THREE DIFFERENT VIEWS ON BARRIER FUNCTIONS IN CONIC OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

A smooth enough function $f : D \to \mathbb{R}$ defined on a domain $D \subset V$ in a real vector space is an analytic object which defines two different geometric objects. On the one hand, one may consider its level hypersurfaces as affine hypersurface immersions in V, on the other hand the graph of its gradient ∇f as a Lagrangian immersion in the product $V \times V^*$ of the real space with its dual. It turns out that when f is a logarithmically homogeneous barrier on a conic domain D, then the three objects are almost equivalent to each other in the sense that they contain nearly the same information and can be recovered from each other. Properties of the barrier such as self-concordance and self-scaledness are equivalent to meaningful properties of the geometric objects. This equivalence furnishes new vantage points to study barriers in conic optimization and builds a bridge to other areas of mathematics which open new ways to obtain results in optimization. We describe the links and equivalences between the three different view-points and give examples of results obtained by means of these connections.

1 INTRODUCTION

Consider a convex minimization problem of the form

$$\min_{x \in D} f(x), \tag{1}$$

where D is a domain in a real vector space V of dimension n, f is sufficiently smooth, and the minimum exists. Finding it is equivalent to finding a stationary point. The problem is therefore equivalent to the problem of finding the intersection point of the two sets

$$G = \{(x, p) \in D \times V^* \mid p = f'(x)\}, \qquad L_0 = V \times \{0\}$$

in the direct product $V \times V^*$. Here G is the graph of the gradient of f. When studying optimization methods in different setups, the one or the other viewpoint may be more useful.

We present three different viewpoints, mostly for the setup of conic optimization using logarithmically homogeneous barriers, which are equivalent to each other, but belong to three different mathematical domains. In each of these domains mathematicians have developed their own theories and created a pool of literature. The goal of this paper is to make the relations between the viewpoints precise, to relate properties known in these domains under different names to each other, and thus to make the corresponding bodies of knowledge available to research in optimization. We complement the paper with examples of results inspired by theories which were initially non-intersecting with optimization, but which have been made available by the developments presented here.

The three viewpoints which are presented here are

- the analytic viewpoint familiar in optimization
- affine differential geometry, studying hypersurface immersions into real vector spaces
- the theory of Lagrangian submanifolds in para-Kähler spaces

In the context of problem equation 1 the hypersurfaces are the level sets of the function f, while the Lagrangian manifold is the graph G of the gradient of f. As an example of the utility of the affine differential geometric viewpoint we present the canonical barrier.

Most results have been scattered in one form or the other in previous publications. It should also be mentioned that the connection between hypersurface immersions and Lagrangian manifolds has been repeatedly noted by geometers in different contexts. We see this work rather as a systematic description of the aforementioned equivalences and connections, to make the geometric body of knowledge more accessible to researchers in optimization.

1.1 Conic programs and self-concordant barriers

In this section we give an account on the classic objects and properties in conic programming.

A *regular* convex cone K is a closed convex cone in a real vector space V, with non-empty interior, and containing no lines. A *conic program* over a regular convex cone K is a problem of the form

$$\min_{x \in K} \langle c, x \rangle \quad : \quad Ax = b. \tag{2}$$

To any conic program over a cone $K \subset V$ one can associate a *dual program*, which is a conic program over the dual cone

$$K^* = \{s \in V^* \mid \langle x, s \rangle \geq 0 \; \forall \, x \in K\}.$$

Well-known classes of conic programs are linear programs (LP), second-order cone programs (SOCP), and semi-definite programs (SDP), which all employ *symmetric cones* Faybusovich (1995).

The main innovation which permitted the design of efficient interior-point methods to solve conic programs was the invention of the *self-concordant barrier* Nesterov & Nemirovskii (1989). These authors recognized two properties of barriers as being sufficient to explain the performance of interior-point methods, namely logarithmic homogeneity and self-concordance. In the book Nesterov & Nemirovskii (1994) a self-contained theory of interior-point methods for conic programs over arbitrary regular convex cones has been elaborated.

Definition 1.1. *Nesterov & Nemirovskii (1994) Let $K \subset V$ be a regular convex cone. A function $F : \operatorname{int} K \to \mathbb{R}$ is called* logarithmically homogeneous *with parameter ν if*

$$F(tx) = -\nu \log t + F(x) \quad \forall \, t > 0, \; x \in \operatorname{int} K. \tag{3}$$

Let $C \subset V$ be an open convex set. A convex C^3 function $F : C \to \mathbb{R}$ is called self-concordant *if for all $x \in C$ and all $h \in V$ the relation*

$$|F'''(x)[h, h, h]| \leq 2(F''(x)[h, h])^{3/2} \tag{4}$$

holds. It is called a self-concordant barrier *on C if in addition*

$$\lim_{x \to \partial C} F(x) = +\infty.$$

*A ν-*self-concordant barrier *or* self-concordant barrier with parameter ν *on a regular convex cone K is a function combining all these properties, with $C = \operatorname{int} K$.*

In order for an interior-point method to be efficient in solving a conic program over a cone K, an efficiently computable self-concordant barrier for the cone K with a low barrier parameter has to be available. In (Nesterov & Nemirovskii, 1994, Section 2.5) a self-concordant barrier, the *universal barrier*, has been constructed for an arbitrary cone K, with barrier parameter bounded by above n Bubeck & Eldan (2019), but its computation is in general difficult. In the case of symmetric cones the construction reduces to the well-known log-determinantal barriers.

In Nesterov & Todd (1994) it was observed that the logarithmic barriers on symmetric cones have a property in common which barriers on general cones do not have, namely that of being *self-scaled*. In order to introduce this notion we first have to define the dual barrier.

Definition 1.2. *Let F be a self-concordant barrier on a cone K with parameter ν. The dual barrier of F is given by*

$$F_*(s) = \max_{x \in K}(-\langle x, s \rangle - F(x)), \qquad s \in \operatorname{int} K^*.$$

By (Nesterov & Nemirovskii, 1994, Theorem 2.4.4) F_* is indeed a self-concordant barrier on the dual cone K^* with the same parameter ν as F. In fact, $F_*(-p)$, $p \in -\operatorname{int} K^*$, is the Legendre transformation of F.

Definition 1.3. *Let $K \subset \mathbb{R}^n$ be a regular convex cone, let K^* be its dual cone, let F be a self-concordant barrier on K, and let F_* be the dual barrier on K^*. Then F is called* self-scaled *if for every $x, w \in \operatorname{int} K$ we have*

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

A cone K admitting a self-scaled barrier is called self-scaled cone.

For every $x \in \operatorname{int} K$ and $s \in \operatorname{int} K^*$ there actually exists a unique point $w \in \operatorname{int} K$ such that $F''(w)x = s$ (Nesterov & Todd, 1997, Theorem 3.1). The point w is called the *scaling point* of the pair (x, s). Nesterov & Todd (1997; 1998) developed a theory of interior-point methods especially for self-scaled barriers, where the directions of the steps, the so-called *Nesterov-Todd directions*, are computed using the local metric $F''(w)$ at the scaling point. In Nesterov (2006; 2012) the notion of scaling point has been considered for general barriers.

While the property of being self-scaled is primarily a property of the barrier, the property of being symmetric is a property of the cone. Nevertheless, these notions turned out to have a close connection. Several authors proved independently that the symmetric cones and the self-scaled cones form the same class, and provided a full classification of self-scaled barriers Hauser (2000b; 1999; 2000a); Schmieta (2000); Hauser & Lim (2002).

1.2 Geometry of self-concordant barriers

Connections between self-concordant barriers and geometry have been already noticed and employed. In this section we review past developments.

Bayer & Lagarias (1989) and Lagarias (1990) analyzed the vector fields of descent directions for solving LPs. In Karmarkar (1988) the Riemannian manifold defined by the Hessian of the barrier for LP has been studied. The complexity, more precisely the number of steps, grows with the curvature of the trajectories of the aforementioned vector fields, because the curvature determines how well the continuous trajectory can be approximated by discrete steps. Later the iteration complexity was linked to the barrier parameter ν Nesterov & Nemirovskii (1994).

Nesterov and Nemirovski showed that the Legendre transformation which maps the interior of the primal cone K to the interior of the dual cone K^* is an *isometry* when these interiors are equipped with Hessian Riemannian metrics generated by a mutually dual pair of self-concordant barriers F, F_* on K and K^*, respectively (Nesterov & Nemirovskii, 1994, p.45), see also Nesterov & Todd (2002) for an explicit statement. Moreover, the third derivative F''' is mapped to $-F_*'''$ and the first derivative F' to $-F_*'$. This easily follows from the geometric theories presented here, especially that of para-Kähler spaces, where primal and dual space fuse into a single product space.

In Ohara (1999); Ohara & Tsuchiya (2007); Kakihara et al. (2013) parallels have been drawn between the geometry of barriers and information geometry. The presence of a *dually flat structure* has been noted. Such a structure is also known under the name of *Hessian manifold* Shima (2007). It has been shown that the affine scaling vector field is parallel under the dual flat connection, which is an equivalent reformulation of the result in Bayer & Lagarias (1989) on the stratification of this vector field by the Legendre transformation.

2 Mathematical tools

In this section we introduce the employed concepts from differential geometry. Basic notions like Riemannian metrics, (co-)tangent spaces, vector and tensor fields are supposed to be known. Introductions into tensor calculus can be found, e.g., in Spivak (1999); Warner (1983).

2.1 Parallelism and curvature

Let M be a Riemannian manifold with metric tensor $g_{\mu\nu}$. A Riemannian manifold is called *complete* if every geodesic can be prolonged infinitely. The metric defines a parallel transport on M. Informally, a vector field on M is parallel if it is invariant with respect to the geodesic flow. Technically, for a vector field u^α we may define a *covariant derivative*

$$\hat{\nabla}_\beta u^\alpha = \partial_\beta u^\alpha + \Gamma^\alpha_{\beta\gamma} u^\gamma. \tag{5}$$

Here we follow the Einstein convention of summation over repeating indices. The symbol ∂_β denotes the partial derivative with respect to the coordinate x^β on M. Here

$$\Gamma^{\mu}_{\alpha\beta} = \frac{1}{2} g^{\mu\gamma} \left(g_{\alpha\gamma,\beta} + g_{\beta\gamma,\alpha} - g_{\alpha\beta,\gamma} \right) \tag{6}$$

are the *Christoffel symbols*, which are symmetric in the lower indices, and $g^{\mu\gamma}$ is the inverse of the metric tensor. equation 5 informs us how much the component u^α changes compared to a hypothetical vector field $\tilde{u}^\alpha$ which is left unchanged by the geodesic flow in the direction of the basis vector e_β. If $\hat{\nabla}_\beta u^\alpha = 0$ for all α, β, then the vector field u is parallel with respect to the metric.

Parallelism can be defined for any tensor field. A dual vector field v_α is parallel if and only if its convolution $v_\alpha u^\alpha$ vanishes for all parallel vector fields u^α. A tensor field $T^{\beta_1 \cdots \beta_l}_{\alpha_1 \ldots \alpha_k}$ is parallel if and only if its convolution $T^{\beta_1 \cdots \beta_l}_{\alpha_1 \ldots \alpha_k} u^{\alpha_1} \ldots w^{\alpha_k} v_{\beta_1} \ldots z_{\beta_l}$ vanishes for all parallel (co-)vector fields $u, v, \ldots, w, z$.

Let us now consider a Riemannian manifold $\mathcal{M}$ with metric G and a smooth submanifold $M \subset \mathcal{M}$. The restriction g of G to M defines a metric on M and turns M into a Riemannian manifold too. At any point $x \in M$, the tangent space $T_x\mathcal{M}$ can be decomposed into a G-orthogonal sum $T_xM \oplus N_xM$. Here N_xM is defined as the G-orthogonal complement of T_xM in $T_x\mathcal{M}$ and is called the *normal subspace* to M at x. The submanifold M is called *totally geodesic* if every M-geodesic is also an $\mathcal{M}$-geodesic.

The *extrinsic curvature* II of a submanifold M measures how much M deviates from a totally geodesic submanifold. Given a point $x \in M$ and a tangent vector u to M at x, one may compare the M-geodesic to the $\mathcal{M}$-geodesic through x with velocity u. It turns out that these geodesics differ by a vector-valued term whose leading part is quadratic in u and points in a direction orthogonal to M at x. Technically, the extrinsic curvature at x is a symmetric bilinear map $II : T_xM \times T_xM \to N_xM$.

The Christoffel symbols in equation 5 are defined by equation 6 derived from the metric g. One may define the covariant derivative by similar symbols ∇^l_{kj} which are not linked to a metric, in which case one speaks of an *affine connection*. The affine connection generated by the metric is called the *Levi-Civita connection* and is denoted by $\hat{\nabla}$.

A prime example of a connection not derived from a metric is the connection of a real vector space V, usually called D, which in any coordinate system derived from a basis of V is given by the symbols $D^l_{kj} = 0$. Its parallel vector fields are just the constant vector fields.

2.2 Affine hypersurface immersions

Affine hypersurface immersions have first been studied in the pioneering works Tzitzeica (1908) and Blaschke & Reidemeister (1923). This branch of differential geometry studies the structures generated by the connection D of V on smooth hypersurface immersions $M \to V$.

Let M be an $(n-1)$-dimensional differentiable manifold and $f : M \to V$ a smooth hypersurface immersion. Let $\xi : M \to \mathbb{R}^n$ be a smooth transversal vector field on M, i.e., such that at any point $y \in M$, every tangent vector $u \in T_{f(y)}\mathbb{R}^n$ can be decomposed into a sum of a tangential component in $f_*[T_yM]$ and a component parallel to $\xi(y)$. Then there exists a unique affine connection ∇ and a unique symmetric $(0,2)$-tensor field h on M such that for every ∇-geodesic $\sigma(t)$ on M we have

$$\frac{d^2 f(\sigma(t))}{dt^2} + h(\dot{\sigma}(t), \dot{\sigma}(t)) \cdot \xi(\sigma(t)) = 0. \tag{7}$$

This can be written equivalently as $D - \nabla = -h \otimes \xi$, and ∇ can be interpreted as projection of the connection D onto the tangent space to M along ξ. The tensor h is called the *affine fundamental form*, and ∇ is called the *induced connection*. If h is non-degenerate everywhere on M, then f is called *non-degenerate* and h is called the *affine metric*.

Given a hypersurface immersion $f : M \to \mathbb{R}^n$ and a transversal vector field $\xi : M \to \mathbb{R}^n$, we may define another hypersurface immersion $\nu : M \to V^*$ into the dual space, as follows. For $y \in M$ we define the image $\nu(y)$ as the vector $v \in V^*$ which is zero on the tangent space $f_*[T_yM]$ and 1 on the transversal vector $\xi(y)$. It turns out that this *conormal map* defines the same quadratic form h

on M, but a different affine connection $\bar{\nabla}$, which is called the *dual affine connection*. The conormal map of the immersion ν is the original immersion f, so the correspondence $f \leftrightarrow \nu$ is a duality.

The induced connection ∇ on M does in general not coincide with the Levi-Civita connection $\hat{\nabla}$ of the affine metric h. The covariant derivative $C_{\alpha\beta\gamma} = \nabla_\alpha h_{\beta\gamma}$ is called the *cubic form*.

We are interested in *hyperbolic centro-affine hypersurface immersions* $f : M \to \mathbb{R}^n$, whose transversal vector field is defined by the relation $\xi = -f$ and whose affine metric is positive definite. This implies that the image of the immersion is convex such that its convex hull and the origin lie on opposite sides of the surface. For centro-affine immersions the cubic form C is totally symmetric.

An excellent general reference on affine differential geometry is the book Nomizu & Sasaki (1994). Centro-affine hypersurface immersions are treated in detail in its first chapter.

2.3 Para-Kähler spaces

A para-Kähler manifold $\mathcal{M}$ is a $2n$-dimensional manifold equipped with three objects: a pseudo-Riemannian metric G, a symplectic form ω, and a linear involution J acting at each point $x \in \mathcal{M}$ on the tangent space $T_x\mathcal{M}$, satisfying the following conditions. For arbitrary tangent vectors $u, v \in T_x\mathcal{M}$ we have

$$G(u, v) = \omega(Ju, v), \qquad \omega(u, v) = g(Ju, v),$$

the covariant derivative $\hat{\nabla}\omega$ vanishes, and the eigenspaces of J for the eigenvalues ± 1 form integrable distributions, i.e., there are submanifolds of $\mathcal{M}$ with these eigenspaces as tangent spaces.

A simple example is the product $\mathcal{M} = V \times V^*$ of a real vector space V with its dual. Here

- G is defined by $G((u, v), (u', v')) = \frac{\langle u', v\rangle + \langle u, v'\rangle}{2}$
- ω is defined by $\omega((u, v), (u', v')) = \frac{\langle u', v\rangle - \langle u, v'\rangle}{2}$
- $J : (u, v) \mapsto (u, -v)$

The primal and dual subspace hence appear as the eigenspaces of the inversion J with eigenvalues ± 1. Swapping these eigenspaces then corresponds to a multiplication of J by 1. In the theory of para-Kähler spaces J is associated to the *para-complex unit* e, which is similar to the imaginary unit i, but instead of $i^2 = -1$ satisfies the relation $e^2 = 1$. Swapping the primal and dual space is hence equivalent to the para-complex conjugation. The squared distance between two points $(u, v), (u', v') \in \mathcal{M}$ in this space is given by

$$G((u - u', v - v'), (u - u', v - v')) = \langle u - u', v - v'\rangle. \tag{8}$$

A *Lagrangian submanifold* of $\mathcal{M}$ (of dimension $2n$) is a submanifold M of dimension n such that the restriction of the symplectic form ω on M vanishes. The Lagrangian manifold inherits the metric of the ambient space $\mathcal{M}$, denote it by g. We then may define the extrinsic curvature II on M, measuring the deviation of M from a totally geodesic submanifold. Recall that II is a bilinear form on the tangent subspace TM to M with values in the normal subspace NM. Now for Lagrangian manifolds the map J takes NM bijectively to TM. Hence the composition $J \circ II$ is a bilinear map from TM to TM, i.e., a $(1, 2)$ tensor field. It is well-known that lowering the contra-variant index of this tensor by means of the metric g leads to a totally symmetric tri-linear form (see (Chen, 2011, Section 3)). In other words, let $x \in M$ be a point and $u, v, w \in T_xM$ three tangent vectors. Then the tri-linear map

$$\omega_{II} : (u, v, w) \mapsto g((J \circ II)(u, v), w) \tag{9}$$

is totally symmetric.

A submanifold $M \subset \mathcal{M}$ is called *minimal* if the convolution of its extrinsic curvature with the inverse of its metric g vanishes.

3 Barriers from the viewpoint of affine differential geometry

In this section we establish how logarithmically homogeneous barriers relate to centro-affine geometry, which deals with centro-affine hypersurfaces in V.

3.1 SPLITTING OF THE METRIC

The domain on which F is defined is the interior K^o of a regular convex cone, which naturally splits into a disjoint union of rays. Now equation 3 implies that for every $\alpha \in \mathbb{R}$, there exists exactly one point x on each ray such that $F(x) = \alpha$. Hence the domain splits also into a disjoint union of mutually homothetic level hypersurfaces of F. The whole domain can hence be considered as a direct product of $\mathbb{R}$ and a level surface of F. The key result which relates barriers to centro-affine immersions is that the Hessian metric F'' splits accordingly into a direct product of a trivial part and a part proportional to the affine metric.

Let $F : K^o \to \mathbb{R}$ be a smooth locally strongly convex function on the interior of a regular convex cone K satisfying the logarithmic homogeneity equation 3 with some $\nu > 0$. Let $F_\alpha = \{x \in K^o \,|\, F(x) = \alpha\}$ be the level surfaces of F. For every $\alpha \in \mathbb{R}$, define a diffeomorphism $I_\alpha : F_\alpha \times \mathbb{R}_+ \to K^o$ by $I_\alpha : (x, \beta) \mapsto \beta x$.

Proposition 3.1. *(Loftin, 2002, Theorem 1, p.428) The Riemannian manifold (K^o, F''), consisting of the interior of K equipped with the Hessian metric generated by F, is isometric under I_α^{-1} to the product $(F_\alpha, \nu h) \times (\mathbb{R}_+, \nu\beta^{-2} d\beta^2)$, where h is the centro-affine metric of the inclusion $F_\alpha \hookrightarrow V$, considered as centro-affine immersion.*

The Hessian metric F'' hence splits into a direct product of a trivial radial part and a non-trivial transversal part, which turns out to be proportional to the affine metric defined by the level hypersurfaces of F, considered as centro-affine immersions. The splitting result, which itself follows from logarithmic homogeneity, has an interesting consequence. Since the co-vector field F' has constant length $\sqrt{\nu}$ in the Hessian metric F'' (Nesterov & Nemirovskii, 1994, Prop. 2.3.4), is everywhere orthogonal to the non-trivial factor, and constant length vector fields in one dimension are parallel, this field must be parallel with respect to the connection defined by the metric F''. Interestingly, the converse is also true in the following sense.

Theorem 3.2. *(Hildebrand, 2014b, Theorem 4.1) Let $F : U \to \mathbb{R}$ be a C^3 function defined on some simply connected domain $U \subset \mathbb{A}^n$ in an affine space. Suppose that F has a non-degenerate Hessian and denote by $\hat{\nabla}$ the Levi-Civita connection of the Hessian metric F''. Then the first derivative F' is $\hat{\nabla}$-parallel if and only if F is locally logarithmically homogeneous with some homogeneity parameter ν with respect to some central point $c \in \mathbb{A}^n$.*

3.2 SELF-CONCORDANCE AND CUBIC FORM

We now consider how the cubic form C on the level surface relates to the third derivative F'''.

Lemma 3.3. *Assume the notations and conditions of Section 3.1. Then for every $\alpha \in \mathbb{R}$, the cubic form C of the inclusion $F_\alpha \hookrightarrow V$, considered as centro-affine immersion, equals the restriction of the symmetric $(0, 3)$ tensor field $\nu^{-1} F'''$ on F_α.*

Proof. Let $y \in F_\alpha$ be an arbitrary point and let $u \in T_y F_\alpha$ be a tangent vector at y. We shall now compute the value $C[u, u, u]$ of the cubic form on this vector. Let ∇ be the induced centro-affine connection on F_α and D the canonical affine connection on V. We then have $C = \nabla h = \nu^{-1} \nabla F''$ by the definition of the cubic form and by Proposition 3.1.

Let $\sigma(t)$ be the ∇-geodesic through $y = \sigma(0)$ with velocity u. Then by definition $u(t) = \dot{\sigma}(t)$ is the vector obtained by ∇-parallel transport of u from y to $\sigma(t)$ along σ. Applying equation 7 we get

$$\dot{u}(t) - h(u(t), u(t)) \cdot \sigma(t) = 0. \tag{10}$$

Let further $\gamma(t) = F''(\sigma(t))[u(t), u(t)]$ be the value of F'' on the vector $u(t)$. On the one hand, we get by differentiation

$$\begin{aligned} \dot{\gamma}(t) &= F'''(\sigma(t))[\dot{\sigma}(t), u(t), u(t)] + 2F''(\sigma(t))[\dot{u}(t), u(t)] \\ &= F'''(\sigma(t))[u(t), u(t), u(t)] + 2h(u(t), u(t)) F''(\sigma(t))[\sigma(t), u(t)] \\ &= F'''(\sigma(t))[u(t), u(t), u(t)]. \end{aligned}$$

Here the second equality comes from equation 10, and the third equality holds because the position vector $\sigma(t)$ and the tangent vector $u(t)$ are orthogonal by virtue of Proposition 3.1. On the other hand, we have

$$\dot{\gamma}(t) = \nabla_{u(t)} \gamma(t) = (\nabla F'')(\sigma(t))[u(t), u(t), u(t)] = \nu C(\sigma(t))[u(t), u(t), u(t)].$$

Here the first equality holds because the covariant derivative of a scalar equals its partial derivative. The second equality holds by the Leibniz rule, taking into account that the ∇-derivative of u vanishes because $u(t)$ is ∇-parallel. The third equality comes from the definition of the cubic form C.

Equalling the two expressions for $\dot\gamma(t)$ completes the proof. □

Hence, just as the affine metric h is proportional to the restriction of the Hessian F'' on the level surface of F, the cubic form C is proportional to the restriction of the third derivative F''' to this level surface. In particular, self-concordance of F implies that the cubic form is uniformly bounded when measured in the metric h. However, the converse implication is not immediately evident, because the affine metric and the cubic form on the level surface F_α determine the derivatives F'', F''' only on those vectors which are tangent to F_α. The logarithmic homogeneity condition, however, allows to obtain the following result.

Theorem 3.4. *(Hildebrand, 2022, Theorem 1.5) Let $F : K^o \to \mathbb{R}$ be a logarithmically homogeneous convex C^3 function with homogeneity degree ν on a convex cone $K \subset V$, $n \geq 2$. Then F is self-concordant if and only if the cubic form on the level surfaces of F has ∞-norm not exceeding 2ϑ with $\vartheta = \frac{\nu-2}{\sqrt{\nu-1}}$ as measured in the affine metric.*

As a consequence, F is self-concordant if and only if the cubic form is bounded when measured in the affine metric h, and the norm of C is in an explicit monotone correspondence with the self-concordance parameter ν.

The pair (h, C) plays a fundamental role in affine differential geometry. By the fundamental uniqueness result (Nomizu & Sasaki, 1994, Theorem 8.1) the centro-affine hypersurface and hence the barrier and the underlying cone can be recovered from this pair up to linear isomorphism.

3.3 Legendre duality and the conormal map

In this subsection we show that Legendre duality for barriers is closely linked to the duality in affine differential geometry defined by the conormal map. By applying the first order optimality condition to the maximization problem in Definition 1.2 we obtain the condition $s = -F'(x)$. Let us denote the corresponding map from K^o to $\mathbb{R}_n$ by Φ, $\Phi(x) = -F'(x)$. By (Nesterov & Nemirovskii, 1994, Theorem 2.4.4) the image of Φ is exactly the interior of the dual cone K^*. Positive definiteness of the Hessian F'' implies that Φ is actually a bijection between the interiors of K and K^*. Moreover, it is an isometry when the interiors of K, K^* are equipped with the Hessian metrics F'', F''_*, respectively (Nesterov & Nemirovskii, 1994, p.45),Nesterov & Todd (2002). By the Definition 1.2 of the dual barrier we have

$$F(x) + F_*(\Phi(x)) = -\langle x, \Phi(x)\rangle = \langle x, F'(x)\rangle = -\nu, \tag{11}$$

and hence Φ maps level surfaces of F to level surfaces of F_*. As a consequence, the isometry Φ preserves the product structure in Proposition 3.1, but as is easily seen from equation 11, the orientation of the rays is reversed: a ray pointing away from the origin in the primal cone is mapped to a ray pointing towards the origin in the dual cone.

Let us now choose $\alpha \in \mathbb{R}$ and consider the level surface $F_\alpha = \{x \in K^o \mid F(x) = \alpha\}$. By definition the conormal map of the hypersurface immersion $F_\alpha \hookrightarrow \mathbb{R}^n$ maps the point $x \in F_\alpha$ to $p \in \mathbb{R}_n$ such that p is proportional to $F'(x)$ and $\langle p, -x\rangle = 1$. From the identity $F'(x)[x] = -\nu$ we obtain the explicit expression $p = \nu^{-1}F'(x) = -\nu^{-1}\Phi(x)$. We get the following result.

Lemma 3.5. *Let $F : K^o \to \mathbb{R}$ be a self-concordant barrier on a regular convex cone $K \subset \mathbb{R}^n$ with parameter ν. Let F_α be a level surface of F. Then the conormal map of the hypersurface immersion $F_\alpha \hookrightarrow \mathbb{R}^n$ is given by $-\nu^{-1}\Phi$, where Φ is the isometry between the interiors of K and K^* defined by Legendre duality.*

3.4 Self-scaled barriers and parallel cubic forms

Self-scaled barriers are a special kind of barriers existing only on symmetric cones. This property can be described by a parallelism condition on the third derivative F''', much like logarithmic homogeneity is described by a parallelism condition on F'. Therefore from the viewpoint of affine differential geometry the self-scaled barriers are the simplest class of self-concordant logarithmically homogeneous barriers after the familiar log-quadratic barrier on the Lorentz cone, which is characterized by the vanishing of the cubic form itself. We have the following results (Hildebrand, 2014b, Section 5), (Hildebrand, 2015, Section 2).

Theorem 3.6. *Let $K \subset V$ be a regular convex cone and $F : K^o \to \mathbb{R}$ a logarithmically homogeneous self-concordant barrier on K. Then the following conditions are equivalent:*

(a) F is self-scaled;

(b) F satisfies $\hat{D}F''' = 0$, where $\hat{D}$ is the Levi-Civita connection of the Hessian metric F'';

(c) the cubic form C on the level surfaces of F is parallel with respect to the Levi-Civita connection $\hat{\nabla}$ of the centro-affine metric h on these level surfaces.

This provides also a radically different view on self-scaled barriers. Initially they are defined by a global algebraic property, while the parallelism condition is local and geometric. The condition $\hat{D}F''' = 0$ amounts to a quasi-linear fourth order PDE (Hildebrand, 2014b, eq. (2)).

4 Lagrangian submanifolds

Due to the symmetry of the Hessian F'' the gradient graph M of a C^2 function $F : V \supset D \to \mathbb{R}$ is a Lagrangian submanifold of the para-Kähler space $\mathcal{M} = V \times V^*$. Moreover, the metric g induced by G on M coincides with the Hessian metric F'' on D. The scaling point then has the following interpretation.

Lemma 4.1. *Let $D \subset V$ be a conic domain, F a logarithmically homogeneous C^2 function on D, $x \in D$, $s = -F'(x)$, and let $M = \{(y, -F'(y)) \mid y \in D\}$. Let further $w \in D$ be such that $F''(w)x = s$. Then $(w, -F'(w)) \in M$ is a stationary point of the distance function from (x, s) on M, measured in the metric G of the para-Kähler space $\mathcal{M}$.*

Proof. The squared distance from (x, s) to an arbitrary point $(y, -F'(y)) \in M$ is given by

$$\langle x - y, s + F'(y) \rangle.$$

due to equation 8. Differentiating with respect to y, we obtain the condition

$$-s + F''(y)x - F''(y)y - F'(y) = F''(y)x - s = 0$$

for stationary points. Here $F''(y)y + F'(y) = 0$ due to logarithmic homogeneity. □

Hence if (x, s) is close enough to the manifold M, the scaling point can be interpreted as the closest point on M to (x, s).

The connection between Lagrangian submanifolds of the mentioned para-Kähler space and centro-affine immersions has been elaborated in Hildebrand (2011b), and in a more general setup in Hildebrand (2011a). In particular, the totally symmetric form defined by equation 9 is related to the cubic form C, and the metric g to the affine metric h.

5 Canonical barrier

A classical family of hypersurfaces studied already in Blaschke & Reidemeister (1923) are the *affine spheres*. A central conjecture that has been proven by many authors over decades is the *Calabi conjecture*, which states that the interior of every regular convex cone K can be foliated by a unique homothetic family of complete hyperbolic affine spheres. It is then not far-fetched to construct a logarithmically homogeneous function $F : K^o \to \mathbb{R}$ which has exactly these affine spheres as level surfaces and to try it as a candidate for a barrier. Hildebrand (2014c) and Fox (2015) independently showed that such a function is indeed a self-concordant barrier with parameter $\nu \leq n$, now known as the *canonical barrier*.

The canonical barrier is a universal construction. It shares with the universal barrier all invariance properties, but in addition it behaves well under duality. It can be represented as the solution of the PDE $\log \det F'' = 2F$ with boundary condition $F|_{\partial K} = +\infty$. On homogeneous cones its level surfaces equal those of the universal barrier, and on this class of cones the two barriers are essentially the same object. That the universal barrier on homogeneous cones satisfies above PDE has been shown already in (Güler, 1996, Theorem 4.4, p.868). Güler also conjectured that the solution of this PDE defines a self-concordant barrier for general cones.

A different characterization of canonical barriers is that they correspond exactly to the complete minimal Lagrangian surfaces in the para-Kähler space of Hildebrand (2011b).

5.1 Semi-homogeneous cones

If K has a rich automorphism group, then one may use this group to reduce the dimension of the PDE defining the canonical barrier, loosely speaking by considering it as an equation on the orbits with respect to the group. If the orbits in the interior $\operatorname{int} K$ form a 1-parametric family, then the PDE reduces to an ODE. One may then attempt to solve this ODE explicitly. In the case of 3-dimensional cones this program has been successfully accomplished, yielding analytic expressions for the canonical barrier on those cones whose automorphism group has dimension at least two. We call such cones *semi-homogeneous*. Hildebrand (2014a) gave the following classification.

Theorem 5.1. *Let $K \subset \mathbb{R}^3$ be a regular convex cone such that $\dim \operatorname{Aut} K \geq 2$. Then K is linearly isomorphic to exactly one of the following cones.*

1. *the cone obtained by the homogenization of the epigraph of the exponential function,*
2. *the positive orthant $\mathbb{R}^3_+$,*
3. *the cone given by $\{x \mid z \geq -x^{1/p}y^{1/q},\ x \geq 0,\ y \geq 0\}$ for some $p \in [2, \infty)$, $\frac{1}{p} + \frac{1}{q} = 1$,*
4. *the cone given by $\{x \mid -\alpha x^{1/p}y^{1/q} \leq z \leq x^{1/p}y^{1/q},\ x \geq 0,\ y \geq 0\}$ for some $p \in [2, \infty)$, $\frac{1}{p} + \frac{1}{q} = 1$, $\alpha \in (0, 1]$,*
5. *the cone given by $\{x \mid 0 \leq z \leq x^{1/p}y^{1/q},\ x \geq 0,\ y \geq 0\}$ for some $p \in [2, \infty)$, $\frac{1}{p} + \frac{1}{q} = 1$.*

In general, the canonical barrier on these cones can be expressed by Weierstrass special functions Lin & Wang (2016), with simpler expressions for the boundary cases Hildebrand (2014c), Hildebrand (2014a). Case 4 with $\alpha = 1$ corresponds to the well-known power cone which is nowadays available for modeling problems solvable by some conic solvers. The barrier parameter $\nu = \frac{3\max(p,q)}{\max(p,q)+1}$ of the canonical barrier on this cone is better than the best-known values $\nu = 3$ (analytically) and $\nu = 3 - \frac{2}{\max(p,q)}$ (numerically) reported in (Chares, 2008, Section 3.1).

6 Conclusions

We have presented three different viewpoints on logarithmically homogeneous barriers used in conic optimization. The following table gives an overview of the relations established in this paper.

self-concordant barriers	centro-affine immersions	para-Kähler spaces
barrier $F : \operatorname{int} K \to \mathbb{R}$	level surface $\{x \mid F(x) = 0\}$	gradient graph $\{(x, F'(x))\}$
Legendre duality	conormal map	para-complex conjugation
Hessian metric	centro-affine metric	gradient graph submanifold metric
third derivative	cubic form	extrinsic curvature
self-concordance condition	bound on cubic form	bound on extrinsic curvature
self-concordance parameter	∞-norm of the cubic form	∞-norm of extrinsic curvature
canonical barrier	affine hypersphere	minimal Lagrangian submanifold
self-scaled barrier	parallel cubic form	parallel extrinsic curvature

References

David Allen Bayer and Jeffrey Clark Lagarias. The nonlinear geometry of linear programming II. Legendre transform coordinates and central trajectories. *Trans. Amer. Math. Soc.*, 314(2):527–581, 1989.

Wilhelm Blaschke and Kurt Reidemeister. *Vorlesungen über Differentialgeometrie II, Affine Differentialgeometrie*. Springer, Berlin, 1923.

Sébastien Bubeck and Ronen Eldan. The entropic barrier: Exponential families, log-concave geometry, and self-concordance. *Math. Oper. Res.*, 44(1):264–276, 2019.

Robert Chares. *Cones and Interior-Point Algorithms for Structured Convex Optimization involving Powers and Exponentials*. PhD thesis, Université Catholique de Louvain, Louvain-la-Neuve, 2008.

Bang-Yen Chen. Lagrangian H-umbilical submanifolds of para-Kähler manifolds. *Taiwanese J. Math.*, 15(6):2483–2502, 2011.

Leonid Faybusovich. Jordan algebras, symmetric cones and interior-point methods. Technical report, University of Notre Dame, 1995.

Daniel J. Fox. A Schwarz lemma for Kähler affine metrics and the canonical potential of a proper convex cone. *Ann. Mat. Pur. Appl.*, 194(1):1–42, 2015.

Osman Güler. Barrier functions in interior point methods. *Math. Oper. Res.*, 21(4):860–885, 1996.

Raphael A. Hauser. Self-scaled barrier functions: decomposition and classification. Technical Report Numerical Analysis Report DAMTP 1999/NA13, Dept. of Applied Mathematics and Theoretical Physics, Cambridge, 1999.

Raphael A. Hauser. Self-scaled barriers for semidefinite programming. Technical Report Numerical Analysis Report DAMTP 2000/NA02, Dept. of Applied Mathematics and Theoretical Physics, Cambridge, 2000a.

Raphael A. Hauser. *On search directions for self-scaled conic programming*. PhD thesis, School of Operations Research and Industrial Engineering, Cornell University, 2000b.

Raphael A. Hauser and Yongdo Lim. Self-scaled barriers for irreducible symmetric cones. *SIAM J. Optimiz.*, 12(3):715–723, 2002.

Roland Hildebrand. Half-dimensional immersions in para-Kähler manifolds. *Int. Electron. J. Geom.*, 4(2):85–113, 2011a.

Roland Hildebrand. The cross-ratio manifold: a model of centro-affine geometry. *Int. Electron. J. Geom.*, 4(2):32–62, 2011b.

Roland Hildebrand. Analytic formulas for complete hyperbolic affine spheres. *Contributions to Algebra and Geometry*, 55(2):497–520, 2014a.

Roland Hildebrand. Hessian potentials with parallel derivatives. *Results Math.*, 65(3-4):399–413, 2014b.

Roland Hildebrand. Canonical barriers on convex cones. *Math. Oper. Res.*, 39(3):841–850, 2014c.

Roland Hildebrand. Centro-affine hypersurface immersions with parallel cubic form. *Contributions to Algebra and Geometry*, 56(2):593–640, 2015.

Roland Hildebrand. Projectively self-concordant barriers. *Math. Oper. Res.*, 47(3):2444–2463, 2022.

Satoshi Kakihara, Atsumi Ohara, and Takashi Tsuchiya. Information geometry and interior-point algorithms in semidefinite programs and symmetric cone programs. *J. Optim. Theory Appl.*, 157 (3):749–780, 2013.

Narendra Karmarkar. Riemannian geometry underlying interior-point methods for linear programming. In Jeffrey Clark Lagarias and Michael J. Todd (eds.), *Mathematical Developments Arising from Linear Programming*, volume 114 of *Contemporary Mathematics*, pp. 51–75. AMS, Providence, Rhode Island, 1988.

Jeffrey Clark Lagarias. The nonlinear geometry of linear programming III. Projective Legendre transform coordinates and Hilbert geometry. *Trans. Amer. Math. Soc.*, 320(1):193–225, 1990.

Zhicheng Lin and Erxiao Wang. The associated families of semi-homogeneous complete hyperbolic affine spheres. *Acta Math. Sci. (English Ed.)*, 36(3):765–781, 2016.

John C. Loftin. Affine spheres and Kähler-Einstein metrics. *Math. Res. Lett.*, 9(4):425–432, 2002.

Yuri Nesterov. Towards nonsymmetric conic optimization. Discussion paper 2006/28, CORE, Louvain-la-Neuve, 2006.

Yuri E. Nesterov and Michael J. Todd. Self-scaled cones and interior-point methods in nonlinear programming. Discussion paper 1994/62, CORE, Louvain-la-Neuve, 1994. URL `http://alfresco.uclouvain.be/alfresco/service/guest/streamDownload/workspace/SpacesStore/93180cab-8ddc-4dcc-94c7-a065e4f8ecef/coredp_1994_62.pdf?guest=true`.

Yuri E. Nesterov and Michael J. Todd. Self-scaled barriers and interior-point methods for convex programming. *Math. Oper. Res.*, 22:1–42, 1997.

Yuri E. Nesterov and Michael J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM J. Optimiz.*, 8(2):324–364, 1998.

Yuri E. Nesterov and Michael J. Todd. On the Riemannian geometry defined by self-concordant barriers and interior-point methods. *Found. Comput. Math.*, 2:333–361, 2002.

Yurii Nesterov. Towards non-symmetric conic optimization. *Optim. Method Softw.*, 27(5):893–917, 2012.

Yurii Nesterov and Arkadii Nemirovskii. Self-concordant functions and polynomial time methods in convex programming. Technical report, Centr. Econ. & Math. Inst., USSR Acad. Sci., Moscow, 1989.

Yurii Nesterov and Arkadii Nemirovskii. *Interior-point Polynomial Algorithms in Convex Programming*, volume 13 of *SIAM Stud. Appl. Math.* SIAM, Philadelphia, 1994.

Katsumi Nomizu and Takeshi Sasaki. *Affine Differential Geometry: Geometry of Affine Immersions*, volume 111 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1994.

Atsumi Ohara. Information geometric analysis of an interior point method for semidefinite programming. In O. Barndorff-Nielsen and E. Jensen (eds.), *Geometry in Present Day Science*, pp. 49–74. World Scientific, 1999.

Atsumi Ohara and Takashi Tsuchiya. An information geometric approach to polynomial-time interior-point algorithms: complexity bound via curvature integral. Technical Report 1055, Institute of Statistical Mathematics, Tokyo, 2007. Research Memorandum.

Stefan Hans Schmieta. Complete classification of self-scaled barrier functions. Technical Report CORC TR-2000-01, Columbia University, New York, 2000.

Hirohiko Shima. *The Geometry of Hessian structures*. World Scientific, Hackensack, New Jersey, 2007.

Michael Spivak. *A Comprehensive Introduction to Differential Geometry*, volume 1. Publish or Perish, Houston, 3rd edition, 1999.

Gheorghe Tzitzeica. Sur une nouvelle classe de surfaces. *Rend. Circ. Mat. Palermo*, 25:180–187, 1908.

Frank Warner. *Foundations of Differentiable Manifolds and Lie Groups*, volume 94 of *Graduate Texts in Mathematics*. Springer, New York, Berlin, Heidelberg, Tokyo, 1983.