

# Radial Duality

## Part I: Foundations

Benjamin Grimmer\*

### Abstract

Renegar [11] introduced a novel approach to transforming generic conic optimization problems into unconstrained, uniformly Lipschitz continuous minimization. We introduce *radial transformations* generalizing these ideas, equipped with an entirely new motivation and development that avoids any reliance on convex cones or functions. Perhaps of greatest practical importance, this facilitates the development of new families of projection-free first-order methods applicable even in the presence of nonconvex objectives and constraint sets. Our generalized construction of this radial transformation uncovers that it is dual (i.e., self-inverse) for a wide range of functions including all concave objectives. This gives a powerful new duality relating optimization problems to their radially dual problem. For a broad class of functions, we characterize continuity, differentiability, and convexity under the radial transformation as well as develop a calculus for it. This radial duality provides a strong foundation for designing projection-free radial optimization algorithms, which is carried out in the second part of this work.

## 1 Introduction

Renegar [11] introduced a framework for conic programming (and by reduction, convex optimization), which turns such problems into uniformly Lipschitz optimization. After being radially transformed, a simple subgradient method can be applied and analyzed. Notably, even for constrained problems, such an algorithm maintains a feasible solution at each iteration while avoiding the use of orthogonal projections, which can often be a bottleneck for first-order methods. Subsequently, Grimmer [8] showed that in the case of convex optimization, a simplified radial subgradient method can be applied with simpler and stronger convergence guarantees. In [12], Renegar further showed that the transformation of hyperbolic cones is amenable to the application of smoothing techniques, offering notable improvements over radial subgradient methods.

In this paper, we provide a wholly different development and generalization of the ideas behind Renegar's framework, which avoids relying on convex cones or functions as the central object. Instead our approach is based on the following simple projective transformation, which we dub the *radial point transformation*,

$$\Gamma(x, u) = (x, 1)/u$$

---

\*School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14850, USA; [people.orie.cornell.edu/bdg79/](http://people.orie.cornell.edu/bdg79/)

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1650441. This work was partially done while the author was visiting the Simons Institute for the Theory of Computing. It was partially supported by the DIMACS/Simons Collaboration on Bridging Continuous and Discrete Optimization through NSF grant #CCF-1740425.

for any  $(x, u) \in \mathcal{E} \times \mathbb{R}_{++}$ , where  $\mathcal{E}$  is some finite-dimensional Euclidean space and  $\mathbb{R}_{++}$  is the set of positive real numbers. Applying this elementwise to a set  $S \subseteq \mathcal{E} \times \mathbb{R}_{++}$  gives the *radial set transformation*, denoted by

$$\Gamma S = \{\Gamma(x, u) \mid (x, u) \in S\}.$$

To motivate the nomenclature of calling these transformations radial, consider the transformation of a vertical line in  $\mathcal{E} \times \mathbb{R}_{++}$ : for any  $x \in \mathcal{E}$ ,

$$\Gamma\{(x, \lambda) \mid \lambda \in \mathbb{R}_{++}\} = \{\gamma(x, 1) \mid \gamma \in \mathbb{R}_{++}\}.$$

We see that this transformation maps vertical lines into rays extending from the origin (and rays into vertical lines since the point transformation is an involution (i.e., dual)  $\Gamma\Gamma(x, u) = (x, u)$ ).

To extend this set operation to apply to functions, we consider functions  $f: \mathcal{E} \rightarrow \mathbb{R}_{++} \cup \{0, \infty\}$  mapping into the extended positive reals. Then we define the *upper and lower radial function transformations* of  $f$  as<sup>1</sup>

$$f^\Gamma(y) = \sup\{v > 0 \mid (y, v) \in \Gamma(\text{epi } f)\},$$

$$f_\Gamma(y) = \inf\{v > 0 \mid (y, v) \in \Gamma(\text{hypo } f)\}$$

where  $\text{epi } f = \{(x, u) \in \mathcal{E} \times \mathbb{R}_{++} \mid f(x) \leq u\}$  denotes the epigraph of  $f$  and  $\text{hypo } f = \{(x, u) \in \mathcal{E} \times \mathbb{R}_{++} \mid f(x) \geq u\}$  denotes the hypograph of  $f$ . The upper transformation has the interpretation of reshaping the epigraph of  $f$  via the radial set transformation and returning the smallest function whose hypograph contains that set. Likewise, the lower transformation aims to turn the hypograph of  $f$  into the epigraph of a new function. We find that for a wide range of functions, the duality of the point and set transformations carries over to the function transformations (see Theorem 3.2):

$$f^{\Gamma\Gamma} = f.$$

**Connections To Prior Works.** Noting  $f^\Gamma(y) = \sup\{v > 0 \mid v \cdot f(y/v) \leq 1\}$ , this relates to the transformation used by Grimmer [8] and those of Renegar [11, 12]. The upper and lower transformations coincide in the convex settings of these previous works but may diverge in the general setting considered herein. For our analysis, we primarily focus on the upper transformation, but equivalent results always hold for the lower transform. Artstein-Avidan and Milman [4] (and the subsequent [3]) consider the same underlying projective point transformation  $\Gamma$  and the similar but quite different function transformation  $\inf\{v > 0 \mid (y, v) \in \Gamma(\text{epi } f)\}$ . Considering this transformation limits their theory to the restrictive setting of nonnegative convex functions that minimize to value zero at the origin. As a result, their theory does not provide an interesting duality between optimization problems.

We remark on one other way to view the radial function transformation. Denote the *Minkowski gauge* of a set  $S \subseteq \mathcal{E} \times \mathbb{R}_{++}$  at some  $(y, v) \in \mathcal{E} \times \mathbb{R}_{++}$  by  $\gamma_S((y, v)) = \inf\{\lambda > 0 \mid (y, v) \in \lambda S\}$ . Then the lower radial transformation can be restated as the restriction of this gauge to  $v = 1$

$$f_\Gamma(y) = \gamma_{\text{hypo } f}((y, 1)).$$

This relationship to gauges motivates our notation of  $\Gamma$  to denote the radial transformation. From this point of view, a connection can be made between this radial framework and the perspective duality considered by Aravkin et al [1]. They extend the theory of gauge duality developed by Freund [7], which then applies to nonnegative convex functions by considering perspective functions.

---

<sup>1</sup>Since we are considering functions mapping into  $\mathbb{R}_{++} \cup \{0, \infty\}$ , if no  $v > 0$  satisfies  $(y, v) \in \Gamma(\text{epi } f)$ , we have the supremum defining  $f^\Gamma(y)$  equal zero.

The resulting perspective dual optimization problem minimizes the function  $\gamma_{\text{hypo } f^*}((y, v))$  where  $f^*$  is the Fenchel conjugate of  $f$ . Thus their perspective duality can be viewed as a combination of applying Fenchel duality and our radial machinery.

From this connection, we point out a key difference between this radial duality and these previous dualities. The classic theories of Lagrange and gauge duality are based on a conjugate or polar defined as a supremum over the dual vector space. In contrast, the radial dual and the Minkowski gauge are defined by a one-dimensional problem. This difference allows the radial dual to be efficiently computed numerically for generic problems using a linesearch or bisection, whereas evaluating the Fenchel conjugate of a function is as hard as optimizing over it.

**Our Contributions.** This work serves to establish the foundations of radial transformations as a new addition to the optimization toolbox. The second part of this work leverages this machinery to develop new radial optimization algorithms. Our development establishes this tool in the following three ways: (i) The radial transformation is dual and enjoys rich structure stemming from this. (ii) The radial transformation produces a new duality between nonnegative optimization problems. For example, constraints are dually transformed into gauges, which allow algorithms to replace orthogonal projections with potentially cheaper, one-dimensional linesearches. We refer any numerically or algorithmically inclined reader to the motivating example of quadratic programming at the start of the second part of this work [9] to see this machinery fully in action. (iii) The radial transformation is the unique operation of its kind.

**Duality of the radial transformation.** We precisely characterize the family of functions where the duality  $f^{\Gamma} = f$  holds through the star-convexity of their hypograph. Moreover, when this duality holds, we find that a number of important classes of functions are dual to each other or self-dual under the radial transformations. Namely,

$$\begin{aligned} f \text{ is upper semicontinuous} &\iff f^{\Gamma} \text{ is lower semicontinuous,} \\ f \text{ is continuous} &\iff f^{\Gamma} \text{ is continuous,} \\ f \text{ is concave} &\iff f^{\Gamma} \text{ is convex,} \\ f \text{ is quasiconcave} &\iff f^{\Gamma} \text{ is quasiconvex,} \\ f \text{ is } k \text{ times differentiable} &\iff f^{\Gamma} \text{ is } k \text{ times differentiable,} \\ f \text{ is analytic} &\iff f^{\Gamma} \text{ is analytic} \end{aligned}$$

under appropriate regularity conditions. We also derive a calculus for the radial transformations, providing formulas for the (sub)gradients and Hessians of  $f^{\Gamma}$  based on those of  $f$ .

**Radial duality between optimization problems.** For a wide range of functions,  $\Gamma(\text{epi } f)$  is the hypograph of another function, and so  $\text{hypo } f^{\Gamma} = \Gamma(\text{epi } f)$ . As a result, for such functions, points in  $\text{hypo } f$  and  $\text{epi } f^{\Gamma}$  can be directly related by the bijection  $\Gamma$  and its inverse (which is also  $\Gamma$ ). This relation also applies to the maximizers of  $f$  and the minimizers of  $f^{\Gamma}$ . Namely for any function  $f: \mathcal{E} \rightarrow \mathbb{R}_{++} \cup \{0, \infty\}$ , consider the primal problem

$$p^* = \max_{x \in \mathcal{E}} f(x). \tag{1}$$

Then the radially dual problem is given by

$$d^* = \min_{y \in \mathcal{E}} f^{\Gamma}(y) \tag{2}$$

and has  $\operatorname{argmax} f \times \{p^*\} = \Gamma(\operatorname{argmin} f^\Gamma \times \{d^*\})$  under certain regularity conditions (see Proposition 4.9). Hence maximizing  $f$  is equivalent to minimizing  $f^\Gamma$ .

The radially dual problem (2) can exhibit very different behavior than the original problem (1). For example, consider the function  $f(x) = \sqrt{1 - \|x\|^2}$  defined on the unit ball, which has arbitrarily large gradients and Hessians as  $x$  approaches the boundary of the ball. Despite this function's poor behavior,  $f^\Gamma(y) = \sqrt{1 + \|y\|^2}$  has full domain with gradients and Hessians bounded in norm by one everywhere. This structure is very appealing for the analysis of first-order optimization methods which tend to heavily rely on these quantities being bounded. The second part of this work utilizes such structure arising from the radial duality developed herein to propose and analyze projection-free radial optimization methods.

**Uniqueness of the radial transformation.** From our construction of the radial transformation, it is natural to ask if other interesting transformations of optimization problems can be given by reshaping the epigraph of a function. Under some basic assumptions (primarily that the reshaping is invertible and convexity preserving), there are only two transformations of this form, up to affine transformations: the trivial duality between maximizing a function and minimizing its negative and the nontrivial duality given by the radial transformation. These results are similar in spirit to the characterization of order isomorphisms by [3].

**Outline** Section 2 develops theory for the radial point and set transformations on  $\mathcal{E} \times \mathbb{R}_{++}$ . Informed by this, Section 3 derives the core theory establishing the radial function transformations. Then Section 4 develops the calculus and optimality relationships between the primal (1) and radial dual (2). Lastly, Section 5 shows that this radial framework is the unique transformation of nonnegative-valued optimization problems of its type.

## 1.1 Notation

We primarily consider sets in  $\mathcal{E} \times \mathbb{R}_{++}$ , which inherits the standard Euclidean inner product and norm from  $\mathcal{E} \times \mathbb{R}$ . Denote the ball of radius  $r > 0$  around a point  $(x, u) \in \mathcal{E} \times \mathbb{R}$  by

$$B((x, u), r) := \{(x', u') \in \mathcal{E} \times \mathbb{R} \mid \|(x', u') - (x, u)\| \leq r\}.$$

Further, denote orthogonal projection onto a closed set  $S \subseteq \mathcal{E} \times \mathbb{R}$  by

$$\operatorname{proj}_S((x, u)) := \operatorname{argmin}\{\|(x', u') - (x, u)\| \mid (x', u') \in S\}.$$

Note  $\operatorname{proj}_S$  is set valued and may not be a singleton if  $S$  is not convex.

We consider functions  $f: \mathcal{E} \rightarrow \overline{\mathbb{R}}_{++}$ , where  $\overline{\mathbb{R}}_{++} = \mathbb{R}_{++} \cup \{0, +\infty\}$  denotes the “extended positive reals”. Here 0 and  $+\infty$  are the limit objects of  $\mathbb{R}_{++}$ , mirroring the roles of  $-\infty$  and  $+\infty$  in the extended reals. The effective domain of such a function is denoted by

$$\operatorname{dom} f := \{x \in \mathcal{E} \mid f(x) \in \mathbb{R}_{++}\}.$$

Such functions relate to  $\mathcal{E} \times \mathbb{R}_{++}$  through their graphs, epigraphs, and hypographs

$$\begin{aligned} \operatorname{graph} f &:= \{(x, u) \in \mathcal{E} \times \mathbb{R}_{++} \mid f(x) = u\}, \\ \operatorname{epi} f &:= \{(x, u) \in \mathcal{E} \times \mathbb{R}_{++} \mid f(x) \leq u\}, \\ \operatorname{hypo} f &:= \{(x, u) \in \mathcal{E} \times \mathbb{R}_{++} \mid f(x) \geq u\}. \end{aligned}$$

We say a function  $f: \mathcal{E} \rightarrow \overline{\mathbb{R}}_{++}$  is upper (lower) semicontinuous if  $\text{hypo } f$  ( $\text{epi } f$ ) is closed with respect to  $\mathcal{E} \times \mathbb{R}_{++}$ . Equivalently, a function is upper semicontinuous if for all  $x \in \mathcal{E}$ ,  $f(x) = \limsup_{x' \rightarrow x} f(x')$  and lower semicontinuous if  $f(x) = \liminf_{x' \rightarrow x} f(x')$ .

We say a function  $f: \mathcal{E} \rightarrow \overline{\mathbb{R}}_{++}$  is concave (convex) if  $\text{hypo } f$  ( $\text{epi } f$ ) is convex. The set of *convex normal vectors* of a set  $S \subseteq \mathcal{E} \times \mathbb{R}$  at some  $(x, u) \in S$  is denoted by

$$N_S^C((x, u)) := \{(\zeta, \delta) \mid (\zeta, \delta)^T((x, u) - (x', u')) \geq 0 \ \forall (x', u') \in S\}.$$

Then the *convex subdifferential* of a function  $f$  at some  $x \in \text{dom } f$  is denoted by

$$\partial_C f(x) := \{\zeta \mid (\zeta, -1) \in N_{\text{epi } f}^C((x, f(x)))\}.$$

Likewise, the *convex supdifferential* of a function  $f$  at some  $x \in \text{dom } f$  is denoted by

$$\partial^C f(x) := \{\zeta \mid (-\zeta, 1) \in N_{\text{hypo } f}^C((x, f(x)))\}.$$

The elements of these differentials are referred to as convex subgradients or supgradients.

For sets and functions that are not convex, we consider the generalization given by proximal normals and differentials. The set of *proximal normal vectors* of a set  $S \subseteq \mathcal{E} \times \mathbb{R}$  at some  $(x, u) \in S$  is denoted by

$$N_S^P((x, u)) := \{(\zeta, \delta) \mid (x, u) \in \text{proj}_S((x, u) + \epsilon(\zeta, \delta)) \text{ for some } \epsilon > 0\}.$$

Then the *proximal subdifferential* of a function  $f$  at some  $x \in \text{dom } f$  is denoted by

$$\partial_P f(x) := \{\zeta \mid (\zeta, -1) \in N_{\text{epi } f}^P((x, f(x)))\}.$$

Likewise, the *proximal supdifferential* of a function  $f$  at some  $x \in \text{dom } f$  is denoted by

$$\partial^P f(x) := \{\zeta \mid (-\zeta, 1) \in N_{\text{hypo } f}^P((x, f(x)))\}.$$

The elements of these differentials are referred to as proximal subgradients or supgradients.

## 2 The Radial Set Transformation

We begin by observing a number of properties of the radial point and set transformations. Section 2.1 uses these to characterize the convex and proximal normal vectors of a radially transformed set. Then Section 2.2 concludes with a number of examples and pictures illustrating the radial set transformation. A careful understanding of this operation on sets forms the foundation for understanding the radial function transformation.

One can easily check the point transformation is a continuous analytic bijection on  $\mathcal{E} \times \mathbb{R}_{++}$ . Further, both the point and set transformations are dual since

$$\Gamma\Gamma(x, u) = \Gamma \frac{(x, 1)}{u} = \frac{(x/u, 1)}{1/u} = (x, u). \quad (3)$$

Now we observe a few basic properties of the set transformation on any pair of sets  $S, T \subseteq \mathcal{E} \times \mathbb{R}_{++}$ . First, since the point transformation is invertible (in fact, it is its own inverse), the set transformation preserves inclusions between sets, giving

$$S \subseteq T \iff \Gamma S \subseteq \Gamma T. \quad (4)$$

Furthermore, the radial set transformation distributes over unions and intersections, giving

$$\Gamma(S \cap T) = \Gamma S \cap \Gamma T, \quad (5)$$

$$\Gamma(S \cup T) = \Gamma S \cup \Gamma T. \quad (6)$$

Since the radial point transformation is a projective transformation, convex sets, halfspaces, and ellipsoids map into convex sets, halfspaces, and ellipsoids, respectively. We give direct proofs of these results in the appendix yielding simple formulas for radially dual halfspaces and ellipsoids in the latter two cases.

**Proposition 2.1.** *A set  $S \subseteq \mathcal{E} \times \mathbb{R}_{++}$  is convex if and only if  $\Gamma S$  is convex.*

In particular, consider the radial transformation of any halfspace in  $\mathcal{E} \times \mathbb{R}_{++}$ . Direct manipulation of its definition shows that the radial transformation of a halfspace is another halfspace.

**Proposition 2.2.** *A set  $S \subseteq \mathcal{E} \times \mathbb{R}_{++}$  is a halfspace if and only if  $\Gamma S$  is a halfspace. In particular, for any halfspace defined by*

$$S = \left\{ (x', u') \in \mathcal{E} \times \mathbb{R}_{++} \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix}^T \begin{bmatrix} x' - x \\ u' - u \end{bmatrix} \leq 0 \right\},$$

letting  $(y, v) = \Gamma(x, u)$ ,  $\Gamma S$  is the following halfspace

$$\Gamma S = \left\{ (y', v') \in \mathcal{E} \times \mathbb{R}_{++} \mid \begin{bmatrix} \zeta \\ -(\zeta, \delta)^T(x, u) \end{bmatrix}^T \begin{bmatrix} y' - y \\ v' - v \end{bmatrix} \leq 0 \right\}.$$

We say that a set is *polyhedral* if it is the intersection of finitely many halfspaces and  $\mathcal{E} \times \mathbb{R}_{++}$ . Then as an immediate consequence of Proposition 2.2 and (5), being polyhedral is preserved under the radial set transformation.

**Corollary 2.3.** *A set  $S \subseteq \mathcal{E} \times \mathbb{R}_{++}$  is polyhedral if and only if  $\Gamma S$  is polyhedral.*

Lastly, we consider the radial transformation of ellipsoids. A set  $S \subseteq \mathcal{E} \times \mathbb{R}$  is an ellipsoid if for some center  $(x, u)$  and positive definite linear mapping  $H$ ,

$$S = \left\{ (x', u') \in \mathcal{E} \times \mathbb{R} \mid \begin{bmatrix} x' - x \\ u' - u \end{bmatrix}^T H \begin{bmatrix} x' - x \\ u' - u \end{bmatrix} \leq 1 \right\}. \quad (7)$$

Similar to halfspaces, the radial transformation of such an ellipsoid in  $\mathcal{E} \times \mathbb{R}_{++}$  is an ellipsoid in  $\mathcal{E} \times \mathbb{R}_{++}$ . Curiously, the center of  $\Gamma S$  is not  $\Gamma(x, u)$  (as one might expect), but rather the depends on  $H$ .

**Proposition 2.4.** *A set  $S$  is an ellipsoid if and only if  $\Gamma S$  is an ellipsoid.*

## 2.1 Normal Vectors Under the Radial Set Transformation

Now we consider how the normal vectors of a set relate to those of its radial transformation. Proposition 2.2's description of transformed halfspaces characterizes convex normal vectors under the transformation. Combining this result with Proposition 2.4's description of transformed ellipsoids gives a characterization for proximal normal vectors.

**Proposition 2.5.** For any  $S \subseteq \mathcal{E} \times \mathbb{R}_{++}$ , all  $(y, v) \in \Gamma S$  have

$$N_{\Gamma S}^C((y, v)) = \left\{ \left[ -(\zeta, \delta)^T(x, u) \right] \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix} \in N_S^C((x, u)) \right\}$$

where  $(x, u) = \Gamma(y, v)$ .

*Proof.* For any  $(x, u) \in S$ ,  $(\zeta, \delta) \in N_S^C((x, u))$  if and only if

$$S \subseteq \left\{ (x', u') \in \mathcal{E} \times \mathbb{R}_{++} \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix}^T \begin{bmatrix} x' - x \\ u' - u \end{bmatrix} \leq 0 \right\}.$$

Letting  $(y, v) = \Gamma(x, u)$ , Proposition 2.2 and (4) imply

$$\Gamma S \subseteq \left\{ (y', v') \in \mathcal{E} \times \mathbb{R}_{++} \mid \left[ -(\zeta, \delta)^T(x, u) \right]^T \begin{bmatrix} y' - y \\ v' - v \end{bmatrix} \leq 0 \right\}.$$

Thus  $(\zeta, -(\zeta, \delta)^T(x, u)) \in N_{\Gamma S}^C((y, v))$ . This gives the containment

$$N_{\Gamma S}^C((y, v)) \supseteq \left\{ \left[ -(\zeta, \delta)^T(x, u) \right] \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix} \in N_S^C((x, u)) \right\}$$

and repeating the argument, replacing  $S$  by  $\Gamma S$ , gives the radially dual containment

$$N_S^C((x, u)) = N_{\Gamma \Gamma S}^C((x, u)) \supseteq \left\{ \left[ -(\zeta', \delta')^T(y, v) \right] \mid \begin{bmatrix} \zeta' \\ \delta' \end{bmatrix} \in N_{\Gamma S}^C((y, v)) \right\}.$$

Applying these two containments in succession shows the claimed formula

$$\begin{aligned} N_{\Gamma S}^C((y, v)) &\supseteq \left\{ \left[ -(\zeta, \delta)^T(x, u) \right] \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix} \in N_S^C((x, u)) \right\} \\ &\supseteq \left\{ \left[ -(\zeta', -(\zeta', \delta')^T(y, v))^T(x, u) \right] \mid \begin{bmatrix} \zeta' \\ \delta' \end{bmatrix} \in N_{\Gamma S}^C((y, v)) \right\} = N_{\Gamma S}^C((y, v)). \quad \square \end{aligned}$$

**Proposition 2.6.** For any  $S \subseteq \mathcal{E} \times \mathbb{R}_{++}$ , all  $(y, v) \in \Gamma S$  have

$$N_{\Gamma S}^P((y, v)) = \left\{ \left[ -(\zeta, \delta)^T(x, u) \right] \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix} \in N_S^P((x, u)) \right\}$$

where  $(x, u) = \Gamma(y, v)$ .

*Proof.* Consider any  $(x, u) \in S$  and  $(\zeta, \delta) \in N_S^P((x, u))$ . Then for some  $\epsilon > 0$ , the ball

$$E = B\left(\begin{bmatrix} x \\ u \end{bmatrix} + \epsilon \begin{bmatrix} \zeta \\ \delta \end{bmatrix}, \epsilon \left\| \begin{bmatrix} \zeta \\ \delta \end{bmatrix} \right\| \right) \subset \mathcal{E} \times \mathbb{R}_{++}$$

has  $E \cap S = \{(x, u)\}$ . Recall from Proposition 2.4 that  $\Gamma E$  is an ellipsoid. Applying (5) implies  $\Gamma E \cap \Gamma S = \{(y, v)\}$  where  $(y, v) = \Gamma(x, u)$ . Since  $-(\zeta, \delta) \in N_E^C((x, u))$ , Proposition 2.5 implies  $(-\zeta, (\zeta, \delta)^T(x, u)) \in N_{\Gamma E}^C((y, v))$ . Then for sufficiently small  $\epsilon' > 0$ , the ball

$$E' = B\left(\begin{bmatrix} y \\ v \end{bmatrix} + \epsilon' \begin{bmatrix} \zeta \\ -(\zeta, \delta)^T(x, u) \end{bmatrix}, \epsilon' \left\| \begin{bmatrix} \zeta \\ -(\zeta, \delta)^T(x, u) \end{bmatrix} \right\| \right)$$

lies in  $\Gamma E$ , and hence has  $E' \cap \Gamma S = \{(y, v)\}$ . Thus  $(\zeta, -(\zeta, \delta)^T(x, u)) \in N_{\Gamma S}^P((y, v))$ , and so

$$N_{\Gamma S}^P((y, v)) \supseteq \left\{ \left[ -(\zeta, \delta)^T(x, u) \right] \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix} \in N_S^P((x, u)) \right\}.$$

As shown in the proof of Proposition 2.5, the claimed formula follows from this containment and the dual containment given by replacing  $S$  by  $\Gamma S$ .  $\square$

## 2.2 Examples and Pictures

In Figures 1 through 10, we give five examples of pairs of sets in  $\mathbb{R} \times \mathbb{R}_{++}$  radially dual to each other. Each figure includes the horizontal line  $L = \{(x, 1) \mid x \in \mathbb{R}\}$  as a black dashed line. Observe that  $L$  is exactly the set of fixed points of the radial point transformation. Further, points above  $L$  always map into points below  $L$  (and vice versa).

The first two example pairs given in Figures 1 and 2 and Figures 3 and 4 show the radial transformation of a halfspace and a polyhedron (which must be a halfspace and a polyhedron by Proposition 2.2 and Corollary 2.3). Examining the transformation of the horizontal and vertical faces of the square in Figure 3 demonstrates two simple properties of the radial set transformation: (i) horizontal lines map into horizontal lines and (ii) vertical lines map into rays extending away from the origin (and vice versa).

Figures 5 and 6 show the radial transformation of an ellipsoid (which must be an ellipsoid by Proposition 2.4). Figures 7 and 8 consider the radial set transformation of a parabola, which is nearly an ellipsoid in  $\mathbb{R} \times \mathbb{R}_{++}$  but it approaches height 0 at the origin.

Our last pair of examples in Figures 9 and 10 show the radial set transformation of a sine wave. Notice that the resulting set is not the graph of any function. As we now transition to discussing our radial function transformations, considering how graphs, epigraphs, and hypographs behave under the set transformation provides key intuitions. The fact that the epigraph of our example parabola does not transform into the hypograph of another function and that the graph of our example sine wave does not transform into the graph of another function (as we will see) correspond to radial duality not holding for these function.

## 3 The Radial Function Transformation

Recall that we defined two radial function transformations based on the radial set transformation. The *upper radial function transformation* of some  $f: \mathcal{E} \rightarrow \overline{\mathbb{R}}_{++}$  is defined by

$$\begin{aligned} f^\Gamma(y) &= \sup\{v > 0 \mid (y, v) \in \Gamma(\text{epi } f)\} \\ &= \sup\{v > 0 \mid v \cdot f(y/v) \leq 1\}. \end{aligned} \quad (8)$$

This transformation essentially applies  $\Gamma$  to the epigraph of  $f$  and then interprets  $\Gamma(\text{epi } f)$  as the hypograph of a new function. Alternatively, interpreting  $\Gamma(\text{hypo } f)$  as the epigraph of a new function gives the *lower radial function transformation* defined by

$$\begin{aligned} f_\Gamma(y) &= \inf\{v > 0 \mid (y, v) \in \Gamma(\text{hypo } f)\} \\ &= \inf\{v > 0 \mid v \cdot f(y/v) \geq 1\}. \end{aligned} \quad (9)$$

Based on (8) and (9), these transformations can alternatively be defined using the *perspective function* of  $f$ , which we denote by  $f^p(y, v) = v \cdot f(y/v)$  for any  $v > 0$ . It is immediate from this viewpoint that

$$f^\Gamma = f_\Gamma \iff f^p(y, \cdot) \text{ is nondecreasing and strictly increasing on its domain.} \quad (10)$$

Having nondecreasing  $f^p(y, \cdot)$  can be understood in terms of the intersection of rays with the epigraph or hypograph of  $f$ . The following lemma shows that if  $f^p(y, \cdot)$  is nondecreasing, the ray  $\{\lambda(y, 1) \mid \lambda > 0\}$  has  $\lambda(y, 1)$  lie in the hypograph for all  $\lambda < \lambda_0$  and lie in the epigraph for all  $\lambda > \lambda_0$  for some  $\lambda_0 \in \overline{\mathbb{R}}_{++}$ . Thus the hypograph of any such function is star-shaped with respect to the origin.



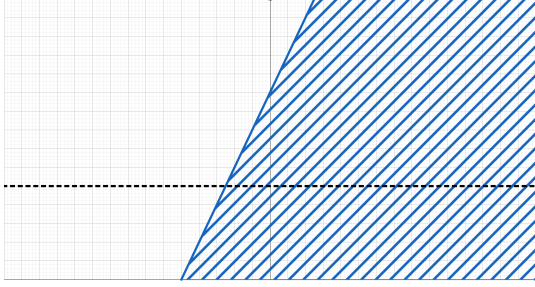


Figure 1: A halfspace.

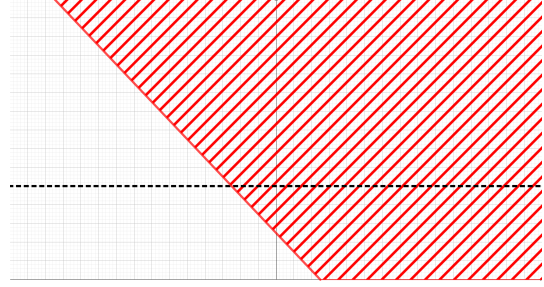


Figure 2: Radially dual halfspace.

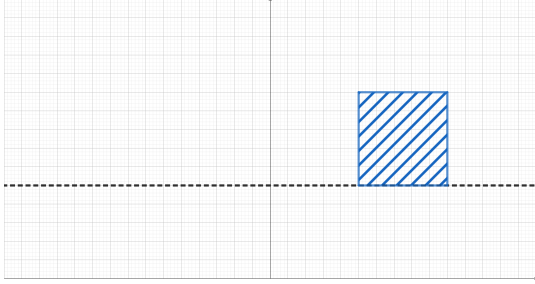


Figure 3: A polyhedron.

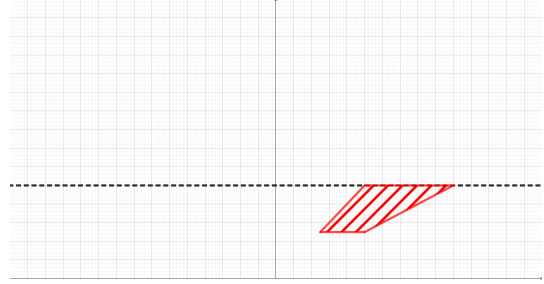


Figure 4: Radially dual polyhedron.

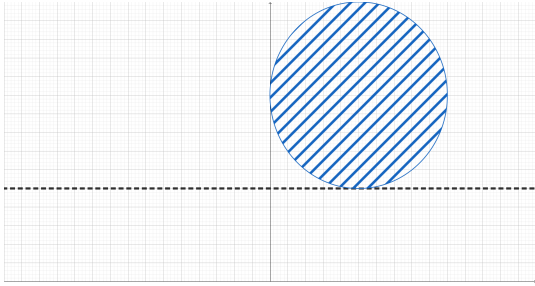


Figure 5: An ellipsoid.

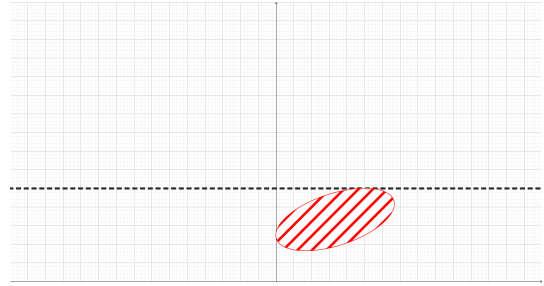


Figure 6: Radially dual ellipsoid.

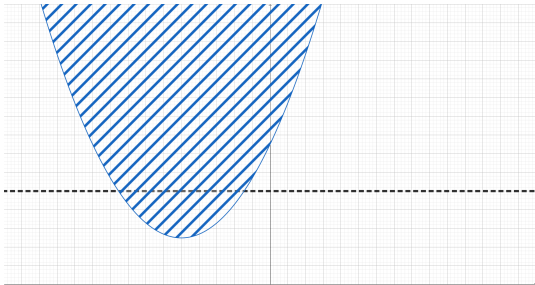


Figure 7: An epigraph of a quadratic.

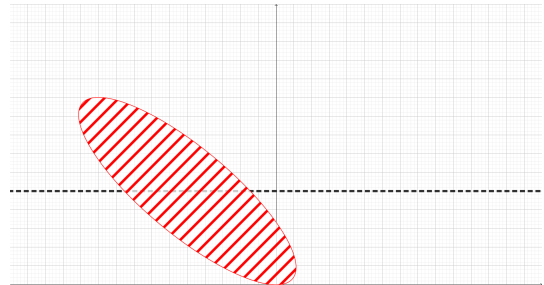


Figure 8: Radial dual of a quadratic.

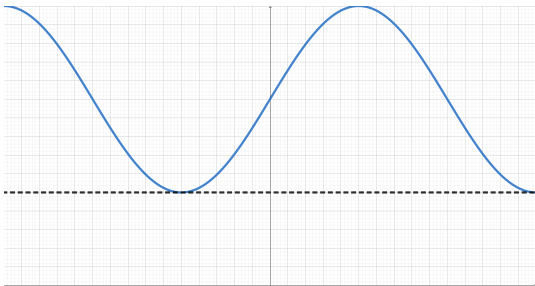


Figure 9: A sine wave.



Figure 10: Radial dual of a sine wave.

**Lemma 3.1.** *The following three conditions are equivalent:*

- (i) *all  $y \in \mathcal{E}$  have  $f^p(y, \cdot)$  nondecreasing,*
- (ii) *all  $(y, v) \in \text{epi } f$  and  $t \leq 1$  have  $(y, v)/t \in \text{epi } f$ ,*
- (iii) *all  $(y, v) \in \text{hypo } f$  and  $t \geq 1$  have  $(y, v)/t \in \text{hypo } f$ .*

*Proof.* First suppose  $f^p(y, \cdot)$  is nondecreasing and consider any  $(y, v) \in \text{epi } f$  and  $t \leq 1$ . Then  $t \cdot f(y/t) \leq f(y) \leq v$ , and so  $(y, v)/t \in \text{epi } f$ . Inversely, suppose some  $t < t'$  has  $f^p(y, t) > f^p(y, t')$ . Then every  $f(y/t') < \alpha < (t/t') \cdot f(y/t)$  must have  $(y/t', \alpha) \in \text{epi } f$ . However, dividing this point by  $t/t' \leq 1$  gives  $(y/t, (t'/t)\alpha) \notin \text{epi } f$ . Hence (i)  $\iff$  (ii). Symmetric arguments show the equivalent hypograph condition as (i)  $\iff$  (iii).  $\square$

We say a function  $f$  is *upper (lower) radial* whenever  $f^p(y, \cdot)$  is nondecreasing and upper (lower) semicontinuous for all  $y \in \mathcal{E}$ . If in addition  $f^p(y, \cdot)$  is strictly increasing on its domain, we say  $f$  is *strictly upper (lower) radial*. The following theorem shows being upper (lower) radial is exactly the condition for the duality of the point and set transformations (3) to carry over to the upper (lower) radial function transformation.

**Theorem 3.2.** *A function  $f$  is upper radial if and only if  $f^{\Gamma\Gamma} = f$ . Likewise<sup>2</sup>, a function  $f$  is lower radial if and only if  $f_{\Gamma\Gamma} = f$ .*

*Proof.* Observe that  $(f^\Gamma)^p(x, \cdot)$  is nondecreasing since it can be written as

$$\begin{aligned} u \cdot f^\Gamma(x/u) &= u \cdot \sup\{v > 0 \mid v \cdot f(x/vu) \leq 1\} \\ &= \sup\{w > 0 \mid w \cdot f(x/w) \leq u\}. \end{aligned}$$

Then the twice radially transformed function equals the following infimum

$$\begin{aligned} f^{\Gamma\Gamma}(x) &= \inf\{u > 0 \mid u \cdot f^\Gamma(x/u) > 1\} \\ &= \inf\{u > 0 \mid \sup\{w > 0 \mid w \cdot f(x/w) \leq u\} > 1\} \\ &= \inf\{u > 0 \mid \exists w > 1 \text{ s.t. } w \cdot f(x/w) \leq u\} \\ &= \inf\{w \cdot f(x/w) \mid w > 1\}. \end{aligned}$$

The claimed duality follows as  $f(x) = \inf\{w \cdot f(x/w) \mid w > 1\}$  if and only if  $w \mapsto w \cdot f(x/w)$  is nondecreasing and upper semicontinuous on  $w > 0$  for all  $x \in \mathcal{E}$ .  $\square$

The radial duality among upper (or lower) radial functions is central to understanding the radial function transformations. In Section 3.1, we begin by characterizing when important classes of functions are upper or lower radial (i.e., semicontinuous, differentiable, convex, and concave functions). Then Section 3.2 shows being radial is preserved under many standard operations (i.e., conic combinations, linear compositions, minimums, and maximums). Sections 3.3, 3.4, 3.5, and 3.6 consider the radial transformations of semicontinuous, piecewise linear, concave/convex, and quasiconcave/quasiconvex functions, respectively. We conclude this section by giving several examples of radial function transformations in Section 3.7

---

<sup>2</sup>Throughout this manuscript, we claim a mirrored results for the lower radial transformation in our theorems and propositions. We omit the proofs for these as they parallel those for the upper radial case.

### 3.1 Characterizing Radial Functions

**3.1.1. Radial Semicontinuous Functions** Here we consider when an upper (or lower) semicontinuous function is upper (or lower) radial, and thus when our duality result holds. Unlike Lemma 3.1 which focuses on rays from the origin, we give a necessary and sufficient condition based on proximal normal vectors of the function's hypograph (or epigraph). We find it suffices to consider whether the origin lies below the hyperplane induced by each proximal normal vector.

**Proposition 3.3.** *An upper semicontinuous  $f$  is upper radial if and only if all  $(x, u) \in \text{hypo } f$  and  $(\zeta, \delta) \in N_{\text{hypo } f}^P((x, u))$  satisfy*

$$(\zeta, \delta)^T(x, u) \geq 0.$$

*Likewise, a lower semicontinuous  $f$  is lower radial if and only if all  $(x, u) \in \text{epi } f$  and  $(\zeta, \delta) \in N_{\text{epi } f}^P((x, u))$  satisfy*

$$(\zeta, \delta)^T(x, u) \leq 0.$$

*Proof.* First suppose  $f$  is upper radial and consider any  $(x, u) \in \text{hypo } f$  and  $(\zeta, \delta) \in N_{\text{hypo } f}^P((x, u))$ . By Lemma 3.1,  $(x, u)/t \in \text{hypo } f$  for all  $t \geq 1$ . Since  $(x, u) \in \text{proj}_{\text{hypo } f}((x, u) + \epsilon(\zeta, \delta))$  for some  $\epsilon > 0$ , all  $t \geq 1$  satisfy

$$\|(x, u) + \epsilon(\zeta, \delta) - (x, u)\|^2 \leq \|(x, u) + \epsilon(\zeta, \delta) - (x, u)/t\|^2.$$

Simplifying this gives

$$0 \leq (1 - 1/t)^2 \|(x, u)\|^2 + 2\epsilon(1 - 1/t)(\zeta, \delta)^T(x, u),$$

and so taking  $t \rightarrow 1$  verifies  $(\zeta, \delta)^T(x, u) \geq 0$ .

Note that  $f^p(y, \cdot)$  is upper semicontinuous by assumption. Now suppose  $f^p(y, \cdot)$  is not nondecreasing. Then Lemma 3.1 guarantees some  $(x, u) \in \text{hypo } f$  and  $\gamma > 1$  has  $(x, u)/\gamma \notin \text{hypo } f$ . The assumed upper semicontinuity guarantees  $\text{hypo } f$  is closed, and thus for some  $\epsilon > 0$ ,

$$B((x, u)/\gamma, \epsilon) \cap \text{hypo } f = \emptyset.$$

Hence the following supremum is well defined

$$\gamma' := \sup\{1 < t \leq \gamma \mid B((x, u)/t, \epsilon/2) \cap \text{hypo } f \neq \emptyset\}.$$

Notice that  $1 < \gamma' < \gamma$ . Further,  $\text{int } B((x, u)/\gamma', \epsilon/2) \cap \text{hypo } f = \emptyset$ . Moreover, since  $\text{hypo } f$  is closed, some  $(x', u') \in \text{hypo } f$  lies on the boundary of this ball – that is,  $\|(x, u)/\gamma' - (x', u')\| = \epsilon/2$ . Then  $\text{hypo } f$  at  $(x', u')$  has the following proximal normal vector

$$(\zeta', \delta') := (x, u)/\gamma' - (x', u') \in N_{\text{hypo } f}^P((x', u')).$$

Since all  $\gamma' < t < \gamma$  have  $(x', u') \notin B((x, u)/t, \epsilon/2)$ ,

$$\begin{aligned} \left(\frac{\epsilon}{2}\right)^2 &\leq \left\| \frac{(x, u)}{t} - (x', u') \right\|^2 \\ &= \left\| \frac{\gamma'}{t}(\zeta', \delta') - \left(1 - \frac{\gamma'}{t}\right)(x', u') \right\|^2 \\ &= \left\| \frac{\gamma'}{t}(\zeta', \delta') \right\|^2 - 2\frac{\gamma'}{t} \left(1 - \frac{\gamma'}{t}\right) (\zeta', \delta')^T(x', u') + \left\| \left(1 - \frac{\gamma'}{t}\right)(x', u') \right\|^2 \end{aligned}$$

$$= \left(\frac{\gamma' \epsilon}{t} \frac{1}{2}\right)^2 - 2\frac{\gamma'}{t} \left(1 - \frac{\gamma'}{t}\right) (\zeta', \delta')^T(x', u') + \left\| \left(1 - \frac{\gamma'}{t}\right)(x', u') \right\|^2.$$

Rearrangement of this inequality gives

$$2(\zeta', \delta')^T(x', u') \leq -\left(\frac{t}{\gamma'} + 1\right) \left(\frac{\epsilon}{2}\right)^2 + \left(\frac{t}{\gamma'} - 1\right) \|(x', u')\|^2.$$

Taking  $t \rightarrow \gamma'$  shows  $(\zeta', \delta')^T(x', u') \leq -\epsilon^2/2 < 0$ .  $\square$

**3.1.2. Radial Differentiable Functions** Now we specialize the previous result for semicontinuous functions to differentiable functions. We want to allow functions like the previously considered example  $f(x) = \sqrt{1 - \|x\|^2}$  (with value 0 whenever  $\|x\| \geq 1$ ) in our theory here. To this end, we say a function is *continuously differentiable* if  $f$  is continuous on  $\mathcal{E}$  and  $\nabla f(x)$  exists and is continuous on its effective domain  $\text{dom } f = \{x \in \mathcal{E} \mid f(x) \in \mathbb{R}_{++}\}$ .

For any such function  $f$  and  $x \in \text{dom } f$ , if some nonzero  $(\zeta, \delta) \in N_{\text{epi } f}^P((x, u))$  exists, then  $u = f(x)$  and  $(\zeta, \delta) = \lambda(\nabla f(x), -1)$  for some  $\lambda \geq 0$ . Further, for a dense subset of  $\text{dom } f$ , the converse holds (which follows from the density theorem of proximal calculus [6, Theorem 1.3.1]). Then the continuity of  $\nabla f$  and Proposition 3.3 imply the following condition is necessary and sufficient for  $f$  to be radial<sup>3</sup>.

**Proposition 3.4.** *A continuously differentiable  $f$  is radial if and only if for all  $x \in \text{dom } f$ ,*

$$(\nabla f(x), -1)^T(x, f(x)) \leq 0.$$

This characterization can be alternatively derived by considering when the partial derivative  $\frac{\partial}{\partial v} f^p(y, v) = f(y/v) - \nabla f(y/v)^T(y/v)$  is nonnegative. Based on this observation, having a positive derivative is sufficient to ensure  $f^p(y, \cdot)$  is strictly increasing on its domain, and thus  $f^\Gamma = f_\Gamma$  by (10).

**Proposition 3.5.** *A continuously differentiable  $f$  is strictly radial if for all  $x \in \text{dom } f$ ,*

$$(\nabla f(x), -1)^T(x, f(x)) < 0.$$

**3.1.3. Radial Convex and Concave Functions** Lastly we consider conditions for convex or concave functions to be upper or lower radial. For convex functions, the proximal subdifferential and convex subdifferential are equal giving the following characterization.

**Proposition 3.6.** *A lower semicontinuous convex  $f$  is lower radial if and only if all  $(x, u) \in \text{epi } f$  and  $(\zeta, \delta) \in N_{\text{epi } f}^C((x, u))$  have*

$$(\zeta, \delta)^T(x, u) \leq 0.$$

Now we consider concave functions, finding that it suffices to have points arbitrarily close to the origin with nonzero function value. As a result, every upper semicontinuous concave function can be translated to become upper radial.

**Proposition 3.7.** *An upper semicontinuous concave  $f$  has  $f^p(y, \cdot)$  nondecreasing if and only if  $0 \in \text{cl } \{x \mid f(x) > 0\}$  or  $f = 0$ .*

---

<sup>3</sup>A continuous functions is (strictly) upper radial if and only if it is (strictly) lower radial. In such cases, we simply say the function is (strictly) radial as a shorthand.

*Proof.* Trivially  $f = 0$  has  $f^p(y, \cdot)$  nondecreasing and so we assume some  $x'$  has  $f(x') > 0$ . Then  $f^p(y, \cdot)$  being nondecreasing implies all  $t \geq 1$  have  $f(x'/t) \geq f(x')/t > 0$ . Taking  $t \rightarrow \infty$  gives a sequence of points verifying  $0 \in \text{cl } \{y \mid f(y) > 0\}$ .

Conversely, suppose  $0 \in \text{cl } \{x \mid f(x) > 0\}$  and consider any  $(x, u) \in \text{hypo } f$ . Since  $\text{hypo } f$  is closed and convex and  $(0, 0) \in \text{cl hypo } f$ , the line segment  $((0, 0), (x, u)]$  must lie in  $\text{hypo } f$ . This is equivalent to  $f^p(y, \cdot)$  being nondecreasing by Lemma 3.1.  $\square$

Furthermore, if the origin lies in the interior of  $\{x \mid f(x) > 0\}$ , we find that  $f^p(x, \cdot)$  is strictly increasing on its domain, and thus  $f^\Gamma = f_\Gamma$  by (10). This condition can easily be attained for any concave  $f$  whenever a point in the interior of the function's domain is known by translating it to the origin. This directly corresponds to the setting assumed by Grimmer [8] and is equivalent to the conic setting assumed by Renegar [11].

**Proposition 3.8.** *A concave  $f$  has  $f^p(y, \cdot)$  strictly increasing on its domain if  $0 \in \text{int } \{x \mid f(x) > 0\}$ .*

*Proof.* Consider any  $y \in \mathcal{E}$  and  $0 < v < v'$  with  $v \cdot f(y/v) \in \mathbb{R}_{++}$ . Since  $(y/v, f(y/v)) \in \text{hypo } f$ , the convexity of  $f$  ensures that the line segment  $((0, 0), (y/v, f(y/v)))$  lies in the interior of the hypograph of  $f$ . In particular,  $(v/v') \cdot (y/v, f(y/v)) = (y/v', (v/v') \cdot f(y/v)) \in \text{int hypo } f$ . Therefore  $f(y/v') > (v/v')f(y/v)$  and so  $f^p(y, v) < f^p(y, v')$ .  $\square$

### 3.2 Closure of Radial Functions Under Common Operations

Building on our characterizations of when important classes of functions are upper or lower radial, here we show that this structure is preserved under many common operations. The following result shows this is the case for any conic combination (that is,  $\sum_{i=1}^k \lambda_i f_i(x)$  with each  $\lambda_i > 0$ ), composition with linear maps, and taking finite minimums and maximums.

**Proposition 3.9.** *For any pair of (strictly) upper radial functions  $f_1, f_2$ , the following functions are also (strictly) upper radial functions:*

- (i) *Positive Rescaling by  $\lambda > 0$ :  $\lambda \cdot f_1$ ,*
- (ii) *Composition with a linear map  $A: \mathcal{E}' \rightarrow \mathcal{E}$ :  $f_1 \circ A$ ,*
- (iii) *Addition:  $f_1 + f_2$ ,*
- (iv) *Minimums:  $\min\{f_1, f_2\}$ ,*
- (v) *Maximums:  $\max\{f_1, f_2\}$ .*

*Likewise, these operations all preserve being (strictly) lower radial.*

*Proof.* Each of these operations preserves upper and lower semicontinuity and being nondecreasing (or strictly increasing). Consequently, they preserve being (strictly) upper or lower radial.  $\square$

Note that (i) and (iii) above together give the claimed result for conic combinations. For all of these operations except addition, we can give simple formulas for the resulting radial transformation, formalized in the following proposition.

**Proposition 3.10.** *For any pair of functions  $f_1, f_2$ , the following identities hold for any  $\lambda > 0$  and linear  $A: \mathcal{E}' \rightarrow \mathcal{E}$*

$$\begin{aligned} (\lambda \cdot f_1)^\Gamma(y) &= f_1^\Gamma(\lambda y)/\lambda, \\ (f_1 \circ A)^\Gamma &= f_1^\Gamma \circ A, \\ \min\{f_1, f_2\}^\Gamma &= \max\{f_1^\Gamma, f_2^\Gamma\}. \end{aligned}$$

Further, if  $f_1, f_2$  are upper radial, then

$$\max\{f_1, f_2\}^\Gamma = \min\{f_1^\Gamma, f_2^\Gamma\}.$$

Likewise, similar identities hold for the lower radial transformation.

*Proof.* The results for positive rescaling by some  $\lambda > 0$ , for composition with a linear map  $A: \mathcal{E}' \rightarrow \mathcal{E}$ , and for minimums follow immediately from the definition of our radial transformation as

$$\begin{aligned} (\lambda \cdot f)^\Gamma(y) &= \sup\{v > 0 \mid \lambda v \cdot f(y/v) \leq 1\} \\ &= \sup\{w > 0 \mid w \cdot f(\lambda y/w) \leq 1\}/\lambda \\ &= f^\Gamma(\lambda y)/\lambda, \\ (f \circ A)^\Gamma(y) &= \sup\{v > 0 \mid v \cdot f(Ay/v) \leq 1\} \\ &= f^\Gamma(Ay), \\ \min\{f_1, f_2\}^\Gamma(y) &= \sup\{v > 0 \mid v \cdot \min\{f_1(y/v), f_2(y/v)\} \leq 1\} \\ &= \max_{i=1,2} \{\sup\{v > 0 \mid v \cdot f_i(y/v) \leq 1\}\} \\ &= \max\{f_1^\Gamma(y), f_2^\Gamma(y)\}. \end{aligned}$$

The claimed formula for maximums follows as

$$\begin{aligned} \max\{f_1, f_2\}^\Gamma(y) &= \sup\{v > 0 \mid v \cdot \max\{f_1(y/v), f_2(y/v)\} \leq 1\} \\ &= \min_{i=1,2} \{\sup\{v > 0 \mid v \cdot f_i(y/v) \leq 1\}\} \\ &= \min\{f_1^\Gamma(y), f_2^\Gamma(y)\} \end{aligned}$$

where the second equality above relies on each  $v \cdot f_i(y/v)$  being nondecreasing.  $\square$

From these simple operations, we can build up to more complex functions that preserve being upper radial. For example, consider the operations taking the  $k$ th largest or smallest element out of a set of  $n$  elements

$$\begin{aligned} k\text{-min}\{x_1, \dots, x_n\} &:= x_{i_k} \text{ where } x_{i_1} \leq x_{i_2} \leq \dots \leq x_{i_n} \\ k\text{-max}\{x_1, \dots, x_n\} &:= x_{i_{n-k+1}} \text{ where } x_{i_1} \leq x_{i_2} \leq \dots \leq x_{i_n}. \end{aligned}$$

and averaging the  $k$  largest or smallest elements

$$\begin{aligned} k\text{-minavg}\{x_1, \dots, x_n\} &:= \frac{1}{k} \sum_{j=1}^k x_{i_j} \text{ where } x_{i_1} \leq x_{i_2} \leq \dots \leq x_{i_n} \\ k\text{-maxavg}\{x_1, \dots, x_n\} &:= \frac{1}{k} \sum_{j=1}^k x_{i_{n-j+1}} \text{ where } x_{i_1} \leq x_{i_2} \leq \dots \leq x_{i_n}. \end{aligned}$$

**Corollary 3.11.** *For any (strictly) upper radial functions  $f_1, \dots, f_n$ , the functions  $k\text{-min}\{f_i(x)\}$ ,  $k\text{-max}\{f_i(x)\}$ ,  $k\text{-minavg}\{f_i(x)\}$ , and  $k\text{-maxavg}\{f_i(x)\}$  are all (strictly) upper radial with*

$$(k\text{-min}\{f_i(x)\})^\Gamma(y) = k\text{-max}\{f_i^\Gamma(y)\}.$$

*Proof.* This follows immediately from Propositions 3.9 and 3.10 since these operations can be described as combinations of minimums, maximums, positive rescaling, and addition

$$\begin{aligned}
k\text{-min}\{f_1(x), \dots, f_n(x)\} &= \min\{\max\{f_i(x) \mid i \in S\} \mid S \subseteq \{1, \dots, n\}, |S| = k\}, \\
k\text{-max}\{f_1(x), \dots, f_n(x)\} &= \max\{\min\{f_i(x) \mid i \in S\} \mid S \subseteq \{1, \dots, n\}, |S| = k\}, \\
k\text{-minavg}\{f_1(x), \dots, f_n(x)\} &= \min\left\{\frac{1}{k} \sum_{i \in S} f_i(x) \mid S \subseteq \{1, \dots, n\}, |S| = k\right\}, \\
k\text{-maxavg}\{f_1(x), \dots, f_n(x)\} &= \max\left\{\frac{1}{k} \sum_{i \in S} f_i(x) \mid S \subseteq \{1, \dots, n\}, |S| = k\right\}. \quad \square
\end{aligned}$$

### 3.3 Radial Transformation of Semicontinuous Functions

Now we turn our focus to understanding how various important families of functions behave under the radial transformation. Considering the transformation of upper and lower semicontinuous functions shows that these become lower and upper semicontinuous, respectively, when  $f$  is appropriately radial.

**Proposition 3.12.** *For any lower semicontinuous, lower radial  $f$ ,  $f^\Gamma$  is upper semicontinuous. Likewise, for any upper semicontinuous, upper radial  $f$ ,  $f_\Gamma$  is lower semicontinuous.*

*Proof.* Consider any  $y \in \mathcal{E}$ . Upper semicontinuity trivially holds at  $y$  if  $f^\Gamma(y) = \infty$ . Now assume  $f^\Gamma(y) < \infty$  and consider any  $\gamma > f^\Gamma(y)$ . Then  $\gamma \cdot f(y/\gamma) > 1$ . From the lower semicontinuity of  $f$ , for some  $\epsilon > 0$ , all  $y' \in B(y, \epsilon)$  satisfy  $\gamma \cdot f(y'/\gamma) > 1$ . Therefore  $f^\Gamma(y') \leq \gamma$ . Taking the limit as  $\gamma$  approaches  $f^\Gamma(y)$  shows  $f^\Gamma(y) = \limsup_{y' \rightarrow y} f^\Gamma(y')$ .  $\square$

These results with upper and lower semicontinuity reversed do not hold in general. However, whenever  $f^\Gamma = f_\Gamma$ , the reversed propositions immediately hold. Thus, for strictly upper (lower) radial functions, upper semicontinuity and lower semicontinuity are dual to each other under the upper (lower) transformation.

**Proposition 3.13.** *For any  $f$  with  $f^p(y, \cdot)$  strictly increasing on its domain for all  $y \in \mathcal{E}$ ,*

$$\begin{aligned}
f \text{ upper semicontinuous} &\implies f^\Gamma = f_\Gamma \text{ lower semicontinuous,} \\
f \text{ lower semicontinuous} &\implies f^\Gamma = f_\Gamma \text{ upper semicontinuous,} \\
f \text{ continuous} &\implies f^\Gamma = f_\Gamma \text{ continuous.}
\end{aligned}$$

*Proof.* Since  $f^\Gamma = f_\Gamma$  by (10), both directions of Proposition 3.12 apply.  $\square$

### 3.4 Radial Transformation of Piecewise Linear Functions

We say a function  $f$  is *convex polyhedral* if  $\text{epi } f$  is the intersection of finitely many halfspaces and  $\mathcal{E} \times \mathbb{R}_{++}$ . Likewise,  $f$  is *concave polyhedral* if  $\text{hypo } f$  is the intersection of finitely many halfspaces and  $\mathcal{E} \times \mathbb{R}_{++}$ . Recall Corollary 2.3 ensures polyhedral sets map to polyhedral sets under the radial set transformation. The following proposition shows how this property is mirrored by the radial function transformation on polyhedral functions.

**Proposition 3.14.** *If  $f$  is convex polyhedral then  $f^\Gamma$  is concave polyhedral. Likewise, if  $f$  is concave polyhedral then  $f_\Gamma$  is convex polyhedral.*

*Proof.* If  $\text{epi } f$  is polyhedral, then Corollary 2.3 implies  $\Gamma(\text{epi } f)$  is also polyhedral. Since  $\Gamma(\text{epi } f)$  is closed with respect to  $\mathcal{E} \times \mathbb{R}_{++}$ , the hypograph of  $f^\Gamma$  can be written as

$$\text{hypo } f^\Gamma = \{(y, v) \mid \exists v' \geq v, (y, v') \in \Gamma(\text{epi } f)\}.$$

Then Fourier-Motzkin ensures  $\text{hypo } f^\Gamma$  is polyhedral.  $\square$

Like the previous results on semicontinuity, the converses do not hold in general. However, whenever  $f^p(y, \cdot)$  is strictly increasing on its domain for all  $y \in \mathcal{E}$ , they immediately hold as  $f^\Gamma = f_\Gamma$ .

### 3.5 Radial Transformation of Concave/Convex Functions

Recall from Proposition 2.1 that the radial set transformation preserves convexity. This structure carries over to the function setting where convex functions become concave and vice versa.

**Proposition 3.15.** *If  $f$  is concave then  $f^\Gamma$  and  $f_\Gamma$  are convex. Likewise, if  $f$  is convex then  $f^\Gamma$  and  $f_\Gamma$  are concave.*

*Proof.* Note that the perspective function  $f^p(y, v)$  is concave (convex) whenever  $f$  is concave (convex) [5]. Supposing  $f$  is concave. Consider any  $(y, v), (y', v') \in \text{epi } f^\Gamma$  and  $0 \leq \lambda \leq 1$ . Note all  $t > v$  and  $t' > v'$  have  $t \cdot f(y/t) > 1$  and  $t' \cdot f(y'/t') > 1$ . Then the concavity of  $f^p$  implies

$$(\lambda t + (1 - \lambda)t') \cdot f\left(\frac{\lambda y + (1 - \lambda)y'}{\lambda t + (1 - \lambda)t'}\right) > 1.$$

Thus  $f^\Gamma(\lambda y + (1 - \lambda)y') \leq \lambda v + (1 - \lambda)v'$  since this holds for all  $t > v$  and  $t' > v'$ .

Now consider any  $(y, v), (y', v') \in \text{epi } f_\Gamma$  and  $0 \leq \lambda \leq 1$ . Note there must exist  $t, t'$  near  $f_\Gamma(y), f_\Gamma(y')$  with  $t \cdot f(y/t) \geq 1$  and  $t' \cdot f(y'/t') \geq 1$ . Then the concavity of  $f^p$  implies

$$(\lambda t + (1 - \lambda)t') \cdot f\left(\frac{\lambda y + (1 - \lambda)y'}{\lambda t + (1 - \lambda)t'}\right) \geq 1.$$

Thus  $f_\Gamma(\lambda y + (1 - \lambda)y') \leq \lambda f_\Gamma(y) + (1 - \lambda)f_\Gamma(y') \leq \lambda v + (1 - \lambda)v'$ .  $\square$

Thus the family of upper radial concave functions is dual to the family of upper radial convex functions. This is particularly interesting because these families of functions are very different due to the symmetry breaking nature of working with the extended positive reals. Proposition 3.7 shows that any concave function can be translated to become radial, whereas no similar operation exists for convex functions. This is a critical algorithmic insight since it allows us to take generic concave maximization problems and transform them into minimization problems that are both convex and upper radial. In the second part of this work, we will see that radially dual minimization problems are very structured, often being globally uniformly Lipschitz continuous, despite us starting with a quite generic maximization problem.

### 3.6 Radial Transformation of Quasi-concave/-convex Functions

Lastly we consider the generalization of concavity and convexity given by quasiconcavity and quasiconvexity. We say a function is *quasiconcave* (*quasiconvex*) if its superlevel sets  $\{x \in \mathcal{E} \mid f(x) \geq z\}$  (sublevel sets  $\{x \in \mathcal{E} \mid f(x) \leq z\}$ ) are convex for all  $z > 0$ . Similar to the previous section's results, we find that quasiconcave functions are dual to quasiconvex functions (although the additional condition that  $f^p(y, \cdot)$  is nondecreasing is needed for our fullest version of this result to hold).



**Proposition 3.16.** *If  $f$  is quasiconcave,  $f^\Gamma$  is quasiconvex. If in addition  $f^p(y, \cdot)$  is nondecreasing,  $f_\Gamma$  is quasiconvex. Likewise, if  $f$  is quasiconvex,  $f_\Gamma$  is quasiconcave. If in addition  $f^p(y, \cdot)$  is nondecreasing,  $f^\Gamma$  is quasiconcave.*

*Proof.* Suppose  $f$  is quasiconcave and fix any level  $z \in \mathbb{R}_{++}$ . Consider any  $0 \leq \lambda \leq 1$  and  $y, y' \in \mathcal{E}$  with  $f^\Gamma(y) \leq z$  and  $f^\Gamma(y') \leq z$ . First, we consider the upper radial transformation. Note all  $\gamma > z$  have  $\gamma \cdot f(y/\gamma) > 1$  and  $\gamma \cdot f(y'/\gamma) > 1$ . Then the quasiconcavity of  $f$  implies

$$\begin{aligned} \gamma \cdot f\left(\frac{\lambda y + (1-\lambda)y'}{\gamma}\right) &\geq \gamma \cdot \min\{f(y/\gamma), f(y'/\gamma)\} \\ &= \min\{\gamma \cdot f(y/\gamma), \gamma \cdot f(y'/\gamma)\} > 1. \end{aligned}$$

Thus  $f^\Gamma(\lambda y + (1-\lambda)y') \leq z$  since this holds for every  $\gamma > z$ .

Now we consider the lower radial transformation and further assume  $f^p(y, \cdot)$  is nondecreasing. Note all  $\gamma > z$  must have  $\gamma \cdot f(y/\gamma) \geq 1$  and  $\gamma \cdot f(y'/\gamma) \geq 1$ . Then the quasiconvexity of  $f$  implies

$$\begin{aligned} \gamma \cdot f\left(\frac{\lambda y + (1-\lambda)y'}{\gamma}\right) &\geq \gamma \cdot \min\{f(y/\gamma), f(y'/\gamma)\} \\ &= \min\{\gamma \cdot f(y/\gamma), \gamma \cdot f(y'/\gamma)\} \geq 1. \end{aligned}$$

Thus  $f_\Gamma(\lambda y + (1-\lambda)y') \leq z$  since this holds for every  $\gamma > z$ . □

### 3.7 Examples and Pictures

In Figures 11 through 20, we give a number of examples of radial function transformations. As done in our illustrations of the radial set transformation, each figure includes the horizontal line  $\{(x, 1) \mid x \in \mathbb{R}\}$  as a black dashed line for reference.

Figures 11, 12, and 13 show the absolute value function and its upper and lower radial transformations respectively. A simple calculation shows

$$\begin{aligned} |\cdot|^\Gamma(y) &= \begin{cases} +\infty & \text{if } -1 \leq y \leq 1 \\ 0 & \text{otherwise,} \end{cases} \\ |\cdot|_\Gamma(y) &= \begin{cases} +\infty & \text{if } -1 < y < 1 \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Note  $|x|$  is both upper and lower radial, but not strictly. As a result, although  $|\cdot|^\Gamma$  and  $|\cdot|_\Gamma$  are not equal, both are dual to  $|x|$  under their respective transformations.

Figures 14 and 15 show the strictly radial, concave function  $\sqrt{1-x^2}$  (with the value at all  $x^2 > 1$  set to 0) and its transformation  $\sqrt{1+y^2}$ . Notice the transformed function is convex as guaranteed by Proposition 3.15. Moreover, it is uniformly Lipschitz and smooth even though the original function possesses neither of these properties.

Figures 16 and 17 show the strictly radial function  $e^{-|x|} + 1/2$  and its transformation. Even though this function is neither concave nor convex, our transformation can still be directly applied. Moreover, this function is quasiconcave and so, as guaranteed by Proposition 3.16, its radial transformation is quasiconvex.

Lastly, Figures 18, 19, and 20 continue the example of transforming the quadratic  $(x+1)^2 + 1/2$  which was used in Figures 7 and 8 to illustrate the radial set transformation. Notice that this quadratic is not upper radial as its epigraph transforms into an ellipsoid-like shape rather than the hypograph of another function. Hence its upper radial transformation is not dual to the original quadratic. Figure 20 shows the result of applying the upper radial transformation to this quadratic twice. In this case, the functions in Figures 19 and 20 are upper radial and thus radially dual.

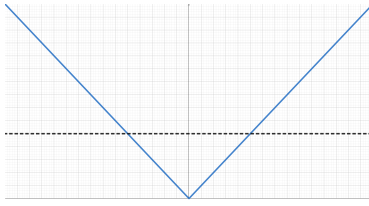


Figure 11:  $|x|$

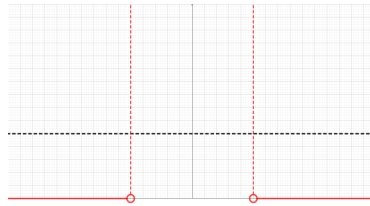


Figure 12:  $|\cdot|^\Gamma(y)$

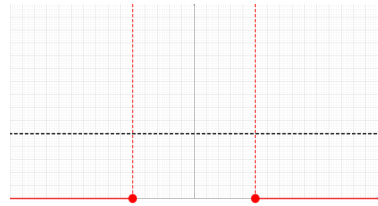


Figure 13:  $|\cdot|_\Gamma(y)$



Figure 14:  $f(x) = \sqrt{1-x^2}$

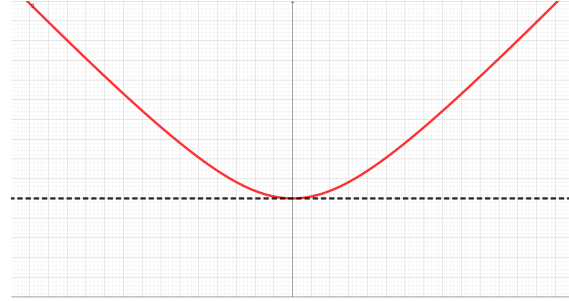


Figure 15:  $f^\Gamma(y) = f_\Gamma(y) = \sqrt{1+y^2}$

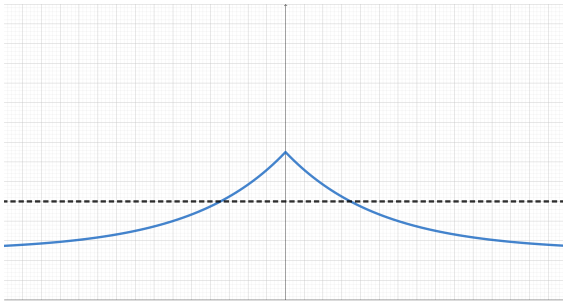


Figure 16:  $g(x) = e^{-|x|} + 1/2$

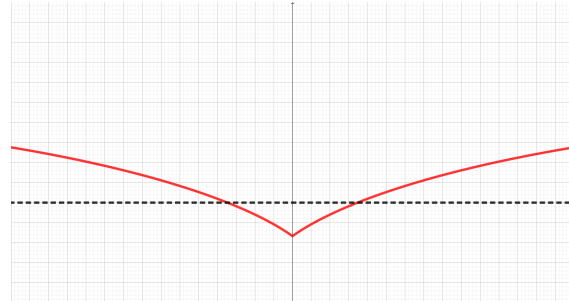


Figure 17:  $g^\Gamma(y) = g_\Gamma(y)$

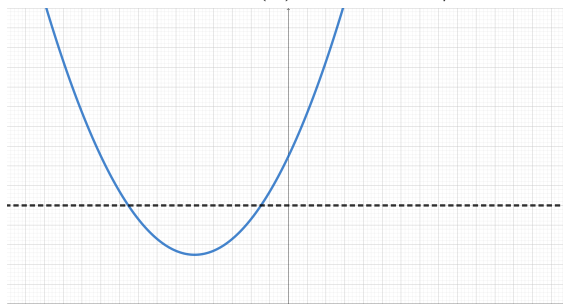


Figure 18:  $h(x) = (x+1)^2 + 1/2$



Figure 19:  $h^\Gamma(y)$

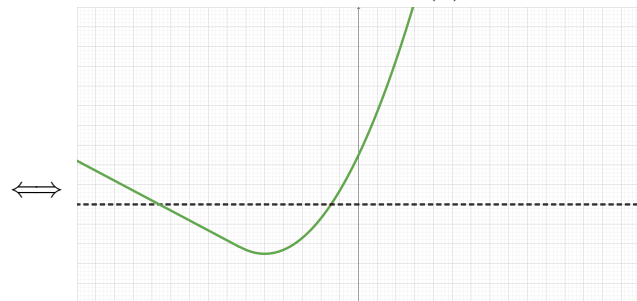


Figure 20:  $h^{\Gamma\Gamma}(x)$

## 4 Optimization Based on Radial Transformations

Here we develop the necessary machinery to propose and analyze optimization methods based on the radial transformation. One powerful facet of the duality between (1) and (2) lies in how constraints are transformed and the subsequent algorithmic gains. Consider maximizing  $f: \mathcal{E} \rightarrow \overline{\mathbb{R}}_{++}$  over some  $S \subseteq \mathcal{E}$ . To move the constraints into the objective, we define the indicator function of a set  $S \subseteq \mathcal{E}$  as

$$\iota_S(x) = \begin{cases} +\infty & \text{if } x \in S \\ 0 & \text{if } x \notin S. \end{cases}$$

Then it is immediate that the initial problem (1) can be written as

$$\max_{x \in S} f(x) = \max_{x \in \mathcal{E}} \min\{f(x), \iota_S(x)\}.$$

Problems of this form have particularly nice radial transformations since it distributes over the minimum by Proposition 3.10. Thus the radially dual problem (2) is

$$\min_{y \in \mathcal{E}} \max\{f^\Gamma(y), \iota_S^\Gamma(y)\}.$$

For convex  $S$  with  $0 \in S$ ,  $\iota_S^\Gamma(y)$  is precisely its gauge  $\gamma_S(y) = \inf\{\lambda \geq 0 \mid y \in \lambda S\}$ . As a result, we find that constraints in the original problem become gauges in the radially dual problem:

$$\iota_S^\Gamma(y) = \sup\{v > 0 \mid v \cdot \iota_S(y/v) \leq 1\} = \sup\{v > 0 \mid y/v \notin S\} = \inf\{\lambda > 0 \mid y \in \lambda S\} = \gamma_S(y).$$

This observation is very useful for designing first-order methods. Typically constrained optimization problems require orthogonal projections onto the feasible region at each iteration (which may be substantially more expensive than computing a single subgradient, often dominating an algorithm's runtime). Evaluating the gauge of a set and computing one of its subgradients can be far cheaper than orthogonal projection as it requires at most a one-dimensional line search and computing a single normal vector of the constraint set. Thus the radially dual problem effectively replaces the need for orthogonal projections with much simpler operations. Observing and taking advantage of this structure in the context of conic programming was a central contribution of [11]. Discussing the fuller implications of transforming constraints into their related gauge function on the design of algorithms is deferred to the second part of this work.

In the remainder of this section, we develop a calculus for the radially dual optimization problem. For any appropriately radial function, formulas for the convex and proximal subdifferentials and supdifferentials of its radial transformation are given in Section 4.1. Further, assuming  $f$  is sufficiently differentiable, Section 4.2 characterizes the gradients and Hessians of its radial transformations. In Section 4.3, we relate the optimal points (minimizers and maximizers) and stationary points of a function and its radial transformations. These calculus and optimality relations form the foundations of relating the pair of radially dual optimization problems (1) and (2).

### 4.1 Convex and Proximal Subgradients and Supgradients

To understand the convex and proximal subdifferentials under radial function transformations, we leverage Propositions 2.5 and 2.6 which described normal vectors under the radial set transformation. The following lemma relates the epigraph and hypograph of radially transformed functions to those of the original function.

**Lemma 4.1.** *For any upper radial  $f$ ,  $\text{epi } f^\Gamma \subseteq \Gamma(\text{hypo } f)$ .*

*Likewise, for any lower radial  $f$ ,  $\text{hypo } f_\Gamma \subseteq \Gamma(\text{epi } f)$ .*

*If  $f^p(y, \cdot)$  is strictly increasing on its domain, equality holds in both cases.*

*Proof.* Noting that  $f^\Gamma(y) \leq v \implies v \cdot f(y/v) \geq 1$  for upper radial  $f$ , this follows directly as

$$\begin{aligned} \Gamma(\text{hypo } f) &= \left\{ \frac{(x, 1)}{u} \mid f(x) \geq u \right\} \\ &= \{(y, v) \mid f(y/v) \geq 1/v\} \\ &= \{(y, v) \mid v \cdot f(y/v) \geq 1\} \\ &\supseteq \{(y, v) \mid f^\Gamma(y) \geq v\} = \text{epi } f^\Gamma. \end{aligned}$$

When  $f$  is strictly upper radial,  $f^\Gamma(y) \leq v \iff v \cdot f(y/v) \geq 1$ , and so equality holds.  $\square$

In light of Lemma 4.1, we can immediately apply our results on normal vectors under the radial set transformation to understand differentials under the function transformation. The following pair of propositions do this for the convex and proximal subdifferential and supdifferential.

**Proposition 4.2.** *For any strictly upper radial  $f$ ,*

$$\partial_C f^\Gamma(y) = \left\{ \frac{\zeta}{(\zeta, \delta)^T(x, u)} \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix} \in N_{\text{hypo } f}^C((x, u)), (\zeta, \delta)^T(x, u) > 0 \right\}$$

where  $(x, u) = \Gamma(y, f^\Gamma(y))$ . Likewise, for any strictly lower radial  $f$ ,

$$\partial^C f_\Gamma(y) = \left\{ \frac{\zeta}{(\zeta, \delta)^T(x, u)} \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix} \in N_{\text{epi } f}^C((x, u)), (\zeta, \delta)^T(x, u) < 0 \right\}.$$

*Proof.* Recall that Propositions 2.5 characterized the convex normal vectors of the radial transformation of a set in terms of the original set. Since the assumed strict increase ensures equality holds in Lemma 4.1, this applies to the epigraph and hypograph of  $f^\Gamma$  and  $f_\Gamma$ , respectively. Thus when  $f$  is strictly upper radial

$$N_{\text{epi } f^\Gamma}^C((y, v)) = \left\{ \begin{bmatrix} \zeta \\ -(\zeta, \delta)^T(x, u) \end{bmatrix} \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix} \in N_{\text{hypo } f}^C((x, u)) \right\} \quad (11)$$

and when  $f$  is strictly lower radial

$$N_{\text{hypo } f_\Gamma}^C((y, v)) = \left\{ \begin{bmatrix} \zeta \\ -(\zeta, \delta)^T(x, u) \end{bmatrix} \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix} \in N_{\text{epi } f}^C((x, u)) \right\}. \quad (12)$$

Then the claimed subgradient and supgradient formulas follow by definition.  $\square$

**Proposition 4.3.** *For any strictly upper radial  $f$ ,*

$$\partial_P f^\Gamma(y) = \left\{ \frac{\zeta}{(\zeta, \delta)^T(x, u)} \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix} \in N_{\text{hypo } f}^P((x, u)), (\zeta, \delta)^T(x, u) > 0 \right\}$$

where  $(x, u) = \Gamma(y, f^\Gamma(y))$ . Likewise, for any strictly lower radial  $f$ ,

$$\partial^P f_\Gamma(y) = \left\{ \frac{\zeta}{(\zeta, \delta)^T(x, u)} \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix} \in N_{\text{epi } f}^P((x, u)), (\zeta, \delta)^T(x, u) < 0 \right\}.$$

*Proof.* Recall that Propositions 2.6 characterized the proximal normal vectors of the radial transformation of a set in terms of the original set. Since the assumed strict increase ensures equality holds in Lemma 4.1, this applies to the epigraph and hypograph of  $f^\Gamma$  and  $f_\Gamma$ , respectively. Thus when  $f$  is strictly upper radial

$$N_{\text{epi } f^\Gamma}^P((y, v)) = \left\{ \left[ -(\zeta, \delta)^T(x, u) \right] \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix} \in N_{\text{hypo } f}^P((x, u)) \right\} \quad (13)$$

and when  $f$  is strictly lower radial

$$N_{\text{hypo } f^\Gamma}^P((y, v)) = \left\{ \left[ -(\zeta, \delta)^T(x, u) \right] \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix} \in N_{\text{epi } f}^P((x, u)) \right\}. \quad (14)$$

Then the claimed subgradient and supgradient formulas follow by definition.  $\square$

## 4.2 Gradients and Hessians for Differentiable Functions

Now we narrow our focus to consider differentiable functions under the radial transformation. Whenever  $f^\Gamma$  is differentiable, a formula for its gradient follows from the subgradient formula in Proposition 4.3. To establish when  $f^\Gamma$  is differentiable, we show that being  $k$  times continuously differentiable (or analytic) is preserved under the radial transformation for appropriate functions. Lastly, we give a formula for the Hessian of the radial transformation of any appropriate twice differentiable function.

As a first step, we give a simple bijection between the graphs (and thus domains) of a function  $f$  and its radial transformation  $f^\Gamma$  whenever  $f$  is continuous and strictly radial.

**Lemma 4.4.** *For any continuous, strictly radial  $f$ ,*

$$\text{graph } f^\Gamma = \Gamma(\text{graph } f).$$

*Hence, if  $y \in \text{dom } f^\Gamma$  then  $y/f^\Gamma(y) \in \text{dom } f$ .*

*Proof.* Noting that  $\text{graph } f = \text{epi } f \cap \text{hypo } f$ , we have

$$\text{graph } f^\Gamma = \text{epi } f^\Gamma \cap \text{hypo } f^\Gamma = \Gamma(\text{epi } f) \cap \Gamma(\text{hypo } f) = \Gamma(\text{epi } f \cap \text{hypo } f) = \Gamma(\text{graph } f)$$

where the second equality follows from Lemma 4.1 and the third follows from (5).  $\square$

This lemma lets us view the graph of the radial transformation as the relation  $\Gamma(\text{graph } f)$ . Applying the implicit function theorem to this relation shows differentiability is preserved under the transformation for appropriate functions. Then leveraging the previous section's results on the proximal subdifferential gives a formula for the gradient of the radial transformation.

**Proposition 4.5.** *Consider any strictly upper radial  $f$  and  $x, y \in \mathcal{E}$  with  $(x, f(x)) = \Gamma(y, f^\Gamma(y))$ . Then  $f$  is  $k$  times continuously differentiable (or analytic) around  $x$  with*

$$(\nabla f(x), -1)^T(x, f(x)) < 0$$

*if and only if  $f^\Gamma = f_\Gamma$  is  $k$  times continuously differentiable (or analytic) around  $y$  with*

$$(\nabla f^\Gamma(y), -1)^T(y, f^\Gamma(y)) < 0$$

*where*

$$\nabla f^\Gamma(y) = \frac{\nabla f(x)}{(\nabla f(x), -1)^T(x, f(x))}.$$

*Proof.* It suffices to only show the forward direction as the duality of the radial function transformation (Theorem 3.2) will then imply the reverse direction. Define the following  $k$  times continuously differentiable (or analytic) function

$$F(y', v') = v' \cdot f(y'/v') - 1.$$

Then from Lemma 4.4, we know  $\text{graph } f^\Gamma = \{(y', v') \in \mathcal{E} \times \mathbb{R}_{++} \mid F(y', v') = 0\}$ . Noting

$$\frac{\partial}{\partial v} F(y, f^\Gamma(y)) = f(y/f^\Gamma(y)) - \nabla f(y/f^\Gamma(y))^T (y/f^\Gamma(y)),$$

we find  $\frac{\partial}{\partial v} F(y, f^\Gamma(y)) = f(x) - \nabla f(x)^T x > 0$ . Thus the implicit function theorem can be applied to produce a  $k$  times continuously differentiable (or analytic) function  $g: U \rightarrow \mathbb{R}_{++}$  for some open neighborhood  $U$  of  $y$  such that

$$\text{graph } f^\Gamma \cap (U \times \mathbb{R}_{++}) = \{(y, g(y)) \mid y \in U\}.$$

As a result,  $f^\Gamma$  must equal  $g$  near  $y$ , and hence is also  $k$  times continuously differentiable (or analytic) near  $y$ .

Now all that remains is to derive our gradient formula and show it satisfies the claimed inequality. Consider any  $y \in \text{dom } f^\Gamma$  and set  $x = y/f^\Gamma(y) \in \text{dom } f$ . The density theorem of proximal calculus [6, Theorem 1.3.1] guarantees a sequence  $y_i \rightarrow y$  exists with all  $\partial_P f^\Gamma(y_i) \neq \emptyset$ . Then letting  $x_i = y_i/f^\Gamma(y_i)$ ,

$$\nabla f^\Gamma(y_i) = \frac{\nabla f(x_i)}{(\nabla f(x_i), -1)^T (x_i, f(x_i))}$$

since  $N_{\text{epi } f}^P(x_i, f(x_i)) \subseteq \{\lambda(\nabla f(x_i), -1) \mid \lambda \geq 0\}$ . Since  $x_i \rightarrow x$ , the continuous differentiability of  $f$  and  $f^\Gamma$  ensures

$$\nabla f^\Gamma(y) = \frac{\nabla f(x)}{(\nabla f(x), -1)^T (x, f(x))}.$$

From this, its immediate that

$$\begin{aligned} (\nabla f^\Gamma(y), -1)^T (y, f^\Gamma(y)) &= \left( \frac{\nabla f(x)}{(\nabla f(x), -1)^T (x, f(x))}, -1 \right)^T \left( \frac{x}{f(x)}, \frac{1}{f(x)} \right) \\ &= \frac{1}{f(x)} \left( \frac{\nabla f(x)^T x}{(\nabla f(x), -1)^T (x, f(x))} - 1 \right) \\ &= \frac{1}{(\nabla f(x), -1)^T (x, f(x))} < 0. \end{aligned} \quad \square$$

We remark that this result does not capture all functions for which the radial transformation is differentiable. For example, consider the strictly upper radial function

$$f(x) = \begin{cases} 1 + \sqrt{1 - x^2} & \text{if } -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

This function is not differentiable everywhere in its domain (namely, it fails at  $x = \pm 1$ ). However, its upper radial transformation is differentiable everywhere as it equals

$$f^\Gamma(y) = \begin{cases} (y^2 + 1)/2 & \text{if } -1 \leq y \leq 1 \\ |y| & \text{otherwise.} \end{cases}$$

Differentiating the gradient formula of Proposition 4.5 directly gives a Hessian formula for the radial transformation of a function.

**Proposition 4.6.** Consider strictly upper radial  $f$  and  $x, y \in \mathcal{E}$  satisfying  $(x, f(x)) = \Gamma(y, f^\Gamma(y))$ . If  $f$  is twice continuously differentiable around  $x$  with

$$(\nabla f(x), -1)^T(x, f(x)) < 0,$$

the Hessian of  $f^\Gamma = f_\Gamma$  at  $y$  is given by

$$\nabla^2 f^\Gamma(y) = \frac{f(x)}{(\nabla f(x), -1)^T(x, f(x))} \cdot J \nabla^2 f(x) J^T$$

where  $J = I - \frac{\nabla f(x)x^T}{(\nabla f(x), -1)^T(x, f(x))}$ .

*Proof.* Denote the bijection relating the domains of  $f^\Gamma$  and  $f$  by  $\pi(y) = y/f^\Gamma(y)$  (as shown by Lemma 4.4). Then the gradient of the radial transformation is

$$\nabla f^\Gamma(y) = \frac{\nabla f(\pi(y))}{(\nabla f(\pi(y)), -1)^T(\pi(y), f(\pi(y)))}.$$

Thus the Jacobian of  $\pi$  is given by

$$\begin{aligned} \nabla \pi(y) &= I/f^\Gamma(y) - y \nabla f^\Gamma(y)^T / f^\Gamma(y)^2 \\ &= \frac{1}{f^\Gamma(y)} \left( I - \frac{y \nabla f(\pi(y))^T}{f^\Gamma(y) (\nabla f(\pi(y)), -1)^T(\pi(y), f(\pi(y)))} \right) \\ &= f(\pi(y)) \left( I - \frac{\pi(y) \nabla f(\pi(y))^T}{(\nabla f(\pi(y)), -1)^T(\pi(y), f(\pi(y)))} \right) \end{aligned}$$

where the third equality uses that  $y/f^\Gamma(y) = \pi(y)$  and  $1/f^\Gamma(y) = f(\pi(y))$  by Lemma 4.4. Let  $g(x) = \nabla f(x)/(\nabla f(x), -1)^T(x, f(x))$  denote  $\nabla f^\Gamma \circ \pi^{-1}$ . Noting that the gradient of  $(\nabla f(x), -1)^T(x, f(x))$  is  $\nabla^2 f(x)^T x$ , the Jacobian of  $g$  is given by

$$\begin{aligned} \nabla g(x) &= \frac{\nabla^2 f(x)}{(\nabla f(x), -1)^T(x, f(x))} - \frac{\nabla f(x)x^T \nabla f^2(x)}{(\nabla f(x), -1)^T(x, f(x))^2} \\ &= \frac{1}{(\nabla f(x), -1)^T(x, f(x))} \left( I - \frac{\nabla f(x)x^T}{(\nabla f(x), -1)^T(x, f(x))} \right) \nabla^2 f(x) \end{aligned}$$

Since  $\nabla f^\Gamma(y) = g(\pi(y))$ , the Hessian of  $f^\Gamma$  is  $\nabla g(\pi(y)) \nabla \pi(y)$ , giving the claimed formula.  $\square$

### 4.3 Optimality Under the Radial Transformation

Before addressing optimality under our radial duality, we observe that inequalities between functions are reversed by applying either radial function transformation. This mirrors (4), where we saw the radial set transformation preserves inclusions between sets. We say  $f \leq g$  if  $f(x) \leq g(x)$  for all  $x \in \mathcal{E}$ .

**Lemma 4.7.** For any functions  $f, g$ , if  $f \leq g$ , then  $g^\Gamma \leq f^\Gamma$  and  $g_\Gamma \leq f_\Gamma$ .

*Proof.* Notice that  $f \leq g$  is equivalent to  $\text{epi } g \subseteq \text{epi } f$ . Then (4) gives  $\Gamma(\text{epi } g) \subseteq \Gamma(\text{epi } f)$ . Therefore  $f^\Gamma(y) \geq g^\Gamma(y)$  for all  $y \in \mathcal{E}$ .  $\square$

Now we consider how the extreme values and points of a function and its radial transformations relate. First, we show for radial functions, the supremum value of  $f$  equals the reciprocal of the infimum value of  $f^\Gamma$  in Proposition 4.8. Then Proposition 4.9 shows the maximizers of  $f$  are related to minimizers of  $f^\Gamma$  by the radial point transformation.

**Proposition 4.8.** *For any function  $f$ ,  $(\inf f) \cdot (\sup f^\Gamma) = 1$  where we let  $\infty \cdot 0 = 0 \cdot \infty = 1$ . Further, if  $f$  is upper radial,  $(\sup f) \cdot (\inf f^\Gamma) = 1$ . Likewise,  $(\sup f) \cdot (\inf f_\Gamma) = 1$  and if  $f$  is lower radial,  $(\inf f) \cdot (\sup f_\Gamma) = 1$ .*

*Proof.* Observe that  $f \geq \inf f$  and so applying Lemma 4.7 implies

$$f^\Gamma \leq (\inf f)^\Gamma = 1/\inf f.$$

Now we show the  $\geq$  inequality (which is trivial if  $\inf f(x) = \infty$ ). Suppose  $\inf f < \infty$ . Let  $x_i$  be a sequence with  $\lim f(x_i) = \inf f$ . Then fix any  $\epsilon > 0$  and set  $y_i = x_i/(f(x_i) + \epsilon)$ . Observe that  $v \cdot f(y_i/v) < 1$  when  $v = 1/(f(x_i) + \epsilon)$ . Therefore  $f^\Gamma(y_i) \geq 1/(f(x_i) + \epsilon)$  and so taking the limit as  $\epsilon \rightarrow 0$  implies  $\sup f^\Gamma \geq 1/\inf f$ . Lastly, supposing  $f$  is upper radial, the upper radial transformation is dual by Theorem 3.2, and so

$$1 = (\inf f^\Gamma) \cdot (\sup f^{\Gamma\Gamma}) = (\inf f^\Gamma) \cdot (\sup f). \quad \square$$

**Proposition 4.9.** *For any upper radial  $f$  with  $\sup f \in \mathbb{R}_{++}$ ,*

$$(\operatorname{argmin} f^\Gamma) \times \{\inf f^\Gamma\} \subseteq \Gamma((\operatorname{argmax} f) \times \{\sup f\}).$$

*Likewise for any lower radial  $f$  with  $\inf f \in \mathbb{R}_{++}$ ,*

$$(\operatorname{argmax} f_\Gamma) \times \{\sup f_\Gamma\} \subseteq \Gamma((\operatorname{argmin} f) \times \{\inf f\}).$$

*If  $f^p(y, \cdot)$  is strictly increasing on its domain, equality holds in both cases.*

*Proof.* Consider any  $(y, v) \in (\operatorname{argmin} f^\Gamma) \times \{\inf f^\Gamma\}$  and set  $(x, u) = \Gamma(y, v)$ . Then  $(y, v) \in \operatorname{epi} f^\Gamma$ , and so Lemma 4.1 ensures  $(x, u) \in \operatorname{hypo} f$ . Therefore  $x$  attains the maximum value of  $f$  by Proposition 4.8. Hence  $(x, u) \in (\operatorname{argmax} f) \times \{\sup f\}$ .

When  $f^p(y, \cdot)$  is strictly increasing, equality holds in Lemma 4.1 and the above argument can be repeated in reverse.  $\square$

For nonconvex optimization problems, finding global solutions is often intractable and so the focus of many optimization methods is on finding stationary points (that is, points with a zero sub(sup)gradient in their sub(sup)differential). Just as optimal solutions were related between the primal and radial dual problems, stationary points are also directly related by the radial point transformation.

**Proposition 4.10.** *For any strictly upper radial  $f$ ,*

$$\{(y, f^\Gamma(y)) \in \mathcal{E} \times \mathbb{R}_{++} \mid 0 \in \partial_P f^\Gamma(y)\} = \Gamma\{(x, f(x)) \in \mathcal{E} \times \mathbb{R}_{++} \mid 0 \in \partial^P f(x)\}.$$

*Likewise for any strictly lower radial  $f$ ,*

$$\{(y, f_\Gamma(y)) \in \mathcal{E} \times \mathbb{R}_{++} \mid 0 \in \partial^P f_\Gamma(y)\} = \Gamma\{(x, f(x)) \in \mathcal{E} \times \mathbb{R}_{++} \mid 0 \in \partial_P f(x)\}.$$

*Proof.* Follows from Proposition 4.3.  $\square$



## 5 Characterizing Epigraph Reshaping Transformations

In this final section, we consider a broader class of transformations given by reshaping a function's epigraph via some mapping  $G$ . In particular, given a generic optimization problem

$$p^* = \max_{x \in \mathcal{E}} f(x), \quad (15)$$

we transform  $f$  by reshaping its epigraph into the hypograph of a new function

$$f^G(y) = \sup\{v \mid (y, v) \in G(\text{epi } f)\}.$$

If  $G(\text{epi } f)$  is indeed the a function's hypograph, the identity hypo  $f^G = G(\text{epi } f)$  holds. Then the transformed optimization problem is defined as

$$d^* = \min_{y \in \mathcal{E}} f^G(y). \quad (16)$$

Paralleling the development of the radial function transformation  $f^\Gamma$ , we would like to relate minimizers of  $f^G$  to maximizers of  $f$  and vice versa through the mapping  $G$ . To this end, we assume **invertibility**:

$$G \text{ is a bijection.} \quad (A1)$$

Whenever the given function  $f$  is concave, we want to preserve this structure by having  $f^G$  be convex. To ensure this, we assume  $G$  is **convexity preserving**: for any  $S \subseteq \text{dom } G$ ,

$$S \text{ is convex} \implies GS \text{ is convex.} \quad (A2)$$

Lastly, we need a relationship between minimizers of  $f$  and maximizers of  $f^G$ . This follows by assuming  $G$  is **height reversing**: for any pairs  $(x, u) = G^{-1}(y, v) \in \text{dom } G$  and  $(x', u') = G^{-1}(y', v') \in \text{dom } G$ ,

$$u \geq u' \implies v \leq v'. \quad (A3)$$

Under these three assumptions, any function  $f$  satisfying the identity hypo  $f^G = G(\text{epi } f)$  must have  $\text{argmax } f \times \{p^*\} = G^{-1}(\text{argmin } f^G \times \{d^*\})$ . Hence the problems (15) and (16) are equivalent and converting points between these problems only requires evaluating  $G$  or its inverse.

First as an example, we consider transforming functions  $f: \mathcal{E} \rightarrow \mathbb{R} \cup \{\pm\infty\}$  mapping into the extended reals. Then  $G$  must map  $\mathcal{E} \times \mathbb{R}$  into  $\mathcal{E} \times \mathbb{R}$ . We may additionally want to impose a condition requiring the function transformation  $f^G$  satisfies hypo  $f^G = G(\text{epi } f)$  for a reasonably large class a functions. Namely, we assume the function transformation is **well-defined**: for all linear  $f: \mathcal{E} \rightarrow \mathbb{R}$ ,

$$\text{hypo } f^G = G(\text{epi } f). \quad (A4)$$

The Fundamental Theorem of Affine Geometry (stated below) gives us an immediate way to characterize what possible transformations satisfy these four assumptions. See [2, 10] as references.

**Theorem 5.1.** *For  $n \geq 2$ , if  $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a bijective, convexity preserving map, then  $F$  is an affine transformation.*

Utilizing this, we find that any transformation satisfying these four assumptions must be producing an affinely shifted version of the original problem (16). Essentially, any duality of this form amounts to the trivial duality between maximizing a function and minimizing its negative. This is proven in Section 5.1.

**Theorem 5.2.** *Consider any map  $G: \mathcal{E} \times \mathbb{R} \rightarrow \mathcal{E} \times \mathbb{R}$  satisfying (A1), (A2), and (A3). Then  $G$  is an affine map*

$$G(x, u) = \begin{bmatrix} A & \alpha \\ 0 & c \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \begin{bmatrix} b \\ d \end{bmatrix}$$

with  $c < 0$ . Furthermore, if (A4) holds, then  $\alpha = 0$  and

$$f^G(y) = cf(A^{-1}(y - b)) + d.$$

Thus there are notable limitations on what a transformation satisfying these four assumptions can accomplish. However, the following theorem provides an alternative to using the Fundamental Theorem of Affine Geometry and facilitates studying more general transformations. See [3] for a reference or [13] for the original version of this result, which takes a more general perspective based in projective spaces.

**Theorem 5.3.** *For  $n \geq 2$ , if for some convex set  $K \subseteq \mathbb{R}^n$  with nonempty interior,  $F: K \rightarrow \mathbb{R}^n$  is an injective, convexity preserving map, then  $F$  is a fractional linear map.*

This indicates that there is more potential for interesting transformations if we can restrict our assumptions to a convex subset of  $\mathcal{E} \times \mathbb{R}$  (in our case,  $\mathcal{E} \times \mathbb{R}_{++}$ ). To this end, we now consider transforming functions  $f: \mathcal{E} \rightarrow \overline{\mathbb{R}}_{++}$  mapping into the extended positive reals.

We also suppose the transformed function  $f^G: \mathcal{E} \rightarrow \overline{\mathbb{R}}_{++}$  maps into the extend positive reals, and so  $G$  maps  $\mathcal{E} \times \mathbb{R}_{++}$  into  $\mathcal{E} \times \mathbb{R}_{++}$ . Our first three assumptions (namely, invertibility (A1), convexity preserving (A2), and height reversing (A3)) extend directly to  $G$  having restricted domain and codomain. We consider the following assumption paralleling (A4) to ensure the function transformation is **well-defined** for a basic class of linear-like functions: for all linear  $f: \mathcal{E} \rightarrow \mathbb{R}$ ,

$$\text{hypo } (f_+)^G = G(\text{epi } f_+) \tag{B4}$$

where  $(\cdot)_+ = \max\{\cdot, 0\}$  denotes nonnegative thresholding. Under these four assumptions, we find that any such transformation of nonnegative-valued optimization problems must produce an affinely shifted version of the upper radial function transformation. This is proven in Section 5.2.

**Theorem 5.4.** *Consider any map  $G: \mathcal{E} \times \mathbb{R}_{++} \rightarrow \mathcal{E} \times \mathbb{R}_{++}$  satisfying (A1), (A2), and (A3). Then  $G$  is a fractional linear map*

$$G(x, u) = \frac{(Ax + \alpha u + b, d)}{u}$$

with  $d > 0$ . Furthermore, if (B4) holds, then  $b = 0$  and the function transformation is given by

$$f^G(y) = \frac{1}{d} f^\Gamma(A^{-1}(y - \alpha)).$$

This result is similar in spirit to [4, Theorem 5], avoiding their reliance on nonnegative convex functions with value 0 at the origin. Thus the radial set transformation  $\Gamma$  and function transformation  $f^\Gamma$  provide the unique mechanism for deriving equivalent nonnegative-valued optimization problems.

## 5.1 Proof of Theorem 5.2

From assumptions (A1) and (A2), Theorem 5.1 immediately implies that

$$G(x, u) = \begin{bmatrix} A & \alpha \\ \beta^T & c \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \begin{bmatrix} b \\ d \end{bmatrix}.$$

Further, (A3) becomes for all  $(x, u), (x', u') \in \mathcal{E} \times \mathbb{R}$ ,

$$u \leq u' \implies \beta^T x + cu \geq \beta^T x' + cu'.$$

Hence, we must have  $\beta = 0$  and  $c \leq 0$ . Moreover, since  $G$  is a bijection,  $c < 0$ .

Now additionally assume (A4), that this transformation is well-defined for all linear functions. Observe that the inverse of  $G$  is given by

$$G^{-1}(y, v) = \begin{bmatrix} A & \alpha \\ 0 & c \end{bmatrix}^{-1} \begin{bmatrix} y - b \\ v - d \end{bmatrix} = \begin{bmatrix} A^{-1}(y - b - \alpha(v - d)/c) \\ (v - d)/c \end{bmatrix}.$$

Suppose for contradiction that  $\alpha \neq 0$ . Then the linear function  $f(x) = -2\alpha^T Ax / \|\alpha\|^2$  has

$$\begin{aligned} G(\text{epi } f) &= \{(y, v) \mid G^{-1}(y, v) \in \text{epi } f\} \\ &= \{(y, v) \mid -2\alpha^T(y - b - \alpha(v - d)/c) / \|\alpha\|^2 \leq (v - d)/c\} \\ &= \{(y, v) \mid -2\alpha^T(y - b) \leq -\|\alpha\|^2(v - d)/c\}. \end{aligned}$$

However, this is not the hypograph of any function as  $(b, d) \in G(\text{epi } f)$ , but  $(b, d - 1) \notin G(\text{epi } f)$ , which contradicts (A4). Thus we must have  $\alpha = 0$ . From this, we have  $G^{-1}(y, v) = (A^{-1}(y - b), (v - d)/c)$ . Then  $f^G(y)$  equals  $cf(A^{-1}(y - b)) + d$  since

$$\text{hypo } f^G = \{(y, v) \mid f(A^{-1}(y - b)) \leq (v - d)/c\} = \{(y, v) \mid G^{-1}(y, v) \in \text{epi } f\} = G(\text{epi } f).$$

## 5.2 Proof of Theorem 5.4

From assumptions (A1) and (A2), Theorem 5.3 immediately implies that

$$G(x, u) = \frac{\begin{bmatrix} A & \alpha \\ \beta^T & c \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \begin{bmatrix} b \\ d \end{bmatrix}}{\eta^T x + gu + h}.$$

Since  $G$  is a bijection, all  $(x, u) \in \mathcal{E} \times \mathbb{R}_{++}$  must have  $G(x, u) \in \mathcal{E} \times \mathbb{R}_{++}$ , and so

$$\frac{\beta^T x + cu + d}{\eta^T x + gu + h} > 0.$$

Thus  $\beta = \eta = 0$ . Then the mapping  $\sigma(u) = (cu + d)/(gu + h)$  must be a bijection from  $\mathbb{R}_{++}$  to  $\mathbb{R}_{++}$ . Therefore  $\sigma$  must be in one of the following two forms: either  $g = 0$  and so  $\sigma(u) = cu/h$  with  $c/h > 0$ , or  $g \neq 0$  and so  $\sigma(u) = d/(gu)$  with  $d/g > 0$ .

From (A3), the latter of these two possibilities must be the case. Thus  $g \neq 0$  and so without loss of generality, we can suppose  $g = 1$  and  $d > 0$ . Hence,

$$G(x, u) = \frac{(Ax + \alpha u + b, d)}{u}.$$

Now additionally assume that this transformation is well-defined for all linear functions (after thresholding to nonnegative values), namely (B4). Observe that the inverse of  $G$  is given by

$$G^{-1}(y, v) = \frac{(A^{-1}(y - bv - \alpha), 1)}{dv}.$$

Suppose for contradiction that  $b \neq 0$ . Then the function  $f(x) = (-2b^T Ax / \|b\|^2)_+$  has

$$\begin{aligned} G(\text{epi } f) &= \{(y, v) \in \mathcal{E} \times \mathbb{R}_{++} \mid G^{-1}(y, v) \in \text{epi } f\} \\ &= \{(y, v) \in \mathcal{E} \times \mathbb{R}_{++} \mid (-2b^T(y - bv - \alpha) / (dv\|b\|^2))_+ \leq 1/(dv)\} \\ &= \{(y, v) \in \mathcal{E} \times \mathbb{R}_{++} \mid -2b^T(y - bv - \alpha) / (dv\|b\|^2) \leq 1/(dv)\} \\ &= \{(y, v) \in \mathcal{E} \times \mathbb{R}_{++} \mid -2b^T(y - \alpha) \leq -\|b\|^2 v\}. \end{aligned}$$

However, this is not the hypograph of any function as  $(b - \alpha, 2) \in G(\text{epi } f)$ , but  $(b - \alpha, 1) \notin G(\text{epi } f)$ , which contradicts (B4). Thus we must have  $b = 0$ . From this, we have  $G^{-1}(y, v) = (A^{-1}(y - \alpha), 1)/(dv)$ . Therefore  $f^G$  must be an affine translation of  $f^\Gamma$  since

$$\begin{aligned} f^G(y) &= \sup\{v > 0 \mid G^{-1}(y, v) \in \text{epi } f\} \\ &= \sup\{v > 0 \mid f(A^{-1}(y - \alpha)/(dv)) \leq 1/(dv)\} \\ &= \frac{1}{d} \sup\{w > 0 \mid f(A^{-1}(y - \alpha)/w) \leq 1/w\} \\ &= \frac{1}{d} f^\Gamma(A^{-1}(y - \alpha)). \end{aligned}$$

**Acknowledgments.** The author thanks Jim Renegar broadly for inspiring this work and concretely for providing feedback on multiple drafts. Further, Jim had the valuable idea to use the implicit function theorem to simplify and improve the proof of Proposition 4.5.

## References

- [1] Aleksandr Y Aravkin, James V Burke, Dmitriy Drusvyatskiy, Michael P Friedlander, and Kellie J MacPhee. Foundations of gauge and perspective duality. *SIAM Journal on Optimization*, 28(3):2406–2434, 2018.
- [2] Emil Artin. *Geometric Algebra*. Interscience. Wiley, 1988.
- [3] Shiri Artstein-Avidan, Dan Florentin, and Vitali Milman. *Order Isomorphisms on Convex Functions in Windows*, pages 61–122. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.
- [4] Shiri Artstein-Avidan and Vitali Milman. Hidden structures in the class of convex functions and a new duality transform. *Journal of the European Mathematical Society*, 013(4):975–1004, 2011.
- [5] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004. p.89.
- [6] Francis H. Clarke, Yuri S. Ledyaev, Ronald J. Stern, and Peter R. Wolenski. *Nonsmooth Analysis and Control Theory*. Springer-Verlag, Berlin, Heidelberg, 1998.
- [7] Robert M. Freund. Dual gauge programs, with applications to quadratic programming and the minimum-norm problem. *Math. Program.*, 38:47–67, 1987.
- [8] Benjamin Grimmer. Radial Subgradient Method. *SIAM Journal on Optimization*, 28(1):459–469, 2018.
- [9] Benjamin Grimmer. Radial Duality Part II: Applications and Algorithms. *arXiv:2104.11185*, April 2021.
- [10] Viktor Prasolov and Vladimir Tikhomirov. *Geometry*. Providence, R.I. : American Mathematical Society ; Oxford University Press, 2001.
- [11] James Renegar. “Efficient” Subgradient Methods for General Convex Optimization. *SIAM Journal on Optimization*, 26(4):2649–2676, 2016.

- [12] James Renegar. Accelerated first-order methods for hyperbolic programming. *Math. Program.*, 173(1-2):1–35, 2019.
- [13] Bernard Shiffman. Synthetic Projective Geometry and Poincaré’s Theorem on Automorphisms of the Ball. *L’Enseignement Mathématique.*, 41:201–215, 1995.

## A Proofs Computing Some Radial Set Transformations

Here we present direct proofs of our claimed Propositions 2.1, 2.2, and 2.4 characterizing the result of applying the projective transformation  $\Gamma$  to convex sets, halfspaces, and ellipsoids, respectively.

### A.1 Proof of Proposition 2.1

It suffices to show  $S$  being convex implies  $\Gamma S$  is convex, since the duality of the radial set transformation (3) will then imply the reverse direction. Consider any  $(y, v), (y', v') \in \Gamma S$ . Let  $(x, u) = \Gamma(y, v)$  and  $(x', u') = \Gamma(y', v')$ . Then since  $(x, u), (x', u') \in S$ , all  $0 \leq \lambda \leq 1$  have

$$\lambda(x, u) + (1 - \lambda)(x', u') \in S.$$

Therefore the line segment between  $(y, v)$  and  $(y', v')$  lies in  $\Gamma S$  as

$$\begin{aligned} \Gamma S &\ni \frac{(\lambda x + (1 - \lambda)x', 1)}{\lambda u + (1 - \lambda)u'} \\ &= \frac{(\lambda y/v + (1 - \lambda)y'/v', 1)}{\lambda/v + (1 - \lambda)/v'} \\ &= \frac{\lambda/v}{\lambda/v + (1 - \lambda)/v'}(y, v) + \frac{(1 - \lambda)/v'}{\lambda/v + (1 - \lambda)/v'}(y', v'). \end{aligned}$$

### A.2 Proof of Proposition 2.2

It suffices to show  $S$  being a halfspace implies  $\Gamma S$  is a halfspace, since the duality of the radial set transformation (3) will then imply the reverse direction. Applying  $\Gamma$  to the definition of  $S$  yields

$$\begin{aligned} \Gamma S &= \left\{ \frac{(x', 1)}{u'} \in \mathcal{E} \times \mathbb{R}_{++} \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix}^T \begin{bmatrix} x' - x \\ u' - u \end{bmatrix} \leq 0 \right\} \\ &= \left\{ (y', v') \in \mathcal{E} \times \mathbb{R}_{++} \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix}^T \begin{bmatrix} y'/v' - x \\ 1/v' - u \end{bmatrix} \leq 0 \right\} \\ &= \left\{ (y', v') \in \mathcal{E} \times \mathbb{R}_{++} \mid \begin{bmatrix} \zeta \\ \delta \end{bmatrix}^T \begin{bmatrix} y' - v'x \\ 1 - v'u \end{bmatrix} \leq 0 \right\} \\ &= \left\{ (y', v') \in \mathcal{E} \times \mathbb{R}_{++} \mid \begin{bmatrix} \zeta \\ -(\zeta, \delta)^T(x, u) \end{bmatrix}^T \begin{bmatrix} y' \\ v' \end{bmatrix} + \delta \leq 0 \right\} \\ &= \left\{ (y', v') \in \mathcal{E} \times \mathbb{R}_{++} \mid \begin{bmatrix} \zeta \\ -(\zeta, \delta)^T(x, u) \end{bmatrix}^T \begin{bmatrix} y' - y \\ v' - v \end{bmatrix} \leq 0 \right\}. \end{aligned}$$

### A.3 Proof of Proposition 2.4

It suffices to show  $S$  being an ellipsoid in  $\mathcal{E} \times \mathbb{R}_{++}$  implies  $\Gamma S$  is an ellipsoid  $\mathcal{E} \times \mathbb{R}_{++}$ , since the duality of the radial set transformation (3) will then imply the reverse direction. Denote the blocks of  $H$  by  $\begin{bmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{bmatrix}$  and define the following matrix

$$G = \begin{bmatrix} H_{11} & -\begin{bmatrix} H_{11} \\ H_{12} \end{bmatrix}^T \begin{bmatrix} x \\ u \end{bmatrix} \\ -\begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} H_{11} \\ H_{12} \end{bmatrix} & \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} - 1 \end{bmatrix},$$

related to the radially dual ellipsoid. In particular, for any ellipsoid  $S \subseteq \mathcal{E} \times \mathbb{R}_{++}$  defined by (7),  $\Gamma S$  is the following ellipsoid in  $\mathcal{E} \times \mathbb{R}_{++}$  with center  $\begin{bmatrix} y \\ v \end{bmatrix} = G^{-1} \begin{bmatrix} -H_{12} \\ H_{12}^T x + H_{22} u \end{bmatrix}$

$$\Gamma S = \left\{ (y', v') \in \mathcal{E} \times \mathbb{R}_{++} \mid \begin{bmatrix} y' - y \\ v' - v \end{bmatrix}^T \left( \frac{G}{\begin{bmatrix} y \\ v \end{bmatrix}^T G \begin{bmatrix} y \\ v \end{bmatrix} - H_{22}} \right) \begin{bmatrix} y' - y \\ v' - v \end{bmatrix} \leq 1 \right\}.$$

First, we observe that  $G$  is indeed positive definite. Since  $H$  is positive definite, considering its Schur complements ensures  $H_{11}$  is positive definite and  $H_{22} - H_{12}^T H_{11}^{-1} H_{12} > 0$ . Likewise,  $G$  is positive definite if  $H_{11}$  is positive definite and

$$\left( \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} - 1 \right) - \left( \begin{bmatrix} H_{11} & H_{12} \end{bmatrix}^T \begin{bmatrix} x \\ u \end{bmatrix} \right)^T H_{11}^{-1} \left( \begin{bmatrix} H_{11} & H_{12} \end{bmatrix}^T \begin{bmatrix} x \\ u \end{bmatrix} \right) > 0.$$

Simplifying this condition for  $G$  to be positive definite yields the equivalent inequality  $u^2(H_{22} - H_{12}^T H_{11}^{-1} H_{12}) > 1$ . This must be the case since  $S \subseteq \mathcal{E} \times \mathbb{R}_{++}$  and so  $H_{22} - H_{12}^T H_{11}^{-1} H_{12} > 1/u^2$ .

Applying  $\Gamma$  to the definition of  $S$  and completing the square gives the claim as

$$\begin{aligned} \Gamma S &= \left\{ \frac{(x', 1)}{u'} \in \mathcal{E} \times \mathbb{R}_{++} \mid \begin{bmatrix} x' - x \\ u' - u \end{bmatrix}^T \begin{bmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{bmatrix} \begin{bmatrix} x' - x \\ u' - u \end{bmatrix} \leq 1 \right\} \\ &= \left\{ (y', v') \in \mathcal{E} \times \mathbb{R}_{++} \mid \begin{bmatrix} y'/v' - x \\ 1/v' - u \end{bmatrix}^T \begin{bmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{bmatrix} \begin{bmatrix} y'/v' - x \\ 1/v' - u \end{bmatrix} \leq 1 \right\} \\ &= \left\{ (y', v') \in \mathcal{E} \times \mathbb{R}_{++} \mid \begin{bmatrix} y' - v'x \\ 1 - v'u \end{bmatrix}^T \begin{bmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{bmatrix} \begin{bmatrix} y' - v'x \\ 1 - v'u \end{bmatrix} \leq v'^2 \right\} \\ &= \left\{ (y', v') \in \mathcal{E} \times \mathbb{R}_{++} \mid \begin{bmatrix} y' \\ v' \end{bmatrix}^T G \begin{bmatrix} y' \\ v' \end{bmatrix} + 2H_{12}^T y' - 2(H_{12}^T x + H_{22} u)v' \leq -H_{22} \right\}. \end{aligned}$$