SELF-SCALED BARRIERS FOR SEMIDEFINITE PROGRAMMING

Raphael A. Hauser*

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

March 31, 2000

Abstract

We show a result that can be expressed in any of the following three equivalent ways: 1. All self-scaled barrier functionals for the cone Σ_+ of symmetric positive semidefinite matrices are homothetic transformation of the universal barrier functional. 2. All self-scaled barrier functionals for Σ_+ can be expressed in the form $X \mapsto -c_1 \ln \det X + c_0$ for some constants $c_1 > 0, c_0 \in \mathbb{R}$. 3. All self-scaled barrier functionals for Σ_+ are isotropic. As a consequence we find that a self-concordant barrier functional H for Σ_+ is self-scaled if and only if $\mathrm{Aut}(\Sigma_+)$ acts as a group of translations on H, and that the closed subgroup of $\mathrm{Aut}(\Sigma_+)$ generated by the set of Hessians of a self-scaled barrier H coincides with the orientation preserving part of $\mathrm{Aut}(\Sigma_+)$.

Key words. semidefinite programming, self-scaled barrier functionals, symmetric cones, interior-point methods.
AMS 1991 subject classification. Primary: 90C25, 52A41, 90C60. Secondary: 90C05, 90C20.

*Research supported in part by the Norwegian Research Council through project No. 127582/410 "Synode II", through the Engineering and Physical Sciences Research Council of the UK under grant No. GR/M30975, and by NSERC of Canada (grants of J. Borwein and P. Borwein).

1 Introduction

1.1 Towards a Full Classification of Self-Scaled Barriers

Semidefinite programs are a special case of self-scaled conic programming problems, a unified class of convex optimization problems that was first introduced by Nesterov and Todd [17, 18]. Interior-point methods applied to this problem class (see e.g. [17, 18, 1, 4] are based on a class of self-concordant barrier functionals [16] that satisfy additional symmetry properties. These so-called *self-scaled* barriers have been investigated by Güler [7]. The characteristic function

$$\varphi_K(x) := \int_{K^\sharp} \exp\{-\langle x; s\rangle\} ds$$

of a regular cone K is defined as the Laplace transform of the uniform measure over the polar cone $K^\sharp := \{s \in E^\sharp : \langle x; s\rangle \geq 0 \quad \forall x \in K\}$. Güler showed that the universal barrier functional U [16] for K is related to φ_K as follows: There exist real numbers $c_1 > 0$ and c_0 such that $U = c_1 \ln \varphi_K + c_0$. Moreover, barrier functionals of this form are self-scaled if K is symmetric [7]. Together with a result by Nesterov and Todd [17] this implies that symmetric cones provide exactly the class of convex cones on which it is possible to define a self-scaled barrier functional.

A complete classification of symmetric cones was developed by Vinberg [25] who showed that any symmetric cone K can be uniquely written as a direct sum $K = \bigoplus_{j\in J} K_j$ of irreducible components, i.e. symmetric cones that cannot be further reduced. There exist exactly six different types of irreducible symmetric cones, one of which is the cone of symmetric positive semidefinite matrices over the reals, henceforth denoted by Σ_+. For further details see e.g. [25, 5, 10].

In [10] we showed that any self-scaled barrier functional F on a symmetric cone K can be decomposed into a direct sum

$$F = \bigoplus_{j\in J} F^{(j)} \tag{1.1}$$

of self-scaled barrier functionals $F^{(j)}$ with domains of definition the interiors of the irreducible components K_j of K. The irreducible decomposition (1.1) of F is unique up to additive constants. We further introduced the notion of *isotropic* self-scaled barrier functionals, defined as being invariant under the action of the *orthogonal group* of K:

$$F \circ \theta = F \qquad \text{for all } \theta \in \mathcal{O}(K) := \{\theta \in \mathcal{O}(E) : \theta K = K\}.$$

In [10] we found that all isotropic self-scaled barrier functionals for irreducible symmetric cones are of the form $c_1 \ln \varphi_K + c_0$, $c_1 > 0, c_0 \in \mathbb{R}$, and hence they are homothetic transformations of the universal barrier [16, 7]. This result allowed us to derive a theorem that says that any isotropic self-scaled barrier functional F is of the form

$$F = \bigoplus_J c_j \ln \varphi_{K_j} + c_0 \tag{1.2}$$

for some constants $c_j > 0 (j \in J), c_0 \in \mathbb{R}$, where $c_i = c_j$ if $K_i \cong K_i$.

In [10] we conjectured that all self-scaled barrier functionals defined on irreducible symmetric cones are isotropic. If this conjecture can be answered in the affirmative, then self-scaled barriers can be completely classified, since all such functionals must then be of the form (1.2) for some constants $c_j > 0 (j \in J), c_0 \in \mathbb{R}$. The purpose of the present article is to prove this conjecture in the case of the symmetric positive semidefinite cone.

As a consequence, we get that all self-scaled barrier functionals for use in semidefinite programming are of the form $X \mapsto -c_1 \ln \det X + c_0$, and that a self-concordant barrier functional defined on the positive semidefinite symmetric cone Σ_+ is self-scaled if and only if the automorphism group of Σ_+ acts as a group of translations under composition with the barrier, i.e. if and only if the barrier is invariant in the sense of Güler and Tunçel [23] under the transitive group $\mathrm{Aut}(\Sigma_+)$.

1.2 Self-Scaled Cones and Barriers

For the purposes of the present paragraph let K be a regular cone in a finite dimensional real vector space E. The dual space, i.e. the space of of linear forms over E will be denoted by $E^\sharp$, and the dual (or polar) of K in $E^\sharp$ is defined as usual as

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

Here $\langle \cdot ; \cdot \rangle$ denotes the canonical bilinear product defined on $E \times E^\sharp$. The following definition and theorem from [17] are among the basic building blocks of the theory of self-scaled conic programming:

Definition 1.1 (Nesterov-Todd). A ν-self-concordant logarithmically homogeneous [16] barrier functional $F \in \mathcal{C}^3(K^\circ, \mathbb{R})$ is said to be *self-scaled* if the following conditions are satisfied:

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

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

It can be shown that if F is self-scaled then its dual

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

is a self-scaled barrier functional as well. If K allows for a self-scaled barrier functional then K is called self-scaled (or symmetric).

Theorem 1.2 (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 $F''_\sharp(t(x,s))s = x$. □

Since the Hessians of the barrier functional F introduced in Definition 1.1 are vector-space isomorphisms that map K onto $K^\sharp$, it is customary to fix a particular element $e \in K^\circ$ and to identify E with $E^\sharp$ and K with $K^\sharp$ via $F''(e)$. Under this identification the bilinear form $\langle \cdot;\cdot \rangle$ becomes a scalar product $(\cdot;\cdot)_e$, which depends on the choice of e, and the Hessians $F''(x)$ become cone automorphisms of K.

1.3 A Few Remarks Regarding Notation

For the rest of this article, the only self-scaled cone we are concerned with is the cone Σ_+ of symmetric positive semidefinite $n \times n$ matrices with real-valued coefficients. Let us endow $M_{n\times n}(\mathbb{R})$ with the trace inner-product $A \bullet B := \operatorname{tr}(AB^T)$. The Euclidean norm corresponding to this inner-product is simply the Frobenius norm. Consider the subspace $E := \Sigma \subset M_{n\times n}(\mathbb{R})$ of symmetric matrices, and endow it with the inner-product defined by the restriction of $\bullet$ to Σ. Thus, contrary to the usual approach to the general case of self-scaled programming, E is already endowed with a natural inner product, and it is easy to check that the cone $K := \Sigma_+ \subset \Sigma$ of symmetric positive semidefinite matrices is self-dual under the canonical isomorphism $E \to E^\sharp$ defined by this inner product, i.e. $K = K^\sharp$. The topological interior of Σ_+, i.e. the set of positive definite symmetric $n \times n$ matrices shall be denoted by Σ_{++}. Whenever we speak of "barrier functionals for Σ_+" it should be understood that these functionals are actually defined on Σ_{++}.

The canonical barrier functional for the positive semidefinite cone is defined by $F(X) = -\ln\det(X)$ for all $X \in \Sigma_{++}$. F is n-self-scaled (see [17]), and the following identities hold:

$$
\begin{aligned}
&F_\sharp : S \mapsto -\ln\big(\det(S)\big) - n \qquad \forall S \in \Sigma_{++},\\
&F'(X) = -X^{-1} \qquad \forall X \in \Sigma_{++},\\
&F''(X)A = X^{-1}AX^{-1} \qquad \forall A \in \Sigma \qquad \text{and}\\
&W(X,S) = S^{-1/2}\big(S^{1/2}XS^{1/2}\big)^{1/2}S^{-1/2}.
\end{aligned}
$$

In the last equation $W(X,S)$ denotes the primal scaling point of $X \in K^\circ$ and $S \in K^{\sharp\circ}$.

2 Representation Theorems for the Group of Automorphisms of the Positive Semidefinite Cone

This section is largely historical and is concerned with the problem of giving a precise characterisation of the set of automorphisms of the linear space Σ that preserve the property of positive semidefiniteness. The set of transformations that have this property forms a closed subgroup of $\mathrm{Aut}(\Sigma)$ and is called the group of automorphisms of the positive semidefinite cone. This group, to which we shall henceforth refer as $\mathrm{Aut}(\Sigma_+)$, can be characterised as the set of conjugations defined by invertible $n \times n$ matrices (see Corollary 2.8 below). We shall summarise the literature pertaining to this result in a way that prepares the path for Section 3.

As Horn and Johnson [12] point out, the problem of characterising the linear transformations of certain matrix spaces that preserve a given feature of a matrix such as the determinant, eigenvalues, rank, positive definiteness, similarity class, congruence class, idempotence etc. has a long history that finds its roots in the work of Frobenius [6] who showed that all eigenvalue-preserving linear transformations on $M_{n\times n}(\mathbb{C})$ are either of the form $X \mapsto SXS^{-1}$ or $X \mapsto SX^TS^{-1}$ for some non-singular $S \in M_{n\times n}(\mathbb{C})$. Frobenius' work was extended by Schur [22], and subsequently an extended literature on linear preservers ensued (see e.g. [15, 19]).

Symmetric cones and their automorphism groups have been investigated in the literature on Jordan algebras. The connected component P of the identity of these automorphism groups is well understood through the concept of so-called stabilisers (see e.g. [13, 2, 21, 25, 5]). Applied to the special case of the positive semidefinite cone, these results imply that all the elements of the connected component of the identity in $\mathrm{Aut}(\Sigma_+)$ are conjugation mappings. However, this result does not describe the discrete group $\mathrm{Aut}(\Sigma_+)/P$ whose elements are the connected components of $\mathrm{Aut}(\Sigma_+)$, and it does not give a characterisation of these components.

We denote the subgroup of $\mathrm{Aut}(\Sigma_+)$ consisting of orthogonal transformations of $(\Sigma, \bullet)$ by $\mathcal{O}(\Sigma_+)$. It is easy to see that the elements of this group act in an eigenvalue-preserving manner on all of Σ (see Lemma 2.4 below). This suggests that a moderate strengthening of Frobenius' result would suffice to show that $\mathcal{O}(\Sigma_+)$ consists of conjugations. Note that the eigenvalue preserving property implies in particular that the elements of $\mathcal{O}(\Sigma_+)$ are rank preservers.

The remarkable work of Lim [14] and Waterhouse [26] completely characterised the set of linear transformations of symmetric matrices over commutative rings that preserve the property of a matrix to be of rank one. In the case where the ring is a field $\mathbb{K}$, all such transformations are conjugations with a non-singular matrix. Hence, all elements of $\mathcal{O}(\Sigma_+)$ are conjugations with a matrix $S \in Gl_n(\mathbb{R})$, and it is an easy task to show that S must be orthogonal. Rothaus' theorem on the polar decomposition (see Theorem 2.7 below) then shows that all elements of $\mathrm{Aut}\,\Sigma_+$ are conjugations. Osman Güler [8] combined these results in a private communication to to Levent Tunçel and obtained the characterisation given in Corollary 2.8 (see also [24]). The automorphism group of the cone of positive semidefinite Hermitian matrices over $\mathbb{C}$ can also be characterised in the same fashion. In fact, the characterisation of rank one preservers over the space of $n \times n$ complex Hermitian matrices in terms of conjugations is a classical result of the literature of linear preservers (see e.g.[12], Theorem 4.5.7).

Let us now develop a few of the aforementioned results in further detail and introduce the notation that we shall use in Section 3.

We begin by a brief review of the Kronecker products of $n \times n$ matrices. Let us consider the canonical vector space isomorphism $\mathbf{vec} : M_{n\times n}(\mathbb{C}) \to \mathbb{C}^{n^2}$ defined by column stacking. Let $A, B \in M_{n\times n}(\mathbb{C})$ and consider the endomorphism on $M_{n\times n}(\mathbb{C})$ defined by

$$X \mapsto BXA^T. \tag{2.1}$$

The $n^2 \times n^2$ matrix corresponding to this endomorphism under the **vec** transformation is called the *Kronecker product* of A and B and denoted by $A \otimes B$, i.e. we have

$$\mathbf{vec}\big(BXA^T\big) = A \otimes B\ \mathbf{vec}(X) \qquad \forall X \in M_{n\times n}(\mathbb{C}).$$

It is well known that the ij-th $n \times n$ block of the matrix $A \otimes B$ is given by $a_{ij}B$ (see for example [3]). Let us now list a few well known properties of Kronecker products:

Lemma 2.1. *Let $A, B, C, D \in M_{n\times n}(\mathbb{C})$ and denote the spectrum of A and B respectively, counting multiplicities, by $\sigma(A)$ and $\sigma(B)$ respectively. Then the following holds true:*

i) $(A \otimes B)^* = A^* \otimes B^*$,

ii) $(A \otimes B)(C \otimes D) = AC \otimes BD$,

iii) $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ *if* $A, B \in Gl_n(\mathbb{C})$,

iv) $A \otimes B$ *is unitary (resp. orthogonal) if* A *and* B *are unitary (resp. orthogonal),*

v) if $\sigma(A) = \{\lambda_1, \dots, \lambda_n\}$ *and* $\sigma(B) = \{\mu_1, \dots, \mu_n\}$ *then*

$$\sigma(A \otimes B) = \{\lambda_i \mu_j : i, j \in \mathbb{N}_n\}.$$

□

Our special interest is in Kronecker products of real-valued matrices corresponding to linear endomorphisms of $M_{n\times n}(\mathbb{R})$ that leave the subspace Σ of symmetric matrices invariant. Let us therefore state a few facts which may easily be verified by the reader. We shall denote both the endomorphism defined in (2.1) and the corresponding matrix by $A \otimes B$, and we shall use the notation

$$E_{ij}^{+} := e_i e_j^T + e_j e_i^T - \delta_{ij} e_i e_j^T \qquad \text{and} \tag{2.2}$$

$$E_{ij}^{-} = e_i e_j^T - e_j e_i^T, \tag{2.3}$$

where e_i denotes the i-th coordinate vector and δ_{ij} denotes the usual Kronecker symbol. Finally we shall write Σ^i for the set of $n \times n$ skew-symmetric matrices over the reals.

Lemma 2.2. *Given the notation we have just introduced, the following holds true:*

i) Σ^i *is the orthogonal complement of* Σ *in* $\big(M_{n\times n}(\mathbb{R}), \bullet\big)$.

ii) Both Σ *and* Σ^i *are invariant under* $A \otimes A$ *for any* $A \in M_{n\times n}(\mathbb{R})$.

iii) Σ_+ *is invariant under* $A \otimes A$ *for any* $A \in Gl_n(\mathbb{R})$.

iv) Suppose A *is non-defective and*

$$PAP^{-1} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Then $PE_{ij}^{+}P^T \in \Sigma$ *is an eigenvector of* $A \otimes A$ *corresponding to the eigenvalue* $\lambda_i \lambda_j$ *for all* $i \leq j$. *Likewise,* $PE_{ji}^{-}P^T \in \Sigma^i$ *is an eigenvector of* $A \otimes A$ *corresponding to the eigenvalue* $\lambda_i \lambda_j$ *for all* $i > j$. *This completely describes the eigenstructure of* $A \otimes A$, $A \otimes A|_{\Sigma}$ *and* $A \otimes A|_{\Sigma^i}$.

$\square$

Together with Lemma 2.1 *ii*), *iii*) the third part of Lemma 2.2 shows that the conjugation mapping

$$\begin{aligned} \mathcal{C} : Gl_n(\mathbb{R}) &\to \mathrm{Aut}(\Sigma_+) \\ A &\mapsto A \otimes A \end{aligned}$$

is a group homomorphism. It is easy to see that its kernel is given by $\{\pm I\}$.

Consider the Euclidean space $\big(\Sigma, \bullet\big)$ and the group $\mathcal{O}(\Sigma)$ of its orthogonal automorphisms. The closed subgroup

$$\mathcal{O}(\Sigma_+) := \big\{\theta \in \mathcal{O}(\Sigma) : \theta\Sigma_+ = \Sigma_+\big\}$$

of orthogonal transformations of Σ that leave the positive semidefinite cone invariant is called the *orthogonal group of* Σ_+. We denote the connected component of the identity in $\mathcal{O}(\Sigma_+)$ by Q. It is straightforward to see that

$$\mathcal{C}|_{\mathcal{O}(n)} : \mathcal{O}(n) \to \mathcal{O}(\Sigma_+)$$

is a group homomorphism. It is also known that

$$\mathcal{C}|_{\mathcal{SO}(n)} : \mathcal{SO}(n) \to Q$$

is surjective (see e.g. [5]). Indeed, this result also follows from Corollary 2.6 below whose proof is independent of the present considerations. Finally, since

$$\ker(\mathcal{C}) = \ker(\mathcal{C}|_{\mathcal{O}(n)}) = \{\pm I\},$$

we conclude that $\mathcal{C}$ covers its image two-fold. Let us summarise these facts:

Remark 2.3.

i) If n is even, then $-I \in \mathcal{SO}(n)$ and hence $\mathcal{SO}(n)$ wraps twice around Q under $\mathcal{C}$. Let J be the identity matrix I with the sign of the first entry reversed. Then Lemma 2.2 iv) shows that $\mathcal{C}(J)$ has an odd number of negative eigenvalues, and hence it must be orientation reversing. Therefore, $\mathcal{O}(n) \setminus \mathcal{SO}(n)$ maps onto the connected component $\mathcal{C}(J)Q$ of $\mathcal{O}(\Sigma_+) \setminus \mathcal{SO}(\Sigma_+)$ under $\mathcal{C}$ and covers it twice (see Figure 1).

ii) If n is odd, then $-I \in \mathcal{O}(n) \setminus \mathcal{SO}(n)$ and hence both connected components of $\mathcal{O}(n)$ map onto Q under $\mathcal{C}$, each covering the image once (see Figure 2). Note that Lemma 2.2 *iv*) confirms that in this case all images of $\mathcal{C}$ have an even number of negative eigenvalues.

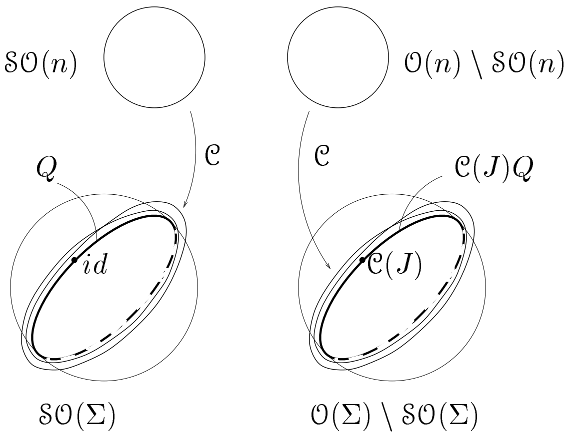

Figure 1: The action of $\mathcal{C}|_{\mathcal{O}(n)}$ when n is even.

It is known that Q is the connected component of the identity in $\mathcal{O}(\Sigma_+)$ (see e.g. [5]). Again, this fact independently follows from Corollary 2.6 below. If Z is a connected component of $\mathcal{O}(\Sigma_+)$ and $\theta \in \mathcal{O}(\Sigma_+)$ then θZ and $Z\theta$ are connected components too. In particular, $\theta Q \theta^{-1}$ is the connected component of the identity and hence $\theta Q \theta^{-1} = Q$ for all $\theta \in \mathcal{O}(\Sigma_+)$. This shows that Q is a normal subgroup of $\mathcal{SO}(\Sigma_+) := \mathcal{O}(\Sigma_+) \cap \mathcal{SO}(\Sigma)$. Since $\mathcal{O}(\Sigma_+)$ is compact, the quotient $\mathcal{SO}(\Sigma_+)/Q$ is a finite group. A priori, the structure described in Remark 2.3 may not be the whole story about $\mathcal{O}(\Sigma_+)$. In fact, the discrete groups $\mathcal{SO}(\Sigma_+)/Q$ may be non-trivial, and the same applies to $\mathcal{O}(\Sigma_+)/\mathcal{SO}(\Sigma_+)$ in the case where n is odd (see Figure 3).

Let $X \in \Sigma_{++}$. The *stabiliser* of X in $\mathrm{Aut}(\Sigma_+)$ is defined as

$$\mathrm{Aut}(\Sigma_+)_X := \{\theta \in \mathrm{Aut}(\Sigma_+) : \theta[X] = X\}.$$

It is known (see e.g. [5] Proposition I.1.8) that the set of stabilisers in $\mathrm{Aut}(\Sigma_+)$ coincides with the set of maximal compact subgroups of $\mathrm{Aut}(\Sigma_+)$. In particular, $\mathcal{O}(\Sigma_+) = \mathrm{Aut}(\Sigma_+)_I$ is the stablizer of the identity matrix I (see e.g. [5] Proposition I.1.9).

Proposition 2.4. *The elements of $\mathcal{O}(\Sigma_+)$ are eigenvalue preservers on Σ.*

Proof. Let $\theta \in \mathcal{O}(\Sigma_+)$. Then $\theta \in \mathrm{Aut}(\Sigma_+)_I$ and hence $\theta[I] = I$. Moreover, $\theta[X] \bullet \theta[Y] = X \bullet Y$ for all $X, Y \in \Sigma$. In particular,

$$\mathrm{tr}\big(\theta[X]^2\big) = \mathrm{tr}\big(X^2\big) \qquad \text{and} \tag{2.4}$$

$$\mathrm{tr}\big(\theta[X]\big) = \theta[X] \bullet \theta[I] = \mathrm{tr}(X) \tag{2.5}$$

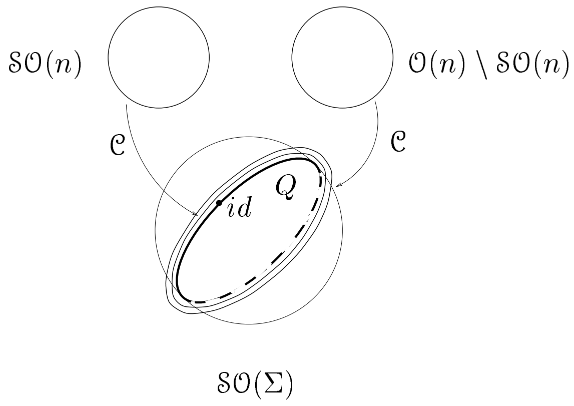

Figure 2: The action of $\mathcal{C}|_{\mathcal{O}(n)}$ when n is odd.

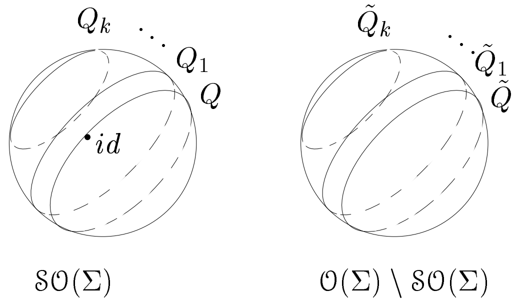

Figure 3: A priori, the discrete group structure of $\mathcal{O}(\Sigma_+) \subset \mathcal{O}(\Sigma)$ may be non-trivial.

for all X in Σ. Let us consider the particular case where $X = E_{ii}^+$ (see (2.2) for notation). Note that $E_{ii}^+ \in \Sigma_+$, whence $\theta\big[E_{ii}^+\big] \in \Sigma_+$ and there exists a vector $\mu^{(i)} \in \mathbb{R}_+^n$ with non-negative entries $\mu_k^{(i)} \geq 0$ and an orthogonal matrix $\Omega^{(i)} \in \mathcal{O}(n)$ such that

$$\theta\big[E_{ii}^+\big] = \mathcal{C}\big(\Omega^{(i)}\big)\big[\mathrm{Diag}(\mu)\big].$$

Equations (2.4) and (2.5) applied to $X = E_{ii}^+$ now imply that

$$\sum \big(\mu_k^{(i)}\big)^2 = 1 \qquad \text{and} \qquad \sum \mu_k^{(i)} = 1.$$

But if more than one of the $\mu_k^{(i)}$ is strictly positive then $\mu_k^{(i)} \in [0, 1)$ for all k and hence $\sum \big(\mu_k^{(i)}\big)^2 < \sum \mu_k^{(i)}$, a contradiction. Therefore, $\mu^{(i)}$ is one of the coordinate vectors and the eigenvalues of E_{ii}^+ are preserved under the action of θ.

Next, note that if $X \bullet Y = 0$ for some $X, Y \in \Sigma_+$ then $XY = 0$. In fact,

$$0 = X \bullet Y = \operatorname{tr}\bigl(X^{1/2} Y X^{1/2}\bigr) = \sum_k \xi_k,$$

where ξ denotes the spectrum of $X^{1/2} Y X^{1/2} \in \Sigma_+$. But since $\xi \in \mathbb{R}_+^n$ this implies that $\xi = 0$ and $X^{1/2} Y X^{1/2} = 0$. Therefore,

$$XY = X^{1/2} X^{1/2} Y X^{1/2} X^{-1/2} = 0$$

as claimed. Applying this idea in conjunction with the fact that $E_{ii}^+ E_{jj}^+ = 0$ for $i \neq j$, we get that

$$\theta\bigl[E_{ii}^+\bigr] \bullet \theta\bigl[E_{jj}^+\bigr] = E_{ii}^+ \bullet E_{jj}^+ = \operatorname{tr}(0) = 0,$$

and hence

$$\theta\bigl[E_{ii}^+\bigr] \theta\bigl[E_{jj}^+\bigr] = 0.$$

Therefore, the $\theta\bigl[E_{ii}^+\bigr]$ all commute with one another and it follows that we can choose $\Omega^{(i)} \equiv \tilde{\Omega} \in \mathcal{O}(n)$ the same for all $i \in \mathbb{N}_n$. It follows that there exists a permutation $\sigma \in \mathcal{S}_n$ and correspondingly a permutation matrix $P \in \mathcal{O}(n)$ such that

$$\theta\bigl[E_{ii}^+\bigr] = \mathcal{C}\bigl(\tilde{\Omega}\bigr)\bigl[E_{\sigma(i)\sigma(i)}^+\bigr] = \mathcal{C}\bigl(\tilde{\Omega} P\bigr)\bigl[E_{ii}^+\bigr] \tag{2.6}$$

for all $i \in \mathbb{N}_n$. Note that this shows that the restriction of θ to the set of diagonal matrices coincides with the restriction of $\mathcal{C}\bigl(\tilde{\Omega} P\bigr)$ to this same set, but this does not necessarily imply that $\theta = \mathcal{C}\bigl(\tilde{\Omega} P\bigr)$. Finally, let X be an arbitrary symmetric matrix with eigenvalues λ_i ($i \in \mathbb{N}_n$). Let $\Omega \in \mathcal{O}(n)$ be an orthogonal matrix such that

$$X = \mathcal{C}(\Omega)\bigl[\operatorname{Diag}(\lambda)\bigr].$$

It follows from Remark 2.3 that if $\mathcal{C}(\Omega) \notin Q$ then n is even and $\mathcal{C}(\Omega) \in \mathcal{C}(J)Q$. But then we can replace Ω by $J\Omega$, and hence we can assume without loss of generality that $\mathcal{C}(\Omega) \in Q$. Since Q is a normal subgroup of $\mathcal{O}(\Sigma_+)$ we find that $\theta \mathcal{C}(\Omega) \theta^{-1} \in Q$ and hence there exists $\hat{\Omega} \in \mathcal{SO}(n)$ such that $\mathcal{C}\bigl(\hat{\Omega}\bigr) = \theta \mathcal{C}(\Omega) \theta^{-1}$. Therefore,

$$\begin{aligned} \theta[X] &= \theta \mathcal{C}(\Omega)\bigl[\operatorname{Diag}(\lambda)\bigr] = \mathcal{C}\bigl(\hat{\Omega}\bigr)\theta\bigl[\operatorname{Diag}(\lambda)\bigr] \\ &\overset{(2.6)}{=} \mathcal{C}\bigl(\hat{\Omega}\tilde{\Omega} P\bigr)\bigl[\operatorname{Diag}(\lambda)\bigr], \end{aligned}$$

and since $\hat{\Omega}\tilde{\Omega} P \in \mathcal{O}(n)$ this shows that the eigenvalues of $\theta[X]$ and X are the same. $\square$

Theorem 2.5. *All rank one preserving endomorphisms of the linear space of symmetric $n \times n$ matrices over $\mathbb{R}$ are of the form $\mathcal{C}(S)$ for some $S \in Gl_n(\mathbb{R})$.*

Proof. This is a special case of a theorem due to Lim [14] a generalization of which was also given by Waterhouse [26] who characterised all rank one preservers of spaces of symmetric matrices with coefficients in commutative rings. □

Corollary 2.6. $\mathcal{O}(\Sigma_+) = \mathcal{C}(\mathcal{O}(n))$.

Proof. Our proof is similar to the argument given by Güler [8]. We already know from Remark 2.3 that $\mathcal{C}(\mathcal{O}(n)) \subseteq \mathcal{O}(\Sigma_+)$. Let $\theta \in \mathcal{O}(\Sigma_+)$. Then θ is eigenvalue preserving on Σ by Lemma 2.4, and hence it maps symmetric rank one matrices to symmetric rank one matrices. It follows from Lim's theorem 2.5 that there exists a non-singular matrix $S \in Gl_n(\mathbb{R})$ such that $\theta = \mathcal{C}(S)$. But since θ is eigenvalue preserving we must have that

$$I = \theta[I] = SIS^T,$$

which shows that $S \in \mathcal{O}(n)$. □

The insight we have gained into the structure of $\mathcal{O}(\Sigma_+)$ can now be used to characterise the structure of $\mathrm{Aut}(\Sigma_{++})$. The following theorem is key:

Theorem 2.7 (Rothaus). *For every $\theta \in \mathrm{Aut}(\Sigma_+)$ there exists a unique $\omega \in \mathcal{O}(\Sigma_+)$ and a unique $W \in \Sigma_{++}$ such that θ has* polar decomposition

$$\theta = \omega \circ \mathcal{C}(W).$$

Proof. This is a special case of a theorem due to Rothaus [21]. □

Corollary 2.8. $\mathrm{Aut}(\Sigma_+) = \mathcal{C}(Gl_n(\mathbb{R}))$. □

Proof. This result follows now trivially from Corollary 2.6, Theorem 2.7 and the fact that $\mathcal{C} : Gl_n(\mathbb{R}) \to \mathrm{Aut}(\Sigma_+)$ is a group homomorphism. □

3 Self-Scaled Barrier Functionals for the Positive Semidefinite Symmetric Cone

Now that we understand the structure of $\mathrm{Aut}(\Sigma_+)$, we can start to characterise the set of self-scaled barrier functionals for the positive semidefinite cone as a subset of the family of functions of the form

$$-c_1 \ln \det + c_0.$$

We achieve this task by comparing the Hessians of an arbitrary self-scaled barrier with the Hessians of the standard logarithmic barrier functional. We shall therefore start by giving a precise characterisation of the subset of $\mathrm{Aut}(\Sigma_+)$ where these Hessians are bound to lie.

3.1 Self-Adjoint Positive Definite Automorphisms of the Positive Semidefinite Symmetric Cone

Lemma 3.1. *The set of self-adjoint positive definite automorphisms of Σ that preserve Σ_+ coincides with $\mathcal{C}(\Sigma_{++})$.*

Proof. Let $\theta \in \mathrm{Aut}(\Sigma_+)$ be a self-adjoint positive definite automorphism of Σ. It follows from Corollary 2.8 that there exists a matrix $A \in Gl_n(\mathbb{R})$ such that $\theta = \mathcal{C}(A)$. But θ is self-adjoint if and only if

$$\mathrm{tr}(AXA^TY) = \mathrm{tr}(XAYA^T) = \mathrm{tr}(A^TXAY)$$

for all $X, Y \in \Sigma$. In particular, for $Y = E^+_{ij}$ (see (2.2) *iv*)) we get

$$\begin{aligned} 2(A^TXA)_{ij} &= (A^TXA)_{ij} + (A^TXA)_{ji} \\ &= \mathrm{tr}(A^TXAY) = \mathrm{tr}(AXA^TY) \\ &= \cdots = 2(AXA^T)_{ij} \end{aligned}$$

for all $X \in \Sigma$ and $i, j \in \mathbb{N}$. Therefore, $A^{-T}AX = XAA^{-T}$ for all $X \in \Sigma$. For $X = I$ we get $A^{-T}A = AA^{-T}$, and hence by the previous equation, $A^{-T}AX = XA^{-T}A$ for all $X \in \Sigma$. It follows that $A^{-T}A = \lambda I$ for some $\lambda \in \mathbb{R} \setminus \{0\}$. Hence, $A = \lambda A^T$ and, upon taking the transpose, $A^T = \lambda A$. Therefore, $A = \lambda^2 A$ and $\lambda \in \{\pm 1\}$. This shows that A must either be symmetric or skew-symmetric. It is easy to check that this condition is also sufficient for self-adjointness. Suppose $A = (a_{ij})$ is skew-symmetric. Since $A \in Gl_n(\mathbb{R})$, there exists a pair of indices $k \neq l \in \mathbb{N}_n$ such that $a_{kl} \neq 0$. But then we have

$$E^+_{kl} \bullet \theta[E^+_{kl}] = -2a^2_{kl} < 0,$$

and this shows that θ is not positive definite on Σ, contrary to what we assumed. Hence, A must be symmetric and is thus non-defective. Lemma 2.2 *iv*) shows that the spectrum of $\mathcal{C}(A)$ is given by $\{\lambda_i\lambda_j : i \leq j\}$ where $\{\lambda_1, \dots, \lambda_n\}$ is the spectrum of A. Therefore, $\mathcal{C}(A)$ is positive definite if and only if all the eigenvalues of A are of equal sign. This shows that $A \in \pm\Sigma_{++}$. But if $A \in -\Sigma_{++}$ then $-A \in \Sigma_{++}$ and clearly $\theta = \mathcal{C}(-A)$. This proves one direction of the claim. The other direction follows immediately from Lemma 2.2 *iv*). □

Note that Lemma 3.1 merely amounts to a weaker form of Theorem 2.7. In fact, for our proof we could have proceeded as follows: Let $\theta \in \mathrm{Aut}\,\Sigma_+$ be self-adjoint positive definite. Then its unique polar decomposition in the sense of Cartan as an automorphism of $(\Sigma, \bullet)$ is given by

$$\theta = id \circ \theta. \tag{3.1}$$

On the other hand, Theorem 2.7 implies that there is a unique polar decomposition of $\theta \in \mathrm{Aut}(\Sigma_+)$ in the sense of Rothaus,

$$\theta = \omega \circ \mathcal{C}(W). \tag{3.2}$$

But (3.2) is also a polar decomposition of θ in the sense of Cartan, since it is easy to check that $\mathcal{C}(W)$ is self-adjoint positive definite, without invoking Lemma 3.1. Hence, it follows from the uniqueness of Cartan's decomposition that the decompositions given in Equations (3.1) and (3.2) coincide. Therefore, ω is the identity mapping and $\theta = \mathcal{C}(W)$ for some $W \in \Sigma_{++}$.

It is interesting to note that, interpreted in this light, Theorem 2.7 says that Cartan's decomposition of a cone automorphism consists of cone automorphisms itself.

3.2 Characterisation of the Hessians of Self-Scaled Barriers for the Positive Semidefinite Symmetric Cone

Let us now compare the Hessians of an arbitrary self-scaled barrier functional H on Σ_{++} with the Hessians of the standard logarithmic barrier functional

$$F : X \mapsto -\ln \det X.$$

Recall from Section 1.3 that for each $X \in \Sigma_{++}$ we have

$$F''(X) = \mathcal{C}\big(X^{-1}\big).$$

Since the Hessians of H are self-adjoint positive definite automorphisms of Σ_+, it follows from Lemma 3.1 that there exists a well-defined mapping $\Upsilon : \Sigma_{++} \to \Sigma_{++}$ such that

$$H''(X) = \mathcal{C}\big(\Upsilon(X)^{-1}\big) \qquad \text{for all } X \in \Sigma_{++}. \tag{3.3}$$

Analogously, if $H_\sharp$ denotes the dual barrier defined by H then there exists a well-defined mapping $\Upsilon_\sharp : \Sigma_{++} \to \Sigma_{++}$ such that

$$H''_\sharp(S) = \mathcal{C}\big(\Upsilon(S)^{-1}\big) \qquad \text{for all } S \in \Sigma_{++}. \tag{3.4}$$

Since both F and H are self-scaled, Theorem 1.2 *iii*) implies that

$$\begin{aligned} &F''(X) = F''(W_F) \circ F''_\sharp(S) \circ F''(W_F) \qquad \text{and} \\ (3.5) \qquad &H''(X) = H''(W_H) \circ H''_\sharp(S) \circ H''(W_H) \qquad \text{for all } X, S \in \Sigma_{++}, \end{aligned}$$

where $W_F = W_F(X,S)$ and $W_H = W_H(X,S)$ denote the (primal) scaling points defined by X and S for the the self-scaled barriers F and H respectively (see Theorem 1.2).

By definition of W_F, W_H and Υ we have

$$\begin{aligned} W_F^{-1} X W_F^{-1} &= F''(W_F)[X] \\ &= S \\ &= H''(W_H)[X] \\ &= \Upsilon(W_H)^{-1} X \Upsilon(W_H)^{-1}. \end{aligned}$$

It thus follows from the uniqueness of $W_F(X,S)$ (see Theorem 1.2) that

$$(3.6) \qquad \Upsilon\big(W_H(X,S)\big) = W_F(X,S) \qquad \forall X, S \in \Sigma_{++}.$$

Equations (3.3), (3.4) and (3.5) now imply that

$$\Upsilon(X)^{-1} = \Upsilon(W_H)^{-1} \Upsilon_\sharp(S)^{-1} \Upsilon(W_H)^{-1}.$$

Together with (3.6) this shows that

$$(3.7) \qquad \Upsilon(X) = W_F \Upsilon_\sharp(S) W_F.$$

In particular, using the fact that $W_F(I,I) = I$ in (3.7) we get

$$(3.8) \qquad \Upsilon_\sharp(I) = \Upsilon(I),$$

and since $W_F(X,I) = X^{1/2}$, (3.7) and (3.8) imply that

$$(3.9) \qquad \Upsilon(X) = X^{1/2} \Upsilon(I) X^{1/2} = \Upsilon_\sharp(X) \qquad \forall X \in \Sigma_{++}.$$

Recall from Section 1.3 that

$$W_F(X,S) = S^{-1/2}\big(S^{1/2}XS^{1/2}\big)^{1/2}S^{-1/2}.$$

Therefore, it is true that

$$\begin{aligned} X^{1/2}\Upsilon(I)X^{1/2} &\overset{(3.9)}{=} \Upsilon(X) \\ &\overset{(3.7)}{=} W_F(X,S)\Upsilon_\sharp(S)W_F(X,S) \\ &\overset{(3.9)}{=} W_F(X,S)S^{1/2}\Upsilon(I)S^{1/2}W_F(X,S), \end{aligned}$$

and hence

$$\Upsilon(I) = \mathcal{C}\Big(X^{-1/2}S^{-1/2}\big(S^{1/2}XS^{1/2}\big)^{1/2}\Big)\big[\Upsilon(I)\big] \qquad \text{for all } X,S \in \Sigma_{++}.$$

Clearly, this condition is equivalent to

$$\Upsilon(I) = \mathcal{C}\Big(N^{-1}\big(NN^T\big)^{1/2}\Big)\big[\Upsilon(I)\big] \qquad \forall N \in \mathcal{N}, \tag{3.10}$$

where $\mathcal{N} := \big\{XS : X,S \in \Sigma_{++}\big\}$ is the set of $n \times n$ matrices that can be written as the product of two symmetric positive definite matrices.

In an e-mail message sent to the interior-point mailing list in September of 1999, Mike Todd asked how to characterise precisely this set, as a toy question to play around with. Various recipients replied to the message and came up with the following solution:

Lemma 3.2. *$\mathcal{N}$ coincides with the set of non-defective $n \times n$ matrices with real coefficients, all of whose eigenvalues are strictly positive real numbers.*

Proof. This is the solution due to Mike Todd: If $N = XS$, then

$$\begin{aligned} N &= X^{1/2}\big(X^{1/2}SX^{1/2}\big)X^{-1/2} \\ &= X^{1/2}\big(QDQ^T\big)X^{-1/2} = PDP^{-1}, \end{aligned}$$

where QDQ^T is the spectral decomposition of the symmetric positive definite matrix $X^{1/2}SX^{1/2}$, and $P := X^{1/2}Q$. This gives the eigenvalue decomposition of N, with eigenvalues the positive entries of D and eigenvectors the columns of P. Conversely, suppose $N = PDP^{-1}$, where D is diagonal with positive diagonal entries. Then we can write $N = \big(PP^T\big)\big(P^{-T}DP^{-1}\big) =: XS$. $\square$

Note that $N^{-1}(NN^T)^{1/2} \in \mathcal{SO}(n)$ for all $N \in \mathcal{N}$. In the next proposition we characterise the closed subgroup of $\mathcal{SO}(n)$ generated by matrices of this form. This result constitutes the mathematical core mechanism of our paper.

Proposition 3.3. *The closed subgroup of $\mathcal{SO}(n)$ generated by the set of orthogonal matrices of the form $N^{-1}(NN^T)^{1/2}$ with $N \in \mathcal{N}$ coincides with $\mathcal{SO}(n)$.*

Proof. Let this closed subgroup be denoted by G, and let $\mathfrak{g}$ be its Lie Algebra. Since $\mathcal{SO}(n)$ is simply connected, it suffices to show that $\mathfrak{g} = \mathfrak{so}(n)$, or in other words that the tangent space of the manifold G at the point $I \in G$ coincides with the set Σ^i of $n \times n$ skew-symmetric matrices over the reals. In fact, use of the exponential mapping

$$\exp : \mathfrak{so}(n) \to \mathcal{SO}(n)$$

shows that G and $\mathcal{SO}(n)$ share a neighborhood $\mathcal{V}$ of I, and parallel transport by left trivialisation shows that G and $\mathcal{SO}(n)$ share the neighborhood $g\mathcal{V}$ of any element $g \in G$. Therefore, G is both open and closed as a subset of $\mathcal{SO}(n)$, and since $\mathcal{SO}(n)$ is simply connected, the result follows.

We now proceed to showing that $\mathfrak{g} = \mathfrak{so}(n)$: Let $\Delta \in \mathcal{N}$ have eigenvalues $\lambda_1, \dots \lambda_n > 0$. Then $\Delta(t) := I + t\Delta \in \mathcal{N}$ for all $t > 0$, since the n linearly independent eigenvectors of Δ are also eigenvectors of $\Delta(t)$ and correspond to the strictly positive eigenvalues $t\lambda_i(t) + 1 > 0$, $(i = 1, \dots, n)$. The Neumann-series development of $\Delta(t)^{-1}$ shows that

$$\Delta(t)^{-1} = I - t\Delta + O(t^2). \tag{3.11}$$

Upon taking squares on both sides of the ansatz

$$(\Delta(t)\Delta(t)^T)^{1/2} = I + tD + O(t^2) \tag{3.12}$$

we get

$$I + t(\Delta + \Delta^T) + O(t^2) = I + 2tD + O(t^2),$$

and hence $D = 1/2(\Delta + \Delta^T)$.

Equations (3.11) and (3.12) thus yield the identity

$$\begin{aligned}\Delta(t)^{-1}\big(\Delta(t)\Delta(t)^T\big)^{1/2} &= \big(I - t\Delta\big)\big(I + \frac{t}{2}(\Delta + \Delta^T)\big) + O(t^2) \\ &= I + \frac{t}{2}(\Delta^T - \Delta) + O(t^2),\end{aligned}$$

and this shows that

$$\Delta^T - \Delta \in \mathfrak{g} \qquad \text{for all } \Delta \in \mathcal{N}. \tag{3.13}$$

Clearly, we have $\Delta^T - \Delta \in \mathfrak{so}(n)$ as expected.

On the other hand, for $P := \left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)$ and $D := \left(\begin{smallmatrix} .5 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ we get $P^{-1}DP = \left(\begin{smallmatrix} .5 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. Hence,

$$(P^{-1}DP)^T - P^{-1}DP = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right).$$

Let P_{ij} be the permutation matrix that permutes the i^{th} and j^{th} variables, and recall the notation introduced in (2.2). Then consider

$$\Delta := \left(P_{1i}P_{2j}\left(\begin{smallmatrix} P^{-1} & \\ & I \end{smallmatrix}\right)\right)\left(\begin{smallmatrix} D & \\ & I \end{smallmatrix}\right)\left(\left(\begin{smallmatrix} P & \\ & I \end{smallmatrix}\right)P_{j2}P_{i1}\right).$$

Clearly, $\Delta \in \mathcal{N}$ and $\Delta^T - \Delta = E^-_{ij}$. But since $\{E^-_{ij} : i, j \in \mathbb{N}_n\}$ forms a basis of $\mathfrak{so}$ we find that the elements of $\mathfrak{g}$ span this whole space. By the remark made at the beginning, this shows the claim. □

Proposition 3.4. *If F is the standard logarithmic barrier functional for the positive semidefinite cone Σ_{++} and if H is an arbitrary self-scaled barrier for the same cone, then there exists a positive constant $\lambda > 0$ such that $H''(X) = \lambda F''(X)$ for all $X \in \Sigma_{++}$.*

Proof. The invariance property (3.10) is clearly preserved under taking compositions and limits. Hence, Lemma 3.2 implies that the symmetric positive definite matrix $\Upsilon(I)$ satisfies the condition $\Upsilon(I) = O\Upsilon(I)O^T$ for all $O \in \mathcal{SO}(n)$. It is a trivial matter to prove that this forces $\Upsilon(I)$ to be a scalar, and the result follows from Equation (3.9). □

3.3 Self-Scaled Barrier Functionals for the Positive Semidefinite Symmetric Cone are Isotropic

The following result constitutes the main theorem of this paper:

Theorem 3.5. *If H is a self-scaled barrier functional for Σ_+ then there exist constants $c_1 > 0$ and $c_0 \in \mathbb{R}$ such that*

$$H : X \mapsto -c_1 \ln \det(X) + c_0 \qquad \forall X \in \Sigma_{++}.$$

Equivalently, we could state that H is of the form $c_1 U + c_0$ for some constants $c_1 > 0$ and $c_0 \in \mathbb{R}$, and where U denotes the universal barrier functional defined on Σ_{++}.

Proof. It follows from Proposition 3.4 and the fundamental theorem of differential and integral calculus that H is of the form $c_1F+\varphi+c_0$, where $c_1=\lambda>0$, $c_0\in\mathbb{R}$ and $\varphi\in\Sigma^\sharp$ is a linear functional on Σ, i.e. there exists a symmetric matrix $Y\in\Sigma$ such that $\varphi: X\mapsto X\bullet Y$ for all $X\in\Sigma$. One of the conditions in the definition of a ν–self-concordant barrier functional B for a convex open domain D is that the length of the Newton step $\nabla^2B(x)^{-1}[-\nabla B(x)]$ at any point $x\in D$ measured in the local Riemannian metric $\|.\|_x$ defined by the Hessian of B at x be uniformly bounded by $\nu^{1/2}$ (see e.g. [16, 17, 20]):

$$\begin{aligned}\|\nabla^2B(x)^{-1}[-\nabla B(x)]\|_x &:= \big(\nabla^2B(x)[-\nabla^2B(x)^{-1}\nabla B(x)];-\nabla^2B(x)^{-1}\nabla B(x)\big)\\ &= \big(\big(\nabla^2B(x)\big)^{-1}[\nabla B(x)];\nabla B(x)\big)\leq\nu^2.\end{aligned}$$

In particular, in the case of H this means that

$$\begin{aligned}\nu &\geq \|\nabla H(X)\|_X^2\\ &= \|Y-\lambda X^{-1}\|_X^2\\ &\overset{C.S.}{\geq} \big(\|Y\|_X-\|\lambda X^{-1}\|_X\big)^2\\ &\overset{3.4}{=} \Big((\lambda^{-1}XYX\bullet Y)^{1/2}-\big(\lambda^{-1}X(\lambda X^{-1})X\bullet(\lambda X^{-1})\big)^{1/2}\Big)^2\\ &= \lambda^{-1}\Big(\big(\mathrm{tr}\big[(XY)^2\big]\big)^{1/2}-\lambda n^{1/2}\Big)^2\end{aligned}$$

for all $X\in\Sigma_{++}$. But clearly, this implies that $Y=0$. $\square$

In Definition 4.1 [10] we defined isotropic barrier functionals as self-scaled and invariant under the orthogonal group of their conic domain of definition. Note that Theorem 3.5 and Proposition 2.4 imply that all self-scaled barriers of Σ_+ are isotropic, since $\det X$ is a function of the eigenvalues of $X\in\Sigma_{++}$.

It is also possible to take the reverse approach and prove Theorem 3.5 via isotropy. In Theorem 4.8 [10] we showed that any self-scaled barrier defined on an irreducible symmetric cone K is of the form

$$x\mapsto c_1\ln\int_{K^\sharp}\exp\big\{\langle x;s\rangle\big\}ds+c_0$$

for some constants $c_1>0$ and $c_0\in\mathbb{R}$. Güler [7] showed that functionals of this form can be equivalently expressed as c_1U+c_0 for some $c_1>0$ and $c_0\in\mathbb{R}$, where U is the universal barrier functional of K (see [16]), and hence they are of the form $-c_1\ln\det+c_0$ in the case where $K=\Sigma_+$. In order to prove Theorem 3.5 it suffices therefore to show that every self-scaled barrier functional H for Σ_+ is isotropic. Let us now present a proof of this claim which relies on arguments that differ somewhat from those given in the proof of Theorem 3.5. This illustrates that there is a certain degree of redundancy in the axiomatic theory of self-scaledness:

Theorem 3.6 (Theorem 3.5 revisited).
Any self-scaled barrier functional for the positive semidefinite cone is isotropic.

Proof. Let H be a self-scaled barrier for Σ_+. It is elementary to see that the logarithmic homogeneity of H implies (see e.g. [17])

$$\begin{aligned} H''(\tau X) &= \tau^{-2} H''(X) \quad \forall \tau > 0 \quad \text{and} \\ -H'(X) &= H''(X)[X] \quad \forall X \in \Sigma_{++}. \end{aligned} \tag{3.14}$$

Lemma 4.19 [10] and Proposition 3.4 imply that H is invariant under the action of any $\theta \in \mathcal{O}(\Sigma_+)$ defined by

$$\begin{aligned} \theta_{(X,S)} := H''_\sharp\big(T_H(\lambda^{1/2} I, -H'(X))\big) \circ H''\big(W_H(X,S)\big) \\ \circ H''_\sharp\big(T_H(\lambda^{1/2} I, S)\big) \circ H''(\lambda^{1/2} I) \end{aligned} \tag{3.15}$$

for some $X, S \in \Sigma_{++}$, where T_H denotes the dual scaling point of X and S (see Theorem 1.2 *iv*)).

For all $X \in \Sigma_{++}$ we have

$$-H'(X) \overset{(3.14),3.4}{=} \lambda X^{-1} X X^{-1} = \lambda X^{-1}.$$

Therefore, Condition (3.15) shows that H is invariant under all orthogonal automorphisms of the form $\mathcal{C}(\Omega)$, where $\Omega \in \mathcal{O}(n)$ is given by

$$\begin{aligned} \Omega &:= \lambda^{-1/4} X^{1/2} \big(S^{1/2} \big(S^{1/2} X S^{1/2}\big)^{-1/2} S^{1/2}\big) \lambda^{1/4} S^{-1/2} I \\ &= X^{1/2} S^{1/2} \big(S^{1/2} X S^{1/2}\big)^{-1/2} \end{aligned}$$

for some $X, S \in \Sigma_{++}$. But the set of all such Ω coincides precisely with

$$\{N^{-1}(NN^T)^{1/2} : N \in \mathcal{N}\},$$

whence it follows from Proposition 3.3 that H is invariant under $\mathcal{C}(\mathcal{SO}(n)) = Q$.

Finally, it was shown in [10] that any self-scaled barrier functional invariant under Q is in fact invariant under the action of any element of $\mathcal{O}(\Sigma_+)$ (since the orbit of the conic hull of a frame of Σ_+ under Q covers all of Σ_+), and hence H is isotropic. $\square$

Note that all our arguments prior to Theorem 3.6 depended only on twice continuous differentiability of H. Hence, we could have replaced the condition $H \in \mathcal{C}^3(K^\circ, \mathbb{R})$ in Definition 1.1 by $H \in \mathcal{C}^2(K^\circ, \mathbb{R})$, and then Theorem 3.5 implies that H is actually in $\mathcal{C}^\infty$. The proof of Theorem 3.6 assumes that $H \in \mathcal{C}^3$, since the classification theory of isotropic barrier functionals depends on this condition.

4 Discussion of Implications

One obvious consequence of Theorem 3.5 is that it simplifies the self-scaled theory [17] for semidefinite programming: Instead of working with an extended set of axioms that define self-scaled barrier functionals one may assume that barrier functionals are given by the explicit expression

$$H = -c_1 \ln \det + c_0. \tag{4.1}$$

Moreover, all results on semidefinite programming that have been derived on the basis of the supposedly stronger assumption (4.1) are seen to hold in the completely general self-scaled setting.

Let us now state a few additional interesting facts that derive from this result:

Corollary 4.1. *All self-scaled barriers for the positive semidefinite cone are smooth.*

Proof. This fact which was already mentioned in Section 3.3 follows immediately from the smoothness of the standard barrier function and from Theorem 3.5. □

Theorem 4.2. *A self-concordant barrier functional H on Σ_{++} is self-scaled if and only if H changes only by an additive constant when composed with an automorphism of Σ_+, i.e. if and only if for each $\theta \in \mathrm{Aut}(\Sigma_+)$ there exists a $C \in \mathbb{R}$ such that*

$$H \circ \theta \equiv H + C. \tag{4.2}$$

Moreover, C as a function of θ is a continuous group homomorphism from $\mathrm{Aut}(\Sigma_+)$ to the additive group $\mathbb{R}$.

Proof. The very last statement of the theorem is trivial to check.

The only if part is a trivial consequence of Theorem 3.5, since the standard logarithmic barrier functional $F = -\ln \circ \det$ satisfies Equation (4.2). In fact, by virtue of Corollary 2.8 it is the case that for every $\theta \in \mathrm{Aut}(\Sigma_+)$ there exists a matrix $A \in Gl_n(\mathbb{R})$ such that $\theta = \mathcal{C}(A)$, and hence,

$$F \circ \theta \equiv F - 2 \det A.$$

For the if part, let us suppose that H is a ν-self-concordant barrier functional for Σ_{++} that satisfies Condition (4.2). Let us then consider the function $C : \mathrm{Aut}(\Sigma_+) \to \mathbb{R}$ defined by this condition, and let us first show that

$$C(\theta) = 0 \qquad \text{for } \theta \in \mathcal{O}(\Sigma_+). \tag{4.3}$$

We follow an argument already given in the proof of Theorem 4.15 in [10]. Suppose to the contrary that $C(\theta) \neq 0$ for some $\theta \in \mathcal{O}(\Sigma_+)$. Then we may assume that $C(\theta) > 0$, since $C(\theta^{-1}) = -C(\theta)$ due to the last statement of the proposition. For a fixed $X \in \Sigma_{++}$ the distance to the boundary of Σ_+ is constant for all the points on the orbit $\{\theta^n[X] : n \in \mathbb{Z}\}$ under the action of the group generated by θ, since all the transformations θ^n lie in $\mathcal{O}(\Sigma_+)$. Hence, the orbit of X is bounded away from the boundary, and yet

$$\lim_{n\to\infty} H\big(\theta^n[X]\big) = H(X) + \lim_{n\to\infty} nC(\theta) = +\infty.$$

It is a well known fact from the theory of self-concordant barrier functionals (see [16] and [20]) that this is impossible. Hence, $C(\theta) = 0$ as claimed.

Next, let $\tau > 0$ and $i \in \mathbb{N}_n$, and let us investigate how C acts on

$$\theta_\tau^{[i]} := \mathcal{C}\big(I + (\tau - 1)E_{ii}^+\big),$$

i.e. $\theta_\tau^{[i]}$ is the conjugation mapping defined by the diagonal matrix all of whose diagonal entries are one except for the i-th entry which is τ. Note that

$$\theta_\tau^{[i]} = \mathcal{C}\big(P_{1i}\big) \circ \mathcal{C}\big(\theta_\tau^{[1]}\big) \circ \mathcal{C}\big(P_{1i}\big),$$

where P_{1i} denotes the permutation matrix that swaps the i-th with the j-th coordinate. Since $P_{1i} \in \mathcal{O}(n)$ it follows from Corollary 2.6 and (4.3) that

$$C\big(\theta_\tau^{[i]}\big) = C\big(\theta_\tau^{[1]}\big) \quad \forall i \in \mathbb{N}_n. \tag{4.4}$$

Let us thus concentrate on $C\big(\theta_\tau^{[1]}\big)$ and use the shorthand notation $C(\tau)$ for this function of $\tau \in \mathbb{R}_+$. Then C is a continuous group homomorphism from the multiplicative group $\mathbb{R}_+$ to the additive group $\mathbb{R}$. Therefore, $C\circ\exp$ is a continuous endomorphism of the additive group $\mathbb{R}$. It is a well known and indeed trivial fact that all such endomorphisms are of the form $x \mapsto -\kappa x$ for some $\kappa \in \mathbb{R}$. This shows that

$$C(\tau) = -\kappa \ln \tau \qquad \forall \tau > 0. \tag{4.5}$$

But since $\theta_\tau^{[1]}$ converges to the boundary of Σ_+ as τ goes to zero, it must be the case that

$$-\lim_{\tau\to 0^+} \kappa \ln \tau = \lim_{\tau\to 0^+} H\big(\theta_\tau^{[1]}\big) = +\infty,$$

and hence we have $\kappa > 0$.

Next, let θ be an arbitrary element of $\mathrm{Aut}(\Sigma_+)$, and let $\theta = \omega \circ \mathcal{C}(W)$ be its unique polar decomposition (see 2.7), where $\omega \in \mathcal{O}(\Sigma_+)$ and $W \in \Sigma_{++}$. Let $W = P\,\mathrm{Diag}\big(\sigma(W)\big)P^T$ be a spectral decomposition of W. Then

$$\theta = \omega \circ \mathcal{C}(P) \circ \prod_{i=1}^{n} \theta^{[i]}_{\sigma(W)_i} \circ \mathcal{C}(P^T),$$

and hence it follows from (4.3),(4.5) and from the fact that C is a group homomorphism that

$$C(\theta) = -\kappa \sum_{i=1}^{n} \ln \sigma(W)_i = -\kappa \ln\big(\det(W)\big). \tag{4.6}$$

Finally, for any $X \in \Sigma_{++}$ we have

$$\begin{aligned} H(X) \quad &= H\big(\mathcal{C}(X^{1/2})[I]\big) \\ &\overset{(4.2),(4.6)}{=} -\kappa \ln\big(\det(X^{1/2})\big) + H(I) \\ &= -c_1 \ln(\det X) + c_2 \end{aligned}$$

with $c_1 = \kappa/2 > 0$ and $c_2 = H(I)$. This proves the claim. On a tangent, let us also note that logarithmic homogeneity implies that $\nu = nc_1$. $\square$

In partially unpublished work, Güler and Tunçel investigated so-called invariant barrier functionals (see e.g. [23], p.124): A self-concordant barrier functional $H : K^\circ \to \mathbb{R}$ for a regular cone is called *invariant* if there exists a transitive subgroup G of $\mathrm{Aut}(K)$ such that for all $\theta \in G$ there exists a constant $C(\theta)$ such that

$$H \circ \theta \equiv H + C(\theta).$$

Invariance can thus be regarded as a weaker notion of self-scaledness. In fact, interpreted in the light of this theory, our Corollary 4.2 says that a self-concordant barrier functional is self-scaled in the sense of Nesterov and Todd if and only if it is invariant under the transitive group $\mathrm{Aut}(\Sigma_+)$.

Let us also take a fresh look at the "only if" part of Corollary 4.2 in this light. It is well known that it follows from Definition 1.1 that the pullback of any self-scaled barrier functional $H : K^\circ \to \mathbb{R}$ for a symmetric cone K under a cone automorphism $\theta := H''(w)$ defined by the Hessian of H at some point $w \in K^\circ$ is a mere translation of H, and it follows from Theorem 1.2 that the set of cone automorphisms defined in this way is transitive. Hence, H is invariant in the sense of Güler and Tunçel under the transitive group given by the subgroup of $\mathrm{Aut}(\Sigma_+)$ generated by the Hessians of H. For the special case where K is the positive semidefinite symmetric cone, the "only if" part of Corollary 4.2 implies the non-trivial fact that H is also invariant under the closure of this group and raises the question of whether this closure is equal to the automorphism group $\mathrm{Aut}(\Sigma_+)$ altogether. We will now give a precise answer to this question:

Corollary 4.3. *Let H be a self-scaled barrier functional for Σ_+. Then the closed subgroup of* $\mathrm{Aut}(\Sigma_+)$ *generated by the set of Hessians of H coincides with the orientation preserving part of* $\mathrm{Aut}(\Sigma_+)$.

Proof. Due to Proposition 3.4 we only need to consider the case where H is the standard barrier functional. But then $H''(X) = \mathcal{C}(X^{-1})$ for all $X \in \Sigma_{++}$. This shows that the set of Hessians of H coincides with the set of conjugations by positive definite matrices. Since all of these automorphisms are orientation preserving, the closed subgroup of $\mathrm{Aut}(\Sigma_+)$ they generate must lie in the orientation preserving connected component of $\mathrm{Aut}(\Sigma_+)$. It follows from Theorem 2.7 that it only remains to show that $\mathcal{SO}(\Sigma_+)$ consists of limits of automorphisms generated by the Hessians of H. But this has already been established in the proof of Theorem 3.6 and hence the claim is true. ☐

5 Concluding Remarks

Our main Theorem 3.5 shows that all self-scaled barrier functionals for use in semidefinite programming are related to the standard logarithmic barrier via a homothetic transformation and a shift in the image space. It is known that amongst the barriers of this form the standard barrier minimises the complexity parameter (see [10]) and in fact does so for a wider class of barrier functionals (see [9]).

Theorem 3.5 also settles Conjecture 4.18 [10] in the case of the positive semidefinite cone. This conjecture constitutes the missing piece in the classification theory of self-scaled barrier functionals. Apart from the case of the positive semidefinite cone treated in the present paper and the trivial case of the positive real line, there are four other non-trivial cases to consider, all of which should in principle be provable along the same lines as our approach in Section 3.

Acknowledgements

I wish to express my warmest thanks to Syvert Nørsett for inviting me to NTNU Trondheim, and to Peter and Jonathan Borwein for inviting me to SFU in Vancouver.

References

[1] F. Alizadeh and S. Schmieta. Optimization with Semidefinite, Quadratic and Linear Constraints. RUTCOR, RRR Report 23-97, November 1997, 640 Bartholomew Rd, Piscataway NJ 08854.

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

[3] G. Golub and C. Van Loan. *Matrix Computations.* Third Edition, The Johns Hopkins University Press, London, 1996.

[4] L. Faybusovich and R. Arana. A long-step primal-dual algorithm for the symmetric programming problem. University of Notre Dame, December 1999.

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

[6] F. G. Frobenius. Ueber die Dartstellung der endlichen Gruppen durch lineare Substitutionen. *Sitzungsberichte der König. Preuss. Akad. der Wissenschaften zu Berlin*, (1897), 944–1015.

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

[8] O. Güler. Private communication. Unpublished, January 1998.

[9] O. Güler and L. Tunçel. Characterization of the barrier parameter of homogeneous convex cones. *Mathematical Programming*, 81:55–76, 1998.

[10] R. A. 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.

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

[12] R. A. Horn and C. R. Johnson *Topics in Matrix Analysis.* Cambridge University Press, Cambridge, England, 1991.

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

[14] M. H. Lim. Linear transformations on symmetric matrices. *Linear and Multilinear Algebra*, 7:47–57, 1979.

[15] M. Marcus and B. Moyls. Linear transformations on algebras of matrices. *Canadian Journal of Mathematics*, 11:61–66,1959.

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

[17] 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.

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

[19] J. de Pillis. Linear transformations which preserve hermitian and positive semidefinite operators. *Pacific Journal of Mathematics*, 23:129–137, 1967.

[20] J. Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization.* SIAM Publications. SIAM, Philadelphia, USA, August 2000.

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

[22] I. Schur. Einige Bemerkungen zur Determinantentheorie. *Sitzungsberichte der Preuss. Akad. der Wissenschaften, Phys.-Math. Kl.*, (1925), 454–463.

[23] L. Tunçel. Convex optimization: Barrier functions and interior-point methods. Technical Report B-336, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, Japan, March 1998.

[24] L. Tunçel. Potential Reduction and Primal-Dual Methods. In H. Wolkowicz, R. Saigal and L. Vandenberghe (eds.), *Handbook of Semidefinite Programming: Theory, Algorithms and Applications.* Kluwer Academic Publishers, Boston–Dordrecht–London, 2000.

[25] 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.

[26] W. C. Waterhouse. Linear Transformations Preserving Symmetric Rank One Matrices. *Journal of Algebra*, 125:502–518, 1989.