Square-Root Fields and the "V-Space" Approach to 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 a convenient framework for the analysis of primal-dual interior-point methods is known as the "V-space" approach. Various generalizations of this framework have recently been proposed for semidefinite and self-scaled conic programming. We propose such a generalization that inherits all the properties that made this approach a successful analytical tool in the linear programming case. Contrary to certain other generalizations, the objects at the center of our own approach, so-called *square-root fields*, are endowed with a differential structure that plays a crucial role in the asymptotic analysis of primal-dual algorithms.

Key words. symmetric cones, self-scaled cones, primal-dual methods, interior-point methods, weighted analytic centers, target-following methods, 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

So-called *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. Primal-dual methods are most conveniently studied in a framework that exhibits the symmetry between the two optimization problems jointly solved by the algorithm. In the linear programming literature such a framework known as the *V-space approach* has been developed by Kojima-Mizuno-Yoshise [10], Monteiro-Adler [16], Mizuno [12], Jansen-Roos-Terlaky-Vial [7, 8], Todd [25] and others. Various generalization of V-space have been developed in the context of semidefinite programming by Sturm-Zhang [23, 24], Monteiro-Zanjacomo [15] and Burer-Monteiro [1], in the more general context of self-scaled conic programming by Tunçel [28], and a related technique has been developed by Tunçel [29] for general convex conic programming problems. The notion of *square-root fields* we present in this article is most closely related to Tunçel's work [28] and can in fact be seen as an extension thereof. In a forthcoming paper we will use the theory of square-root fields developed here to generate so-called *target directions*. These search directions are linked to the unifying theory of target-following algorithms from the linear programming literature and to the notion of weighted analytic centers.

Self-scaled conic programming provides a unified framework for linear- semidefinite- and second-order-cone programming and was first investigated by Nesterov and Todd [19, 20]. One of the fundamental notions of this theory is the concept of *symmetric* (or *self-scaled*) cone. 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 cone in E. The group

$$\mathrm{Aut}(K) := \big\{\theta \in \mathrm{Aut}(E) : \theta K = K\big\},$$

of automorphisms of E that map K onto itself is called the *automorphism group* of K. In this article we are going to consider cones whose automorphism group is particularly rich: A cone $K \subset E$ is called *symmetric* (see e.g. [3]) if it is convex, pointed, has non-empty topological interior $K^\circ \neq \emptyset$ and satisfies the following two conditions:

1. K and $K^\sharp$ are isomorphic and
2. for all $x, z \in K^\circ$ there exists an automorphism $\theta \in \mathrm{Aut}(K)$ of K such that $\theta x = z$.

Clearly because of condition $i)$ the dual of K, defined as usual by

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

is a symmetric cone in $E^\sharp$. Symmetric cones play an important role in the theory of Jordan algebras (see e.g. [9, 3, 22, 30]) and have been completely classified by Vinberg [30]. Three particular examples are given by

1. $K = \mathbb{R}^n_+$ the positive orthant,

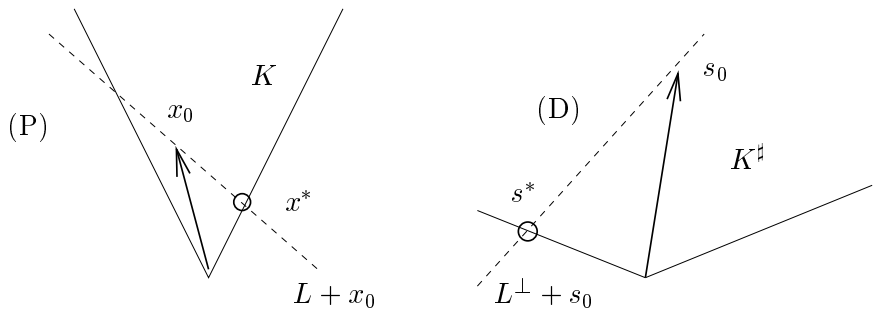

Figure 1: A primal-dual pair of conic optimization problems.

2. $K = \left\{ \binom{\tau}{x} \in \mathbb{R}^{n+1} : \tau \geq \|x\|_2 \right\}$ the Lorentz cone and

3. $K = \mathrm{S}^{n\times n}(\mathbb{R})_+$ the cone of positive semidefinite symmetric matrices over the reals.

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

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

where $K \subset E$ is a symmetric cone (see Figure 1). Note that (D) is actually an equivalent reformulation of the dual problem in its usual form, where we got rid of Lagrange multiplier variables.
Let us mention that under the choice $K = \mathbb{R}^n_+$ in (1.1) one retrieves the general problem of linear programming, the choice K equal to the Lorentz cone yields the general problem of second-order cone programming (which contains the general problem of convex quadratic programming with convex quadratic constraints as a special case), while the choice $K = \mathrm{S}^{n\times n}(\mathbb{R})_+$ yields the general problem of semidefinite programming.

Nesterov and Todd [19, 20] who first studied interior-point algorithms for self-scaled conic programming considered a special set of self-concordant logarithmically homogeneous barrier functionals (see [18]) which they call *self-scaled* barrier functionals (see [19] and Definition 2.2. The Hessians of these barrier functionals are closely connected to the automorphism group $\mathrm{Aut}(K)$ of their (symmetric conic) domain of definition (see Theorem 2.6). As usual, we endow the interior of the dual cone $K^\sharp$ with the *dual barrier functional*

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

This is the Legendre-Fenchel transform of F with the sign of its argument inverted. If F is self-scaled then $F_\sharp$ is a self-scaled barrier functional too (see [19]). 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 a primal-dual pair of self-scaled conic programming problems 1.1, and let the symmetric cone K be endowed with a self-scaled barrier functional $F \in \mathcal{C}^3\big(K^\circ, \mathbb{R}\big)$. 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 2.6). Let us assume that we are given a primal-dual strictly feasible point (x, s), i.e. $x \in K^\circ \cap \big(L + x_0\big)$ and $s \in K^{\sharp\circ} \cap \big(L^\perp + s_0\big)$. In order to solve the perturbed problems $\min\big\{\langle x; s_0\rangle + \mu F(x) : x \in K \cap \big(L + x_0\big)\big\}$ and $\min\big\{\langle x_0; s\rangle + \mu F_\sharp(s) : s \in K^\sharp \cap \big(L^\perp + s_0\big)\big\}$ for some positive real number $\mu > 0$, i.e. in order to solve the central path equations

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

Nesterov and Todd [19] 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 $(x^*, s^*) \in \partial K \times \partial K^\sharp$.

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 it is possible to get rid of both problems by multiplying equation (1.3a) by the operator $X = \mathrm{Diag}(x)$ on the left, and we end up with the equivalent equation

$$Sd_x + Xd_x = \mu e_V - \gamma(x,s), \tag{1.4}$$

where $S := \mathrm{Diag}(s)$, $e_V := (1,\ldots,1)^T$ and where $\gamma(x,s) := \big(x_1 s_1,\ldots,x_n s_n\big)^T$ is called the *target map*. Equation (1.4) is the Newton system corresponding to the non-linear equation

$$\gamma(x,s) = \mu e_V, \qquad \text{with } (x,s) \in \mathcal{F}^\circ(PD), \tag{1.5}$$

which is the central path equation (1.2) in disguised form. One may interpret X and S as linear operators $X \in \mathrm{Aut}(E^\sharp, V)$ and $S \in \mathrm{Aut}(E, V^\sharp)$ whose image space V is a n-dimensional Euclidean space lying "half-way" between E and $E^\sharp$. V is a space where the primal and the dual problems meet and can be treated in a truly joint fashion.
The purpose of this paper is to describe generalizations of the operator pairs (X,S) and of the target map γ in the framework of self-scaled conic programming. We choose X a square-root of the operator $\big(F''(x)\big)^{-1}$ and S a square-root of the operator $\big(F_\sharp''(s)\big)^{-1}$ which are linked up of some sorts. Tunçel [28] studied almost exactly the same operators, but he defines the pair (X,S) for each admissible pair (x,s) separately. In this article we study the case where the dependence of the operator pairs (X,S) on the defining point (x,s) is sufficiently smooth, i.e. we study differentiable operator fields

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

and their differentiable structures. We call these operator fields *square-root fields*.
The material of this article is organized as follows: In Section 2 we give a brief summary of prerequisite results that will be used throughout the text. In Section 3 we present the basic definitions, establish the existence of square-root fields (see Theorem 3.16) and show how intricate examples can be constructed from simple ones. The results of Section 5 illustrate why many of the difficulties associated with the theory of square-root fields collapse in the linear programming case, and that the simplicity of the linear programming case cannot be preserved in the general case.

In a forthcoming paper we will show how square-root fields can be used to define generalized target maps and that in the generalization of equation (1.4) to self-scaled conic programming none of the terms blows up in a neighborhood of the optimal solution. Generalized target maps define an infinite new family of Newton direction to solve problem (1.2).

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 2.11. 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 $v \in \mathbb{R}^n$. Identity maps defined on any set will always be denoted by the symbol $\imath$.

2 Prerequisite results, examples 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 [18] or [5], Definition 2.1 from which our notation is adopted.

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

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 [19, 20]. 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.

Example 2.3 (Positive orthant).
We endow the spaces $E = E^\sharp = V := \mathbb{R}^n$ with their canonical basis and scalar product and choose $K = K^\sharp := \mathbb{R}^n_+$ and

$$F : x \mapsto -\sum_{i=1}^{n} \ln(x_i) \qquad \text{for all } x \in \mathbb{R}^n_{++}.$$

Then F is n-self-scaled and the following identities hold true:

$$\begin{aligned} &F_\sharp : s \mapsto -\sum_{i=1}^{n} \ln(s_i) - n \qquad \text{for all } s \in \mathbb{R}^n_{++}, \\ &F'(x) = -\big(x_1^{-1}, \ldots, x_n^{-1}\big)^T \qquad \text{and} \\ &F''(x) = \operatorname{Diag}(x)^{-2}. \end{aligned}$$

Example 2.4 (Symmetric positive semidefinite cone).
We endow $E = E^\sharp = V := \mathrm{S}^{n\times n}(\mathbb{R})$ with the trace inner product

$$\begin{aligned} &\big\langle x; s\big\rangle = x \bullet s := \operatorname{tr}(xs) \qquad \text{for all } x, s \in \mathrm{S}^{n\times n}(\mathbb{R}), \\ &\big(v_1; v_2\big) = v_1 \bullet v_2 \qquad \text{for all } v_1, v_2 \in \mathrm{S}^{n\times n}(\mathbb{R}). \end{aligned}$$

On $\mathrm{S}^{n\times n}(\mathbb{R})$ let us choose the canonical basis and consider the cone $K = K^\sharp := \mathrm{S}^{n\times n}(\mathbb{R})_+ \subset E$ and the functional

$$F : x \mapsto -\ln(\det(x)) \qquad \text{for all } x \in \mathrm{S}^{n\times n}(\mathbb{R})_{++}.$$

Then F is n-self-scaled and the following identities hold true:

$$\begin{aligned} &F_\sharp : s \mapsto -\ln\big(\det(s)\big) - n \qquad \forall\, s \in \mathrm{S}^{n\times n}(\mathbb{R})_{++}, \\ &F'(x) = -x^{-1} \qquad \text{and} \\ &F''(x)h = x^{-1}hx^{-1} \qquad \forall\, h \in \mathrm{S}^{n\times n}(\mathbb{R}). \end{aligned}$$

Example 2.5 (Lorentz cone).
We endow $E = E^\sharp = V := \mathbb{R}^{n+1}$ with its canonical basis and scalar product, and define

$$\begin{aligned} &K := \{\left(\begin{smallmatrix}\tau\\ x\end{smallmatrix}\right) \in E : \tau \geq \|x\|\}, \\ &F : \left(\begin{smallmatrix}\tau\\ x\end{smallmatrix}\right) \mapsto -\ln\Big(\left(\begin{smallmatrix}\tau\\ x\end{smallmatrix}\right)^T \left(\begin{smallmatrix}\tau\\ -x\end{smallmatrix}\right)\Big) \qquad \text{for all } \left(\begin{smallmatrix}\tau\\ x\end{smallmatrix}\right) \in K^\circ. \end{aligned}$$

Then F is 2-self-scaled and the following equalities hold true:

$$\begin{aligned} &K^\sharp = \{\left(\begin{smallmatrix}\rho\\ s\end{smallmatrix}\right) \in E^\sharp : \rho \geq \|s\|\}, \\ &F_\sharp : \left(\begin{smallmatrix}\rho\\ s\end{smallmatrix}\right) \mapsto -\ln\Big(\left(\begin{smallmatrix}\rho\\ s\end{smallmatrix}\right)^T \left(\begin{smallmatrix}\rho\\ -s\end{smallmatrix}\right)\Big) - 2 + 2\ln 2 \qquad \text{for all } \left(\begin{smallmatrix}\rho\\ s\end{smallmatrix}\right) \in K^{\sharp\circ}, \\ &F'\Big(\left(\begin{smallmatrix}\tau\\ x\end{smallmatrix}\right)\Big) = -\frac{2}{\left(\begin{smallmatrix}\tau\\ x\end{smallmatrix}\right)^T \left(\begin{smallmatrix}\tau\\ -x\end{smallmatrix}\right)} \begin{pmatrix}\tau\\ -x\end{pmatrix}, \\ &F''\Big(\left(\begin{smallmatrix}\tau\\ x\end{smallmatrix}\right)\Big) = \frac{2}{\left(\begin{smallmatrix}\tau\\ x\end{smallmatrix}\right)^T \left(\begin{smallmatrix}\tau\\ -x\end{smallmatrix}\right)} \begin{pmatrix}-1 & 0\\ 0 & I_n\end{pmatrix} + \left[\frac{2}{\left(\begin{smallmatrix}\tau\\ x\end{smallmatrix}\right)^T \left(\begin{smallmatrix}\tau\\ -x\end{smallmatrix}\right)}\right]^2 \begin{pmatrix}\tau\\ -x\end{pmatrix}\begin{pmatrix}\tau\\ -x\end{pmatrix}^T. \end{aligned}$$

Theorem 2.6 (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 2.7. *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 $x = y$.* □

Example 2.8. This is Example 2.3 continued. The primal scaling point can be computed as follows in this case:

$$w(x,s) = \big(x_1^{1/2} s_1^{-1/2}, \ldots, x_n^{1/2} s_n^{-1/2}\big)^T.$$

Example 2.9. This is Example 2.4 continued. In this case the following formulas hold for the primal scaling point:

$$w(x,s) = x^{1/2}\left(x^{1/2}sx^{1/2}\right)^{-1/2}x^{1/2} = s^{-1/2}\left(s^{1/2}xs^{1/2}\right)^{1/2}s^{-1/2},$$
$$w(e,-F'(x)) = x^{1/2}.$$

Example 2.10. This is Example 2.5 continued. Computing the primal scaling point explicitly in this case involves solving a system of multivariate polynomial equations explicitly. We spare the reader the computational details and indicate the resulting formulas which can be checked directly by first assuming that $\tau^2 - \|x\|^2 = 1$ and $\rho^2 - \|s\|^2 = 1$, and then by rescaling and making use of Proposition 2.1 v) & vi). The scaling point $\binom{\sigma}{w}$ of $\left(\binom{\tau}{x},\binom{\rho}{s}\right)$ and the Hessian of F at $\binom{\sigma}{w}$ are given by the expressions

$$\binom{\sigma}{w} = \frac{\binom{\tau}{x} + \sqrt{\frac{\binom{\tau}{x}^T\binom{\tau}{-x}}{\binom{\rho}{s}^T\binom{\rho}{-s}}}\binom{\rho}{-s}}{\sqrt{\binom{\tau}{x}^T\binom{\rho}{s} + \sqrt{\binom{\tau}{x}^T\binom{\tau}{-x}}\sqrt{\binom{\rho}{s}^T\binom{\rho}{-s}}}} \quad ,$$

$$\begin{aligned} F''\left(\binom{\sigma}{w}\right) =& \sqrt{\frac{\binom{\rho}{s}^T\binom{\rho}{-s}}{\binom{\tau}{x}^T\binom{\tau}{-x}}}\begin{pmatrix}-1 & 0\\ 0 & I_n\end{pmatrix} \\ &+ \frac{\left[\binom{\rho}{s}^T\binom{\rho}{-s}\right]\binom{\tau}{-x}\binom{\tau}{-x}^T}{\binom{\tau}{x}^T\binom{\tau}{-x}\left[\binom{\tau}{x}^T\binom{\rho}{s} + \sqrt{\binom{\tau}{x}^T\binom{\tau}{-x}}\sqrt{\binom{\rho}{s}^T\binom{\rho}{-s}}\right]} \\ &+ \frac{\binom{\rho}{s}\binom{\rho}{s}^T}{\binom{\tau}{x}^T\binom{\rho}{s} + \sqrt{\binom{\tau}{x}^T\binom{\tau}{-x}}\sqrt{\binom{\rho}{s}^T\binom{\rho}{-s}}} \\ &+ \frac{\sqrt{\binom{\rho}{s}^T\binom{\rho}{-s}}\left[\binom{\rho}{s}\binom{\tau}{-x}^T + \binom{\tau}{-x}\binom{\rho}{s}^T\right]}{\sqrt{\binom{\tau}{x}^T\binom{\tau}{-x}}\left[\binom{\tau}{x}^T\binom{\rho}{s} + \sqrt{\binom{\tau}{x}^T\binom{\tau}{-x}}\sqrt{\binom{\rho}{s}^T\binom{\rho}{-s}}\right]} \quad . \end{aligned}$$

Definition 2.11. 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 $\left(\cdot;\cdot\right)_e$ by the rule

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

We say that $(\cdot;\cdot)_e$ 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 $B^\sharp = \{e_1^\sharp, \ldots, e_n^\sharp\}$, 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 $\imath : E \to E$.

Güler[4] and Nesterov-Todd [19] showed that the notion of self-scaled cone is equivalent to the notion of *symmetric cone* (see e.g. [3]) developed in the context of the theory of Jordan algebras. An important theorem [30] of this theory says that symmetric cones can be uniquely decomposed into irreducible symmetric cones of which there exist exactly six types, and any direct sum of symmetric cones is a symmetric cone. For details see e.g. [3] or [5] Theorem 2.22, from which we adopt notation.

Lemma 2.12. *All notation is as in [5], Example 2.6. Let* $K = \mathrm{S}^{n\times n}(\mathbb{R})_+ \subset E = \mathrm{S}^{n\times n}(\mathbb{R})$ *be provided with the standard barrier functional* $F : x \mapsto -\ln(\det(x))$*, and provide* E *with the standard Euclidean structure, i.e. with the Euclidean structure defined by* $e = I$*. Then the following is a surjective group homomorphism:*

$$\begin{aligned} \mathcal{N} &: \mathcal{SO}(n) \to \mathcal{Q}(\mathrm{S}^{n\times n}(\mathbb{R})) \\ \mathcal{N}(o) &: h \mapsto o^T h o \qquad \text{for all } h \in \mathrm{S}^{n\times n}(\mathbb{R}). \end{aligned}$$

Proof. See [5], Lemma 4.6. □

3 Square-root fields

Note that all results of this section have dual analogues which we shall usually not state explicitly, but tacit use will be made of them wherever needed. The first two lemmas are well known, yet we mention them for the sake of completeness.

Lemma 3.1. $\cdot^{-1} : Gl_n(\mathbb{R}) \to Gl_n(\mathbb{R})$ *is a smooth mapping.*

Proof. This follows immediately from the Neumann series expansion of $\cdot^{-1}$ about any $G \in Gl_n(\mathbb{R})$. □

Lemma 3.2. $\cdot^{1/2} : \mathrm{S}^{n\times n}(\mathbb{R})_{++} \to \mathrm{S}^{n\times n}(\mathbb{R})_{++}$ *is a well defined and smooth mapping.*

Proof. It is a standard fact from Lie theory that $\exp : M_{n\times n}(\mathbb{R}) \to Gl_n(\mathbb{R})$ is an analytic mapping, and likewise that $\exp|_{\mathrm{S}^{n\times n}(\mathbb{R})} : \mathrm{S}^{n\times n}(\mathbb{R}) \to \mathrm{S}^{n\times n}(\mathbb{R})_{++}$ is a homoeomorphism of manifolds; see for instance [6], Proposition I.2.4 and Satz I.2.10. Therefore, $\exp|_{\mathrm{S}^{n\times n}(\mathbb{R})} : \mathrm{S}^{n\times n}(\mathbb{R}) \to \mathrm{S}^{n\times n}(\mathbb{R})_{++}$ is a $\mathcal{C}^\infty$-diffeomorphism, and $\cdot^{1/2} \equiv \exp|_{\mathrm{S}^{n\times n}(\mathbb{R})} \circ 1/2 \cdot \exp|^{-1}_{\mathrm{S}^{n\times n}(\mathbb{R})}$ is well defined and smooth. □

Remark 3.3. Some of the definitions in this section 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 2.11 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 $\left\{v_1, \ldots, v_n\right\}$.

Definition 3.4. Suppose $B \in \mathcal{B}$ and $x \in K^\circ$. Let $F_B^{\prime\prime 1/2}(x) \in \mathrm{Lin}(E, V)$ and $F_B^{\prime\prime -1/2}(x) \in \mathrm{Lin}(E^\sharp, V)$ 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 $\nabla^2_B F(x)^{1/2}$ and $\nabla^2_B F(x)^{-1/2}$ respectively (see Definition 2.11 for notation). Let us then define the following sets of linear operators:

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

Remark 3.5. Note that $\left(F_B^{\prime\prime 1/2}(x)\right)^{-1} : V \to E$ and $F_B^{\prime\prime -1/2}(x) : E^\sharp \to V$ 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:

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

The next proposition shows that $F^{\prime\prime 1/2}(x)$ and $F^{\prime\prime -1/2}(x)$ are the sets of square-roots of the operators $F''(x)$ and $F^{\prime\prime -1}(x)$ respectively:

Proposition 3.6. *Suppose $B \in \mathcal{B}$ and $x \in K^\circ$. Then*

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

Proof. Suppose that $B \in \mathcal{B}^{(e)}$ for some $e \in K^\circ$ and $B' \in \mathcal{B}^{(e')}$ for some $e' \in K^\circ$. Let $M \in Gl_n(\mathbb{R})$ be the matrix that describes the change of coordinates between objects represented with respect to B and B'. Then

$$\nabla^2_{B'} F(x) = M^T \nabla^2_B F(x) M.$$

Let $OS = \nabla^2_B F(x)^{1/2} M$ be the Cartan decomposition of the matrix $\nabla^2_B F(x)^{1/2} M$, i.e. O is orthogonal and S symmetric. Then

$$S^2 = SO^T OS = M^T \nabla^2_B F(x) M = \nabla^2_{B'} F(x),$$

and this shows that $S = \nabla^2_{B'} F(x)^{1/2}$. If $\Omega \in \mathcal{O}(V)$ is the operator that has O^T as its matrix representation with respect to $\{v_1, \ldots, v_n\}$, then

$$F'^{1/2}_{B'}(x) = \Omega \circ F'^{1/2}_B(x).$$

Note that the matrix O and hence the operator Ω depend on the particular x. So far we have shown that

$$F'^{1/2}(x) \subset \mathcal{O}(V) \circ F'^{1/2}_B(x).$$

For the reverse inclusion it suffices to show that exhaustion of possibilities for the choice of $M \in Gl_n(\mathbb{R})$ leads to exhaustion of $\mathcal{O}(n)$ by the ensuing orthogonal matrices O. But for any $O \in \mathcal{O}(n)$, if we choose $M = O$ in particular we get a basis that is again in $\mathcal{B}^{(e)}$ and the Cartan decomposition of $\nabla^2_B F(x)^{1/2} M$ is $O\big(O^T \nabla^2_B F(x)^{1/2} O\big)$. Hence, O appears among the orthogonal matrices covered by this procedure. In order to see the second equality, let $L \in \mathrm{Lin}(E, V)$ be such that $L^* \circ L = F'(x)$ and suppose L_B is the matrix representation of L with respect to $(B, \{v_1, \ldots, v_n\})$. Let $OS = L_B$ be the Cartan decomposition of L_B and let $\Omega \in \mathcal{O}(V)$ be the operator that has O as its matrix representation with respect to $\{v_1, \ldots, v_n\}$. Again we can argue that $S = \nabla^2_B F(x)^{1/2}$ and $L = \Omega \circ F'^{1/2}_B(x)$. On the other hand, it was already mentioned in Remark 3.5 that $L^* \circ L = F'(x)$ for all $L \in F'^{1/2}(x)$. The equalities pertaining to $F'^{-1/2}$ can be established in the same fashion. □

Lemma 3.7. *Recall that we denote the scalar product defined on V by $(\cdot;\cdot)$ and the bilinear form defined on $E \times E^\sharp$ by $\langle \cdot;\cdot \rangle$. Let us fix a basis $B \in \mathcal{B}$. Then, for all $x \in K^\circ$ we have*

$$(\cdot;\cdot) = \langle F'^{-1/2}_B(x)^* \cdot\,; F'^{1/2}_B(x)^* \cdot \rangle.$$

Proof. This follows trivially from Remark 3.5. □

Lemma 3.8. *Suppose that the self-scaled barrier functional $F : K \to \mathbb{R}$ is k times continuously differentiable. Then the following mapping is in $\mathcal{C}^{(k-2)}$:*

$$\begin{aligned} K^\circ \times K^{\sharp\circ} &\to E \otimes E \\ (x, s) &\mapsto F'^{-1}(w(x, s)). \end{aligned}$$

Proof. Let $B \in \mathcal{B}$. It follows from Proposition 3.6 and Theorem 2.6 *iii)* that for all $x \in K^\circ$ and $s \in K^{\sharp\circ}$,

$$\left(F_B''^{1/2}(x) \circ F''^{-1}(w) \circ F_B''^{1/2}(x)^*\right)^2 = F_B''^{1/2}(x) \circ F_\sharp''(s) \circ F_B''^{1/2}(x)^*$$

and hence,

$$F''^{-1}(w) = F_B''^{1/2}(x)^{-1} \circ \left(F_B''^{1/2}(x) \circ F_\sharp''(s) \circ F_B''^{1/2}(x)^*\right)^{1/2} \circ F_B''^{-1/2}(x).$$

When represented with respect to the basis $\{v_1, \ldots, v_n\}$ the operator accounting for the center term on the right-hand side is expressed as a symmetric matrix C, and the square-root appearing in the equation should be interpreted as the operator whose matrix representation with respect to the same basis is the unique symmetric square-root of C. All right hand terms in the last equation are of class $\mathcal{C}^{(k-2)}$ in view of Lemmas 3.1 and 3.2. □

Lemma 3.9. *Let $B \in \mathcal{B}$ and suppose the self-scaled barrier functional $F : K \to \mathbb{R}$ is k times continuously differentiable. Then the following mapping is well defined and lies in $\mathcal{C}^{(k-2)}$:*

$$\begin{aligned} K^\circ \times K^{\sharp\circ} &\to \mathcal{O}(V) \\ (x, s) &\mapsto F_B''^{-1/2}(x) \circ F''(w(x,s)) \circ F_{\sharp B}''^{1/2}(s)^*. \end{aligned}$$

Proof. For all $x \in K^\circ$ and $s \in K^{\sharp\circ}$,

$$\begin{aligned} &\left[F_B''^{-1/2}(x) \circ F''(w(x,s)) \circ F_{\sharp B}''^{1/2}(s)^*\right]^* \circ F_B''^{-1/2}(x) \circ F''(w(x,s)) \circ F_{\sharp B}''^{1/2}(s)^* \\ &\qquad \overset{3.5}{=} F_{\sharp B}''^{1/2}(s) \circ F''(w(x,s)) \circ F''^{-1}(x) \circ F''(w(x,s)) \circ F_{\sharp B}''^{1/2}(s)^* \\ &\qquad\qquad \overset{2.6\,iii)}{=} F_{\sharp B}''^{1/2}(s) \circ F_\sharp''^{-1}(s) \circ F_{\sharp B}''^{1/2}(s)^* \overset{3.5}{=} \imath. \end{aligned}$$

This shows that the image of the mapping under consideration lies indeed in $\mathcal{O}(V)$. Moreover, $(k-2)$-fold continuous differentiability follows from Lemmas 3.1, 3.2 and 3.8. □

Corollary 3.10. *The following is a set of self-dual regular cones in V, mutually congruent and isomorphic to K:*

$$\left\{L(K) : L \in F''^{1/2}(x), x \in K^\circ\right\} \cup \left\{L(K^\sharp) : L \in F_\sharp''^{1/2}(s), s \in K^{\sharp\circ}\right\}$$

Proof. Self-duality follows from the fact that $F''(x) : K \to K^\sharp$ is a cone isomorphism and from

$$(\cdot;\cdot) \overset{3.7}{=} \langle F_B''^{1/2}(x)^{-1} \cdot ; F_B''^{1/2}(x)^* \cdot \rangle \overset{3.5}{=} \langle F_B''^{1/2}(x)^{-1} \cdot ; F''(x) \circ F_B''^{1/2}(x)^{-1} \cdot \rangle.$$

Moreover, for any $x \in K^\circ$, $s \in K^{\sharp\circ}$ and $B \in \mathcal{B}$,

$$\begin{aligned} F_B''^{1/2}(x)K &\overset{3.5}{=} F_B''^{-1/2}(x) \circ F''(x)K = F_B''^{-1/2}(x)K^\sharp \\ &= F_B''^{-1/2}(x) \circ F''(w(x,s)) \circ F_\sharp''(s)K^\sharp \\ &= F_B''^{-1/2}(x) \circ F''(w(x,s)) \circ F_{\sharp B}''^{1/2}(s)^* \circ F_{\sharp B}''^{1/2}(s)K^\sharp \\ &\in \mathcal{O}(V) \circ F_{\sharp B}''^{1/2}(s)K^\sharp \qquad \text{by Lemma 3.9.} \end{aligned}$$

The result now follows trivially from Proposition 3.6. □

Definition 3.11. Let $e_V \in V$ be a fixed vector of norm $\nu^{1/2}$. For $x \in K^\circ$ and $s \in K^{\sharp\circ}$ the following sets of square-roots of $F''^{-1}(x)$ and $F_\sharp''^{-1}(s)$ will be relevant for us:

$$\begin{aligned} \mathcal{R}_{(x,e_V)} &:= \big\{X \in F''^{-1/2}(x) : X[-F'(x)] = e_V\big\}, \\ \mathcal{R}_{(s,e_V)} &:= \big\{S \in F_\sharp''^{-1/2}(s) : S[-F_\sharp'(s)] = e_V\big\}. \end{aligned}$$

Example 3.12. Consider the case of linear programming with the log-barrier functional, and set $e_V = (1, \ldots, 1)^T$. Then $X := \mathrm{Diag}(x) \in \mathcal{R}_{(x,e_V)}$ for all $x \in \mathbb{R}^n$. □

Proposition 3.13. *Let $x \in K^\circ$ and let $e_V \in V$ be a fixed vector of norm $\nu^{1/2}$. Let $\mathcal{O}(V, e_V)$ be the group of orthogonal transformations of V that leave e_V invariant. Then $\mathcal{R}_{(x,e_V)}$ is non empty. Moreover, if $X \in \mathcal{R}_{(x,e_V)}$ then*

i) $\mathcal{R}_{(x,e_V)} = \mathcal{O}(V, e_V) \circ X$,

ii) $X^{-*} \in \mathcal{R}_{(-F'(x),e_V)}$, *where* $X^{-*} := \big(X^*\big)^{-1}$,

iii) $\mathcal{R}_{(\tau x,e_V)} = \tau\mathcal{R}_{(x,e_V)}$ *for all* $\tau > 0$.

Proof. Let R be an arbitrary element of $F''^{-1/2}(x)$. Then

$$\|R[-F'(x)]\|^2 \overset{3.7}{=} \big\langle F''^{-1}(x)[-F'(x)]; -F'(x)\big\rangle \overset{2.1i)}{=} \big\langle x; -F'(x)\big\rangle \overset{2.1iv)}{=} \nu \tag{3.1}$$

There exists therefore an orthogonal transformation $\Omega \in \mathcal{O}(V)$ such that $\Omega \circ R[-F'(x)] = e_V$, and we have $\Omega \circ R \in \mathcal{R}_{(x,e_V)} \neq \emptyset$. Now suppose $X, Y \in \mathcal{R}_{(x,e_V)}$. Then it follows from Proposition 3.6 that $YX^{-1} \in \mathcal{O}(V)$, and

$$YX^{-1}e_V = YX^{-1}X[-F'(x)] = Y[-F'(x)] = e_V.$$

On the other hand, it is immediate that $\Omega \circ X \in \mathcal{R}_{(x,e_V)}$ for any $\Omega \in \mathcal{O}(V, e_V)$. This proves part *i)*.

Part *ii*) follows from the following identities:

$$\left(X^{-*}\right)^* \circ X^{-*} = \left(X^* \circ X\right)^{-1} = F''(x) = \left(F''_\sharp(-F'(x))\right)^{-1},$$
$$X^{-*}\left[-F'_\sharp(-F'(x))\right] \overset{2.1}{=} X^{-*}x = X \circ F''(x)x \overset{2.1}{=} X[-F'(x)] = e_V.$$

Finally, part *iii*) is immediate from Proposition 2.1. □

Proposition 3.14. *Let $x \in K^\circ$ and $s \in K^{\sharp\circ}$. Then a canonical isomorphism between $\mathcal{R}_{(x,e_V)}$ and $\mathcal{R}_{(s,e_V)}$ is given by the mapping*

$$\begin{aligned} \mathcal{R}_{(x,e_V)} &\to \mathcal{R}_{(s,e_V)} \\ X &\mapsto X^\sharp := X \circ F''(w(x,s)). \end{aligned}$$

If $X \in \mathcal{R}_{(x,e_V)}$ and $\Omega \in \mathcal{O}(V, e_V)$ then $(\Omega \circ X)^\sharp = \Omega \circ X^\sharp$. Moreover, $X^{\sharp\sharp} = X$.

Proof. Let $X \in \mathcal{R}_{(x,e_V)}$. Then, since $F''(w)$ is self-adjoint, we have

$$\begin{aligned} X^{\sharp *} \circ X^\sharp &= F''(w(x,s))^* \circ X^* \circ X \circ F''(w(x,s)) \\ &= F''(w(x,s)) \circ F''^{-1}(x) \circ F''(w(x,s)) = F''^{-1}_\sharp(s), \\ X^\sharp\left(-F'_\sharp(s)\right) &= X \circ F''(w(x,s))[-F'_\sharp(s)] \overset{2.6i)}{=} X[-F'(x)] = e_V. \end{aligned}$$

These identities and Proposition 3.6 imply that $X^\sharp \in \mathcal{R}_{(s,e_V)}$ and the isomorphism is well defined. The identity $(\Omega \circ X)^\sharp = \Omega \circ X^\sharp$ is trivially true. Finally, the identity $X^{\sharp\sharp} = X$ follows from the fact that $F''_\sharp(t(x,s)) = F''^{-1}(w(x,s))$; see Proposition 2.1 *iii*) and Theorem 2.6 *iv*). □

Definition 3.15 (Square-root fields).
Let $K \subset E$ be a symmetric cone with ν-self-scaled barrier functional $F \in \mathcal{C}^k\left(K^\circ, \mathbb{R}\right)$, 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

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

such that the following conditions are satisfied for all $x \in K^\circ$ and $s \in K^{\sharp\circ}$:

i) $X(x,s) \in \mathcal{R}_{(x,e_V)}$ and

ii) $S(x,s) = X^\sharp(x,s)$.

This is equivalent to requiring that Diagram 3.2 commutes for all (x,s):

(3.2a)

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

$$
\begin{array}{ccc}
-F'_\sharp(s) & \xrightarrow{F''(w(x,s))} & -F'(x) \\
\Big\downarrow{\scriptstyle S(x,s)} & & \Big\downarrow{\scriptstyle X(x,s)} \\
e_V & = & e_V
\end{array}
\tag{3.2b}
$$

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 $(x,s) \mapsto F_\sharp''^{-1}(s)$ respectively. Let $B \in \mathcal{B}$ (see Definition 2.11). Then it follows from Proposition 3.6 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

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

These identities combined with Corollary 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 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 3.16 (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 $\big(F, e_V\big)$ 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 $\big(F, e_V, K_V\big)$ 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 $T \equiv \Omega \circ S$.

Proof. Consider an arbitrary element e of K° and an operator $R \in F''^{1/2}(e)$. By virtue of equation (3.1) we may assume without loss of generality that $e_V = Re$. Let us consider the operator fields

$$
\begin{aligned}
X : K^\circ \times K^{\sharp\circ} &\to \mathrm{Lin}(E^\sharp, V), \\
(x,s) &\mapsto R \circ F''^{-1}\big(w(e, -F'(x))\big)
\end{aligned}
$$

and $S := X^\sharp$. Then Lemma 3.8 implies that X and S are in $\mathcal{C}^{k-2}$. Moreover, it follows from Remark 3.5 and the fact that operators of the form $F''(w)$ are self-adjoint that the following identities apply for all $(x,s) \in K^\circ \times K^{\sharp\circ}$:

$$\begin{aligned} X(x,s)^* \circ X(x,s) &= F''^{-1}\big(w(e,-F'(x))\big) \circ F''(e) \circ F''^{-1}\big(w(e,-F'(x))\big) \\ &\overset{2.6iii)}{=} F''_\sharp(-F'(x)) \overset{2.1iii)}{=} F''^{-1}(x), \\ X(x,s)[-F'(x)] &= R \circ F''^{-1}\big(w(e,-F'(x))\big)[-F'(x)] = R[e] = e_V, \\ X(x,s)K^\sharp &\equiv R \circ F''^{-1}\big(w(e,-F'(x))\big)K^\sharp \equiv R[K] =: K_V. \end{aligned}$$

This shows that (X,S) is a congruent square-root field for (F, e_V, K_V). The rest follows easily from Proposition 3.13 *i*). □

On direct sums of symmetric cones square-root fields can be constructed by taking direct sums of square-root fields for their constituent cones:

Theorem 3.17 (Direct sums of square-root fields).
Consider a finite set of 8-tuples

$$\big\{\big(E_i, K_i, F_i, \nu_i, V_i, e_{V_i}, X_i, S_i\big) : i \in I\big\}$$

such that for each $i \in I$ $K_i \subset E_i$ is a symmetric cone, $F_i \in \mathcal{C}^k\big(K_i^\circ, \mathbb{R}\big)$ is a ν_i-self-scaled barrier functional for K_i, V_i is a Euclidean space of equal dimension as E_i, $e_{V_i} \in V_i$ is a vector of length $\|e_{V_i}\|_2 = \nu_i^{1/2}$ and $\big(X_i, S_i\big)$ is a square-root field for $\big(F_i, e_{V_i}\big)$. Let us define $E := \bigoplus_I E_i$, $K := \bigoplus_I K_i$, $F := \bigoplus_I F_i$, $\nu := \sum_I \nu_i$, $V := \bigoplus_I V_i$, $e_V := \bigoplus_I e_{V_i}$, $X := \bigoplus_I X_i$ and $S := \bigoplus_I S_i$. Then F is a ν-self-scaled barrier functional for K, $e_V \in V$ has length $\|e_V\|_2 = \nu^{1/2}$ and (X,S) is a square-root field for $\big(F, e_V\big)$. If the $\big(X_i, S_i\big)$ are congruent then (X,S) is congruent too.

Proof. It is straightforward to check that F is a ν-self-scaled barrier functional for K and that $e_V \in V$ has length $\|e_V\|_2 = \nu^{1/2}$. In order to check that (X,S) is a square-root field for $\big(F, e_V\big)$, it suffices to keep in mind that Diagram 3.2 is the direct sum of the diagrams corresponding to the square-root fields $\big(X_i, S_i\big)$. Also, it is immediate to check that if $S_i K_i \equiv K_{V_i}$ for all $i \in I$ then $SK \equiv \bigoplus_I K_{V_i}$. □

4 Examples

In order to keep things simple we will introduce fixed coordinate systems and confound linear operators and their matrix representations in the examples that follow.

Example 4.1 (Linear programming). Let us construct a square-root field for the self-scaled cone and barrier functional from Example 2.8: We may choose $e = e_V := (1, \ldots, 1)^T$

and $R := I \in F''^{-1/2}(e)$. Consider the square-root field (X, S) that was constructed in the proof of Theorem 3.16:

$$X : (x, s) \mapsto R \circ F''^{-1}\big(w(e, -F'(x))\big) = F''^{-1}(\sqrt{x}) = \mathrm{Diag}(x),$$
$$S = X^\sharp : (x, s) \mapsto X \circ F''(w(x, s)) = \mathrm{Diag}(x)\,\mathrm{Diag}(x)^{-1}\,\mathrm{Diag}(s) = \mathrm{Diag}(s).$$

Note that this is the canonical square-root field for linear programming that was already mentioned in the introduction. Since this is the example that motivated the definition of square-root fields at the outset we will henceforth refer to it as the *primordial square-root field*. From Theorem 3.16 and the fact that $\mathcal{O}(V, \mathbb{R}^n_+)$ is the group of coordinate permutations it follows that for a fixed K_V there are only $n!$ distinct congruent square-root fields for (F, e_V). The "theory" of square-root fields is trivial in this case due to the polyhedral nature of the cone. □

Example 4.2 (Semidefinite programming I). Let us construct a square-root field for the self-scaled cone and barrier functional from Example 2.9: Proceeding as in the proof of Theorem 3.16 we may choose $e := I$, $R = \imath \in F''^{1/2}(e)$, $e_V := R[e] = I$ and construct a square-root field as follows: For all $h \in \mathrm{S}^{n\times n}(\mathbb{R})$,

$$X(x, s)h = R \circ F''^{-1}\big(w(e, -F'(x))\big)h = x^{1/2}hx^{1/2},$$
$$S(x, s)h = X^\sharp(x, s)h = X \circ F''(w(x, s))h = \big(x^{1/2}sx^{1/2}\big)^{1/2}x^{-1/2}hx^{-1/2}\big(x^{1/2}sx^{1/2}\big)^{1/2}.$$

Note that (X, S) is not primal-dual symmetric, a defect that can be fixed as the next example will show. □

Example 4.3 (Semidefinite programming II). This example was found by Mike Todd. We will make use of Theorem 3.16 and Lemma 2.12 to compose the previous example with an orthogonal field

$$\Omega \in \mathcal{C}^{k-2}\big(K^\circ \times K^{\sharp\circ}, \mathcal{O}(\mathrm{S}^{n\times n}(\mathbb{R})_+, I)\big)$$

so as to retrieve a primal-dual symmetric square-root field for F as a result. For $x, s \in \mathrm{S}^{n\times n}(\mathbb{R})_{++}$ let us define $v(x, s) := w(x, s)^{-1/2}xw(x, s)^{-1/2}$. This concept is originally due to Kojima-Megiddo-Noma-Yoshise [11] for LP and to Sturm-Zhang [24] for SDP; see also Todd [26]. The map $v : \mathrm{S}^{n\times n}(\mathbb{R})^2_{++} \to V$ is primal-dual symmetric. In fact, since the dual scaling point satisfies the equation $t(x, s) = w(x, s)^{-1}$ we have

$$v(x, s) = w^{-1/2}xw^{-1/2} = t^{-1/2}w^{-1}xw^{-1}t^{-1/2} = t^{-1/2}st^{-1/2} = v(s, x).$$

Let $n(x, s) := v^{-1/2}w^{-1/2}x^{1/2}$. Then $n(x, s)n(x, s)^T = v^{-1/2}w^{-1/2}xw^{-1/2}v^{-1/2} = I$, i.e. $n(x, s) \in \mathcal{O}(n)$. Let (X, S) denote the square-root field found in Example 4.2, define

$$\Omega := \mathcal{N} \circ n : \mathrm{S}^{n\times n}(\mathbb{R})^2_{++} \to \mathcal{O}(\mathrm{S}^{n\times n}(\mathbb{R})_+, I)$$

and consider the the tuple $(\Omega \circ X, \Omega \circ S)$ which according to Theorem 3.16 is a congruent square-root field for (F, I). We contend that it is primal-dual symmetric. In fact, for all $h \in \mathrm{S}^{n\times n}(\mathbb{R})$,

$$
\begin{aligned}
\big(\Omega \circ X\big)(x,s)h &= v^{-1/2}w^{-1/2}xhxw^{-1/2}v^{-1/2},\\
\big(\Omega \circ S\big)(x,s)h &= \\
&= v^{-1/2}w^{-1/2}x^{1/2}\big(x^{1/2}sx^{1/2}\big)^{1/2}x^{-1/2}hx^{-1/2}\big(x^{1/2}sx^{1/2}\big)^{1/2}x^{1/2}w^{-1/2}v^{-1/2}\\
&= v^{-1/2}w^{-1/2}xw^{-1}hw^{-1}xw^{-1/2}v^{-1/2}\\
&= v^{-1/2}t^{-1/2}w^{-1}xw^{-1}hw^{-1}xw^{-1}t^{-1/2}v^{-1/2}\\
&= v^{-1/2}t^{-1/2}shst^{-1/2}v^{-1/2}.
\end{aligned}
$$

□

Example 4.4 (Second-Order Cone Programming). Let us construct a square-root field for the self-scaled cone and barrier functional from Example 2.5: Let $e = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$. Then $F''(e) = 2I_{n+1}$, and proceeding as in Theorem 3.16 we may choose $R = \sqrt{2}I_{n+1} \in F''^{1/2}(e)$ and $e_V = R[e] = \left(\begin{smallmatrix}\sqrt{2}\\0\end{smallmatrix}\right)$. Next we need to compute the inverse of the Hessian of F at the primal scaling point of e and $-F'\left(\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right)\right)$ for each $\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right) \in K^\circ$. But thanks to Theorem 2.6 *iv)* and Proposition 2.1 *iii)* this amounts to computing the Hessian of $F_\sharp$ at the dual scaling point $\left(\begin{smallmatrix}\zeta\\t\end{smallmatrix}\right)$ of e and $-F'\left(\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right)\right)$. Using the formulas derived in Example 2.10 we get

$$
\begin{aligned}
\zeta &= \sqrt{2}\,\sqrt{\frac{\tau + \sqrt{\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right)^T\left(\begin{smallmatrix}\tau\\-x\end{smallmatrix}\right)}}{\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right)^T\left(\begin{smallmatrix}\tau\\-x\end{smallmatrix}\right)}},\\
t &= -\sqrt{\frac{2}{x^Tx}}\,\sqrt{\frac{\tau - \sqrt{\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right)^T\left(\begin{smallmatrix}\tau\\-x\end{smallmatrix}\right)}}{\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right)^T\left(\begin{smallmatrix}\tau\\-x\end{smallmatrix}\right)}}\;x.
\end{aligned}
$$

Finally, defining $X : K^\circ \times K^{\sharp\circ} \to \mathrm{Lin}(E^\sharp, V)$ as in the proof of Theorem 3.16 we get

$$
\begin{aligned}
X : \Big(\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right), \left(\begin{smallmatrix}\rho\\s\end{smallmatrix}\right)\Big) &\mapsto R \circ F''_\sharp\Big(\left(\begin{smallmatrix}\zeta\\t\end{smallmatrix}\right)\Big)\\
&= \sqrt{\frac{\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right)^T\left(\begin{smallmatrix}\tau\\-x\end{smallmatrix}\right)}{2}}\begin{pmatrix}-1 & 0\\ 0 & I_n\end{pmatrix} + \frac{1}{\sqrt{2}}\begin{pmatrix}\tau + \sqrt{\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right)^T\left(\begin{smallmatrix}\tau\\-x\end{smallmatrix}\right)} & x^T\\ x & \dfrac{\tau - \sqrt{\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right)^T\left(\begin{smallmatrix}\tau\\-x\end{smallmatrix}\right)}}{x^Tx}\,xx^T\end{pmatrix}.
\end{aligned}
$$

It is easily checked that

$$
\begin{aligned}
X\Big(\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right), \left(\begin{smallmatrix}\rho\\s\end{smallmatrix}\right)\Big)\Big[-F'\Big(\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right)\Big)\Big] &\equiv \left(\begin{smallmatrix}\sqrt{2}\\0\end{smallmatrix}\right), \qquad \text{and}\\
X\Big(\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right), \left(\begin{smallmatrix}\rho\\s\end{smallmatrix}\right)\Big)^* \circ X\Big(\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right), \left(\begin{smallmatrix}\rho\\s\end{smallmatrix}\right)\Big) &\equiv \frac{\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right)^T\left(\begin{smallmatrix}\tau\\-x\end{smallmatrix}\right)}{2}\begin{pmatrix}-1 & 0\\ 0 & I_n\end{pmatrix} + \begin{pmatrix}\tau\\x\end{pmatrix}\begin{pmatrix}\tau\\x\end{pmatrix}^T\\
&= F''_\sharp\Big(-F'\Big(\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right)\Big)\Big) = F''^{-1}\Big(\left(\begin{smallmatrix}\tau\\x\end{smallmatrix}\right)\Big)
\end{aligned}
$$

as required. The operator field $S : K^\circ \times K^{\sharp\circ} \to \mathrm{Lin}(E, V)$ could now be computed in the usual fashion:

$$S : \left(\begin{pmatrix} \tau \\ x \end{pmatrix}, \begin{pmatrix} \rho \\ s \end{pmatrix} \right) \mapsto X\left(\begin{pmatrix} \tau \\ x \end{pmatrix}, \begin{pmatrix} \rho \\ s \end{pmatrix} \right) \circ F''\left(\begin{pmatrix} \sigma \\ w \end{pmatrix} \right),$$

but explicit calculation of this expression reveals once again that the square-root field (X, S) is not primal-dual symmetric. In a forthcoming paper we will show how this flaw can be fixed. □.

5 Univariate square-root fields

In this section we will see why many of the difficulties associated with the theory of square-root fields collapse in the linear programming case, and why there is no hope that such a collapse can take place in the general case.

Definition 5.1. Let K, F be a ν-self-scaled cone and barrier functional, and suppose (X, S) is a square-root field for F such that $X(x, s) \equiv X(x)$ and $S(x, s) \equiv S(s)$. Then (X, S) is called a *univariate* square-root field.

The canonical square-root field for linear programming

$$X : (x, s) \mapsto \mathrm{Diag}(x) \qquad S : (x, s) \mapsto \mathrm{Diag}(s) \qquad \text{for all } x, s \in \mathbb{R}^n_{++}$$

is univariate, and this leads to a substantial simplification of the theory in this case, so much that one can safely say that if univariate square-root fields existed in the general case there would be little interest in developing a theory for the non-univariate case. Unfortunately, univariate square-root fields do not exist in general.

Suppose $F \in \mathcal{C}^3(K^\circ, \mathbb{R})$ is ν-self-scaled and that there exists a univariate square-root field (X, S) for (F, e_V). Let us choose a basis $B \in \mathcal{B}$ of E, and let $\Omega_x := \Omega_{(x,B)}$, $\Omega_s := \Omega_{(x,B)}$ (see Definition 3.15). Then it must be true that $\Omega_x(x, s) \equiv \Omega_x(x)$ and $\Omega_s(x, s) \equiv \Omega_s(s)$. Moreover,

$$\Omega_s(s) \circ F''^{-1/2}_{\sharp B}(s) \equiv S(x, s) \equiv X^\sharp(x, s) \equiv \Omega_x(x) \circ F''^{-1/2}_B(x) \circ F''(w(x, s)),$$
$$\text{and hence,} \qquad \frac{\partial}{\partial x}\left[\Omega_x(x) \circ F''^{-1/2}_B(x) \circ F''(w(x, s))\right] \equiv 0.$$

Therefore, we get that for all $d \in E$

$$\frac{\partial}{\partial x} F''(w(x, s))[d] \circ F''^{-1}(w(x, s)) = \tag{5.1}$$
$$= -F''^{1/2}_B(x)^* \circ \Omega_x(x)^T \circ \frac{\partial}{\partial x}\left[\Omega_x(x) \circ F''^{-1/2}_B(x)\right][d].$$

Since the right hand side of (5.1) does not depend on the slack variable s, the same must hold true for the left hand side, an expression that is *independent* of the particular square-root field chosen at the outset. But this yields a conceptually simple sufficient criterion for

the non-existence of univariate square-root fields:

Lemma 5.2. *Let $K \subset E$ be a self-scaled cone and $F \in \mathcal{C}^3(K^\circ, \mathbb{R})$ a self-scaled barrier functional. If the mapping*

$$\begin{aligned} \varsigma : K^\circ \times K^{\sharp\circ} \times E &\to \mathrm{End}(E^\sharp), \\ (x, s, d) &\mapsto \frac{\partial}{\partial x} F''(w(x,s))[d] \circ F''^{-1}(w(x,s)). \end{aligned}$$

functionally depends on the variable s, then no univariate square-root field exists for F. □

For linear programming we do know of the existence of a univariate square-root field. A cross check should thus reveal that ς *is* independent of s in this case. Indeed,

$$\varsigma(x,s,d) = \frac{\partial}{\partial x} \operatorname{Diag}\big(s_i/x_i\big)[d] \operatorname{Diag}\big(x_i/s_i\big) = -\operatorname{Diag}\big(d_i/x_i\big).$$

We now contend that for semidefinite and second-order cone programming with the canonical barrier functionals ς is *not* independent of s and hence by Lemma 5.2 there exists no univariate square-root field in these cases. In later sections we are going to restrict the domain of definition of the square-root fields under consideration to the set of strictly feasible points $\mathcal{F}^\circ(P) \times \mathcal{F}^\circ(D)$ of a particular dual pair of problem instances (1.1). One might then hope that restricting the domain of definition of ς to $\mathcal{F}^\circ(P) \times \mathcal{F}^\circ(D)$ renders it independent of s. However, our proof will show that in general this situation does not occur either.

Theorem 5.3. *In the case of semidefinite programming there is no univariate square-root field for the canonical barrier functional. In general this is even true if the range is restricted to the primal-dual feasible domain.*

Proof. Let us begin by choosing a specific dual pair of problem instances (1.1): Let $\mathbb{O}$ be the $(n-2) \times (n-2)$-zero matrix, $\mathbb{I} := I_{n-2}$ and

$$x_0 := \begin{pmatrix} 1 & 0 & \\ 0 & 1 & \\ & & \mathbb{I} \end{pmatrix}, \quad s_0 := \begin{pmatrix} 2 & 0 & \\ 0 & 2 & \\ & & \mathbb{I} \end{pmatrix}, \quad d^{[x]} := \begin{pmatrix} 0 & 1 & \\ 1 & 1 & \\ & & \mathbb{O} \end{pmatrix}, \quad d^{[s]} := \begin{pmatrix} 1 & 0 & \\ 0 & 0 & \\ & & \mathbb{O} \end{pmatrix}, \quad s_1 := \begin{pmatrix} 8 & 0 & \\ 0 & 2 & \\ & & \mathbb{I} \end{pmatrix}$$

$$\text{and} \quad h := \begin{pmatrix} -1 & 0 & \\ 0 & 1 & \\ & & \mathbb{O} \end{pmatrix} \in E = \mathrm{S}^{n\times n}(\mathbb{R}).$$

Let $L := \operatorname{span}\{d^{[x]}\}$. Then $d^{[s]} \in L^\perp$ and $s_1 = s_0 + 6d^{[s]}$ is strictly feasible for (D'), hence (x_0, s_0), $(x_0, s_1) \in \mathcal{F}^\circ(P) \times \mathcal{F}^\circ(D)$. We are going to show that

$$\varsigma(x_0, s_0, d^{[x]})h = -\begin{pmatrix} 0 & 0 & \\ 0 & 1 & \\ & & \mathbb{O} \end{pmatrix} \neq -\begin{pmatrix} 0 & \frac{1}{3} & \\ \frac{1}{3} & 1 & \\ & & \mathbb{O} \end{pmatrix} = \varsigma(x_0, s_1, d^{[x]})h.$$

Let us start by showing how ς can be computed via Lyapunov equations. Recall that $w(x,s) = s^{-1/2}\big(s^{1/2}xs^{1/2}\big)^{1/2}s^{-1/2}$ and $F''(w)h = w^{-1}hw^{-1}$. Now, considering the expression $\big(s^{1/2}(x+d)s^{1/2}\big)^{1/2}$ as a function of $d \in E$ and developing this function into its Taylor series around $d = 0$ we get

$$\big(s^{1/2}(x+d)s^{1/2}\big)^{1/2} = \big(s^{1/2}xs^{1/2}\big)^{1/2} + d^{(1)} + O(\|d\|^2),$$

where $d^{(1)} = O(\|d\|)$ is the first order term. Taking squares on both sides of this equation and collecting first order terms we get the Lyapunov equation

$$\big(s^{1/2}xs^{1/2}\big)^{1/2}d^{(1)} + d^{(1)}\big(s^{1/2}xs^{1/2}\big)^{1/2} = s^{1/2}ds^{1/2}.$$

Let $o \in \mathcal{O}(n)$ be an orthogonal matrix such that $os^{1/2}xs^{1/2}o^T = \mathrm{Diag}(\lambda)^2$ with $\lambda \in \mathbb{R}^N_+$. Then $d^{(1)} = o^T d^{(2)} o$, where

$$\mathrm{Diag}(\lambda)d^{(2)} + d^{(2)}\,\mathrm{Diag}(\lambda) = os^{1/2}ds^{1/2}o^T =: \Delta, \qquad d^{(2)}_{ij} = \frac{\Delta_{ij}}{\lambda_i + \lambda_j}. \tag{5.2}$$

Hence,

$$\begin{aligned} w(x+d,s) &= w(x,s) + s^{-1/2}o^T d^{(2)} o s^{-1/2} + O(\|d\|^2),\\ w^{-1}(x+d,s) &= w^{-1}(x,s) - w^{-1}s^{-1/2}o^T d^{(2)} o s^{-1/2} w^{-1}(x,s) + O(\|d\|^2)\\ &= w^{-1} - s^{1/2}o^T\,\mathrm{Diag}(\lambda)^{-1}d^{(2)}\,\mathrm{Diag}(\lambda)^{-1}os^{1/2} + O(\|d\|^2)\\ &=: w^{-1} + d^{(3)} + O(\|d\|^2). \end{aligned}$$

Therefore, for any $h \in E = \mathrm{S}^{n\times n}(\mathbb{R})$,

$$\begin{aligned} F''(w(x+d,s)) \circ F'^{-1}(w(x,s))h &= w^{-1}(x+d,s)w(x,s)hw(x,s)w^{-1}(x+d,s)\\ &= h + d^{(3)}w(x,s)h + hw(x,s)d^{(3)} + O(\|d\|^2), \end{aligned}$$

and it follows that

$$\begin{aligned} \varsigma(x,s,d)h = \frac{\partial}{\partial x}F''(w(x,s))[d] \circ F'^{-1}(w(x,s))h &= \\ = d^{(3)}w(x,s)h + hw(x,s)d^{(3)} &\\ = -\big[s^{1/2}o^T\,\mathrm{Diag}(\lambda)^{-1}d^{(2)}os^{-1/2}h + hs^{-1/2}o^T d^{(2)}\,\mathrm{Diag}(\lambda)^{-1}os^{1/2}\big]. & \end{aligned} \tag{5.3}$$

Let us now apply this formula to compute $\varsigma(x_0, s_0, d^{[x]})h$. Since

$$s_0^{1/2}x_0 s_0^{1/2} = \begin{pmatrix} \sqrt{2} & 0 & & \\ 0 & \sqrt{2} & & \\ & & \mathbb{I} & \end{pmatrix}^2,$$

expressions (5.2) and (5.3) simplify:

$$d^{(2)}_{ij} = \frac{\Delta_{ij}}{\lambda_i + \lambda_j} = \frac{\lambda_i\lambda_j}{\lambda_i + \lambda_j}d^{[x]}_{ij}, \qquad d^{(2)} = \frac{1}{\sqrt{2}}d^{[x]},$$

$$\begin{aligned} \varsigma(x_0,s_0,d^{[x]})h = -\big[d^{(2)}s^{-1/2}h + hs^{-1/2}d^{(2)}\big] &= \\ = \frac{-1}{\sqrt{2}\sqrt{2}}\left[\begin{pmatrix} 0 & 1 & \\ 1 & 1 & \\ & & \mathbb{O} \end{pmatrix}\begin{pmatrix} -1 & 0 & \\ 0 & 1 & \\ & & \mathbb{O} \end{pmatrix} + \begin{pmatrix} -1 & 0 & \\ 0 & 1 & \\ & & \mathbb{O} \end{pmatrix}\begin{pmatrix} 0 & 1 & \\ 1 & 1 & \\ & & \mathbb{O} \end{pmatrix}\right] &= -\begin{pmatrix} 0 & 0 & \\ 0 & 1 & \\ & & \mathbb{O} \end{pmatrix}. \end{aligned}$$

In order to compute $\varsigma(x_0, s_1, d^{[x]})h$ we once again observe that

$$s_1^{1/2} x_0 s_1^{1/2} = \begin{pmatrix} 2\sqrt{2} & 0 & \\ 0 & \sqrt{2} & \\ & & \mathbb{O} \end{pmatrix}^2 \quad \text{is diagonal, hence} \quad d^{(2)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & \frac{4}{3} & \\ \frac{4}{3} & 1 & \\ & & \mathbb{O} \end{pmatrix} \quad \text{and}$$

$$\begin{aligned} \varsigma(x_0, s_1, d^{[x]})h &= -\left[d^{(2)} s^{-1/2} h + h s^{-1/2} d^{(2)}\right] \\ &= \frac{-1}{\sqrt{2}\sqrt{2}} \left[\begin{pmatrix} 0 & \frac{4}{3} & \\ \frac{4}{3} & 1 & \\ & & \mathbb{O} \end{pmatrix} \begin{pmatrix} -\frac{1}{2} & 0 & \\ 0 & 1 & \\ & & \mathbb{O} \end{pmatrix} + \begin{pmatrix} -\frac{1}{2} & 0 & \\ 0 & 1 & \\ & & \mathbb{O} \end{pmatrix} \begin{pmatrix} 0 & \frac{4}{3} & \\ \frac{4}{3} & 1 & \\ & & \mathbb{O} \end{pmatrix} \right] = - \begin{pmatrix} 0 & \frac{1}{3} & \\ \frac{1}{3} & 1 & \\ & & \mathbb{O} \end{pmatrix}. \end{aligned}$$

The rest follows from Lemma 5.2. □

Theorem 5.4. *In the case of second-order cone programming there is no univariate square-root field for the canonical barrier functional. In general this is even true if range is restricted to the primal-dual feasible domain.*

Proof. Using the formulas from Example 2.10, we can compute ς explicitly with elementary but very tedious algebra. After a lengthy calculation one finds

$$\begin{aligned} \varsigma\left(\begin{pmatrix} \tau \\ x \end{pmatrix}, \begin{pmatrix} \rho \\ s \end{pmatrix}, \begin{pmatrix} \delta \\ d \end{pmatrix}\right) =& \frac{\partial}{\partial \begin{pmatrix} \tau \\ x \end{pmatrix}} F''\left(\begin{pmatrix} \sigma \\ w \end{pmatrix}\right)\left[\begin{pmatrix} \delta \\ d \end{pmatrix}\right] \circ F''^{-1}\left(\begin{pmatrix} \sigma \\ w \end{pmatrix}\right) \\ =& -\frac{\begin{pmatrix} \tau \\ x \end{pmatrix}^T \begin{pmatrix} \delta \\ -d \end{pmatrix}}{\begin{pmatrix} \tau \\ x \end{pmatrix}^T \begin{pmatrix} \tau \\ -x \end{pmatrix}} I_{n+1} \\ &+ \frac{\begin{pmatrix} \tau \\ x \end{pmatrix}^T \begin{pmatrix} \delta \\ -d \end{pmatrix} \left[\begin{pmatrix} \rho \\ s \end{pmatrix} \begin{pmatrix} \tau \\ x \end{pmatrix}^T - \begin{pmatrix} \tau \\ -x \end{pmatrix} \begin{pmatrix} \rho \\ -s \end{pmatrix}^T \right]}{\begin{pmatrix} \tau \\ x \end{pmatrix}^T \begin{pmatrix} \tau \\ -x \end{pmatrix} \left[\begin{pmatrix} \tau \\ x \end{pmatrix}^T \begin{pmatrix} \rho \\ s \end{pmatrix} + \sqrt{\begin{pmatrix} \tau \\ x \end{pmatrix}^T \begin{pmatrix} \tau \\ -x \end{pmatrix}} \sqrt{\begin{pmatrix} \rho \\ s \end{pmatrix}^T \begin{pmatrix} \rho \\ -s \end{pmatrix}} \right]} \\ &+ \frac{\left[\begin{pmatrix} \delta \\ -d \end{pmatrix} \begin{pmatrix} \rho \\ -s \end{pmatrix}^T - \begin{pmatrix} \rho \\ s \end{pmatrix} \begin{pmatrix} \delta \\ d \end{pmatrix}^T \right]}{\left[\begin{pmatrix} \tau \\ x \end{pmatrix}^T \begin{pmatrix} \rho \\ s \end{pmatrix} + \sqrt{\begin{pmatrix} \tau \\ x \end{pmatrix}^T \begin{pmatrix} \tau \\ -x \end{pmatrix}} \sqrt{\begin{pmatrix} \rho \\ s \end{pmatrix}^T \begin{pmatrix} \rho \\ -s \end{pmatrix}} \right]} \\ &+ \frac{\sqrt{\begin{pmatrix} \rho \\ s \end{pmatrix}^T \begin{pmatrix} \rho \\ -s \end{pmatrix}} \left[\begin{pmatrix} \delta \\ -d \end{pmatrix} \begin{pmatrix} \tau \\ x \end{pmatrix}^T - \begin{pmatrix} \tau \\ -x \end{pmatrix} \begin{pmatrix} \delta \\ d \end{pmatrix}^T \right]}{\sqrt{\begin{pmatrix} \tau \\ x \end{pmatrix}^T \begin{pmatrix} \tau \\ -x \end{pmatrix}} \left[\begin{pmatrix} \tau \\ x \end{pmatrix}^T \begin{pmatrix} \rho \\ s \end{pmatrix} + \sqrt{\begin{pmatrix} \tau \\ x \end{pmatrix}^T \begin{pmatrix} \tau \\ -x \end{pmatrix}} \sqrt{\begin{pmatrix} \rho \\ s \end{pmatrix}^T \begin{pmatrix} \rho \\ -s \end{pmatrix}} \right]}. \end{aligned}$$

It is easy to construct specific data $\begin{pmatrix} \tau_0 \\ x_0 \end{pmatrix}, \begin{pmatrix} \rho_0 \\ s_0 \end{pmatrix}, \begin{pmatrix} \rho_1 \\ s_1 \end{pmatrix}, \begin{pmatrix} \delta \\ d \end{pmatrix}$ such that $\left(\begin{pmatrix} \tau_0 \\ x_0 \end{pmatrix}, \begin{pmatrix} \rho_0 \\ s_0 \end{pmatrix}\right)$ and $\left(\begin{pmatrix} \tau_0 \\ x_0 \end{pmatrix}, \begin{pmatrix} \rho_1 \\ s_1 \end{pmatrix}\right)$ are strictly primal and dual feasible yet $\varsigma\left(\begin{pmatrix} \tau_0 \\ x_0 \end{pmatrix}, \begin{pmatrix} \rho_0 \\ s_0 \end{pmatrix}, \begin{pmatrix} \delta \\ d \end{pmatrix}\right) \neq \varsigma\left(\begin{pmatrix} \tau_0 \\ x_0 \end{pmatrix}, \begin{pmatrix} \rho_1 \\ s_1 \end{pmatrix}, \begin{pmatrix} \delta \\ d \end{pmatrix}\right)$. The result now follows from Lemma 5.2. □

Acknowledgements

Many thanks go to Mike Todd for having given me the impulse that led to this work, and for the many fruitful conversations we had about this paper.

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

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

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

[5] 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.

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

[7] 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.

[8] 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).

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

[10] 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.

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

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

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

[14] 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.

[15] 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.

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

[17] 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.

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

[19] 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.

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

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

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

[23] 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.

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

[25] 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.

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

[27] 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).

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

[29] 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.

[30] 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.

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