Target Directions for Primal-Dual Interior-Point Methods for Self-Scaled Conic Programming

Raphael A. Hauser*

Department of Applied Mathematics and Theoretical Physics,
University of Cambridge,Silver Street,Cambridge CB3 9EW,England.
rah48@damtp.cam.ac.uk

December 18, 1999

Abstract

The theory of self-scaled conic programming provides a unified framework for the theories of linear programming, semidefinite programming and convex quadratic programming with convex quadratic constraints. In the linear programming literature there exists a unifying framework for the analysis of various important classes of interior-point algorithms, known under the name of *target-following* algorithms. This article is a step towards combining these two unifying theories in that we develop an infinite new family of Newton directions for self-scaled conic programming which inherit the properties of the search-direction that made target-following algorithms possible in the LP case. These so-called *target directions* are close relatives of the Nesterov-Todd direction and lend themselves to the construction of predictor-corrector methods. Moreover, target directions are closely connected to the notion of weighted analytic centers.

Key words. symmetric cones, self-scaled cones, primal-dual methods, interior-point methods, target-following methods, weighted analytic centers, convex programming, self-scaled programming, linear programming, semidefinite programming, second-order cone programming, quadratic programming.

AMS 1991 subject classification. Primary: 90C25, 90C05, 49M45. Secondary: 90C20, 68Q25.

*Research supported by the Swiss National Science Foundation through project "Newton Directions for Interior-Point Methods in Self-Scaled Convex Programming", by the National Science Foundation through grant DMS-9505155 and by the Office of Naval Research through grant N00014-96-1-0050.

1 Introduction

Self-scaled conic programming provides a unified framework for linear- semidefinite- and second-order-cone programming and was first investigated by Nesterov-Todd [23, 24]. In [9] we began the generalization of the so-called "V-space" approach, a theory that allows the unified analysis of various important classes of primal-dual interior-point methods for linear programming (see e.g. Kojima-Mizuno-Yoshise [14], Monteiro-Adler [20], Mizuno [16], Jansen-Roos-Terlaky-Vial [11, 12], Todd [29]). Related generalizations to semidefinite programming have been analyzed by Sturm-Zhang [27, 28], Monteiro-Zanjacomo [19] and Burer-Monteiro [1]. Primal-dual interior point methods for convex optimization problems are designed to solve a problem and its dual jointly by making use of convex duality theory. The paradigm for such algorithms usually is to reduce the duality gap between primal and dual approximate solutions to zero. In the V-space approach, both the primal and the dual problem are mapped into some Euclidean space V lying "half-way" between the spaces E and $E^\sharp$ in which the primal and dual problems live. In [9] we introduced a class of differentiable operator fields that achieve a useful transformation into such a V-space. We called these operator fields *square-root fields*. These objects are closely related to similar transformations defined by Tunçel for self-scaled conic programming in [32] and for general convex conic programming in [33]. In the present article we take the generalization of the V-space approach a step further, by introducing *target maps* and *target directions*, which allow for the construction of so-called *target-following algorithms* in the context of self-scaled conic programming. These notions are defined in terms of square-root fields. Target-directions constitute a new infinite class of Newton directions for primal-dual interior-point methods for self-scaled conic programming. These direction fields are close relatives of the Nesterov-Todd direction fields, and they all lend themselves to the construction of so-called *predictor-corrector methods*. Moreover, target-maps are closely connected to the notion of *weighted analytic centers*.
Weighted centers were introduced to the linear programming literature by Kojima, Mizuno and Yoshise [14]. There they play an important role as a unifying framework for the analysis of interior-point methods, an approach that was developed by Kojima et al. [15] and Jansen et al. [11]. For a complete treatment see also [25]. Considerable interest has arisen recently to generalize weighted centers to semidefinite programming, and several competing approaches have been developed by Monteiro-Pang [21], Sturm-Zhang [28] and Monteiro-Zanjacomo [19]. In the present section we present yet another generalization of weighted centers to SDP that extends to the more general framework of self-scaled conic programming.

Self-scaled conic programming is extensively discussed in the two seminal papers [23, 24] of Nesterov-Todd. See also Güler [7] and Hauser [8]. The extent to which this theory is used in this article and all the notation applied here can be found in Hauser [8, 9].

Let E be a n-dimensional real vector space and $E^\sharp$ its dual. Throughout this article we will reserve super- or subscript-$\sharp$ for duals, superscript-$*$ for adjoints and superscript-$\circ$ for topological interiors. Let $K \subset E$ be a regular cone in E, i.e. convex, pointed and with

non-empty interior $K^\circ \neq \emptyset$. As usual, we define the dual of K as follows:

$$K^\sharp := \{ s \in E^\sharp : \langle x; s \rangle \geq 0 \quad \forall\, x \in K \}.$$

By $\langle \cdot; \cdot \rangle$ let us denote the canonical bilinear product defined on $E \times E^\sharp$. Let $(Y, Y^\sharp, \langle \cdot; \cdot \rangle)$ be another bilinear system of m-dimensional real vector spaces, and let $A : E \to Y^\sharp$ be a surjective vector space homomorphism. Let $b \in Y^\sharp$ and $c \in E^\sharp$ be fixed vectors. Let us make the usual regularity assumption that there *exist* vectors $x_0 \in K^\circ$ and $(y_0, s_0) \in Y \times K^{\sharp\circ}$ such that $Ax_0 = b$ and $A^* y_0 + s_0 = c$ respectively. Let L be the linear subspace of E defined by $L := \ker(A)$ and $L^\perp := \mathrm{im}(A^*)$ its orthogonal complement with respect to the bilinear form $\langle \cdot; \cdot \rangle$. The general primal-dual pair of self-scaled conic programming problems in primal-dual symmetric form is given by

$$\begin{array}{lll} (1.1) & \text{(P)} \quad \inf \langle x; s_0 \rangle & \text{(D)} \quad \inf \langle x_0; s \rangle \\ & \qquad x \in (L + x_0) \cap K & \qquad s \in (L^\perp + s_0) \cap K^\sharp, \end{array}$$

where $K \subset E$ is a symmetric cone (see e.g. [13, 6, 26, 34]). Under the above-mentioned regularity assumption of primal-dual strict feasibility, self-scaled conic programming contains as special cases linear programming (by choosing $K = \mathbb{R}^n_+$), semidefinite programming (by choosing $K = \mathrm{S}^{n\times n}(\mathbb{R})_+$, the cone of positive semidefinite symmetric $n \times n$ matrices) and second-order cone programming (by choosing K equal to the $n+1$-dimensional Lorentz cone).

Let the symmetric cone K be endowed with a *self-scaled barrier functional* F (see [23]). As usual, we endow $K^\sharp$ with the *dual barrier functional*

$$\begin{aligned} F_\sharp : K^{\sharp\circ} &\to \mathbb{R} \\ s &\mapsto \max\{ -\langle x; s \rangle - F(x) : x \in K^\circ \}, \end{aligned}$$

which is also self-scaled (see [23]). The reader may want to think of the linear programming case, where the standard logarithmic barrier functional

$$F(x) : x \mapsto -\sum \ln(x_i)$$

and its dual barrier functional

$$F_\sharp(s) : s \mapsto -\sum \ln(s_i) - n$$

are self-scaled. Let us consider the primal-dual pair of self-scaled conic programming problems 1.1. The fact that F is self-scaled implies among other things that at any point $w \in K^\circ$ the mapping $F''(w) : E \to E^\sharp$ is an isomorphism between K and $K^\sharp$, and that for any pair of points $(x, s) \in K^\circ \times K^{\sharp\circ}$ there exists a unique *scaling point* $w(x, s) \in K^\circ$ such that $F''(w)x = s$ (see Theorem 2.3). Let us assume that we are given a primal-dual strictly feasible point (x, s), i.e. $x \in K^\circ \cap (L + x_0)$ and $s \in K^{\sharp\circ} \cap (L^\perp + s_0)$. In order to solve the perturbed problems $\min\{ \langle \tilde{x}; s_0 \rangle + \mu F(\tilde{x}) : \tilde{x} \in K \cap (L + x_0) \}$ and $\min\{ \langle x_0; \tilde{s} \rangle + \mu F_\sharp(\tilde{s}) : \tilde{s} \in K^\sharp \cap (L^\perp + s_0) \}$ for some positive real number $\mu > 0$, i.e. in order to solve the central path equations

$$(1.2) \qquad \tilde{s} = -\mu F'(\tilde{x}), \quad \tilde{x} \in K \cap (L + x_0),\ \tilde{s} \in K^\sharp \cap (L^\perp + s_0),$$

Nesterov and Todd [23] propose to generate a search direction that arises as the solution to the following system of equations (the so-called *N-T equations*):

$$F''(w)d_x + d_s = -\mu F'(x) - s, \tag{1.3a}$$

$$d_x \in L, \tag{1.3b}$$

$$d_s \in L^\perp. \tag{1.3c}$$

Two aspects of the pseudo-Newton equation (1.3a) are unpleasant in its current form: First, this equation does not reflect the symmetry between the primal and the dual problems. Second, some of its terms blow up near the optimal solution $(x^*, s^*) \in \partial K \times \partial K^\sharp$.

In [9] we introduced *square-root fields* to overcome these problems. Square-root fields are differentiable fields of non-singular operators

$$\begin{aligned} X:\ & K^\circ \times K^{\sharp\circ} \to \mathrm{Aut}(E^\sharp, V) \qquad \text{and} \\ S:\ & K^\circ \times K^{\sharp\circ} \to \mathrm{Aut}(E, V) \end{aligned}$$

that are linked to one another in a certain fashion (see Definition 3.15, [9]). In the present article we will show that when equation (1.3a) is multiplied by $X(x,s)$ on the left, we end up with a primal-dual symmetric system

$$S(x,s)d_x + X(x,s)d_s = \mu e_V - \gamma(x,s), \qquad d_x \in L, \quad d_s \in L^\perp$$

all of whose terms remain bounded in terms of the *duality gap* $\langle x; s \rangle$, where $e_V \in V$ is a fixed vector (see Definition 3.15, [9]) and the *target map* γ is defined as

$$\begin{aligned} \gamma : K^\circ \times K^{\sharp\circ} &\to V \\ (x,s) &\mapsto X(x,s)s = S(x,s)x. \end{aligned}$$

The reader may for example think of the linear programming case, where the Nesterov-Todd equations reduce to the well known canonical system for search direction generation

$$\begin{pmatrix} \frac{s_1}{x_1} & & \\ & \ddots & \\ & & \frac{s_n}{x_n} \end{pmatrix} d_x + d_s = \begin{pmatrix} \frac{\mu}{x_1} - s_1 \\ \vdots \\ \frac{\mu}{x_n} - s_n \end{pmatrix},$$

$$Ad_x = 0,$$

$$d_s = -A^* d_y \qquad \text{for some } d_y \in \mathbb{R}^m.$$

In this case a square-root field can be chosen as follows (see Example 3.18, [9]):

$$\begin{aligned} X &: (x,s) \mapsto \mathrm{Diag}(x), \\ S &: (x,s) \mapsto \mathrm{Diag}(s). \end{aligned} \tag{1.4}$$

Left multiplication of equation (1.3a) by the operator X yields

$$Sd_x + Xd_x = \mu e_V - \gamma(x,s),$$

where $S := \mathrm{Diag}(s)$, $e_V := (1, \ldots, 1)^T$ and where $\gamma(x,s) := \big(x_1 s_1, \ldots, x_n s_n\big)^T$.

The central path equation (1.2) can be reformulated in terms of the target map as follows:

$$\gamma(x,s) = \mu e_V, \qquad x \in L + x_0, \quad s \in L^\perp + s_0.$$

The corresponding Newton equation is

$$\gamma'(x,s)\begin{pmatrix} d_x \\ d_s \end{pmatrix} = S(x,s)d_x + X(x,s)d_x + R(x,s)\begin{pmatrix} d_x \\ d_s \end{pmatrix} = \mu e_V - \gamma(x,s), \tag{1.5}$$

where $R \in \mathcal{C}^1\big(K^\circ \times K^{\sharp\circ}, \mathrm{Lin}(L \oplus L^\perp, V)\big)$ is a differentiable operator field that vanishes on the central path (see Proposition 3.9). The solution to the Newton system (1.5) is the *target direction* defined by the square-root field (X, S). Leaving out the residual part

$$R(x,s)\begin{pmatrix} d_x \\ d_s \end{pmatrix}$$

in equation (1.5) yields the Nesterov-Todd system again, and hence the two systems asymptotically coincide in a neighborhood of the central path. In Section 4 this allows us to prove that the Nesterov-Todd method converges at a quadratic rate towards points on the central path (see Theorem 4.1). Finally, in Section 5 we show that a certain class of target maps generalizes the notion of *weighted analytic centers* to the framework of self-scaled conic programming. This concept plays an important role in the analysis of interior-point methods for linear programming (see e.g. Kojima-Mizuno-Yoshise [14]). Various generalizations to the framework of semidefinite programming have been proposed by Monteiro-Pang [21], Sturm-Zhang [28] and Monteiro-Zanjacomo [19].

Let us conclude this section with some remarks regarding notation. The set of *primal-dual strictly feasible solutions* and the set of *primal-dual feasible solutions* respectively are defined as follows:

$$\begin{aligned} \mathcal{F}^\circ(PD) &:= \Big(\big(L + x_0\big) \cap K^\circ\Big) \times \Big(\big(L^\perp + s_0\big) \cap K^{\sharp\circ}\Big) \qquad \text{and} \\ \mathcal{F}(PD) &:= \Big(\big(L + x_0\big) \cap K\Big) \times \Big(\big(L^\perp + s_0\big) \cap K^{\sharp}\Big). \end{aligned}$$

On the same note we define the following sets whose nomenclature is clear by analogy to what was just said:

$$\begin{aligned} \mathcal{F}(P) &:= \big(L + x_0\big) \cap K && \textit{the primal feasible domain} \\ \mathcal{F}^\circ(P) &:= \big(L + x_0\big) \cap K^\circ && \textit{the primal strictly feasible domain} \\ \mathcal{F}(D) &:= \big(L^\perp + s_0\big) \cap K^\sharp && \textit{the dual feasible domain} \\ \mathcal{F}^\circ(D) &:= \big(L^\perp + s_0\big) \cap K^{\sharp\circ} && \textit{the dual strictly feasible domain} \end{aligned}$$

The group of orthogonal transformations of the Euclidean space V will play an important role, and we denote it by $\mathcal{O}(V)$. The subgroup whose action leaves the element $e_V \in V$ invariant will be denoted by $\mathcal{O}(V, e_V)$. Also, we will consider a regular cone $K_V \subset V$ and denote the subgroup of $\mathcal{O}(V)$ whose elements map K_V onto itself by $\mathcal{O}(K_V)$, and

the subgroup of $\mathcal{O}(K_V)$ whose action leaves e_V invariant by $\mathcal{O}(K_V, e_V)$. Often the need arises to distinguish between linear operators and their matrix representation with respect to a particular choice of coordinates. In particular, we will write $\nabla^2 F(x)$ for the matrix representation of $F''(x)$. See also Remark 2.5. The sets of symmetric matrices, symmetric positive semidefinite and symmetric positive definite matrices over $\mathbb{R}$ will be denoted by $\mathrm{S}^{n\times n}(\mathbb{R})$, $\mathrm{S}^{n\times n}(\mathbb{R})_+$ and $\mathrm{S}^{n\times n}(\mathbb{R})_{++}$ respectively. For the set of diagonal $n \times n$-matrices over $\mathbb{R}$ we will use the notation $\mathrm{Diag}_n(\mathbb{R})$. By $\mathrm{diag}(D)$ we denote the element of $\mathbb{R}^n$ defined by the diagonal of matrix D, and conversely we write $\mathrm{Diag}(v)$ for the diagonal $n \times n$-matrix defined by $v \in \mathbb{R}^n$. The spectrum of a matrix M will be denoted by $\sigma(M)$, while we refer to the smallest and largest eigenvalues of M as $\sigma_{\min}(M)$ and $\sigma_{\max}(M)$ respectively. Identity maps will always be denoted by the symbol $\imath$. $S^{(n-1)}(H)$ denotes the unit sphere in a n-dimensional normed vector space H.

2 Prerequisite results and additional notation

Modern interior-point theory is based on the foundation of Nesterov & Nemirovskii's theory of self-concordant functionals. For a definition of ν-logarithmically homogeneous self-concordant barrier functionals see [22] or [8], Definition 2.1 from which we adopt notation.

Proposition 2.1 (Nesterov & Nemirovskii). *Let F be a ν-self-concordant logarithmically homogeneous barrier functional on the regular cone $K \subset E$, and let $x \in K^\circ$. Then*

i) $-F'(x) = F''(x)x \in K^{\sharp\circ}$,

ii) $-F_\sharp'(-F'(x)) = x$,

iii) $F_\sharp''(-F'(x)) = F''(x)^{-1}$,

iv) $\langle x; -F'(x)\rangle = \nu$,

v) for all $\tau > 0:\ F'(\tau x) = \tau^{-1}F'(x)$,

vi) for all $\tau > 0:\ F''(\tau x) = \tau^{-2}F''(x)$ *and*

vii) $F_\sharp\big(-F'(x)\big) = -\nu - F(x)$. □

The theory of self-scaled conic programming is founded on two seminal papers by Nesterov and Todd [23, 24]. The following definition and theorem are among the basic building blocks of this theory.

Definition 2.2 (Nesterov-Todd). A regular cone $K \subset E$ is *self-scaled* if it admits a *ν-self-scaled* barrier functional F, defined as follows:

i) $F \in \mathcal{C}^3(K^\circ, \mathbb{R})$ is a ν-self- concordant logarithmically homogeneous barrier functional for K,

ii) $F''(w)x \in K^{\sharp\circ}$ for all $x,\, w \in K^\circ$ and

iii) $F_\sharp\big(F''(w)x\big) = F(x) - 2F(w) - \nu$ for all $x,\, w \in K^\circ$.

(Compare with *i)* and *vii)* of Proposition 2.1 which show that these last two conditions hold for $w = x$ in the general case.) It can be shown that if F is self-scaled then its dual $F_\sharp$ is self-scaled as well.

For examples see Nesterov-Todd [23]. See also Hauser [8, 9] from which we will adopt all notation.

Theorem 2.3 (Nesterov-Todd). *Let K be a self-scaled cone with ν-self-scaled barrier functional F, and suppose $x \in K^\circ$, $s \in K^{\sharp\circ}$. Then there exists a unique* scaling point *$w = w(x,s) \in K^\circ$ such that $s = F''(w)x$. Furthermore, the following properties hold:*

i) $w(-F_\sharp'(s), -F'(x)) = w(x,s)$,

ii) $F''(w)$ *is a cone isomorphism* $K \to K^\sharp$,

iii) $F''(x) = F''(w) \circ F_\sharp''(s) \circ F''(w)$,

iv) The dual scaling point $t(x,s) = -F'(w)$ *satisfies* $F_\sharp''(t(x,s))s = x$. □

Corollary 2.4. *Let K be a self-scaled cone with ν-self-scaled barrier functional F, and suppose $x, y \in K^\circ$ are such that $F''(x)h = F''(y)h$ for some $h \in K^\circ$. Then $x = y$.* □

Definition 2.5. Let K, F be self-scaled. In the interior-point literature it is customary to view E as a Euclidean space by identifying E with $E^\sharp$ via $F''(e)$ for some fixed $e \in K^\circ$, i.e. we define a scalar product $\big(\cdot;\cdot\big)_e$ by the rule

$$\big(x_1; x_2\big)_e := \big\langle x_1; F''(e)x_2 \big\rangle \qquad \forall x_1, x_2 \in E.$$

We say that $\big(\cdot;\cdot\big)_e$ provides E with the *Euclidean structure defined by* e. If E is endowed with the basis $B \in \mathcal{B}^{(e)}$ then we endow the space $E^\sharp$ with the dual basis $B^\sharp = \big\{e_1^\sharp, \ldots, e_n^\sharp\big\}$, where $e_i^\sharp := F''(e)e_i$ for all $i \in \mathbb{N}_n$. Whenever we represent objects in matrix form it will notationally suffice to refer only to the basis B chosen on the primal space E. For example, the notation $\nabla_B^k F(x)$ shall refer to the matrix representation of the tensor $F^{(k)}(x) \in \otimes^k E^\sharp$ with respect to the bases B and $B^\sharp$. In particular, if $B \in \mathcal{B}^{(e)}$ then $\nabla_B^2 F(e) = I$. Once E has been endowed with a Euclidean structure, we denote the group of orthogonal transformations that map K onto itself by $\mathcal{O}(K)$ and call this the orthogonal group of K. By $\mathcal{Q}(K)$ let us denote the connected component of $\mathcal{O}(K)$ that contains the identity $\imath : E \to E$.

Remark 2.6. Some of our definitions will depend on the choice of coordinates on the spaces $E, E^\sharp$ and V. For coordinates on E and $E^\sharp$ we will stick to the conventions of Definition 2.5 by choosing a basis $B \in \mathcal{B}^{(e)}$ for E for some $e \in K^\circ$ and providing $E^\sharp$ with the dual basis $B^\sharp$. The objects defined in terms of $\left(B, B^\sharp\right)$ will usually bear a subscript B in their notation. V is always a Euclidean space of same dimension as E, and we will assume that an orthogonal basis $\left\{v_1, \ldots, v_n\right\} \subset V$ has been fixed for this space. Therefore, we will not usually refer to this basis in our notation for objects whose definitions depend on $\left\{v_1, \ldots, v_n\right\}$.

All the results and definitions of Hauser [9] will be heavily used. References to that paper will be indicated by the string "H2." followed by the number of the cited result, i.e. H2.3.11 refers to Definition 3.11, [9]. In the remainder of this section we cite just a few results from that paper that should enable the reader to get going.

Definition 2.7. Suppose $B \in \mathcal{B}$ and $x \in K^\circ$. Let $F_B''^{1/2}(x) \in \mathrm{Lin}(E, V)$ and $F_B''^{-1/2}(x) \in \mathrm{Lin}(E^\sharp, V)$ be the unique vector space homomorphisms whose matrix representations with respect to the bases tuple $\left(B, \left\{v_1, \ldots, v_n\right\}\right)$ are $\nabla_B^2 F(x)^{1/2}$ and $\nabla_B^2 F(x)^{-1/2}$ respectively (see Definition 2.5 for notation). Let us then define the following sets of linear operators:

$$F''^{1/2}(x) := \{F_B''^{1/2}(x) : B \in \mathcal{B}\} \qquad \text{and}$$
$$F''^{-1/2}(x) := \{F_B''^{-1/2}(x) : B \in \mathcal{B}\}.$$

Remark 2.8. Note that $\left(F_B''^{1/2}(x)\right)^{-1} : V \to E$ and $F_B''^{-1/2}(x) : E^\sharp \to V$ are two distinct operators. In fact, it is straightforward to see that these are adjoints of one another and that the following diagram is commutative:

$$\begin{array}{ccccccc}
E & = & E & \xrightarrow{F_B''(x)} & E^\sharp & = & E^\sharp \\
\uparrow{\scriptstyle F_B''^{-1/2}(x)^*} & & \downarrow{\scriptstyle F_B''^{1/2}(x)} & & \downarrow{\scriptstyle F_B''^{-1/2}(x)} & & \uparrow{\scriptstyle F_B''^{1/2}(x)^*} \\
V & = & V & = & V & = & V
\end{array}$$

Proposition 2.9. *Suppose $B \in \mathcal{B}$ and $x \in K^\circ$. Then*

$$F''^{1/2}(x) = \mathcal{O}(V) \circ F_B''^{1/2}(x) = \{L \in \mathrm{Lin}(E, V) : L^* \circ L = F''(x)\},$$
$$F''^{-1/2}(x) = \mathcal{O}(V) \circ F_B''^{-1/2}(x) = \{L \in \mathrm{Lin}(E^\sharp, V) : L^* \circ L = F''^{-1}(x)\}.$$

□

Definition 2.10 (Square-root fields).
Let $K \subset E$ be a symmetric cone with ν-self-scaled barrier functional $F \in \mathcal{C}^k\big(K^\circ, \mathbb{R}\big)$, and let $e_V \in V$ be a vector of length $\|e_V\|_2 = \nu^{1/2}$. A *square-root field* for (F, e_V) consists of a pair of differentiable operator fields

$$(X, S) \in \mathcal{C}^{k-2}\big(K^\circ \times K^{\sharp\circ}, \mathrm{Lin}(E^\sharp, V)\big) \times \mathcal{C}^{k-2}\big(K^\circ \times K^{\sharp\circ}, \mathrm{Lin}(E, V)\big)$$

such that the following diagram commutes for all (x, s):

$$\begin{array}{ccccccc} E^\sharp & \xrightarrow{F''_\sharp(s)} & E & \xrightarrow{F''(w(x,s))} & E^\sharp & \xleftarrow{F''(x)} & E \\ \uparrow{\scriptstyle S^*(x,s)} & & \downarrow{\scriptstyle S(x,s)} & & \downarrow{\scriptstyle X(x,s)} & & \uparrow{\scriptstyle X^*(x,s)} \\ V & = & V & = & V & = & V \end{array} \tag{2.1a}$$

$$\begin{array}{ccc} -F'_\sharp(s) & \xrightarrow{F''(w(x,s))} & -F'(x) \\ \downarrow{\scriptstyle S(x,s)} & & \downarrow{\scriptstyle X(x,s)} \\ e_V & = & e_V \end{array} \tag{2.1b}$$

Our choice of nomenclature reflects the fact that X and S are square-roots of the tensor fields $(x, s) \mapsto F''^{-1}(x)$ and $(x, s) \mapsto F''^{-1}_\sharp(s)$ respectively. Let $B \in \mathcal{B}$ (see Definition 2.5). Then it follows from Proposition 2.9 that there exist orthogonal fields

$$\Omega_{(x,B)}, \Omega_{(s,B)} \in \mathcal{C}^{k-2}\big(K^\circ \times K^{\sharp\circ}, \mathcal{O}(V)\big)$$

such that

$$\begin{aligned} X(x, s) &\equiv \Omega_{(x,B)}(x, s) \circ F''^{-1/2}_B(x), \\ S(x, s) &\equiv \Omega_{(s,B)}(x, s) \circ F''^{-1/2}_{\sharp B}(s), \\ \Omega_{(s,B)} &\equiv \Omega_{(x,B)} \circ F''^{-1/2}_B(x) \circ F''(w) \circ F''^{1/2}_{\sharp B}(s)^*. \end{aligned}$$

These identities combined with Corollary H2.3.10 show that the cones $X(x, s)K^\sharp$ are self-dual under the Euclidean structure defined on V and mutually congruent for all values of (x, s). If all of these cones coincide in a fixed cone $K_V \subset V$ then we say that (X, S) is a *congruent* square-root field for $\big(F, e_V, K_V\big)$. Finally, we shall refer to a square-root field (X, S) as *homogeneous* if $X(\tau x, \eta s) = \tau X(x, s)$ for all $x \in K^\circ, s \in K^{\sharp\circ}$ and $\tau, \eta > 0$ (compare with Proposition H2.3.13*iii*)). Since $w(\tau x, \eta s) = \tau^{1/2}\eta^{-1/2}w(x, s)$ it follows that $S(\tau x, \eta s) = \eta S(x, s)$ in this case. By abuse of language we shall often simply speak of a (congruent) square-root field for F, and it is then tacitly understood that a vector $e_V \in V$ (and a cone $K_V \subset V$) has been fixed.

Theorem 2.11 (Existence of square-root fields).
Let $K \subset E$ and $F \in \mathcal{C}^k\big(K^\circ, \mathbb{R}\big)$ be self-scaled. Then the following applies:

i) The set of congruent square root fields for F is nonempty.

ii) If (X, S) is a square-root field for (F, e_V) then any pair (Y, T) of operator fields is a square-root field for (F, e_V) if and only if there exists a mapping $\Omega \in \mathcal{C}^{k-2}\big(K^\circ \times K^{\sharp\circ}, \mathcal{O}(V, e_V)\big)$ such that $Y \equiv \Omega \circ X$ and $T \equiv \Omega \circ S$.

iii) If (X, S) is a congruent square-root field for (F, e_V, K_V) then any pair (Y, T) of operator fields is a congruent square-root field for (F, e_V, K_V) if and only if there exists a mapping $\Omega \in \mathcal{C}^{k-2}\big(K^\circ \times K^{\sharp\circ}, \mathcal{O}(K_V, e_V)\big)$ such that $Y \equiv \Omega \circ X$ and $T \equiv \Omega \circ S$.

□

3 Target maps

In the linear programming literature the target map is defined as follows:

$$\begin{aligned} \gamma : K^\circ \times K^{\sharp\circ} &\to V \\ (x, s) &\mapsto \big(x_1 s_1 \dots x_n s_n\big)^T. \end{aligned}$$

We will now generalize this notion to self-scaled conic programming and prove bounds on $\|\gamma(x, s)\|$ in terms of the duality gap $\langle x; s \rangle$ (see Theorem 3.3). We will also show that optimal solutions (x^*, s^*) to the optimization problems (1.1) correspond to $\gamma(x, s) = 0$. This suggests the intuitive interpretation of $\|\gamma(x, s)\|$ as a version of the duality gap with directional information about where an optimum can be found from the viewpoint of the current solution.

Definition 3.1. Let K, F be self-scaled and (X, S) a square root field for F. The *target map* defined by (X, S) is given by the function

$$\begin{aligned} \gamma : K^\circ \times K^{\sharp\circ} &\to V \\ (x, s) &\mapsto X(x, s)s = S(x, s)x. \end{aligned}$$

γ is well defined, since for all (x, s), $S(x, s)x = X(x, s) \circ F''(w)x = X(x, s)s$.

Example 3.2. The target map defined by our primordial square-root field (1.4) for the linear programming case coincides with the classical target map defined in the theory of linear programming:

$$\begin{aligned} (x, s) \mapsto \gamma(x, s) &= X(x)s = \mathrm{Diag}(x)s \\ &= \big(x_1 s_1 \dots x_n s_n\big)^T. \end{aligned}$$

□

As usual in the theory of self-concordant barrier functionals, the norm given on E by the Euclidean structure defined by $x \in K^\circ$ (see Definition 2.5) will be denoted by $\|\cdot\|_x$. Since this Euclidean structure is based on identifying $E^\sharp$ with E via the isomorphism $F''^{-1}(x)$, $\|\cdot\|_x$ is also defined on $E^\sharp$. Likewise, we could start with an element $s \in K^{\sharp\circ}$ and define the norm $\|\cdot\|_s$ by the Euclidean structure defined by s on $E^\sharp$ and E. These definitions are standard (see e.g. [23]), and we have $\|\cdot\|_x = \|\cdot\|_{-F'(x)}$ for all $x \in K^\circ$. We shall thus use whichever notation is more convenient in a given circumstance. The norm $\|\cdot\|_2$ on V is fixed by the Euclidean structure defined on this space. For the ensuing operator norms on $\mathrm{Lin}(E,V)$ and $\mathrm{Lin}(E^\sharp,V)$ respectively let us use the notation $\|\cdot\|_e := \|\cdot\|_{\|\cdot\|_e,\|\cdot\|_2}$.

Theorem 3.3. *Let $K \subset E$ be a self-scaled cone, F a ν-self-scaled barrier functional, $e_V \in V$ a fixed vector of norm $\nu^{1/2}$ and (X,S) a square-root field for (F,e_V). Let γ be the target map defined by (X,S), and let e be a fixed element of K°. Then*

i) there exists a positive real number $\rho_1 > 0$ such that for all $(x,s) \in K^\circ \times K^{\sharp\circ}$,

$$\rho_1\langle x;s\rangle \geq \|\gamma(x,s)\| = \|s\|_x = \|x\|_s \geq \nu^{-1/2}\langle x;s\rangle. \tag{3.1}$$

Moreover, the last inequality in (3.1) holds at equality if and only if $s = -\mu F'(x)$ for some positive real number $\mu > 0$, or equivalently if and only if $x = -\mu F_\sharp'(s)$.

ii) there exists a positive real number $\rho_2 > 0$ such that for all $(x,s) \in K^\circ \times K^{\sharp\circ}$ it is true that

$$\|X\|_e \leq \rho_2\|x\|_e \qquad \textit{and} \tag{3.2}$$

$$\|S\|_e \leq \rho_2\|s\|_e. \tag{3.3}$$

Proof. Note that since by Theorem 2.11 distinct square-root fields for F differ only by an orthogonal factor, the whole statement of Theorem 3.3 is independent of the particular choice of (X,S). We may thus assume without loss of generality that (X,S) is the square-root field identified in the proof of Theorem 2.11:

$$X(x,s) = R \circ F''^{-1}\big(w(e,-F'(x))\big)$$

for some $R \in F''^{1/2}(e)$. Likewise, we may assume without loss of generality that $e_V = R[e]$.

i) The Cauchy-Schwartz inequality and Proposition H2.3.13*ii)* imply that

$$\langle x;s\rangle = \big(X(x,s)s; X^{-*}(x,s)x\big) \overset{H2.3.13}{=} \big(\gamma(x,s);e_V\big) \overset{C.S.}{\leq} \nu^{1/2}\|\gamma(x,s)\|_2. \tag{3.4}$$

This shows the last inequality in (3.1). Equality holds if and only if the Cauchy-Schwartz inequality holds at equality, i.e. exactly when $\gamma(x,s) = \mu e_V$ for some positive real number $\mu > 0$. This is equivalent to the condition that

$$s = X^{-1}\gamma = \mu X^{-1}e_V = -\mu F'(x).$$

The equalities in the middle part of (3.1) can be seen to hold as follows:

$$\begin{aligned}\|s\|_x^2 &= \langle F''^{-1}(x)s; s\rangle = \langle X^* \circ Xs; s\rangle \\ &= (Xs; Xs) = \|\gamma\|_2^2 \\ &= (Sx; Sx) = \cdots = \|x\|_s^2.\end{aligned}$$

Now, for all $(x, s) \in K^\circ \times K^{\sharp\circ}$ we have $X(x, s)K^\sharp = R[K]$. By virtue of Corollary H2.3.10 $R[K]$ is a regular self-dual cone containing $e_V = R[e]$ in its interior. Therefore, $\{v \in R[K] : (e_V; v) = 1\}$ is a compact convex set of affine dimension $(n-1)$ that doesn't meet the origin. Let ρ_1 be the maximum 2-norm of a vector in this set. Then equation (3.1) is implied as follows

$$\|\gamma(x, s)\|_2 \leq \rho_1 \cdot (e_V; \gamma(x, s)) = \rho_1 \cdot \langle x; s\rangle.$$

ii) Let $B \in \mathcal{B}^{(e)}$ (see Definition 2.5). Since $X(x, s) = \Omega_{(x,B)} \circ F_B''^{-1/2}(x)$ (see Definition 2.10), it is enough to show that

$$\|F_B''^{-1/2}(x)\|_e < \rho_2 \cdot \|x\|_e. \tag{3.5}$$

Let $S_e^{(n-1)}(E)$ and $S_e^{(n-1)}(E^\sharp)$ respectively denote the unit spheres around the origin in E and $E^\sharp$, taken with respect to the norm $\|\cdot\|_e$. It suffices to show that inequality (3.5) holds for $x \in K^\circ \cap S_e^{(n-1)}(E)$, since $F''(\tau x) = \tau^{-2}F''(x)$ for all $\tau > 0$ (see Proposition 2.1 *vi*)). $K \cap S_e^{(n-1)}(E)$ is a compact set because K is closed.
Let $\ker e^\sharp$ denote the kernel of the linear form form $e^\sharp$ on E. Then, for every $x \in K$ the set $(x + \ker e^\sharp) \cap K$ is compact and the function

$$\begin{aligned}\rho : K &\to \mathbb{R} \\ x &\mapsto \max\{\|z\|_e : z \in (x + \ker e^\sharp) \cap K\}\end{aligned}$$

is continuous. Let $\rho_2 := \max\{\rho(x) : x \in K \cap S_e^{(n-1)}(E)\} + 1$.
Now, consider a vector $x \in K^\circ \cap S_e^{(n-1)}(E)$. Since $\nabla_B^2 F(x)$ is positive definite symmetric we have

$$\sigma_{\max}\big(\nabla_B^2 F(x)^{-1}\big) = \Big[\sigma_{\min}\big(\nabla_B^2 F(x)\big)\Big]^{-1},$$

where $\sigma_{\max}$ and $\sigma_{\min}$ denote the largest and smallest eigenvalues respectively. It follows that

$$\begin{aligned}\|F_B''^{-1/2}(x)\|_e^2 &= \max\{\langle F''^{-1}(x)q; q\rangle : q \in S_e^{(n-1)}(E^\sharp)\} \\ &= \sigma_{\max}\big(\nabla_B^2 F(x)^{-1}\big) = \Big[\sigma_{\min}\big(\nabla_B^2 F(x)\big)\Big]^{-1} \\ &= \Big(\min\{\langle p; F''(x)p\rangle : p \in S_e^{(n-1)}(E)\}\Big)^{-1} \\ &= \langle p_0; F''(x)p_0\rangle^{-1}\end{aligned}$$

for some $p_0 \in S_e^{(n-1)}(E)$. But then we have $\langle p_1; F''(x)p_1 \rangle = 1$ for

$$p_1 := \|F_B''^{-1/2}(x)\|_e p_0,$$

and since F is self-concordant it follows from Definition H1.2.1 that $x + p_1$ and $x - p_1$ are elements of K. At least one of them must lie in

$$K \cap \{p \in E : \langle p; e^\sharp \rangle \leq \langle x; e^\sharp \rangle\} \subset \bar{B}_{\rho_2 - 1}(0),$$

where the ball $B_{\rho_2 - 1}$ is taken with respect to the norm $\|\cdot\|_e$ of E. Therefore,

$$\begin{aligned} \|F_B''^{-1/2}(x)\|_e &= \|p_1\|_e \\ &\leq \rho_2 - 1 + \|x\|_e = \rho_2, \end{aligned}$$

and this shows that equation (3.5) holds for $x \in K^\circ \cap S_e^{(n-1)}(E)$ indeed. □

Consider the dual pair of optimization problems (P), (D) (1.1). We are now going to show how (P) and (D) can be treated jointly in the framework of V-space. As a result we obtain a new family of Newton directions for the central path equations. Recall the notation $\mathcal{F}(P), \mathcal{F}^\circ(P)$ etc. from Section 1. Recall also that we fixed elements $x_0 \in \mathcal{F}^\circ(P)$ and $s_0 \in \mathcal{F}^\circ(D)$ in (1.1). We start by embedding the set of primal-dual feasible tuples $(x, s) \in \mathcal{F}(P) \times \mathcal{F}(D)$ in a n-dimensional Euclidean space. Let $e \in K^\circ$ be fixed and let us use the following notation:

1. $Z := L \oplus L^\perp$ for the *primal-dual* space,
2. $\pi_L : Z \to L$ for the projection of Z along $L^\perp$ onto L,
3. $\pi_{L^\perp} = \imath - \pi_L$ for the projection of Z along L onto $L^\perp$,
4. for the Euclidean structure on Z defined by e:

$$\begin{aligned} \big(\cdot;\cdot\big)_e : Z \times Z &\to \mathbb{R}, \\ (z_1, z_2) &\mapsto \langle \pi_L z_2; F''(e)\pi_L z_1 \rangle + \langle F''^{-1}(e)\pi_{L^\perp} z_2; \pi_{L^\perp} z_1 \rangle, \end{aligned}$$

5. for the primal part of z:

$$\begin{aligned} x : Z &\to E, \\ z &\mapsto x_0 + \pi_L z, \end{aligned}$$

6. for the dual part of z:

$$\begin{aligned} s : Z &\to E^\sharp \\ z &\mapsto s_0 + \pi_{L^\perp} z, \end{aligned}$$

7. for the duality gap $\delta \in \mathrm{Lin}(Z, \mathbb{R}) : z \mapsto \delta(z) := \langle x(z); s(z) \rangle$,

8. $\mathcal{F}(PD) := (x,s)^{-1}(\mathcal{F}(P) \times \mathcal{F}(D))$ for the primal-dual feasible domain,

9. $\mathcal{F}^\circ(PD) := (x,s)^{-1}(\mathcal{F}^\circ(P) \times \mathcal{F}^\circ(D))$ for the primal-dual strictly feasible domain.

In particular, L and $L^\perp$ are perpendicular in the Euclidean structure defined by e on E.

Lemma 3.4. *$\mathcal{F}(PD)$ is a convex set whose interior with respect to the topology defined by $(\cdot;\cdot)$ on E is equal to $\mathcal{F}^\circ(PD)$. The mapping*

$$\begin{aligned} \mathfrak{F} : \mathcal{F}^\circ(PD) &\to \mathbb{R} \\ z &\mapsto F\big(x(z)\big) + F_\sharp\big(s(z)\big) \end{aligned}$$

is a ν-self-concordant barrier functional for $\mathcal{F}^\circ(PD)$. Moreover, the following primal-dual optimization problem is equivalent to the problems (P) *and* (D) *from* (1.1)*:*

$$\text{(PD)} \qquad \inf\{\delta(z) : z \in \mathcal{F}(PD)\}.$$

Finally, if η is a positive real number, then $\mathcal{F}(PD) \cap \delta^{-1}\big([0,\eta]\big)$ and $\mathcal{F}(PD) \cap \delta^{-1}\big(\{\eta\}\big)$ are compact subsets of Z.

Proof. All statements are well known and completely trivial to prove. □

We shall henceforth restrict our attention to feasible points (x,s) and use $\mathcal{F}(PD) \subset Z$ for the parameterization of all the relevant functions. For example, $w(z) := w\big(x(z), s(z)\big)$ for all $z \in \mathcal{F}^\circ(PD)$. Let us extend the functional $\mathfrak{F}$ to Z in the usual fashion, i.e. by setting $\mathfrak{F}(z) := +\infty$ outside of $\mathcal{F}^\circ(PD)$. The guiding principle of the barrier method is to replace the constrained problem (PD) by the directed set of perturbed unconstrained problems $\{(\text{PD})_\mu : \mu > 0\}$ where

$$(\text{PD})_\mu \qquad \min\{\delta(z) + \mu\mathfrak{F}(z) : z \in Z\}.$$

As usual, the directed set of optimal solutions of $\{(\text{PD})_\mu : \mu > 0\}$ parameterized by μ, $\mu \mapsto z_\mu$ is called the *central path* for the tuple $\big((\text{PD}), \mathfrak{F}\big)$. The central paths $\mu \mapsto x_\mu$ and $\mu \mapsto s_\mu$ for the tuples $\big((\text{P}), F\big)$ and $\big((\text{D}), F_\sharp\big)$ are defined in the same fashion. It is well known that for $\mu > 0$ the pair $\big(x_\mu, s_\mu\big)$ is characterized as the unique feasible solution of the central path equation

$$x = -\mu F_\sharp'(s), \tag{3.6}$$

or equivalently of the equation $s = -\mu F'(x)$. Let (X,S) be a square-root field for F, and let γ be the associated target map. Multiplication of equation (3.6) by S on the left yields the primal-dual central path equation

$$\gamma(z) = \mu e_V. \tag{3.7}$$

Proposition 3.5. *Let $\mu > 0$ and $z \in \mathcal{F}^\circ(PD)$. The following conditions are equivalent:*

i) $z = z_\mu$,

ii) $x_\mu = x(z)$ *and* $s_\mu = s(z)$,

iii) z *solves equation* (3.7),

iv) z *minimizes the function* $\nu^{1/2}\|\gamma\|_2 + \mu\mathfrak{F}$.

Proof. It is well-known and easy to prove that *i)* $\Leftrightarrow$ *ii)*, and since (3.6) characterizes (x_μ, s_μ) and S is non-singular, clearly *ii)* $\Leftrightarrow$ *iii)*. It follows from part *ii)*, equation (3.6) and Theorem 3.3 that the inequality

$$\nu^{1/2}\|\gamma(z)\|_2 + \mu\mathfrak{F}(z) \geq \delta(z) + \mu\mathfrak{F}(z),$$

holds at equality if and only if $z = z_{\hat{\mu}}$ for some $\hat{\mu} > 0$, and since the right-hand side of (3) is minimized uniquely at $z = z_\mu$ this shows that *iv)* $\Leftrightarrow$ *i)*. □

Let us extend γ by continuity to all the boundary points $z \in \partial\mathcal{F}(PD)$ where the limit of γ is uniquely defined. In particular it follows from Theorem 3.3 that this is possible for all optimal solutions of (PD) if there are any, and that the following extension of Proposition 3.5 to the parameter value $\mu = 0$ holds true:

Proposition 3.6. *The vector* $z \in \mathcal{F}(PD)$ *is optimal for (PD) if and only if* $\gamma(z) = 0$. □

If the barrier functional F is of class $\mathcal{C}^k$ then the target map lies in $\mathcal{C}^{k-2}\big(\mathcal{F}^\circ(PD), V\big)$. Hence, for $k \geq 3$ we can apply the Newton method to solve equation (3.7). For a fixed $\mu > 0$ the corresponding Newton system reads

$$\gamma'(z)d_z = \mu e_V - \gamma(z). \tag{3.8}$$

This equation has a unique solution at least in a neighborhood of the central path, as follows from continuity of γ' and the following lemma:

Lemma 3.7. *Let* (X, S) *be a square-root field for* $F \in \mathcal{C}^3(K^\circ, \mathbb{R})$, *and let* $\gamma : \mathcal{F}^\circ(PD) \to V$ *be its associated target map. If* $z \in \mathcal{F}^\circ(PD)$ *lies on the central path then* $\gamma'(z)$ *is non-singular.*

Proof. Consider the identity $X(x,s)[-F'(x)] = e_V$ which must hold identically by Definitions H2.3.11 and 2.10. Taking the partial derivative with respect to s we get

$$\Big(\frac{\partial}{\partial s}X[d_s]\Big)[-F'(x)] = 0 \qquad \text{for all } d_s \in E^\sharp, \tag{3.9}$$

whereas taking the partial derivative with respect to x yields

$$\Big(\frac{\partial}{\partial x}X[d_x]\Big)[-F'(x)] = X \circ F''(x)[d_x] \qquad \text{for all } d_x \in E. \tag{3.10}$$

The derivative of $\gamma(z)$ in the direction $d_z \in Z$ is given by

$$\begin{aligned}\gamma'(z)[d_z] =& \Big(\frac{\partial}{\partial x} X\big(x(z), s(z)\big)[\pi_L d_z]\Big) s(z) \\ &+ \Big(\frac{\partial}{\partial s} X\big(x(z), s(z)\big)[\pi_{L^\perp} d_z]\Big) s(z) + X(z)\pi_{L^\perp} d_z.\end{aligned} \tag{3.11}$$

But for the point $z = z_\mu$ on the central path corresponding to the parameter value $\mu > 0$ it is true that $s(z) = -\mu F'(x(z))$ (see (3.6)). We can therefore substitute equations (3.10) and (3.9) in (3.11) to get

$$\gamma'(z_\mu)[d_z] = X(z_\mu) \circ \big(\mu F''(x_\mu) \circ \pi_L + \pi_{L^\perp}\big).$$

Since $X(z_\mu)$ is non-singular it only remains to show that $\mu F''(x_\mu) \circ \pi_L + \pi_{L^\perp}$ is invertible. Suppose $d_z \in Z$ is in the kernel of this operator. Then

$$\|\pi_L d_z\|_{x_\mu}^2 = \big\langle \pi_L d_z; F''(x_\mu) \circ \pi_L d_z \big\rangle = -\mu^{-1} \big\langle \pi_L d_z; \pi_{L^\perp} d_z \big\rangle = 0.$$

Hence, $\pi_L d_z = 0$, and this implies that

$$d_z = \pi_L d_z + \pi_{L^\perp} d_z = \big(\imath - \mu F''(x_\mu)\big)\pi_L d_z = 0.$$

□

Let us now further investigate the structure of the Jacobian γ'. Let us choose a basis $B \in \mathcal{B}^{(e)}$ for some fixed element $e \in K^\circ$ and recall from Definition 2.10 that $X \equiv \Omega_{(x,B)} \circ F''^{-1/2}(x)$ and $S \equiv \Omega_{(s,B)} \circ F_\sharp''^{-1/2}(s)$. For all $z \in \mathcal{F}^\circ(PD)$ and $d_z \in Z$ it is therefore true that

$$\begin{aligned}\gamma'(z)[d_z] =& \gamma'(z)[\pi_L d_z] + \gamma'(z)[\pi_{L^\perp} d_z] \\ =& S(z) \circ \pi_L[d_z] + \Omega'_{(s,B)}(z)[\pi_L d_z] \circ \Omega^*_{(s,B)}(z) \circ \gamma(z) \\ &+ X(z) \circ \pi_{L^\perp}[d_z] + \Omega'_{(x,B)}(z)[\pi_{L^\perp} d_z] \circ \Omega^*_{(x,B)}(z) \circ \gamma(z).\end{aligned}$$

Definition 3.8. Assume the self-scaled barrier functional F is of class $\mathcal{C}^k$. Then we can define operator fields $M, R \in \mathcal{C}^{k-2}\big(\mathcal{F}^\circ(PD), \mathrm{Lin}(Z, V)\big)$ by

$$\begin{aligned}M(z) : d_z &\mapsto S(z) \circ \pi_L[d_z] + X(z) \circ \pi_{L^\perp}[d_z], \\ R(z) : d_z &\mapsto \big(\Omega'_{(s,B)}(z)[\pi_L d_z] \circ \Omega^*_{(s,B)}(z) + \Omega'_{(x,B)}(z)[\pi_{L^\perp} d_z] \circ \Omega^*_{(x,B)}(z)\big)\gamma(z),\end{aligned}$$

so that $\gamma' \equiv M + R$. We call M and R the main and residual part of the Jacobian γ' respectively.

Proposition 3.9. *Let us use the same notation as in Definition 3.8. Then for all $z \in \mathcal{F}^\circ(PD)$ it is true that*

i) $M(z)$ is a non-singular operator,

ii) $\operatorname{im}\big(R(z)\big) \subseteq \operatorname{span}\{e_V, \gamma(z)\}^\perp$,

iii) if z *lies on the central path, then* $R(z) = 0$,

iv) $\dim\big(\ker R(z)\big) \geq 2$.

Proof. *i)* As we shall see in (4.1), part *i)* is equivalent to claiming that the Nesterov-Todd system is non-singular. This is indeed a well-known result (see [31]), but we shall include a proof for the sake of completeness. Suppose $d_z \in \ker\big(M(z)\big)$. Then $S(z) \circ \pi_L[d_z] = -X(z) \circ \pi_{L^\perp}[d_z]$. Recall from the introduction that we defined the operator $A \in \operatorname{Lin}(E, Y^\sharp)$ and $L := \ker(A)$. Likewise, we defined $L^\perp := \operatorname{im}\big(A^*\big)$, where $A^* : Y^\sharp \to E^\sharp$ is the adjoint of A. Therefore, we get

$$\begin{aligned} 0 = A \circ \pi_L[d_z] &= A \circ S^{-1}(z) \circ S(z) \circ \pi_L[d_z] \\ &= -A \circ S^{-1}(z) \circ X(z) \circ \pi_{L^\perp}[d_z]. \end{aligned}$$

But since $L^\perp$ was defined as the image of A^* there exists an element $y \in Y$ such that $\pi_{L^\perp} d_z = A^* y$. Now, $S^{-1}(z) \circ X(z) = F''_\sharp\big(t(z)\big)$ is a positive definite operator, and so is $A \circ F''_\sharp\big(t(z)\big) \circ A^*$, since A^* is injective. But $\big\langle y; A \circ F''_\sharp\big(t(z)\big) \circ A^* y\big\rangle = 0$, and hence $y = 0$. This implies that $\pi_{L^\perp} d_z = 0$ and $S(z) \circ \pi_L[d_z] = -X(z)0 = 0$. Since $S(z)$ is non-singular it follows that $\pi_L d_z = 0$ and $d_z = \big(\pi_L + \pi_{L^\perp}\big)[d_z] = 0$. This implies the claim.

ii) For all $d_z \in Z$ we have

$$\begin{aligned} \big(R(z)[d_z]; e_V\big) + \big(M(z)[d_z]; e_V\big) &= \big(\gamma'(z)[d_z]; e_V\big) = \frac{d}{dz}\big(\gamma(z); e_V\big)[d_z] \\ &\overset{(3.4)}{=} \frac{d}{dz}\delta(z)[d_z] = \big\langle x(z); \pi_{L^\perp} d_z\big\rangle + \big\langle \pi_L d_z; s(z)\big\rangle \\ &= \big(X^{-*}(z)x(z); X(z) \circ \pi_{L^\perp}[d_z]\big) + \big(S(z) \circ \pi_L[d_z]; S^{*-1}(z)s(z)\big) = \\ &\overset{H2.3.13ii)}{=} \big((S(z) \circ \pi_L + X(z) \circ \pi_{L^\perp})[d_z]; e_V\big) = \big(M(z)[d_z]; e_V\big). \end{aligned}$$

Comparing both sides of this equation we get that $\big(R(z)[d_z]; e_V\big) = 0$. But since this holds for all d_z it follows that $\operatorname{im}\big(R(z)\big) \perp e_V$. On the other hand,

$$\begin{aligned} \big(R(z)[d_z]; \gamma(z)\big) + \big(M(z)[d_z]; \gamma(z)\big) &= \big(\gamma'(z)[d_z]; \gamma(z)\big) \\ &= \frac{1}{2}\frac{d}{dz}\|\gamma(z)\|^2[d_z] \\ &\overset{(3.1)}{=} \frac{1}{2}\frac{d}{dz}\|x(z)\|^2_{s(z)}\big[\pi_L d_z\big] + \frac{1}{2}\frac{d}{dz}\|s(z)\|^2_{x(z)}\big[\pi_{L^\perp} d_z\big] \\ &= \big\langle x(z); F''^{-1}_\sharp(s(z)) \circ \pi_L[d_z]\big\rangle + \big\langle F''^{-1}(x(z)) \circ \pi_{L^\perp}[d_z]; s(z)\big\rangle \\ &\overset{H2.3.16}{=} \big\langle x(z); S(z)^* \circ S(z) \circ \pi_L[d_z]\big\rangle + \big\langle X(z)^* \circ X(z) \circ \pi_{L^\perp}[d_z]; s(z)\big\rangle \\ &= \big((S(z) \circ \pi_L + X(z) \circ \pi_{L^\perp})[d_z]; \gamma(z)\big) \\ &= \big(M(z)[d_z]; \gamma(z)\big). \end{aligned}$$

Comparing both sides we get $\big(R(z)[d_z]; \gamma(z)\big) = 0$, and since d_z was arbitrary this shows that $\operatorname{im}\big(R(z)\big) \perp \gamma(z)$.

iii) Let $z = z_\mu$ for some $\mu > 0$. Lemma 3.7 says that the Jacobian $\gamma'(z)$ is non-singular, and since γ is continuously differentiable the inverse mapping theorem implies that there exist open neighborhoods $z \in O_z \subset \mathcal{F}^\circ(PD)$ and $\mu e_V \in O_{\mu e_V} \subset V$ such that $\gamma|_{O_z} : O_z \to O_{\mu e_V}$ is a homoeomorphism. Now, choose two arbitrary vectors $d \in Z$ and $0 \neq v \in V$. We can choose a sequence $\big(t_k\big)_{\mathbb{N}} \subset \mathbb{R}_{++}$ such that $\lim_{k\to\infty} t_k = 0$ and $\mu e_V + t_k v \in O_{\mu e_V}$ for all $k \in \mathbb{N}$. Let

$$z_k := \big(\gamma|_{O_z}\big)^{-1}(\mu e_V + t_k v).$$

Since $\big(\gamma|_{O_z}\big)^{-1}$ is continuous it must be true that $\lim_{k\to\infty} z_k = z_\mu$. But part *ii*) says that

$$R(z_k)d \in \operatorname{span}\{e_V, \mu e_V + t_k v\}^\perp = \operatorname{span}\{e_V, v\}^\perp \qquad \text{for all } k \in \mathbb{N}.$$

Since R is a continuous operator field on $\mathcal{F}^\circ(PD)$, and since $\operatorname{span}\{e_V, v\}^\perp$ is a closed subset of V we get

$$R(z)[d] = \lim_{k\to\infty} R(z_k)[d] \in \operatorname{span}\{e_V, v\}^\perp.$$

But v was chosen arbitrarily, whence $R(z)[d] = 0$. Finally, since d was chosen arbitrarily we get $R(z) \equiv 0$.

iv) For z not on the central path this follows immediately from part *ii*), whereas for z on the central path the claim follows from part *iii*). □

Next we are going to show that the search directions obtained from the Newton system $\gamma' d_z = \mu e_V - \gamma$ lend themselves as building blocks in predictor-corrector algorithms. For this it is important to note that equation (3.4) says that $\delta(z) = \big(\gamma(z); e_V\big)$, whence for a fixed μ we have

$$\gamma\big(\delta^{-1}(\mu) \cap \mathcal{F}^\circ(PD)\big) \subset \big\{v \in V : \big(v; e_V\big) = \mu\big\}. \tag{3.12}$$

For all $z \in \mathcal{F}^\circ(PD)$ let us define

$$\mu(z) := \nu^{-1}\delta(z). \tag{3.13}$$

Then $\mu(z)$ is the unique value of the parameter $\mu > 0$ such that the corresponding point on the central path satisfies $\delta(z_\mu) = \delta(z)$. In fact,

$$\delta\big(z_{\mu(z)}\big) = \big(\gamma(z_{\mu(z)}); e_V\big) = \big(\mu(z)e_V; e_V\big) = \mu(z)\nu = \delta(z).$$

Equation (3.12) implies that

$$\gamma'(z)\big(\ker\delta\big) \subset \{e_V\}^\perp,$$

and in the case where $\gamma'(z)$ is non-singular we get $\gamma'(z) \in \operatorname{Iso}\big(\ker\delta, \operatorname{span}\{e_V\}^\perp\big)$ simply by an argument on the dimensions of these spaces. Since we cannot assume that $\gamma'(z)$ is non-singular in general, it will be interesting to note that

$$\gamma'(z)^{-1}\big(\operatorname{span}\{e_V\}^\perp\big) \subset \ker\delta$$

(see Proposition 3.11 *i)* below). This shows that if the Newton system

$$\gamma'(z)d_z = \mu(z)e_V - \gamma(z) \tag{3.14}$$

has a solution d_z then the duality gap remains constant if we take a step in this direction. This has the following important consequence:

Remark 3.10. System (3.14) is often used in algorithms when the current iterate falls too far from the central path where only small improvements in the objective function can be made in general. A few iterations in the direction defined by (3.14) bring the iterates back toward the central path without a change in the objective function:

$$\mu\big(z^+\big) = \mu(z), \qquad\qquad \delta\big(z^+\big) = \delta(z). \tag{3.15}$$

Algorithms that exploit this idea are often called "predictor-corrector" methods.

Proposition 3.11. *Let (X, S) be a square-root field for (F, e_V), let γ be the associated target map, and let $z \in \mathcal{F}^\circ(PD)$. Then it is true that*

i) $\gamma'^{-1}(z)\big(\operatorname{span}\{e_V\}^\perp\big) \subset \ker\delta$,

ii) $M(z) \in \operatorname{Iso}\big(\ker\delta, \operatorname{span}\{e_V\}^\perp\big)$,

iii) $\ker\big(\gamma'(z)\big) \subset \ker\delta$,

iv) $\dim\big(\operatorname{im}(\gamma'(z))\big) \geq 2$.

Proof. i) Suppose $d_z \in Z$ is such that $\gamma'(z)[d_z] \in \operatorname{span}\{e_V\}^\perp$. Then, for all $\tau > 0$ small enough $z + \tau d_z \in \mathcal{F}^\circ(PD)$. Hence,

$$\begin{aligned}
\delta(d_z) &\overset{\delta\in Z^\sharp}{\equiv} \tau^{-1}\big[\delta(z+\tau d_z) - \delta(z)\big] \\
&\overset{(3.4)}{\equiv} \tau^{-1}\big[\big(\gamma(z+\tau d_z); e_V\big) - \big(\gamma(z); e_V\big)\big] \\
&= \lim_{\tau\to 0} \tau^{-1}\big(\gamma(z+\tau d_z) - \gamma(z); e_V\big) \\
&= \big(\gamma'(z)[d_z]; e_V\big) = 0.
\end{aligned}$$

ii) This part extends the well known identity

$$\mu\big(z + M(z)^{-1}[\mu(z)e_V - \gamma(z)]\big) = \mu(z) \tag{3.16}$$

which guarantees that the objective function doesn't change during a corrector phase of a Nesterov-Todd type algorithm. Let $d_z \in Z$ such that $M(z)[d_z] \perp e_V$. Then, by Proposition 3.9 *ii)* we have $\gamma'(z)[d_z] = M(z)[d_z] + R(z)[d_z] \perp e_V$. Hence, part *i)* implies that $d_z \in \ker\delta$, and it follows that $M(z)^{-1}\big(\operatorname{span}\{e_V\}^\perp\big) \subset \ker\delta$. But $\dim\big(\ker\delta\big) = n - 1 = \dim\big(\operatorname{span}\{e_V\}^\perp\big)$, and since $M(z)$ is non-singular by Proposition 3.9 *i)* it follows that the claim is true.

iii) Since $0 \perp e_V$ it follows from part *i)* that $\ker\big(\gamma'(z)\big) = \gamma'^{-1}(z)\big(\{0\}\big) \subset \ker\delta$.

iv) This part follows trivially from Proposition 3.9 *i)* & *ii)* and the identity $\gamma' \equiv M + R$. □

4 Some remarks concerning the Nesterov-Todd method

Let us now compare the Newton system (3.8) with the Nesterov-Todd (N-T) equations (1.3a)–(1.3c). Equations (1.3b) and (1.3c) can be rewritten as $(d_x^{(\mu)} \in L, d_s^{(\mu)} \in L^\perp)$ or equivalently, $d_z^{(\mu)} := \big(d_x^{(\mu)}, d_s^{(\mu)}\big) \in Z$. Let us choose a square-root field (X, S) for (F, e_V) and multiply equation (1.3a) by the operator $X(x, s)$ from the left. If we use the parameterization in Z-space, the rescaled Nesterov-Todd system becomes

$$M(z)d_z^{(\mu)} = \mu e_V - \gamma(z), \quad (d_z^{(\mu)} \in Z). \tag{4.1}$$

On the other hand, writing the Newton-system (3.8) in terms of $x(z), s(z)$ and using $d_z^{(\mu)} = d_x^{(\mu)} \oplus d_s^{(\mu)} \in L \oplus L^\perp$ we get

$$\Big(M(z) + R(z)\Big)d_z^{(\mu)} = \mu e_V - \gamma(z), \quad (d_z^{(\mu)} \in Z). \tag{4.2}$$

Hence, the only difference between the Nesterov-Todd system (4.1) and the Newton system (4.2) is the term $R(z)d_z^{(\mu)}$ on the left-hand side of (4.2). Let us note two interesting consequences of this observation:

First, it follows from Proposition 3.9 *iii)* that the N-T direction and the Newton direction corresponding to the solution of (4.2) converge to one another as (x, s) approaches any point on the central path. In a forthcoming paper we will show that the square-root field (X, S) can be chosen so that the N-T direction lies in the kernel of $R(z)$ at all feasible points $z \in \mathcal{F}^\circ(PD)$ (recall 3.9 *iv)*), and hence it follows that the N-T directions are Newton directions because they solve system (4.2) for this particular choice of the target map. It is however possible to investigate the convergence rates of the N-T process to points on the central path solely by comparing equations (4.1) and (4.2). Let us assume that $F \in \mathcal{C}^4$ and suppose that for a fixed $\mu > 0$ we aim at solving the problem $(PD)_\mu$ by updating iterates by the rule $z^+ = z + u^{(\mu)}(z)$, where $u^{(\mu)}(z) \in Z$ solves system (4.2). Then γ is twice continuously differentiable, and the standard theory for the convergence of the Newton method says that if the process is started at z_0 close enough to z_μ then the resulting sequence of iterates converges to this point at a quadratic rate (see, e.g., [5] Theorem 5.2.1). If instead we use the update $z^+ = z + d^{(\mu)}(z)$, where $d^{(\mu)}(z) \in Z$ is the Nesterov-Todd direction obtained as the solution to system (4.1), then the resulting iterative process will also converge to the point z_μ at a quadratic rate (see Theorem 4.1 below). This result, already established by Nesterov and Todd [24] in a slightly different form, follows trivially from Proposition 3.9 and a second-order Taylor development via a procedure that is standard in the theory of approximate Newton processes.

Theorem 4.1 (Quadratic Convergence of N-T Process). *Let $F \in \mathcal{C}^4$ be a self-scaled barrier functional for the self-scaled cone K. Let $\mu > 0$ be fixed. For each $z \in \mathcal{F}^\circ(PD)$ let $d^{(\mu)}(z) \in Z$ be the solution to the N-T system* (4.1). *Then $d^{(\mu)}(z) = z_\mu - z + O\big(\|z_\mu - z\|^2\big)$.*

Proof. Let (X, S) be a square-root field for (F, e_V) for some $e_V \in V$. Consider the

associated target map $\gamma : Z \to V$. Then $\gamma \in \mathcal{C}^2$. Therefore,

$$\begin{aligned}
d^{(\mu)}(z) &\stackrel{def}{=} M^{-1}(z)[\mu e_V - \gamma(z)] \\
&\stackrel{\gamma\in\mathcal{C}^2}{=} M^{-1}(z)\big[\gamma(z_\mu) - \big(\gamma(z_\mu) + \gamma'(z_\mu)[z - z_\mu] + O(\|z_\mu - z\|^2)\big)\big] \\
&\stackrel{M^{-1}\in\mathcal{C}^0}{=} -M^{-1}(z) \circ \gamma'(z_\mu)[z - z_\mu] + O(\|z_\mu - z\|^2) \\
&\stackrel{M^{-1}\in\mathcal{C}^1}{=} -\big(M^{-1}(z_\mu) + O(\|z_\mu - z\|)\big) \circ \big(M(z_\mu) + R(z_\mu)\big)[z - z_\mu] + O(\|z_\mu - z\|^2) \\
&\stackrel{3.9iii)}{=} z_\mu - z + O(\|z_\mu - z\|^2).
\end{aligned}$$

□

A second implication of equation (4.1) concerns the numerical stability of the N-T equations. In its standard form equation (1.3a) blows up near the boundary of the feasible domain. For instance, if $x \to \partial\mathcal{F}^\circ(P)$ then $\|F'\|, \|F''(w)\| \to \infty$. Note that as in the primordial linear programming example this problem is taken care of if we consider the N-T equations in their rescaled form (4.1). In fact, let $e \in K^\circ$ be fixed. Then Theorem 3.3 shows that all terms of equation (4.1) are bounded in $\|x\|_e$ and $\|s\|_e$:

$$\begin{gathered}
\|M(z)\|_e \le \|S(z)\|_e + \|X(z)\|_e \stackrel{(3.2),(3.3)}{\le} \rho_2\big(\|x(z)\|_e + \|s(z)\|_e\big), \\
\|\mu e_V\| = \mu\nu^{1/2}, \\
\|\gamma(z)\| \stackrel{(3.1)}{\le} \rho_1\delta(z) \stackrel{C.S.}{\le} \rho_1\|s(z)\|_e\|x(z)\|_e.
\end{gathered}$$

5 Weighted analytic centers

Consider the target map $\gamma : \mathcal{F}^\circ(PD) \to V$ again, and recall from Section 3 that in the general case we are only guaranteed that the Jacobian γ' is non-singular in a neighborhood of the central path; see Lemma 3.7. Let us now investigate the case where γ' is nonsingular everywhere on $\mathcal{F}^\circ(PD)$ and where γ is defined by a congruent square-root field. In this case the target map generalizes the notion of weighted centers to self-scaled conic programming. The main result of this section is Theorem 5.3.

Weighted centers were introduced to the linear programming literature by Kojima, Mizuno and Yoshise [14]. There they play an important role in a unifying framework for the analysis of interior-point methods, an approach that was developed by Kojima et al. [15] and Jansen et al. [11]. For a complete treatment see also [25]. Considerable interest has arisen recently in generalizing weighted centers to semidefinite programming, and several competing approaches have been developed by Monteiro-Pang [21], Sturm-Zhang [28], Monteiro-Zanjacomo [19], Burer-Monteiro [1] and Burer-Monteiro-Zhang [2, 3]. In this section we present a generalization of weighted centers to the more general framework of self-scaled conic programming.

Lemma 5.1. *Let (X, S) be a congruent square-root field for the ν-self-scaled barrier functional $F \in \mathcal{C}^3(K^\circ, \mathbb{R})$. Suppose $(x_k, s_k)_\mathbb{N} \subset K^\circ \times K^{\sharp\circ}$ is a sequence such that $\lim_{k\to\infty}(x_k, s_k) \in \partial\big(K \times K^\sharp\big)$ and $\lim_{k\to\infty} \gamma(x_k, s_k) = v$ exist. Then $v \in \partial K_V$.*

Proof. Let $x := \lim x_k$ and $s := \lim s_k$. Since $(x, s) \in \partial\,\big(K \times K^\sharp\big)$, either $x \in \partial K$ or $s \in \partial K^\sharp$ or both. Without loss of generality we may assume that $s \in \partial K^\sharp$. Let $e \in K^\circ$ be fixed and choose a basis $B \in \mathcal{B}^{(e)}$ for E, orthogonal for the Euclidean structure defined by e. Consider the union of spectra

$$\Sigma := \bigcup_{k\in\mathbb{N}} \sigma\big(\nabla^2_B F(x_k)^{-1/2}\big).$$

Then there exists a positive real number $\rho > 0$ such that Σ is contained in the compact set $\bar{B}_\rho(0) \cap \mathbb{R}^n_+ \subset \mathbb{R}^n$. In fact, $(x_k)_\mathbb{N} \cup \{x\}$ is contained in the compact set $K \cap \bar{B}_{\rho'}(0)$ where $\rho' := \sup \|x_k\|_e$, and hence, it follows from Proposition 3.3 *ii)* that

$$\begin{aligned}\sup\big\{\sigma_{\max}\big(\nabla^2_B F(x_k)^{-1/2}\big) : k \in \mathbb{N}\big\} &= \sup\big\{\big\|X(x_k, s_k)\big\|_e : k \in \mathbb{N}\big\} \\ &\le \rho_2 \sup \|x_k\|_e = \rho_2\rho' =: \rho,\end{aligned}$$

where ρ_2 is defined as in equation (3.2). Let $C := \big\{v_1, \ldots, v_n\big\}$ be an orthonormal basis of the Euclidean space V. Proposition 2.9, Definition 2.10 and the fact that $\nabla^2_B F(x_k)^{-1/2}$ is symmetric for all $k \in \mathbb{N}$ imply that there exist operators $\Omega_k \in \mathcal{O}(V)$, $\Delta_k \in \mathrm{Lin}\big(E^\sharp, V\big)$ and $\Xi_k \in \mathcal{O}(E^\sharp)$ such that $X(x_k, s_k) = \Omega_k \circ \Delta_k \circ \Xi_k$ for all k and so that the matrix representation of Δ_k with respect to $(B^\sharp, C)$ is given by $\mathrm{Diag}\big(\sigma\big(\nabla^2_B F(x_k)^{-1/2}\big)\big)$, where the eigenvalues are ordered in descending order. Let $\mathcal{G}$ be the set of vector space homomorphisms $E^\sharp \to V$ whose matrix representation with respect to the matrices $(B^\sharp, C)$, is given by $\mathrm{Diag}(z)$ for some $z \in \mathbb{R}^n_+ \cap \bar{B}_\rho(0)$. Since $\mathbb{R}^n_+ \cap \bar{B}_\rho(0)$ is a compact set and the correspondence between matrices and operators a homoeomorphism, it follows that $\mathcal{G}$ is a compact set of operators. On the other hand, it is a well known fact (see e.g. [10]) that $\mathcal{O}(V)$ and $\mathcal{O}(E^\sharp)$ are compact groups. Hence, Tychonov's theorem implies that the sequence of operators $\big(X(x_k, s_k)\big)_\mathbb{N}$ has an accumulation point X and we may assume without loss of generality that $\lim X(x_k, s_k) = X$ (convergence in operator norm). Since $s \in \partial K^\sharp$ we have $X\big(x_k, s_k\big)s \in \partial K_V$ for all $k \in \mathbb{N}$, and since ∂K_V is a closed set it follows that

$$v = \lim_{k\to\infty} X\big(x_k, s_k\big)s_k = Xs = \lim_{k\to\infty} X\big(x_k, s_k\big)s \in \partial K_V.$$

□

Definition 5.2. Let $F \in \mathcal{C}^3(K^\circ, \mathbb{R})$ be a ν-self-scaled barrier functional for the cone $K \subset E$, and consider the optimization problem (PD) from Lemma 3.4. Let (X, S) be a square-root field for F and let γ be the associated target map. (X, S) is *regular* for (PD) if it is congruent and $\gamma'(z)$ is non-singular for all $z \in \mathcal{F}^\circ(PD)$. Note that regularity of (X, S) is a notion that depends on the affine subspaces $x_0 + L \subset E$ and $s_0 + L^\perp \subset E^\sharp$ underlying the definition of Z.

Theorem 5.3. *Let* (PD) *be a problem instance over the self-scaled cone* K *and let* (X, S) *be a square-root field for the* ν*-self-scaled barrier functional* $F \in \mathcal{C}^3(K^\circ, \mathbb{R})$*. If* (X, S) *is regular for* (PD)*, then the associated target map* $\gamma : \mathcal{F}^\circ(PD) \to K_V^\circ$ *is a diffeomorphism.*

Proof. All we need to show is that γ is bijective, since it follows from the inverse mapping theorem and the assumption that γ' is nonsingular everywhere that γ is locally diffeomorphic everywhere. Let us begin with surjectivity. Since γ is congruent its image $\mathrm{im}(\gamma)$ must lie in K_V°. For $z \in \mathcal{F}^\circ(PD)$ the inverse mapping theorem says that there exist open neighborhoods $z \in O_1 \subset \mathcal{F}^\circ(PD)$ and $\gamma(z) \in O_2 \subset K_V$ such that $\gamma|_{O_1}$ is a homoeomorphism between O_1 and O_2. Hence, $\mathrm{im}(\gamma)$ is an open subset of K°. Now suppose $(v_k)_\mathbb{N} \subset \mathrm{im}(\gamma)$ is a sequence that converges to a boundary point $v \in \partial\,\mathrm{im}(\gamma)$. For each $k \in \mathbb{N}$ we can choose a $z_k \in \gamma^{-1}(v_k)$. Let $k_0 \in \mathbb{N}$ such that $\|v_k\|_2 \leq \|v\|_2 + 1$ for all $k \geq k_0$. By Theorem 3.3 and Lemma 3.4 $z_k \in \delta^{-1}\big([0, \|v\|_2 + 1]\big) \cap \mathcal{F}(PD)$ for all $k \geq k_0$, and since Lemma 3.4 says that this set is compact, (z_k) must have an accumulation point. Without loss of generality we may thus assume that (z_k) converges to some point $z \in \mathcal{F}(PD)$. But then $z \in \partial\mathcal{F}(PD)$, since otherwise $v = \gamma(z) \in K_V^\circ$ would hold true by the arguments above, contrary to our assumption. Hence, the sequence $\big(x(z_k), s(z_k)\big)_\mathbb{N}$ satisfies the conditions of Lemma 5.1, and it follows that $v \in \partial K_V$. This shows that $\partial\,\mathrm{im}(\gamma) \subset \partial K_V$ and that the image of γ is closed in the subspace topology of K° inherited from V. We have shown that $\mathrm{im}(\gamma)$ is both open and closed and hence a connected component of the topological space K_V°. But $\mathrm{im}(\gamma)$ is non-empty and K_V° simply connected, hence

$$\mathrm{im}(\gamma) = K_V^\circ.$$

It remains to prove that γ is injective. Let $U := \{v \in K_V^\circ : \#\gamma^{-1}(v) = 1\}$ be the set of elements in K_V° covered only once by γ. It was established in Proposition 3.5 *iii)* that $\{\mu e_V : \mu > 0\} \subset U$. Suppose $v \in K_V^\circ \setminus U$. By surjectivity of γ it must be true that $\#\gamma^{-1}(v) \geq 2$, and hence we can choose two distinct points $z_1, z_2 \in \gamma^{-1}(v)$. The inverse mapping theorem says that there are open neighborhoods $z_i \in O_Z^{(i)} \subset \mathcal{F}^\circ(PD)$ and $v \in O_V^{(i)} \subset K_V^\circ$ for $i = 1, 2$ such that

$$\gamma|_{O_Z^{(i)}} : O_Z^{(i)} \to O_V^{(i)}$$

are homoeomorphisms for $i = 1, 2$. Without loss of generality we may assume that $O_Z^{(1)}$ and $O_Z^{(2)}$ are disjoint. But this means that $O_V^{(1)} \cap O_V^{(2)} \subset K_V^\circ \setminus U$, whence v lies in the interior of $K_V^\circ \setminus U$. Since v was arbitrary we must have that U is a closed subset of the topological space K_V°. Next we will show that it is also open. Suppose $v \in K_V^\circ \cap \partial U$. Then $v \in U$, since U is closed, and there exists a sequence $(v_k)_\mathbb{N} \subset K_V^\circ \setminus U$ that converges to v. Let $z := \gamma^{-1}(v)$. The inverse mapping theorem says that there exist open neighborhoods $z \in O_Z \subset Z$ and $v \in O_V \subset V$ such that $\gamma|_{O_Z} : O_Z \to O_V$ is a homoeomorphism. Since $\#\gamma^{-1}(v_k) \geq 2$ for all $k \in \mathbb{N}$ there exists a sequence $(z_k)_\mathbb{N} \subset \mathcal{F}^\circ(PD) \setminus O_Z$ such that $\gamma(z_k) = v_k$ for all k. The following limit exists:

$$\lim_{k\to\infty} \delta(z_k) \overset{(3.4)}{=} \lim_{k\to\infty} \big(\gamma(z_k); e_V\big) = \big(v; e_V\big).$$

Let $k_0 \in \mathbb{N}$ be such that $\delta(z_k) \leq \big(v; e_V\big) + 1$ for all $k \geq k_0$. Since the set

$$\delta^{-1}\big([0, \big(v; e_V\big) + 1]\big) \cap \mathcal{F}(PD)$$

is compact by Lemma 3.4, it follows that $(z_k)_{\mathbb{N}}$ has an accumulation point $\hat{z} \in \mathcal{F}(PD)$. Without loss of generality we may assume that $(z_k)_{\mathbb{N}}$ converges to $\hat{z}$. Now, since $z_k \notin O_Z$ for all $k \in \mathbb{N}$ it cannot be that $\hat{z}$ and z are equal. If it were true that $\hat{z} \in \mathcal{F}^\circ(PD)$, then $\gamma(z) = v$ would have to hold by continuity of γ, contradicting the fact that $\#\gamma^{-1}(v) = 1$. Therefore, $\hat{z} \in \partial\mathcal{F}(PD)$, Lemma 5.1 applies to $\big(x(z_k), s(z_k)\big)_{\mathbb{N}}$, and it follows that $v \in \partial K_V$, contradicting the assumption that $v \in K_V^\circ$. Therefore, $\partial U \subseteq \partial K_V$, which shows that U is open in the topological space K_V°. Once again, U is a non-empty connected component of the simply connected topological space K_V°, thus equal to K_V°, and hence all elements of K_V° are covered exactly once by γ. □

Definition 5.4. Let $\gamma : \mathcal{F}^\circ(PD) \to K_V^\circ$ be the target map associated with the square-root field (X, S) for $F \in \mathcal{C}^3(K^\circ, \mathbb{R})$, and suppose (X, S) is regular for the problem instance (PD) over K. Let $v \in K_V^\circ$ and $z = \gamma^{-1}(v)$. Then the the points $\big(x(z), s(z)\big)$ are called the *weighted analytic centers* of v for the target map γ.

Acknowledgements

I would like to thank Mike Todd for having given me the impulse that led to this work, and for the countless hours we spent discussing this material.

References

[1] S. Burer and R.D.C. Monteiro. A General Framework for Establishing Polynomial Convergence of Long-Step Methods for Semidefinite Programming. Working paper, School of ISyE, Georgia Tech, USA, August 1999.

[2] S. Burer, R. Monteiro and Y. Zhang. Solving Semidefinite Programs via Nonlinear Programming. Part I: Transformations and Derivatives. Technical Report TR99-17, Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77005, USA.

[3] S. Burer, R. Monteiro and Y. Zhang. Solving Semidefinite Programs via Nonlinear Programming. Part II: Interior Point Methods for a Subclass of SDPs. Technical Report TR99-23, Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77005.

[4] E. Cartan. Domaines bornés homogènes de l'espace de n variables complexes. *Abh. Math. Sem. Hamburg Univ.* 11:116–162, 1935.

[5] J. E. Dennis and R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations.* SIAM Publications. SIAM, Philadelphia, USA, 1996.

[6] J. Faraut and A. Korányi. *Analysis on Symmetric Cones.* Oxford Mathematical Monograph, Oxford University Press, New York, USA, 1994.

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

[8] R. Hauser. Self-Scaled Barrier Functions: Decomposition and Classification. Numerical Analysis Report DAMTP 1999/NA13, Department of Applied Mathematics and Theoretical Physics, Silver Street, Cambridge, England CB3 9EW.

[9] R. Hauser. Square-Root Fields and the "V-Space" Approach to Primal-Dual Interior-Point Methods for Self-Scaled Conic Programming. Numerical Analysis Report DAMTP 1999/NA14, Department of Applied Mathematics and Theoretical Physics, Silver Street, Cambridge, England CB3 9EW.

[10] J. Hilgert and K. H. Neeb. *Lie Gruppen und Lie Algebren.* Vieweg Verlag, Braunschweig/Wiesbaden, Germany, 1991.

[11] B. Jansen, C. Roos, T. Terlaky and J. P. Vial. Primal-dual target-following algorithms for linear programming. *Annals of Operations Research*, 62:197–232, 1996.

[12] B. Jansen, C. Roos, T. Terlaky and J. P. Vial. Long-step primal-dual target-following algorithms for linear programming. *Zeitschrift für Operations Research - Mathematical Methods of Operations Research*, 44:11–30 (1996).

[13] M. Köcher. Positivitätsbereiche im $\mathbb{R}^n$. *Amer. Jour. of Math.*, 79:575–596, 1957.

[14] M. Kojima, S. Mizuno and A. Yoshise. An $O(\sqrt{n}L)$ iteration potential reduction algorithm for linear complementarity problems. *Mathematical Programming*, 50:331–342, 1991.

[15] M. Kojima, N. Megiddo, T. Noma and A. Yoshise. *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*. Springer Verlag, Berlin, 1991.

[16] S. Mizuno. A new polynomial time method for a linear complementarity problem. *Mathematical Programming*, 56:31–43, 1992.

[17] R.D.C. Monteiro. Primal-dual path following algorithms for semidefinite programming. *SIAM Journal on Optimization*, 7:663–678, 1997.

[18] R.D.C. Monteiro and Y. Zhang. A unified analysis for a class of path-following primal-dual interior-point algorithms for semidefinite programming. *Mathematical Programming*, 81:281–299, 1998.

[19] R.D.C. Monteiro and P. Zanjacomo. General Interior-Point Maps and Existence of Weighted Paths for Nonlinear Semidefinite Complementarity Problems. Working paper, School of ISyE, Georgia Tech, USA, April 1998.

[20] R. D. C. Monteiro and I. Adler. Interior path following primal-dual algorithms: Part I: Linear Programming. *Mathematical Programming*, 44:27–41, 1989.

[21] R.D.C. Monteiro and J.-S. Pang. On two interior point mappings for nonlinear semidefinite complementarity problems. *Mathematics of Operations Research*, 23:39–60, 1998.

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

[23] Yu. E. Nesterov and M. J. Todd. Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations Research*, 22:1–42, 1997.

[24] Yu. E. Nesterov and M. J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM Journal on Optimization*, 8:324–364, 1998.

[25] C. Roos, T. Terlaky and J.-Ph. Vial. *Theory and Algorithms for Linear Optimization*. Wiley-Interscience, John Wiley & Sons Ltd, West Sussex, England, 1997.

[26] O. S. Rothaus. Domains of positivity. *Abh. Math. Sem. Univ. Hamburg*, 24:189–235, 1960.

[27] J. F. Sturm and S. Zhang. Symmetric primal-dual path following algorithms for semidefinite programming. Technical Report 9554/A, Econometric Institute, Erasmus University, Rotterdam, The Netherlands, 1995.

[28] J. F. Sturm and S. Zhang. On weighted centers for semidefinite programming. Technical Report 9636/A, Econometric Institute, Erasmus University, Rotterdam, The Netherlands, 1996.

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

[30] M. J. Todd. On search-directions in interior-point methods for semidefinite programming. Revised Preprint, 1997.

[31] M. J. Todd, K.-C. Toh and R.H. Tütüncü. On the Nesterov-Todd direction in semidefinite programming. *SIAM Journal on Optimization*, 8:769–796, (1998).

[32] L. Tunçel. Primal-dual symmetry and scale invariance of interior-point algorithms for convex optimization. *Mathematics of Operations Research*, 23, 1998.

[33] L. Tunçel. Generalization of primal-dual interior-point methods to convex optimization problems in conic form. Research Report CORR 99-35, Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada, August 1999.

[34] E. B. Vinberg. The theory of convex homogeneous cones. *Trudy Moskov. Mat. Obsc.*, 12:303–358, 1963; Trans. Moscow Math. Soc., 13:340–403, 1963.

[35] Y. Zhang. On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming. *SIAM Journal on Optimization*, 8:365-386, 1998