UNIVERSITY OF CAMBRIDGE

Numerical Analysis Reports

SELF–SCALED BARRIER FUNCTIONS ON SYMMETRIC CONES AND THEIR CLASSIFICATION

Raphael Hauser and Osman Güler

DAMTP 2001/NA03
March 2001

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

SELF–SCALED BARRIER FUNCTIONS ON SYMMETRIC CONES AND THEIR CLASSIFICATION

Raphael A. Hauser *, Osman Güler †

March 28, 2001

Abstract

Self–scaled barrier functions on self–scaled cones were introduced through a set of axioms in 1994 by Y. E. Nesterov and M. J. Todd as a tool for the construction of long–step interior point algorithms. This paper provides firm foundation for these objects by exhibiting their symmetry properties, their intimate ties with the symmetry groups of their domains of definition, and subsequently their decomposition into irreducible parts and algebraic classification theory. In a first part we recall the characterisation of the family of self–scaled cones as the set of symmetric cones and develop a primal–dual symmetric viewpoint on self-scaled barriers, results that were first discovered by the second author. We then show in a short, simple proof that any pointed, convex cone decomposes into a direct sum of irreducible components in a unique way, a result which can also be of independent interest. We then show that any self–scaled barrier function decomposes in an essentially unique way into a direct sum of self–scaled barriers defined on the irreducible components of the underlying symmetric cone. Finally, we present a complete algebraic classification of self–scaled barrier functions using the correspondence between symmetric cones and Euclidean Jordan algebras.

Key words. Self–scaled barrier functions, symmetric cones, decomposition of convex cones, Jordan algebras, universal barrier function, and interior–point methods.
Abbreviated title. Self–scaled barrier functions.
AMS(MOS) subject classifications: primary 90C25, 90C60, 52A41; secondary 90C06, 52A40.

* Department of Applied Mathematics and Theoretical Physics (DAMTP), University of Cambridge, Silver Street, Cambridge CB3 9EW, England. E-mail: *rah48@damtp.cam.ac.uk*. Research supported by the Swiss National Science Foundation and the Swiss Academy of Technical Sciences, and by the Engineering and Physical Sciences Research Council of the UK, EPSRC grant GR/M30975.

† Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, Maryland 21250, USA. E-mail: *guler@math.umbc.edu*. Research on this paper was conducted when the author was visiting the Industrial Engineering Department at Bilkent University, Ankara, Turkey. Research partially supported by the National Science Foundation under grant DMS–0075722.

1 Introduction

In recent years a theory of interior–point methods for linear, semidefinite, and second–order cone programming has been developed within the unified framework of *self–scaled conic programming*. The origins of this theory can be traced to the works [14, 15, 4]. The importance of the problems which can be cast in this framework, and the fact that it is possible to develop efficient *long–step* interior–point methods for these problems have contributed to the popularity of the subject in the optimization community and beyond.

In order to facilitate our exposition, we consider the following pair of convex programs in conic duality

$$\begin{array}{llll} \text{(P)} & \inf \langle x, s_0 \rangle & \text{(D)} & \inf \langle x_0, s \rangle \\ & x \in \big(L + x_0\big) \cap K & & s \in \big(L^{\perp} + s_0\big) \cap K^*. \end{array} \tag{1.1}$$

Here E is a finite dimensional Euclidean space equipped with an inner product $\langle \cdot, \cdot \rangle$, L is a linear subspace of E, and $L^{\perp}$ its orthogonal complement. The cone K is a *regular* (closed, convex, pointed, solid) cone, $x_0 \in \operatorname{int}(K)$, and $s_0 \in \operatorname{int}(K^*)$, where K^* is the dual cone

$$K^* := \{s \in E : \langle x, s \rangle \geq 0, \quad \forall x \in K\}. \tag{1.2}$$

Interior–point algorithms can be used to solve (1.1) over *any* regular cone, provided one has a *self–concordant* barrier function $F(x)$ defined over the interior $\operatorname{int}(K)$ of K. The reader may consult the authoritative monographs [13, 17] for a detailed treatment of self–concordant functions and interior–point methods. In a generic self–concordant barrier function, one has control over the behaviour of the Hessians $F''(y)$ only when y lies in the local ball $\big\{y : \langle F''(x)(y - x), y - x \rangle < 1\big\}$, leading to *short–step* interior–point methods. Although these methods have a polynomial running–time guarantee, they tend to be less efficient linear programming solvers in practice than *long–step* interior–point methods. The theoretical basis for this latter type of algorithm is the fact that the self–concordant barrier function

$$x \mapsto -\sum_{i=1}^{n} \ln x_i$$

has additional properties which make it possible to control it in *all* of $\operatorname{int}(K)$.

In [14], Nesterov and Todd isolate two properties of the barrier $-\sum_{i=1}^{n} \ln x_i$ which they identify as being responsible for making the long–step approach succeed in linear programming. They generalise these properties (see (1.4) and (1.5)) and call self–concordant barrier functions satisfying these conditions *self–scaled*. Since these properties also impose certain conditions on the domain of definition of such functions, Nesterov and Todd call the closures of such domains *self–scaled cones*. For convenience, we recall these concepts here:

Definition 1.1. Let $K \subseteq E$ be a regular cone. A self–concordant barrier function $F : \operatorname{int}(K) \to \mathbb{R}$ is called *self–scaled* if $F''(x)$ is non–singular for every $x \in K$, F is logarithmically homogeneous, that is, there exists a constant $\nu > 0$ such that

$$F(tx) = F(x) - \nu \ln t, \qquad \forall x \in \operatorname{int}(K),\ t > 0, \tag{1.3}$$

and if F satisfies the following two properties

$$F''(w)x \in \operatorname{int}(K^*), \qquad \forall x, w \in \operatorname{int}(K), \tag{1.4}$$

$$F_*\big(F''(w)x\big) = F(x) - 2F(w) - \nu, \qquad \forall x, w \in \operatorname{int}(K). \tag{1.5}$$

If K allows such a barrier function, then K is called a *self–scaled cone*.

The *dual barrier* $F_* : \operatorname{int}(K) \to \mathbb{R}$ that appears in the last Axiom (1.5) is defined as $F_*(s) := \sup\{-\langle x, s\rangle - F(x) : x \in \operatorname{int}(K)\}$. Theorem 3.1 in [14] states that (1.4) can be strengthened to

Theorem 1.2. *If $x \in \operatorname{int}(K)$ and $y \in \operatorname{int}(K^*)$, then there exists a unique point $w \in \operatorname{int}(K)$ such that*

$$F''(w)x = y.$$

Moreover, if $w \in \operatorname{int}(K)$ then $F''(w)(K) = K^*$. □

The point w is called the *scaling point* of x and y. The last statement is a consequence of the first part of the theorem and of Equation (3.2) from Nesterov and Todd's paper [14]. We reproduce this formula here for convenience:

$$F''(x) = F''(w)F_*''\big(F''(w)x\big)F''(w). \tag{1.6}$$

See also Lemma 2.3 below, where this formula reappears and where we give a proof of this important identity.

We would like to mention that Rothaus [18], using rather elementary tools, proves a number of results which are useful in Section 3 of this paper. Two *key* results are [18], Theorem 3.12 and Corollary 3.15, p. 205. These results imply Theorem 1.2 for the universal barrier function, a special self–concordant barrier function defined by Nesterov and Nemirovskii [13] which is further discussed below. Theorem 1.2 is an independently discovered extension of Rothaus's result to all self–scaled functions. It can be shown that, when suitably modified, all results of Section III in [18] can be extended to general self–scaled barriers. In particular this is true for Theorem 1.2.

Nesterov and Todd [14, 15] demonstrate that self–scaled barrier functions can indeed be used to develop various long–step interior–point methods for linear optimization over self–scaled cones, in particular for semidefinite programming and for convex quadratic programming with convex quadratic constraints.

Inspired by the paper of Vinberg [20], O. Güler [4] develops the relationship between the *universal barrier function* of Nesterov and Nemirovskii [13] and the *characteristic function* of the cone K,

$$\varphi_K(x) := \int_{K^*} \mathrm{e}^{-\langle x,y\rangle}\, dy, \tag{1.7}$$

introduced in 1957 by Koecher, see [10]. In particular, Güler [4] shows that if K is a regular cone, then the universal barrier function $U(x)$ for K satisfies the equation

$$U(x) = c_1 \ln \varphi_K(x) + c_2, \tag{1.8}$$

for some constants $c_1 > 0$ and c_2.

Through Güler's paper [4] the concepts of *homogeneous cones*, *homogeneous self–dual cones* (or *symmetric cones*), *Euclidean Jordan algebras* and *Siegel domains* as well as the classification theory of symmetric cones and Euclidean Jordan algebras, known to mathematicians since 1960 and 1934 respectively, were first introduced into the in the interior–point literature. The interested reader can find a complete treatment of these classification results in the book of J. Faraut and A. Korányi [1]. See also [10] for a different treatment of some of the same topics. Because of their importance for this paper, we recall some of the concepts mentioned above.

Definition 1.3. Let $K \subseteq E$ be a regular cone. The *automorphism group* of K is the set of all non–singular linear maps $A : E \to E$ that leave K invariant, i.e.,

$$\operatorname{Aut}(K) := \{A \in \operatorname{GL}(E) : A(K) = K\}$$

The cone K is called *homogeneous* if $\mathrm{Aut}(K)$ acts transitively on K, that is, given arbitrary points $x, y \in K$, there exists a map $A \in \mathrm{Aut}(K)$ such that $Ax = y$. The cone K is called *self–dual* if E can be endowed with an inner product such that $K^* = K$ where K^* is defined with respect to this inner product, see (1.2). The cone K is called *homogeneous self–dual* if K is both homogeneous and self–dual.

Homogeneous self–dual cones are also called *symmetric cones* in [1], a terminology which we shall adopt in this paper.

The motivation behind [14, 15] in contrast to [4] is rather different: While the first two papers deal with long–step interior–point methods and regular cones on which such methods can be designed, the latter one deals with the universal barrier function and the symmetry properties of regular cones, both in the group theoretic and the duality theoretic sense. Shortly after the announcement of the paper [14] Güler [3] discovers that the families of self–scaled cones and of symmetric cones are identical, thus establishing a connection between these two previously distinct ideas.

As mentioned earlier, symmetric cones are fully classified in the theory of Euclidean Jordan algebras, see [10, 1] and the references therein. According to this theory, each symmetric cone has a unique decomposition into a direct sum of elementary building blocks, so–called *irreducible* symmetric cones, of which there exist only five types. Three examples of symmetric cones are of particular interest to the optimization community: The non–negative orthant $K = \mathbb{R}^n_+$, the cone $K = \Sigma^+_n$ of $n \times n$ symmetric, positive semidefinite matrices over the real numbers, and the Lorentz cone $K = \left\{ \left(\begin{smallmatrix} \tau \\ x \end{smallmatrix}\right) \in \mathbb{R}^{n+1} : \tau \geq \|x\|_2 \right\}$. The general self–scaled conic optimization problems associated with these cones are respectively linear programming, semidefinite programming and second order cone programming. The latter can be seen as a reformulation of convex quadratic programming with convex quadratic constraints. Considering more general symmetric cones, one can treat linear optimization problems with mixed linear, semidefinite and convex quadratic constraints in a single unified framework.

Motivated by [4] and by the fact that only a small number of examples of self-scaled barrier functions are explicitly known, Hauser develops a partial algebraic classification theory for self–scaled barrier functions in a chapter of his thesis [7] and later announces a report [8] based on these result. Hauser shows that any self–scaled barrier over a symmetric cone K has an essentially unique decomposition into a direct sum of self–scaled barriers defined on the irreducible summands of K. Using this decomposition, he classifies in particular the family of *isotropic* self–scaled barrier functions which are characterised by rotational invariance. The insight gained from a lemma leads Hauser to conjecture that all self–scaled barrier functions defined on *irreducible* symmetric cones must be isotropic. Hauser points out that this conjecture, if true, would settle the classification problem of general self–scaled barrier functions. This conjecture also implies that all self–scaled barriers over irreducible symmetric cones are of the form $c_1 \ln \varphi_K + c_2$, where $c_1 > 0$ and c_2 are constants. It follows from this theory that the full set of self–scaled barrier functions is readily known, and that all of these functions are just minor transformations of the universal barrier function.

In a second report [9], Hauser proves this conjecture in the special case where K is the positive semidefinite cone, i.e., he shows that all self–scaled barrier functions for use in semidefinite programming are isotropic and essentially identical to the standard logarithmic barrier function. Starting from first principles, Hauser shows that the orientation preserving part of the automorphism group of the positive semidefinite symmetric cone is generated by the Hessians of an arbitrary self–scaled barrier function defined on this cone, see [9], Corollary 4.3. Hauser's solution of the isotropy conjecture in this special case relies on Proposition 3.3, which also forms the key mechanism in the proof of Corollary 4.3 in [9].

Hauser's Corollary 4.3 is essentially a rediscovery of a result by Koecher, Theorem 4.9 (b), pp. 88–89 [10], for the special case of the cone of positive semidefinite symmetric matrices. Shortly after the announcement of Hauser's report [9] in March of 2000, Y. Lim [11] settles the general case of the isotropy conjecture by generalising Hauser's proof while refereeing the paper. Subsequently, both O. Güler [5] and S. Schmieta [19] independently of each other and independently of Y. Lim prove the isotropy conjecture in the general irreducible case. It is interesting to note that all three approaches to the general case of the conjecture, as well as Hauser's approach to the special case relevant to semidefinite programming, rely on the same deeper principle provided by Koecher's Theorem 4.9 (b) cited above. See also Remark 5.2 of this article.

The present article is a revision of Hauser's original report [8] which is based on parts of his thesis [7], while incorporating the solution to the general problem using Güler's approach. Schmieta's report [19] constitutes the first document where a proof of the general classification Theorem 5.5 became publicly available.

The rest of the paper is organised as follows. In Section 2 we reconsider self–scaled cones and self–scaled barriers from a symmetric point of view. Section 3 is devoted to certain properties of self–scaled barriers which link self–scaled barriers to the symmetry group of their domain of definition. These results are needed in later sections. In Section 4, we show that any pointed, convex cone has a unique decomposition as a direct sum of irreducible components. This result, of which we manage to locate only technically more involved generalisations, may be of independent interest. We therefore include a simple proof. We then use this decomposition result to show that any self–scaled barrier defined on a symmetric cone K decomposes in an essentially unique way into a direct sum of self–scaled barriers defined on the irreducible components of K, which also shows that the irreducible components are symmetric cones themselves. This decomposition reduces the problem of classifying self-scaled barriers to the case where the domain of definition is irreducible, a problem we solve in Section 5. Theorem 5.5 constitutes the main and final result of this paper. We thus present all the essential elements of the theory of self–scaled barrier functions in a single document.

The following basic properties of ν–self–concordant logarithmically homogeneous barrier functions and their duals will be used frequently in later sections. These properties are easy consequences of logarithmic homogeneity, see [13] or [14]:

Proposition 1.4. *Let F be a ν-self-concordant logarithmically homogeneous barrier function on the regular cone $K \subset E$, and let $x \in \operatorname{int}(K)$, $s \in \operatorname{int}(K^*)$. Then*

$$
\begin{array}{rlrl}
i) & -F'(x) = F''(x)x \in \operatorname{int}(K^*), & ii) & -F'_*\big(-F'(x)\big) = x, \\
iii) & F''_*(-F'(x)) = F''(x)^{-1}, & iv) & \langle x, -F'(x)\rangle = \nu, \\
v) & F^{(k)}(tx) = t^{-k}F^{(k)}(x),\ \forall\, t>0, k=1,2, & vi) & F_*\big(-F'(x)\big) = -\nu - F(x), \\
vii) & F(x) + F_*(s) \geq -\nu - \nu\log\nu - \nu\log\langle x, s\rangle, & &
\end{array}
$$

where $F^{(1)}(x) = F'(x)$, $F^{(2)}(x) = F''(x)$ in v). □

2 A Symmetric View on Self–Scaledness

In this section we undertake a study of self–scaled cones and barrier functions while emphasising their symmetry properties in a duality-theoretic sense.

Let F be a self–scaled barrier function on a regular cone K in a finite dimensional Euclidean space E equipped with an inner product $\langle \cdot, \cdot \rangle$. With a given arbitrary point $e \in \operatorname{int}(K)$ we associate an inner product

$$\langle u,v\rangle_e := \langle F''(e)u,v\rangle.$$

The following result is due to Güler [3]:

Theorem 2.1. *The cone K is symmetric, and F is self–scaled under $\langle\cdot,\cdot\rangle_e$.*

Proof. We have

$$\begin{aligned}K_e^* &:= \big\{y : \langle x,y\rangle_e \geq 0,\ \forall x\in K\big\} = \big\{y : \langle F''(e)x,y\rangle \geq 0,\ \forall x\in K\big\}\\ &= \big\{y : \langle z,y\rangle \geq 0,\ \forall z\in K^*\big\} = (K^*)^* = K,\end{aligned}$$

where the third equality follows from Theorem 1.2. Note that

$$\langle F''(x)u,v\rangle = D^2F(x)[u,v] = \langle F_e''(x)u,v\rangle_e = \langle F''(e)F_e''(x)u,v\rangle$$

yields $F''(x) = F''(e)F_e''(x)$, or

$$F_e''(x) = F''(e)^{-1}F''(x). \tag{2.1}$$

Theorem 1.2 implies that $F_e''(x)(K) = F''(e)^{-1}F''(x)(K) = F''(e)^{-1}(K^*) = K$, so that $F_e''(x)\in \mathrm{Aut}(K)$. Theorem 1.2 also shows that, given any two points $u,v\in\mathrm{int}(K)$, we can find a (unique) point $z\in K$ such that $F''(z)u = F''(e)v\in K^*$. Therefore, $F_e''(x)(u) = v$, which shows that the set of linear operators $\{F_e''(x) : x\in\mathrm{int}(K)\}$ acts transitively on K. Hence, K is a symmetric cone.

For the second assertion, note that if $s\in K_e^* = K$, then

$$(F_e)_*(s) := \sup_{x\in K}\{-\langle x,s\rangle_e - F(x)\} = \sup_{x\in K}\{-\langle x,F''(e)s\rangle - F(x)\} = F_*(F''(e)s).$$

For $x,z\in\mathrm{int}(K)$, we thus have

$$(F_e)_*(F_e''(z)x) = F_*(F''(e)F_e''(z)x) = F_*(F''(z)x) = F(x) - 2F(z) - \nu,$$

where the second and last equalities follow from (2.1) and (1.5), respectively. Consequently, F is self–scaled under $\langle\cdot,\cdot\rangle_e$. $\square$

Remark 2.2. Theorem 2.1 shows that from here on we may assume without loss of generality that K is a symmetric cone and that there exists a (unique) point $e\in\mathrm{int}(K)$ such that $F''(e) = \mathrm{I}$.

Together with Equation (1.5) this implies that

$$F_*(x) = F(x) - 2F(e) - \nu = F(x) + \mathit{const}, \tag{2.2}$$

and invoking (1.5) once more this implies the identity

$$F(F''(w)x) = F(x) - 2F(w) + 2F(e), \qquad \forall x,w\in\mathrm{int}(K). \tag{2.3}$$

Note that (2.3) is a criterion that involves only the primal barrier F. Indeed, this identity allows one to characterise self–scaled barrier functions without invoking F_*, see Lemma 2.5 below. Changing a barrier function by an additive constant is of no real consequence, as interior–point methods rely on gradient and Hessian information. Therefore, we can assume without loss of generality that $F(e) = 0$. For the same reason, Equation (2.2) makes it possible to think of F and F_* as the same function. Hence, we no longer need to distinguish between primal and dual quantities – between F and F_*, the primal and dual scaling points and so forth.

We next prove a property of the Hessian $F''(w)$ which will become an essential tool for our classification of self–scaled barriers. For all $y\in\mathrm{int}(K)$ let us define

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

Lemma 2.3. *For all $x, w \in \operatorname{int}(K)$ it is true that*

$$P(P(w)x) = P(w)P(x)P(w). \tag{2.4}$$

Proof. Let us define $z = P(w)x$. Equation (2.3) implies that for any $h \in E$ we have $F(F''(w)(z + th)) = F(z + th) - 2F(w) + 2F(e)$. Expanding both sides of this equation and comparing the t^2 terms one gets

$$D^2F(F''(w)z)[F''(w)h, F''(w)h] = D^2F(z)[h, h],$$

or $\langle F''(x)F''(w)h, F''(w)h\rangle = \langle F''(z)h, h\rangle$. Thus, $F''(w)F''(x)F''(w) = F''(z) = F''(P(w)x)$, and (2.4) follows. $\square$

In the proof above, we need only the weaker condition $F(F''(w)x) = F(x) + c(w)$ where $c(w)$ is a constant dependent on w. (However, Lemma 2.5 below shows that this is equivalent to (2.3).) Equation (2.4) is a symmetric version of a Formula (3.2) from [14], see also Equation (1.6) above. In accordance with the established tradition in the theory of Jordan algebras we call (2.4) the *fundamental formula*.

Remark 2.4. We do not have a Jordan algebra connected to F yet, but the fundamental formula leads one to suspect that there might be one. In spring 2000, inspired by the work of Petersson [16], Güler [5] proves that this is indeed the case. Subsequently, S. Schmieta [19] independently discovers the same result, following essentially the same steps. Schmieta uses this fact as an essential tool to classify self–scaled barriers. As it turns out, the Jordan algebra connected to F is already discovered by McCrimmon in his thesis [12], even without the assumption of convexity for F. His proof in turn is a generalisation of Koecher's ideas [10] on *ω–domains.* Reading both works is instructive in delineating the role of convexity.

The following result provides an alternative definition of self–scaled barrier functions.

Lemma 2.5. *Let K be a regular, self–dual cone. A logarithmically homogeneous self–concordant barrier function F on $\operatorname{int}(K)$ is self-scaled if and only if*

$$F''(w)x \in \operatorname{int}(K), \qquad \forall\, x, w \in \operatorname{int}(K), \tag{2.5}$$

$$F(F''(w)x) = F(x) + c(w), \qquad \forall\, x, w \in \operatorname{int}(K), \tag{2.6}$$

where $c(w)$ is a constant that depends on w.

Proof. Since K is self–dual, Equation (2.5) is equivalent to Axiom (1.4). If F is self–scaled, then Equation (2.6) follows from (2.3). Conversely, assume that F satisfies Equation (2.6). Let $x, s \in \operatorname{int}(K)$ be arbitrary points.

We claim that there exists a point $w \in \operatorname{int}(K)$ such that $F''(w)x = s$. Towards proving the claim, we consider the optimization problem $\min\{\langle z^*, x\rangle : \langle z, s\rangle = 1\}$, where $z^* = -F'(z)$. It is well known that the feasible region is bounded, see [1], Corollary I.1.6, p. 4. We have $F(x) + F_*(z^*) \geq -\nu - \nu\log\nu - \nu\log\langle z^*, x\rangle$ (see Proposition 1.4 *vii*)), and $F(z) + F_*(z^*) = -\nu$ (see Proposition 1.4 *vi*)), which imply $F(x) - F(z) \geq -\nu\log\nu - \nu\log\langle z^*, x\rangle$. These imply that the objective function of the optimization problem goes to infinity as z approaches the boundary of the feasible region, and thus the optimization problem has a minimizer $\hat{z} \in \operatorname{int}(K)$ satisfying $F''(\hat{z})x = \lambda s$ for some scalar λ. Since $F''(\hat{z})x, s \in \operatorname{int}(K)$, we have $\lambda > 0$. The point $w = \sqrt{\lambda}\hat{z}$ satisfies $F''(w)x = s$ (see Proposition 1.4 *v*)), and proves the claim.

Next, we claim that

$$c(w) = -2F(w) + 2F(e). \tag{2.7}$$

Let $u \in \operatorname{int}(K)$ be a point satisfying $F''(u)w = e$. The fundamental formula (2.4) is a consequence of (2.6) and gives $F''(u)P(w)F''(u) = \mathrm{I}$, or equivalently $F''(w) = F''(u)^2$. From (2.6), we obtain

$$F(e) + c(w) = F(F''(w)e) = F(F''(u)^2 e) = F(e) + 2c(u),$$

or $c(w) = 2c(u)$. Equation (2.6) also implies that

$$F(e) = F(F''(u)w) = F(w) + c(u) = F(w) + \frac{1}{2}c(w),$$

hence proving the claim.

Using logarithmic homogeneity alone one can prove that $F_*(w^*) = -\nu - F(w)$ where $w^* := -F'(w)$ (see Proposition 1.4 vi)). Proposition 1.4 ii) shows that the mapping $w \mapsto w^*$ is involutive, that is, $w^{**} = w$. These imply $F_*(w) = F_*(w^{**}) = -\nu - F(w^*)$. Since $F''(w)w = w^*$ by Proposition 1.4 i), we have

$$-\nu - F_*(w) = F(w^*) = F(F''(w)w) = F(w) - 2F(w) + 2F(e),$$

which is to say that $F_*(w) = F(w) - 2F(e) - \nu$. This implies

$$F_*(F''(w)x) = F(F''(w)x) - 2F(e) - \nu = F(x) - 2F(w) - \nu,$$

where the last equality follows from Equations (2.6) and (2.7). This concludes the proof. $\square$

Note that together with Equation (2.2), Lemma 2.5 implies that we can replace Axiom (1.5) of the original definition of a self–scaled barrier function by the requirement $F_*(F''(w)x) = F(x) + C(w)$ for some constant $C(w)$ which depends on w. This fact is already known, see [17].

3 Group–Theoretic Aspects of Self–Scaledness

In this section we explore the relationship between the Hessians of self–scaled barrier functions and the symmetry group of their domain of definition. Though we present these results primarily for the purposes of later sections they are also of independent interest.

The universal barrier function $U(x)$ defined in (1.8) plays an important role in the context of this section. The choice of the inner product $\langle \cdot, \cdot \rangle$ used in the definition of the characteristic function $\varphi_K(x)$ via (1.7) is irrelevant, since φ_K changes only by an additive constant under a change of $\langle \cdot, \cdot \rangle$. Güler [4] shows that the universal barrier function $U(x)$ is self–scaled, see Equation (13) and Theorem 4.4 in [4]. For all $x \in \operatorname{int}(K)$ let

$$Q(x) := U''(x)^{-1},$$

and let $f \in \operatorname{int}(K)$ be the point characterised by the equation

$$Q(f) = \mathrm{I}.$$

Remark 3.1. It follows from Theorem 1.2 that f is unique. The existence of such a point is also well known, see for example page 17 of [1].

The point f is the "unit" associated with the self–scaled barrier $U(x)$, see [1] Proposition I.3.5, p. 14, and it is also the unit of the Jordan algebra associated with U.

The following lemma is Theorem 3.17, pp. 205–206 in [18]. We include a short proof of this result because these ideas play an important role in later sections. See also [1], Proposition I.4.3, p. 18 for a different approach to proving this result.

Lemma 3.2. *The orthogonal subgroup* $\mathrm{O}\big(\mathrm{Aut}(K)\big) \subseteq \mathrm{Aut}(K)$ *coincides with the stabiliser group at* f*, that is,*

$$\mathrm{O}\big(\mathrm{Aut}(K)\big) = \{H \in \mathrm{Aut}(K) : Hf = f\}.$$

Proof. If $A \in \mathrm{Aut}(K)$, then

$$D^2U(Af)[Ah, Ah] = D^2U(f)[h, h]$$

for every vector $h \in E$. That is, $A^*Q(Af)^{-1}A = \mathrm{I}$, or $Q(Af) = AA^*$, see for example [4], Equation (11). This shows that A is orthogonal if and only if $\mathrm{I} = Q(Af)$. The uniqueness of f implies that this condition is equivalent to $Af = f$. □

Next we note that the elements of $\mathrm{Aut}(K)$ have a unique *polar decomposition*, see [18], Theorem 3.18, p. 206. For the sake of completeness we give a short proof.

Lemma 3.3. *Let* $A \in \mathrm{Aut}(K)$*. There exists a unique vector* $u \in \mathrm{int}(K)$ *and a unique orthogonal cone automorphism* $H \in \mathrm{O}\big(\mathrm{Aut}(K)\big)$ *such that*

$$A = Q(u)H.$$

Proof. By virtue of Theorem 1.2, there exists a unique point $u \in \mathrm{int}(K)$ such that $Q(u)f = Af$. Then $H := Q(u)^{-1}A$ satisfies $Hf = Q(u)^{-1}Af = f$, which implies that H is orthogonal by Lemma 3.2. Since H is orthogonal and $Q(u)$ is symmetric, $A = Q(u)H$ is indeed a polar decomposition of A.

Suppose now that $A = Q_1H_1 = Q_2H_2$ where Q_i is symmetric and H_i is orthogonal, $i = 1, 2$. Then, $H := Q_2^{-1}Q_1 = H_2H_1^{-1}$ is orthogonal, and we have $\mathrm{I} = HH^* = Q_2^{-1}Q_1^2Q_2^{-1}$, or $Q_2^2 = Q_1^2$. Since Q_1 and Q_2 are symmetric, we have $Q_1 = Q_2$ and $H_1 = H_2$. □

The following result will play a key role in Section 5 where we classify self–scaled barriers.

Lemma 3.4. *The sets of inverse Hessians of* F *and* U *coincide, that is,*

$$\{P(v) : v \in \mathrm{int}(K)\} = \{Q(u) : u \in \mathrm{int}(K)\},$$

and for all $x \in \mathrm{int}(K)$ *it is true that*

$$P(x) = Q(Q(x)^{1/2}e^{-1}) = Q(x)^{1/2}Q(e)^{-1}Q(x)^{1/2}, \tag{3.1}$$

where $e^{-1} \in \mathrm{int}(K)$ *is characterised by the equation* $Q(e^{-1}) = Q(e)^{-1}$.

The point e^{-1} is the inverse of e in the Jordan algebra associated with $U(x)$. Note that Proposition 1.4 *iii)* shows that $e^{-1} = -F'(e)$.

Proof. If $v \in \mathrm{int}(K)$, then Lemma 3.3 implies that we can write $P(v) = Q(u)H$ for some $u \in \mathrm{int}(K)$ and $H \in \mathrm{O}\big(\mathrm{Aut}(K)\big)$. By the uniqueness of the polar decompositions $P(v) = P(v)\,\mathrm{I}$ and $P(v) = Q(u)H$ it must be true that $P(v) = Q(u)$. Thus,

$$\{P(v) : w \in \mathrm{int}(K)\} \subseteq \{Q(u) : u \in \mathrm{int}(K)\}.$$

Conversely, let $u \in \mathrm{int}(K)$. By Theorem 1.2, there exists a point $v \in \mathrm{int}(K)$ such that $P(v)f = Q(u)f$. But this implies that $Q(u)^{-1}P(v)f = f$. By virtue of Lemma 3.2 $H := Q(u^{-1})P(v)$ is therefore orthogonal. This means that $P(v)$ has the polar decompositions $P(v) = P(v)\,\mathrm{I}$ and

$P(v) = Q(u)H$. The uniqueness part of Lemma 3.2 then implies that $Q(u) = P(v)$ and $H = \mathrm{I}$. This proves the first statement of the lemma.

Now let $x \in \operatorname{int}(K)$ and define u by $x = Q(u)f$, see Theorem 1.2. We have

$$Q(x) = Q\big(Q(u)f\big) = Q(u)Q(f)Q(u) = Q(u)^2,$$

where the second equality follows from the fundamental Formula (2.4). In a similar vein, taking the first part of this lemma into account we obtain

$$P(x) = P\big(Q(u)f\big) = Q(u)P(f)Q(u).$$

These two equations imply that $Q(u) = Q(x)^{1/2}$ and

$$P(x) = Q(x)^{1/2}P(f)Q(x)^{1/2}.$$

In particular, setting $x = e$ yields $\mathrm{I} = Q(e)^{1/2}P(f)Q(e)^{1/2}$, that is, $P(f)Q(e) = \mathrm{I}$, and $P(f) = Q(e)^{-1} = Q(e^{-1})$. The lemma follows, since this implies that

$$P(x) = Q(x)^{1/2}Q(e^{-1})Q(x)^{1/2} = Q(Q(x)^{1/2}e^{-1}).$$

□

Although it does not have a direct bearing on later results, the following proposition already shows that the self–scaled barrier function F is intimately connected to the universal barrier function.

Proposition 3.5. *There exist constants $\alpha_1 > 0$ and α_2 such that*

$$U(x) = \alpha_1 \ln \det F''(x) + \alpha_2.$$

Proof. ¿From Equation (3.1) we see that $\det P(x) = \det Q(e)^{-1} \det Q(x)$, implying that $\det F''(x) = c_1 \det U''(x)$ for some constant $c_1 > 0$. Theorem 4.4 in [4] shows that the function $u(x) = \ln \varphi_K(x)$ satisfies the equation $u(x) = c_2 + \frac{1}{2} \ln \det u''(x)$ for some constant c_2. These facts combined with (1.8) imply the proposition. □

4 Decomposition of Cones and Barrier Functions

In this section, we prove two related results. A cone is called *pointed* if it does not contain any whole lines. First, we show that any pointed, convex cone decomposes into a direct sum of indecomposable or *irreducible* components in a unique fashion. This result, which is of independent interest, is essentially a special case of Corollary 1 in [2]. Gruber's paper is the earliest mentioning of this result we could find, though it may have been derived several times independently. Gruber's original result addresses a more general affine setting which renders his proof more technically involved. Therefore, we include a simple and accessible proof. Second, we use this decomposition to write any self–scaled barrier function defined on the interior of a symmetric cone K as a direct sum of self–scaled barriers defined on the irreducible components of K.

Recall that the Minkowski sum of a set $\{A_i\}_{i=1}^k$ of subsets of E is defined as

$$A_1 + \cdots + A_k := \Big\{\sum_{i=1}^k x_i : x_i \in A_i\Big\}.$$

If all of the A_i are linear subspaces $\{0\} \neq E_i \subseteq E$ which satisfy $E = E_1 + \cdots + E_k$ and $E_i \cap (\sum_{j \neq i} E_j) = \{0\}$, then we say that the sum $E = E_1 + \cdots + E_m$ is *direct* and write

$$E = E_1 \oplus E_2 \oplus \cdots \oplus E_m.$$

Definition 4.1. Let $K \subseteq E$ be a pointed, convex cone. K is called *decomposable* if there exist cones $\{K_i\}_{i=1}^m$, $m \geq 2$, such that $K = K_1 + \cdots + K_m$, where each K_i $(i = 1, \ldots, m)$ lies in a linear subspace $E_i \subset E$, and where the spaces $\{E_i\}_{i=1}^m$ decompose E into a direct sum $E = E_1 \oplus E_2 \oplus \cdots \oplus E_m$. Each K_i is called a *direct summand* of K, and K is called the *direct sum* of the $\{K_i\}$. We write

$$K = K_1 \oplus K_2 \oplus \cdots \oplus K_m \tag{4.1}$$

to denote this relationship between K and $\{K_i\}_{i=1}^m$. K is called *indecomposable* or *irreducible* if it cannot be decomposed into a non-trivial direct sum.

Let us define $\hat{E}_i := \oplus_{j \neq i} E_j$ and $\hat{K}_i := \oplus_{j \neq i} K_j$. If K is the direct sum (4.1) then every $x \in K$ has a unique representation $x = x_1 + \cdots + x_m$ with $x_i \in K_i \subseteq E_i$. Thus, $x_i = \pi_{E_i} x$, where π_{E_i} is the projection of E onto E_i along $\hat{E}_i$. Also, since $0 \in K_i$, we have $K_i = K_i + \sum_{j \neq i} \{0\} \subseteq \sum_{j=1}^m K_j = K$. Therefore,

$$\pi_{E_i} K = K_i \subseteq K.$$

This implies that $K_i = \pi_{E_i} K$ is a convex cone. Similarly, we have

$$(\mathrm{I} - \pi_{E_i}) K = \pi_{\hat{E}_i} K = \hat{K}_i \subseteq K.$$

We first prove a useful technical result:

Lemma 4.2. *Let K be a pointed, convex cone which decomposes into the direct sum* (4.1). *If $x \in K_i$ is a sum $x = x_1 + \cdots + x_k$ of elements $x_j \in K$, then each $x_j \in K_i$.*

Proof. We have $0 = \pi_{\hat{E}_i} x = \pi_{\hat{E}_i} x_1 + \ldots + \pi_{\hat{E}_i} x_k$. Each term $\hat{x}_j := \pi_{\hat{E}_i} x_j \in \hat{K}_i \subseteq K$, therefore we have $\hat{x}_j \in K$ and $-\hat{x}_j = \sum_{l \neq j} \hat{x}_l \in K$. Since K contains no lines it must be true that $\hat{x}_j = 0$, that is, $x_j = \pi_{E_i} x_j \in K_i$, $(j = 1, \ldots, k)$. ∎

Theorem 4.3. *Let $K \subseteq E$ be a decomposable, pointed, convex cone. The irreducible decompositions of K are identical modulo indexing, that is, the set of cones $\{K_i\}_{i=1}^m$ is unique. Moreover, the subspaces E_i corresponding to the non–zero cones K_i are also unique. In particular, if K is solid, then all the cones K_i are non–zero and the subspaces E_i are unique.*

Proof. Suppose that K admits two irreducible decompositions

$$K = \bigoplus_{i=1}^m K_i \subseteq \bigoplus_{i=1}^m E_i \quad \text{and}$$
$$K = \bigoplus_{j=1}^q C_j \subseteq \bigoplus_{j=1}^q F_j.$$

Note that each non–zero summand in either decomposition of K must lie in $\mathrm{span}(K)$ and that the subspace corresponding to each zero summand must be one–dimensional, for otherwise the summand would be decomposable. This implies that the number of zero summands in both decompositions is $\mathrm{codim}\big(\mathrm{span}(K)\big)$. We may thus concentrate our efforts on $\mathrm{span}(K)$, that is, we can assume that K is solid and all the summands of both decompositions of K are non–zero. By (4.1), each $x \in C_j \subseteq K$ has a unique representation $x = x_1 + \cdots + x_m$ where $x_i = \pi_{E_i} x \in K_i \subseteq K$. Also, Lemma 4.2 implies that $x_i \in C_j$, and hence $x_i \in K_i \cap C_j$. Consequently, every $x \in C_j$ lies in the set $(K_1 \cap C_j) + \cdots + (K_m \cap C_j)$. Conversely, we have $K_i \cap C_j \subseteq C_j$, implying that $(K_1 \cap C_j) + \cdots + (K_m \cap C_j) \subseteq C_j$. Therefore, we have

$$C_j = (K_1 \cap C_j) + \cdots + (K_m \cap C_j).$$

We have $K_i \cap C_j \subseteq E_i \cap F_j$, $F_j = (E_1 \cap F_j) + \cdots + (E_m \cap F_j)$, and the intersection of any two distinct summands in the last sum is the trivial subspace $\{0\}$. The above decompositions of F_j and C_j are thus direct sums. Since C_j is indecomposable, exactly one of the summands in the decomposition of C_j is non–trivial. Thus, $C_j = K_i \cap C_j$, and hence $C_j \subseteq K_i$ for some i. Arguing symmetrically, we also have $K_i \subseteq C_l$ for some l, implying that $C_j \subseteq C_l$. Therefore, $j = l$ for else $C_j \subseteq F_j \cap F_l = \{0\}$, contradicting our assumption above. This shows that $C_j = K_i$. The theorem is proved by repeating the above arguments for the cone $\hat{K}_i = \oplus_{k \neq i} K_k = \oplus_{l \neq j} C_l$. $\square$

Next, we show that self–scaled barrier functions have irreducible decompositions as well. Let F be defined on $\mathrm{int}(K)$ where K is a symmetric cone with irreducible decomposition (4.1). For $i = 1, \dots, m$, let F_i be a function defined on $\mathrm{ri}(K_i)$, the relative interior of K_i in E. If $F(x) = \sum_{i=1}^m F_i(x_i)$ for every $x = \sum_{i=1}^m x_i \in \oplus_{i=1}^m \mathrm{ri}(K_i) = \mathrm{int}(K)$, then we say that F is the direct sum of the F_i and write

$$F = \bigoplus_{i=1}^m F_i. \tag{4.2}$$

Theorem 4.4. *Let K be a symmetric cone with irreducible decomposition* (4.1)*. The irreducible components K_i are symmetric cones. Let $F(x)$ be a self–scaled barrier for K. There exist self–scaled barrier functions F_i for the cones K_i such that*

$$F = F_1 \oplus \cdots \oplus F_m.$$

The functions F_i are unique up to additive constants.

Proof. Recall that the universal barrier function $U(x)$ is given in (1.8). Changing the inner product used in the definition of the characteristic function $\varphi_K(x)$ changes $U(x)$ only by an additive constant, hence we may assume that $\langle x, y \rangle = \sum_{i=1}^m \langle x_i, y_i \rangle_{E_i}$ for the purposes of this definition. Here, $\langle \cdot, \cdot \rangle_{E_i}$ is an inner product defined on E_i chosen so that $U_i''(f_i) = id_{E_i}$ for some elements $f_i \in \mathrm{ri}(K_i)$ where U_i denotes the universal barrier function defined on $\mathrm{ri}(K_i)$. Then we have $Q(f) = \mathrm{I}$ for $f = \oplus_{i=1}^m f_i \in \mathrm{int}(K)$, in full consistency with our previous notation. Moreover, K is self–dual under $\langle \cdot, \cdot \rangle$, since $K^* = Q(f)K = K$. Hence, we may choose the vector $e \in \mathrm{int}(K)$ specified in Remark 2.2 as the unique element in $\mathrm{int}(K)$ such that $F''(e) = \mathrm{I}$ under $\langle \cdot, \cdot \rangle$, see Remark 3.1.

Thus, U can be written as the direct sum $U(x) = \oplus_{i=1}^m U_i(x_i)$ and $Q(x)$ has block structure corresponding to the subspaces E_i, $Q(x) = \oplus_{i=1}^m Q_i(x_i)$ where $Q_i(x_i) = U_i''(x_i)^{-1}$. Consequently, Equation (3.1) implies that $P(x)$ also has the same block structure,

$$P(x) = P_1(x_1) \oplus \cdots \oplus P_m(x_m), \tag{4.3}$$

where $P_i(x) = Q_i\big(Q_i(x_i)^{1/2} e_i^{-1}\big) \in \mathrm{Aut}(K_i)$ with $e_i^{-1} = \pi_{E_i}\big(-F'(e)\big)$, π_{E_i} being the projection at the beginning of this section.

So far we know that $P(x)$ has block structure corresponding to the direct sum $E = \oplus_{i=1}^m E_i$, but it is not a priori clear that $P_i(x_i)^{-1}$ is the Hessian of a function defined on $\mathrm{ri}(K_i)$. Let the spaces $\hat{E}_i$ be defined as earlier in this section, and let us consider the vector fields $v_i : x \mapsto \pi_{E_i} F'(x)$, defined on $\mathrm{int}(K)$ and taking values in E_i for all i, $i = 1, \dots, m$. We claim that v_i depends only on $x_i = \pi_{E_i} x$. In fact, for any two vectors $x, y \in \mathrm{int}(K)$ such that $x_i = y_i$ we have

$$\begin{aligned} v_i(y) = \pi_{E_i} F'(y) &= \pi_{E_i} \Big[F'(x) + \int_0^1 F''\big(ty + (1-t)x\big)[y - x] dt \Big] \\ &= v_i(x) + \int_0^1 P_i\big(\pi_{E_i}[ty + (1-t)x]\big)^{-1} \pi_{E_i}[y - x] dt = v_i(x) + \int_0^1 0\, dt, \end{aligned}$$

which shows our claim. Hence, the quotient vector fields

$$\begin{aligned} \hat{v}_i &: \operatorname{int}(K)/\hat{E}_i \to E_i, \\ x &\mod \hat{E}_i \mapsto \pi_{E_i} F'(x) \end{aligned}$$

are well defined and can be identified with vector fields $\hat{v}_i$ defined on the cones $\operatorname{ri}(K_i)$. The direct sum of these vector fields amounts to the gradient field

$$F' = \hat{v}_i \oplus \cdots \oplus \hat{v}_m : \operatorname{int}(K) \to E. \tag{4.4}$$

F' being conservative, the $\hat{v}_i$ must be conservative too, implying that these are the gradient vector fields of some functions F_i defined on $\operatorname{ri}(K_i)$ which are uniquely determined up to additive constants. We may choose these constants so that $F = \oplus_{i=1}^m F_i$. Clearly, we have $F_i''(x_i) = P_i(x)^{-1}$ for any $x \in \operatorname{int}(K)$.

Using Equation (4.3), it is straightforward to check that the F_i are self-concordant, see [13]. Applying Proposition 1.4 *i*) and *iv*) to F, using Equation (4.4) and considering variations of $x \in \operatorname{int}(K)$ only in the part $x_i = \pi_{E_i} x$, we obtain that $\langle x_i, -F_i'(x_i)\rangle = \nu_i$ for some number $\nu_i > 0$. Moreover, applying Proposition 1.4 *v*) to F and using Equation (4.4) we get $F_i'(\tau x_i) = \tau^{-1} F_i'(x_i)$ for all $\tau > 0$. Hence,

$$\begin{aligned} F_i(\tau x_i) &= F_i(x_i) + \int_1^\tau \frac{d}{d\xi} F_i(\xi x_i)\, d\xi = F_i(x_i) + \int_1^\tau \big\langle x_i, F_i'(\xi x_i)\big\rangle d\xi \\ &= F_i(x_i) - \int_1^\tau \xi^{-1} \big\langle x_i, -F_i'(x_i)\big\rangle d\xi = F_i(x_i) - \nu_i \int_1^\tau \xi^{-1} d\xi \\ &= F_i(x_i) - \nu_i \ln \tau. \end{aligned}$$

This shows that the functions F_i are ν_i-logarithmically homogeneous. It is a well–known fact that any logarithmically homogeneous self–concordant function is also a barrier function, see for example [13] or [17]. It remains to show that the functions F_i are self–scaled. As previously noted, Condition (2.5) is satisfied, since $P_i(x_i) \in \operatorname{Aut}(K_i)$ for all i. Finally, Condition (2.6) holds for F_i because we can apply this condition to F, choosing $w = w_i \oplus \big(\oplus_{j \neq i} e_i\big)$ and $x = x_i \oplus \big(\oplus_{j \neq i} e_i\big)$ and using the block structures of F and F''.

Note that the irreducible components K_i of K must be symmetric cones, since the F_i are self–scaled barriers defined on $\operatorname{ri}(K_i)$. The symmetry of the K_i can also be directly derived from the block structure of $Q(x) = \oplus_{i=1}^m Q_i(x_i)$ and the fact that the set of cone automorphisms $\big\{Q(x) : x \in \operatorname{int}(K)\big\}$ acts transitively on K. $\square$

The decomposition Theorem 4.4 shows that for the purposes of classifying self–scaled barriers we may concentrate our efforts on irreducible cones.

5 Classification of Self–Scaled Barriers

In this section, we give a complete classification of self–scaled barrier functions on the symmetric cone K.

The definition of a self–scaled barrier function F requires that F changes only by an additive constant under the action of symmetric cone automorphisms $\{P(u) : u \in \operatorname{int}(K)\}$, see Equation (2.3). However, it is not a priori known how F behaves under the action of an arbitrary element of $\operatorname{Aut}(K)$. Note that this is in marked contrast to the case where F is the universal barrier function U, which is known to change only by an additive constant under the action of any element of the symmetry group of K. This explains in a sense the main difficulty one faces when proving the results below.

The next result is key in resolving this difficulty and is just a slight reformulation of the conjecture raised in [8], according to which self-scaled barriers on irreducible symmetric cones are isotropic. Let us denote by $\mathrm{Aut}(K)_0$ the connected component of the identity in $\mathrm{Aut}(K)$.

Theorem 5.1. *Let K be a symmetric cone. If $H \in \mathrm{Aut}(K)_0$ is orthogonal, then $F(Hx) = F(x)$ for all $x \in \mathrm{int}(K)$.*

Proof. Koecher [10] proves that if K is a symmetric cone, then $\mathrm{Aut}(K)_0$ is generated by $\{Q(u) : u \in \mathcal{V}\}$ where $\mathcal{V}$ is a neighbourhood of the identity f, see [10], Theorem 4.9 (b), pp. 88–89. Koecher's proof exploits the fact that all *derivations* of the Jordan algebra associated with $U(x)$ are *inner*. An accessible proof for the case where K is irreducible is given in [1], Lemma VI.1.2, pp. 101–102, based on certain non–trivial results from the theory of Jordan algebras. If $H \in \mathrm{Aut}(K)_0$ is orthogonal, it follows from Koecher's result that

$$H = \prod_1^l Q(u_i) = \prod_1^l P(v_i),$$

for some $u_i, v_i \in \mathrm{int}(K)$, $(i = 1, \dots, l)$. Here the second equality follows from Lemma 3.4. Therefore, it follows from Equation (2.3) that

$$F(Hx) = F(\prod_1^l P(v_i)x) = F(x) - 2\sum_1^l F(v_i).$$

Since $Hf = f$, setting $x = f$ above yields $\sum_1^l F(v_i) = 0$ and settles the claim of the theorem. □

The group $\mathrm{Aut}(K)_0$ already acts transitively on K, see [1], page 5. Thus, the above result is significant.

Remark 5.2. Hauser's approach [9] to solving the isotropy conjecture for the cone of positive semidefinite symmetric matrices is based on similar ideas. Hauser essentially rediscovers Koecher's Theorem 4.9 (b) in this special case, see [9] Proposition 3.3 and Corollary 4.3. He then uses Proposition 3.3 in conjunction with the fundamental formula as the key mechanism in the proof of the conjecture. Y. Lim [11] generalised this approach to arbitrary irreducible symmetric cones while refereeing the paper [9], thus completing the classification of self–scaled barriers. However, since Lim was an anonymous referee, his result was not publicly announced.

Now, assume that K is irreducible and let U be the universal barrier function for K. Let k be the rank of the Jordan algebra associated with $U(x)$. Let x be an arbitrary point in $\mathrm{int}(K)$. Then there exists an orthogonal frame $\{e_1, \dots, e_k\}$ such that $x = \sum_{i=1}^k \lambda_i e_i$, $\lambda_i > 0$, $i = 1, \dots, k$, see [1], Theorem III.1.2, pp. 44–45. By C let us denote the cone generated by this frame, that is,

$$C = \{\sum_{i=1}^k \lambda_i e_i : \lambda_i \geq 0\}.$$

Note that C is a direct sum of the half–lines $\{\lambda_i e_i : \lambda_i \geq 0\}$ and thus a symmetric cone.

Lemma 5.3. *If F is a self–scaled barrier function defined on the interior of the irreducible symmetric cone K, and if C is as defined above, then*

$$F(\sum_{i=1}^k \alpha_i e_i) = -\frac{\nu}{k}\sum_{i=1}^k \log \alpha_i, \qquad \alpha_i > 0 \quad (i = 1, \dots, k).$$

Proof. Let σ be any permutation of $\{1, \ldots, k\}$. Theorem IV.2.5 in [1] implies that there exists an orthogonal automorphism $H \in Aut(K)_0$ such that $He_i = e_{\sigma(i)}$ $(i = 1, \ldots, k)$. Using Theorem 5.1, we then obtain

$$F(\sum_{i=1}^{k} \alpha_i e_{\sigma(i)}) = F(\sum_{i=1}^{k} \alpha_i e_i), \qquad \forall \alpha_i > 0 \quad (i = 1, \ldots, k).$$

Define $g(\alpha_1, \ldots, \alpha_k) := F(\sum_{i=1}^{k} \alpha_i e_i)$. Note that g is a symmetric function. Consider a point $e + \sum_{i=1}^{k} \beta_i e_i = \sum_{i=1}^{k} \alpha_i e_i \in \mathrm{ri}(C)$, with arbitrary $\beta_i \geq 0$ and $\alpha_i := 1 + \beta_i$. Applying Theorem 5.1 in [14] repeatedly, we obtain

$$F(e + \sum_{i=1}^{k} \beta_i e_i) - F(e) = \sum_{i=1}^{k} (F(e + \beta_i e_i) - F(e)).$$

Using the symmetry of g and $F(e) = 0$, the above equation translates into

$$g(\alpha_1, \ldots, \alpha_k) = \sum_{i=1}^{k} g(\alpha_i, 1, \ldots, 1), \qquad \forall \alpha_i \geq 1 \quad (i = 1, \ldots, k).$$

If $\alpha_1 = \ldots = \alpha_k = \alpha$ above, we have $g(\alpha, \ldots, \alpha) = F(\alpha e) = F(e) - \nu \log \alpha = -\nu \log \alpha$. This gives $g(\alpha, 1, \ldots, 1) = -\frac{\nu}{k} \log \alpha$ for all $\alpha \geq 1$. Consequently, we have

$$g(\alpha_1, \ldots, \alpha_k) = -\frac{\nu}{k} \sum_{i=1}^{k} \log \alpha_i, \qquad \forall \alpha_i \geq 1 \quad (i = 1, \ldots, k). \tag{5.1}$$

Now, if $\alpha_i > 0$ $(i = 1, \ldots, k)$ are arbitrary, choose $t > 0$ such that $\hat{\alpha}_i = \alpha_i / t \geq 1$. Since F is logarithmically homogeneous, we have $g(\alpha_1, \ldots, \alpha_k) = g(\hat{\alpha}_1, \ldots, \hat{\alpha}_k) - \nu \log t$ by the logarithmic homogeneity of F. A simple calculation shows that (5.1) holds true for all $\alpha_i > 0$. The lemma is proved. □

The following theorem classifies self–scaled barrier functions for irreducible symmetric cones.

Theorem 5.4. *Let K be an irreducible symmetric cone, and let F be a self–scaled barrier function defined on* $\mathrm{int}(K)$. *Then there exist constants $\alpha > 0$ and β such that*

$$F(x) = \alpha U(x) + \beta,$$

where $U(x)$ is the universal barrier function on $\mathrm{int}\, K$.

Proof. Lemma 5.3 describes the restriction of F on $\mathrm{ri}(C)$. Since the universal barrier function is also self–scaled, the same considerations apply to $U(x) - U(e)$. Thus, the functions F and U are homothetic transformations of each other on $\mathrm{ri}(C)$, that is, there exist $\alpha > 0$, β such that

$$F(x) = \alpha U(x) + \beta. \tag{5.2}$$

Let $y \in \mathrm{int}(K)$ be an arbitrary point with the spectral decomposition $y = \sum_{i=1}^{k} \nu_i d_i$. Corollary IV.2.7 in [1] implies that there exists $A \in O(\mathrm{Aut}(K)_0)$ such that $Ae_i = d_i$ $(i = 1, \ldots, k)$. We have $y = Ax$ where $x = \sum_{i=1}^{k} \nu_i e_i \in \mathrm{ri}(C)$. Theorem 5.1 gives $F(y) = F(x)$ and $U(y) = U(x)$, and hence the Identity (5.2) extends to all of $\mathrm{int}(K)$. □

We are now ready to give the final classification theorem for self–scaled barrier functions on arbitrary symmetric cones. This theorem shows that all self–scaled barrier functions are related to the standard logarithmic or the universal barrier via homothetic transformations.

Theorem 5.5. *Let F be a self–scaled barrier function for a symmetric cone K with irreducible decomposition* (4.1)*. Then there exist constants c_0 and $c_1 \geq 1, \ldots, c_m \geq 1$ such that*

$$F = c_0 - \bigoplus_{i=1}^{m} c_i \ln \det{}_{K_i},$$

see Equation (4.2)*. Here $\det_{K_i} x_i$ denotes the determinant of $x_i \in \mathrm{ri}(K_i)$ in the Jordan algebraic sense, see [1], Chapter 2. Conversely, any function of this form is a self–scaled barrier for K.*

Proof. Theorems 4.4 and 5.4 imply that there exist constants d_0 and $d_1 > 0, \ldots, d_m > 0$ such that

$$F(x) = d_0 + d_1 u_1(x_1) + \ldots + d_m u_m(x_m),$$

where $u_i(x_i) = \ln \varphi_{K_i}(x_i)$. It is known that $u_i(x_i) = \mathit{const} - n_i/r_i \ln \det x_i$, where r_i is the rank of the Jordan algebra associated with $u_i(x)$, and n_i is the dimension of the cone K_i, see [1], Proposition III.4.3, p. 53. Finally, Theorem 4.1 in [6] implies that the function $-\alpha \ln \det x_i$ is self–concordant if and only if $\alpha \geq 1$. □

6 Acknowledgements

R. Hauser conducted his research on this paper partially while writing his Ph.D. at the School of Operations Research and Industrial Engineering of Cornell University, and partially as a post-doctoral fellow at the University of Cambridge. He wishes to thank Mike Todd, Jim Renegar and Arieh Iserles for numerous valuable discussions and for their continuous support. This research was supported by a “bourse pour chercheur débutant” from the Swiss National Science Foundation and the Swiss Academy of Technical Sciences, and by the Engineering and Physical Sciences Research Council of the UK, EPSRC grant GR/M30975.

The research of O. Güler on this paper was conducted at the Industrial Engineering Department at Bilkent University, Ankara, Turkey, while on a sabbatical leave from the Department of Mathematics and Statistics, University of Maryland, Baltimore County, Baltimore, Maryland. The author thanks the Industrial Engineering Department at Bilkent University for providing him a congenial atmosphere and excellent working conditions. This research was partially supported by the National Science Foundation under grant DMS–0075722.

References

[1] J. Faraut, and A. Korányi, *Analysis on Symmetric Cones*, Oxford University Press, Oxford, 1994.

[2] P. M. Gruber, Zur Charakterisierung konvexer Körper. Über einen Satz von Rogers und Shepard II, *Mathematische Annalen* **184** (1970), 79–105.

[3] O. Güler, Private communication to Y. E. Nesterov and M. J. Todd, April 1994.

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

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

[6] O. Güler and L. Tunçel, Characterization of the barrier parameter of homogeneous convex cones, *Mathematical Programming (Series A)* **81** (1998) 55–76.

[7] 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.

[8] 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.

[9] 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.

[10] M. Koecher, *The Minnesota Notes on Jordan Algebras and their Applications*, Lecture Notes in Mathematics 1710, Springer–Verlag, New York, 1999.

[11] Y. Lim, Private communication to L. Faybusovich, Spring 2000.

[12] K. McCrimmon, Norms and noncommutative Jordan algebras, *Pacific Journal of Mathematics* **15** (1965), 925–956.

[13] Y. E. Nesterov and A. S. Nemirovskii, *Interior-point Polynomial Algorithms in Convex Programming*, SIAM Publications, Philadelphia, 1994.

[14] Y. E. Nesterov and M. J. Todd, Self–scaled barriers and interior–point methods for convex programming, *Mathematics of Operations Research* **22** (1997), 1–42.

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

[16] H. P. Petersson, Max Koecher's work on Jordan algebras. In *Jordan Algebras, W. Kaup, K. McCrimmon, and H. L. Petersen, Eds.*, pp. 187–195. Walter de Gruyter, Berlin, (1995).

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

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

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

[20] È. B. Vinberg, The theory of homogeneous convex cones, *Transactions of Moscow Mathematical Society* **12** (1963), 340–403.