HYPERBOLIC PROGRAMS, AND THEIR DERIVATIVE RELAXATIONS

JAMES RENEGAR

Abstract. We study the algebraic and facial structures of hyperbolic programs, and examine natural relaxations of hyperbolic programs, the relaxations themselves being hyperbolic programs.

1. Introduction

Hyperbolic programming was introduced by Güler [6] in the context of interior-point methods. His inspiration drew partly from work arising in the study of hyperbolic pde's; in particular, from work of Gårding [5].

The richness of hyperbolic programming was further explored by Bauschke, Güler, Lewis and Sendov [1]. They initiated an intriguing theory in the vein of general convex analysis.

We continue the exploration of hyperbolic programming, influenced greatly by the above works. The present paper lays out some of the basic structure of hyperbolic programs.

For coherence, we reprove some results found in the above papers. Perhaps noteworthy in this regard is that we reprove Gårding's key results, with arguments that while entirely inspired by his proofs, are considerably briefer.

2. Fundamentals

Let $\mathcal{E}$ denote a finite-dimensional Euclidean space.

> A homogeneous polynomial $p : \mathcal{E} \to \mathbb{R}$ is said to be *hyperbolic* if there exists a direction $e \in \mathcal{E}$, $p(e) \neq 0$, with the property that for each $x \in \mathcal{E}$, the univariate polynomial $t \mapsto p(x + te)$ has only real roots (i.e., each root has no imaginary part). The polynomial is said to be hyperbolic *in direction* e.

1991 *Mathematics Subject Classification.* 90C05, 90C22, 90C25, 52A41, 52B15.

Key words and phrases. hyperbolicity cone, hyperbolic polynomial, conic programming, convex optimization.

The pdf file is hyperlinked.

I especially wish to thank Chek Beng Chua for illuminating discussions during the formative stages of this work.

Research supported by NSF Grant #CCR-9901941.

In this work, p always denotes a hyperbolic polynomial of degree at least 1. We assume $p(e) > 0$ (replace p with $-p$ if necessary).

Particularly relevant examples pertain to:

- Linear Programming (LP): $\mathcal{E} = \mathbb{R}^n$, $p(x) = x_1 \cdots x_n$,
  $e =$ any vector with only positive coordinates.
- Semi-Definite Programming (SDP):
  $\mathcal{E} = \mathbb{S}^{n \times n}$ (vector space of $n \times n$ symmetric matrices), $p(x) = \det(x)$,
  $e =$ any symmetric matrix with only positive eigenvalues.

Clearly, if p_1 and p_2 are hyperbolic in direction e, then so is $p_1 p_2$.

The reader should keep in mind that because of homogeneity, we could equivalently define a hyperbolic polynomial by replacing "for each $x \in \mathcal{E}$, $t \mapsto p(x + te)$ has only real roots," with "for each $x \in \mathcal{E}$, $t \mapsto p(e + tx)$ has only real roots."

For motivation, we rely on terminology familiar from SDP.

> The univariate functional $\lambda \mapsto p(\lambda e - x)$ is the *characteristic polynomial* of x (with respect to p, in direction e). The roots of the characteristic polynomial are the *eigenvalues* of x.

Thus, a hyperbolic polynomial is a homogeneous polynomial with the property that for each x, all of the eigenvalues are real (with respect to some direction e, where $p(e) > 0$).

Let n denote the degree of p. Write the eigenvalues of x as $\lambda_1(x) \leq \ldots \leq \lambda_n(x)$, counting multiplicities. For clarity, we sometimes write the smallest eigenvalue as $\lambda_{\min}(x)$.

Two observations:

$$\lambda_j(sx + te) = \begin{cases} s\lambda_j(x) + t & \text{if } s \geq 0 \\ s\lambda_{n-j}(x) + t & \text{if } s \leq 0, \end{cases}$$

$$p(x) = p(e) \prod_j \lambda_j(x).$$

The eigenvalues are continuous in x. Indeed, for any family of univariate polynomials $\lambda \mapsto \sum_i a_i(x)\lambda^i$ in which the coefficients $a_i(x)$ vary continuously with x, and in which the leading coefficient is constant, the (complex) roots vary continuously with x (c.f., [9], Thm. 1.3.1).

> The set $\Lambda_{++} := \{x : \lambda_{\min}(x) > 0\}$ is the *hyperbolicity cone* (for p in direction e).

Obviously, $e \in \Lambda_{++}$. Note that if $x \in \Lambda_{++}$, then $p(x) > 0$ (because $p(x) = p(e) \prod_j \lambda_j(x)$). Also, observe Λ_{++} is indeed a cone, i.e., if $x \in \Lambda_{++}$, then $tx \in \Lambda_{++}$ for all $t > 0$.

In LP, where $p(x) = x_1 \cdots x_n$ and e is any vector each of whose coordinates is positive, $\Lambda_{++} = \mathbb{R}^n_{++}$, the strictly positive orthant. In SDP, where $p(x) = \det(x)$ and e is any $n \times n$ symmetric matrix each of whose eigenvalues is positive, $\Lambda_{++} = \mathbb{S}^{n\times n}_{++}$, the set of all $n \times n$ symmetric matrices each of whose eigenvalues is positive.

Proposition 1. *The hyperbolicity cone is the connected component of* $\{x : p(x) \neq 0\}$ *containing* e.

Proof. Let S denote the connected component containing e.

Since x has 0 as an eigenvalue only if $p(x) = 0$, and since $e \in \Lambda_{++}$, it follows from continuity of eigenvalues that $S \subseteq \Lambda_{++}$.

To understand why Λ_{++} is a subset of S, consider $x \in \Lambda_{++}$ and let ℓ be the line segment with endpoints x, e. For sufficiently large $\bar{t} > 0$, all $y \in \ell$ satisfy $p(y + \bar{t}e) > 0$. Also, since $x, e \in \Lambda_{++}$, we know $x + te, e + te \in \Lambda_{++}$ whenever $t \geq 0$, implying $p(x + te), p(e + te) > 0$ whenever $t \geq 0$. Thus, the segments $\{x + te : 0 \leq t \leq \bar{t}\}$, $\{y + \bar{t}e : y \in \ell\}$, and $\{e + te : 0 \leq t \leq \bar{t}\}$ form a path from x to e on which p remains strictly positive. □

Define

$$\Lambda_+ := \{x : \lambda_{\min}(x) \geq 0\}.$$

This is the closure of Λ_{++}. Indeed, if $x \in \Lambda_+$, then $x + te \in \Lambda_{++}$ for all $t > 0$, showing Λ_+ is contained in the closure. That the eigenvalues vary continuously with x implies Λ_+ contains the closure.

The following theorem is the cornerstone of hyperbolic programming.

Theorem 2 (Gårding). *Hyperbolicity cones are convex.*

Since Λ_+ is the closure of Λ_{++}, it follows that Λ_+ is convex, too.

We provide a simplified version of Gårding's proof. The arguments go by way of imaginary numbers. It would be nice if a proof could be made which is more consistent with the spirit of optimization.

Theorem 2 is a corollary of the following result (which has uses beyond establishing Theorem 2).

Theorem 3 (Gårding). *If* $x \in \Lambda_{++}$, *then* p *is hyperbolic in direction* x. *Moreover, the hyperbolicity cone in direction* x *is* Λ_{++} *(i.e., the same cone as in direction* e*).*

Proof. Assume $x \in \Lambda_{++}$ and let y be an arbitrary point. We know $p(x) > 0$. It remains to show $r \mapsto p(rx + y)$ has only real roots.

Let $i := \sqrt{-1}$ and fix $\alpha > 0$. We claim that for all non-negative real numbers s, all roots of $r \mapsto p(\alpha ie + rx + sy)$ have negative imaginary part. This is true for $s = 0$ due to $x \in \Lambda_{++}$ and homogeneity (in fact, for $s = 0$, all roots are purely negative imaginary). Consequently, if for some $s > 0$, a root of $r \mapsto p(\alpha ie + rx + sy)$ had non-negative imaginary part, then by the continuity of roots with respect to s, there would be an intermediate value $0 < s' \leq s$ for which a root r' is real. Clearly, s' would be a root of $s \mapsto p(\alpha i e + r'x + sy)$, implying $t = \alpha i$ would be a root of $t \mapsto p(te + z)$, where $z := r'x + s'y$. However, with r' real, z would be a real vector. In other words, there would exist $z \in \mathcal{E}$ for which $t \mapsto p(te + z)$ has a non-real root, contradicting hyperbolicity of p in direction e. Hence, for each $s \geq 0$, all of the roots of $r \mapsto p(\alpha i e + rx + sy)$ indeed have negative imaginary part.

In particular, all roots of $r \mapsto p(\alpha i e + rx + y)$ have negative imaginary part, regardless of the particular positive value α. Consequently, letting α go to 0, continuity of roots with respect to α implies all roots of $r \mapsto p(rx + y)$ have non-positive imaginary part. However, $r \mapsto p(rx + y)$ is a real polynomial, and the non-real roots of real polynomials occur in conjugate pairs. Since no roots of $r \mapsto p(rx + y)$ have positive imaginary part, all roots must thus be real. As y was an arbitrary point, we have that p is hyperbolic in direction x.

The final statement of the theorem is immediate from Proposition 1. $\square$

Proof of Theorem 2. For $x, y \in \Lambda_{++}$ and $r, s > 0$, we wish to show $rx + sy \in \Lambda_{++}$. Theorem 3 implies that without loss of generality, we may assume $y = e$. However, the eigenvalues of $rx + se$ in direction e are $r\lambda_i(x) + s > 0$. $\square$

Later, we consider various directions $e \in \Lambda_{++}$, as is allowed due to Theorem 3. When required for clarity, we make dependence on e explicit, writing, for example, $\lambda_{j,e}(x)$.

The following corollary records a fact immediately evident from the preceeding results. We bother to state the corollary only because doing so provides an expedient way to refer to the fact in later arguments.

Corollary 4. *For every $e \in \Lambda_{++}$ and for every point x, the univariate polynomial $t \mapsto p(e + tx)$ has only real roots.*

The results above provide a mechanism for passing between hyperbolicity developed in the homogeneous (i.e., conic) setting – the approach we pursue – and hyperbolicity developed affinely. For the affine setting, one defines a (not necessarily homogeneous) polynomial q to be hyperbolic if there is a point d, $q(d) \neq 0$, with the property that for all y, the univariate polynomial $t \mapsto q(d + ty)$ has only real roots. One then obtains a polynomial which is hyperbolic according to our definition simply by homogenizing q, that is, by introducing a new variable t and multiplying all terms of q by the appropriate power of t so as to obtain a polynomial $p(x)$ which is homogeneous, where $x = (y, t)$. In particular, p is hyperbolic in direction $e := (d, 1)$.

Consequently, from the existence of a point d, $q(d) \neq 0$, with the property that $t \mapsto q(d + ty)$ has only real roots for each y, it follows – by applying preceeding results after homogenization – that the same property is possessed by all points

y in the connected component of $\{y : q(y) \neq 0\}$ containing d. Convexity of the connected component is also immediate.

Henceforth, we refer only the homogeneous setting.

Corollary 5 (Gårding)**.** *The functional* $x \mapsto \lambda_{\min}(x)$ *is concave.*

Proof. Clearly, for $\alpha \in \mathbb{R}$,

$$\{x : \lambda_{\min}(x) \geq \alpha\} = \alpha e + \Lambda_+,$$

which by Theorem 2 is a convex set. Consequently, if $\lambda_{\min}(x) = \alpha = \lambda_{\min}(y)$, then

$$\lambda_{\min}(tx + (1-t)y) \geq \alpha = t\lambda_{\min}(x) + (1-t)\lambda_{\min}(y) \quad \text{for } 0 \leq t \leq 1.$$

The corollary is thus proven for the special case $\lambda_{\min}(x) = \alpha = \lambda_{\min}(y)$.

The general case reduces to the special case by using

$$\lambda_{\min}(\alpha z + \beta e) = \alpha\lambda_{\min}(z) + \beta \quad \text{for } \alpha \geq 0.$$

Details are left to the reader. □

The extent of generality of hyperbolicity cones is unknown. Indeed, it has been conjectured that for each hyperbolicity cone, there is a "slice" of some SDP cone to which the hyperbolicity cone is linearly isomorphic; more specifically, the conjecture is that for each hyperbolicity cone Λ_{++}, there exist n, a subspace $S \subseteq \mathbb{S}^{n\times n}$, and an isomorphism L from $\mathcal{E}$ onto S, under which Λ_{++} is the inverse image of $\mathbb{S}^{n\times n}_{++} \cap S$.

This was a conjecture even for $\mathcal{E} = \mathbb{R}^3$ until recently [7]. The conjecture in this special case came from Peter Lax. The general case, too, has been called "the Lax conjecture." Now that the special case is resolved, we refer to the unrestricted setting as "the general Lax conjecture."

The most significant result bearing on the general Lax conjecture was accomplished by Chua [2] (related work is [4]). He showed that each homogeneous cone is a slice of an SDP cone (a homogeneous cone is a convex cone whose automorphism group – the group of linear isomorphisms that map the cone onto itself – acts transitively on the cone's interior). Homogeneous cones are hyperbolicity cones, as was established in [6].

To end the section in the spirit of the present work, we note that in defining hyperbolic polynomials, there is the requirement $p(e) \neq 0$, which feels to be more for convenience than for substance. To realize otherwise, consider the homogeneous polynomial

$$p(x_1, x_2, x_3) = x_1x_2x_3 - \tfrac{1}{9}(x_1 + x_2 + x_3)(x_1x_2 + x_2x_3 + x_3x_1)$$

and direction $e = (1, 1, 1)$. Each x has only real "eigenvalues" (indeed, $\lambda \mapsto p(\lambda e - x)$ is a non-constant linear map for all $x \neq 0$). However, neither of the two connected components of $\{x : p(x) \neq 0\}$ is convex.

3. Boundary Basics

For $x \in \mathcal{E}$, define the *multiplicity* of x – denoted $\mathrm{mult}(x)$ – to be the multiplicity of 0 as an eigenvalue of x.

Later, the value $\mathrm{mult}(x)$ is proven to be independent of the derivative direction $e \in \Lambda_{++}$ (Proposition 22).

For non-negative integers m, define

$$\partial^m \Lambda_+ := \{x \in \Lambda_+ : \mathrm{mult}(x) = m\}.$$

Thus, $\partial \Lambda_+$ – the boundary of Λ_+ – is partitioned by the sets $\partial^m \Lambda_+$, $m \geq 1$.

Theorem 6. *The set $\partial^1 \Lambda_+$, if non-empty, is a codimension-1 analytic submanifold of $\mathcal{E}$.*

The proof relies on two lemmas:

Lemma 7. *Assume $x \in \partial \Lambda_+$. Then $x \in \partial^1 \Lambda_+$ iff $Dp(x) \not\equiv 0$.*

Proof. We have

$$\tfrac{d}{d\lambda} p(\lambda e - x)|_{\lambda=0} = Dp(-x)e = (-1)^{n-1} Dp(x)e. \tag{1}$$

Consequently, if $Dp(x) \equiv 0$, then $\mathrm{mult}(x) > 1$.

On the other hand, if $Dp(x) \not\equiv 0$, then $\{y : Dp(x)y = 0\}$ is the supporting hyperplane to the cone Λ_+ at x. Thus, $Dp(x)e \neq 0$, because $e \in \Lambda_+^\circ$. Hence, by (1), $\mathrm{mult}(x) = 1$. □

Lemma 8. *For each m, the set $\{x : \mathrm{mult}(x) \geq m\}$ is closed, as is the set*

$$\{x \in \Lambda_+ : \mathrm{mult}(x) \geq m\} \cup \{x \notin \Lambda_+ : \mathrm{mult}(x) \geq m-1\}.$$

Proof. A straightforward consequence of the continuity of eigenvalues. □

Proof of Theorem 6: Let U denote the complement in $\mathcal{E}$ of the set

$$\{x \in \Lambda_+ : \mathrm{mult}(x) \geq 2\} \cup \{x \notin \Lambda_+ : \mathrm{mult}(x) \geq 1\}.$$

By Lemma 8, U is open.

Consider the map $x \mapsto p(x)$ restricted to U, denoted $p|_U$. Clearly, $(p|_U)^{-1}(0) = \partial^1 \Lambda_+$. Since $D\left(p|_U\right)(x) \equiv Dp(x) \not\equiv 0$ for each $x \in \partial^1 \Lambda_+$ (Lemma 7), the Implicit Function Theorem thus shows $\partial^1 \Lambda_+$, if non-empty, is a codimension-1 submanifold of U (hence, of $\mathcal{E}$). □

For $x \in \partial^1 \Lambda_+$, let T_x denote the tangent space at x.

Proposition 9. *If $x \in \partial^1 \Lambda_+$ and $v \in T_x$, then $D^2 p(x)[v, v] \leq 0$.*

Proof. Of course, $x + T_x$ $(= T_x)$ is a supporting hyperplane to the convex cone Λ_+ at x. Since for any point $y \in \mathcal{E}$, $y + te \in \Lambda_+$ for sufficiently large t, it follows that for each $v \in T_x$, there exists $t(v) \geq 0$ satisfying $p(x + v + t(v)e) = 0$. Hence, $\lambda_{\min}(x + v) \leq 0$.

For v near 0, the point $x + v$ can not have more than one non-positive eigenvalue, because x has only positive eigenvalues other than the simple eigenvalue 0. Consequently, for $v \in T_x$ near 0,

$$p(x + v) = p(e)\lambda_{\min}(x + v) \prod_{j>1} \lambda_j(x + v) \leq 0.$$

Since $p(x) = 0$ and $Dp(x)v = 0$, it follows that $D^2p(x)[v, v] \leq 0$. $\square$

Proposition 9 implies that the Hessian $\nabla^2 p(x)$ has at most one positive eigenvalue. In fact, there is a positive eigenvalue, as will be apparent in §5.

The following theorem shows that in those directions for which $\partial^1\Lambda_+$ is not linear, the boundary has definite curvature.

Theorem 10. *If $x \in \partial^1\Lambda_+$ and $v \in T_x$, then one of the following holds:*

- *$p(x + tv) = 0$ for all t, and there exists $\epsilon > 0$ such that $x + tv \in \partial^1\Lambda_+$ whenever $-\epsilon < t < \epsilon$;*
- *$D^2p(x)[v, v] < 0$.*

Proof. Assume $x \in \partial^1\Lambda_+$, and assume $v \in T_x$ does not satisfy $D^2p(x)[v, v] < 0$. By Proposition 9, $D^2p(x)[v, v] = 0$.

To prove the theorem, it suffices to show $t \mapsto p(x + tv) \equiv 0$. Indeed, continuity of eigenvalues and $\lambda_1(x) < \lambda_2(x)$ then imply $0 = \lambda_1(x + tv) < \lambda_2(x + tv)$ for all t near 0.

To show $t \mapsto p(x + tv) \equiv 0$, we assume otherwise and obtain a contradiction.

For $s \in \mathbb{R}$, let $\phi_s(t) := p(x + se + tv)$. Note $\phi_0 \not\equiv 0$ (by assumption).

Since $p(x) = 0 = Dp(x)v = D^2p(x)[v, v]$, the multiplicity of $t = 0$ as a root of ϕ_0 is $m \geq 3$.

Choose $\delta > 0$ such that $t = 0$ is the only root $t \in \mathbb{C}$ satisfying $|t| < \delta$. By continuity of roots, there exists $\epsilon > 0$ such that whenever $0 \leq s \leq \epsilon$, the polynomial ϕ_s has precisely m roots $t \in \mathbb{C}$ – counting multiplicities – satisfying $|t| < \delta$, and has no roots satisfying $|t| = \delta$.

Of course $x + se \in \Lambda_{++}$ whenever $s > 0$. Thus, by Corollary 4, all roots of ϕ_s are real when s is positive. Consequently, for $0 < s \leq \epsilon$, the polynomial ϕ_s has $m \geq 3$ roots – counting multiplicities – in $(-\delta, \delta)$, and $\phi_s(-\delta) \neq 0 \neq \phi_s(\delta)$. In particular, these properties hold for $s = \epsilon$.

Trivially, $t = 0$ is not a root of ϕ_ϵ (because $x + \epsilon e \in \Lambda_{++}$). Hence, ϕ_ϵ has at least two roots in the open interval $(0, \delta)$, or at least two in the open interval $(-\delta, 0)$. Without loss of generality, assume there are at least two roots in $(0, \delta)$.

Consider the line segments

$$\ell_1 := \{x + tv : 0 \le t \le \delta\} \quad \text{and} \quad \ell_2 := \{x + se + \delta v : 0 \le s \le \epsilon\}.$$

These two segments create a path between x and $x + \epsilon e + \delta v$. By choice of ϵ and δ, the only point y on the path which satisfies $p(y) = 0$ is $y = x$.

For each point y on the path, define $w(y) := (x + \epsilon e) - y$. Consider $\psi_y(t) := p(y + tw(y))$, the univariate polynomial obtained by restricting p to the line through y and $x + \epsilon e$. Since $x + \epsilon e \in \Lambda_{++}$, ψ_y has only real roots (Corollary 4).

We know $\psi_y(0) \neq 0$ when $y \neq x$, i.e., we know x is the only zero of p on the path. Also, for all y on the path, $\psi_y(1) = p(x + \epsilon e) \neq 0$, because $x + \epsilon e \in \Lambda_{++}$.

When $y = x + \epsilon e + \delta v$, $\psi_y(t) = \phi_\epsilon(\delta(1 - t))$. Hence, for this choice of y, ψ_y has at least two roots in the interval $0 < t < 1$.

It follows – by continuity of roots in y – that for each $y \neq x$ on the path, the polynomial ψ_y has at least 2 roots in the open interval $(0, 1)$. Hence, in the limit, ψ_x has at least 2 roots in the closed interval $[0, 1]$. But $\psi_x(t) \neq 0$ whenever $0 < t \le 1$, because then $x + tw(x) = x + t\epsilon e \in \Lambda_{++}$. Consequently, $t = 0$ is a root of ψ_x of multiplicity at least 2.

Finally, observe $p(\lambda e - x) = (-1)^n \psi_x(-\frac{1}{\epsilon}\lambda)$, and hence $\lambda = 0$ is an eigenvalue of x with multiplicity at least 2, contradicting $x \in \partial^1 \Lambda_+$. $\square$

Recall that the *lineality space* of a closed convex cone is the maximal subspace contained in the cone.

Proposition 11. *The lineality space of Λ_+ is precisely $\partial^n \Lambda_+$.*

Proof. Assume $x \in \partial^n \Lambda_+$. Thus, $p(\lambda e - x) = p(e)\lambda^n$. Hence, by homogeneity, $p(e - \gamma x) = p(e) \neq 0$ for all γ. Consequently, $\{e - \gamma x : \gamma \in \mathbb{R}\} \subset \Lambda_{++}$, showing x is in the lineality space of Λ_+.

Conversely, assume x is in the lineality space. Thus, $\{e - \gamma x : \gamma \in \mathbb{R}\} \subset \Lambda_{++}$, implying $p(e - \gamma x) > 0$ for all γ. However, the polynomial $\gamma \mapsto p(e - \gamma x)$ has only real roots (Corollary 4). Consequently, it must be a constant, i.e., $p(e - \gamma x) = p(e)$ for all γ. By homogeneity, $p(\lambda e - x) = p(e)\lambda^n$, that is, $x \in \partial^n \Lambda_+$. $\square$

4. The Derivative Cone

Continuity implies, of course, that between any two roots of $\phi(t) := p(x + te)$, there lies a root of $\phi'(t) = \frac{d}{dt} p(x + te) = Dp(x + te)e$. Since ϕ' has precisely one less root (counting multiplicities) than does ϕ, it follows that ϕ having only real roots implies ϕ' has only real roots. Consequently, the multivariate polynomial

$$p'(x) := \tfrac{d}{dt}p(x+te)|_{t=0} = Dp(x)e$$

is hyperbolic in direction e. This is the *derivative polynomial* (of p in direction e). When necessary for clarity, we write p'_e.

Denote the hyperbolicity cone of p' by Λ'_{++} (or by $\Lambda'_{++,e}$), the *derivative cone.*

Generally, the cones $\Lambda'_{++,e}$ vary with e. For example, when $\mathcal{E} = \mathbb{R}^2$ and $p(x_1, x_2) = x_1x_2$, the derivative cone is the open halfspace with boundary orthogonal to e (for any vector e whose coordinate are non-zero).

Observe

$$p'(\lambda e - x) = \tfrac{d}{d\lambda}p(\lambda e - x).$$

The eigenvalues $\lambda'_i(x)$ with respect to p' thus interlace the eigenvalues with respect to p:

$$\lambda_1(x) \leq \lambda'_1(x) \leq \lambda_2(x) \leq \ldots \leq \lambda'_{n-1}(x) \leq \lambda_n(x),$$

where

$$\big(\lambda_j(x) = \lambda'_j(x) \;\text{ or }\; \lambda'_j(x) = \lambda_{j+1}(x)\big) \quad \Leftrightarrow \quad \lambda_j(x) = \lambda'_j(x) = \lambda_{j+1}(x). \tag{2}$$

As a simple consequence of the interlacing, we have

$$\Lambda_{++} \subseteq \Lambda'_{++},$$

i.e., the derivative cone Λ'_{++} is a relaxation of Λ_{++}.

Theorem 12. *For integers $m \geq 2$,*

$$\partial^m \Lambda'_+ = \partial^{m+1}\Lambda_+.$$

Also,

$$(\partial^1 \Lambda'_+) \cap \Lambda_+ = \partial^2 \Lambda_+.$$

Proof. Straightforward consequences of the interlacing of eigenvalues and the equivalence (2). □

A closed, convex cone is *regular* if both it has non-empty interior and its lineality space is $\{0\}$. Trivially, however, Λ_+ has non-empty interior ($e \in \Lambda_{++}$).

Proposition 13. *If $n \geq 3$, then the lineality spaces of Λ_+ and Λ'_+ are identical (thus, if $n \geq 3$ and Λ_+ is regular, Λ'_+ is regular, too).*

Proof. Immediate from Theorem 12 and Proposition 11. □

To understand the structure of hyperbolicity cones, attention need only be given to when Λ_+ is regular. Indeed, by choosing a subspace S which both complements the lineality space and contains e, the restriction of p to S is a hyperbolic polynomial whose hyperbolicity cone $S \cap \Lambda_+$ is regular. The faces of Λ_+, for example, are precisely Minkowski sums of the lineality space with the faces of $S \cap \Lambda_+$.

Theorem 14. *Assume Λ_+ is a regular and $n \geq 3$.*
Assume $x \in (\partial^1 \Lambda'_+) \setminus \Lambda_+$ and $v \in T'_x$ (tangent space to $\partial^1 \Lambda'$).
If v is not a scalar multiple of x, then $D^2(x)[v,v] < 0$.

(That is, the boundary of Λ'_+ has strict curvature at x except in the single direction $v = x$ where, as a cone, the boundary must be linear.)

Proof. Assume x_1, x_2 are linearly independent and have the property that the line segment connecting them – denoted $[x_1, x_2]$ – lies entirely in $\partial^1 \Lambda'_+$. To prove the theorem, it suffices – by Theorem 10 applied to p' – to show $x_1, x_2 \in \Lambda_+$.

Since Λ'_+ is regular (Proposition 13), there exist positive values t_1, t_2 for which the line through $t_1 x_1$ and $t_2 x_2$ intersects Λ'_+ in a segment of finite length. Let y_1, y_2 be the endpoints of the segment. To show $x_1, x_2 \in \Lambda_+$, it suffices to show $y_1, y_2 \in \Lambda_+$ (because Λ_+ is a convex cone).

For $t \in \mathbb{R}$, consider the univariate polynomial

$$\phi_t(\lambda) := p'(\lambda e - t y_1 - (1-t) y_2).$$

We claim $\lambda = 0$ is a root of ϕ_t for all t. Indeed, $[y_1, y_2] \subset \partial \Lambda'_+ \subseteq \{y : p'(y) = 0\}$ and hence $\phi_t(0) = 0$ for all $0 \leq t \leq 1$ (in particular, for infinitely many values of t). But $t \mapsto \phi_t(0)$ is a polynomial, and thus has finitely many roots or is identically 0. Thus, $t \mapsto \phi_t(0) \equiv 0$, establishing the claim.

If $t \notin [0,1]$, then $t y_1 + (1-t) y_2 \notin \Lambda'_+$, and hence ϕ_t has a negative root. Thus, since $\phi_t(0) = 0$, ϕ_t has at least 2 non-positive roots if $t \notin [0,1]$.

By continuity of roots, ϕ_0 and ϕ_1 thus each have at least 2 non-positive roots, counting multiplicities. However, since $y_1, y_2 \in \Lambda'_+$, all roots of ϕ_0 and ϕ_1 are non-negative. Hence, 0 is a root of multiplicity at least 2 for each of ϕ_0 and ϕ_1. That is, 0 is an eigenvalue – with respect to p' – of multiplicity at least 2 for both y_1 and y_2. Theorem 12 now shows $y_1, y_2 \in \Lambda_+$, completing the proof. $\square$

Corollary 15. *Assume Λ_+ is regular and $n \geq 3$. If $x \in (\partial \Lambda'_+) \setminus \Lambda_+$, then $x \in \partial^1 \Lambda'_+$ and x is an (exposed) extreme direction of Λ'_+.*

Proof. Immediate from Theorems 12 and 14. $\square$

5. Higher Derivatives

Taking the derivative repeatedly gives a sequence of hyperbolic polynomials $p^{(0)} = p, p^{(1)} = p', p^{(2)}, \ldots, p^{(n-1)}$ with nested hyperbolicity cones

$$\Lambda_{++} = \Lambda_{++}^{(0)} \subseteq \Lambda_{++}^{(1)} \subseteq \ldots \subseteq \Lambda_{++}^{(n-1)}.$$

The cone $\Lambda_{++}^{(n-1)}$ is an open halfspace.

Trivially, results from §4 can be generalized by induction. For example, if $m \geq 2$, then by Theorem 12,

$$\partial^m \Lambda_+^{(i)} = \partial^{m+1} \Lambda_+^{(i-1)} = \ldots = \partial^{m+i} \Lambda_+^{(0)}.$$

For this result, and for the following proposition and its corollary, the derivatives need not all be in the same direction. That is, one can choose a sequence of directions $e_1, \ldots, e_n$, let $p^{(1)} := p'_{e_1}$, and inductively define $p^{(i+1)} := (p^{(i)})'_{e_{i+1}}$, the only requirement being that e_{i+1} lie in the hyperbolicity cone for $p^{(i)}$.

Proposition 16. *If $x \in (\partial \Lambda_+^{(i)}) \cap \Lambda_+^{(i-1)}$ for some $1 \leq i \leq n-1$, then $x \in \Lambda_+$.*

Proof. Under the hypothesis, Theorem 12 applied to $\Lambda_+^{(i-1)}$ shows $x \in \partial^m \Lambda_+^{(i-1)}$ for some $m \geq 2$. Thus, $x \in \partial^{m+i-1} \Lambda_+$, also by Theorem 12. $\square$

Corollary 17. *Assume Λ_+ is regular and $1 \leq i \leq n-2$. If $x \in (\partial \Lambda_+^{(i)}) \setminus \Lambda_+$, then $x \in \partial^1 \Lambda_+^{(i)}$ and x is an (exposed) extreme direction of $\Lambda_+^{(i)}$.*

Proof. Under the hypothesis, $\Lambda_+^{(i)}$ is regular (Proposition 13), and $x \in (\partial \Lambda_+^{(i)}) \setminus \Lambda_+^{(i-1)}$ (Proposition 16). Consequently, Corollary 15 can be applied to $\Lambda_+^{(i)}$, concluding the proof. $\square$

Hereafter, we assume there is a single derivative direction e, not a sequence of directions.

Let σ_k denote the elementary symmetric polynomial of degree k,

$$\sigma_k(\lambda_1, \ldots, \lambda_n) = \sum_{j_1 < \cdots < j_k} \lambda_{j_1} \cdots \lambda_{j_k}.$$

For convenience, define $\sigma_0 \equiv 1$.

Proposition 18. *For each $0 \leq i \leq n$,*

$$p^{(i)}(x) = i!\, p(e) \sigma_{n-i}(\lambda(x))$$

(consequently, $x \mapsto \sigma_k(\lambda(x))$ is a hyperbolic polynomial of degree k).

Proof. The proposition is immediate from $p^{(i)}(x) = \frac{d^i}{dt^i} p(x+te)|_{t=0}$ and

$$\begin{aligned} p(x+te) &= p(te - (-x)) \\ &= p(e) \prod_j (t - \lambda_j(-x)) \\ &= p(e) \prod_j (t + \lambda_j(x)) \\ &= p(e) \sum_{k=0}^{n} \sigma_{n-k}(\lambda(x)) t^k. \end{aligned}$$

$\square$

Let $\lambda_1^{(i)}(x) \leq \ldots \leq \lambda_{n-i}^{(i)}(x)$ denote the eigenvalues of x in direction e with respect to $p^{(i)}$.

Corollary 19. *For $0 \leq i \leq n$ and $0 \leq k \leq n-i$,*

$$\binom{n}{i}\sigma_k(\lambda^{(i)}(x)) = \binom{n-k}{i}\sigma_k(\lambda(x))$$

(where on the left, σ_k is in $n-i$ variables).

Proof. Follows easily from Proposition 18 □

The following theorem provides a useful characterization of Λ_+. Essentially, the theorem is only a restatement of the standard fact that a univariate real polynomial with only real roots has only negative roots iff all of its coefficients are positive.

Theorem 20. $\Lambda_+ = \{x : \sigma_k(\lambda(x)) \geq 0,\ k = 1, \ldots, n\}$

Proof. Trivially, if $x \in \Lambda_+$, then $\sigma_k(\lambda(x)) \geq 0$ for all k, giving the inclusion "$\subseteq$". For the reverse inclusion, note that if $\sigma_k(\lambda(x)) \geq 0$ for all k, and if $\lambda < 0$, then

$$\begin{aligned} p(\lambda e - x) &= p(e)\prod_j(\lambda - \lambda_j(x)) \\ &= p(e)\sum_k(-1)^k\sigma_k(\lambda(x))\lambda^{n-k} \\ &= (-1)^n p(e)\sum_k \sigma_k(\lambda(x))|\lambda|^{n-k} \\ &\neq 0. \end{aligned}$$

Since λ is an arbitrary negative number, all eigenvalues of x must thus be non-negative, i.e., $x \in \Lambda_+$. □

Of the inequalities $\sigma_k(\lambda(x)) \geq 0$ appearing in the characterization provided by Theorem 20, $\sigma_n(\lambda(x)) \geq 0$ is most crucial. Indeed, Λ_{++} is the connected component of $\{x : \sigma_n(\lambda(x)) > 0\}$ containing e (Proposition 1). The role of the remaining inequalities (i.e., $\sigma_k(x) \geq 0,\ k < n$) is only to isolate the particular connected component.

The theorem yields the characterizations

$$\begin{aligned} \Lambda_{++} &= \{x : \sigma_n(\lambda(x)) > 0 \text{ and } \sigma_k(\lambda(x)) \geq 0,\ 1 \leq k < n\} \\ &= \{x : \sigma_k(\lambda(x)) > 0,\ k = 1, \ldots, n\}, \end{aligned}$$

where the second identity is due to the first identity and the positivity of eigenvalues for $x \in \Lambda_{++}$. Furthermore, applying the theorem to $\Lambda_+^{(i)}$ and relying on Corollary 19,

$$\Lambda_+^{(i)} = \{x : \sigma_k(\lambda(x)) \geq 0,\ k = 1, \ldots, n-i\},$$

$$\partial\Lambda_+^{(i)} = \{x : \sigma_{n-i}(\lambda(x)) = 0 \text{ and } \sigma_k(\lambda(x)) \geq 0,\ 1 \leq k < n-i\}. \tag{3}$$

Corollary 21. *For $i = 1, \ldots, n$,*

$$\partial^m \Lambda_+ = \{x \in \Lambda_+ : \sigma_{n-m}(\lambda(x)) > 0 \text{ and } \sigma_{n-m+1}(\lambda(x)) = 0\}.$$

Proof. Clearly,

$$\partial^m \Lambda_+ = \Lambda_+ \cap \left((\partial \Lambda_+^{(m-1)}) \setminus (\partial \Lambda_+^{(m)}) \right).$$

However, from (3),

$$(\partial \Lambda_+^{(m-1)}) \setminus (\partial \Lambda_+^{(m)}) = \{x \in \Lambda_+^{(n-m+1)} : \sigma_{n-m+1}(\lambda(x)) = 0 \text{ and } \sigma_{n-m}(\lambda(x)) > 0\},$$

yielding the corollary (because $\Lambda_+ \subseteq \Lambda_+^{(n-m+1)}$). □

In relation to the corollary, it should be noted that for $x \in \Lambda_+$, if $\sigma_{n-m+1}(\lambda(x)) = 0$, then $\sigma_k(\lambda(x)) = 0$ for all $k \geq n - m + 1$ (simply because the first equality is equivalent, when all eigenvalues are non-negative, to 0 being an eigenvalue for x of multiplicity at least m). Similarly, if $x \in \Lambda_+$ and $\sigma_{n-m}(\lambda(x)) > 0$, then $\sigma_k(\lambda(x)) > 0$ for all $k \leq n - m$.

The remainder of the section is devoted to establishing two claims made in §2.

Just after Proposition 9, we claimed that $\nabla^2 p(x)$ has a positive eigenvalue when $x \in \partial^1 \Lambda_+$ (the proposition implied at least $n-1$ of the eigenvalues to be non-positive). Establishing that there is a positive eigenvalue can now be accomplished with brevity:

$$\langle e, \nabla^2 p(x) e \rangle = p^{(2)}(x) = 2p(e)\sigma_2(\lambda(x)) > 0,$$

the inequality coming from Theorem 20 and $p^{(2)}(x) \neq 0$ (because $x \in \partial^1 \Lambda_+$).

The other claim from §2 which remains unjustified is handled by the following proposition.

Proposition 22. *For $x \in \Lambda_+$, the value* $\mathrm{mult}(x)$ *is independent of the direction $e \in \Lambda_{++}$.*

Proof. Assume $e, \tilde{e} \in \Lambda_{++}$. For arbitrary $x \in \Lambda_+$, let m (resp., $\tilde{m}$) denote the multiplicity of 0 as an eigenvalue with respect to the direction e (resp., $\tilde{e}$). We assume $m \leq \tilde{m}$ and proceed by induction.

Trivially, if $m = 0$ – that is, if $p(x) \neq 0$ – then $\tilde{m} = 0$.

Assume $m \geq 1$. Clearly, x is a simple root for $p_e^{(m-1)}$ in direction e. Since $e, \tilde{e} \in \Lambda_{++} \subseteq \Lambda_{++,e}^{(m-1)}$, Lemma 7 thus implies x is a simple root for $p_e^{(m-1)}$ in direction $\tilde{e}$ as well as in direction e. However,

$$(p_e^{(m-1)})_{\tilde{e}}^{(1)} = (p_{\tilde{e}}^{(1)})_e^{(m-1)},$$

and hence, $(p_{\tilde{e}}^{(1)})_e^{(m-1)}(x) \neq 0$, that is, for the hyperbolic polynomial $p_{\tilde{e}}^{(1)}$, the point x is a root of multiplicity at most $m-1$ for direction e. As $e, \tilde{e} \in \Lambda_{++} \subseteq \Lambda_{++,\tilde{e}}^{(1)}$, inductive hypothesis thus applies to x and $p_{\tilde{e}}^{(1)}$, yielding

$$(p_{\tilde{e}}^{(1)})_{\tilde{e}}^{l-1}(x) \neq 0 \quad \text{for some } l \leq m.$$

Since $(p_{\tilde{e}}^{(1)})_{\tilde{e}}^{l-1} = p_{\tilde{e}}^{(l)}$ and $m \leq \tilde{m}$, it follows that $m = \tilde{m}$. □

6. Faces Exposed

Truong and Tunçel [12] showed that all boundary faces[1] of homogeneous cones are exposed, i.e., the faces coincide precisely with the sets $H \cap K$, where H ranges over all supporting hyperplanes for the cone K. Later, the exposure became a corollary to the exposure of all faces of $\mathbb{S}_+^{n \times n}$, when Chua established that each homogeneous cone is a slice of an SDP cone. The following theorem would likewise become a corollary if the general Lax conjecture was established as true.

Theorem 23. *All boundary faces of Λ_+ are exposed.*

Towards proving the theorem, we introduce two propositions.

Proposition 24. *Assume Λ_+ is regular. For $i = 0, \ldots, n-2$, each boundary face of $\Lambda_+^{(i)}$ either is a face of Λ_+, or is an exposed extreme direction not contained in Λ_+.*

Proof. Let $F^{(i)}$ be a boundary face of $\Lambda_+^{(i)}$. Since Corollary 17 asserts that each $x \in (\partial \Lambda_+^{(i)}) \setminus \Lambda$ is an exposed extreme direction of $\Lambda_+^{(i)}$, we may assume $F^{(i)} \subset \Lambda_+$. However, then $F^{(i)}$ is trivially a face of Λ_+, because $\Lambda_+ \subseteq \Lambda_+^{(i)}$. □

For a face F, let $\mathrm{relint}(F)$ denote its relative interior.

Proposition 25. *Assume F is a boundary face of Λ_+ other than the lineality space. If $x \in \mathrm{relint}(F)$ and $m := \mathrm{mult}(x)$, then F is a face of $\Lambda_+^{(m-1)}$.*

Proof. We may assume Λ_+ is regular (via the observations just prior to Theorem 14).

Assume $x \in \mathrm{relint}(F)$. Thus, F is the unique face of Λ_+ containing x in its relative interior.

Since F is not the lineality space, x is not in the lineality space. Hence, $m := \mathrm{mult}(x) \leq n-1$ (Proposition 11).

Let G be the face of $\Lambda_+^{(m-1)}$ containing x in its relative interior. Since $m-1 \leq n-2$, Proposition 24 shows G is a face of Λ_+. By uniqueness of F, we conclude $G = F$. □

---

[1]The faces of a convex set S are its subsets F with the property that for each open line segment in S that intersects F, the closure of the segment lies in F. (For results on the facial structure of general convex sets, we recommend [11], Chaps. 1, 2.)

Proof of Theorem 23: The lineality space is, of course, an exposed face. Thus, we may assume F is a boundary face other than the lineality space.

Choose $x \in \mathrm{relint}(F)$. By Proposition 25, F is a face of $\Lambda_+^{(m-1)}$, where $m = \mathrm{mult}(x)$. Consequently, since x has multiplicity 1 in $\Lambda_+^{(m-1)}$, Theorem 10 shows the tangent space to $\Lambda_+^{(m-1)}$ at x exposes F as a face of $\Lambda_+^{(m-1)}$. Since $\Lambda_+ \subseteq \Lambda_+^{(m-1)}$, F also is thusly exposed as a face of Λ_+. $\square$

For faces F of Λ_+, define

$$\mathrm{mult}(F) := \min_{x \in F} \mathrm{mult}(x).$$

The next result gives some insight into the algebraic structure of faces.

Theorem 26. *Assume F is a face of Λ_+ and $x \in F$. Then*

$$\mathrm{mult}(x) = \mathrm{mult}(F) \quad \text{iff} \quad x \in \mathrm{relint}(F).$$

Proof. Trivially, the statement is true for $F = \Lambda_+$.

If F is the lineality space, then $F = \mathrm{relint}(F)$; hence, the statement of the theorem is immediate from Proposition 11.

Finally, assume F is a boundary face other than the lineality space.

Assume $x \in \mathrm{relint}(F)$. By Proposition 25, $F \subset \partial\Lambda_+^{(m-1)}$, where $m := \mathrm{mult}(x)$. Since, trivially, $x \notin \partial\Lambda_+^{(m)}$, we have $\mathrm{mult}(x) = \mathrm{mult}(F)$.

Now assume x is in F, but not in the relative interior. Let L denote a line through x which intersects the relative interior.

For all $y \in L \cap \mathrm{relint}(F)$, we know, from above, that $\mathrm{mult}(y) = \mathrm{mult}(F)$. Consequently, for all these points y, we have $p^{(i)}(y) = 0$ if $i < \mathrm{mult}(F)$. Since there are infinitely many such points, the polynomials $p^{(i)}$, $i < \mathrm{mult}(F)$, must thus be identically 0 on L.

On the other hand, $p^{(i)}(x) > 0$ for $i \geq \mathrm{mult}(x)$ (i.e., $x \in \Lambda_{++}^{(i)}$ for $i \geq \mathrm{mult}(x)$).

If it was the case that $\mathrm{mult}(x) = \mathrm{mult}(F)$, these observations would imply, for all $i = 1, \ldots, n$, that $p^{(i)}(z) \geq 0$ for all z in an open interval of L containing x. The open interval would be contained in Λ_+ (Proposition 18 and Theorem 20), and thus, contained in any face of Λ_+ containing x – in particular, contained in F. However, x is an endpoint of $L \cap F$. Hence, $\mathrm{mult}(x) > \mathrm{mult}(F)$. $\square$

7. Hyperbolic Programs

A *hyperbolic program* – or "hyperbolic instance" – is an optimization problem instance of the form

$$\begin{array}{ll} \min & c^* x \\ \text{s.t.} & Ax = b \\ & x \in \Lambda_+. \end{array}$$

Define

$$\mathrm{trace}(x) := \sum_j \lambda_j(x),$$

a functional which, by Proposition 18, is linear. As this functional – like all of the hyperbolic polynomials $x \mapsto \sigma_k(\lambda(x))$ for $1 \le k < n$ – depends on the derivative direction e, we sometimes write trace_e for clarity.

Let Λ_+^* denote the dual cone (the set of linear functionals which are non-negative everywhere on Λ_+). The following theorem shows that under the standard optimization assumption of strict dual feasibility[2], c^*x can be replaced by trace_e for some $e \in \Lambda_+^\circ$. Under slightly different guise (made apparent by the proof), the theorem is central to the general duality theory of interior-point methods ([8]; also see [10]).

Theorem 27. *If Λ_+ is regular, then*

$$(\Lambda_+^*)^\circ = \{\mathrm{trace}_e : e \in \Lambda_{++}\}.$$

Proof. The mapping $\Lambda_{++} \to \mathbb{R}$ given by $e \mapsto -\ln p(e)$ is a self-concordant barrier functional (c.f., [10], p.70), and consequently, $e \mapsto D\ln p(e)$ is a bijection between Λ_{++} and $(\Lambda_+^*)^\circ$ (c.f., [10], Prop. 3.3.2). However, for all x,

$$D\ln p(e)x = \tfrac{1}{p(e)}\tfrac{d}{dt}p(e+tx)|_{t=0}$$

and

$$p(e+tx) = p(e)\prod_j \lambda_j(e+tx) = p(e)\prod_j (1+t\lambda_j(x)),$$

implying $D\ln p(e)x = \mathrm{trace}_e(x)$. □

8. Relaxations

When $c^* = \mathrm{trace} = \mathrm{trace}_e$, the derivative cones $\Lambda_+^{(i)} = \Lambda_{+,e}^{(i)}$ provide natural relaxations to a hyperbolic program, the relaxations themselves being hyperbolic programs:

$$\begin{array}{ll} \min & \mathrm{trace}(x) \\ \text{s.t.} & Ax = b \\ & x \in \Lambda_+^{(i)}. \end{array}$$

[2]For readers unfamiliar with the notion of strict dual feasibility: A hyperbolic instance equivalent to HP is obtained by replacing c^*x with s^*x, where s^* is any linear functional for which there exists y^* satisfying $y^*A + s^* = c^*$. Indeed, $c^*(x_1 - x_2) = s^*(x_1 - x_2)$ for all x_1, x_2 satisfying $Ax = b$, so the ordering on feasible points induced by the objective function is identical for the two instances. The instance HP is said to be *strictly dual feasible* if s^* can be chosen from $(\Lambda_+^*)^\circ$, the interior of the dual cone.

The sequence of relaxations obviously is ripe for induction, especially in light of the relations

$$\text{trace}(x) := \sum_{j=1}^{n} \lambda_j(x) = \tfrac{n}{n-i} \sum_{j=1}^{n-i} \lambda_j^{(i)}(x)$$

(by Proposition 15). We give one example of results that can be derived from the preceeding development.

Let $\Omega^{(i)}$ denote the optimal solution set of the i^{th} relaxation. Thus, the optimal solution set for the original hyperbolic program is $\Omega^{(0)}$.

Theorem 28. *Assume Λ_+ is regular, and assume the original hyperbolic program is feasible. Then $\Omega^{(i)} \neq \emptyset$, $i = 0, \ldots, n-1$. Furthermore, for $i = 0, \ldots, n-2$,*

$$\begin{array}{rl} \textit{either} & \Omega^{(i)} \textit{ consists of a unique point, and } \Omega^{(i)} \cap \Omega^{(j)} = \emptyset \textit{ for } j \neq i, \\ \textit{or} & \Omega^{(i)} = \Omega^{(i-1)} = \ldots = \Omega^{(0)}. \end{array}$$

Proof. Trivially, all of the relaxations are feasible.

Proposition 13 implies $\Lambda_+^{(i)}$ is regular for $i = 0, \ldots, n-2$. Consequently, since $\text{trace} = \sum_j \lambda_j$ is a positive multiple of $\sum_j \lambda_j^{(i)}$, Theorem 27 applied to $\Lambda_+^{(i)}$ shows the i^{th} relaxation to be strictly dual feasible for $i = 0, \ldots, n-2$.

Being both feasible and strictly dual feasible, the i^{th} relaxation, $i = 0, \ldots, n-2$, has at least one optimal solution, a consequence of standard duality theory for convex optimization (c.f., [10], §3.2).

By Proposition 24, $\Omega^{(i)}$ $(i = 0, \ldots, n-2)$ thus either consists of a unique point not in Λ_+, or is a face of Λ_+. Since $\Lambda_+ \subseteq \Lambda_+^{(i)}$, in the latter case it is easily argued that $\Omega^{(i)} = \Omega^{(0)}$.

Finally, $\Lambda_+^{(n-1)} = \{x : \text{trace}(x) \geq 0\}$ (Corollary 19 and Theorem 20), from which it trivially follows that $\Omega^{(n-1)} \neq \emptyset$. □

9. Of Computation

It might appear that most higher derivatives would be exorbitantly expensive to compute, in which case the relaxed hyperbolic programs would be practically useless. In LP, for example, where $p(x) = x_1 \cdots x_n$, we have that $p^{(i)}(x)$ is a constant multiple of $\sigma_{n-i}(\frac{x_1}{e_1}, \ldots, \frac{x_n}{e_n})$, a polynomial with $\binom{n}{i}$ terms (exponentially many terms when i is in the mid-range between 0 and n).

Effective computation of the derivatives can be made, however, via interpolation, assuming the desired outcome to be, say, values of derivatives – or values of their gradients and Hessians – at a specified point x.

To illustrate, first recall that a univariate polynomial $\phi(t) = \sum_{i=0}^{n} a_i t^i$ is determined by the values $\phi(t_j)$ for any set of $n+1$ distinct complex numbers $t_0, \ldots, t_n$.

Indeed, assuming $t_0 = 0$ (so $a_0 = \phi(t_0)$), the coefficients a_i, $i = 1, \ldots, n$, are the unique solution for the following linear equations involving a Vandermonde matrix:

$$\begin{bmatrix} t_1^n & t_1^{n-1} & \cdots & t_1 \\ \vdots & \vdots & \ddots & \vdots \\ t_n^n & t_n^{n-1} & \cdots & t_n \end{bmatrix} \begin{bmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_1 \end{bmatrix} = \begin{bmatrix} \phi(t_1) - \phi(0) \\ \vdots \\ \phi(t_n) - \phi(0) \end{bmatrix}.$$

This already makes apparent that computing all of the coefficients a_i – equivalently, computing all of the derivative values $\phi^{(i)}(0)$ – requires no more work than solving an $n \times n$ system of linear equations (plus the effort required for the evaluations $\phi(0), \phi(t_1), \ldots, \phi(t_n)$).

By choosing $t_j = \omega^j$, where ω is a primitive n^{th} root of unity, the inverse of the Vandermonde matrix becomes of particularly simple form:

$$\frac{1}{n} \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ \omega & \omega^2 & \cdots & \omega^{2(n-1)} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \omega^{n-1} & \omega^{(n-1)2} & \cdots & \omega^{(n-1)(n-1)} & 1 \end{bmatrix}.$$

(there are only n distinct entries in the matrix, the n^{th} roots of unity). Now, determining the coefficients a_k is only a matter of matrix-vector multiplication.

Cooley and Tukey [3] showed even further reduction in computational effort can be made due to the pattern of entries in the inverse when n is a power of 2 (one can always pad a polynomial to make its degree be a power of 2). The resulting algorithm – the (inverse) Discrete Fourier Transform – computes all of the coefficients a_k with $O(n \log^2 n)$ arithmetic operations (given the values $\phi(0)$, $\phi(\omega^j)$).

To relate this to hyperbolic programming, let $\phi(t) := p(x + te)$, where x is specified. Then, $a_i = \frac{1}{i!} p^{(i)}(x)$. Consequently, the values of all derivatives at x are efficiently computable if the values $p(x)$ and $p(x + \omega^j e)$ can be quickly determined. (In LP, for example, the values certainly can be quickly determined; $p(x + \omega^j e) = (x_1 + \omega^j) \cdots (x_n + \omega^j)$.)

This strategy extends to computing not only the values of the derivative polynomials $p^{(i)}$, but also their gradients and Hessians. Indeed, noting for $i = 1, \ldots, n-1$ that $\sum_{j=1}^n \omega^{ij} = 0$ (because $\sum_{j=1}^n t^j = t(t^n - 1)/(t-1)$), the matrix-vector multiplication gives

$$p^{(i)}(x) = \tfrac{i!}{n} \sum_{j=1}^n \omega^{ij} p(x + \omega^j e) \qquad i = 1, \ldots, n-1,$$

and hence,

$$\nabla p^{(i)}(x) = \tfrac{i!}{n} \sum_{j=1}^{n} \omega^{ij} \nabla p(x + \omega^j e),$$

$$\nabla^2 p^{(i)}(x) = \tfrac{i!}{n} \sum_{j=1}^{n} \omega^{ij} \nabla^2 p(x + \omega^j e).$$

References

[1] H. H. Bauschke et al., Hyperbolic polynomials and convex analysis, Canad. J. Math. **53** (2001), no. 3, 470–488. MR1827817 (2002c:90099)

[2] C.B. Chua, Relating Homogeneous Cones and Positive Definite Cones via T-algebras, SIAM J. Optim. **14** (2003) no. 2, 500–506.

[3] J. W. Cooley and J. W. Tukey, An algorithm for the machine calculation of complex Fourier series, Math. Comp. **19** (1965), 297–301. MR0178586 (31 #2843)

[4] L. Faybusovich, On Nesterov's approach to semi-infinite programming, Acta Appl. Math. **74** (2002), no. 2, 195–215. MR1935854 (2003i:90094)

[5] L. Gårding, "An inequality for hyperbolic polynomials," *J. Math. Mech.* 8 (1959) 957–965.

[6] O. Güler, Hyperbolic polynomials and interior point methods for convex programming, Math. Oper. Res. **22** (1997), no. 2, 350–377. MR1450796 (98d:90084)

[7] A.S. Lewis, P.A. Parrilo, M.V. Ramana, The Lax conjecture is true, technical report, Dept. of Math., Simon Fraser Univ., Canada. Submitted to Proc. Amer. Math. Soc.

[8] Y. Nesterov and A. Nemirovskii, *Interior-point polynomial algorithms in convex programming*, SIAM, Philadelphia, PA, 1994; MR 94m:90005.

[9] Q. I. Rahman and G. Schmeisser, *Analytic theory of polynomials*, Oxford Univ. Press, Oxford, 2002. MR1954841 (2004b:30015)

[10] J. Renegar, *A mathematical view of interior-point methods in convex optimization*, SIAM, Philadelphia, PA, 2001; MR 2002g:90002

[11] R. Schneider, *Convex bodies: the Brunn-Minkowski theory*, Cambridge Univ. Press, Cambridge, 1993; MR 94d:52007

[12] V.A. Truong, L. Tunçel, Geometry of homogeneous cones, duality mapping and optimal self-concordant barriers, Math. Prog., Ser. A (2003).

[13] C. K. Yap, *Fundamental problems of algorithmic algebra*, Oxford Univ. Press, New York, 2000. MR1740761 (2000m:12014)

School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY, 14853, U.S.

E-mail address: `renegar@mac.com`