UNIVERSITY OF CAMBRIDGE

Numerical Analysis Reports

SELF–SCALED BARRIERS FOR IRREDUCIBLE SYMMETRIC CONES

Raphael Hauser and Yongdo Lim

DAMTP 2001/NA04
April 2001

Department of Applied Mathematics and Theoretical Physics
Silver Street
Cambridge
England CB3 9EW

SELF–SCALED BARRIERS FOR IRREDUCIBLE SYMMETRIC CONES

RAPHAEL A. HAUSER, YONGDO LIM

April 2, 2001

Abstract. Self–scaled barrier functions are fundamental objects in the theory of interior–point methods for linear optimization over symmetric cones, of which linear and semidefinite programming are special cases. We are classifying all self–scaled barriers over irreducible symmetric cones and show that these functions are merely homothetic transformations of the universal barrier function. Together with a decomposition theorem for self–scaled barriers this concludes the algebraic classification theory of these functions. After introducing the reader to the concepts relevant to the problem and tracing the history of the subject, we start by deriving our result from first principles in the important special case of semidefinite programming. We then generalise these arguments to irreducible symmetric cones by invoking results from the theory of Euclidean Jordan algebras.

Key Words

Semidefinite programming, self–scaled barrier functions, interior–point methods, symmetric cones, Euclidean Jordan algebras.

AMS 1991 Subject Classification

Primary 90C25, 52A41, 90C60. Secondary 90C05, 90C20.

Contact Information and Credits

Raphael Hauser, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, England. *rah48@damtp.cam.ac.uk*. Research supported in part by the Norwegian Research Council through project No. 127582/410 “Synode II”, by 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.

Yongdo Lim, Department of Mathematics, Kyungpook National University, Taegu 702–701, Korea. *ylim@knu.ac.kr*. Research supported in part by the Basic Research Program of the Korea Science and Engineering Foundation through project No. 2000-1-10100-007-3.

1. Introduction

Self–scaled barriers are a special class of self–concordant barrier functions [16] introduced by Nesterov and Todd [17] for the purpose of extending long–step primal–dual symmetric interior–point methods from linear and semidefinite programming to more general convex optimization problems [17, 18], see Definition 1 below.

The domain of definition of a self–scaled barrier F is an proper open convex cone Ω lying in a real Euclidean space $\big(V, \langle\cdot,\cdot\rangle\big)$. By abuse of language one often refers to F as a self–scaled barrier for the topological closure $\bar{\Omega}$ of Ω. Not every proper open convex cone Ω allows a self–scaled barrier, and for those who do $\bar{\Omega}$ is called a *self–scaled cone* in the terminology of Nesterov and Todd [17]. Güler [3] found that the family of interiors of self–scaled cones is identical to the set of *symmetric cones* studied in the theory of Euclidean Jordan algebras. Due to this discovery, Jordan algebra theory became an important analytic tool in the theory of semidefinite programming and its natural generalisation, self–scaled programming.

Self–scaled barriers were studied by Nesterov–Todd [17, 18], Güler [3], Güler–Tunçel [21] (see p. 124 and related material), Hauser [5] and others. Though an axiomatic theory of these functions exists, the only known examples are trivially related to the characteristic function of the cone Ω,

$$\varphi_\Omega(x) := \int_{\Omega^\sharp} \mathrm{e}^{-\langle x,s\rangle}\, ds, \tag{1.1}$$

where $\Omega^\sharp := \{s \in V : \langle x,s\rangle > 0\ \forall x \in \Omega\}$ is the polar of Ω with respect to $\langle\cdot,\cdot\rangle$. This was first discovered by Güler who showed that the universal barrier U [16] of a symmetric cone is self–scaled and is a homothetic transformation $U = c_1 \ln \varphi_\Omega + c_2$ of the characteristic function φ_Ω, where $c_1 \geq 1$ and c_2 are constants. More generally, every symmetric cone Ω has a decomposition, unique up to indexing, into a direct sum of irreducible symmetric cones

$$\Omega = \Omega_1 \oplus \cdots \oplus \Omega_m,$$

where the Ω_i lie in subspaces $V_i \subseteq V$ decomposing V into a direct sum $V = \oplus_{i=1}^m V_i$. The *irreducible summands* Ω_i can be classified into five different types, see [2] and references therein. All known self–scaled barrier functions for Ω are of the form

$$H = c_0 + \bigoplus_{i=1}^m c_i \ln \varphi_{\Omega_i}, \tag{1.2}$$

where $c_1, \dots, c_m \geq 1$. We use the direct sum notation for H in Equation (1.2) and elsewhere to indicate that each $x = \oplus_{i=1}^m x_i \in \oplus_{i=1}^m \Omega_i = \Omega$ is mapped to $H(x) = c_0 + \sum_{i=1}^m c_i \ln \varphi_{\Omega_i}(x_i)$. It is also well–known that each function of the form (1.2) is a self–scaled barrier for Ω.

The natural question arises as to whether *all* self–scaled barrier functions are of the form (1.2). Early dents into this question were made by Güler and Tunçel when considering *invariant barriers*, see [21] page 124 and related material. In a chapter of his thesis [5] and in a subsequent report [6], Hauser showed that any self–scaled barrier H over a symmetric cone Ω decomposes into a direct sum $H = \oplus_{i=1}^m H_i$ of self–scaled barriers H_i over the irreducible components Ω_i, and that any *isotropic*, i.e., rotationally invariant, self–scaled barrier H_i on Ω_i is of the form $c_1 \ln \varphi_{\Omega_i} + c_2$ with $c_1 \geq 1$. Hauser also observed that any self–scaled barrier H on Ω is invariant under a rich class of rotations of Ω, i.e., elements of $\mathrm{O}(\Omega)$, see (1.3) below, where these particular rotations are defined in terms of the Hessians of H, see Lemma 2.2.19 [5]. Suspecting that in the case where Ω is irreducible this family of rotations is rich enough to generate all of $\mathrm{O}(\Omega)$, Hauser [5, 6] conjectured that all self–scaled barriers on irreducible symmetric cones are isotropic. This conjecture, whose correctness to prove is the primary objective of this paper, shows that all self–scaled barriers are indeed of the form (1.2) and concludes their algebraic classification.

In a second report [7] Hauser showed the correctness of the isotropy conjecture for the special case of the cone of positive semidefinite symmetric matrices. The proof follows exactly the path suggested by Lemma 2.2.19 [5], and as outlined above. Shortly after Hauser's report [7] was announced, Lim [15] generalised Hauser's arguments to general irreducible symmetric cones and settled the isotropy conjecture. This report is a joint publication consisting of a revision of Hauser's report [7] and of Lim's generalisation [15].

Subsequently, both Schmieta [22] and Güler [4] independently of each other and independently of Lim also proved the isotropy conjecture. Schmieta's report [22] was the first publication where the full classification result became available. Güler's approach [4] was later incorporated in a joint publication with Hauser [8] which started as a major revision of the report [6].

It is interesting to note that, though the approaches of Lim, Güler and Schmieta differ in important details, all three involve two key mechanisms: The so–called *fundamental formula* on the one hand, see (3.17) below, and Koecher's Theorem 4.9 (b) [12] on the other hand. Already Hauser's approach [7] to solving the special case of the positive semidefinite cone was based on fundamentally the same ideas, as his Proposition 3.3 was essentially an independent rediscovery of Koecher's theorem in this particular case (c.f. Corollary 4.3 [7]).

The main part of this paper is organised in two sections addressing different communities: Section 2 treats the case of the positive semidefinite cone only. Readers interested in semidefinite programming and lacking a background in Jordan algebra theory will find it easy to read this section, the results being derived from first principles. All the qualitative features of the general approach already appear in the restricted framework, and it is possible to understand some of the essential ideas behind Koecher's theorem by reading the proof of Proposition 6. Section 3 on the other hand treats the general case and addresses primarily readers with a background in Jordan algebra theory. It is possible to understand this section without prior lecture of Section 2.

We conclude this section by giving the essential definitions and identities that form the basis of the theory of self–scaled barriers. Recall that we introduced the notation Ω and V above. The set of vector space automorphisms of V that leave Ω invariant is called the *symmetry group* of Ω, and we denote it by

$$G(\Omega) := \{\theta \in \mathrm{GL}(V) : \theta\Omega = \Omega\}.$$

The inner product on V defines a notion of adjoint of an endomorphism and hence an orthogonal group $\mathrm{O}(V) = \{\theta \in \mathrm{GL}(V) : \theta^* = \theta^{-1}\}$. The subgroup

$$\mathrm{O}(\Omega) := G(\Omega) \cap \mathrm{O}(V) \tag{1.3}$$

is called the *orthogonal group* of Ω.

Definition 1 ([17])**.** A ν–self–concordant logarithmically homogeneous [16] barrier functional $H \in \mathcal{C}^3(\Omega, \mathbb{R})$ is said to be *self–scaled* if the following conditions are satisfied:

(a) $H''(w)x \in \Omega^\sharp$ for all $x, w \in \Omega$ and
(b) $H_\sharp\big(H''(w)x\big) = H(x) - 2H(w) - \nu$ for all $x, w \in \Omega$.

The function $H_\sharp : s \mapsto \max\big\{-\langle x, s\rangle - H(x) : x \in \Omega\big\}$ is a self–scaled barrier defined on $\Omega^\sharp$ [17]. It is assumed in the definition of a self–concordant function, see [16], that the Hessians $H''(x)$ are non–singular for all $x \in \Omega$. The next theorem is a compilation of several separate results of Nesterov and Todd [17]:

Theorem 2 ([17])**.** *Let $H \in \mathcal{C}^3(\Omega, \mathbb{R})$ be self–scaled and $x \in \Omega, s \in \Omega^\sharp$. Then there exists a unique* scaling point $w_H(x, s) \in \Omega$ *such that $s = H''\big(w(x, s)\big)x$. Furthermore, the following properties hold:*

(a) $H''(w) \in \mathrm{Iso}\big(\Omega, \Omega^\sharp\big) \quad \forall w \in \Omega,$
(b) $H''(x) = H''\big(w(x, s)\big) \circ H''_\sharp(s) \circ H''\big(w(x, s)\big).$ □

It is customary to change the inner product $\langle \cdot, \cdot\rangle$ to $\langle \cdot, \cdot\rangle_f := \langle H''(f)\cdot, \cdot\rangle$ where f is a fixed element in Ω. We will always assume that $\langle \cdot, \cdot\rangle$ is already of this kind, i.e., that there exists an element $f \in \Omega$ such that $H''(f) = \mathrm{id}_\Omega$ is the identity when the Hessian is computed with respect to this inner product. Under this assumption it is easy to show that $\Omega^\sharp = \Omega$ and $H_\sharp = H + \mathrm{cst}$. Hence, in this framework we can reformulate Parts (a) and (b) of Theorem 2 as follows:

$$H''(w) \in G(\Omega), \tag{1.4}$$
$$H''(x) = H''\big(w_H(x, s)\big) \circ H''(s) \circ H''\big(w_H(x, s)\big). \tag{1.5}$$

Theorem 2 reveals that Ω can allow a self–scaled barrier only if it is a *symmetric cone*, c.f. Section 3. As mentioned above, this fact was first observed by Güler [3] who showed that the relation also holds in the opposite direction.

Equation (1.5) is a reformulation of an identity which is called *fundamental formula* in Jordan algebra theory, see Equation (3.17) below. This identity is one of the keys to proving the isotropy conjecture, as it allows to express all the Hessians of H in terms of the Hessian $H''(e)$ at a single element $e \in \Omega$ and of the Hessians of the standard logarithmic barrier function

$$F(x) = \ln \varphi_\Omega(x), \tag{1.6}$$

see Equation (1.1), which is self–scaled [3]. Rothaus [20] proved the following result that will be important for our purposes:

Theorem 3 ([20])**.** *For every $\theta \in G(\Omega)$ there exists a unique $\omega \in \mathrm{O}(\Omega)$ and a unique $w \in \mathrm{int}(\Omega)$ such that θ has the* polar decomposition *$\theta = \omega \circ F''(w)$.* □

Since $\mathrm{O}(\Omega) \subset \mathrm{O}(V)$ and $F''(w)$ is a self–adjoint positive definite automorphism of V, Rothaus' polar decomposition is identical to Cartan's polar decomposition. Theorem 3 shows the non–trivial fact that both factors lie in $G(\Omega)$. The uniqueness of Cartan's polar decomposition trivially implies the following lemma which will be useful in later sections:

Lemma 4. *The set of self–adjoint positive definite automorphisms of V that preserve Ω coincides with $\{F''(w) : w \in \Omega\}$.* □

2. Self–Scaled Barriers for Semidefinite Programming

This section is limited to semidefinite programming and provides readers who are unfamiliar with Jordan algebra terminology access to the main ideas behind the mechanism that forces self–scaled barriers to be essentially unique.

In this framework it is customary to write variable names with capitalised letters. V is the space $\mathrm{S}ym(n,\mathbb{R})$ of symmetric $n \times n$ matrices endowed with the trace inner product $\langle X, S\rangle = \mathrm{tr}\big(X^* S\big)$ which corresponds to the Frobenius norm. Ω is the cone $\mathrm{S}ym(n,\mathbb{R})^+$ of positive definite symmetric $n \times n$ matrices. The following identities hold for the standard logarithmic barrier function $F(X) = -\ln\det(X)$:

$$F''(X)A = X^{-1}AX^{-1} \qquad \forall A \in \mathrm{S}ym(n,\mathbb{R}) \quad \text{and} \tag{2.7}$$

$$W_F(X,S) = S^{-1/2}\big(S^{1/2}XS^{1/2}\big)^{1/2}S^{-1/2}. \tag{2.8}$$

Let us assume that H is an arbitrary self–scaled barrier function for $\mathrm{S}ym(n,\mathbb{R})^+$. Applying Equation (1.5) to both H and F, we can derive a series of expressions that will allow us to relate Hessians of H to Hessians of F. Since $H''(X)$ is a self–adjoint positive definite automorphism of $\mathrm{S}ym(n,\mathbb{R})^+$ it follows from Lemma 4 that there exists a well–defined mapping $\Upsilon : \mathrm{S}ym(n,\mathbb{R})^+ \to \mathrm{S}ym(n,\mathbb{R})^+$ such that

$$H''(X) = F''\big(\Upsilon(X)\big) \tag{2.9}$$

for all $X \in \mathrm{S}ym(n,\mathbb{R})^+$. We claim that Υ is a scalar function, i.e., there exists a number $\lambda > 0$ such that $\Upsilon = \lambda \cdot \mathrm{id}_{\mathrm{S}ym(n,\mathbb{R})}$. The proof of this claim is going to occupy us until Corollary 7.

Let $W_F = W_F(X,S)$ and $W_H = W_H(X,S)$ denote the scaling points of $X, S \in \mathrm{S}ym(n,\mathbb{R})^+$ defined by F and H, see Theorem 2. Equation (2.9) implies that $F''\big(\Upsilon(W_H)\big)X = S = F''\big(W_F\big)X$, and it follows from the uniqueness part of Theorem 2 that $\Upsilon\big(W_H\big) = W_F$. Therefore, Equation (1.5) applied to H shows that

$$H''(X) = F''(W_F) \circ H''(S) \circ F''(W_F), \quad \text{and} \tag{2.10}$$

$$\Upsilon(X) = W_F\Upsilon(S)W_F. \tag{2.11}$$

Note that $F''(\mathrm{I}) = \mathrm{id}_{\mathrm{S}ym(n,\mathbb{R})}$. Therefore, Equation (1.5) applied to F and $S = \mathrm{I}$ shows that $F''\big(W_F(X,\mathrm{I})\big) = F''(X)^{-1/2}$. Using this fact in conjunction with Equations (2.7) and (2.10) we get

$$\Upsilon(X) = X^{1/2}\Upsilon(\mathrm{I})X^{1/2} \quad \forall X \in \mathrm{S}ym(n,\mathbb{R})^+. \tag{2.12}$$

Therefore,

$$\begin{aligned} X^{1/2}\Upsilon(I)X^{1/2} &\overset{(2.12),(2.11)}{=} W_F(X,S)\Upsilon(S)W_F(X,S) \\ &\overset{(2.12)}{=} W_F(X,S)S^{1/2}\Upsilon(I)S^{1/2}W_F(X,S), \end{aligned}$$

and by virtue of (2.8) this implies that

$$\Upsilon(I) = X^{-1/2}S^{-1/2}\big(S^{1/2}XS^{1/2}\big)^{1/2}\Upsilon(I)S^{-1/2}\big(S^{1/2}XS^{1/2}\big)^{1/2}X^{-1/2} \tag{2.13}$$

for all $X, S \in \mathrm{S}ym(n,\mathbb{R})^+$. Clearly, this condition is equivalent to

$$\Upsilon(I) = N^{-1}\big(NN^{\mathrm{T}}\big)^{1/2}\Upsilon(I)\big(NN^{\mathrm{T}}\big)^{1/2}N^{-1} \quad \forall N \in \mathcal{K}, \tag{2.14}$$

where $\mathcal{K} := \big\{XS : X, S \in \mathrm{S}ym(n,\mathbb{R})^+\big\}$ is the set of $n \times n$ matrices that can be written as the product of two symmetric positive definite matrices. The following characterisation of this set is due to Mike Todd [23]:

Lemma 5. *$\mathcal{K}$ coincides with the set of non–defective $n \times n$ matrices with real coefficients, all of whose eigenvalues are strictly positive real numbers.*

Proof. If $N = XS$, then

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

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 $N = \big(PP^{\mathrm{T}}\big)\big(P^{-T}DP^{-1}\big) =: XS$. □

Note that $N^{-1}(NN^{\mathrm{T}})^{1/2} \in \mathrm{SO}(n)$ for all $N \in \mathcal{K}$. In the next proposition we characterise the closed subgroup of $\mathrm{SO}(n)$ generated by matrices of this form. This constitutes the mathematical core mechanism of our proof and is a close relative of Koecher's Theorem 4.9 (b) [12].

Proposition 6. *The closed subgroup of* $\mathrm{SO}(n)$ *generated by the set of orthogonal matrices of the form $N^{-1}(NN^{\mathrm{T}})^{1/2}$ with $N \in \mathcal{K}$ coincides with* $\mathrm{SO}(n)$.

Proof. Let this closed subgroup be denoted by $\mathcal{G}$, and let $\mathfrak{g}$ be its Lie Algebra. Since $\mathrm{SO}(n)$ is connected, it suffices to show that $\mathfrak{g} = \mathfrak{so}(n)$, or in other words that the tangent space of the manifold $\mathcal{G}$ at the point $\mathrm{I} \in \mathcal{G}$ coincides with the set of $n \times n$ skew–symmetric matrices over the reals. In fact, use of the exponential mapping $\exp : \mathfrak{so}(n) \to \mathrm{SO}(n)$ shows that $\mathcal{G}$ and $\mathrm{SO}(n)$ share a neighbourhood $\mathcal{V}$ of I, and parallel transport by left trivialisation shows that $\mathcal{G}$ and $\mathrm{SO}(n)$ share the neighbourhood $g\mathcal{V}$ of any element $g \in \mathcal{G}$. Therefore, $\mathcal{G}$ is both open and closed as a subset of $\mathrm{SO}(n)$, and since $\mathrm{SO}(n)$ is connected, the result follows.

It remains to show that $\mathfrak{g} = \mathfrak{so}(n)$: Let $\Delta \in \mathcal{K}$ have eigenvalues $\lambda_1, \dots \lambda_n > 0$. Then $\Delta(t) := I + t\Delta \in \mathcal{K}$ 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

$$\Delta(t)^{-1} = I - t\Delta + O(t^2). \tag{2.15}$$

Upon taking squares on both sides of the ansatz

$$(\Delta(t)\Delta(t)^{\mathrm{T}})^{1/2} = I + tD + O(t^2) \tag{2.16}$$

we get $I + t(\Delta + \Delta^{\mathrm{T}}) + O(t^2) = I + 2tD + O(t^2)$, and hence $D = 1/2(\Delta + \Delta^{\mathrm{T}})$. Equations (2.15) and (2.16) thus yield the identity

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

and this shows that $\Delta^{\mathrm{T}} - \Delta \in \mathfrak{g}$ for all $\Delta \in \mathcal{K}$. Clearly, we have $\Delta^{\mathrm{T}} - \Delta \in \mathfrak{so}(n)$ as expected. Now, 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, $(P^{-1}DP)^{\mathrm{T}} - P^{-1}DP = \left(\begin{smallmatrix}0 & 1\\ -1 & 0\end{smallmatrix}\right)$. Let P_{ij} be the permutation matrix that permutes the i^{th} and j^{th} variables, and let $E^-_{ij} := e_ie_j^{\mathrm{T}} - e_je_i^{\mathrm{T}}$, where e_i is the i–th element in the canonical basis of $\mathbb{R}^n$. Then consider

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

Clearly, $\Delta \in \mathcal{K}$ and $\Delta^{\mathrm{T}} - \Delta = E^-_{ij}$. But since $\big\{E^-_{ij} : i, j \in \{1, \ldots, n\}\big\}$ forms a basis of $\mathfrak{so}$ we find that the elements of $\mathfrak{g}$ span this whole space. This shows the claim. □

Corollary 7. *There exists a positive constant* $\lambda > 0$ *such that* $H''(X) = \lambda F''(X)$ *for all* $X \in \mathrm{S}ym(n, \mathbb{R})^+$.

Proof. The invariance property (2.14) is clearly preserved when taking compositions and limits. Hence, Lemma 5 implies that the symmetric positive definite matrix $\Upsilon(\mathrm{I})$ satisfies the condition $\Upsilon(I) = \Omega\Upsilon(\mathrm{I})\Omega^{\mathrm{T}}$ for all $\Omega \in \mathrm{SO}(n)$. It is a trivial matter to prove that this forces $\Upsilon(\mathrm{I})$ to be a scalar, and the result now follows from (2.12). □

Theorem 8. *If* H *is a self–scaled barrier functional for the cone of symmetric positive semidefinite* $n \times n$ *matrices then there exist constants* $c_1 > 0$ *and* $c_0 \in \mathbb{R}$ *such that*

$$H : X \mapsto c_0 - c_1 \ln\det(X) \quad \forall X \in \mathrm{S}ym(n, \mathbb{R})^+.$$

Proof. It follows from Corollary 7 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 $\varphi \in \mathrm{S}ym(n, \mathbb{R})^\sharp$ is a linear form on $\mathrm{S}ym(n, \mathbb{R})$, i.e. there exists a symmetric matrix $Y \in \mathrm{S}ym(n, \mathbb{R})$ such that $\varphi : X \mapsto \mathrm{tr}(Y^{\mathrm{T}}X)$ for all $X \in \mathrm{S}ym(n, \mathbb{R})$. 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 $B''(x)^{-1}[-B'(x)]$ at $x \in D$, measured in the Riemannian metric $\|.\|_x$ defined by $B''(x)$ be uniformly bounded by $\nu^{1/2}$, see e.g. [16, 17, 19], i.e.,

$$\begin{aligned}\|B''(x)^{-1}[-B'(x)]\|_x^2 &:= \big\langle B''(x)\big[-B''(x)^{-1}[B'(x)]\big], -B''(x)^{-1}B'(x)\big\rangle\\ &= \big\langle B'(x), \big(B''(x)\big)^{-1}[B'(x)]\big\rangle \le \nu.\end{aligned}$$

In particular, in the case of H this means that

$$\begin{aligned}\nu &\ge \|H'(X)\|_X^2 = \|Y - \lambda X^{-1}\|_X^2 \overset{C.S.}{\ge} \big(\|Y\|_X - \|\lambda X^{-1}\|_X\big)^2\\ &\overset{\mathrm{C}or.7}{=} \Big(\mathrm{tr}(\lambda^{-1}XYX \cdot Y^{\mathrm{T}})^{1/2} - \big(\lambda^{-1}X(\lambda X^{-1})X \cdot (\lambda X^{-1})^{\mathrm{T}}\big)^{1/2}\Big)^2\\ &= \lambda^{-1}\Big(\big(\mathrm{tr}\big[(XY)^2\big]\big)^{1/2} - \lambda n^{1/2}\Big)^2\end{aligned}$$

for all $X \in \mathrm{S}ym(n, \mathbb{R})^+$. But clearly, this implies that $Y = 0$. □

Remark 9. It is possible to prove Theorem 8 by invoking Lemma 2.2.19 [5], see [7] for details. This approach is interesting, since it follows the path traced by the intuition that first led to the isotropy conjecture. However, the proof we gave above fits better into the mainstream literature and is therefore easier to understand.

3. Self–Scaled Barriers for Irreducible Symmetric Cones

An open convex cone Ω in a real Euclidean space V that is self–dual with respect to the given inner product and is homogeneous in the sense that the group

$$G(\Omega) := \{g \in \mathrm{GL}(V) : g(\Omega) = \Omega\}$$

acts transitively on Ω is called a *symmetric cone*. The theory of symmetric cones is closely tied to that of Euclidean Jordan algebras. We recall certain basic notions and well–known facts concerning Jordan algebras from the book [2] by J. Faraut and A. Korányi.

A *Jordan algebra* V over the field $\mathbb{R}$ or $\mathbb{C}$ is a commutative algebra satisfying $x^2(xy) = x(x^2y)$ for all $x, y \in V$. We also assume the existence of a multiplicative identity e. Denote by L the left translation $L(x)y = xy$, and P by the quadratic representation $P(x) = 2L(x)^2 - L(x^2)$ for $x \in V$. An alternate statement of the Jordan algebra law is $(xy)x^2 = x(yx^2)$, a weak associativity condition that is strong enough to ensure that the subalgebra generated by $\{e, x\}$ in V is associative. An element $x \in V$ is said to be *invertible* if there exists an element y in the subalgebra generated by x and e such that $xy = e$. It is known that an element x in V is invertible if and only if $P(x)$ is invertible. In this case, $P(x)^{-1} = P(x^{-1})$. If x and y are invertible, then $P(x)y$ is invertible and $(P(x)y)^{-1} = P(x^{-1})y^{-1}$. Furthermore, the fundamental formula

$$P(P(x)y) = P(x)P(y)P(x) \tag{3.17}$$

holds for any elements $x, y \in V$, see Proposition II.3.2 (iii) [2].

A *Euclidean* Jordan algebra is a finite–dimensional real Jordan algebra V equipped with an *associative* inner product $\langle\cdot|\cdot\rangle$, i.e., satisfying $\langle xy|z\rangle = \langle y|xz\rangle$ for all $x, y, z \in V$. The space $\mathrm{S}ym(n, \mathbb{R})$ of $n \times n$ real symmetric matrices is a Euclidean Jordan algebra with Jordan product $(1/2)(XY + YX)$ and inner product $\langle X, Y\rangle = \mathrm{tr}(XY)$. The spectral theorem (see Theorem III.1.2 of [2]) of a Euclidean Jordan algebra V states that for $x \in V$ there exist a Jordan frame, i.e., a complete system of orthogonal primitive idempotents $c_1, \ldots, c_r$, where r is the rank of V, and real numbers $\lambda_1, \ldots, \lambda_r$, the eigenvalues of x, such that $x = \sum_{i=1}^{r} \lambda_i c_i$. Due to the the power associative property $x^p \cdot x^q = x^{p+q}$, see [2], the exponential map $\exp : V \to V$

$$\exp : x \mapsto \sum_{n=0}^{\infty} x^n/n!$$

is well defined. Likewise as for the special case $V = \mathrm{S}ym(n, \mathbb{R})$, the Jordan algebra exponential map is a bijection between its domain of definition V and its image $\Omega := \exp V$. In fact, Ω coincides with the interior of the set of square elements of V, and this is equal to the set of invertible squares of V. A fundamental theorem of Euclidean Jordan algebras asserts that (i) Ω is a symmetric cone, and (ii) every symmetric cone in a real Euclidean space arises in this way. In the case of the Jordan algebra $V = \mathrm{S}ym(n, \mathbb{R})$ the Jordan algebra exponential is simply the matrix exponential map, and hence the corresponding symmetric cone is $\Omega = Sym(n,\mathbb{R})^+$, the open convex cone of positive definite $n \times n$ matrices. Irreducible symmetric cones have been completely classified, and the remaining cases consist of (i) the cones of positive definite Hermitian and Hermitian quaternion $n \times n$ matrices, (ii) the Lorentzian cones, and (iii) a 27–dimensional exceptional cone. General symmetric cones are Cartesian products of these. The connected component $\mathrm{Aut}(V)_\circ$ of the identity id_V in the Jordan algebra automorphism group $\mathrm{Aut}(V)$ is a subgroup of $\mathrm{O}(\Omega)$. Ω is irreducible if and only if V is simple, and in this case we have $\mathrm{Aut}(V)_\circ = \mathrm{O}(\Omega)$. For all of these statements, see [2] and the references therein.

In the special case $V = Sym(n,\mathbb{R})$ the general formula $P(x)y = 2x(xy) - x^2y$ reduces to $P(X)Y = XYX$, where the multiplication in the right hand side of this equation is the usual matrix multiplication, not the Jordan multiplication. A key tool for generalising $Sym(n,\mathbb{R})$ to arbitrary Euclidean Jordan algebras is to consistently replace expressions of the form XYX by $P(x)y$. Throughout this section we will assume that V is a simple Euclidean Jordan algebra with the associative inner product $\langle x|y\rangle = \mathrm{tr}(xy)$ and that Ω is the symmetric cone associated with V.

The symmetric cone Ω carries a $G(\Omega)$–invariant Riemannian metric defined by

$$\gamma_x(u,v) = \langle P(x^{-1})u|v\rangle, x \in \Omega, u, v \in V$$

for which the Jordan inversion $x \to x^{-1}$ on Ω is an involutive isometry fixing e. The curve $t \mapsto P(a^{1/2})(P(a^{-1/2})b)^t$ is the unique geodesic that joins a to b in Ω, and the Riemannian distance $\delta(a,b)$ is given by $\delta(a,b) = \left(\sum_{i=1}^n \log^2 \lambda_i\right)^{1/2}$, where the λ_i are the eigenvalues of $P(a^{-1/2})b$. The *geometric mean* $a\#b$ of two elements $a, b \in \Omega$ is defined by

$$a\#b = P(a^{1/2})(P(a^{-1/2})b)^{1/2}.$$

This is the unique midpoint – or geodesic middle – of a and b for the Riemannian distance δ. The metric δ is known as a Bruhat–Tits metric, i.e., a complete metric satisfying the semi–parallelogram law, with midpoint $a\#b$. See [13] for more details. If $V = Sym(n,\mathbb{R})$ then the geometric mean $A\#B$ of positive definite matrices A and B is given by $A\#B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$. The following basic properties of geometric means will be useful for our purpose:

Proposition 10 ([14])**.** *Let a and b be elements of Ω. Then*

(a) *$a\#b$ is a unique solution belonging to Ω of the quadratic equation $P(x)a^{-1} = b$.*
(b) *(The commutativity property) $a\#b = b\#a$.*
(c) *(The inversion property) $(a\#b)^{-1} = a^{-1}\#b^{-1}$.*
(d) $P(a\#b) = P(a)\#P(b) = P(a^{1/2})\left(P(a^{-1/2})P(b)P(a^{-1/2})\right)^{1/2} P(a^{1/2})$.
(e) *(The transformation property) $g(a\#b) = g(a)\#g(b)$ for all $g \in G(\Omega)$.*

Let $F(x) = -\ln \det(x)$ be the standard logarithmic barrier functional on the symmetric cone Ω, where det is the determinant function of the Jordan algebra V, see [2]. Then one can see that $F'(x) = -x^{-1}$ and the Hessian of F is given by

$$F''(x) = P(x^{-1}).$$

Proposition 10 implies that the geometric mean $a\#b^{-1}$ is the scaling point of a and $b \in \Omega$ defined by F. Indeed, $F''(a\#b^{-1})a = P\big((a\#b^{-1})^{-1}\big)a = P(a^{-1}\#b)a = b$, that is, $w_F(a,b) = a\#b^{-1}$.

Lemma 11. *Let Ω be an irreducible symmetric cone and $\alpha : \Omega \to \Omega$ a function such that*

$$x^{-1}\#y = \alpha(x)^{-1}\#\alpha(y) \tag{3.18}$$

for all $x, y \in \Omega$. Then $\alpha = \lambda \cdot \mathrm{id}_\Omega$ for some positive real number λ.

Proof. Upon exchanging the roles of x and y, Proposition 10(a) implies that Condition (3.18) is equivalent to

$$\alpha(x) = P(y^{-1}\#x)\alpha(y). \tag{3.19}$$

Setting $y = e$ and using $e^{-1}\#x = x^{1/2}$ in (3.19), we get

$$\alpha(x) = P(x^{1/2})\alpha(e) \tag{3.20}$$

for all $x \in \Omega$. Let us show that $k(\alpha(e)) = \alpha(e)$ for all $k \in \mathrm{Aut}(V)_\circ$. Applying (3.19) and (3.20) both to x and y we get

$$P(x^{1/2})\alpha(e) = P(y^{-1}\#x)\alpha(y) = P(y^{-1}\#x)\big(P(y^{1/2})\alpha(e)\big),$$

and hence we obtain the identity $\alpha(e) = \big(P(x^{-1/2})P(y^{-1}\#x)P(y^{1/2})\big)\alpha(e)$ for all $x, y \in \Omega$, which generalises (2.13). Set

$$K := \{P(x^{-1/2})P(y^{-1}\#x)P(y^{1/2}) : x, y \in \Omega\}.$$

It follows from the definition of the geometric mean and from the fundamental formula that

$$P(a\#b) = P(a^{1/2})\Big(P\big(P(a^{-1/2})b\big)\Big)^{1/2}P(a^{1/2}).$$

Together with Proposition 10 (b) this implies

$$\begin{aligned}
P(x^{-1/2})P(y^{-1}\#x)P(y^{1/2}) &= P(x^{-1/2})P(x\#y^{-1})P(y^{1/2}) \\
&= P(x^{-1/2})\Big(P(x^{1/2})P\big(P(x^{-1/2})y^{-1}\big)^{1/2}P(x^{1/2})\Big)P(y^{1/2}) \\
&= P\big(P(x^{1/2})y\big)^{-1/2}P(x^{1/2})P(y^{1/2}).
\end{aligned}$$

Therefore, the set K can be written as $K = \big\{P\big(P(x)y^2\big)^{-1/2}P(x)P(y) | x, y \in \Omega\big\}$. By Koecher's Theorem 4.9 (b) [12], K generates $\mathrm{Aut}(V)_\circ$. This implies that the point $\alpha(e)$ is fixed by all Jordan automorphisms $k \in \mathrm{Aut}(V)_\circ$.

Finally, Corollary IV.2.7 [2] (in which the assumption of irreducibility for Ω is essential) says that the group $\mathrm{Aut}(V)_\circ$ acts transitively on the set of primitive idempotents of V. The spectral theorem applied to $\alpha(e)$ therefore implies that $\alpha(e) = \lambda e$ for some positive real number λ. Together with (3.20) this implies that $\alpha(x) = P(x^{1/2})(\lambda e) = \lambda P(x^{1/2})(e) = \lambda x$ for all $x \in \Omega$. □

Corollary 12. *Let H be an arbitrary self–scaled barrier for the irreducible symmetric cone Ω. Then there exists a positive constant λ such that $H''(x) = \lambda \cdot F''(x)$ for all $x \in \Omega$.*

Proof. Since the Hessian $H''(x)$ are positive definite cone automorphisms, Lemma 4 implies that there exists a well–defined function $\Upsilon : \Omega \to \Omega$ such that

$$H''(x) = P(\Upsilon(x)^{-1}). \tag{3.21}$$

Since H is self-scaled, we have

$$\begin{aligned}P(\Upsilon(x)^{-1}) &\overset{(3.21),(1.5)}{=} H''(w_H) \circ H''(y) \circ H''(w_H)\\ &= P(\Upsilon(w_H)^{-1}) \circ P(\Upsilon(y)^{-1}) \circ P(\Upsilon(w_H)^{-1})\\ &\overset{(3.17)}{=} P\Big(P(\Upsilon(w_H)^{-1})\Upsilon(y)^{-1}\Big)\end{aligned}$$

for all $x, y \in \Omega$, where $w_H = w_H(x, y)$ denotes the scaling point of x and y for the self–scaled barrier H. The quadratic representation P is injective on Ω, see Lemma 2.3 [14]. Therefore, the identity above shows that

$$\Upsilon(x)^{-1} = P\big(\Upsilon(w_H)^{-1}\big)\Upsilon(y)^{-1}$$

for all $x, y \in \Omega$. By Proposition 10, we have

$$\Upsilon(w_H)^{-1} = \Upsilon(y)\#\Upsilon(x)^{-1} = \Upsilon(x)^{-1}\#\Upsilon(y) \tag{3.22}$$

for all $x, y \in \Omega$. Now, $y = H''(w_H)(x) = P(\Upsilon(w_H)^{-1})(x)$ by definition of w_H, and Proposition 10 (a) shows that we have $\Upsilon(w_H)^{-1} = x^{-1}\#y$, which together with (3.22) shows that

$$x^{-1}\#y = \Upsilon(x)^{-1}\#\Upsilon(y)$$

for all $x, y \in \Omega$. The proof is now completed by Lemma 11. $\square$

Theorem 13. *If H is a self–scaled barrier functional for Ω then there exist constants $c_1 > 0$ and $c_0 \in \mathbb{R}$ such that*

$$H : x \to -c_1 \ln\ \det(x) + c_0, \qquad \forall x \in \Omega.$$

Proof. Similar to that of Theorem 8. $\square$

4. Acknowledgments

Both authors would like to thank Professor Leonid Faybusovich for suggesting this collaboration and establishing the contact between them. Raphael Hauser also wishes to express his warmest thanks to Syvert Nørsett for inviting him to NTNU Trondheim, and to Peter and Jonathan Borwein for inviting him to SFU in Vancouver, where part of this research was done.

References

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

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

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

[4] O. Güler, Interior–point methods on symmetric and homogeneous cones. Invited Presentation, *17th International Symposium on Mathematical Programming*, Atlanta, Georgia, 7–11 August, 2000.

[5] R. A. Hauser, *On Search Directions for Self–Scaled Conic Programming*, Ph.D. thesis. School of Operations Research and Industrial Engineering, Cornell University. January 2000.

[6] 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, October 1999.

[7] R. A. Hauser, Self–scaled barriers for semidefinite programming, *Numerical Analysis Report DAMTP 2000/NA02*, Department of Applied Mathematics and Theoretical Physics, Silver Street, Cambridge, England CB3 9EW, March 2000.

[8] R. A. Hauser and O. Güler. *Self–Scaled Barrier Functions on Symmetric Cones and Their Classification. Numerical Analysis Report DAMTP 2001/NA??*, Department of Applied Mathematics and Theoretical Physics, Silver Street, Cambridge, England CB3 9EW, March 2001.

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

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

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

[12] M. Koecher. *Jordan algebras and their applications*. Lectures Notes, Univ. of Minnesota, Minneapolis, 1962.

[13] J. D. Lawson and Y. Lim. The geometric mean, matrices, metrics, and more. *Amer. Math. Monthly*, to appear.

[14] Y. Lim. Geometric means on symmetric cones. *Arch. Math.*, 75: 39–45, 2000.

[15] Y. Lim. Private communication to L. Faybusovich.

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

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

[21] 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.

[22] S. H. Schmieta. Complete classification of self–scaled barrier functions. Technical Report, Department of IEOR, Columbia University, New York, NY 10027, July 2000.

[23] M. J. Todd. Email message to the interior–point mailing list, September 1999.

[24] 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.