, [4, 5], and [7]).

Recently in [6], a technique has been developed for establishing *lower* complexity bounds for short-step path-following schemes. This technique is based on a combination of the theory of self-concordant functions [5] with the concepts of *Riemannian geometry* (see [1–3, 8, 9] for the main definitions and their applications in optimization).

Let f be a nondegenerate self-concordant function on an open convex domain $Q \subseteq E$. We can use the Hessian of f in order to define a Riemannian structure on Q. Namely, for a continuously differentiable curve $\gamma(t) \in Q$, $t_0 \le t \le t_1$, we define the length of this curve as follows:

$$\rho\big[\gamma(\cdot), t_0, t_1\big] = \int_{t_0}^{t_1} \big\langle f''(\gamma(t))\gamma'(t), \gamma'(t)\big\rangle^{1/2} dt.$$

Definition 1.1 The infimum of the quantities $\rho[\gamma(\cdot), t_0, t_1]$ over all continuously differentiable curves $\gamma(\cdot)$ in Q linking a given pair of points x and y of the set (i.e., $\gamma(t_0) = x, \gamma(t_1) = y$) is called the *Riemannian distance* between x and y. We denote it by $\sigma(x, y)$.

Clearly, $\sigma(x, y)$ is a distance on Q. For $\widehat{Q} \subset Q$, we define $\sigma(x, \widehat{Q}) = \inf_y\{\sigma(x, y): y \in \widehat{Q}\}$.

Note that the Riemannian metric defined by $f''(x)$ arises very naturally in the theory of interior-point schemes. Indeed, it can be shown that the *Dikin ellipsoid*

$$W_r(x) = \big\{y : \big\langle f''(x)(y - x), y - x\big\rangle \le r^2\big\}$$

centered at an arbitrary point $x \in \operatorname{int} Q$ always belongs to Q for $r < 1$. Moreover, *short-step* interior-point methods usually generate a sequence of points $\{x_k\}_{k=0}^{\infty}$ satisfying the condition

$$\big\langle f''(x_k)(x_{k+1} - x_k), x_{k+1} - x_k\big\rangle \le r^2 < 1, \quad \forall k \ge 0.$$

Hence, given a starting point $\bar{x}$, a target x_*, both in $\operatorname{int} Q$, and a step-size bound $r \in (0, 1)$, it would be interesting to estimate from below and from above the *minimal number of steps* $N_f(\bar{x}, x_*)$, which is sufficient for connecting the end points by a short-step sequence. If x_* is an approximate solution of certain optimization problem, then an *upper bound* for $N_f(\bar{x}, x_*)$ can be extracted from the complexity estimate of some numerical scheme. Thus, a *lower bound* for $N_f(\bar{x}, x_*)$ delivers a lower complexity bound for corresponding methods.

In [6], it was shown that the value $O(\sigma(x, y))$ provides us with a lower bound for $N_f(x, y)$ (see [6], Sect. 3). Moreover, in [6], Sect. 5, it was shown that the upper bounds for this value delivered by primal-dual feasible and infeasible path-following schemes coincide with the lower bounds up to a constant factor. Thus, for the corresponding problem setting, these methods are *optimal*.

Note that the conclusion of [6] was derived from some special symmetry of the primal–dual feasible cone $\mathcal{K} \times \mathcal{K}_*$. This technique is not applicable to the pure primal setting. Actually, it is easy to see that in this case the central path trajectories can be very bad. Let us look at the following example.

Example 1.1 Let $Q \equiv R^n_+$ be the positive orthant in R^n. Let us endow Q with the standard self-concordant barrier

$$f(x) = -\sum_{i=1}^n \ln x^{(i)}, \qquad \nu = n.$$

Then the Riemannian distance in Q is defined as follows (see (6.20), [6]):

$$\sigma(x, y) = \left[\sum_{i=1}^n \ln^2 \frac{x^{(i)}}{y^{(i)}} \right]^{1/2}.$$

Let us form a central path, which connects the point $x_0 = e = (1, \dots, 1)^T \in R^n$ with the simplex

$$\Delta(\beta) = \left\{ x \in Q : \sum_{i=1}^n x^{(i)} = n \cdot (1 + \beta) \right\}, \qquad \beta > 0.$$

That is a solution of the following problem

$$x(t) = \arg \min_{x \in \Delta(t)} f(x), \qquad 0 \le t \le \beta.$$

Clearly, $x(t) = (1 + t) \cdot e$. Thus, using the central path, we can travel from $x_0 = e$ to the set $\Delta(\beta)$ in $O(\sigma(e, x(\beta)))$ iterations of a path-following scheme with

$$\sigma\big(e, x(\beta)\big) = \sqrt{n} \ln(1 + \beta).$$

However, it is easy to see that there exists a shortcut:

$$y = e + n\beta e_1 \in \Delta(\beta), \qquad \sigma(e, y) = \ln(1 + n\beta),$$

where $e_1 = (1, 0, \dots, 0)^T \in R^n$.

Despite referring to a slightly different problem setting,[1] the above observation is quite discouraging. Fortunately, the situation is not always so bad. The main goal of this paper is to show that for several important problems the primal central paths are $O(\nu^{1/4})$-geodesic (we use terminology of [6]).

[1] In Example 1.1, we speak about Riemannian distance between a point and a *convex set*.

The paper is organized as follows. In Sect. 2, we recall for the benefit of the reader the main facts on the theory of self-concordant functions and prove several new inequalities, which are necessary for our analysis. In Sect. 3, we establish different lower bounds on the Riemannian distances in convex sets in terms of the local norms defined by a self-concordant function. In Sect. 4, we prove the main result of this paper. That is an upper bound on the Riemannian length of a segment of a central path in terms of variation of the value of corresponding self-concordant function and the logarithm of the variation of the path parameter. In Sect. 5, we apply this result to different problem instances: finding a minimum of self-concordant function (Sect. 5.1), feasibility problem (Sect. 5.2) and the standard minimization problem (Sect. 5.3). We show that in Example 1.1 the presence of the unpleasant factor $O(\sqrt{\nu}/\ln\nu)$ in the ratio of the length of the central path and the corresponding Riemannian distance is due to unboundedness of the basic feasible set Q. If Q is bounded, this ratio can be at most of the order $O(\nu^{1/4})$.

2 Self-Concordant Functions and Barriers

In order to make the paper self-contained, in this section, we summarize the main results on self-concordant functions, which can be found in [5], Chap. 2, and in [4], Chap. 4. We prove also some new inequalities, which are useful for working with the Riemannian distances.

Let E be a finite-dimensional real vector space and $Q \subset E$ be an open convex domain. A three-times continuously differentiable convex function

$$f(x): \quad Q \to R$$

is called *self-concordant*, if the sets $\{x \in Q : f(x) \le a\}$ are closed for every $a \in R$ and

$$\frac{d^3}{dt^3} f(x+th)\bigg|_{t=0} \le 2\left(\frac{d^2}{dt^2} f(x+th)\bigg|_{t=0}\right)^{3/2}, \quad \forall x \in Q,\ h \in E. \tag{2.1}$$

Such a function is called *nondegenerate*, if its Hessian is positive definite at some (and then—at every) point of Q. This happens, for example, if Q contains no straight line. An equivalent condition to (2.1) can be written in terms of relation between the second and third differentials:

$$D^3 f(x)[h_1,h_2,h_3] \le 2\prod_{i=1}^{3}\langle f''(x)h_i,h_i\rangle^{1/2}, \quad \forall x \in Q,\ h_1,h_2,h_3 \in E. \tag{2.2}$$

Denote by E^* the space *dual* to E. For $h \in E$ and $g \in E^*$ denote by $\langle g,h\rangle$ the value of the linear function g on the vector h. Let f be a nondegenerate self-concordant function on Q. For every $x \in Q$, we have $f'(x) \in E^*$. Thus, the Hessian $f''(x)$ defines a nondegenerate linear operator:

$$h \mapsto f''(x)h \in E^*, \quad \forall h \in E.$$

Hence, we can define a local *primal* norm:

$$\|h\|_x = \langle f''(x)h, h\rangle^{1/2} : E \to R,$$

and, using the standard definition $\|\eta\|_x^* = \max_{h:\|h\|_x \le 1}\langle \eta, h\rangle$, the corresponding *local dual* norm:

$$\|g\|_x^* = \langle g, [f''(x)]^{-1} g\rangle^{1/2}: E^* \to R.$$

Denote by

$$\lambda(x) = \langle f'(x), [f''(x)]^{-1} f'(x)\rangle^{1/2} = \| f'(x)\|_x^*$$

the local norm of the gradient $f'(x)$. A nondegenerate self-concordant function f is called a *self-concordant barrier* with parameter ν, if

$$\lambda^2(x) \equiv \langle f'(x), [f''(x)]^{-1} f'(x)\rangle \le \nu, \quad \forall x \in Q.$$

Note that ν cannot be smaller than one.

Let us mention first the well-known facts.

Proposition 2.1 *Let f be a nondegenerate self-concordant function on an open convex domain $Q \subset E$. Then*

(i) *For every $x \in Q$ and $r \in [0, 1)$, the ellipsoid $W_r(x) \equiv \{y : \|y - x\|_x < r\}$ is contained in Q. For any $y \in W_r(x)$ and any $h \in E$, we have*

$$\big(1 - \|y - x\|_x\big)\|h\|_x \le \|h\|_y \le \frac{\|h\|_x}{1 - \|y - x\|_x}. \tag{2.3}$$

Moreover, for any x and y from Q

$$\|y - x\|_y \ge \frac{\|y - x\|_x}{1 + \|y - x\|_x}, \tag{2.4}$$

and

$$\langle f'(x) - f'(y), x - y\rangle \ge \frac{\|x - y\|_x^2}{1 + \|x - y\|_x}. \tag{2.5}$$

(ii) *The following facts are related to existence of a minimizer x_f of $f(x)$ on Q.*

- *f attains its minimum on Q if and only if it is below bounded.*
- *f attains its minimum on Q if and only if the set $\{x : \lambda(x) < 1\}$ is nonempty.*
- *If $\lambda(x) < 1$, then $f(x) - f(x_f) \le -\lambda(x) - \ln(1 - \lambda(x))$.*
- *If x_f exists, then it is unique.*
- *For every $\rho < 1$, the set $\{x \in Q : \lambda(x) \le \rho\}$ is compact.*
- *Denote $r_f(x) = \|x - x_f\|_{x_f}$. If $\lambda(x) < 1$, then*

$$r_f(x) - \ln\big(1 + r_f(x)\big) \le -\lambda(x) - \ln\big(1 - \lambda(x)\big).$$

Hence, the set $\{x : \lambda(x) \le \frac{1}{2}\}$ is contained in the ellipsoid $\{x : \|x - x_f\|_{x_f} \le \frac{3}{4}\}$.

(iii) *For every $x \in Q$, the damped Newton iterate*

$$x_+ = x - \frac{1}{1+\lambda(x)}\big[f''(x)\big]^{-1} f'(x)$$

belongs to Q, and

$$\begin{aligned} &f(x_+) \le f(x) - \big[\lambda(x) - \ln\big(1+\lambda(x)\big)\big], \\ &\lambda(x_+) \le 2\lambda^2(x). \end{aligned} \tag{2.6}$$

Hence, the damped Newton method converges quadratically as $\lambda(x) < \frac{1}{2}$.

(iv) *The domain of the Legendre transformation*

$$f_*(\xi) = \sup_x \big[\langle \xi, x\rangle - f(x)\big]: \quad E^* \to R \cup \{+\infty\}$$

of f is an open convex set which is exactly the image of Q under the one-to-one C^2 mapping

$$f'(x): \quad Q \to E^*.$$

The function f_* is a nondegenerate self-concordant function on its domain, and the Legendre transformation of f_* is f. If x_f exists, then by (2.3) as applied to f_*, we have

$$\big(1-\lambda(x)\big)\|g\|_x^* \le \|g\|_{x_f}^* \le \frac{\|g\|_x^*}{1-\lambda(x)}, \quad \forall x \in Q : \lambda(x) < 1, \forall g \in E^*. \tag{2.7}$$

(v) *Let f be ν-self-concordant barrier for* cl Q. *Let x and y belong to Q. Then*

$$\big\langle f'(x), y - x\big\rangle \le \nu.$$

If in addition, $\langle f'(x), y-x\rangle \ge 0$, then $\|y-x\|_x \le \nu + 2\sqrt{\nu}$.

Let us prove now some new inequalities. Let f be a nondegenerate self-concordant function with dom $f \subseteq E$. Denote

$$\zeta(t) = \ln(1+t) - \frac{t}{1+t}, \quad t > -1,$$

$$\zeta_*(t) = \ln(1-t) + \frac{t}{1-t}, \quad t < 1.$$

In what follows, we assume that these functions are equal $+\infty$ outside their natural domains.

Lemma 2.1 *For every $x, y \in$ dom f we have:*

$$f(y) \ge f(x) + \big\langle f'(x), y-x\big\rangle + \zeta\big(\|y-x\|_y\big), \tag{2.8}$$

$$f(y) \le f(x) + \big\langle f'(x), y-x\big\rangle + \zeta_*\big(\|y-x\|_y\big), \tag{2.9}$$

$$f(x) \ge f(y) + \langle f'(y), x-y\rangle + \zeta\bigl(\|f'(x)-f'(y)\|_y^*\bigr), \tag{2.10}$$

$$f(x) \le f(y) + \langle f'(y), x-y\rangle + \zeta_*\bigl(\|f'(x)-f'(y)\|_y^*\bigr). \tag{2.11}$$

Proof For $t \in [0, 1]$ consider the function

$$\phi(t) = f(y) - f\bigl(y + t(x-y)\bigr) + \bigl\langle f'\bigl(y+t(x-y)\bigr), t(x-y)\bigr\rangle.$$

Note that $\phi(0) = 0$. Then

$$\phi'(t) = t\bigl\langle f''\bigl(y+t(x-y)\bigr)(x-y), x-y\bigr\rangle.$$

Denote $r = \|x-y\|_y$. Since f is self-concordant, we have

$$\bigl\langle f''\bigl(y+t(x-y)\bigr)(x-y), x-y\bigr\rangle \ge \frac{r^2}{(1+tr)^2}$$

(see (2.4)). Hence,

$$f(y) - f(x) + \langle f'(x), x-y\rangle = \phi(1) - \phi(0) = \int_0^1 \phi'(t)\,dt \ge \int_0^1 \frac{tr^2\,dt}{(1+tr)^2} = \zeta(r),$$

as required in (2.8). If $r < 1$, we have also

$$\bigl\langle f''\bigl(y+t(x-y)\bigr)(x-y), x-y\bigr\rangle \le \frac{r^2}{(1-tr)^2}$$

(see (2.3)). Hence,

$$f(y) - f(x) + \langle f'(x), x-y\rangle = \phi(1) - \phi(0) = \int_0^1 \phi'(t)\,dt \le \int_0^1 \frac{tr^2\,dt}{(1-tr)^2} = \zeta_*(r),$$

and that is (2.9).

In order to prove two others inequalities, note that the Legendre transformation of f

$$f_*(s) = \max_x\bigl[\langle s, x\rangle - f(x)\bigr]$$

is nondegenerate and self-concordant along with f (Proposition 2.1(iv)). Therefore, using (2.8) and (2.9) at the points

$$u = f'(x), \qquad v = f'(y)$$

we have

$$f_*(v) - f_*(u) - \bigl\langle f_*'(u), v-u\bigr\rangle \ge \zeta\bigl(\|v-u\|_v\bigr),$$

$$f_*(v) - f_*(u) - \bigl\langle f_*'(u), v-u\bigr\rangle \le \zeta_*\bigl(\|v-u\|_v\bigr).$$

It remains to note that

$$f_*(u) = \langle f'(x), x \rangle - f(x), \quad f'_*(u) = x,$$

$$f_*(v) = \langle f'(y), y \rangle - f(y), \quad f'_*(v) = y,$$

and that $\|h\|_v = \langle f''_*(v)h, h\rangle^{1/2} = \langle [f''(y)]^{-1}h, h\rangle^{1/2} = \|h\|^*_y$ for $h \in E^*$. □

We will need some bounds on the variation of the gradient of a self-concordant barrier in terms of a Minkowski function. Let $\pi_z(x)$ be the Minkowski function of Q with the pole at $z \in Q$:

$$\pi_z(x) = \inf\{t > 0 \mid z + t^{-1}(x - z) \in Q\}.$$

Lemma 2.2 *Let u be an arbitrary point in Q. Then for any $v \in Q$, we have*

$$\|f'(v)\|^*_u \le \frac{\nu}{1 - \pi_u(v)}. \tag{2.12}$$

Moreover, if $\langle f'(u), v - u\rangle \ge 0$ for some $v \in Q$, then

$$\|f'(v)\|^*_u \ge \frac{\pi_u(v)}{(\nu + 2\sqrt{\nu})(1 - \pi_u(v))}. \tag{2.13}$$

Finally, if $\langle f'(u), v - w\rangle = 0$ for some $v, w \in Q$, then

$$1 - \pi_u(v) \ge \frac{1 - \pi_u(w)}{1 + \nu + 2\sqrt{\nu}}. \tag{2.14}$$

Proof The set cl Q contains a $\|\cdot\|_u$-ball of radius 1, which is centered at u. Consequently, this set contains a $\|\cdot\|_u$-ball B of radius $1 - \pi_u(v)$, which is centered at v. Since f is a ν-self-concordant barrier, from Proposition 2.1(v), we have

$$\langle f'(v), x - v \rangle \le \nu, \quad \forall x \in B \subseteq \mathrm{cl}\, Q,$$

and (2.12) follows.

In order to prove (2.13), let us choose $v \in Q$ such that

$$\langle f'(u), v - u \rangle \ge 0,$$

and let $r = \|v - u\|_u$. The case of $r = 0$ is trivial. Assuming $r > 0$, let

$$\phi(t) = f\big(u + tr^{-1}(v - u)\big), \quad t \in \Delta = \{t \mid u + tr^{-1}(v - u) \in Q\}.$$

Note that ϕ is self-concordant barrier for Δ and the right endpoint of Δ is the point $T = r/\pi_u(v)$. By Proposition 2.1(i), Δ contains the set $\{s : (s - t)^2\phi''(t) < 1\}$, whence $t + (\phi''(t))^{-1/2} \le T$ for all $t \in \Delta$. Thus, for $t \in \Delta$, $t \ge 0$, we have

$$\phi''(t) \ge \frac{1}{(T - t)^2}. \tag{2.15}$$

Combining (2.15) with the relation $\phi'(0) = r^{-1}\langle f'(u), v-u\rangle \geq 0$, we get

$$\phi'(r) \geq \int_0^r \frac{1}{(T-t)^2}dt = \frac{r}{T(T-r)} = \frac{\pi_u^2(v)}{r(1-\pi_u(v))}.$$

On the other hand, $\phi'(r) = \langle f'(v), v-u\rangle r^{-1} \leq \|f'(v)\|_u^*$, and we come to

$$\big\| f'(v)\big\|_u^* \geq \frac{\pi_u(v)}{1-\pi_u(v)} \cdot \frac{\pi_u(v)}{r}. \tag{2.16}$$

Setting $x = u + \pi_u^{-1}(v)(v-u)$, we have $x \in \operatorname{cl} Q$ and $\langle f'(u), x-u\rangle \geq 0$, whence by Proposition 2.1(v) we get

$$\frac{r}{\pi_u(v)} = \|x-u\|_u \leq \nu + 2\sqrt{\nu},$$

which combined with (2.16) implies (2.13).

To prove (2.14), let $\langle f'(v), w-v\rangle = 0$, and let $Q_v = \{x \in Q \mid \langle f'(v), x-v\rangle = 0\}$. Since v is the minimizer of a ν-self-concordant barrier on Q_v, the set Q_v, regarded as a full-dimensional subset of its affine span, contains an ellipsoid centered at v, and is contained in a $(\nu + 2\sqrt{\nu})$ times larger concentric ellipsoid (Proposition 2.1(i), (v)). It follows that there exists $x \in \operatorname{cl} Q_{v,}$ such that

$$v = \frac{1}{1+\nu+2\sqrt{\nu}}w + \frac{\nu+2\sqrt{\nu}}{1+\nu+2\sqrt{\nu}}x.$$

Thus,

$$\pi_u(v) \leq \frac{1}{1+\nu+2\sqrt{\nu}}\pi_u(w) + \frac{\nu+2\sqrt{\nu}}{1+\nu+2\sqrt{\nu}}\pi_u(x).$$

Note that $\pi_u(x) \leq 1$. Hence,

$$1 - \pi_u(v) \geq \frac{1-\pi_u(w)}{1+\nu+2\sqrt{\nu}},$$

as required in (2.14). □

We conclude this section with two lower bounds on the size of the gradient of self-concordant function computed with respect to the local norm defined by its minimizer.

Lemma 2.3 *Assume that there exists a minimizer x_f of a ν-self-concordant barrier $f(x)$. Then for any $\bar{x} \in \operatorname{dom} f$ we have*

$$f(\bar{x}) - f(x_f) \leq \nu \ln\big(1 + 3\big(1+2\nu^{-1/2}\big)\big\| f'(\bar{x})\big\|_{x_f}^*\big) \leq \nu\ln\big(1 + 9\big\| f'(\bar{x})\big\|_{x_f}^*\big).$$

Proof Denote

$$\delta = \bar{x} - x_f, \qquad r = \|\delta\|_{x_f}, \qquad \phi(t) = f\left(x_f + \frac{t}{r}\delta\right), \qquad \Delta = \operatorname{dom}\phi.$$

Then $\phi(\cdot)$ is a ν-self-concordant barrier for the segment $\operatorname{cl}\Delta$. Note that $r \in \Delta$ and ϕ attains its minimum at $t = 0$. Moreover, $\phi''(0) = 1$. By Proposition 2.1(v), we have

$$\phi'(t)(r-t) \le \nu, \quad \forall t \in [0, r],$$

and $r \le \nu + 2\sqrt{\nu}$. Besides this,

$$0 \le \phi'(t) \le \phi'(r), \quad \forall t \in [0, r].$$

Thus,

$$\begin{aligned} f(\bar{x}) - f(x_f) &= \phi(r) - \phi(0) = \int_0^r \phi'(t)\,dt \\ &\le \int_0^r \min\bigl[\phi'(r), \nu(r-t)^{-1}\bigr]\,dt \\ &= \begin{cases} r\phi'(r), & \text{if } \phi'(r)r \le \nu, \\ \nu(1+\ln(r\phi'(r)/\nu)), & \text{otherwise} \end{cases} \\ &\le \nu \ln\left(1 + 3\frac{r\phi'(r)}{\nu}\right) \\ \left(\phi'(r) = \frac{1}{r}\langle f'(\bar{x}), \delta\rangle\right) &\le \nu\ln\bigl(1 + 3\bigl(1+2\nu^{-1/2}\bigr)\|f'(\bar{x})\|^*_{x_f}\bigr). \end{aligned}$$

□

Lemma 2.4 *Assume that there exists a minimizer x_f of a self-concordant barrier $f(x)$. Then for any $x \in \operatorname{dom} f$, we have*

$$\|x - x_f\|_x \le \bigl(1 + 2\|x - x_f\|_{x_f}\bigr) \cdot \bigl\|f'(x)\bigr\|^*_{x_f}. \tag{2.17}$$

If in addition f is a ν-self-concordant barrier, then

$$\|x - x_f\|_x \le \bigl(2\nu + 4\sqrt{\nu} + 1\bigr) \cdot \bigl\|f'(x)\bigr\|^*_{x_f}. \tag{2.18}$$

Proof Denote

$$\delta = x - x_f, \qquad r = \|\delta\|_{x_f}, \qquad \phi(t) = f\left(x_f + \frac{t}{r}\delta\right).$$

Note that ϕ is self-concordant and $\phi'(0) = 0$. Therefore, in view of (2.5), we have

$$\phi'(r) \ge \frac{r\phi''(r)}{1 + r\sqrt{\phi''(r)}},$$

whence

$$\phi''(r) \le \bigl(\phi'(r)\bigr)^2\left(1 + \frac{1}{r\sqrt{\phi''(r)}}\right)^2. \tag{2.19}$$

Further, $\phi''(t) = \frac{1}{r^2}\langle f''(x_f + \frac{t}{r}\delta)\delta, \delta\rangle$. Thus, $\phi''(0) = 1$ and $\phi''(r) \geq (1+r)^{-2}$ by (2.4). Therefore, (2.19) implies that

$$\phi''(r) \leq \big(\phi'(r)\big)^2\big(2 + r^{-1}\big)^2.$$

Note that

$$\phi'(r) = r^{-1}\big\langle f'(x), \delta\big\rangle \leq \big\| f'(x)\big\|^*_{x_f}.$$

Combining the two last inequalities, we get

$$\big\langle f''(x)\delta, \delta\big\rangle = r^2\phi''(r) \leq (1+2r)^2\big(\phi'(r)\big)^2 \leq (1+2r)^2\big(\big\| f'(x)\big\|^*_{x_f}\big)^2,$$

as required in (2.17). Inequality (2.18) follows from (2.17) and the second part of Proposition 2.1(v). □

3 Lower Bounds on Riemannian Distances

Let us establish lower bounds for the Riemannian distance between two points in Q in terms of different local norms defined by $f(x)$.

Lemma 3.1 *Let u and v belong to Q. Then for any $h \in E\backslash\{0\}$, we have*

$$\left| \ln \frac{\|h\|_u}{\|h\|_v} \right| \leq \sigma(u, v), \tag{3.1}$$

and for any $\eta \in E^*\backslash\{0\}$

$$\left| \ln \frac{\|\eta\|^*_u}{\|\eta\|^*_v} \right| \leq \sigma(u, v). \tag{3.2}$$

Moreover,

$$\left| \ln \frac{\lambda(u)+1}{\lambda(v)+1} \right| \leq \sigma(u, v), \tag{3.3}$$

and

$$\ln \frac{\| f'(u)\|^*_v + 1}{\| f'(v)\|^*_v + 1} \leq \sigma(u, v). \tag{3.4}$$

Besides this, if f is a ν-self-concordant barrier for cl Q, *then*

$$\big| f(u) - f(v)\big| \leq \sqrt{\nu}\,\sigma(u, v). \tag{3.5}$$

Proof By continuity reasons, we may assume that both u, v are distinct from x_f. Let us fix $\epsilon > 0$. Consider a C^1-curve $\gamma(t) \in Q$, $0 \leq t \leq 1$, which satisfies the following conditions:

$$\gamma(0)=u, \qquad \gamma(1)=v,$$

$$\gamma(t)\neq x_f, \quad \forall t\in[0,1],$$

$$\rho\big[\gamma(\cdot),0,1\big]\leq\sigma(u,v)+\epsilon.$$

To prove (3.1), let us fix $h\in E\setminus\{0\}$ and set $\psi(t)=\langle f''(\gamma(t))h,h\rangle$. Then

$$\big|\psi'(t)\big|=\big|D^3f(\gamma(t))\big[\gamma'(t),h,h\big]\big|\overset{(2.2)}{\leq}2\big\langle f''\big(\gamma(t)\big)\gamma'(t),\gamma'(t)\big\rangle^{1/2}\psi(t),$$

whence

$$\left|\ln\frac{\langle f''(u)h,h\rangle^{1/2}}{\langle f''(v)h,h\rangle^{1/2}}\right|=\left|\frac{1}{2}\ln\frac{\psi(1)}{\psi(0)}\right|\leq\rho\big[\gamma(\cdot),0,1\big]\leq\sigma(u,v)+\epsilon,$$

and (3.1) follows. Relation (3.2) can be derived from (3.1) using the definition of the dual norm.

Further, denoting $\psi(t)=\lambda(\gamma(t))$, we have

$$\begin{aligned}\left|\frac{d}{dt}\psi^2(t)\right|&=\big|-D^3f(\gamma(t))\big[\gamma'(t),\big[f''\big(\gamma(t)\big)\big]^{-1}f'\big(\gamma(t)\big),\big[f''\big(\gamma(t)\big)\big]^{-1}f'\big(\gamma(t)\big)\big]\\&\quad+2\big\langle\big[f''\big(\gamma(t)\big)\big]^{-1}f''\big(\gamma(t)\big)\gamma'(t),f'\big(\gamma(t)\big)\big\rangle\big|\\&\leq2\big\langle f''\big(\gamma(t)\big)\gamma'(t),\gamma'(t)\big\rangle^{1/2}\\&\quad\times\big\langle f''\big(\gamma(t)\big)\big[f''\big(\gamma(t)\big)\big]^{-1}f'\big(\gamma(t)\big),\big[f''\big(\gamma(t)\big)\big]^{-1}f'\big(\gamma(t)\big)\big\rangle\\&\quad+2\big\langle\big[f''\big(\gamma(t)\big)\big]^{-1}f'\big(\gamma(t)\big),f'\big(\gamma(t)\big)\big\rangle^{1/2}\big\langle f''\big(\gamma(t)\big)\gamma'(t),\gamma'(t)\big\rangle^{1/2}\\&=2\psi(t)\big(\psi(t)+1\big)\big\langle f''\big(\gamma(t)\big)\gamma'(t),\gamma'(t)\big\rangle^{1/2}.\end{aligned}$$

Since $\psi(\cdot)>0$ on $[0,1]$, we get

$$\left|\frac{d}{dt}\ln\big(1+\psi(t)\big)\right|\leq\big\langle f''\big(\gamma(t)\big)\gamma'(t),\gamma'(t)\big\rangle^{1/2},$$

whence

$$\left|\ln\frac{\lambda(u)+1}{\lambda(v)+1}\right|\leq\rho\big[\gamma(\cdot),0,1\big]\leq\sigma(u,v)+\epsilon,$$

and (3.3) follows.

To prove (3.4), let $\psi(t)=\|f'(\gamma(t))\|_v^*$ and $r(t)=\rho[\gamma(\cdot),0,t]$. Then

$$\begin{aligned}\frac{d}{dt}\psi^2(t)&=\frac{d}{dt}\bigl\langle f'\bigl(\gamma(t)\bigr),\bigl[f''(v)\bigr]^{-1}f'\bigl(\gamma(t)\bigr)\bigr\rangle\\&=2\bigl\langle f''\bigl(\gamma(t)\bigr)\gamma'(t),\bigl[f''(v)\bigr]^{-1}f'\bigl(\gamma(t)\bigr)\bigr\rangle\\&\le 2\underbrace{\|\gamma'(t)\|_{\gamma(t)}}_{r'(t)}\times\bigl\|\bigl[f''(v)\bigr]^{-1}f'\bigl(\gamma(t)\bigr)\bigr\|_{\gamma(t)}\end{aligned}$$

$$(\text{by (3.1)})\le 2r'(t)e^{r(t)}\times\bigl\|\bigl[f''(v)\bigr]^{-1}f'\bigl(\gamma(t)\bigr)\bigr\|_v=2r'(t)e^{r(t)}\psi(t),$$

whence $\psi'(t)\le r'(t)e^{r(t)}$, so that

$$\begin{aligned}\bigl\|f'(u)\bigr\|_v^*+1=\psi(1)+1&\le\psi(0)+1+\bigl[e^{r(1)}-e^{r(0)}\bigr]\\&\le\bigl\|f'(v)\bigr\|_v^*+\exp\bigl\{\sigma(u,v)+\epsilon\bigr\}\\&\le\bigl(\bigl\|f'(v)\bigr\|_v^*+1\bigr)\exp\bigl\{\sigma(u,v)+\epsilon\bigr\},\end{aligned}$$

and (3.4) follows.

Finally, to prove (3.5), it suffices to note that if f is a ν-self-concordant barrier, then

$$\begin{aligned}\left|\frac{d}{dt}f\bigl(\gamma(t)\bigr)\right|&=\bigl|\bigl\langle f'\bigl(\gamma(t)\bigr),\gamma'(t)\bigr\rangle\bigr|\le\lambda\bigl(\gamma(t)\bigr)\cdot\bigl\langle f''\bigl(\gamma(t)\bigr)\gamma'(t),\gamma'(t)\bigr\rangle^{1/2}\\&\le\sqrt{\nu}\cdot\bigl\langle f''\bigl(\gamma(t)\bigr)\gamma'(t),\gamma'(t)\bigr\rangle^{1/2},\end{aligned}$$

whence

$$\bigl|f(u)-f(v)\bigr|\le\sqrt{\nu}\rho\bigl[\gamma(\cdot),0,1\bigr]\le\sqrt{\nu}\bigl[\sigma(u,v)+\epsilon\bigr],$$

and (3.5) follows. □

4 Riemannian Length of Central Path

Let f be a nondegenerate self-concordant function with domain $Q\subseteq E$. Given a nonzero vector $e\in E^*$, consider the associated *central path*

$$x(t)=\operatorname*{argmin}_x\bigl[-t\langle e,x\rangle+f(x)\bigr].\tag{4.1}$$

By Proposition 2.1(iv), the domain of this curve is an open interval Δ on the axis. From now on, we assume that this interval contains a given segment $[t_0,t_1]$ with

$0 \le t_0 < t_1 < \infty$. Note that by the implicit function theorem the path $x(t)$ is continuously differentiable on its domain and satisfies the relations

$$f'(x(t)) = te; \qquad x'(t) = [f''(x(t))]^{-1} e = t^{-1}[f''(x(t))]^{-1} f'(x(t)). \tag{4.2}$$

Our goal now is to obtain a useful upper bound on the Riemannian length $\rho[x(\cdot), t_0, t_1]$ of the central path.

Lemma 4.1 (i) *We always have*

$$\rho[x(\cdot), t_0, t_1] \le \sqrt{[f(x(t_1)) - f(x(t_0))] \ln \frac{t_1}{t_0}}. \tag{4.3}$$

If in addition $\lambda(x(t_0)) \ge \frac{1}{2}$, *then*

$$\ln \frac{t_1}{t_0} = \ln \frac{\|f'(x(t_1))\|^*_{x(t_0)}}{\|f'(x(t_0))\|^*_{x(t_0)}} \le \sigma(x(t_0), x(t_1)) + \ln 3. \tag{4.4}$$

(ii) *If* $\lambda(x(t_0)) < \frac{1}{2}$, *then*

$$\rho[x(\cdot), t_0, t_1] \le \ln 2 + \sqrt{[f(x(t_1)) - f(x(\widehat{t}))] \ln \frac{t_1}{\widehat{t}}}, \tag{4.5}$$

where $\widehat{t}$ *is the largest* $t \in [t_0, t_1]$ *such that* $\lambda(x(t)) \le \frac{1}{2}$, *and*

$$\ln \frac{t_1}{\widehat{t}} \le \ln(\max\{1, 4\|f'(x(t_1))\|^*_{x_f}\}) \le \sigma(x(t_0), x(t_1)) + \ln 12. \tag{4.6}$$

Proof Let

$$\phi(t) = f(x(t)), \qquad r(t) = \rho[x(\cdot), t_0, t], \quad t_0 \le t \le t_1.$$

Since from (4.2) $x'(t) = t^{-1}[f''(x(t))]^{-1} f'(x(t))$, we have

$$\phi'(t) = \langle f'(x(t)), x'(t)\rangle = t^{-1}\langle [f''(x(t))]^{-1} f'(x(t)), f'(x(t))\rangle = t^{-1}\lambda^2(x(t)),$$

$$r'(t) = \langle [f''(x(t))] x'(t), x'(t)\rangle^{1/2} = t^{-1}\langle [f''(x(t))]^{-1} f'(x(t)), f'(x(t))\rangle^{1/2}$$
$$= t^{-1}\lambda(x(t)),$$

whence by the Cauchy inequality

$$\rho^2[x(\cdot), t_0, t_1] = \left(\int_{t_0}^{t_1} t^{-1}\lambda(x(t))\, dt\right)^2 \le \left(\int_{t_0}^{t_1} t^{-1}\lambda^2(x(t))\, dt\right)\left(\int_{t_0}^{t_1} t^{-1}\, dt\right)$$
$$= (\phi(t_1) - \phi(t_0)) \ln \frac{t_1}{t_0},$$

as required in (4.3).

Let us prove now inequality (4.4). Assume that $\lambda(x(t_0)) \geq \frac{1}{2}$. Since

$$f'\big(x(t_1)\big) = \frac{t_1}{t_0} f'\big(x(t_0)\big),$$

using (3.4) and the fact that $\lambda(x(t_0)) = \| f'(x(t_0))\|^*_{x(t_0)} \geq \frac{1}{2}$, we get

$$\ln \frac{t_1}{t_0} = \ln \frac{\| f'(x(t_1))\|^*_{x(t_0)}}{\| f'(x(t_0))\|^*_{x(t_0)}} \leq \ln(3 \frac{\| f'(x(t_1))\|^*_{x(t_0)} + 1}{\| f'(x(t_0))\|^*_{x(t_0)} + 1}) \leq \sigma\big(x(t_0), x(t_1)\big) + \ln 3,$$

as required in (4.4).

Now assume that $\lambda(x(t_0)) < \frac{1}{2}$. By Proposition 2.1(ii), under this assumption f attains its minimum on Q at a unique point x_f, the quantity

$$T = \max\big\{t \geq 0 : t \in \Delta, \lambda\big(x(t)\big) \leq \tfrac{1}{2}\big\}$$

is well defined and $\lambda(x(T)) = \frac{1}{2}$. Note that the multiplication of vector e in (4.1) by an appropriate constant changes only the scale of the time and does not change the trajectory. Hence, for the sake of notation, we may assume that $T = 1$. Since $\lambda(x(t_0)) < \frac{1}{2}$, we have

$$t_0 < \widehat{t} \leq T = 1.$$

Let us first prove that

$$\rho\big[x(\cdot), 0, 1\big] \leq \ln 2. \tag{4.7}$$

Note that $e = f'(x(1))$. Denote by f_* the Legendre transformation of f. Since f attains it minimum on Q, we have $0 \in \operatorname{dom} f_*$ and since $1 = T \in \Delta$, we have $e \in \operatorname{dom} f_*$. Besides this, f_* is nondegenerate and self-concordant on its domain in view of Proposition 2.1(iv). Thus,

$$\big\langle e, f''_*(e)e\big\rangle = \big\langle f'\big(x(1)\big), \big[f''\big(x(1)\big)\big]^{-1} f'\big(x(1)\big)\big\rangle = \lambda^2\big(x(1)\big) = \frac{1}{4},$$

whence, by Proposition 2.1(i)

$$\big\langle e, f''_*(te)e\big\rangle \leq \frac{\lambda^2(x(1))}{(1 - (1-t)\lambda(x(1)))^2}, \quad 0 \leq t \leq 1.$$

Hence,

$$\int_0^1 \big\langle e, f''_*(te)e\big\rangle^{1/2} dt \leq \int_0^1 \frac{\lambda(x(1))}{1 - (1-t)\lambda(x(1))} dt = -\ln\big(1 - \lambda\big(x(1)\big)\big) = \ln 2.$$

On the other hand, we have $f'(x(t)) = t f'(x(1))$, whence

$$f''\big(x(t)\big)x'(t) = f'\big(x(1)\big) = t^{-1} f'\big(x(t)\big),$$

and consequently

$$\begin{aligned}\rho\big[x(\cdot),0,1\big] &= \int_0^1 \big\langle f''\big(x(t)\big)x'(t),x'(t)\big\rangle^{1/2}\,dt\\ &= \textstyle\int_0^1 \big\langle f'\big(x(1)\big),\big[f''\big(x(t)\big)\big]^{-1}f'\big(x(1)\big)\big\rangle^{1/2}\,dt\\ &= \int_0^1 \big\langle e, f''_*\big(f'\big(x(t)\big)\big)e\big\rangle^{1/2}\,dt = \int_0^1 \big\langle e, f''_*(te)e\big\rangle^{1/2}\,dt \le \ln 2,\end{aligned}$$

as required in (4.7).

Now let us prove (4.5). In the case of $t_1 \le 1\,[\equiv T]$, inequality (4.5) immediately follows from (4.7). In the case of $1 < t_1$, we have $\widehat{t} = 1$, and

$$\rho\big[x(\cdot),t_0,t_1\big] \le \rho\big[x(\cdot),0,t_1\big] = \rho\big[x(\cdot),0,1\big] + \rho\big[x(\cdot),1,t_1\big] \le \ln 2 + \rho\big[x(\cdot),1,t_1\big].$$

Bounding $\rho[x(\cdot),1,t_1] = \rho[x(\cdot),\widehat{t},t_1]$ from above by (4.3), we get (4.5).

It remains to prove (4.6). There is nothing to prove if $\widehat{t} = t_1$. Thus, we may assume that $\widehat{t} < t_1$, whence, in particular, $\lambda(x(\widehat{t})) = \frac{1}{2}$. In view of the latter observation, we can apply (4.4) and get

$$\ln\frac{t_1}{\widehat{t}} = \ln\frac{\|f'(x(t_1))\|^*_{x(\widehat{t})}}{\|f'(x(\widehat{t})\|^*_{x(\widehat{t})}} \le \sigma\big(x(\widehat{t}),x(t_1)\big) + \ln 3,$$

or, which is the same,

$$\ln\frac{t_1}{\widehat{t}} = \ln\big(2\big\|f'\big(x(t_1)\big)\big\|^*_{x(\widehat{t})}\big) \le \sigma\big(x(\widehat{t}),x(t_1)\big) + \ln 3. \tag{4.8}$$

By (4.7), we have

$$\sigma\big(x(t_0),x(\widehat{t})\big) \le \rho\big[x(\cdot),t_0,\widehat{t}\big] \le \rho\big[x(\cdot),0,\widehat{t}\big] \le \rho\big[x(\cdot),0,1\big] \le \ln 2, \tag{4.9}$$

whence by triangle inequality

$$\sigma\big(x(\widehat{t}),x(t_1)\big) \le \ln 2 + \sigma\big(x(t_0),x(t_1)\big).$$

Note that by (2.7)

$$\frac{1}{2}\big\|f'\big(x(t_1)\big)\big\|^*_{x_f} \le \big\|f'\big(x(t_1)\big)\big\|^*_{x(\widehat{t})} \le 2\big\|f'\big(x(t_1)\big)\big\|^*_{x_f}.$$

Combining these relations and (4.8), we arrive at (4.6). □

Remark 4.1 Note that in the proof of (4.3), we did not use the fact that f is self-concordant.

We can establish now the $O(\nu^{1/4})$-*geodesic* property of the central path associated with a self-concordant barrier (see [6] for terminology).

Theorem 4.1 *Let f be a nondegenerate ν-self-concordant barrier for* $\operatorname{cl} Q$, *and let $x(t)$ be the central path given by*

$$f'\big(x(t)\big) = te, \quad 0 \le t_0 \le t \le t_1.$$

Then

$$\rho\big[x(\cdot), t_0, t_1\big] \le \ln 2 + \nu^{1/4}\sqrt{\sigma\big(x(t_0), x(t_1)\big)\big[\sigma\big(x(t_0), x(t_1)\big) + \ln 12\big]}. \tag{4.10}$$

If $\lambda(x(t_0)) \ge \frac{1}{2}$, then

$$\rho\big[x(\cdot), t_0, t_1\big] \le \nu^{1/4}\sqrt{\sigma\big(x(t_0), x(t_1)\big)\big[\sigma\big(x(t_0), x(t_1)\big) + \ln 3\big]}. \tag{4.11}$$

Proof It suffices to combine (4.3), (4.5), (3.5), (4.4), and (4.6). □

Recall that the primal-dual central paths are $\sqrt{2}$-geodesic (see Theorem 5.2 in [6]).

5 Applications

In this section, we apply the results of Sect. 4 to the analysis of several short-step interior-point methods. We consider the following three problems:

1. Finding an approximation to the minimizer of a self-concordant function f.
2. Finding a point in a nonempty intersection of a bounded convex open domain Q and an affine plane. We assume that Q is represented by a ν-self-concordant barrier f with $\operatorname{dom} f = Q$, and that the minimizer x_f is known.
3. Finding an ϵ-solution to the optimization problem $\min_{\operatorname{cl} Q}\langle c, x\rangle$, with Q represented in the same way as in item 2.

Our main results state that in problems 2 and 3 (as in problem 1, provided that f is a ν-self-concordant barrier), appropriate well-known short-step path-following methods are "suboptimal" within the factor $\nu^{1/4}$. Namely, the number of Newton steps, which is required by the methods, coincides, up to a factor $O(\nu^{1/4})$, with the Riemannian distance between the starting point and the set of solutions. Recall that this distance is a natural *lower* bound on the number of iterations of the short-step interior-point methods.

5.1 Minimization of a Self-Concordant Function

Consider the following problem:

$$\min_{x\in Q} f(x), \tag{5.1}$$

where f is a nondegenerate bounded-below self-concordant function with $\operatorname{dom} f = Q$. Assume that we have a starting point $\bar{x} \in Q$. Let us analyze the efficiency of a short-step path-following method $\mathcal{M}(\bar{x})$, which traces the central path $x(t)$:

$$f'\big(x(t)\big) = tf'(\bar{x}), \quad t \in [0, 1]. \tag{5.2}$$

as t decreases.[2] By Proposition 2.1(ii), $0 \in \operatorname{dom} f_*$, and of course $f'(\bar{x}) \in \operatorname{dom} f_*$. Consequently, the path is well defined (Proposition 2.1(iv)), and $x(0)$ is the minimizer x_f of $f(x)$ over Q.

To avoid trivialities, we assume that $\lambda(\bar{x}) > \frac{1}{2}$, i.e., that $\bar{x}$ does not belong to the domain of quadratic convergence of the Newton method as applied to f. Then

$$f(\bar{x}) - f(x_f) \geq O(1), \qquad \|f'(\bar{x})\|^*_{x_f} \geq O(1), \qquad \sigma(x_f, \bar{x}) \geq O(1) \tag{5.3}$$

(from now on, $O(1)$'s are positive absolute constants). Indeed, the first inequality is readily given by Proposition 2.1(iii); the second inequality follows from the first one in view of (2.11) applied with $x = \bar{x}$, $y = x_f$, and the third inequality follows from the second one in view of (3.4).

Using Lemma 4.1(ii) with $t_0 = 0$, $t_1 = 1$, in view of (5.3), we get the following result.

Theorem 5.1 *Let $\lambda(\bar{x}) > \frac{1}{2}$. Then the short-step method $\mathcal{M}(\bar{x})$ justifies the following upper bound*:

$$\begin{aligned} N_f(\bar{x}, x_f) &\leq O(1)\sqrt{\big[f(\bar{x}) - f(x_f)\big]\ln\big(1 + \|f'(\bar{x})\|^*_{x_f}\big)} \\ &\leq O(1)\sqrt{\big[f(\bar{x}) - f(x_f)\big]\sigma(x_f, \bar{x})}. \end{aligned} \tag{5.4}$$

Let us discuss the bound (5.4).

1. The only previously known efficiency estimate for the problem (5.1) was

$$O(1)\big(f(\bar{x}) - f(x_f)\big) \tag{5.5}$$

Newton steps (recall that we do not assume f to be a self-concordant barrier). The simplest method providing us with this estimate is the usual damped Newton method. However, using a standard argument, it is not difficult to see that the short-step path-following scheme as applied to (5.1) also shares the same estimate (5.5). Let us demonstrate that the bound (5.4) is sharper. Indeed, using inequality (2.10), we have

$$f(\bar{x}) - f(x_f) \geq \ln\big(1 + \|f'(\bar{x})\|^*_{x_f}\big) - 1,$$

so that the bound (5.5) on $N_f(\bar{x}, x_f)$ follows from the first inequality in (5.4) combined with initial conditions (5.3).

Note that the bound (5.4) can be much smaller than (5.5).

Example 5.1 Let B_n be the unit n-dimensional box:

$$B_n = \big\{x \in R^n : \big|x^{(i)}\big| \leq 1, i = 1, \ldots, n\big\},$$

[2] For a precise definition of the notion of short-step path-following method, see Definition 3.1 in [6].

and $f(x) = -\sum_{i=1}^{n} \ln(1-(x^{(i)})^2)$. Without loss of generality, we may assume that $\bar{x}^{(i)} \geq 0$, so that $\bar{x}^{(i)} = [1-\epsilon_i]^{1/2}$, for some $\epsilon_i \in (0,1]$, $i = 1, \ldots, n$. Then

$$f(\bar{x}) = \sum_{i=1}^{n} \ln \frac{1}{\epsilon_i}.$$

At the same time,

$$\left(\| f'(\bar{x}) \|_{x_f}^{*}\right)^2 = \frac{1}{2} \sum_{i=1}^{n} \frac{4(\bar{x}^{(i)})^2}{(1-(x^{(i)})^2)^2} = 2 \sum_{i=1}^{n} \frac{1-\epsilon_i}{\epsilon_i^2} \leq 2 \sum_{i=1}^{n} \frac{1}{\epsilon_i^2} \leq \frac{2n}{\min_{1\leq i \leq n} \epsilon_i^2}.$$

It follows that the ratio of the complexity bound (5.4) to the bound (5.5) does not exceed

$$O(1)\left(\left[1 + \max_{1\leq i\leq n} \ln \frac{n}{\epsilon_i}\right] \Big/ \left[1 + \sum_{i=1}^{n} \ln \frac{1}{\epsilon_i}\right]\right)^{1/2},$$

and the latter quantity can be arbitrary close to $n^{-1/2}$.

2. Consider a particular case of problem (5.1), with f being a ν-self-concordant barrier. In this case, the known complexity estimate for a short-step path-following scheme is

$$O(1)\sqrt{\nu} \ln\left(1 + \nu \| f'(\bar{x}) \|_{x_f}^{*}\right) \tag{5.6}$$

(see, for example, [4]). However, from Lemma 2.3, it is clear that the estimate (5.4) is sharper.

3. As it is shown in [6], a natural lower bound on the number of Newton steps in every short-step interior-point method for solving (5.1) is the Riemannian distance from $\bar{x}$ to x_f. In the case when f is a ν-self-concordant barrier, the performance of the path-following scheme, in view of Theorem 4.1 (or in view of (5.4) combined with (3.5)), is at most $O(\nu^{1/4})$ times worse than this lower bound.

5.2 Finding a Feasible Point

Consider the following problem:

$$\text{Find a point } \bar{x} \in \mathcal{F} = \{x \in Q,\ Ax = b\}, \tag{5.7}$$

where Q is an open and bounded convex set endowed with a ν-self-concordant barrier $f(x)$, and $A : x \mapsto Ax$ is a linear mapping from E onto a linear space F. Since the mapping A is onto, the conjugate mapping $A^* : F^* \to E^*$ is an embedding. From now on, we assume that problem (5.7) is feasible, and that we know the minimizer x_f of f on Q. Without loss of generality, we assume that $x_f = 0$.

Let f_* be the Legendre transformation of f. Since $x_f = 0$, we have

$$f_*'(0) = 0. \tag{5.8}$$

Note that $f_*(s)$ is a self-concordant function with $\operatorname{dom} f_* = E^*$ such that

$$\langle s, f''_*(s)s \rangle \le \nu, \quad \forall s \in E^* \tag{5.9}$$

(see [5], Theorem 2.4.2). In order to avoid trivial cases, let us assume that

$$\{x : Ax = b\} \cap \{x : \|x\|_0 < 1\} = \emptyset \tag{5.10}$$

(recall that $\|\cdot\|_0 \equiv \|\cdot\|_{x_f}$). Indeed, otherwise a solution to (5.7) can be found by projecting the origin onto the plane $Ax = b$ in the Euclidean metric $\|\cdot\|_0$ (see Proposition 2.1(i)).

In order to solve (5.7), we can trace the path of minimizers of f on the sets E_t,

$$x(t) = \operatorname*{argmin}_{E_t} f(x), \quad E_t = \{x \in Q \mid Ax = tb\},$$

as t varies from 0 to 1. Note that this path is well defined: E_0 and E_1 are nonempty and bounded by assumption. Therefore, all E_t, $t \in (0, 1)$ are also nonempty and bounded.

As it is shown in [6], the path $x(\cdot)$ can be traced by an appropriate short-step path-following sequence, which length is proportional to $\rho[x(\cdot), 0, 1]$. Thus, in order to establish the complexity of our problem, we need to find some bounds on the Riemannian length of the path $x(t)$.

Observe that tracing $x(t)$ as t varies from 1 to 0 is equivalent to tracing a dual central path

$$y(t): \quad h'(y(t)) = th'(0), \tag{5.11}$$

but with t varying from 0 to 1. This path is associated with a nondegenerate (since A^* is an embedding) self-concordant function

$$h(y) = f_*(A^*y) - \langle y, b\rangle : F^* \to R.$$

The relation between the paths $x(t)$ and $y(t)$ is given by the following lemma.

Lemma 5.1 *For any* $t \in [0, 1]$, *we have*

$$x(1-t) = f'_*(A^*y(t)). \tag{5.12}$$

Proof Indeed, from the origin of $x(1-t)$, it follows that $f'(x(1-t)) \in (\operatorname{Ker} A)^\perp$. This means that $f'(x(1-t)) = A^*y(t)$ for certain uniquely defined $y(t)$, and this is equivalent to (5.12). Besides this, $Ax(1-t) = (1-t)b$, whence $Af'_*(A^*y(t)) = (1-t)b$, or

$$h'(y(t)) = Af'_*(A^*y(t)) - b = -tb = th'(0),$$

where the concluding equality is readily given by (5.8). We have arrived at (5.11). □

It turns out that the Riemannian length $\rho[y(\cdot),0,1]$ of the path $y(\cdot)$ in the Riemannian structure given by $h(\cdot)$ is equal to $\rho[x(\cdot),0,1]$:

$$\begin{aligned}\rho\big[x(\cdot),0,1\big] &= \int_0^1 \big\langle f''\big(x(t)\big)x'(t), x'(t)\big\rangle^{1/2}\,dt\\ &= \int_0^1 \big\langle f''\big(f'_*\big(A^*y(1-t)\big)\big) f''_*\big(A^*y(1-t)\big)A^*y'(1-t),\\ &\quad f''_*\big(A^*y(1-t)\big)A^*y'(1-t)\big\rangle^{1/2}\,dt\\ &= \int_0^1 \big\langle A^*y'(1-t), f''_*\big(A^*y(1-t)\big)A^*y'(1-t)\big\rangle^{1/2}\,dt\\ &= \int_0^1 \big\langle h''\big(y(1-t)\big)y'(1-t), y'(1-t)\big\rangle^{1/2}\,dt = \rho\big[y(\cdot),0,1\big].\end{aligned}$$

Observe that by (5.11), $y(1)=0$, while $y(0)$ is the minimizer of $h(\cdot)$. In view of the latter fact and Lemma 4.1, we have

$$\rho\big[x(\cdot),0,1\big] = \rho\big[y(\cdot),0,1\big] \le O(1)\left[\ln 2 + \sqrt{\big[h\big(y(1)\big)-h\big(y(0)\big)\big]\ln\frac{1}{\widehat{t}}}\right], \tag{5.13}$$

where $\widehat{t}$ is the largest $t\in[0,1]$ such that $\lambda_h(y(t))\le\frac{1}{2}$, $\lambda_h(y)$ being the local norm of the gradient of $h(\cdot)$ at $y\in\operatorname{dom}h(\cdot)\equiv F^*$.

Since $y(1)=0$, we have $h'(y(1))=-b$. Moreover, the point

$$z = \big[f''(0)\big]^{-1}A^*\big[A\big[f''(0)\big]^{-1}A^*\big]^{-1}b$$

clearly belongs to E_1, so that by (5.10), we have

$$\begin{aligned}1 \le \|z\|_0^2 &= \big\langle b, \big[A\big[f''(0)\big]^{-1}A^*\big]^{-1}b\big\rangle = \big\langle b, \big[Af''_*(0)A^*\big]^{-1}b\big\rangle\\ &= \big\langle h'(0), \big[h''(0)\big]^{-1}h'(0)\big\rangle = \lambda_h^2(0) = \lambda_h^2\big(y(1)\big).\end{aligned}$$

From $\lambda_h(y(1))\ge 1$ and $\lambda_h(y(0))=0$, it follows that

$$\begin{gathered}h\big(y(1)\big)-h\big(y(0)\big) = h\big(y(1)\big) - \min h \ge 1-\ln 2,\\ \lambda_h(\widehat{t}) = \frac{1}{2},\\ \ln\frac{1}{\widehat{t}} \ge O(1)\end{gathered} \tag{5.14}$$

(see Proposition 2.1). Consequently, (5.13) implies that

$$\rho\big[y(\cdot),0,1\big] \le \sqrt{\big[h\big(y(1)\big)-h\big(y(0)\big)\big]\ln\frac{1}{\widehat{t}}}. \tag{5.15}$$

The next statement expresses the complexity bound (5.15) in terms of f.

Theorem 5.2 *For every* $x \in \mathcal{F} \equiv E_1 \cap Q$, *we have*

$$\rho\big[x(\cdot), 0, 1\big] \le O(1)\sqrt{\big[f(x) - f(x_f)\big]\big[O(1) + \ln \nu + \ln \|f'(x)\|^*_{x_f}\big]} \\ \le O(1)\sqrt{\nu}\ln\big(O(1)\nu \|f'(x)\|^*_{x_f}\big). \tag{5.16}$$

Proof Recall that we have assumed $x_f = 0$. Observe first that

$$h\big(y(1)\big) = h(0) = f_*(0) = -\min_Q f = -f(0).$$

Therefore, using (5.12), we get

$$\begin{aligned} h\big(y(0)\big) &= f_*\big(A^* y(0)\big) - \big\langle y(0), b\big\rangle \\ &= \big\langle A^* y(0), f'_*\big(A^* y(0)\big)\big\rangle - f\big(f'_*\big(A^* y(0)\big)\big) - \big\langle y(0), b\big\rangle \\ &= \big\langle y(0), Ax(1) - b\big\rangle - f\big(x(1)\big) = -f\big(x(1)\big). \end{aligned}$$

Thus,

$$h\big(y(1)\big) - h\big(y(0)\big) = f\big(x(1)\big) - \min_Q f = f\big(x(1)\big) - f(0). \tag{5.17}$$

Let us prove now that

$$\ln \frac{1}{\hat{t}} \le O(1) \ln\big\langle f''\big(x(1)\big)x(1), x(1)\big\rangle. \tag{5.18}$$

Denote $d = \|h'(y(1))\|^*_{y(0)}$. Then, either $d \ge 1$, or $d < 1$. In the latter case applying (2.11) with h playing the role of f and $x = y(1)$, $y = y(0)$, we get

$$h\big(y(1)\big) - h\big(y(0)\big) \le \ln(1 - d) + \frac{d}{1 - d},$$

which combined with the first relation in (5.14) results in

$$\ln(1 - d) + \frac{d}{1 - d} \ge O(1),$$

whence in any case $d \ge O(1)$.

The last conclusion combined with Lemma 4.1(ii), (see (4.6)) implies that

$$\ln \frac{1}{\hat{t}} \le O(1) \ln d. \tag{5.19}$$

Recalling that $y(1) = 0$, $h'(0) = -b$, we get

$$\begin{aligned} d^2 &= \big\langle h'\big(y(1)\big), \big[h''\big(y(0)\big)\big]^{-1} h'\big(y(1)\big)\big\rangle = \big\langle b, \big[Af''_*\big(A^* y(0)\big)A^*\big]^{-1} b\big\rangle \\ &= \big\langle Ax(1), \big[A\big[f''\big(x(1)\big)\big]^{-1} A^*\big]^{-1} Ax(1)\big\rangle \le \big\langle f''\big(x(1)\big)x(1), x(1)\big\rangle, \end{aligned}$$

and (5.18) follows from (5.19).

Combining (5.15), (5.17), (5.18), and (2.18), we come to the following inequality

$$\rho\big[x(\cdot),0,1\big] \le O(1)\sqrt{\big[f\big(x(1)\big)-f(0)\big]\ln\big(O(1)\nu\big\|f'\big(x(1)\big)\big\|_0^*\big)}. \tag{5.20}$$

Let us derive inequality (5.16) from (5.20). Since $x(1)$ is the minimizer of f on $\mathcal{F}$ and $x(0)=0$, for every $x\in\mathcal{F}$, we have

$$f\big(x(1)\big)-f\big(x(0)\big)\le f(x)-f(0)\le\big\langle f'(x),x\big\rangle\le\|x\|_0\big\|f'(x)\big\|_0^*.$$

We already have mentioned that $\|x\|_0\le\nu+2\sqrt{\nu}$ for every $x\in Q$. Thus, the first part of (5.16) is proved. Further, let $x\in\mathcal{F}$. Applying (2.12) with $u=0$ and $v=x(1)$, we have

$$\begin{aligned}\big\|f'\big(x(1)\big)\big\|_0^* &\le \frac{\nu}{1-\pi_0(x(1))}\\ [(2.14)\text{ with } u=0,\, v=x(1),\, w=x] &\le \frac{\nu(1+\nu+2\sqrt{\nu})}{1-\pi_0(x)}\\ [(2.13)\text{ with } u=0,\, v=x] &\le \frac{\nu(\nu+2\sqrt{\nu})(1+\nu+2\sqrt{\nu})}{\pi_0(x)}\big\|f'(x)\big\|_0^*\\ &\le O(1)\nu^4\big\|f'(x)\big\|_0^*,\end{aligned}$$

the concluding inequality being given by (5.10) combined with the fact that Q is contained in $\|\cdot\|_0$-ball of the radius $\nu+2\sqrt{\nu}$ centered at 0 (Proposition 2.1(v)). Thus, we get the second part of inequality (5.16). □

Corollary 5.1 *Under Assumption* (5.10), *the Riemannian length of the central path can be bounded from above as*

$$\rho\big[x(\cdot),0,1\big]\le O(1)\nu^{1/4}\big(\sigma(0,\mathcal{F})+\ln\nu\big), \tag{5.21}$$

where $\sigma(x,X)$ is the infimum of Riemannian lengths of curves starting at a point x and ending at a point from the set X.

Proof Let $x\in\mathcal{F}$. By (3.5), we have

$$f(x)-f(0)\le\sqrt{\nu}\sigma(0,x),$$

while by (3.4) we have

$$\ln\big(1+\big\|f'(x)\big\|_0^*\big)\le\sigma(0,x).$$

Combining these inequalities and (5.16), we get

$$\rho\big[x(\cdot),0,1\big]\le O(1)\nu^{1/4}\big(\sigma(0,x)+\ln\nu\big);$$

since the resulting inequality is valid for every $x\in\mathcal{F}$ and $\sigma(0,x)\ge O(1)$ by (5.10), (5.21) follows. □

Note that in view of Example 1.1 this statement is not valid for an unbounded Q.

5.3 Central Path in Standard Minimization Problem

Now let us look at the optimization problem

$$\min\bigl\{\langle c, x\rangle : x \in \mathrm{cl}\, Q\bigr\}, \tag{5.22}$$

where Q is a bounded open convex set endowed with a ν-self-concordant barrier f. Assuming that the minimizer x_f of f on Q is available, consider the standard f-generated short-step path-following method for solving (5.22), where one traces the central path

$$x(t) : f'\bigl(x(t)\bigr) = -tc, \quad t \geq 0,$$

as $t \to \infty$. From the standard complexity bound for the number of Newton steps required to trace the segment $0 < t_0 \leq t \leq t_1$ of the path, we can derive that

$$N_f\bigl(x(t_0), x(t_1)\bigr) \leq O(1)\left(1 + \sqrt{\nu}\ln\frac{t_1}{t_0}\right). \tag{5.23}$$

On the other hand, Lemma 4.1 combined with approach [6] leads to another bound:

$$\begin{aligned} &N_f\bigl(x(t_0), x(t_1)\bigr) \leq O(1)\bigl(1 + \rho\bigl[x(\cdot), t_0, t_1\bigr]\bigr),\\ &\rho\bigl[x(\cdot), t_0, t_1\bigr] \leq \sqrt{\bigl[f\bigl(x(t_1)\bigr) - f\bigl(x(t_0)\bigr)\bigr]\ln\frac{t_1}{t_0}}. \end{aligned} \tag{5.24}$$

Moreover, the value $\rho[x(\cdot), t_0, t_1]$ is an *exact two-side estimate* (up to constant factors) for the number of steps of any short-step path following scheme (see [6]).

Note that the new bound (5.24) is sharper than the old one. Indeed, by (3.5), we have

$$f\bigl(x(t_1)\bigr) - f\bigl(x(t_0)\bigr) \leq \sqrt{\nu}\sigma\bigl(x(t_0), x(t_1)\bigr) \leq \sqrt{\nu}\rho\bigl[x(\cdot), t_0, t_1\bigr],$$

so that the second inequality in (5.24) says that $\rho[x(\cdot), t_0, t_1] \leq \sqrt{\nu}\ln\frac{t_1}{t_0}$. With this upper bound for $\rho[x(\cdot), t_0, t_1]$, the first inequality in (5.24) implies (5.23). As it is shown in Example 5.1, the ratio of the right-hand side of (5.24) to that of (5.23) can be as small as $O(\frac{1}{\nu^{1/2}})$.

Bound (5.24) allows to obtain certain result on suboptimality of the path-following method in the family of all short-step interior point methods associated with the same self-concordant barrier. Namely, consider a segment $t \in [t_0, t_1]$, $t_0 > 0$, $t_1 < \infty$, of the central path and assume that at certain moment we stand at the point $x(t_0)$. Starting from this moment, the path-following method reaches the point $x(t_1)$ with the value of the objective $\langle c, x(t_1)\rangle < \langle c, x(t_0)\rangle$ in $O(\rho[x(\cdot), t_0, t_1])$ Newton steps. Let us ask ourselves whether it is possible to reach a point with the value of the objective at least $\langle c, x(t_1)\rangle$ by a short-step interior point method, associated with f, for which the number of iterations is essentially smaller than $O(\rho[x(\cdot), t_0, t_1])$. As it is shown in [6], the number of steps of any competing method is bounded below by $O(\sigma(x(t_0), Q_{t_1}))$, where

$$Q_{t_1} = \bigl\{x \in Q \mid \langle c, x\rangle \leq \bigl\langle c, x(t_1)\bigr\rangle\bigr\}.$$

Thus, the aforementioned question can be posed as follows:

(Q) *How large could be the ratio* $\rho[x(\cdot), x(t_0), x(t_1)]/\sigma(x(t_0), Q_{t_1})$*?*

The answer is given by the following theorem.

Theorem 5.3 *Assume that the ellipsoid* $W_{\frac{1}{10}}(x(t_0))$ *does not intersect* Q_{t_1}.[3] *Then*

$$\rho\big[x(\cdot), t_0, t_1\big] \le O(1)\nu^{1/4}\sqrt{\sigma\big(x(t_0), Q_{t_1}\big)\big[\sigma\big(x(t_0), Q_{t_1}\big) + \ln\nu\big]}. \tag{5.25}$$

Proof (1) Let $H = \{x \in Q \mid \langle c, x\rangle = \langle c, x(t_1)\rangle\}$. Since $x(t_0) \in Q\backslash Q_{t_1}$, we clearly have

$$\sigma\big(x(t_0), H\big) \le \sigma\big(x(t_0), Q_{t_1}\big). \tag{5.26}$$

Note that f attains its minimum on H at the point $x(t_1)$, so that

$$\big\langle f'\big(x(t_1)\big), x - x(t_1)\big\rangle = 0, \quad \forall x \in H. \tag{5.27}$$

Since $f'(x(t_1)) = -t_1 c$, $f'(x(t_0)) = -t_0 c$ and $\langle c, x(t_1)\rangle < \langle c, x(t_0)\rangle$, we have also

$$\big\langle f'\big(x(t_0)\big), x - x(t_0)\big\rangle = \big\langle f'\big(x(t_0)\big), x(t_1) - x(t_0)\big\rangle \ge 0, \quad \forall x \in H. \tag{5.28}$$

(2) Let $x \in H$. Since f attains its minimum on H at the point $x(t_1)$ and in view of (3.5), we have

$$f\big(x(t_1)\big) - f\big(x(t_0)\big) \le f(x) - f\big(x(t_0)\big) \le \sqrt{\nu}\sigma\big(x(t_0), x\big). \tag{5.29}$$

Furthermore, by (3.4)

$$\begin{aligned}\big\| f'(x)\big\|^*_{x(t_0)} + 1 &\le \exp\big\{\sigma\big(x(t_0), x\big)\big\}\big(\big\| f'\big(x(t_0)\big)\big\|^*_{x(t_0)} + 1\big) \\ &\le \exp\big\{\sigma\big(x(t_0), x\big)\big\}\big(\sqrt{\nu} + 1\big).\end{aligned} \tag{5.30}$$

Using (2.12) with $u = x(t_0)$ and $v = x(t_1)$ we get

$$\big\| f'\big(x(t_1)\big)\big\|^*_{x(t_0)} \le \frac{\nu}{1 - \pi_{x(t_0)}(x(t_1))}$$

[(2.14) with $u = x(t_0)$, $v = x(t_1)$, $w = x$ and (5.27)]
$$\le \frac{\nu(1 + \nu + 2\sqrt{\nu})}{1 - \pi_{x(t_0)}(x)}$$

[(2.13) with $u = x(t_0)$, $v = x$ and (5.28)]
$$\le \frac{\nu(\nu + 2\sqrt{\nu})(1 + \nu + 2\sqrt{\nu})}{\pi_{x(t_0)}(x)}\big\| f'(x)\big\|^*_{x(t_0)}$$

$$\le O(1)\nu^4\big\| f'(x)\big\|^*_{x(t_0)},$$

[3] It is easily seen that the case when $W_{\frac{1}{10}}(x(t_0))$ intersects Q_{t_1} is trivial: $\rho[x(\cdot), t_0, t_1] \le O(1)$.

where the concluding inequality follows from the fact that, on one hand, $W_{\frac{1}{10}}(x(t_0))$ does not intersect H, whence $\|x - x(t_0)\|_{x(t_0)} \geq \frac{1}{10}$, and, on the other hand, the $\|\cdot\|_{x(t_0)}$-distance from $x(t_0)$ to the boundary of Q in the direction $x - x(t_0)$ does not exceed $\nu + 2\sqrt{\nu}$ in view of (5.28) and Proposition 2.1(v).

Thus,

$$\big\| f'\big(x(t_1)\big)\big\|^*_{x(t_0)} \leq O(1)\nu^4 \big\| f'(x)\big\|^*_{x(t_0)} \leq O(1)\nu^5 \exp\big\{\sigma\big(x(t_0), x\big)\big\}, \tag{5.31}$$

the concluding inequality being given by (5.30).

Since the ellipsoid $W_{\frac{1}{10}}(x(t_0))$ does not intersect Q_{t_1} and $x \in Q_{t_1}$, we have

$$\sigma\big(x(t_0), x\big) \geq O(1). \tag{5.32}$$

(3) Assume first that $\lambda(x(t_0)) \geq \frac{1}{2}$. Then by Lemma 4.1, we have

$$\rho\big[x(\cdot), t_0, t_1\big] \leq \sqrt{\big[f\big(x(t_1)\big) - f\big(x(t_0)\big)\big] \ln \frac{\|f'(x(t_1))\|^*_{x(t_0)}}{\|f'(x(t_0))\|^*_{x(t_0)}}}$$

$$[\text{by (5.29) and } \lambda(x(t_0)) \geq \tfrac{1}{2}] \leq \sqrt{\nu^{1/2}\sigma\big(x(t_0), x\big) \ln\big(2\big\| f'\big(x(t_1)\big)\big\|^*_{x(t_0)}\big)}$$

$$[\text{by (5.31) and (5.32)}] \leq O(1)\sqrt{\nu^{1/2}\sigma\big(x(t_0), x\big)\big(\sigma\big(x(t_0), x\big) + \ln \nu\big)}. \tag{5.33}$$

Now let $\lambda(x(t_0)) \leq \frac{1}{2}$. Then by (2.7), we have

$$\frac{1}{2} \leq \|g\|^*_{x(t_0)}/\|g\|^*_{x_f} \leq 2, \quad \forall g \in E^*\backslash\{0\}. \tag{5.34}$$

Now by Lemma 4.1(ii), we have

$$\rho\big[x(\cdot), t_0, t_1\big] \leq \ln 2 + \sqrt{\big[f\big(x(t_1)\big) - f\big(x(t_0)\big)\big] \ln\big(\max\big\{1, 4\big\| f'\big(x(t_1)\big)\big\|^*_{x_f}\big\}\big)}$$

$$[\text{by (5.29), (5.34)}] \leq \ln 2 + \sqrt{\nu^{1/2}\sigma\big(x(t_0), x\big) \ln\big(\max\big\{1, 8\big\| f'\big(x(t_1)\big)\big\|^*_{x(t_0)}\big\}\big)}$$

$$[\text{by (5.31), (5.32)}] \leq \ln 2 + O(1)\sqrt{\nu^{1/2}\sigma\big(x(t_0), x\big)\big(\sigma\big(x(t_0), x\big) + \ln \nu\big)}.$$

Recalling (5.32), we conclude that

$$\rho\big[x(\cdot), t_0, t_1\big] \leq O(1)\nu^{1/4}\sqrt{\sigma\big(x(t_0), x\big)\big(\sigma\big(x(t_0), x\big) + \ln \nu\big)}.$$

Since $x \in H$ is arbitrary, we get

$$\rho\big[x(\cdot), t_0, t_1\big] \leq O(1)\nu^{1/4}\sqrt{\sigma\big(x(t_0), H\big)\big(\sigma\big(x(t_0), H\big) + \ln \nu\big)}. \qquad \square$$

We conclude that the ratio in (Q) is up to logarithmic in ν terms, at most $\nu^{1/4}$.

Acknowledgements We would like to thank the anonymous referees for very useful comments and suggestions.

References

1. A. Edelman, T. Arias, S.T. Smith, The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.* **20**, 303–353 (1998)
2. S. Helgason, *Differential Geometry and Symmetric Spaces* (American Mathematical Society, Providence, 2000)
3. M. Koecher, Positivitätsbereiche im R^n. *Am. J. Math.* **79**, 575–596 (1957)
4. Yu. Nesterov, *Introductory Lectures on Convex Optimization. A Basic Course* (Kluwer, Boston, 2004)
5. Yu. Nesterov, A. Nemirovskii, *Interior Point Polynomial Methods in Convex Programming: Theory and Applications* (SIAM, Philadelphia, 1994)
6. Yu. Nesterov, M.J. Todd, On the Riemannian geometry defined by self-concordant barriers and interior-point methods. *Found. Comput. Math.* **2**(4), 333–361 (2002)
7. Yu. Nesterov, M.J. Todd, Y. Ye, Infeasible-start primal-dual methods and infeasibility detectors. *Math. Program.* **84**, 227–267 (1999)
8. O. Rothaus, Domains of positivity. *Abh. Math. Sem. Univ. Hamburg* **24**, 189–235 (1960)
9. K. Tanabe, Geometric method in nonlinear programming. *J. Optim. Theory Appl.* **30**, 181–210 (1980)
10. S.J. Wright, *Primal–Dual Interior-Point Methods* (SIAM, Philadelphia, 1997)