On the Riemannian Geometry Defined by Self-Concordant Barriers and Interior-Point Methods *†

Yu. E. Nesterov ‡and M. J. Todd§

October 29, 2002

Abstract

We consider the Riemannian geometry defined on a convex set by the Hessian of a self-concordant barrier function, and its associated geodesic curves. These provide guidance for the construction of efficient interior-point methods for optimizing a linear function over the intersection of the set with an affine manifold. We show that algorithms that follow the primal-dual central path are in some sense close to optimal. The same is true for methods that follow the shifted primal-dual central path among certain infeasible-interior-point methods. We also compute the geodesics in several simple sets.

 Foundations of Computational Mathewmatics 2 (2002), 333–361.

†AMS Classification Numbers: 52A41, 53C22, 68Q25, 90C22, 90C25, 90C51, 90C60.

‡CORE, Catholic University of Louvain, Louvain-Neuve, Belgium (nesterov@core.ucl.ac.be). This author was supported in part by the Mathematical Sciences Research Institute, UC Berkeley, USA.

§School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York 14853, USA (miketodd@cs.cornell.edu). This author was supported in part by NSF through grant DMS-9805602 and ONR through grant N00014-96-1-0050, and also by the Mathematical Sciences Research Institute, UC Berkeley, USA.

1 Introduction

In the last fifteen years, interior-point methods have become one of the most intensely studied methods for both linear programming and certain convex programming problems, because of their excellent computational and theoretical properties; see, e.g., Karmarkar [4], Nesterov and Nemirovskii [8], M. Wright [13], and S. Wright [14]. Perhaps their greatest successes have been for linear and semidefinite programming (in the latter, a linear function is optimized over the set of symmetric matrices satisfying linear equations and the further requirement that the matrix be positive semidefinite). However, the approach is applicable in principle to any finite-dimensional convex programming problem [8] (here viewed as the optimization of a linear function of a decision variable subject to linear equality constraints and the constraint that the variable lie in a closed convex set with nonempty interior). The key ingredient identified by Nesterov and Nemirovskii is the construction of a so-called self-concordant barrier function for the constraint set (without the equality constraints), whose derivatives provide provably good local ellipsoidal approximations to the constraint set that avoid the combinatorial complexities of linear programming or the nonlinear complexities in other applications.

There are many versions of interior-point methods, but all are based on removing the constraint set and adding a multiple of the barrier function to (some version of) the objective function. Then a step is taken towards a minimizer of the resulting simpler problem. A basic version of this procedure uses the objective function itself, and changes the multiple from one iteration to the next. The set of minimizers as the multiplier changes forms the so-called central path for the problem. This approach can be taken just for the original problem, giving the primal central path, or for the original problem and its dual simultaneously, yielding the primal-dual central path.

In the brief discussion above we assumed that we had available a strictly feasible point, satisfying all the linear equality constraints as well as lying in the interior of the constraint set. This is of course a strong assumption, and in practice modified methods are employed; one such uses the shifted primal-dual central path [9].

The archetypical versions of these algorithms use short steps to guarantee polynomial complexity. At each iteration, a step is taken that has some bounded length less than one in the norm defined by the Hessian of the barrier function at the current iterate. The beautiful properties of self-concordant barrier functions then assure that good progress can be made in moving towards the appropriate minimizer. Hence the algorithm as a whole takes a sequence of short steps, each measured in the local norm associated to that iterate. It is therefore natural to consider the corresponding Riemannian metric defined on the interior of the constraint set, and in particular shortest paths in this metric, or geodesic curves. Algorithms whose iterates lie on or close to such geodesics are presumably efficient. (We note that in practice longer steps are often taken if they yield good iterates, but the directions suggested by short-step methods are likely to be good for taking longer steps also.)

The first discussion of these ideas in relation to the complexity of optimization algorithms appears in Karmarkar [5] for linear programming. A different view, relating to homotopy algorithms in general, is discussed in Todd [12]. Here we consider arbitrary self-concordant barriers and the resulting Riemannian geometry, concentrating on geodesics and their relation to interior-point methods.

We note that the Riemannian geometry of homogeneous self-dual cones has been studied extensively in the mathematics literature, as it relates to the classification of these cones via Jordan algebras – see Koecher [6] and Rothaus [10] as well as Güler [2], who brought this line of research to the attention of optimizers. But their concerns were far removed from the complexity of optimization algorithms. Also, differential and Riemannian geometry has been studied in optimization in relation to the need to remain feasible – see, e.g., Tanabe [11] and Edelman et al. [1]. In this paper, however (at least in feasible methods), the algorithms move in the intersection of an affine manifold and the interior of the constraint set, and thus maintaining feasibility is not an issue. The Riemannian geometry is instead used to induce good search directions, and hopefully produce "optimal" algorithms.

In the next section we introduce self-concordant barriers, the associated Riemannian geometry, and geodesics. Section 3 introduces some terminology allowing us to relate geodesic curves to finite iterative processes. In the following section we consider curves in the direct product of several sets equipped with Riemannian metrics, and relate the resulting distances and geodesic curves to those in the constituent sets. Direct products arise in many applications, and also when considering primal-dual methods. Section 5 contains our main results. We show that the primal-dual central path comes within a factor of $\sqrt{2}$ of being geodesic, and hence that paths that follow it are close to "optimal." For infeasible-interior-point methods, we show that the shifted primal-dual central path is also within a factor of $\sqrt{2}$ of being geodesic. Finally, in Section 6 we give geodesic curves for several simple sets: the nonnegative orthant, a hypercube, the semidefinite and Lorentz cones, and the epigraph of the parabola in two dimensions, with their usual self-concordant barriers.

2 Definition and main properties

Let $Q \subset E$ be a closed convex subset of a finite-dimensional real vector space E. Let us assume that the interior of Q is nonempty and that it contains no straight line. Then, Q admits a self-concordant barrier $F(x)$ with a finite parameter value ν (Nesterov and Nemirovskii [8]). This means that F is a three-times continuously differentiable convex function defined on the interior of Q, tending to ∞ as its argument approaches a point in the boundary of Q, and satisfying the two key inequalities

$$\begin{aligned} D^3F(x)[h,h,h] &\leq 2(D^2F(x)[h,h])^{3/2} \\ \text{and} \qquad & \\ DF(x)[h] &\leq (\nu D^2F(x)[h,h])^{1/2} \end{aligned} \tag{2.1}$$

for any $x \in \operatorname{int} Q$ and $h \in E$. We call F a ν-self-concordant barrier. (In the terminology of [8], this is a ν-strongly self-concordant barrier.)

We define $F'(x) \in E^*$, the dual of E, by $\langle F'(x), h\rangle := DF(x)[h]$, $F''(x) : E \to E^*$ by $\langle F''(x)h_1, h_2\rangle := D^2F(x)[h_1,h_2]$, and $F'''(x) : E \to L(E,E^*)$ by $\langle (F'''(x)h_1)h_2, h_3\rangle := D^3F(x)[h_1,h_2,h_3]$. With a slight abuse of notation, we also write $F'''(x)[h_1,h_2,h_3]$ for $D^3F(x)[h_1,h_2,h_3]$, and $F'''(x)[h_1,h_2]$ for $(F'''(x)h_1)h_2 \in E^*$.

If Q is a cone K, then F is called ν-logarithmically homogeneous if

$$F(\tau x) = F(x) - \nu \ln \tau$$

for any $x \in \operatorname{int} Q$ and $\tau > 0$. Then, by differentiating this equation once, twice, and three times, we find

$$\begin{array}{rclcrcl} F'(\tau x) & = & \tau^{-1} F'(x), & \langle F'(x), x \rangle & = & -\nu, \\ F''(x)x & = & -F'(x), & \langle F''(x)x, x \rangle & = & \nu, \\ F'''(x)[x] & = & -2F''(x). & & & & \end{array} \tag{2.2}$$

It is easy to see from these relations that the second inequality in (2.1) always holds.

We will assume that F is smooth (infinitely differentiable). This barrier defines a local norm at any $x \in \operatorname{int} Q$:

$$\|u\|_x := \langle F''(x)u, u \rangle^{1/2}.$$

By Theorem 2.1.1 in [8], $\operatorname{int} Q$ contains the open unit ball $\{y \in E : \|y - x\|_x < 1\}$ around any such x. Let us fix two points x_0, $x_1 \in \operatorname{int} Q$. Consider a piecewise smooth curve $\xi(t)$, $t \in \Delta := [0, 1]$, such that

$$\xi(0) = x_0, \quad \xi(1) = x_1,$$

and $\xi(t) \in \operatorname{int} Q$ for all $t \in \Delta$. We call ξ a curve from x_0 to x_1. Then the Riemannian length of ξ, or the Riemannian distance between x_0 and x_1 along ξ, is defined as follows:

$$\rho_\xi(x_0, x_1) := \int_0^1 \|\xi'(t)\|_{\xi(t)} dt.$$

Let us denote the whole family of such curves by $\mathcal{C}(x_0, x_1)$. Then the Riemannian distance between x_0 and x_1 is

$$\rho(x_0, x_1) := \inf_{\xi \in \mathcal{C}(x_0, x_1)} \rho_\xi(x_0, x_1).$$

We call a curve from x_0 to $x_1 \in \operatorname{int} Q$ *normal* if

$$\|\xi'(t)\|_{\xi(t)} \text{ is constant for almost all } t \in \Delta.$$

Thus, for a normal curve ξ from x_0 to x_1 we have

$$\|\xi'(t)\|_{\xi(t)} \equiv \rho_\xi(x_0, x_1), \qquad \text{for almost all } t \in \Delta.$$

Lemma 2.1 *Any differentiable curve can be reparametrized in normal form.*

Proof: Consider a curve $\xi(t)$ with endpoints x_0 and x_1, assumed distinct. Let us define a new trajectory $\hat{\xi}(\tau(t)) := \xi(t)$, where $\tau(t)$ is a solution of the following differential equation:

$$\tau'(t) = \frac{\|\xi'(t)\|_{\xi(t)}}{\rho_\xi(x_0, x_1)}, \quad \tau(0) = 0.$$

Then $\|\hat{\xi}'(\tau)\|_{\hat{\xi}(\tau)} = \rho_\xi(x_0, x_1)$ for almost all $\tau \in \Delta$. Therefore $\hat{\xi}(1) = x_1$. □

One reason for using piecewise smooth rather than smooth paths is that reparametrization may make a smooth path only piecewise smooth. Indeed, if $\xi(t) := 1 + (1 - 2t)^2$ is a smooth path from 2 to itself in $\operatorname{int}[0, \infty)$ with the standard logarithmic barrier $F(x) := -\ln x$, then the reparametrized normal curve is $\hat{\xi}(\tau) = 2^{|1-2\tau|}$, which is not differentiable at $\tau = 1/2$. (Another reason is that two piecewise smooth paths can be concatenated to give another such.)

Henceforth, most of the curves we consider will be normal.

Theorem 2.1 *(a)* $(\operatorname{int} Q, \rho)$ *is a complete metric space.*

(b) For any x_0 *and* $x_1 \in \operatorname{int} Q$*, there is a smooth normal path* $\hat{\xi}$ *from* x_0 *to* x_1 *with* $\rho_{\hat{\xi}}(x_0, x_1) = \rho(x_0, x_1)$*.*

(c) Let $\hat{\xi}$ *be as in part (b). Then*

$$F'''(\hat{\xi}(t))\hat{\xi}''(t) + \tfrac{1}{2}F'''(\hat{\xi}(t))[\hat{\xi}'(t), \hat{\xi}'(t)] = 0, \quad \text{for all } t \in \Delta. \tag{2.3}$$

(d) Consider $\hat{\xi}(t)$ *as in part (b). If* Q *is a cone* K *and* F *is logarithmically homogeneous, then the function* $\langle F'(\hat{\xi}(t)), \hat{\xi}'(t)\rangle$ *is constant in* t*.*

Proof: That ρ defines a metric on $\operatorname{int} Q$ is a standard result; see for example Helgason [3], p. 51. The completeness is a consequence of Lemma 3.1 below, which relates $\rho(y, x)$ to $\|y - x\|_x$ when either is small (see Corollary 3.1). Then part (b) follows; see Theorem 10.4 of [3].

Part (c) is also a standard result: this turns out to be the "geodesic equation" after computing the appropriate Christoffel symbols, see [3], (3) on p. 30, in our terminology. We give an argument here for completeness. Let η be any smooth mapping from Δ to E vanishing at 0 and 1, and consider the path $\xi_s := \hat{\xi} + s\eta$ from x_0 to x_1, which lies in $\operatorname{int} Q$ for sufficiently small (in absolute value) s. Then, since $\hat{\xi}$ is the shortest path from x_0 to x_1, the derivative of $\rho_{\xi_s}(x_0, x_1)$ with respect to s at 0 must be zero, so

$$\int_0^1 \frac{1}{2\|\hat{\xi}'(t)\|_{\hat{\xi}(t)}} \frac{d}{ds}(F''(\hat{\xi}(t) + s\eta(t))[\hat{\xi}'(t) + s\eta'(t), \hat{\xi}'(t) + s\eta'(t)])\Big|_{s=0} dt = 0,$$

or

$$\int_0^1 \frac{1}{2\|\hat{\xi}'(t)\|_{\hat{\xi}(t)}} \left[F'''(\hat{\xi}(t))[\hat{\xi}'(t), \hat{\xi}'(t), \eta(t)] + 2F''(\hat{\xi}(t))[\hat{\xi}'(t), \eta'(t)]\right] dt = 0. \tag{2.4}$$

From our assumption that $\hat{\xi}$ is normal, we know that $\|\hat{\xi}'(t)\|_{\hat{\xi}(t)}$ is constant in t (not just almost everywhere, since $\hat{\xi}$ is smooth by part (b)) and equal to $\rho(x_0, x_1)$.

Now we evaluate the second term in (2.4) using partial integration, to get

$$\rho(x_0, x_1) \int_0^1 \frac{1}{\|\hat{\xi}'(t)\|_{\hat{\xi}(t)}} F''(\hat{\xi}(t))[\hat{\xi}'(t), \eta'(t)]dt = \int_0^1 F''(\hat{\xi}(t))[\hat{\xi}'(t), \eta'(t)]dt$$

$$= \left[F''(\hat{\xi}(t))[\hat{\xi}'(t)]\eta(t)\right]_0^1 - \int_0^1 \left[F'''(\hat{\xi}(t))[\hat{\xi}'(t), \hat{\xi}'(t)] + F''(\hat{\xi}(t))[\hat{\xi}''(t)]\right] \eta(t)dt.$$

The first term on the right hand side vanishes by the boundary conditions on η. Hence, from (2.4) we deduce

$$-\frac{1}{\rho(x_0, x_1)} \int_0^1 \left[F''(\hat{\xi}(t))[\hat{\xi}''(t)] + \frac{F'''(\hat{\xi}(t))[\hat{\xi}'(t), \hat{\xi}'(t)]}{2}\right] \eta(t)dt = 0. \tag{2.5}$$

Since this holds for all η satisfying the boundary conditions, by choosing the support of η to be clustered about any $t \in (0, 1)$ and $\eta(t)$ to point in an arbitrary direction, we conclude that the expression in brackets vanishes, so that (2.3) holds at this t, for all $t \in (0, 1)$.

For part (d), we use the properties $F'(x) = -F''(x)x$ and $F'''(x)x = -2F''(x)$ of logarithmically homogeneous barriers. Then we obtain

$$\begin{aligned}
&\frac{d}{dt}\langle F'(\hat{\xi}(t)), \hat{\xi}'(t)\rangle \\
&= \langle F''(\hat{\xi}(t))\hat{\xi}'(t), \hat{\xi}'(t)\rangle + \langle F'(\hat{\xi}(t)), \hat{\xi}''(t)\rangle \\
&= \langle F''(\hat{\xi}(t))\hat{\xi}'(t), \hat{\xi}'(t)\rangle + \langle -F''(\hat{\xi}(t))\hat{\xi}(t), \hat{\xi}''(t)\rangle \\
&= \langle F''(\hat{\xi}(t))\hat{\xi}'(t), \hat{\xi}'(t)\rangle - \langle F''(\hat{\xi}(t))\hat{\xi}''(t), \hat{\xi}(t)\rangle \\
&= \langle F''(\hat{\xi}(t))\hat{\xi}'(t), \hat{\xi}'(t)\rangle + \frac{1}{2}F'''(\hat{\xi}(t))[\hat{\xi}'(t), \hat{\xi}'(t), \hat{\xi}(t)] \text{ (by (2.3))} \\
&= \langle F''(\hat{\xi}(t))\hat{\xi}'(t), \hat{\xi}'(t)\rangle - \langle F''(\hat{\xi}(t))\hat{\xi}'(t), \hat{\xi}'(t)\rangle = 0 \text{ (by the last line in (2.2)).}
\end{aligned}$$

□

In particular, when Q is a cone, Item (d) of the above theorem implies that $F(x)$ is a *linear* function of time along any geodesic curve. (As pointed out by a referee, this is related to the fact that $\operatorname{int} Q$, as a Riemannian manifold, is isomorphic to the product of the hypersurface $\{x \in \operatorname{int} Q : F(x) = 0\}$ and the real axis with its usual Riemannian structure.)

Any normal curve as in part (b) above we call a *geodesic curve* from x_0 to x_1. Note that in the literature, e.g., in [3], a geodesic refers to a normal curve that satisfies the necessary condition (2.3); such curves always give shortest paths between any two sufficiently close points lying on them, but not necessarily globally – consider great circles on a sphere. We will call these curves "local geodesics," reserving the term geodesic for shortest paths, which are our main interest.

Note that the Riemannian metric has the following properties.

1. Clearly, we have the triangle inequality:

$$\rho(x_0, x_2) \le \rho(x_0, x_1) + \rho(x_1, x_2), \quad x_0, x_1, x_2 \in \operatorname{int} Q,$$

and symmetry:

$$\rho(x_0, x_1) = \rho(x_1, x_0), \quad x_0, x_1 \in \operatorname{int} Q.$$

2. The metric is invariant with respect to nonsingular linear transformations of variables. Indeed, let the matrix A be nonsingular. Then we can define the set

$$\bar{Q} := \{y : \ Ay \in Q\}.$$

For this set a natural barrier function is as follows:

$$\bar{F}(y) := F(Ay).$$

Then it is easy to see that the Riemannian length of any curve $\xi(t)$ in $\bar{Q}$ using the barrier $\bar{F}$ is the same as the Riemannian length of the curve $A\xi(t)$ in Q using the barrier F. If A is an automorphism of Q such that $\bar{F}$ differs by a constant from F, then we have just one Riemannian metric defined on $\operatorname{int} Q = \operatorname{int} \bar{Q}$ using either F or $\bar{F}$, and it is invariant under transformation of its endpoints by A; that is, A is an isometry.

3. If $Q = K$ is a convex cone, equipped with a logarithmically homogeneous barrier, then the Riemannian distance is homogeneous of degree zero:

$$\rho(\tau x_0, \tau x_1) = \rho(x_0, x_1), \quad \tau > 0.$$

4. The Riemannian metric is also dual invariant. Namely, consider the conjugate function

$$F_*(s) = \max_s [-\langle s, x\rangle - F(x)].$$

Denote by $Q^* \subseteq E^*$ its domain (i.e., the set of all s satisfying $F_*(s) < \infty$). Then, using the local norm defined by $F_*''(s)$, $s \in \operatorname{int} Q^*$, we can define a Riemannian metric ρ_* in Q^*.

Proposition 2.1 *(a) For any $x_0, x_1 \in \operatorname{int} Q$ we have*

$$\rho(x_0, x_1) = \rho_*(-F'(x_0), -F'(x_1)).$$

(b) If $\xi(t)$ is a geodesic in Q, then $\xi_*(t) = -F'(\xi(t))$ is a geodesic in Q^*. Moreover,

$$F''(\xi(t))\xi''(t) = \xi_*''(t), \quad t \in \Delta.$$

Proof: Note that the derivatives of the primal and dual barrier are related in the following way:

$$\begin{aligned} F_*'(-F'(x)) &= -x, \\ F_*''(-F'(x)) &= [F''(x)]^{-1}, \\ F_*'''(-F'(x))[h] &= [F''(x)]^{-1} F'''(x) \left[[F''(x)]^{-1} h\right] [F''(x)]^{-1}. \end{aligned} \tag{2.6}$$

Let $\xi(t)$ be an arbitrary curve in $\operatorname{int} Q$. Denote $\xi_*(t) = -F'(\xi(t))$. Then

$$\xi_*'(t) = -F''(\xi(t))\xi'(t), \tag{2.7}$$

$$\xi_*''(t) = -F''(\xi(t))\xi''(t) - F'''(\xi(t))[\xi'(t), \xi'(t)]. \tag{2.8}$$

The equation (2.7) shows that

$$\begin{aligned} \|\xi_*'(t)\|^2_{\xi_*(t)} &= \langle F''(\xi(t))\xi'(t), F_*''(-F'(\xi(t)))F''(\xi(t))\xi'(t)\rangle \\ &= \langle F''(\xi(t))\xi'(t), [F''(\xi(t))]^{-1} F''(\xi(t))\xi'(t)\rangle = \|\xi'(t)\|^2_{\xi(t)}. \end{aligned}$$

Hence $\rho^*_{\xi_*}(-F'(x_0), -F'(x_1)) = \rho_\xi(x_0, x_1)$. We also note that this process is reversible: for any curve $\xi_*(t)$ in $\operatorname{int} Q^*$, we can define $\xi(t) := -F_*'(\xi_*(t))$ and the above analysis still holds. That proves part (a) of the proposition, and the first part of (b).

Now let $\xi(t)$ be a geodesic. Then (2.8) and (2.3) give

$$\xi_*''(t) = -F''(\xi(t))\xi''(t) - F'''(\xi(t))[\xi'(t), \xi'(t)] = F''(\xi(t))\xi''(t).$$

□

If Q is a cone K and F is ν-logarithmically homogeneous, then so is F_* on $Q^* = K^* := \{s \in E^* : \langle s, x\rangle \geq 0 \text{ for all } x \in K\}$. Moreover, we have

$$F_*(-F'(x)) = -\langle -F'(x), x\rangle - F(x) = -F(x) - \nu \tag{2.9}$$

in this case.

5. Sometimes the following trivial lower bound is useful.

Lemma 2.2 *For any $x_0, x_1 \in \operatorname{int} Q$ we have*

$$\rho(x_0, x_1) \geq \frac{1}{\sqrt{\nu}} \mid F(x_0) - F(x_1) \mid .$$

Proof: Indeed, since $F(x)$ is a self-concordant barrier, for any curve $\xi(t)$ in Q the Cauchy-Schwarz inequality yields:

$$\begin{aligned} \|\xi'(t)\|_{\xi(t)} &\geq \frac{|\langle F'(\xi(t)), \xi'(t)\rangle|}{\langle [F''(\xi(t))]^{-1} F'(\xi(t)), F'(\xi(t))\rangle^{1/2}} \geq \frac{1}{\sqrt{\nu}} \mid \langle F'(\xi(t)), \xi'(t)\rangle \mid \\ &\geq \pm \frac{1}{\sqrt{\nu}} (F(\xi(t))'_t. \end{aligned}$$

□

3 Some terminology

The main goal of this paper is to provide further justification for some trajectories and iterative schemes used in interior-point methods. The existing complexity analyses provide strong motivation: our results show that, in some sense, they cannot be improved. First of all let us show how the Riemannian metric can be used in the analysis of iterative proceses. We need the following lemma, which also enables us to show that our Riemannian metric space is complete.

Lemma 3.1 *Let $x, y \in \operatorname{int} Q$ with $r := \|y - x\|_x < 1$. Then*

$$r - \tfrac{1}{2} r^2 \leq \rho(x, y) \leq -\ln(1 - r). \tag{3.10}$$

Also, if $\rho(x, y) \leq \kappa - \frac{1}{2}\kappa^2$ for some $\kappa \in [0, 1)$, then $\|y - x\|_x \leq \kappa < 1$ and hence the bounds above hold.

Proof: Consider the trajectory $\eta(t) := x + t(y - x)$, $t \in \Delta$. Then

$$\begin{aligned} \rho(x, y) &\leq \int_0^1 \|\eta'(t)\|_{\eta(t)} dt = \int_0^1 \langle F''(\eta(t))(y - x), y - x\rangle^{1/2} dt \\ &\leq \int_0^1 \frac{\langle F''(x)(y-x), y-x\rangle^{1/2}}{1 - \|\eta(t) - x\|_x} dt = \int_0^1 \frac{r}{1 - tr} dt = -\ln(1 - r), \end{aligned}$$

where the second inequality follows from Theorem 2.1.1 in [8].

In order to prove the lower bound, consider any geodesic ξ from x to y. Denote by $\bar{t}$ the moment when this curve hits the boundary of the ellipsoid $\{z : \| z - x\|_x \leq r\}$ for the first time. Clearly, such $\bar{t}$ is well defined since $\|\xi(1) - x\|_x = r$. Then, for the function $\delta(t) = \|\xi(t) - x\|_x$ we have by the Cauchy-Schwarz inequality

$$\delta'(t) = \frac{1}{\delta(t)} \langle F''(x)(\xi(t) - x), \xi'(t)\rangle \leq \|\xi'(t)\|_x.$$

Therefore, again using Theorem 2.1.1 in [8], we have

$$\begin{aligned}\rho(x,y) &\geq \int_0^{\bar{t}} \|\xi'(t)\|_{\xi(t)} dt \geq \int_0^{\bar{t}} (1-\delta(t))\delta'(t) dt \\ &= \delta(\bar{t}) - \tfrac{1}{2}\delta^2(\bar{t}) = r - \tfrac{1}{2}r^2.\end{aligned}$$

This completes the proof of the first part.

For the last part, assume to the contrary that $r := \|y-x\|_x > \kappa$, and choose $\kappa' \in (\kappa, \min\{1,r\})$. Then any geodesic from x to y must hit the boundary of the ellipsoid $\{z : \| z-x\|_x \leq \kappa'\}$, say for the first time at y'. Hence, by the first part of the lemma, since $\|y'-x\|_x = \kappa' < 1$, $\rho(x,y') \geq \kappa' - \frac{1}{2}(\kappa')^2 > \kappa - \frac{1}{2}\kappa^2$. But this contradicts $\rho(x,y') \leq \rho(x,y) \leq \kappa - \frac{1}{2}\kappa^2$. □

Corollary 3.1 $(\operatorname{int} Q, \rho)$ *is a complete metric space*

Proof: Let $\{x_k\}$ be a Cauchy sequence, and choose K so that $\rho(x_l, x_m) \leq \frac{3}{8}$ for all $l,m \geq K$. Then $\|x_m - x_l\|_{x_l} \leq 1/2$ for such l,m from the lemma, and in particular, $\|x_k - x_K\|_{x_K} \leq 1/2$ for all $k \geq K$. This implies that the sequence has at least one limit point. If there were two distinct limit points, then it would be easy to reach a contradiction by choosing points x_l and x_m, with l,m arbitrarily large, arbitrarily close to these two limit points, and using similar arguments. □

Let us introduce now the notion of a short-step sequence. Let $X_N \equiv \{x_k\}_{k=0}^N \subset \operatorname{int} Q$ be a finite sequence. We call the value N the *length* of the sequence X_N.

Definition 3.1 *The sequence X_N is called a* κ-short-step *sequence if its* upper step length

$$\kappa(X_N) := \max_{0 \leq k \leq N-1} \|x_{k+1} - x_k\|_{x_k}$$

is bounded away from 1 by κ: $\kappa(X_N) \leq \kappa < 1$.

We call a short-step sequence X_N α-regular *if its* lower step length

$$\alpha(X_N) := \min_{1 \leq k \leq N-2} \|x_{k+1} - x_k\|_{x_k}$$

is bounded away from 0 by α: $\alpha(X_N) \geq \alpha > 0$. (We allow an exception at the endpoints because it may be necessary to allow small steps in the neighbourhood of x_0 and x_N.)

Our interest in sequences of this type is motivated by the fact that the best complexity results in interior-point theory are justified by short-step methods. Moreover, long-step methods are often motivated by short-step algorithms, maybe using the same directions but allowing line searches. Hence analysis for short-step methods can also help in designing better long-step approaches. In our terminology, such methods generate regular short-step sequences.

Let us relate the length of a short-step sequence with the Riemannian distance between its initial and final points.

Lemma 3.2 *For any κ-short-step sequence X_N we have*

$$\rho(x_0, x_N) \leq -N \ln(1-\kappa).$$

Proof: In view of Lemma 3.1 we have:

$$\rho(x_0, x_N) \leq \sum_{i=0}^{N-1} \rho(x_i, x_{i+1}) \leq -N \ln(1-\kappa(X_N)).$$

□

This trivial result has an important consequence.

Corollary 3.2 *The length of an arbitrary κ-short-step sequence X_N going from $x \in \operatorname{int} Q$ to $y \in \operatorname{int} Q$ is bounded below:*

$$N \geq \frac{\rho(x,y)}{-\ln(1-\kappa)}.$$

□

Conversely, we have

Lemma 3.3 *For any $\kappa \in (0,1)$ and any x, $y \in \operatorname{int} Q$, there is an α-regular κ-short-step sequence X_N from x to y with $\alpha = \left(1 - e^{-1/2}\right)\kappa$, whose length can be bounded as follows:*

$$N \leq \frac{\rho(x,y)}{\kappa - \frac{1}{2}\kappa^2} + 2.$$

Proof: Divide any geodesic from x to y into segments using points x_1, $x_2, \ldots, x_{N-1}$ so that, with $x_0 := x$ and $x_N := y$, $\rho(x_k, x_{k+1}) \leq \kappa - \frac{1}{2}\kappa^2$ for $k = 0, \ldots, N-1$, with equality for all except possibly $k = N-1$. Clearly N can be bounded as stated. Moreover, by Lemma 3.1, for $r_k = \|x_k - x_{k+1}\|_{x_k}$ we have $r_k \leq \kappa$ and, for $k < N-1$, $\kappa - \frac{1}{2}\kappa^2 \leq -\ln(1-r_k)$ which gives

$$r_k \geq 1 - e^{-\kappa + \frac{1}{2}\kappa^2} \geq \left(1 - e^{-1/2}\right)\kappa.$$

□

The above inequalities suggest the following definition of optimal sequences.

Definition 3.2 *Given positive absolute constants c_1 and c_2, we call a short-step sequence X_N (c_1, c_2)-*optimal *(or just* optimal*) for going from x_0 to x_N if*

$$N \leq c_1 \cdot \rho(x_0, x_N) + c_2.$$

Many interior-point schemes are based on approximately following certain curves in the feasible set. It is convenient to have some relations between the length of such sequences and the Riemannian length of the curves.

Lemma 3.4 *Let an α-regular short-step sequence X_N follow a curve $\xi(t)$ at the proximity level β, i.e.,*

$$\|x_k - \xi(t_k)\|_{x_k} \leq \beta, \quad k = 0, \ldots, N,$$

for some $t_0 < t_1 < \ldots < t_N$, where $\beta \in [0,1)$. Then

$$\rho_\xi(\xi(t_0), \xi(t_N)) \geq (N-1)[\alpha - \tfrac{1}{2}\alpha^2 + 2\ln(1-\beta)].$$

Proof: Indeed,

$$\begin{aligned}\rho_\xi(\xi(t_0),\xi(t_N)) &= \sum_{i=0}^{N-1} \rho_\xi(\xi(t_i),\xi(t_{i+1})) \\ &\geq \sum_{i=0}^{N-2} \rho(\xi(t_i),\xi(t_{i+1})) \\ &\geq \sum_{i=0}^{N-2} [\rho(x_i,x_{i+1}) - \rho(\xi(t_i),x_i) - \rho(x_{i+1},\xi(t_{i+1}))] \\ &\geq (N-1)[\alpha - \tfrac{1}{2}\alpha^2 + 2\ln(1-\beta)].\end{aligned}$$

□

Thus, we can estimate the length of a regular short-step path-following sequence by the Riemannian length of the corresponding curve. Therefore we need the following definition.

Definition 3.3 *A curve $\xi(t)$, $t \in \Delta$ is called* τ-geodesic *for some $\tau \geq 1$ if*

$$\rho_\xi(\xi(t_0),\xi(t_1)) \leq \tau \cdot \rho(\xi(t_0),\xi(t_1))$$

for any $t_0, t_1 \in \Delta$. If τ is an absolute constant, then we call the curve sub-geodesic.

Note that $\tau = 1$ corresponds to a geodesic curve.

Finally, we establish the following result.

Theorem 3.1 *Let an α-regular short-step sequence X_N follow a τ-geodesic curve at the proximity level β. Assume that these parameters satisfy the relations:*

$$\alpha - \tfrac{1}{2}\alpha^2 + 2\ln(1-\beta) \geq \gamma_1, \quad \beta \leq 1 - \gamma_2$$

where τ, γ_1, and γ_2 are positive absolute constants. Then X_N is (c_1, c_2)-optimal for going from x_0 to x_N for $c_1 := \tau/\gamma_1$ and $c_2 := 1 - 2c_1 \ln \gamma_2$.

Proof: Indeed, let X_N follow a τ-geodesic curve $\xi(t)$. Then

$$N \leq \frac{\rho_\xi(\xi(t_0),\xi(t_N))}{\alpha - \frac{1}{2}\alpha^2 + 2\ln(1-\beta)} + 1 \leq \frac{\tau}{\gamma_1}\rho(\xi(t_0),\xi(t_N)) + 1.$$

It remains to note that $\rho(x_i,\xi(t_i)) \leq -\ln(1-\beta) \leq -\ln\gamma_2$ for $i = 0$ and $i = N$ and use $\rho(\xi(t_0),\xi(t_N)) \leq \rho(x_0,x_N) + \rho(x_0,\xi(t_0)) + \rho(x_N,\xi(t_N))$ from the triangle inequality. □

4 Curves in direct products

In what follows we often work with curves in a direct product of convex sets. Therefore we need to present some properties of the Riemannian metric in such a product. Note that in this section, subscripts refer to the components of a point or a curve in the different spaces rather than points in a single space in some sequence.

Normal curves allow us to compute easily the metric in the direct product of sets:

Lemma 4.1 *Let the Riemannian metrics ρ_i in the convex sets Q_i be defined from the self-concordant barriers F_i, $i = 1, \ldots, m$, and let the Riemannian metric ρ in the convex set $Q := \Pi_{i=1}^m Q_i$ be defined from the self-concordant barrier $F(z) := \sum_{i=1}^m F_i(z_i)$, $z = (z_1, \ldots, z_m) \in \operatorname{int} Q$. Consider m normal curves $\xi_i(t)$ connecting x_i, $y_i \in \operatorname{int} Q_i$, and let $x = (x_1, \ldots, x_m)$, $y = (y_1, \ldots, y_m)$.*

(a) The curve $\xi(t) = (\xi_1(t), \ldots, \xi_m(t))$ from x to y is normal for the set $Q = \Pi_{i=1}^m Q_i$, with respect to the barrier F. Moreover, we have

$$\rho_\xi(x, y) = \left[\sum_{i=1}^m \rho_{\xi_i}^2(x_i, y_i)\right]^{1/2}.$$

(b) $\rho(x, y) = \left[\sum_{i=1}^m \rho_i^2(x_i, y_i)\right]^{1/2}$.

(c) If the curves $\xi_i(t)$ are τ-geodesic in Q_i, for any $\tau \geq 1$, then the curve $\xi(t)$ is τ-geodesic in Q.

Proof: Indeed, since all ξ are normal, we have:

$$\begin{aligned} \|\xi'(t)\|_{\xi(t)}^2 &= \sum_{i=1}^m \langle F_i''(\xi_i(t))\xi_i'(t), \xi_i'(t)\rangle \\ &= \sum_{i=1}^m \rho_{\xi_i}^2(x_i, y_i), \end{aligned}$$

which is independent of t, for almost all $t \in \Delta$.

For part (b), consider any path η from x to y in $\operatorname{int} Q$ and let η_i denote its components in the Q_i's. Let (v_i) denote the m-vector with components v_i. Then, with obvious notation,

$$\begin{aligned} \rho_\eta(x, y) &= \int \|\eta'(t)\|_{\eta(t)} dt \\ &= \int \|(\|\eta_i'(t)\|_{\eta_i(t)})\|_2 dt \\ &\geq \|\int (\|\eta_i'(t)\|_{\eta_i(t)}) dt\|_2 \\ &\geq \|(\rho_i(x_i, y_i))\|_2, \end{aligned}$$

so that $\rho(x, y) \geq \|(\rho_i(x_i, y_i))\|_2$. However, the first statement shows that we achieve equality when each η_i is a geodesic curve from x_i to y_i in $\operatorname{int} Q_i$. This proves part (b).

The third statement is evident. □

For a geodesic curve we can prove a converse statement.

Lemma 4.2 *Let the Riemannian metrics in the convex sets Q_i be defined from the self-concordant barriers F_i, $i = 1, \ldots, m$, and similarly in the set $Q = \Pi_{i=1}^m Q_i$ equipped with the barrier*

$$F(z) = \sum_{i=1}^m F_i(z_i), \quad z = (z_1, \ldots, z_m) \in \operatorname{int} Q.$$

Then any geodesic curve $\xi(t) = (\xi_1(t), \ldots, \xi_m(t))$ in Q is composed of curves $\xi_i(t)$ that are geodesic in Q_i.

Proof: Let ξ go from x to y, and ξ_i from x_i to y_i for each i. Then, with the same notation as in the previous lemma,

$$\begin{aligned}
\rho(x,y) &= \int \|\xi'(t)\|_{\xi(t)} dt \\
&= \int \|(\|\xi_i'(t)\|_{\xi_i(t)})\|_2 dt \\
&\geq \|(\int \|\xi_i'(t)\|_{\xi_i(t)}) dt\|_2 \\
&\geq \|(\rho_i(x_i, y_i))\|_2 \\
&= \rho(x,y).
\end{aligned}$$

Hence we have equality throughout, which implies that the curves $\xi_i(t)$ are geodesic in Q_i. □

Although we have stated them for Riemannian metrics derived from self-concordant barrier functions, Lemmas 4.1 and 4.2 hold for general products of Riemannian manifolds.

In the following, we are interested in the product of just two Riemannian manifolds, corresponding to the interiors of the primal and dual cones in a primal-dual pair of optimization problems, and in the primal, dual, and primal-dual central paths. It is important to note that, even in this restricted setting, the conclusion of Lemma 4.2 fails to hold for sub-geodesic curves (see Section 6.1). Some intuition that the projections of sub-geodesic curves can be arbitrarily bad follows from the following result.

Theorem 4.1 *Let K be a closed, pointed, and solid cone equipped with a ν-self-concordant logarithmically homogeneous barrier $F(x)$. Let $x(t)$, $t \in \Delta$, be an arbitrary piecewise smooth primal curve in* $\operatorname{int} K$. *Define a dual curve by*

$$s(t) := -\omega(t) F'(x(t))$$

in the dual cone K^*, where $\omega(t)$ is a smooth strictly positive function whose derivative is positive (or negative) in Δ.

If the curves $x(t)$ and $s(t)$ are orthogonal:

$$\langle s'(t), x'(t) \rangle = 0, \quad \text{for almost all } t \in \Delta, \tag{4.11}$$

then the primal-dual curve $z(t) := (x(t), s(t))$ is $\sqrt{2}$-geodesic in $\hat{K} := K \times K^*$ with respect to the barrier

$$\hat{F}(z) := F(x) + F_*(s).$$

The length of this curve is given by the following expression:

$$\rho_{z(\cdot)}(z(t_0), z(t_1)) = \sqrt{\nu} \mid \ln \frac{\omega(t_1)}{\omega(t_0)} \mid, \quad t_0, t_1 \in \Delta.$$

Proof: Without loss of generality, we can assume that $\omega'(t) > 0$, $t \in \Delta$. Note that (almost everywhere in Δ)

$$s'(t) = -\omega'(t)F'(x(t)) - \omega(t)F''(x(t))x'(t),$$

which gives

$$F'(x(t)) = -\frac{1}{\omega'(t)}s'(t) - \frac{\omega(t)}{\omega'(t)}F''(x(t))x'(t). \tag{4.12}$$

At the same time, since $F(x)$ is logarithmically homogeneous, we have $x(t) = -\omega(t)F'_*(s(t))$ (see (2.6)). Thus,

$$F'_*(s(t)) = -\frac{1}{\omega'(t)}x'(t) - \frac{\omega(t)}{\omega'(t)}F''_*(s(t))s'(t). \tag{4.13}$$

Further,

$$\hat{F}(z(t)) = F(x(t)) + F_*(s(t)) = F(x(t) + F_*(-\omega(t)F'(x(t)) = -\nu - \nu\ln\omega(t)$$

using (2.9). Therefore,

$$\langle F'(x(t)), x'(t)\rangle + \langle s'(t), F'_*(s(t))\rangle = -\nu\frac{\omega'(t)}{\omega(t)}.$$

Using now the relations (4.12), (4.13) and (4.11), we get the following:

$$\begin{aligned}
\nu\tfrac{\omega'(t)}{\omega(t)} &= \langle \tfrac{1}{\omega'(t)}s'(t) + \tfrac{\omega(t)}{\omega'(t)}F''(x(t))x'(t), x'(t)\rangle + \langle s'(t), \tfrac{1}{\omega'(t)}x'(t) + \tfrac{\omega(t)}{\omega'(t)}F''_*(s(t))s'(t)\rangle \\
&= \tfrac{\omega(t)}{\omega'(t)}\left(\langle F''(x(t))x'(t), x'(t)\rangle + \langle s'(t), F''_*(s(t))s'(t)\rangle\right) \\
&= \tfrac{\omega(t)}{\omega'(t)}\|z'(t)\|^2_{z(t)}.
\end{aligned}$$

Thus, $\|z'(t)\|_{z(t)} = \sqrt{\nu}\omega'(t)/\omega(t)$. Note that $\hat{F}(z)$ is a 2ν-self-concordant barrier. Therefore, in view of Lemma 2.2, for any t_0, $t_1 \in \Delta$, $t_0 < t_1$, we obtain:

$$\begin{aligned}
\rho_z(z(t_0), z(t_1)) &= \int\limits_{t_0}^{t_1} \|z'(t)\|_{z(t)}dt = \sqrt{\nu}[\ln\omega(t_1) - \ln\omega(t_0)] \\
&= \sqrt{\nu}\left[\tfrac{1}{\nu}\left(\hat{F}(z(t_0)) - \hat{F}(z(t_1))\right)\right] = \sqrt{2}\left[\tfrac{1}{\sqrt{2\nu}}\left(\hat{F}(z(t_0)) - \hat{F}(z(t_1))\right)\right] \\
&\leq \sqrt{2}\rho(z(t_0), z(t_1)).
\end{aligned}$$

□

Note that the condition (4.11) in Theorem 4.1 can be achieved for a rather wide family of curves in K just by choosing an appropriate function $\omega(t)$.

Proposition 4.1 *Let $x(t)$ be a smooth curve in $\operatorname{int} K$ along which the derivative of $F(x(t))$ is always positive (or always negative). Define the function $\omega(t)$ as a solution of the following differential equation:*

$$\frac{\omega'(t)}{\omega(t)} = -\frac{\langle F''(x(t))x'(t), x'(t)\rangle}{\langle F'(x(t)), x'(t)\rangle}.$$

Then for the dual trajectory $s(t) = -\omega(t)F'(x(t))$ *we have*

$$\langle s'(t), x'(t) \rangle \equiv 0.$$

Proof: Indeed, $s'(t) = -\omega'(t)F'(x(t)) - \omega(t)F''(x(t))x'(t)$. □

5 Lower bounds on the complexity of short-step methods

Now we are ready to analyse some important trajectories in interior-point path-following schemes. Let us start with the primal-dual central path.

5.1 Feasible primal-dual central path

Consider the following conic minimization problem:

$$\begin{aligned} \min_x \quad & \langle c, x \rangle, \\ \text{s.t.} \quad & Ax = b \in R^m, \\ & x \in K \subset R^n, \end{aligned} \tag{5.14}$$

where K is a closed, pointed, and solid convex cone and the matrix A has full row rank. Let $F(x)$ be a ν-self-concordant logarithmically homogeneous barrier for K. The *primal central path* is the set of points $x(t)$ defined as follows:

$$x(t) := \arg\min_x \{t\langle c, x \rangle + F(x) : \ Ax = b\}, \ t \in R_+.$$

The problem dual to (5.14) has the following form:

$$\begin{aligned} \max_{s,y} \quad & \langle b, y \rangle, \\ \text{s.t.} \quad & A^T y + s = c, \\ & s \in K^*. \end{aligned}$$

For this problem we also can introduce a central path, consisting of the points

$$(s(t), y(t)) := \arg\min_{s,y} \{-t\langle b, y \rangle + F_*(s) : \ A^T y + s = c, \ y \in R^m\}, \ t \in R_+.$$

Now we can consider the following primal-dual problem:

$$\begin{aligned} \min_{x,s,y} \quad & \langle c, x \rangle - \langle b, y \rangle, \\ \text{s.t.} \quad & Ax = b, \\ & A^T y + s = c, \\ & x \in K, \ s \in K^*, \ y \in R^m. \end{aligned} \tag{5.15}$$

Then, the *primal-dual central path* is just the concatenation of the primal and dual paths: the set of points

$$z(t) := (x(t), s(t)), \quad t \in R_+.$$

Here we can omit the component $y(t)$ since it can be found from the system of linear equations of the primal-dual problem.

Note that $z(t) \in \operatorname{int} \hat{K} = \operatorname{int}(K \times K^*)$. The barrier for the cone $\hat{K}$ is

$$\hat{F}(z) := F(x) + F_*(s), \quad z := (x, s) \in \operatorname{int} \hat{K}.$$

Note that the value of parameter of this barrier is $\hat{\nu} = 2\nu$.

Let us present a well-known duality theorem (see, for example, [7]).

Theorem 5.1 *(a) If the primal-dual feasible set in (5.15) has a strictly feasible primal-dual point (i.e., $x \in \operatorname{int} K$ and $s \in \operatorname{int} K^*$), then the optimal value of this problem is zero and the primal-dual central path is well defined.*

(b) Along the primal-dual central path we have the following identity:

$$s(t) = -\frac{1}{t} F'(x(t)). \tag{5.16}$$

(c) $z(t) = \arg \min_{z=(x,s),y} \{\hat{F}(z) : \ Ax = b, \ A^T y + s = c, \ \langle c, x \rangle - \langle b, y \rangle \leq \frac{\nu}{t}\}$.

□

Now we can use Theorem 4.1 to prove the following statement.

Theorem 5.2 *Any segment of the primal-dual central path is $\sqrt{2}$-geodesic in $\hat{K}$ with respect to $\hat{F}$.*

Proof: Indeed, in view of the system of linear equations in (5.15) we have

$$Ax'(t) = 0, \quad A^T y'(t) + s'(t) = 0.$$

Therefore

$$\langle s'(t), x'(t) \rangle = -\langle A^T y'(t), x'(t) \rangle = -\langle Ax'(t), y'(t) \rangle = 0,$$

and, noting (5.16), we can apply Theorem 4.1. □

Corollary 5.1 *If our starting point is close to the primal-dual central path, then any regular short-step central-path-following scheme is optimal for traveling to the set*

$$\{z = (x, s) \in \hat{K} : \ \text{for some } y, \ Ax = b, \ A^T y + s = d, \ \langle c, x \rangle - \langle b, y \rangle \leq \epsilon\}.$$

Proof: The proof follows from Theorem 5.2, Theorem 3.1, and part (c) of Theorem 5.1. □

Note that the statement of the theorem is applicable to *all* sequences in $\hat{K}$; they are allowed to be infeasible for the system of linear equations of the problem (5.15).

Unfortunately, we do not see how to prove such a result for the primal central path with respect to the cone K.

5.2 Shifted primal-dual central path

This path arises in infeasible-start interior-point primal-dual schemes. The idea of this approach (see [9]) is to solve the following minimization problem:

$$\min \quad F(x) + F_*(s) - \ln\tau - \ln\kappa,$$

$$(x, y, s, \tau, \kappa) \in \mathcal{F} := \begin{cases} Ax = r_x + \tau b, \\ A^T y + s = r_s + \tau c, \\ \langle c, x\rangle - \langle b, y\rangle + \kappa = r_c, \\ x \in K,\ s \in K^*,\ y \in R^m,\ \tau \geq 0,\ \kappa \geq 0. \end{cases} \tag{5.17}$$

It can be proved that the feasible set of this problem $\mathcal{F}$ is always unbounded. Therefore, the minimization of the self-concordant barrier for this set (the objective function in (5.17)) gives in the limit a recession direction $(\bar{x}, \bar{y}, \bar{s}, \bar{\tau}, \bar{\kappa})$. If in this direction the component $\bar{\tau}$ is positive, then the point $(\bar{x}/\bar{\tau}, \bar{y}/\bar{\tau}, \bar{s}/\bar{\tau})$ is an optimal solution to the primal-dual problem (5.15). If the component $\bar{\kappa}$ is positive, then we can get some infeasibility certificates for (5.15) (see [9]) for details).

The advantage of the formulation (5.17) is that we can always choose the corresponding residuals r_x, r_s and r_c to have our starting point

$$(x_0, y_0, s_0, \tau_0, \kappa_0) \in \operatorname{int}\bar{K} = \operatorname{int}(K \times R^m \times K^* \times R_+ \times R_+)$$

feasible for the system of linear equations in $\mathcal{F}$:

$$r_x := Ax_0 - \tau_0 b, \quad r_s := A^T y_0 + s_0 - \tau_0 c, \quad r_c := \langle c, x_0\rangle - \langle b, y_0\rangle + \kappa_0.$$

For the sake of simplicity, let us introduce new notation:

$$\bar{x} := (x, \tau) \in \bar{K} := K \times R_+, \quad \bar{F}(\bar{x}) := F(x) - \ln\tau,$$

$$\bar{s} := (s, \kappa) \in \bar{K}^* := K^* \times R_+, \quad \bar{F}_*(\bar{s}) := F_*(s) - \ln\kappa.$$

$$\bar{z} := (\bar{x}, \bar{s}), \quad \hat{F}(\bar{z}) := \bar{F}(\bar{x}) + \bar{F}_*(\bar{s}),$$

and let us suppress y from the variables of the problem (5.17). Then this problem can be written as

$$\begin{aligned} \min \quad & \hat{F}(\bar{z}), \\ \text{s.t.} \quad & B\bar{z} = r, \\ & \bar{z} \in \bar{K} \times \bar{K}^*. \end{aligned} \tag{5.18}$$

Denote

$$l(\bar{z}) := \langle s_0, x\rangle + \langle s, x_0\rangle + \tau_0\kappa + \kappa_0\tau.$$

Now we can define the shifted primal-dual central path:

$$\bar{z}(t) := \arg\min\{\hat{F}(\bar{z}) :\ B\bar{z} = r,\ l(\bar{z}) = (\nu + 1)(t + 1)\}.$$

Theorem 5.3 ([9]). *Let us choose $\bar{z}_0$ as follows:*

$$x_0 \in \operatorname{int} K, \quad s_0 = -F'(x_0), \quad \tau_0 = \kappa_0 = 1, \quad y_0 = 0.$$

Then the shifted primal-dual central path is well defined and $\bar{z}(1) = \bar{z}_0$. Moreover, along this path we have the following identities:

$$s(t) = -tF'(x(t)), \quad \tau(t)\kappa(t) = t. \tag{5.19}$$

□

Using this statement it is easy to prove the following important fact.

Theorem 5.4 *Any segment of the shifted primal-dual central path is $\sqrt{2}$-geodesic in the cone $\bar{K} \times \bar{K}^*$ with respect to the barrier $\hat{F}(\bar{z})$.*

Proof: Indeed, the equations (5.19) can be written in the following form:

$$\bar{s}(t) = -t\bar{F}'(\bar{x}(t)).$$

On the other hand, in view of the structure of the linear system in $\mathcal{F}$ we have the following:

$$Ax'(t) = \tau'(t)b, \quad s'(t) + A^T y'(t) = \tau'(t)c, \quad \langle c, x'(t)\rangle - \langle b, y'(t)\rangle + \kappa'(t) = 0,$$

where $y(t)$ corresponds to $s(t)$. Therefore

$$\begin{aligned}
\langle \bar{s}'(t), \bar{x}'(t)\rangle &= \langle s'(t), x'(t)\rangle + \tau'(t)\kappa'(t) \\
&= \langle \tau'(t)c - A^T y'(t), x'(t)\rangle + \tau'(t)(\langle b, y'(t)\rangle - \langle c, x'(t)\rangle) \\
&= -\langle Ax'(t), y'(t)\rangle + \tau'(t)\langle b, y'(t)\rangle = 0.
\end{aligned}$$

Thus, using Theorem 4.1 we get the result. □

Hence, regular short-step path-following strategies for the shifted primal-dual central path are optimal for reaching the set $\{\bar{z} : B\bar{z} = r,\ l(\bar{z}) = (\nu+1)M\}$ for any $M > 1$ (see [9] for discussion).

6 Geodesic curves for simple sets

In this section we derive the form of geodesic curves for some standard convex sets. Some of these sets are direct products of simple one-dimensional sets. In such cases we apply Lemmas 4.1 and 4.2, which tell us how to construct geodesic curves for direct products.

6.1 Positive orthant

We can represent the positive orthant R^n_+ as a direct product of one-dimensional rays. Choosing the barrier function $f(\tau) = -\ln\tau$ as a barrier for a positive ray, we obtain from (2.3) the following equation for its geodesic curves:

$$\xi(t)\xi''(t) = (\xi'(t))^2.$$

Moreover, this holds componentwise for any geodesic curve in $\operatorname{int} R^n_+$. Let us change variables: $v = \ln x$ (componentwise). Then for the curve $\beta(t) := \ln\xi(t)$ we have

$$\beta''(t) = 0, \quad t \in \Delta.$$

Thus, any geodesic curve connecting two points x and y in $\operatorname{int} R^n_+$ has the form

$$\xi(t) = \exp[(1-t)\ln x + t\ln y] = x^t y^{1-t}$$

(componentwise). It is clear that this does indeed give the unique geodesic curve for the case $n = 1$, and then Lemma 4.1, part (c), and Lemma 4.2 show that it also gives the unique geodesic curve for general n. In view of Theorem 2.1 and Lemma 4.1, we have

$$\begin{aligned} \rho^2(x,y) &= \langle F''(x)\xi'(0), \xi'(0)\rangle \\ &= \sum_{i=1}^n \tfrac{1}{x_i^2}\left(x_i \ln\tfrac{y_i}{x_i}\right)^2 = \sum_{i=1}^n (\ln y_i - \ln x_i)^2 . \end{aligned} \tag{6.20}$$

Thus the transformation $x \to v := \ln x$ is an isometry between R^n_+ with the Riemannian metric ρ and R^n with the Euclidean metric. Note that, in v-space, Lemma 2.2 corresponds to $\|\Delta v\|_2 \geq |\sum \Delta v_i| / \sqrt{n}$, a consequence of the Cauchy-Schwarz inequality.

We can use this example to show that projections of sub-geodesic curves are not necessarily sub-geodesic. Consider the path in $\operatorname{int} R^2_+$ defined by

$$x(t) := \begin{cases} \begin{pmatrix} \exp[\alpha(1/2-t)] \\ \exp[t-1/2] \end{pmatrix}, & 0 \leq t \leq 1/2, \\ \begin{pmatrix} \exp[t-1/2] \\ \exp[\alpha(1/2-t)] \end{pmatrix}, & 1/2 \leq t \leq 1, \end{cases}$$

where $\alpha > 1$. In v-space, this corresponds to a line segment from $(\alpha/2, -1/2)^T$ to the origin followed by one from the origin to $(1/2, -\alpha/2)^T$. It is easy to see that $\langle F'(x(t)), x'(t)\rangle = \alpha - 1 > 0$ for $t \in \Delta$, $t \neq 1/2$, so that F is increasing. Using Proposition 4.1 we find a corresponding path $s(t)$ in $\operatorname{int} R^2_+$ which satisfies the hypotheses of Theorem 4.1. Indeed, we obtain

$$s(t) := \begin{cases} \begin{pmatrix} \exp[-\alpha/2 - (\beta-\alpha)t)] \\ \exp[1/2 - (\beta+1)t] \end{pmatrix}, & 0 \leq t \leq 1/2, \\ \begin{pmatrix} \exp[1/2 - (\beta+1)t] \\ \exp[-\alpha/2 - (\beta-\alpha)t) \end{pmatrix}, & 1/2 \leq t \leq 1, \end{cases}$$

with $\beta := (\alpha^2+1)/(\alpha-1)$. Then the path $(x(t), s(t))$ is $\sqrt{2}$-geodesic. However, using the corresponding path in v-space, we find that $\rho_x(x(0), x(1)) = \sqrt{\alpha^2+1}$ while

$$\rho(x(0), x(1)) = \|((\alpha/2, -1/2)^T - (1/2, -\alpha/2)^T\|_2 = (\alpha-1)/\sqrt{2},$$

and the ratio is unbounded as α tends to 1 from above.

6.2 The hypercube

In order to simplify our calculations, consider the n-dimensional hypercube in the following form:

$$Q := \{x \in R^n : \mid x_i \mid \leq \frac{\pi}{2},\ i = 1, \ldots, n\}.$$

(An arbitrary hyperrectangle can be treated by a trivial affine transformation.) As before, we can represent Q as a direct product of the intervals $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Let us introduce for this interval the following barrier function:

$$f(\tau) := -\ln \cos \tau.$$

Note that

$$f'(\tau) = \tan \tau, \quad f''(\tau) = \sec^2 \tau, \quad f'''(\tau) = 2 \tan \tau \sec^2 \tau.$$

Clearly, this function is a 1-self-concordant barrier for the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

In view of equation (2.3), any geodesic curves for this interval must satisfy the following differential equation:

$$\xi''(t) + \tan \xi(t)(\xi'(t))^2 = 0,$$

or $(\ln \xi'(t))' = (\ln \cos \xi(t))'$. Therefore

$$\xi'(t) = c \cdot \cos \xi(t), \tag{6.21}$$

where c is a positive constant. Consider the function

$$\psi(\tau) := \ln(\sec \tau + \tan \tau).$$

Note that $\psi'(\tau) = [\tan \tau \sec \tau + \sec^2 \tau]/[\sec \tau + \tan \tau] = \sec \tau$. Thus, the equation (6.21) can be rewritten as follows:

$$(\psi(\xi(t)))' = c.$$

Therefore, any geodesic curve connecting x and y in $(-\frac{\pi}{2}, \frac{\pi}{2})$ has the following form:

$$\xi(t) = \psi^{-1}(\psi(x) + t(\psi(y) - \psi(x))), \quad t \in \Delta.$$

It is clear that this does indeed give the unique geodesic curve for the case $n = 1$, and then Lemma 4.1, part (c), and Lemma 4.2 show that applying the formula componentwise also gives the unique geodesic curve for general n.

Using Theorem 2.1 and Lemma 4.1 we get the following:

$$\begin{aligned}
\rho^2(x, y) &= \langle F''(x)\xi'(0), \xi'(0)\rangle \\
&= \sum_{i=1}^{n} \left[\frac{(\psi^{-1}(\psi(x_i)+t(\psi(y_i)-\psi(x_i))))'_{t=0}}{\cos x_i}\right]^2 \\
&= \sum_{i=1}^{n} \left[\psi(y_i) - \psi(x_i)\right]^2.
\end{aligned}$$

Again, ψ applied componentwise is an isometry between Q with the Riemannian metric ρ and R^n with the Euclidean metric.

6.3 Cone of positive semidefinite matrices

Let S^n be the space of symmetric positive semidefinite matrices of order n. Consider $Q = K := S^n_+$, the cone of symmetric positive semidefinite matrices in S^n. Let us endow this cone with the natural n-self-concordant logarithmically homogeneous barrier

$$F(X) := -\ln \det X.$$

Theorem 6.1 *Let $X, Y \in \operatorname{int} S^n_+$. Denote*

$$\lambda_i(X,Y) = \lambda_i[X^{-1/2} Y X^{-1/2}], \quad i = 1, \ldots, n,$$

where $\lambda_i[A]$ is the ith eigenvalue of the symmetric matrix $A \in S^n$. Then

$$\rho(X,Y) = \left[\sum_{i=1}^n \ln^2 \lambda_i(X,Y)\right]^{1/2}$$

and the shortest path from X to Y is

$$\xi(t) := X^{1/2}(X^{-1/2} Y X^{-1/2})^t X^{1/2}.$$

Proof: Note that the Riemannian metric is invariant with respect to appropriate non-singular affine transformations of variables. Therefore, for any X, $Y \succ 0$ we have

$$\rho(X,Y) = \rho(I, X^{-1/2} Y X^{-1/2}),$$

where I is the identity matrix. A further affine transformation diagonalizes $X^{-1/2} Y X^{-1/2}$. Thus, we need to compute the distance between I and an arbitrary diagonal matrix $Z \succ 0$. Note that the characteristic equation (2.3) in our case has the following form:

$$\zeta^{-1}(t)\zeta''(t)\zeta^{-1}(t) - \zeta^{-1}(t)\zeta'(t)\zeta^{-1}(t)\zeta'(t)\zeta^{-1}(t) = 0.$$

That is

$$\zeta''(t) = \zeta'(t)\zeta^{-1}(t)\zeta'(t), \quad \zeta(0) = I,\ \zeta(1) = Z.$$

It is easy to see that

$$\bar{\zeta}(t) = Z^t, \quad t \in \Delta,$$

is a solution to this equation (since Z is diagonal, this can be viewed as the diagonal matrix whose diagonal entries are the tth powers of those of Z). In order to compute its Riemannian length, we note that $\bar{\zeta}'(t) = \ln(Z)\bar{\zeta}(t)$, so that

$$\langle F''(\bar{\zeta}(t)\bar{\zeta}'(t), \bar{\zeta}'(t)\rangle = \langle \ln(Z), \ln(Z)\rangle = \sum_{i=1}^n \ln^2 \lambda_i(Z).$$

(Again, $\ln(Z)$ can be viewed as the diagonal matrix whose diagonal entries are the logarithms of those of Z.) Thus $\rho_{\bar{\zeta}}(I, Z) = \| \ln \lambda(Z)\|_2$.

We have shown that $\bar{\zeta}$ is a local geodesic from I to Z. However, in this example, the Riemannian metric resulting from F coincides (up to a constant multiple) with that arising from the negative of the Hessian of the logarithm of the *norm* given by Rothaus

[10] for the interior of this cone (see Güler [2]). Corollary 4.13 in [10] shows that there is a *unique* local geodesic connecting any two points, and since any geodesic must be a local geodesic, $\bar{\zeta}$ is the unique geodesic from I to Z.

The corresponding path from I to $X^{-1/2}YX^{-1/2}$ is $\xi'(t) := (X^{-1/2}YX^{-1/2})^t$, where for an arbitrary symmetric matrix W, W^t is the matrix QD^tQ^T, with QDQ^T the eigenvalue decomposition of W and D^t defined as above for the diagonal matrix D. Finally, the corresponding path from X to Y is then $\xi(t)$ as given in the theorem. □

6.4 Lorentz cone

.

Let us consider $Q = K$, where K is the Lorentz cone:

$$K := \{z = (\tau, x) \in R^{1+n} : \ \tau \geq \|x\|_2\}.$$

Let us endow this cone with the 2-self-concordant logarithmically homogeneous barrier

$$F(z) := -\ln(\tau^2 - \|x\|_2^2), \quad z \in \operatorname{int} K.$$

For any z_0, $z_1 \in \operatorname{int} K$ denote

$$\sigma(z_0, z_1) := \max\{\lambda : \ z_0 - \lambda z_1 \in K\}.$$

Proposition 6.1 *For any $z_0, z_1 \in \operatorname{int} K$ we have*

$$\rho(z_0, z_1) = [\ln^2 \sigma(z_0, z_1) + \ln^2 \sigma(z_1, z_0)]^{1/2}.$$

Proof: Note that K is a homogeneous cone. Therefore there exists an automorphism A which maps z_0 to $e_0 = (1, 0_n)$: $e_0 = Az_0$. Indeed, if $z_0 = (\tau_0, x_0)$ and $\rho_0 := \sqrt{\tau_0^2 - \|x_0\|_2^2}$, then A is represented by the matrix

$$\frac{1}{\rho_0}\begin{pmatrix} -\frac{\tau_0}{\tau_0+\rho_0} & 0 \\ 0 & I \end{pmatrix} + \frac{1}{\rho_0(\tau_0+\rho_0)}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} \tau_0 \\ -x_0 \end{pmatrix}^T + \begin{pmatrix} \tau_0 \\ -x_0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}^T + \frac{1}{\rho_0}\begin{pmatrix} \tau_0 \\ -x_0 \end{pmatrix}\begin{pmatrix} \tau_0 \\ -x_0 \end{pmatrix}^T\right).$$

Denote $w_1 = Az_1$ and assume that $w_1 \neq e_0$. Then

$$\rho(z_0, z_1) = \rho(e_0, w_1).$$

As in the previous example, the metric here coincides (up to a constant multiple) with that arising from Rothaus's [10] norm, and so the geodesic is unique. In view of the symmetry of the cone K, the geodesic curve from e_0 to w_1 must belong to the plane spanned by these vectors, since, if not, its reflection in this plane would be another geodesic. Because of the rotational symmetry of the barrier function, we can assume that $n = 2$. Denote

$$e_1 = e_0 - \sigma(e_0, w_1)w_1, \quad e_2 = w_1 - \sigma(w_1, e_0)e_0.$$

Then in the basis (e_1, e_2) our vectors have the following representation:

$$\begin{aligned} e_0 &= \left(\tfrac{1}{1-\sigma(e_0,w_1)\sigma(w_1,e_0)}, \tfrac{\sigma(e_0,w_1)}{1-\sigma(e_0,w_1)\sigma(w_1,e_0)}\right), \\ w_1 &= \left(\tfrac{\sigma(w_1,e_0)}{1-\sigma(e_0,w_1)\sigma(w_1,e_0)}, \tfrac{1}{1-\sigma(e_0,w_1)\sigma(w_1,e_0)}\right). \end{aligned}$$

Note that the denominator in the above expressions is strictly positive since the nonzero vectors e_1 and e_2 lie in K and for any $s \in \operatorname{int} K^*$ we have

$$\begin{aligned} 0 < \langle s, e_1\rangle &= \langle s, e_0 - \sigma(e_0, w_1)w_1\rangle = \langle s, e_0 - \sigma(e_0, w_1)(e_2 + \sigma(w_1, e_0)e_0)\rangle \\ &= (1 - \sigma(e_0, w_1)\sigma(w_1, e_0))\langle s, e_0\rangle - \sigma(e_0, w_1)\langle s, e_2\rangle. \end{aligned}$$

Observe that the points e_1 and e_2 are positive multiples of the extreme rays of the cone $(1, 1)$ and $(1, -1)$. Now the barrier for a point $(\tau, x) =: \lambda_1 e_1 + \lambda_2 e_2$ is

$$-\ln(\tau^2 - x^2) = -\ln(\tau + x) - \ln(\tau - x) = -\ln\lambda_1 - \ln\lambda_2 + \kappa$$

for some constant κ, and so the Riemannian metric coincides with that for R^2_+ when points are represented in terms of the basis (e_1, e_2). Thus, in view of relation (6.20) we have

$$\rho^2(e_0, w_1) = \ln^2\sigma(e_0, w_1) + \ln^2\sigma(w_1, e_0).$$

It remains to note that

$$\sigma(e_0, w_1) = \sigma(z_0, z_1), \quad \sigma(w_1, e_0) = \sigma(z_1, z_0).$$

□

Given the explicit form of A and the known geodesics in R^2_+, the form of the geodesic can be computed in a straightforward if tedious way.

6.5 Epigraph of the parabola

In this section we derive the family of geodesic curves for the following two-dimensional set

$$Q := \{z = (x, y) \in R^2 : \ y \geq x^2\}.$$

The Riemannian metric in this set is defined using the natural 1-self-concordant barrier

$$F(z) := -\ln(y - x^2).$$

Proposition 6.2 *For any points $z_0 = (x_0, y_0)$, $z_1 = (x_1, y_1)$ from $\operatorname{int} Q$ with $x_0 \neq x_1$ the corresponding geodesic curve can be written in the form*

$$y = \frac{1}{2}x^2 + ax + b,$$

for some constants a, b. If $x_0 = x_1$ then the unique geodesic curve is a straight vertical line.

Proof: In order to find the form of geodesic curves, we need to solve equation (2.3). Denote $\omega(z) := y - x^2$. Consider the two functions

$$\begin{aligned} g_x(z) &:= \tfrac{\partial F(z)}{\partial x} = \tfrac{2x}{\omega(z)}, \\ g_y(z) &:= \tfrac{\partial F(z)}{\partial y} = -\tfrac{1}{\omega(z)}. \end{aligned}$$

In order to work with equation (2.3), we need first to compute the directional derivatives of these functions. For a direction $h = (h_x, h_y) \in R^2$ we have

$$\omega'(z)[h] = h_y - 2xh_x, \quad \omega''(z)[h,h] = -2(h_x)^2. \tag{6.22}$$

Therefore

$$g_x'(z)[h] = \frac{2h_x}{\omega(z)} - \frac{2x}{\omega^2(z)}\omega'(z)[h] = \frac{2}{\omega^2(z)}(\omega(z)h_x - x\omega'(z)[h]), \tag{6.23}$$

$$\begin{aligned} g_x''(z)[h,h] &= -\frac{4h_x}{\omega^2(z)}\omega'(z)[h] + \frac{4x}{\omega^3(z)}(\omega'(z)[h])^2 - \frac{2x}{\omega^2(z)}\omega''(z)[h,h] \\ &= -\frac{4}{\omega^3(z)}\left[\omega(z)\omega'(z)[h]h_x - x(\omega'(z)[h])^2 + \tfrac{1}{2}x\omega(z)\omega''(z)[h,h]\right], \end{aligned} \tag{6.24}$$

$$g_y'(z)[h] = \frac{\omega'(z)[h]}{\omega^2(z)} = \frac{1}{\omega^2(z)}(h_y - 2xh_x), \tag{6.25}$$

$$\begin{aligned} g_y''(z)[h,h] &= \frac{\omega''(z)[h,h]}{\omega^2(z)} - 2\frac{(\omega'(z)[h])^2}{\omega^3(z)} \\ &= -\frac{2}{\omega^3(z)}\left[(\omega'(z)[h])^2 - \tfrac{1}{2}\omega(z)\omega''(z)[h,h]\right]. \end{aligned} \tag{6.26}$$

Now let z be a function of t: $z(t) = (x(t), y(t))$. Let us rewrite the equations (6.23) and (6.25) using the direction $h = z''(t)$. We have

$$\omega'(z(t))[z''(t)] = y''(t) - 2x(t)x''(t).$$

Therefore (in the following we frequently omit the arguments for simplicity)

$$g_x'(z(t))[z''(t)] = \frac{2}{\omega^2}\left[\omega x'' - xy'' + 2x^2x''\right], \tag{6.27}$$

$$g_y'(z(t))[z''(t)] = \frac{1}{\omega^2}\left[y'' - 2xx''\right]. \tag{6.28}$$

For the direction $h = z'(t)$ we have

$$\begin{aligned} \omega'(z(t))[z'(t)] &= y' - 2xx', \\ \omega''(z(t))[z'(t), z'(t)] &= -2(x')^2. \end{aligned}$$

Therefore, the relations (6.24) and (6.26) can be rewritten as follows:

$$g''_x(z(t))[z'(t), z'(t)] = -\frac{4}{\omega^3}\left[\omega(y' - 2xx')x' - x(y' - 2xx')^2 - x\omega(x')^2\right], \quad (6.29)$$

$$g''_y(z(t))[z'(t), z'(t)] = -\frac{2}{\omega^3}\left[(y' - 2xx')^2 + \omega(x')^2\right]. \quad (6.30)$$

Note that from equation (2.3) we must have

$$g'_x(z(t))[z''(t)] = -\tfrac{1}{2}g''_x[z'(t), z'(t)].$$

In view of equations (6.27) and (6.29) this is

$$\tfrac{2}{\omega^2}\left[\omega x'' - xy'' + 2x^2x''\right] = \tfrac{2}{\omega^3}\left[\omega(y' - 2xx')x' - x(y' - 2xx')^2 - x\omega(x')^2\right].$$

Similarly, since we need

$$g'_y(z(t))[z''(t)] = -\tfrac{1}{2}g''_y[z'(t), z'(t)],$$

in view of equations (6.28) and (6.30) we have

$$\tfrac{1}{\omega^2}\left[(y'' - 2xx''\right] = \tfrac{1}{\omega^3}\left[(y' - 2xx')^2 + \omega(x')^2\right].$$

These equations are equivalent to

$$\begin{aligned} \omega^2x'' - \omega x(y'' - 2xx'') &= \omega(y' - 2xx')x' - x[(y' - 2xx')^2 + \omega(x')^2], \\ \omega(y'' - 2xx'') &= (y' - 2xx')^2 + \omega(x')^2, \end{aligned}$$

which are in turn equivalent to

$$\begin{aligned} \omega[y'' - 2xx'' - (x')^2] &= (y' - 2xx')^2, \\ \omega x'' &= (y' - 2xx')x'. \end{aligned} \quad (6.31)$$

Let us suppose first that both $x'(t)$ and $y'(t) - 2x(t)x'(t)$ are nonzero for a particular t. Then, equating the two expressions for $\omega(z(t))$, we find $(y' - 2xx')x'' = [y'' - 2xx'' - (x')^2]x'$ so that $(y' - xx')x'' = [y'' - xx'' - (x')^2]x'$ or

$$\left[\frac{(y - \frac{1}{2}x^2)'}{x'}\right]' = \left[\frac{y' - xx'}{x'}\right]' = \frac{y'' - xx'' - (x')^2}{x'} - \frac{y' - xx'}{(x')^2}x'' = 0.$$

Hence locally $(y - \frac{1}{2}x^2)' = ax'$ for some constant a, and thus $y = \frac{1}{2}x^2 + ax + b$ for some constants a and b.

Now we deal with the exceptional cases. First note that the condition that the local norm of $z'(t)$ be one becomes $x'(t)g'_x(z(t))[z'(t)] + y'(t)g'_y(z(t))[z'(t)] = 1$, or

$$\frac{(y' - 2xx')^2}{\omega^2} + \frac{2(x')^2}{\omega} = 1.$$

If $y'(t) - 2x(t)x'(t) = 0$ for some t, then (6.31) gives $x'' = 0$ and $y'' - 2xx'' - (x')^2 = 0$ at t. Then the norm condition above shows that $(x')^2 = \omega/2$ and hence $y'' = (x')^2 = \omega/2$. But

then $y'(s) - 2x(s)x'(s) = 0$ cannot hold in a neighborhood of t, since then its derivative is zero, but at t this is $y'' - 2xx'' - 2(x')^2 = -(x')^2 = -\omega/2 < 0$. (This condition is possible at a point: it holds for $x(t) = a$ in the parabolic solution above.)

If $x'(t) = 0$ for some t, then (6.31) gives $x'' = 0$, the norm condition implies $y' = y' - 2xx' = \omega$, and so a solution to this is $x(t) = \xi$ constant, $y(t) - \xi^2 = \omega(z(t)) = \alpha \exp(t)$ for some constant α. This solution gives a vertical line. Since any geodesic is smooth, it cannot switch from the parabolic solution to the straight-line solution, so the proposition is proved. □

References

[1] A. Edelman, T. Arias, and S. T. Smith, The geometry of algorithms with orthogonality constraints, *SIAM Journal on Matrix Analysis and Applications* **20** (1998), 303–353.

[2] O. Güler, Barrier functions in interior-point methods, *Mathematics of Operations Research* **21** (1996), 860–885.

[3] S. Helgason, *Differential Geometry and Symmetric Spaces*, American Mathematical Society, Providence, RI, 2000.

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

[5] N. K. Karmarkar, Riemannian geometry underlying interior–point methods for linear programming, in *Mathematical Developments Arising from Linear Programming* (J. C. Lagarias and M. J. Todd, eds.), Contemporary Mathematics, Vol. 114, American Mathematical Society, Providence, RI, 1990, pp. 51–75.

[6] M. Koecher, Positivitätsbereiche im R^n, *American Journal of Mathematics* **79** (1957), 575–596.

[7] Yu. E. Nesterov, Long-step strategies in interior-point primal-dual methods, *Mathematical Programming* **76** (1996), 47–94.

[8] Yu. E. Nesterov and A. S. Nemirovskii, *Interior Point Polynomial Methods in Convex Programming : Theory and Algorithms*, SIAM, Philadelphia, PA, 1994.

[9] Yu. E. Nesterov, M. J. Todd, and Y. Ye, Infeasible-start primal-dual methods and infeasibility detectors for nonlinear programming problems, *Mathematical Programming* **84** (1999), 227–267.

[10] O. Rothaus, Domains of positivity, *Abh. Math. Sem. Univ. Hamburg* **24** (1960), 189–235.

[11] K. Tanabe, Geometric method in nonlinear programming, *Journal of Optimization Theory and Applications* **30** (1980), 181–210.

[12] M. J. Todd, On adjusting parameters in homotopy methods for linear programming, in *Approximation Theory and Optimization* (M. Buhmann and A. Iserles, eds.), Cambridge University Press, 1997, pp. 201–220.

[13] M. H. Wright, Interior methods for constrained optimization, in *Acta Numerica 1992* (A. Iserles, ed.), Cambridge University Press, New York, 1992, pp. 341–407.

[14] S. J. Wright, *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, PA, 1997.