CENTRAL SWATHS
(A GENERALIZATION OF THE CENTRAL PATH)

JAMES RENEGAR

Abstract. We develop a natural generalization to the notion of the central path – a notion that lies at the heart of interior-point methods for convex optimization. The generalization is accomplished via the "derivative cones" of a "hyperbolicity cone," the derivatives being direct and mathematically-appealing relaxations of the underlying (hyperbolic) conic constraint, be it the non-negative orthant, the cone of positive semidefinite matrices, or other.

We prove that a dynamics inherent to the derivative cones generates paths always leading to optimality, the central path arising from a special case in which the derivative cones are quadratic. Derivative cones of higher degree better fit the underlying conic constraint, raising the prospect that the paths they generate lead to optimality quicker than the central path.

1. Introduction

Let $\mathcal{E}$ denote a finite-dimensional Euclidean space and let $p : \mathcal{E} \to \mathbb{R}$ be a hyperbolic polynomial, that is, a homogeneous polynomial for which there is a designated direction vector e satisfying $p(e) > 0$ and having the property that for all $x \in \mathcal{E}$, the univariate polynomial $t \mapsto p(x + te)$ has only real roots. Thus, p is "hyperbolic in direction e."

Let Λ_{++} denote the hyperbolicity cone – the connected component of $\{x : p(x) > 0\}$ containing e. Let Λ_+ be the closure.

A simple example is $p(x) = \prod_j x_j$ and $e = (1, \ldots, 1)$, in which case Λ_{++} is the strictly positive orthant and Λ_+ is the non-negative orthant. Perhaps the most fundamental example, however, is $p(X) = \det(X)$, where X ranges over $n \times n$ symmetric matrices and $e = I$, the identity matrix. Here, Λ_{++} is the cone of (strictly-)positive definite (pd) matrices and Λ_+ is the positive semidefinite (psd) cone.

Gårding [6] showed for each hyperbolic polynomial p that every $\hat{e} \in \Lambda_{++}$ is a hyperbolicity direction, i.e., for every x, all of the roots of $t \mapsto p(x + t\hat{e})$ are real. One of several remarkable corollaries Gårding established is that Λ_{++} is convex. (Of course Λ_+ is thus convex, too.) (See §2 of [13] for simplified proofs.)

The combination of convexity and rich algebraic structure make hyperbolicity cones promising objects for study in the context of optimization, as was first made evident by Güler [7], who developed a rich theory of interior-point methods for hyperbolic programs, that is, for problems of the form

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

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

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

Research supported in part by NSF Grant #CCF-0430672.

Thanks to Chek Beng Chua and Yuriy Zinchenko for many helpful conversations.

– linear programming, second-order programming and semidefinite programming being particular cases. Key to Güler's development is that the function $x \mapsto -\ln p(x)$ is a self-concordant barrier for Λ_+; thus the general theory of Nesterov and Nemirovskii [12] applies.

A primary purpose of the present paper is to use the viewpoints provided by hyperbolic programming to develop a natural generalization to the notion of the central path[1] (a notion that lies at the heart of interior-point method theory). This is accomplished via derivative cones, which are direct relaxations of the underlying convex conic constraint (be it the non-negative orthant, the cone of positive semidefinite matrices, ...).

However, perhaps more important than the "natural generalization to the notion of the central path" is our "use (of) the viewpoints provided by hyperbolic programming" in developing the generalization. Indeed, it is our conviction that even if results about hyperbolic programming never find application more general than linear programming and semidefinite programming, the setting of hyperbolic programming is favorable for engendering intriguing algorithmic ideas that otherwise would have been unrealized (or at least considerably delayed).

Familiarity with the central path is not required to readily understand our results. (The central path simply provides an initial anchor with which many readers *are* familiar.)

The literature focusing on hyperbolic polynomials is relatively small but its growth is accelerating and its quality in general is distinctly impressive. Although the nature of our results is such that during the development we have occasion to cite only a few works, we take the opportunity before beginning the development to draw the reader's attention to the bibliography, which includes a variety of notable papers appearing in recent years. In particular, an appreciation of the breadth and quality of research ideas surrounding hyperbolic polynomials can be fostered by browsing [1], [2], [8], [9], and [10].

2. **Overview of Results**

Let ϕ be a univariate polynomial all of whose coefficients are real. Between any two real roots of ϕ there lies, of course, a root of ϕ'. Consequently, because ϕ' is of degree one less than the degree of ϕ, a simple counting argument shows that if all of the roots of ϕ are real, then so are all of the roots of ϕ'.

In particular, if $\phi(t) := p(x+te)$ where p is a polynomial hyperbolic in direction e (and where x is an arbitrary point), then all the roots of $t \mapsto \phi'(t) = Dp(x+te)[e]$ are real, where $Dp(x)$ denotes the differential of p at x. Hence, the polynomial $p'_e(x) := Dp(x)[e]$ is, like p, hyperbolic in direction e.

For example, if $p(x) = \prod_j x_j$ and all coordinates of e are positive, then $p'_e(x) = \sum_i e_i \prod_{j \neq i} x_j$ is hyperbolic in direction e.

We refer to p'_e as the "derivative polynomial (in direction e)," and denote its hyperbolicity cone by $\Lambda'_{++,e}$. The fact that the roots of ϕ' lie between the roots of ϕ implies $\Lambda_{++} \subseteq \Lambda'_{++,e}$ (see §4 of [13]), – in words, $\Lambda'_{++,e}$ is a *relaxation* of Λ_{++}.

Of course one can in turn take the derivative in direction e of the hyperbolic polynomial p'_e, thereby obtaining yet another polynomial – $(p'_e)'_e(x) = D^2p(x)[e,e]$ – hyperbolic in direction e. Letting n denote the degree of p, repeated differentiation in direction e results in a sequence of hyperbolic polynomials

$$p_e^{(1)} = p'_e,\, p_e^{(2)}, \ldots, p_e^{(n-1)} \;,$$

[1]Central Path $:= \{x(\eta) : \eta > 0\}$ where $x(\eta)$ solves $\min_x \eta c^* x - \ln p(x)$, s.t. $Ax = b$.

where $\deg(p_e^{(i)}) = n - i$. (For convenience, let $p_e^{(0)} := p$.) The associated hyperbolicity cones $\Lambda_{++,e}^{(i)}$ – and their closures $\Lambda_{+,e}^{(i)}$ – form a nested sequence of relaxations of the original cone:

$$\Lambda_+ \subseteq \Lambda_{+,e}^{(1)} \subseteq \Lambda_{+,e}^{(2)} \subseteq \cdots \subseteq \Lambda_{+,e}^{(n-1)} .$$

The final relaxation, $\Lambda_{+,e}^{(n-1)}$, is a halfspace, because $p_e^{(n-1)}$ is linear.

The cones become tamer as additional derivatives are taken. The halfspace $\Lambda_{+,e}^{(n-1)}$ is as tame as a cone can be, but extremely tame also is the second-order cone $\Lambda_{+,e}^{(n-2)}$ (no cone with curvature could be nicer). As one moves along the nesting towards the original cone Λ_+, the boundary structures gain more and more corners. For example, when $p(x) = \prod_j x_j$ and all coordinates of e are positive, the boundary $\partial\Lambda_{+,e}^{(i)}$ contains all of the non-negative orthant's faces of dimension less than $n - i$ (hence *lots* of corners when i is small and n large). On the other hand, everywhere else, $\partial\Lambda_{+,e}^{(i)}$ has nice curvature properties (no corners). (See [13].)

Before moving to discussion of hyperbolic programs, we record a characterization of the derivative cones that is useful both conceptually and in proofs:

$$\Lambda_{+,e}^{(i)} = \{x : p_e^{(j)}(x) \geq 0 \text{ for all } j = i, \ldots, n-1\} . \tag{2.1}$$

The characterization is immediate from Proposition 18 and Theorem 20 of [13].

✦ ✦ ✦

Consider a hyperbolic program

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

and its derivative relaxations in direction e,

$$\left.\begin{array}{rl} \min & c^*x \\ \text{s.t.} & Ax = b \\ & x \in \Lambda_{+,e}^{(i)} \end{array}\right\} \text{HP}_e^{(i)} \quad (i = 1, \ldots, n-1) .$$

Strictly speaking, "min" should be replaced with "inf," but we focus on instances where a minimizer exists. The optimal values for the derivative relaxations $\text{HP}_e^{(i)}$ form a decreasing sequence in i, due to the nesting of the derivative cones.

Let Feas (resp., $\text{Feas}_e^{(i)}$) denote the feasible region of HP (resp., $\text{HP}_e^{(i)}$) – the set of points satisfying the constraints. Let Opt ($\text{Opt}_e^{(i)}$) denote the set of optimal points – a.k.a. optimal solutions – and let val denote the optimal value of HP.

We shall only be concerned with derivative directions e satisfying $Ae = b$. Thus, $e \in \text{Feas}_e^{(i)}$ for all i – in particular, $\text{HP}_e^{(i)}$ is "strictly" feasible, as $e \in \Lambda_{++} \subseteq \Lambda_{++,e}^{(i)}$.

We assume Λ_+ is a regular cone (i.e., contains no line (infinite in both directions) in addition to having non-empty interior). Then, for $1 \leq i \leq n-2$, $\Lambda_{+,e}^{(i)}$ is regular, too ([13], Proposition 13).

We assume Opt is a non-empty bounded set, but we make no assumptions regarding $\text{Opt}_e^{(i)}$ for $1 \leq i \leq n-1$. Depending on the choice of the direction e, it can happen that $\text{Opt}_e^{(i)} = \emptyset$, either because $\text{HP}_e^{(i)}$ has unbounded optimal value or because the (infimum) optimal value is bounded but unattained. This leads us to make the following definition.

Definition. *For $1 \leq i \leq n-1$, the i^{th}* central swath *is the set*

$$\text{Swath}(i) = \{e \in \Lambda_{++} : Ae = b \text{ and } \text{Opt}_e^{(i)} \neq \emptyset\} .$$

The swaths are nested:

$$\mathrm{Swath}(1) \supseteq \cdots \supseteq \mathrm{Swath}(n-1) \ , \tag{2.2}$$

as is proven[2] in §3.12.

Whereas a path is narrow (one-dimensional), swaths can be broad, just as Swath(i) generally fills most of the feasible region for HP when i is small. But why do we use the terminology "*central* swaths" rather than simply "swaths"? The following elementary theorem (proven in §3.2) gives a reason.

Theorem 1. $\mathrm{Swath}(n-1) = \mathrm{Central\,Path}$

The central path – that is, Swath($n-1$) – leads to optimality for HP, but it is not immediately evident how Swath(i) might lead to optimality when $i < n-1$. In this regard, consider the idealized setting where given $e \in \mathrm{Swath}(i)$, a point $x(e) \in \mathrm{Opt}_e^{(i)}$ is easily determined. Clearly, if it happens that $x(e)$ is in Λ_+, then $x(e)$ is optimal for HP as well as for $\mathrm{HP}_e^{(i)}$, and thus our goal of solving HP has been accomplished. If, on the other hand, $x(e)$ is not in Λ_+, we might attempt to change the derivative direction e in a way that somehow moves us closer to optimality for HP.

Improving the objective value of the derivative direction is easy. Indeed, just move e towards $x(e)$, thereby obtaining a new derivative direction $\bar{e} = (1-t)e + tx(e)$ for some $0 < t < 1$. Being careful to choose t sufficiently small to ensure that $\bar{e} \in \Lambda_{++}$, we thereby obtain a new point $\bar{e} \in \mathrm{relint}(\mathrm{Feas})$ (relative interior) with smaller objective value than the old point e.

As it happens, this strategy makes good sense for the ultimate goal of solving HP – good sense, that is, at least when the movement is done infinitesimally. To make this claim precise requires first that the strategy be formalized. The following theorem (proven in §3.3) provides the foundations.

Define

$$\mathrm{Core}(i) := \{e \in \mathrm{Swath}(i) : \mathrm{Opt}_e^{(i)} = \mathrm{Opt}\} \ .$$

(From $\mathrm{Feas} \subseteq \mathrm{Feas}_e^{(i)} \subseteq \mathrm{Feas}_e^{(i+1)}$ easily follows the nesting $\mathrm{Core}(i) \supseteq \mathrm{Core}(i+1)$.)

Theorem 2. *Assume* $1 \le i \le n-2$.

(A) *The set* $\mathrm{Swath}(i) \setminus \mathrm{Core}(i)$ *is open in the relative topology of* $\mathrm{relint}(\mathrm{Feas})$.

(B) *If* $e \in \mathrm{Swath}(i) \setminus \mathrm{Core}(i)$ *then* $\mathrm{Opt}_e^{(i)}$ *consists of a single point* $x_e^{(i)}$. *Moreover,* $x_e^{(i)} \notin \Lambda_{+,e}^{(i-1)}$ *(hence* $x_e^{(i)}$ *is not in* Λ_+*, thus not in* Opt*), but* $x_e^{(i)} \in \Lambda_{++,e}^{(i+1)}$.

(C) *The map* $e \mapsto x_e^{(i)}$ *is analytic on* $\mathrm{Swath}(i) \setminus \mathrm{Core}(i)$.

We use the notation $x_e^{(i)}$ only when both $1 \le i \le n-2$ and $e \in \mathrm{Swath}(i) \setminus \mathrm{Core}(i)$.

The strategy of moving e towards its optimal solution $x(e)$ is formalized by the dynamics

$$\dot{e} = x_e^{(i)} - e$$

and by considering the maximal trajectories $\{e(t) : 0 \le t < T\}$ that result. (If $T < \infty$, the trajectory started at $e = e(0)$ reaches the boundary of $\mathrm{Swath}(i) \setminus \mathrm{Core}(i)$ in finite time.) Before stating our main theorem regarding these dynamics, we give a simple motivational theorem (proven in §3.4) which provides another reason for appropriateness of the terminology "*central* swaths."

Theorem 3. *Assume* $i = n-2$. *If* $e(0)$*, the initial derivative direction, is on the central path, then the trajectory* $t \mapsto e(t)$ *exactly follows the central path.*

---

[2] The nesting almost is trivial, due to the containments $\mathrm{Feas}_e^{(i)} \subseteq \mathrm{Feas}_e^{(i+1)}$, but the case has to be ruled out that $\mathrm{Opt}_e^{(i+1)}$ is an unbounded set and, simultaneously, $\mathrm{HP}_e^{(i)}$ has unattained (but finite) optimal value.

Main Theorem (Part I). *Assume* $1 \le i \le n-2$ *and let* $\{e(t) : 0 \le t < T\}$ *be a maximal trajectory generated by the dynamics* $\dot{e} = x_e^{(i)} - e$*.*

(A) *The trajectory* $\{e(t) : 0 \le t < T\}$ *forms a bounded set, and* $t \mapsto c^* x_{e(t)}^{(i)}$ *is strictly increasing, with* val *(the optimal value of* HP*) as the limit.*
(B) *If* $T = \infty$*, then every limit point of the trajectory* $t \mapsto e(t)$ *is optimal for* HP*.*
(C) *If* $T < \infty$*, then the trajectory* $t \mapsto e(t)$ *has a unique limit point* $\bar{e}$ *and* $\bar{e} \in \mathrm{Core}(i)$*; moreover, the path* $t \mapsto x_{e(t)}^{(i)}$ *is bounded and each of its limit points is optimal for* HP*.*

The Main Theorem (Part I) is proven in §3.8.

From the two theorems above, we see that the central path is but one trajectory in a rich spectrum of paths. Moreover, the central path is at the far end of the spectrum, where the cone Λ_+ is relaxed to second-order cones $\Lambda_{+,e}^{(n-2)}$. Second-order cones have nice curvature properties but are far cruder approximations to Λ_+ than are cones $\Lambda_{+,e}^{(i)}$ for small i. This raises the interesting prospect that algorithms more efficient than interior-point methods can be devised by relying on a smaller value of i, or on a range of values of i in addition to $i = n-2$.

Some exploration in this vein has been made by Zinchenko ([14],[16]), who showed for linear programs satisfying standard non-degeneracy conditions that if i is chosen appropriately and the initial derivative direction $e(0)$ is within a certain region, then a particular algorithm based on discretizing the flow $\dot{e} = x_e^{(i)} - e$ converges R-quadratically to an optimal solution.

The Main Theorem potentially can be embellished without restricting the general setting. For example, in Part I(B) there is no statement that limit points of the path $t \mapsto x_{e(t)}^{(i)}$ are optimal solutions for HP – there is not even a statement that the path is bounded. This omission seems odd given that the trajectory $t \mapsto e(t)$ is following the path $t \mapsto x_{e(t)}^{(i)}$ (according to the dynamics $\dot{e} = x_e^{(i)} - e$) and given that the theorem states limit points of the trajectory are optimal for HP. Intuitively, it seems the path would converge to optimality and do so even more quickly than the trajectory. The intuition is correct for a wide variety of "nondegenerate" problems (indeed, quicker convergence of the path than the trajectory underlies Zinchenko's speedup), but no proof – or counterexample – has been found for the general setting of the theorem.

To gain a sense of the difficulties (and how intuition can mislead), consider that for any value $0 < T \le \infty$, it is straightforward to define a dynamics $\dot{a}, \dot{b} \in \mathbb{R}^2$ that generates a pair of paths $a(t)$, $b(t)$ for which $\dot{b}(t) = a(t) - b(t)$, $a(t)$ spirals outward to infinity, and $b(t)$ spirals inward to a point, as $t \to T$.

The theorem potentially can be embellished also in that, when $T = \infty$, it does not assert the trajectory $t \mapsto e(t)$ has a unique limit point, as does the central path. Proving uniqueness would be easy if the trajectory were known to be a semialgebraic set (the central path *is* semialgebraic). The theorem leaves open for the general setting the possibility that when $T = \infty$ (resp., $T < \infty$), some trajectories $t \mapsto e(t)$ (resp., some paths $t \mapsto x_{e(t)}^{(i)}$) have non-trivial limit cycles.

✦ ✦ ✦

Of course the dynamics of moving e towards an optimal solution $x(e)$ can also be done for $e \in \mathrm{Core}(i)$. As a matter of formalism, it would be nice to know doing so retains optimality in that the resulting derivative direction also is in $\mathrm{Core}(i)$. The following theorem (proven in §3.12) establishes a bit more.

Theorem 4. *Assume* $e \in \mathrm{Core}(i)$*,* $x \in \mathrm{Opt}$ *and let* $\mathcal{A}$ *be the minimal affine space containing both* e *and* Opt*. Then*

$$\mathcal{A} \cap \Lambda_{++} \subseteq \mathrm{Core}(i) \; .$$

In the following conjecture, the empty set is taken, by default, to be convex.

Conjecture. $\mathrm{Core}(i)$ *is convex.*

✦ ✦ ✦

Much work remains in order to transform the ideas captured in the Main Theorem (Part I) into general and efficient algorithms. For example, devising and analyzing efficient methods for computing $x_e^{(i)}$ given e is, in the general case, a challenging research problem (and is the subject of an upcoming paper). However, computing $x_e^{(n-2)}$ (that is, the case $i = n-2$) amounts simply to solving a least-squares problem and using the quadratic formula. Here, Chua [4], starting with – and extending – ideas similar to ones above, devised and analyzed an algorithm for semidefinite programming (and, more generally, for symmetric cone programming) with complexity bounds matching the best presently known – $O(\sqrt{n}\log(1/\epsilon))$ iterations to reduce the duality gap by a factor ϵ when Λ_+ is the cone of $n \times n$ sdp matrices.

Although in the present work we do not analyze methods for efficiently computing $x_e^{(i)}$, we now present a few results relevant to algorithm design. These results also are important to the proof of the Main Theorem.

The next two theorems apply to any hyperbolic polynomial, not just those that are derivative polynomials $p_e^{(i)}$. Thus, we phrase the discussion using p rather than $p_e^{(i)}$, and HP instead of $\mathrm{HP}_e^{(i)}$. For motivation, think of the situation where one has an approximation to an optimal solution for HP and the goal is to compute a better approximation.

Let $q_e := p/p_e'$, a rational function. The natural domain for q_e is $\Lambda'_{++,e}$, because $p_e'(x) > 0$ for $x \in \Lambda'_{++,e}$ and $p_e'(x) = 0$ for x in the boundary of $\Lambda'_{++,e}$ (unless x happens also to be in the boundary of Λ_{++}, a case we'll need not consider).

The following result is proven in §3.6.

Theorem 6. *The function* $q_e : \Lambda'_{++,e} \to \mathbb{R}$ *is concave.*

Previously, q_e was known to be concave on the smaller cone Λ_{++} (see [1]). For us, the significance of the function being concave on the larger cone Λ'_{++} is that, as the following theorem illustrates, HP can then be reformulated as a linearly-constrained convex optimization problem (no explicit conic constraint). The theorem is proven in §3.7.

Theorem 7. *The optimal solutions* $x \in \Lambda'_{++,e}$ *for* HP *are the same as for the convex optimization problem*

$$\begin{array}{ll} \min_x & -\ln c^*(e-x) - q_e(x) \\ \text{s.t.} & Ax = b \; . \end{array}$$

Since, by Theorem 2(B), $x_e^{(i)} \in \Lambda_{++,e}^{(i+1)}$ whenever $e \in \mathrm{Swath}(i) \setminus \mathrm{Core}(i)$, the following corollary is immediate, except for the assertion regarding the second differential, which is a straightforward consequence of Theorem 14 in [13].

Corollary 8. *If $1 \le i \le n-2$, $e \in \mathrm{Swath}(i) \setminus \mathrm{Core}(i)$ and*

$$f(x) := -\ln c^*(e-x) - \frac{p_e^{(i)}(x)}{p_e^{(i+1)}(x)} \ ,$$

then $x_e^{(i)}$ is the unique optimal solution for the convex optimization problem

$$\begin{array}{rl} \min_x & f(x) \\ \text{s.t.} & Ax = b \ . \end{array}$$

Moreover, $D^2 f(x_e^{(i)})[v,v] > 0$ if v is not a scalar multiple of $x_e^{(i)}$ (in particular, if $v \ne 0$ satisfies $Av = 0$).

A consequence of the assertion regarding the second differential is that Newton's method will converge (quadratically) to $x_e^{(i)}$ if initiated nearby. (Which is not to say that Newton's method is the algorithm of choice for this problem.)

✦ ✦ ✦

Our final result relates the dynamics $\dot{e} = x_e^{(i)} - e$ to the optimization problem dual to HP:

$$\left.\begin{array}{rl} \sup_{y,s} & yb \\ \text{s.t.} & A^*y + s = c^* \\ & s \in \Lambda_+^* \end{array}\right\} \mathrm{HP}^* \ ,$$

where Λ_+^* is the cone dual to Λ_+.[3] A pair (y,s) satisfying the constraints is said to be "strictly" feasible if $s \in \mathrm{int}(\Lambda_+^*)$ (interior).

Letting val^* denote the optimal value of HP^* ($\mathrm{val}^* = -\infty$ if HP^* is infeasible), we have, just as a matter of tracing definitions, the standard result known as "weak duality"[4]: $\mathrm{val}^* \le \mathrm{val}$.

Optimizers expect that if a dynamics provides a natural path-following framework for solving an optimization problem, then not only do the dynamics generate paths leading to (primal) optimality, also the dynamics somehow generate paths leading to dual optimality.

For $e \in \mathrm{Swath}(i) \setminus \mathrm{Core}(i)$, define

$$s_e^{(i)} := \frac{c^*(e-x)}{p_e^{(i+1)}(x)} Dp_e^{(i)}(x) \quad \text{where } x = x_e^{(i)} \ .$$

The following result is proven in §3.11.

Main Theorem (Part II). *Let $1 \le i \le n-2$ and let $\{e(t) : 0 \le t < T\}$ be a maximal trajectory for the dynamics $\dot{e} = x_e^{(i)} - e$. Then $A^*y + s_{e(t)}^{(i)} = c^*$ has a (unique) solution $y = y_{e(t)}^{(i)}$, and the pair $(y_{e(t)}^{(i)}, s_{e(t)}^{(i)})$ is strictly feasible for* HP^*. *Moreover,*

$$y_{e(t)}^{(i)} b = c^* x_{e(t)}^{(i)} \xrightarrow[t \to T]{} \mathrm{val}$$

(in fact, increases to val *strictly monotonically) and the path $t \mapsto (y_{e(t)}^{(i)}, s_{e(t)}^{(i)})$ is bounded.*

---

[3]The dual cone Λ^* consists of the linear functionals $s : \mathcal{E} \to \mathbb{R}$ that have non-negative value everywhere on Λ_+.

[4]This is proven simply by observing that for feasible x and (y,s),

$$c^*x = x(A^*y + s) = yb + xs \ge yb \ .$$

Consequences, of course, are that $\mathrm{val}^* = \mathrm{val}$ ("strong duality") and that the limit points of $t \mapsto (y^{(i)}_{e(t)}, s^{(i)}_{e(t)})$ form a nonempty set, each of whose elements is optimal for HP*.

Thus, although the path $t \mapsto x^{(i)}_{e(t)}$ is infeasible for HP, and the feasible trajectory $t \mapsto e(t)$ can potentially not converge to optimality (it might instead converge to $\bar{e} \in \mathrm{Core}(i)$), there is naturally generated a path $t \mapsto (y^{(i)}_{e(t)}, s^{(i)}_{e(t)})$ that both is feasible for HP* and converges to optimality.

Perhaps, then, the algorithmic framework we have posed as being for the primal optimization problem would be better posed as being for the dual, as then there would be a single path generated for the primal, and that path would be both feasible and converge to optimality.

✦ ✦ ✦

Lastly, we mention that in defining the sequence of derivative polynomials $p^{(i)}_e$ $(i = 1, \ldots, n-1)$, we could have used various derivative directions $e_1, \ldots, e_{n-1}$, choosing e_1 from the hyperbolicity cone for p and defining $p'_{e_1}(x) := Dp(x)[e_1]$, choosing e_2 from the (larger) hyperbolicity cone for p'_{e_1} and defining $p''_{e_1,e_2}(x) := Dp'_{e_1}(x)[e_2] = D^2p(x)[e_1, e_2]$, and so on. Several results in the following pages can be extended to this more general setting. However, computing multidirectional derivatives could be prohibitively expensive, even for the innocuous-appearing hyperbolic polynomials $p(x) = \prod_{i=1}^n a_i^T x$ naturally arising from polyhedral cones $\{x \in \mathbb{R}^n : a_i^T x \geq 0 \text{ for all } i = 1, \ldots, n\}$; indeed, choosing $e_1, \ldots, e_n$ to be the standard basis, $D^n p(x)[e_1, \ldots, e_n]$ is the permanent of the matrix whose i^{th} column is a_i.

By contrast, if the same direction e is used for all derivatives $(i = 1, 2, \ldots)$, the resulting polynomials $p^{(i)}_e$ (resp., their gradients, their Hessians) can be efficiently evaluated at any point if the initial polynomial p (resp., its gradient, its Hessian) can be efficiently evaluated at any point. This is straightforwardly accomplished by interpolation, and can be sped up via the (inverse) Discrete Fourier Transform (see §9 of [13] for some discussion). As a primary motivation for the present paper is designing efficient algorithms, it thus is sensible to restrict consideration to the same direction e being used for all derivatives $(i = 1, \ldots, n-1)$.

3. **Proofs**

3.1. **Preliminary Remarks.** The theorems are proven in the order in which they were stated with the exceptions of Main Theorem (Part I) and Theorem 4. The proof of the first is delayed because it depends on theorems that were stated later. The proof of Theorem 4 is delayed, until the end, due to the combination of the proof being long and the theorem being less important than others.

Several proofs rely fundamentally on results from [13], a paper on structural aspects of hyperbolicity cones (and hyperbolic programs), a paper for which a primary goal was to provide a ready reference of "lower-level" details so that subsequent papers (such as the present one) could avoid drawn-out proofs. The relevant results for a proof are mentioned before the proof gets underway, with the exception of the proof of Theorem 2 (a proof which amounts to little more than stringing together results from [13]).

One result from [13] is used time and again – the characterization we recorded as (2.1), that is,

$$\Lambda^{(i)}_{+,e} = \{x : p^{(j)}_e(x) \geq 0 \text{ for all } j = i, \ldots, n-1\} \,. \tag{3.1.1}$$

For ease of reference we record the following two characterizations that are very similar to the one above:

$$\Lambda^{(i)}_{++,e} = \{x : p^{(j)}_e(x) > 0 \text{ for all } j = i, \dots, n-1\} \;, \tag{3.1.2}$$

$$\Lambda^{(i)}_{++,e} = \{x : p^{(i)}_e(x) > 0 \text{ and } p^{(j)}_e(x) \geq 0 \text{ for all } j = i+1, \dots, n-1\} \;. \tag{3.1.3}$$

As was noted earlier, (3.1.1) is immediate from Proposition 18 and Theorem 20 in [13]. The representations (3.1.2) and (3.1.3) are established in the two paragraphs following that theorem.

Readily proven from the above characterizations are that for $1 \leq i \leq n-1$,

$$x \in \Lambda^{(i)}_{+,e} \setminus \Lambda^{(i-1)}_{+,e} \quad \Rightarrow \quad p^{(i-1)}_e(x) < 0 \tag{3.1.4}$$

and

$$x \in \partial \Lambda^{(i)}_{+,e} \quad \Rightarrow \quad p^{(i-1)}_e(x) \leq 0 \;. \tag{3.1.5}$$

A fact used extensively, and which is easily proven (by, say, induction), comes from $p^{(i)}_e$ being homogeneous of degree $n-i$: If $0 \leq j \leq k \leq n-i$ then

$$D^k p^{(i)}_e(x)[\underbrace{x, \dots, x}_{j \text{ times}}] = \tfrac{(n-i-k+j)!}{(n-i-k)!} D^{k-j} p^{(i)}_e(x) \;. \tag{3.1.6}$$

A consequence of (3.1.6) used occasionally is that if $0 \leq k \leq n-i$, then

$$D^k p^{(i)}_e(e) = D^{k+i} p(e)[\underbrace{e, \dots, e}_{i \text{ times}}] = \tfrac{(n-k)!}{(n-i-k)!} D^k p(e) \;; \tag{3.1.7}$$

in particular, $D^k p^{(i)}_e(e)$ is a positive multiple of $D^k p(e)$.

Finally, we mention that the nesting of the derivative cones is often used implicitly. Two examples of assertions made without the nesting being mentioned: "If $x \in \mathrm{Opt}^{(i)}_e$ and $x \in \mathrm{Feas}$, then $x \in \mathrm{Opt}$." "If $x \in \Lambda^{(i)}_{++,e}$ then $p^{(i+1)}_e(x) > 0$."

3.2. **Proof of Theorem 1.** Since $x \mapsto p^{(n-1)}_e(x)$ is linear, we have $p^{(n-1)}_e(x) = Dp^{(n-1)}_e(z)[x]$ for every point z – in particular for $z = e$ – and thus

$$\Lambda^{(n-1)}_{+,e} = \{x : Dp^{(n-1)}_e(e)[x] \geq 0\} \;.$$

Clearly, then, $\mathrm{Opt}^{(n-1)}_e \neq \emptyset$ iff (if and only if)

$$c^* = \lambda D p^{(n-1)}_e(e) + A^* y \quad \text{for some } \lambda > 0 \text{ and } y \;.$$

However, by (3.1.7), $Dp^{(n-1)}_e(e)$ is a positive multiple of $Dp(e)$. Hence, $\mathrm{Opt}^{(n-1)}_e \neq \emptyset$ iff

$$c^* = \lambda D p(e) + A^* y \quad \text{for some } \lambda > 0 \text{ and } y \;.$$

However, this is precisely the first-order condition for e to be on the central path. The condition is sufficient (as well as necessary) because it comes from a linearly-constrained *convex* optimization problem (specifically, $\min_x \eta(c^*x) - \ln p(x)$, s.t. $Ax = b$, with $\eta = 1/\lambda$).

3.3. **Proof of Theorem 2.** We forewarn the reader that this proof, more than any of the others, frequently invokes results from [13].

Proposition 24 in [13] establishes that for all $e \in \Lambda_{++}$, each boundary face of $\Lambda^{(i)}_{+,e}$ is either a face of Λ_+ or is a ray not contained in Λ_+ (this relies on $1 \leq i \leq n-2$ and Λ_+ being regular). Consequently, for e satisfying $\mathrm{Opt}^{(i)}_e \neq \emptyset$, either $\mathrm{Opt}^{(i)}_e = \mathrm{Opt}$ or $\mathrm{Opt}^{(i)}_e$ consists of a single point $x^{(i)}_e$ not contained in Λ_+ (using that $b \neq 0$ and hence at most one point on any ray satisfies $Ax = b$). Since we are assuming $e \in \mathrm{Swath}(i) \setminus \mathrm{Core}(i)$, the latter case holds.

Since $x_e^{(i)}$ is in the boundary $\partial\Lambda_{+,e}^{(i)}$ but not in Λ_+, it holds that $x_e^{(i)} \notin \Lambda_{+,e}^{(i-1)}$ (by [13], Proposition 16) and that $x_e^{(i)} \notin \partial\Lambda_{+,e}^{(i+1)}$ (by the same proposition but with $i+1$ substituted for i). We have now proven part (B) of the theorem.

To complete the arguments we rely on the following general proposition, whose proof (left to the reader) follows entirely standard lines, and is primarily an application of the Implicit Function Theorem to optimality conditions (see, for example, §2.4 of [5] for similar results).

Proposition 3.3.1. *Let x_z be a local optimum for*

$$\begin{array}{rl} \min_x & c^* x \\ s.t. & Ax = b \\ & x \in S_z \ , \end{array} \tag{3.3.1}$$

where S_z is the closure of a connected component of $\{x : f_z(x) > 0\}$, and where $(x, z) \mapsto f_z(x)$ is analytic. Assume A is of full rank, and c^* is not in the image of A^*.

Assume $Df_z(x_z) \neq 0$, and assume $D^2 f_z(x_z)[v, v] < 0$ for all $v \neq 0$ that satisfy both $Df_z(x_z)[v] = 0$ and $Av = 0$. Then there exists an open neighborhood U of z, and an analytic function $\bar{z} \mapsto x(\bar{z})$ defined on U and satisfying $x(z) = x_z$, such that for all $\bar{z} \in U$, the point $x(\bar{z})$ is an isolated local optimum for the optimization problem (3.3.1) obtained upon substitution of $\bar{z}$ for z.

Let $x = x_e^{(i)}$. Due to the proposition, to conclude the proof of the theorem we need only show $Dp_e^{(i)}(x) \neq 0$ and $D^2 p_e^{(i)}(x)[v, v] < 0$ for all $v \neq 0$ satisfying both $Dp_e^{(i)}(x)[v] = 0$ and $Av = 0$.

Since $x \notin \partial\Lambda_{+,e}^{(i+1)}$ (as we saw at the beginning of the proof), we have $p_e^{(i+1)}(x) \neq 0$, and thus $Dp_e^{(i)}(x) \neq 0$ (because $0 \neq p_e^{(i+1)}(x) = Dp_e^{(i)}(x)[e]$).

Assume $v \neq 0$ satisfies $Dp_e^{(i)}(x)[v] = 0$ and $Av = 0$. By the latter condition, v is not a scalar multiple of x (because $Ax = b \neq 0$). Since $x \notin \Lambda_{+,e}^{(i-1)}$ (as we saw at the beginning of the proof), Theorem 14 of [13] thus gives $D^2 p_e^{(i)}(x)[v, v] < 0$ (where the theorem is applied with $\Lambda_{+,e}^{(i)}$ (resp., $\Lambda_{+,e}^{(i-1)}$) substituted for Λ'_+ (resp., Λ_+) – that theorem's requirement "$x \in \partial^1 \Lambda'_+$" then is simply the combination of conditions $x \in \partial\Lambda_{+,e}^{(i)}$ and $Dp_e^{(i)}(x) \neq 0$).

3.4. **Proof of Theorem 3.** For any $e \in \Lambda_{++}$, the polynomial $p_e^{(n-2)}$ is homogeneous and quadratic, and thus $p_e^{(n-2)}(x) = \frac{1}{2} x^T M_e x$ for some matrix M_e. Since $\Lambda_{+,e}^{(n-2)}$ is regular (as, was noted in §2, follows from regularity of Λ_+ ([13], Proposition 13)), M_e is invertible. Consequently, $\mathrm{HP}_e^{(n-2)}$ has either a unique optimal solution $x_e^{(n-2)}$ or no optimal solution.

Of course $M_e = D^2 p_e^{(n-2)}(x)$ for all x – in particular for $x = e$. On the other hand, according to (3.1.7), $D^2 p_e^{(n-2)}(e)$ is a positive multiple of $D^2 p(e)$. Also, according to (3.1.6), $D^2 p(e)[e]$ is a positive multiple of $Dp(e)$. Thus, M_e (resp., $M_e e$) is a positive multiple of $D^2 p(e)$ (resp., $Dp(e)$).

Let e be on the central path. Since $M_e e$ is a positive multiple of $Dp(e)$, from the first-order conditions for optimality,

$$\lambda c = M_e e + A^T y \quad \text{for some } \lambda > 0 \text{ and } y \ .$$

Let v be the vector for which there exists w satisfying

$$\begin{array}{c} c = M_e v + A^T w \\ Av = 0 \ . \end{array}$$

Using that M_e is a positive multiple of $D^2p(e)$, one readily verifies v is tangent to the central path and points in the direction of decreasing objective value (points in the direction that interior-point methods traverse the central path).

Since $v^T M_e v = c^T v < 0$ (as v is direction of decreasing objective value), and $e^T M_e e = Dp(e)[e] = np(e) > 0$, there exists $t > 0$ satisfying $(e+tv)^T M_e (e+tv) = 0$; let $\hat{t}$ be the smallest such positive value. Let $\hat{x} = e + \hat{t}v$. Clearly,

$$\begin{aligned} (\lambda + \hat{t})c &= M_e \hat{x} + A^T(y + \hat{t}w) \\ A\hat{x} &= b \\ \hat{x} &\in \partial\Lambda_{+,e}^{(n-2)} \ . \end{aligned}$$

Thus, $\hat{x}$ is the unique point satisfying the first-order optimality conditions for $\mathrm{HP}_e^{(n-2)}$ (uniqueness is a consequence of the invertibility of M_e); hence, $\hat{x} = x_e^{(n-2)}$. Since $\hat{x} - e$ is a positive multiple of v, the proof is complete.

For reference in the following subsection, we record that we have proven

$$e \in \text{Central Path} \quad \Rightarrow \quad e \in \text{Swath}(n-2) \ . \tag{3.4.1}$$

3.5. **Proof of Nesting of Swaths.** Here we prove the nesting claimed in (2.2), that is, $\text{Swath}(i) \subseteq \text{Swath}(i-1)$ for $i = 2, \ldots, n-1$.

First assume $2 \leq i \leq n-2$, and $e \in \text{Swath}(i)$. By Theorem 2, either $\text{Opt}_e^{(i)} = \text{Opt}$ or $\text{Opt}_e^{(i)}$ consists of a single point. Thus, $\text{Opt}_e^{(i)}$ is bounded as well as nonempty, which together with $\emptyset \neq \text{Feas}_e^{(i-1)} \subseteq \text{Feas}_e^{(i)}$ implies by standard convexity arguments that $\text{Opt}_e^{(i-1)} \neq \emptyset$, that is, $e \in \text{Swath}(i-1)$.

On the other hand, if $i = n-1$, that is, if $e \in$ Central Path (Theorem 1), then according to (3.4.1), $e \in \text{Swath}(n-2)$.

3.6. **Proof of Theorem 6.** Introduce a new variable t and let $P(x,t) := t\,p(x)$, a polynomial that is easily seen to be hyperbolic in direction $E = (e,1)$. Let K' be the hyperbolicity cone for the derivative polynomial P_E' – thus, K' is the connected component of $S := \{(x,t) : p(x) + tp_e'(x) > 0\}$ containing E. We claim K' is precisely the epigraph of $-q_e : \Lambda_{++,e}' \to \mathbb{R}$. Since K', being a hyperbolicity cone, is convex, establishing the claim will establish the theorem.

If $x \in \partial\Lambda_{+,e}'$ then $p_e'(x) = 0$ and, according to (3.1.5), $p(x) \leq 0$, so for no t do we have $(x,t) \in K'$. Thus, now assuming $x \in \Lambda_{++,e}'$ and $t > -q_e(x)$, to establish the claim it clearly suffices to show there is a path in S from (x,t) to $E = (e,1)$.

Choose $\bar{t}$ satisfying $\bar{t} > \max\{-q_e(sx + (1-s)e) : 0 \leq s \leq 1\}$ – the maximum exists because the line segment connecting x to e is contained in the convex set $\Lambda_{++,e}'$ and because p_e' is positive everywhere in $\Lambda_{++,e}'$. It is easily verified that a path in S from (x,t) to $(e,1)$ is obtained with three line segments; the line segment between (x,t) and $(x,\bar{t})$, the line segment between $(x,\bar{t})$ and $(e,\bar{t})$, and the line segment between $(e,\bar{t})$ and $(e,1)$.

3.7. **Proof of Theorem 7.** For x in $\Lambda_{++,e}'$ we have $Dp(x) \neq 0$ (because $0 < p_e'(x) = Dp(x)[e]$). Consequently, necessary and sufficient conditions for $x \in \Lambda_{++,e}'$ to be optimal for the convex optimization problem HP are

$$\begin{aligned} \lambda c^* &= Dp(x) + A^* y \qquad \text{for some } \lambda > 0 \text{ and } y \\ Ax &= b \\ x &\in \partial\Lambda_+ \ . \end{aligned}$$

Observe that these conditions and homogeneity of p give

$$\lambda c^*(e - x) = Dp(x)[e - x] = p_e'(x) - n\,p(x) = p_e'(x) \ ,$$

that is,

$$\lambda = \frac{p_e'(x)}{c^*(e-x)} \; .$$

On the other hand, necessary and sufficient conditions for $x \in \Lambda'_{++,e}$ to solve the convex optimization problem

$$\begin{array}{ll} \min & -\ln c^*(e-x) - q_e(x) \\ \text{s.t.} & Ax = b \end{array} \tag{3.7.1}$$

are

$$\begin{array}{c} \frac{1}{c^*(e-x)} c^* - Dq_e(x) = A^*w \qquad \text{for some } w \\ Ax = b \; . \end{array}$$

Observe that these conditions along with homogeneity of p and p_e' give

$$\begin{aligned} 1 &= \tfrac{1}{c^*(e-x)} c^*(e-x) \\ &= Dq_e(x)[e-x] \\ &= \tfrac{1}{p_e'(x)} \Big(Dp(x)[e-x] - \tfrac{p(x)}{p_e'(x)} Dp_e'(x)[e-x] \Big) \\ &= \tfrac{1}{p_e'(x)} \Big(p_e'(x) - n\, p(x) - \tfrac{p(x)}{p_e'(x)} \left(p_e''(x) - (n-1)\, p_e'(x) \right) \Big) \\ &= 1 - \frac{p(x)}{p_e'(x)} \left(1 + \frac{p_e''(x)}{p_e'(x)} \right) \; . \end{aligned}$$

Thus, either $p(x) = 0$ or $p_e''(x) = -p_e'(x)$. However, the domain of the optimization problem (3.7.1) is $\Lambda'_{++,e}$, and $\Lambda'_{++,e} \subseteq \Lambda''_{++,e}$; hence, both $p_e'(x)$ and $p_e''(x)$ are positive. Consequently, any point $x \in \Lambda'_{++,e}$ satisfying the conditions will satisfy $p(x) = 0$ and, thus, $x \in \partial \Lambda_+$ (by (3.1.1) and (3.1.2)). Consequently, since $p(x) = 0$ also implies $Dq_e(x) = \frac{1}{p_e'(x)} Dp(x)$, the optimality conditions can be rewritten as

$$\begin{array}{c} \frac{1}{c^*(e-x)} c^* - \frac{1}{p_e'(x)} Dp(x) = A^*w \qquad \text{for some } w \\ Ax = b \\ x \in \partial \Lambda_+ \; . \end{array}$$

Clearly, now, the first-order conditions for the two convex optimization problems are identical when $x \in \Lambda'_{++,e}$. The proof is complete.

3.8. **Proof of Main Theorem (Part I).** We begin by recording that Theorem 2(B) and (3.1.4) easily give

$$e \in \mathrm{Swath}(i) \setminus \mathrm{Core}(i) \quad \Rightarrow \quad \Big(p_e^{(i-1)}(x_e^{(i)}) < 0 \Big) \wedge \Big(p_e^{(i+1)}(x_e^{(i)}) > 0 \Big) \; . \tag{3.8.1}$$

Fix $1 \leq i \leq n-2$ and assume $\{e(t) : 0 \leq t < T\}$ is a maximal trajectory generated by the dynamics $\dot{e} = x_e^{(i)} - e$. Trivially, $t \mapsto c^*e(t)$ is a decreasing function. Consequently, since the trajectory is contained in Feas, the trajectory is bounded due to our assumption that the nonempty set Opt is bounded.[5]

For the remainder of the proof let $x(t) := x_{e(t)}^{(i)}$.

We now show $t \mapsto c^*x(t)$ is strictly increasing. Here and later in the paper we must consider $p_e^{(i)}(x)$ as a function of e as well as of x. In particular, we need to compute differentials wrt (with respect to) e. But this is easy. Indeed, for nonnegative integers j and k, letting $D_e(D^j p_e^{(k)}(x))$ denote the differential wrt e of

[5]Simple convexity arguments show that Opt being nonempty and bounded implies $\{x \in \text{Feas} : c^*x \leq \alpha\}$ is bounded for each value α.

$e \mapsto D^j p_e^{(k)}(x)$ (where D^j is the j^{th} differential wrt x), we have

$$\begin{aligned} D_e(D^j p_e^{(k)}(x)) &= D_e(D^{j+k}p(x)[\underbrace{e,\ldots,e}_{k \text{ times}}]) \\ &= kD^{j+k}p(x)[\underbrace{e,\ldots,e}_{k-1 \text{ times}}] \\ &= kD^{j+1}p_e^{(k-1)}(x) \ . \end{aligned} \tag{3.8.2}$$

From $t \mapsto p_{e(t)}^{(i)}(x(t)) \equiv 0$ we find for $e = e(t)$, $x = x(t)$ that

$$\begin{aligned} 0 &= \tfrac{d}{dt} p_{e(t)}^{(i)}(x(t)) \\ &= Dp_e^{(i)}(x)[\dot{x}] + D_e(p_e^{(i)}(x))[\dot{e}] \\ &= Dp_e^{(i)}(x)[\dot{x}] + iDp_e^{(i-1)}(x)[\dot{e}] \\ &= Dp_e^{(i)}(x)[\dot{x}] + i\left((n-i+1)p_e^{(i-1)}(x) - p_e^{(i)}(x)\right) \ , \end{aligned}$$

using $\dot{e} = x - e$ and $Dp_e^{(i-1)}(x)[x] = (n-i+1)p_e^{(i-1)}(x)$ (by homogeneity). Thus, since $p_e^{(i)}(x) = 0$ and $p_e^{(i-1)}(x) < 0$ (by (3.8.1)), we have

$$Dp_e^{(i)}(x)[\dot{x}] > 0 \ . \tag{3.8.3}$$

On the other hand, $x = x(t)$ satisfies the first-order condition

$$\lambda c^* = Dp_e^{(i)}(x) + A^* y \qquad \text{for some } \lambda > 0 \text{ and } y \ ,$$

where $e = e(t)$. Applying both sides to $\dot{x}$, using $A\dot{x} = 0$ and substituting (3.8.3), shows $t \mapsto c^* x(t)$ indeed is strictly increasing.

We have now established Part I(A) of the Main Theorem, except for showing that $c^* x(t)$ converges to val, the optimal value for HP.

To complete the proof of Part I of the Main Theorem for the case $T = \infty$ (i.e., when the timeline for the maximal trajectory is infinite), note it trivially follows from $\dot{e}(t) = x(t) - e(t)$ and the relations $c^* x(t) \le \text{val} \le c^* e(t)$ that $c^* e(t)$ converges to val, and by also using the monotonicity of $t \mapsto c^* x(t)$, that $c^* x(t)$ converges to val. Moreover, since the trajectory $t \mapsto e(t)$ is contained in Feas, of course the trajectory's limit points are thus optimal for HP. This completes consideration of the case $T = \infty$.

For the much harder case $T < \infty$ (that is, when the trajectory reaches the boundary of $\text{Swath}(i) \setminus \text{Core}(i)$ in finite time), we split the analysis into two propositions.

Proposition 3.8.1. *If $T < \infty$ then one of the following two cases holds.*

(A) *The trajectory $t \mapsto e(t)$ has a unique limit point $e \in \text{Core}(i)$, and the path $t \mapsto x(t)$ is bounded, with all limit points contained in* Opt.

(B) *It holds that $\liminf_{t \to T} \|x(t)\| = \infty$. Moreover, there exists a constant $\alpha < 0$ such that for all $0 \le t < T$,*

$$p_{e(t)}^{(i-1)}(x(t)) \le \alpha \|x(t)\|^{n-i+1} \ .$$

Proposition 3.8.2. *Regardless of whether T is finite,*

$$\frac{d}{dt} \ln \frac{\left(-p_{e(t)}^{(i-1)}(x(t))\right)^{n-i}}{\left(p_{e(t)}^{(i+1)}(x(t))\right)^{n-i+1}} \le C_1 \frac{p_e^{(i-1)}(x)}{p_e^{(i+1)}(x)} + C_2 \frac{p_e^{(i-2)}(x)}{p_e^{(i-1)}(x)} + C_3 \ ,$$

where $e = e(t)$, $x = x(t)$,

$$C_1 = i(n-i+1)^2 \ , \quad C_2 = (i-1)(n-i)(n-i+2) \quad \text{and} \quad C_3 = 2n - i + 1$$

(and where for the case $i = 1$ we define $p_e^{(-1)} \equiv 0$).

The propositions are proven in the two subsequent subsections. Assume their validity for now, and assume that the conditions of case (B) in Proposition 3.8.1 hold. Clearly, then, to finish proving the Part I of the Main Theorem, it suffices to use Proposition 3.8.2 to obtain a contradiction.

Let K be a compact set containing the entire (bounded) trajectory $t \mapsto e(t)$ and for $j = i-2, i-1, i+1$, define

$$\gamma_j := \max\{|p_e^{(j)}(x)| : e \in K \text{ and } \|x\| = 1\}$$

(let $\gamma_{i-2} := 0$ when $i = 1$). Since $p_e^{(j)}$ is homogeneous of degree $n-j$, we have

$$|p_e^{(j)}(x)| \leq \gamma_j \, \|x\|^{n-j} \quad \text{for all } e \in K \text{ and } x \, .$$

Using the assumed bound $p_{e(t)}^{(i-1)}(x(t)) \leq \alpha \|x(t)\|^{n-i+1}$ – keep in mind especially that α is negative – and using the positivity of $p_{e(t)}^{(i+1)}(x(t))$ (according to (3.8.1)), from the inequality of Proposition 3.8.2 we find

$$\frac{d}{dt} \ln \frac{\left(-p_{e(t)}^{(i-1)}(x(t))\right)^{n-i}}{\left(p_{e(t)}^{(i+1)}(x(t))\right)^{n-i+1}} \leq C_1 \, \frac{\alpha}{\gamma_{i+1}} \, \|x(t)\|^2 + C_2 \, \frac{\gamma_{i-2}}{|\alpha|} \, \|x(t)\| + C_3 \, .$$

Hence, there exists a value r for which

$$\|x(t)\| > r \quad \Rightarrow \quad \frac{d}{dt} \ln \frac{\left(-p_{e(t)}^{(i-1)}(x(t))\right)^{n-i}}{\left(p_{e(t)}^{(i+1)}(x(t))\right)^{n-i+1}} < 0 \, .$$

Since $\liminf_{t \to T} \|x(t)\| = \infty$ (by assumption), it follows that

$$\sup_{0 \leq t < T} \frac{\left(-p_{e(t)}^{(i-1)}(x(t))\right)^{n-i}}{\left(p_{e(t)}^{(i+1)}(x(t))\right)^{n-i+1}} < \infty \, .$$

However,

$$\frac{\left(-p_{e(t)}^{(i-1)}(x(t))\right)^{n-i}}{\left(p_{e(t)}^{(i+1)}(x(t))\right)^{n-i+1}} \geq \frac{\left(|\alpha| \, \|x(t)\|^{n-i+1}\right)^{n-i}}{\left(\gamma_{i+1} \|x(t)\|^{n-i-1}\right)^{n-i+1}}$$

$$= \frac{|\alpha|^{n-i}}{\gamma_{i+1}^{n-i-1}} \, \|x(t)\|^{n-i+1} \to \infty \, ,$$

a contradiction, thus concluding the proof (except for proving the two propositions).

3.9. **Proof of Proposition 3.8.1.** We make use of the fact that, by Proposition 24 of [13], each boundary face of $\Lambda_{+,e}^{(i)}$ is either a face of Λ_+ or a single ray not contained in Λ_+ (this relies on $i \leq n-2$ and Λ_+ being regular). We also make use of the fact that, by Proposition 16 of [13],

$$x \in (\partial \Lambda_{+,e}^{(i)}) \cap \Lambda_{+,e}^{(i-1)} \quad \Rightarrow \quad x \in \Lambda_+ \, . \tag{3.9.1}$$

Let $1 \leq i \leq n-2$, assume $\{e(t) : 0 \leq t < T\}$ is a maximal trajectory for the dynamics $\dot{e} = x_e^{(i)} - e$, and assume $T < \infty$.

We begin by showing that all limit points of the trajectory lie in Λ_{++}. (Thus, when the trajectory exits Swath$(i) \setminus$ Core(i), the exit is not due to leaving the feasible region for HP.) For this it clearly suffices to show $p(e(t)) \geq p(e(0)) e^{-nt}$ for all t, and hence suffices to show $\frac{d}{dt} p(e(t)) \geq -np(e(t))$, that is, suffices to show $Dp(e(t))[\dot{e}(t)] \geq -np(e(t))$. Since $\dot{e}(t) = x(t) - e(t)$ (where $x(t) := x_{e(t)}^{(i)}$) and since $Dp(e(t))[e(t)] = np(e(t))$ (by (3.1.6)), our current task is reduced to showing $Dp(e(t))[x(t)] \geq 0$.

However, by (3.1.7), for all e, $Dp(e)$ is a positive multiple of $Dp_e^{(n-1)}(e)$. Since for all x, $Dp_e^{(n-1)}(e)[x] = p_e^{(n-1)}(x)$ (because $x \mapsto p_e^{(n-1)}(x)$ is linear), we have $Dp(e)[x] \geq 0$ iff x is in the cone $\Lambda_{+,e}^{(n-1)}$. But $x(t) \in \Lambda_{+,e(t)}^{(i)} \subseteq \Lambda_{+,e(t)}^{(n-1)}$, thus concluding the proof that all limit points of the trajectory $t \mapsto e(t)$ lie in Λ_{++} (under the assumption $T < \infty$).

Always assuming $1 \leq i \leq n-2$, we now turn to establishing two lemmas from which the proposition will follow.

Lemma 3.9.1. *Let $\{e_j\}$ be a sequence of derivative directions converging to $e \in \Lambda_{++}$, and let x_j be optimal for $\mathrm{HP}_{e_j}^{(i)}$.*

(A) *If $x_j \to x$ then $x \in \mathrm{Opt}_e^{(i)}$.*
(B) *If $\{x_j\}$ is an unbounded set then $\mathrm{Opt}_e^{(i)} = \emptyset$ and $\liminf \|x_j\| = \infty$.*
(C) *If $\{x_j\}$ is an unbounded set and $\limsup c^* x_j > -\infty$, then $x_j/\|x_j\|$ has exactly one limit point d; moreover, d satisfies $p_e^{(i-1)}(d) < 0$.*

Proof. First assume $x_j \to x$. Then $p_e^{(k)}(x) = \lim p_{e_j}^{(k)}(x_j)$ for all k. However, by (3.1.1), $p_{e_j}^{(k)}(x_j) \geq 0$ for all $k = i, \ldots, n-1$. Thus, $p_e^{(k)}(x) \geq 0$ for all $k = i, \ldots, n-1$. Hence, again invoking (3.1.1), $x \in \Lambda_{+,e}^{(i)}$. Since, trivially, $Ax = b$, we thus see that $x \in \mathrm{Feas}_e^{(i)}$.

To prove that x not only is feasible but is optimal, we assume otherwise and obtain a contradiction. Thus, assume $\bar{x} \in \mathrm{Feas}_e^{(i)}$ satisfies $c^*\bar{x} < c^*x$. By nudging $\bar{x}$ towards the strictly feasible point e, we may assume $\bar{x}$ is strictly feasible in addition to satisfying $c^*\bar{x} < c^*x$. However, by (3.1.2), strict feasibility of $\bar{x}$ implies $p_e^{(k)}(\bar{x}) > 0$ for all $k = i, \ldots, n-1$. Since $p_{e_j}^{(k)}(\bar{x}) \to p_e^{(k)}(\bar{x})$, we thus have, again using (3.1.2), that $\bar{x} \in \mathrm{Feas}_{e_j}^{(i)}$ for $j > J$ (some J). But then

$$c^*x > c^*\bar{x} \geq \lim c^*x_j = c^*x \ ,$$

a contradiction. Hence, x is optimal, and assertion (A) of the lemma is established.

Now assume $\{x_j\}$ is unbounded. Choose a subsequence $\{x_{j_\ell}\}$ for which $\liminf_{\ell\to\infty} \|x_{j_\ell}\| = \infty$ and let d be a limit point of $\{x_{j_\ell}/\|x_{j_\ell}\|\}$. Then d is a feasible direction for $\mathrm{HP}_e^{(i)}$, that is, satisfies $Ad = 0$ and $d \in \Lambda_{+,e}^{(i)}$. (That $Ad = 0$ is trivial. That $d \in \Lambda_{+,e}^{(i)}$ follows from $p_e^{(i)}(d) = \lim_{\ell\to\infty} p_{e_{j_\ell}}^{(i)}(x_{j_\ell}/\|x_{j_\ell}\|) \geq 0$ for $k = i, \ldots n-1$, the inequality being due to $x_{j_\ell} \in \Lambda_{+,e_{j_\ell}}^{(i)}$ along with homogeneity of $p_{e_{j_\ell}}^{(i)}$.) The feasible direction d satisfies $c^*d \leq 0$ (because $e \in \mathrm{Feas}_{e_{j_\ell}}^{(i)}$ – and hence $\limsup_{\ell\to\infty} c^*x_{j_\ell} < \infty$ – and because $\liminf_{\ell\to\infty} \|x_{j_\ell}\| = \infty$).

Clearly, now, if $\mathrm{Opt}_e^{(i)} \neq \emptyset$ then $\mathrm{Opt}_e^{(i)}$ is unbounded (in direction d), implying by Theorem 2(B) that $\mathrm{Opt}_e^{(i)} = \mathrm{Opt}$. But this would contradict our assumption that Opt is bounded. Thus, $\mathrm{Opt}_e^{(i)} = \emptyset$.

To conclude the proof of assertion (B) of the lemma, it remains to show $\liminf \|x_j\| = \infty$. But if this identity did not hold then $\{x_j\}$ would have a limit point, and thus, by assertion (A) of the lemma, we would have $\mathrm{Opt}_e^{(i)} \neq \emptyset$, contradicting what we just proved. The proof of assertion (B) of the lemma is now complete.

Now we prove assertion (C). Assume $\{x_j\}$ is unbounded and $\limsup c^*x_j > -\infty$.

With $\{x_j\}$ unbounded, we already know from above that every limit point d of $\{x_j/\|x_j\|\}$ is a feasible direction for $\mathrm{HP}_e^{(i)}$ (that is, satisfies $Ad = 0$ and $d \in \Lambda_{+,e}^{(i)}$) and $c^*d \leq 0$.

We claim, however, that every feasible direction d for $\mathrm{HP}_e^{(i)}$ satisfies $c^*d \geq 0$. Indeed, otherwise the optimal objective value of $\mathrm{HP}_e^{(i)}$ would be unbounded and hence there would exist $\bar{x} \in \mathrm{relint}(\mathrm{Feas}_e^{(i)})$ satisfying $c^*\bar{x} < \limsup c^*x_j$ (using the assumption $\limsup c^*x_j > -\infty$). But $\bar{x} \in \mathrm{Feas}_{e_j}^{(i)}$ for sufficiently large j (because $0 < p_e^{(i)}(\bar{x}) = \lim p_{e_j}^{(i)}(\bar{x})$ for $k = i, \ldots, n-1$), contradicting that x_j is optimal for $\mathrm{HP}_{e_j}^{(i)}$ for all j.

Clearly, from the two preceding paragraphs, every limit point of $\{x_j/\|x_j\|\}$ is contained in $D := \{d \in \Lambda_{+,e}^{(i)} : Ad = 0 \text{ and } c^*d = 0\}$. To show $\{x_j/\|x_j\|\}$ has a unique limit point it thus suffices to show D consists of a single ray.

We claim $D \cap \Lambda_{++,e}^{(i)} = \emptyset$. Otherwise there would exist $d \in D$ satisfying $p_e^{(i)}(d) > 0$ for $k = i, \ldots, n-1$, which from continuity would imply d to be a feasible direction for $\mathrm{HP}_{e_j}^{(i)}$ for all $j > J$ (some J). But then $\mathrm{Opt}_{e_j}^{(i)}$ would be an unbounded set for $j > J$, implying by Theorem 2(B) that $\mathrm{Opt}_{e_j}^{(i)} = \mathrm{Opt}$ and thus contradicting boundedness of Opt. Hence, it is indeed the case that $D \cap \Lambda_{++,e}^{(i)} = \emptyset$.

It is clear now that the subspace D is contained in a boundary face of $\Lambda_{+,e}^{(i)}$. Hence, D consists of a single ray or is a subset of Λ_+ (according to a fact we stated in the first paragraph of this subsection). But if D were a subset of Λ_+ then each $d \in D$ would be a feasible direction for HP, which, with $c^*d = 0$, would imply Opt to be unbounded, a contradiction. Thus, D consists of a single ray, and so $\{x_j/\|x_j\|\}$ has a unique limit point d (and d satisfies $d \in \partial\Lambda_{+,e}^{(i)}$, $Ad = 0$, $c^*d = 0$).

Lastly, we prove $p_e^{(i-1)}(d) < 0$.

As $x_j \in \partial\Lambda_{+,e_j}^{(i)}$, we have $p_{e_j}^{(i-1)}(x_j) \leq 0$ (by (3.1.5)). Thus, by homogeneity and continuity, $p_e^{(i-1)}(d) \leq 0$. However, since $d \in \partial\Lambda_{+,e}^{(i)}$, if d satisfied $p_e^{(i-1)}(d) = 0$, then we would have $d \in \Lambda_{+,e}^{(i-1)}$, and hence $d \in \Lambda_+$ (by (3.9.1)) – so d would be a feasible direction for HP, and hence Opt would be unbounded (using $c^*d = 0$), a contradiction. Thus, $p_e^{(i-1)}(d) < 0$, completing the proof of the lemma. □

Lemma 3.9.2. *Let $e \in \Lambda_{++}$. Assume there exists a sequence $e_j \in \mathrm{Swath}(i)$ satisfying $e_j \to e$ and for which the set $\{x_j\}$ is unbounded for some choice of points $x_j \in \mathrm{Opt}_{e_j}^{(i)}$.*

(A) *There exists a constant $\alpha < 0$ and an open neighborhood U of e ($U \neq \emptyset$) such that for all $\bar{e} \in U \cap \mathrm{Swath}(i)$ and $\bar{x} \in \mathrm{Opt}_{\bar{e}}^{(i)}$,*

$$p_{\bar{e}}^{(i-1)}(\bar{x}) \leq \alpha \|\bar{x}\|^{n-i+1}.$$

(B) *For each value ℓ there exists an open neighborhood $V(\ell)$ of e ($V(\ell) \neq \emptyset$) such that for all $\bar{e} \in V(\ell) \cap \mathrm{Swath}(i)$ and $\bar{x} \in \mathrm{Opt}_{\bar{e}}^{(i)}$, it holds that $\|\bar{x}\| \geq \ell$.*

Proof. Parts (A) and (B) of Lemma 3.9.1 together imply that if for one sequence $\{e_j\} \subset \mathrm{Swath}(i)$ there exist $x_j \in \mathrm{Opt}_{e_j}^{(i)}$ for which $\{x_j\}$ is unbounded – where $e_j \to e$ – then for every sequence $\{e_j\} \subset \mathrm{Swath}(i)$ converging to e, and every choice of $x_j \in \mathrm{Opt}_{e_j}^{(i)}$, we have $\liminf \|x_j\| = \infty$. Part (B) of the present lemma easily follows.

For part (A), choose, say, $\alpha = \frac{1}{2}\, p_e^{(i-1)}(d)$, where d is the limit vector in part (C) of Lemma 3.9.1, noting that although d is defined there with respect to fixed sequences $\{e_j\}$ and $\{x_j\}$, the arbitrariness of those sequences – i.e., subject only to the conditions $e_j \to e$ and $x_j \in \mathrm{Opt}_{e_j}^{(i)}$ – and the uniqueness of d, together imply d to be independent of particular sequences $\{e_j\}$, $\{x_j\}$. Part (A) of the present lemma is then a consequence of the continuity of the function $(\bar{e}, \bar{x}) \mapsto p_{\bar{e}}^{(i-1)}(\bar{x})$, and the fact that $p_{\bar{e}}^{(i-1)}$ is homogeneous of degree $n - i + 1$ when $\bar{e} \in \Lambda_{++}$. □

In completing the proof of Proposition 3.8.1, let $\mathcal{L}$ be the set of limit points for the trajectory $t \mapsto e(t)$ as $t \to T$. We know (from the beginning of the proof) that $\mathcal{L} \subset \Lambda_{++}$. Thus, since the trajectory is maximal in $\mathrm{Swath}(i) \setminus \mathrm{Core}(i)$, if $e \in \mathcal{L}$ then either $e \in \mathrm{Core}(i)$ (that is, $\mathrm{Opt}_e^{(i)} = \mathrm{Opt}$) or $e \in \Lambda_{++} \setminus \mathrm{Swath}(i)$ (in which case $\mathrm{Opt}_e^{(i)} = \emptyset$).

Assume first that $\mathcal{L}$ consists of a single point e.

By Lemma 3.9.1(A), each limit point of the path $t \mapsto x(t)$ lies in $\mathrm{Opt}_e^{(i)}$. If the path is bounded, the set of limit points is nonempty, hence $\mathrm{Opt}_e^{(i)} \neq \emptyset$, and thus, from above, $e \in \mathrm{Core}(i)$. Clearly, then, if the path $t \mapsto x(t)$ is bounded, case (A) of the proposition holds.

On the other hand, if the path $t \mapsto x(t)$ is unbounded, Lemma 3.9.2 implies that case (B) of the proposition holds.

Thus, for the remainder of the proof assume $\mathcal{L}$ contains more than one point. We show case (B) of the proposition holds. For this it suffices to show $\mathcal{L}$ is compact, and that each $e \in \mathcal{L}$ satisfies the hypothesis of Lemma 3.9.2, as then $\mathcal{L}$ can be covered by finitely many of the open neighborhoods of the lemma, from which the establishment of case (B) easily follows.

That $\mathcal{L}$ is compact is trivial – indeed, the trajectory $t \mapsto e(t)$ is bounded.

Since $\mathcal{L}$ consists of more than one point, for each $e \in \mathcal{L}$ and each open neighborhood W of e, the Euclidean arc length of $\{e(t) : 0 \le t < T\} \cap W$ is infinite. Thus, since $T < \infty$, for each $e \in \mathcal{L}$ there exists an infinite sequence $t_1 < t_2 < \ldots$ satisfying $e(t_j) \to e$ and $\|\dot{e}(t_j)\| \to \infty$. But $\dot{e}(t_j) = x(t_j) - e(t_j)$, so $\|x(t_j)\| \to \infty$, and hence e satisfies the hypothesis of Lemma 3.9.2.

3.10. **Proof of Proposition 3.8.2.** Fix $1 \le i \le n-2$ and let $q_e := p_e^{(i)}/p_e^{(i+1)}$.

Recall the identity given in (3.8.2), that is,

$$D_e(D^j p_e^{(k)}(x)) = k D^{j+1} p_e^{(k-1)}(x) \ . \tag{3.10.1}$$

Lemma 3.10.1. *If x is any point satisfying $p_e^{(i)}(x) = 0 \neq p_e^{(i+1)}(x)$ then*

$$\begin{aligned} D_e(Dq_e(x))[x-e] = {} & Dq_e(x) \\ & + \frac{i(n-i)}{p_e^{(i+1)}(x)} Dp_e^{(i-1)}(x) \\ & - \frac{i(n-i+1)p_e^{(i-1)}(x)}{(p_e^{(i+1)}(x))^2} Dp_e^{(i+1)}(x) \ . \end{aligned}$$

Proof. For an arbitrary vector Δe and for any x at which q_e is defined (i.e., any x satisfying $p_e^{(i+1)}(x) \neq 0$), use of (3.10.1) easily gives

$$\begin{aligned} D_e(Dq_e(x))[\Delta e] = {} & \tfrac{i}{p_e^{(i+1)}(x)} D^2 p_e^{(i-1)}(x)[\Delta e] \\ & - \tfrac{i+1}{(p_e^{(i+1)}(x))^2} (Dp_e^{(i)}(x)[\Delta e]) Dp_e^{(i)}(x) \\ & - \tfrac{i}{(p_e^{(i+1)}(x))^2} (Dp_e^{(i-1)}(x)[\Delta e]) Dp_e^{(i+1)}(x) \\ & - \tfrac{2(i+1)p_e^{(i)}(x)}{(p_e^{(i+1)}(x))^3} (Dp_e^{(i)}(x)[\Delta e]) Dp_e^{(i+1)}(x) \\ & - \tfrac{(i+1)p_e^{(i)}(x)}{(p_e^{(i+1)}(x))^2} D^2 p_e^{(i)}(x)[\Delta e] \ . \end{aligned}$$

Substitute $\Delta e = x - e$, then use obvious identities to get rid of "$[e]$" (e.g., $Dp_e^{(i)}(x)[e] = p_e^{(i+1)}(x)$, $D^2 p_e^{(i-1)}(x)[e] = Dp_e^{(i)}(x)$), and use homogeneity to get rid of "$[x]$" – specifically, use (3.1.6). Finally, substitute $p_e^{(i)}(x) = 0$, and $\frac{1}{p_e^{(i+1)}(x)} Dp_e^{(i)}(x) = Dq_e(x)$ (because $p_e^{(i)}(x) = 0$), thereby concluding the proof. □

For $e \in \mathrm{Swath}(i) \setminus \mathrm{Core}(i)$, let $x_e := x_e^{(i)}$ and $\dot{x}_e := \frac{d}{dt} x_{e(t)}^{(i)}|_{t=0}$, where $t \mapsto e(t)$ is the trajectory satisfying $e(0) = e$.

Lemma 3.10.2. *If $e \in \mathrm{Swath}(i) \setminus \mathrm{Core}(i)$ there exists y satisfying*

$$\begin{aligned} & D^2 q_e(x_e)[\dot{x}_e] - \frac{c^* \dot{x}_e}{(c^*(e - x_e))^2} c^* \\ & = \frac{i p_e^{(i-1)}(x_e)}{p_e^{(i+1)}(x_e)} \left((n-i+1) \frac{D p_e^{(i+1)}(x_e)}{p_e^{(i+1)}(x_e)} - (n-i) \frac{D p_e^{(i-1)}(x_e)}{p_e^{(i-1)}(x_e)} \right) + A^* y \ . \end{aligned}$$

Proof. Corollary 8 shows $x_{e(t)}$ is optimal for the convex optimization problem

$$\begin{array}{rl} \min_x & -\ln c^*(e(t) - x) - q_{e(t)}(x) \\ \text{s.t.} & Ax = b \end{array}$$

and hence there exists $w(t)$ satisfying the first-order condition

$$\frac{1}{c^*(e(t) - x_{e(t)})} c^* - D q_{e(t)}(x_{e(t)}) = A^* w(t) \ . \tag{3.10.2}$$

Differentiating in t, then setting t equal to 0, gives

$$\frac{c^*(\dot{e} - \dot{x}_e)}{(c^*(x - e))^2} c^* + D^2 q_e(x_e)[\dot{x}_e] + D_e(D q_e(x_e))[\dot{e}] = A^* v \ ,$$

where $v := -\frac{d}{dt} w(t)|_{t=0}$. To complete the proof, substitute

$$\dot{e} = x_e - e \quad \text{and} \quad \frac{1}{c^*(e - x_e)} c^* = D q_e(x_e) + A^* w(0) \ \ (\text{by (3.10.2)}) \ ,$$

and then use Lemma 3.10.1 to substitute for $D_e(D q_e(x_e))[x_e - e]$. □

Now we are ready to prove the result to which this subsection is devoted, Proposition 3.8.2. Of course we need only consider the initial point $e = e(0)$ of the maximal trajectory $\{e(t) : 0 \leq t < T\}$, since any point $e(t)$ on the trajectory is itself the initial point for the maximal trajectory $\{\bar{e}(s) : 0 \leq s < T - t\}$ with $\bar{e}(s) = e(s + t)$.

Observe that

$$\begin{aligned} & \frac{d}{dt} \ln \frac{\left(-p_{e(t)}^{(i-1)}(x_{e(t)})\right)^{n-i}}{\left(p_{e(t)}^{(i+1)}(x_{e(t)})\right)^{n-i+1}} \Bigg|_{t=0} \\ & = (n-i) \ \tfrac{d}{dt} \ln \left(-p_{e(t)}^{(i-1)}(x_{e(t)})\right)\Big|_{t=0} - (n-i+1) \ \tfrac{d}{dt} \ln p_{e(t)}^{(i+1)}(x_{e(t)})\Big|_{t=0} \\ & = \tfrac{n-i}{p_e^{(i-1)}(x_e)} \left(D p_e^{(i-1)}(x_e)[\dot{x}] + D_e(p_e^{(i-1)}(x_e))[\dot{e}] \right) \\ & \qquad - \tfrac{n-i+1}{p_e^{(i+1)}(x_e)} \left(D p_e^{(i+1)}(x_e)[\dot{x}] + D_e(p_e^{(i+1)}(x_e))[\dot{e}] \right) \ . \end{aligned}$$

Using Lemma 3.10.2 to substitute for $\frac{n-i}{p_e^{(i-1)}(x_e)} Dp_e^{(i-1)}(x_e) - \frac{n-i+1}{p_e^{(i+1)}(x_e)} Dp_e^{(i+1)}(x_e)$ gives

$$\begin{aligned}
&\frac{d}{dt} \ln \frac{\left(-p_{e(t)}^{(i-1)}(x_{e(t)})\right)^{n-i}}{\left(p_{e(t)}^{(i+1)}(x_{e(t)})\right)^{n-i+1}}\Bigg|_{t=0} \\
&\quad = -\frac{p_e^{(i+1)}(x_e)}{ip_e^{(i-1)}(x_e)} \left(D^2 q_e(x_e)[\dot{x}_e, \dot{x}_e] - \left(\frac{c^* \dot{x}_e}{c^*(e-x_e)} \right)^2 \right) \\
&\qquad + (n-i) \frac{D_e(p_e^{(i-1)}(x_e))[\dot{e}]}{p_e^{(i-1)}(x_e)} - (n-i+1) \frac{D_e(p_e^{(i+1)}(x_e))[\dot{e}]}{p_e^{(i+1)}(x_e)} \\
&\quad \leq \frac{p_e^{(i+1)}(x_e)}{ip_e^{(i-1)}(x_e)} \left(\frac{c^* \dot{x}_e}{c^*(e-x_e)} \right)^2 + (n-i) \frac{D_e(p_e^{(i-1)}(x_e))[\dot{e}]}{p_e^{(i-1)}(x_e)} - (n-i+1) \frac{D_e(p_e^{(i+1)}(x_e))[\dot{e}]}{p_e^{(i+1)}(x_e)} ,
\end{aligned}$$

where the inequality is due to the combination of $p_e^{(i+1)}(x_e)/p_e^{(i-1)}(x_e)$ being negative (according to (3.8.1)) and $D^2 q_e(x_e)$ being negative semidefinite (Theorem 6).

To complete the proof, simply substitute

$$\begin{aligned}
D_e(p_e^{(i-1)}(x_e))[\dot{e}] &= (i-1) Dp_e^{(i-2)}(x_e)[x_e - e] \\
&= (i-1) \left((n-i+2) p_e^{(i-2)}(x_e) - p_e^{(i-1)}(x_e) \right) ,
\end{aligned}$$

$$\begin{aligned}
D_e(p_e^{(i+1)}(x_e))[\dot{e}] &= (i+1) Dp_e^{(i)}(x_e)[x_e - e] \\
&= (i+1) \left(0 - p_e^{(i+1)}(x_e) \right) ,
\end{aligned}$$

and

$$\begin{aligned}
\frac{c^* \dot{x}_e}{c^*(e-x_e)} &= Dq_e(x_e)[\dot{x}_e] \qquad \text{(by (3.10.2))} \\
&= \frac{Dp_e^{(i)}(x_e)[\dot{x}_e]}{p_e^{(i+1)}(x_e)} \qquad \text{(using } p_e^{(i)}(x_e) = 0\text{)} \\
&= -\frac{D_e(p_e^{(i)}(x_e))[\dot{e}]}{p_e^{(i+1)}(x_e)} \qquad \text{(because } \tfrac{d}{dt} p_{e(t)}^{(i)}(x_{e(t)}) = 0\text{)} \\
&= -i \, \frac{Dp_e^{(i-1)}(x_e)[x_e - e]}{p_e^{(i+1)}(x_e)} \\
&= -i \, \frac{(n-i+1) p_e^{(i-1)}(x_e) - 0}{p_e^{(i+1)}(x_e)} .
\end{aligned}$$

3.11. **Proof of Main Theorem (Part II).** In order to prove $(y_{e(t)}^{(i)}, s_{e(t)}^{(i)})$ is strictly feasible for HP* (not just feasible), we make use of the fact that, by Corollary 17 of [13], if $e \in \Lambda_{++}$ then every vector in $(\partial \Lambda_{+,e}^{(i)}) \setminus \Lambda_+$ is an exposed extreme direction of $\Lambda_{+,e}^{(i)}$. In particular, $(\partial \Lambda_{+,e}^{(i)}) \setminus \Lambda_+$ contains no line segments of positive length other than those lying in rays $\{sx : s > 0\}$.

Now we begin the proof of Part II of the Main Theorem.

Assume $e \in \text{Swath}(i) \setminus \text{Core}(i)$. By Corollary 8, $x_e^{(i)}$ is optimal for the convex optimization problem

$$\begin{array}{ll}
\min_x & -\ln c^*(e-x) - \frac{p_e^{(i)}(x)}{p_e^{(i+1)}(x)} \\
\text{s.t.} & Ax = b ,
\end{array}$$

and hence there exists $y = y_e^{(i)}$ satisfying the (rearranged) first-order condition

$$A^* y + s_e^{(i)} = c \ ,$$

where, letting $q_e := p_e^{(i)}/p_e^{(i+1)}$ and $x_e = x_e^{(i)}$,

$$s_e^{(i)} := (c^*(e - x_e))\, Dq_e(x_e) = \frac{c^*(e - x_e)}{p_e^{(i+1)}(x_e)} Dp_e^{(i)}(x_e),$$

(using $p_e^{(i)}(x_e) = 0$). Thus, to show the pair $(y_e^{(i)}, s_e^{(i)})$ is strictly feasible for HP*, it suffices to show $s_e^{(i)} \in \mathrm{int}(\Lambda_+^*)$, that is, assuming $z \in \Lambda_+$, $z \neq 0$, it suffices to show $Dq_e(x_e)[z] > 0$.

Since $(\partial \Lambda_{+,e}^{(i)}) \setminus \Lambda_+$ contains no line segments of positive length other than those lying in rays $\{sx : s > 0\}$, the line segment connecting x_e to z lies entirely within $\Lambda_{++,e}^{(i)}$ with the exception of the point x_e and possibly the point z. Thus, $z(s) := (1-s)z + sx$ satisfies $q_e(z(s)) > 0$ for $0 < s < 1$; fix such an s.

On the other hand, concavity of q_e (Theorem 6)) and $q_e(x_e) = 0$ imply $q_e(z(s)) \leq Dq_e(x_e)[z(s) - x_e]$, whereas $p_e^{(i)}(x_e) = 0$ and homogeneity of $p_e^{(i)}$ give $Dq_e(x_e)[x_e] = 0$. Thus,

$$0 < q(z(s)) \leq Dq_e(x_e)[z(s) - x_e] = (1-s) Dq_e(x_e)[z] \ ,$$

completing the proof that $(y_e^{(i)}, s_e^{(i)})$ is strictly feasible for HP*. Additionally observing

$$\begin{aligned} y_e^{(i)} b &= y_e^{(i)} A x_e \\ &= c^* x_e - s_e^{(i)} x_e \\ &= c^* x_e - (c^*(e - x_e))\, Dq_e(x_e)[x_e] \\ &= c^* x_e \ , \end{aligned}$$

we thus see for a maximal trajectory $\{e(t) : 0 \leq t < T\}$, each pair $(y_{e(t)}^{(i)}, s_{e(t)}^{(i)})$ is strictly feasible for HP*, and

$$\begin{aligned} y_{e(t)}^{(i)} b &= c^* x_{e(t)}^{(i)} \\ &\quad \to \mathrm{val} \quad \text{strictly monotonically} \quad \text{(by Part I(A) of the Main Theorem)} \ . \end{aligned}$$

It only remains to show the path $t \mapsto (y_{e(t)}^{(i)}, s_{e(t)}^{(i)})$ is bounded, for which it suffices to show $t \mapsto s_{e(t)}^{(i)}$ is bounded, because A is of full rank. In turn, because $e(0) \in \Lambda_{++}$ and $s_{e(t)}^{(i)} \in \Lambda_+^*$, it suffices to show the linear functional $s_{e(t)}^{(i)}$ applied to $e(0)$ is bounded from above independent of t. However,

$$\begin{aligned} s_{e(t)}^{(i)} e(0) &= c^* e(0) - (A^* y_{e(t)}^{(i)}) e(0) \\ &= c^* e(0) - y_{e(t)}^{(i)} b \\ &\qquad \to \ c^* e(0) - \mathrm{val} \ , \end{aligned}$$

concluding the proof.

3.12. **Proof of Theorem 4.** For $e \in \Lambda_{++}$ and an arbitrary point x, let $M_e(x)$ denote the number of non-negative roots of $t \mapsto p(x + te)$. Late in the proof we use the fact that $M_e(x)$ is independent of $e \in \Lambda_{++}$, as was established by Harvey and Lawson ([9], Theorem 2.12). (In [13], the fact was (essentially) established *only* when $x \in \Lambda_+$.)

Perhaps worth recording is that the independence can be readily verified with a most useful tool from the hyperbolic polynomial literature:

Helton-Vinnikov Theorem. *Assume $e \in \Lambda_{++}$. For any points x and z, there exist $n \times n$ symmetric matrices X and Z such that*

$$(r,s,t) \mapsto p(rx + sz + te) = p(e) \det(rX + sZ + tI) \; .$$

This formulation of the theorem comes from [11], where it was obtained by straightforward homogenization of the original result in [10]. (The initial importance of the homogeneous version was that it affirmatively settled the "Lax Conjecture" – see [11] for discussion.) (See [3] for recent negative results on possibilities of extensions to more than three variables r, s, t.)

To verify the independence of $M_e(x)$ from $e \in \Lambda_{++}$, first observe that because all roots of univariate polynomials vary continuously in the coefficients so long as the leading coefficient does not vanish, it suffices to show $\mathrm{mult}_e(x)$ is independent of $e \in \Lambda_{++}$, where $\mathrm{mult}_e(x)$ is the multiplicity of 0 as a root of $t \mapsto p(x+te)$. However, for $e, z \in \Lambda_{++}$, using the Helton-Vinnikov Theorem,

$$\mathrm{mult}_e(x) = n - \mathrm{rank}(X) = n - \mathrm{rank}(Z^{-1/2} X Z^{-1/2}) = \mathrm{mult}_z(x) \; ,$$

and thus the independence is verified.

Below we use the fact that

$$(\partial \Lambda^{(i)}_{+,e}) \cap \Lambda_+ \text{ is independent of } e \in \Lambda_+ \; . \tag{3.12.1}$$

This follows from Theorem 12 of [13] by straightforward induction.

Now we begin the proof of Theorem 4.

We claim it suffices to prove for $e \in \mathrm{Core}(i)$ and $x \in \mathrm{Opt}$ that the (infinite) line $\mathcal{L}$ through e and x satisfies $\mathcal{L} \cap \Lambda_{++} \subseteq \mathrm{Core}(i)$. Indeed, as Opt is convex, the relative interior of the convex hull $\mathcal{H}(e)$ for $\mathrm{Opt} \cup \{e\}$ will then clearly be a subset of $\mathrm{Core}(i)$, and hence, for any $\bar{x} \in \mathrm{Opt}$, so will be the set $\mathcal{K}(e, \bar{x})$ consisting of all points in Λ_{++} on the line through $\bar{x}$ and $\bar{e}$ for some $\bar{e} \in \mathrm{relint}(\mathcal{H}(e))$. However, if $\bar{x}$ is chosen from the relative interior of (the convex set) Opt, then $\mathcal{K}(e, \bar{x}) = \mathcal{A} \cap \Lambda_{++}$, where $\mathcal{A}$ is the smallest affine space containing e and Opt. The claim is thus established.

Fix $e \in \mathrm{Core}(i)$. Thus, $\mathrm{Core}(i)$ is nonempty, implying $1 \leq i \leq n-2$.[6] Let x denote a point in Opt ($= \mathrm{Opt}^{(i)}_e$).

Let $e(t) := (1-t)e + tx$, and let T denote the open interval consisting of t for which $e(t) \in \Lambda_{++}$. Always assume $t \in T$. We know from above that to prove the theorem, it suffices to show $e(t) \in \mathrm{Core}(i)$ for all $t \in T$.

Observe that a consequence of (3.12.1) and $\mathrm{Opt}^{(i)}_e = \mathrm{Opt}$ is

$$\mathrm{Opt} \subset \partial \Lambda^{(i)}_{+,e(t)} \text{ for all } t \in T \; . \tag{3.12.2}$$

Let $\mathcal{D} := \{d : Ad = 0 \text{ and } c^* d = 0\}$, and let $x + \mathcal{D}$ be the set consisting of all points $x + d$ where $d \in \mathcal{D}$.

Since $\mathrm{relint}(\mathrm{Feas}^{(i)}_{e(t)}) = (x + \{d : Ad = 0\}) \cap \Lambda^{(i)}_{++,e(t)}$ and $x \in \mathrm{Feas}^{(i)}_{e(t)}$ (because $x \in \mathrm{Opt}$), we have

$$x \in \mathrm{Opt}^{(i)}_{e(t)} \quad \Leftrightarrow \quad (x + \mathcal{D}) \cap \Lambda^{(i)}_{++,e(t)} = \emptyset \; .$$

---

[6]A quick way to see $\mathrm{Core}(n-1) = \emptyset$ is to observe that whereas Opt is bounded, $\mathrm{Opt}^{(n-1)}_e$ is either empty or unbounded, so it cannot possibly happen that $\mathrm{Opt}^{(n-1)}_e = \mathrm{Opt}$. This, however, is unsatisfying in that it leaves open the possibility Opt is a proper subset of $\mathrm{Opt}^{(n-1)}_e$, which we know cannot happen for other values of i (Theorem 2). A more fulfilling argument comes from: (i) $b \neq 0$; (ii) Λ_+ is regular (so, lineality space $= \{0\}$); and (iii) for all $e \in \Lambda_{++}$, $(\partial \Lambda^{(n-1)}_{+,e}) \cap \Lambda_+$ is precisely the lineality space of Λ_+ (by induction using Proposition 11 and Theorem 12 from [13]).

On the other hand, x is in $\mathrm{Opt}_{e(t)}^{(i)}$, as well as in Opt, iff $e(t) \in \mathrm{Core}(i)$ (by Theorem 2(B)). Consequently,

$$e(t) \in \mathrm{Core}(i) \quad \Leftrightarrow \quad (x+\mathcal{D}) \cap \Lambda_{++,e(t)}^{(i)} = \emptyset \, .$$

Thus, our goal – proving $e(t) \in \mathrm{Core}(i)$ for all $t \in T$ – will be accomplished if we prove $(x+\mathcal{D}) \cap \Lambda_{++,e(t)}^{(i)} = \emptyset$ for all $t \in T$. Hence, fixing arbitrary $d \in \mathcal{D}$, letting $\mathrm{ray}(d) := \{sd : s \geq 0\}$ and

$$T(d) := \{t \in T : (x + \mathrm{ray}(d)) \cap \Lambda_{++,e(t)}^{(i)} = \emptyset\} \, ,$$

our goal is to show $T(d) = T$.

Clearly, however, due to the characterization (3.1.2), $T \setminus T(d)$ is an open subset of the open interval T. Hence, since $T(d) \neq \emptyset$ (indeed, $0 \in T(d)$), to complete the proof of the theorem, it suffices to show that $T(d)$ is open. As e is an arbitrary element of $\mathrm{Core}(i)$ – in particular, e could be replaced with any $e(t)$ that happens to be in $\mathrm{Core}(i)$ – our goal has become:

$$\text{Show } T(d) \text{ contains an open interval including } 0. \tag{3.12.3}$$

For non-negative integers j and k, define

$$\alpha_{jk}(t) := D^k p_{e(t)}^{(j)}(x)[\underbrace{d, \ldots, d}_{k \text{ times}}]$$

and

$$\beta_j(t) := \begin{cases} 0 & \text{if } \alpha_{jk}(t) = 0 \text{ for all } k \, ; \\ \alpha_{j,k(j,t)} & \text{otherwise, where } k(j,t) := \min\{k : \alpha_{jk}(t) \neq 0\} \, . \end{cases}$$

Note that if $\beta_j(t) = 0$, then the polynomial $s \mapsto p_{e(t)}^{(j)}(x+sd)$ is identically zero, whereas if $\beta_j(t) \neq 0$, the polynomial evaluated at small, positive s has the same sign as $\beta_j(t)$. Thus, since x lies in the convex set $\Lambda_{+,e(t)}^{(i)}$, we have by the characterization (3.1.2) that

$$(x + \mathrm{ray}(d)) \cap \Lambda_{++,e(t)}^{(i)} \neq \emptyset \quad \Leftrightarrow \quad \beta_j(t) > 0 \text{ for all } j = i, \ldots, n-1 \, .$$

Our goal (3.12.3) has now been reduced to:

$$\text{Show } \exists \epsilon > 0 \text{ such that } |t| < \epsilon \Rightarrow \beta_j(t) \leq 0 \text{ for some } j \in \{i, \ldots, n-1\} \, . \tag{3.12.4}$$

For this we consider two cases.

First, assume $\beta_j(0) \geq 0$ for all $j = i, \ldots, n-1$. Then $x + sd \in \mathrm{Feas}_e^{(i)}$ for all sufficiently small, positive s. Thus, since $x \in \mathrm{Opt}_e^{(i)} = \mathrm{Opt}$ and $c^*d = 0$, we have $x + sd \in \mathrm{Opt}$ for all sufficiently small, positive s. Hence, by (3.12.2), for each $t \in T$, the univariate polynomial $s \mapsto p_{e(t)}^{(i)}(x+se)$ has value zero on an open interval, and thus is identically zero, so $\beta_i(t) = 0$ for all $t \in T$, (more than) accomplishing (3.12.4) for the first case.

Now consider the remaining case, that is, assume $\beta_j(0) < 0$ for some $j \in \{i, \ldots, n-1\}$; fix such a j. To accomplish (3.12.4), of course it suffices to show for this fixed value of j that $\beta_j(t) < 0$ for all t in an open interval containing 0. In turn, by definition of $\beta_j(t)$, it suffices to:

$$\text{Show } \exists \epsilon > 0 \text{ such that } |t| < \epsilon \quad \Rightarrow \quad D^{k(j,0)} p_{e(t)}^{(j)}(x)[\underbrace{\, d, \ldots, d \,}_{k(j,0) \text{ times}}] < 0 \tag{3.12.5}$$

and

$$\text{show } (t \in T) \wedge (k < k(j,0)) \quad \Rightarrow \quad D^k p_{e(t)}^{(j)}(x)[\underbrace{d, \ldots, d}_{k \text{ times}}] = 0 \, . \tag{3.12.6}$$

The existence of ϵ as in (3.12.5) is simply a matter of continuity and $\beta_j(0) < 0$. For accomplishing (3.12.6), observe

$$\begin{aligned} D^k p_{e(t)}^{(j)}(x) &= D^{k+j} p(x)[\underbrace{(1-t)e + tx, \ldots, (1-t)e + tx}_{j \text{ times}}] \\ &= \sum_{\ell=0}^{j} \binom{j}{\ell} (1-t)^{\ell} t^{j-\ell} D^{k+j-\ell} p_e^{(\ell)}(x)[\underbrace{x, \ldots, x}_{\ell - j \text{ times}}] \\ &= \sum_{\ell=0}^{j} \binom{j}{\ell} (1-t)^{\ell} t^{j-\ell} \, \frac{(n-k-\ell)!}{(n-k-j)!} D^k p_e^{(\ell)}(x) \quad (\text{using } (3.1.6)) \; . \end{aligned}$$

Consequently, (3.12.6) is immediately accomplished by the following proposition, thus concluding the proof of the theorem (except for proving the proposition).

Proposition 3.12.1. *Assume $e \in \Lambda_{++}$, $x \in \Lambda_+$ and let d be a vector. Assume non-negative integers j and k satisfy $D^k p_e^{(j)}(x)[\underbrace{d, \ldots, d}_{k \text{ times}}] \neq 0$; let $k(j)$ be the smallest such k for j and assume $k(j) > 0$.*

Then

$$D^k p_e^{(\ell)}(x)[\underbrace{d, \ldots, d}_{k \text{ times}}] = 0 \quad \textit{for all } \ell = 0, \ldots, j \textit{ and } k = 0, \ldots, k(j) - 1 \; .$$

The proof of the proposition makes use of the following lemma.

For arbitrary z, let $M(z)$ denote the number of non-negative roots of $t \mapsto p(z + te)$, a quantity that is independent of $e \in \Lambda_{++}$ (as we saw at the beginning of the subsection).

Lemma 3.12.2. *Assume $x \in \Lambda_{++}$, let $d \neq 0$, and let $s_1 \leq s_2 \leq \cdots \leq s_k$ denote the positive roots (including multiplicities) of $s \mapsto p(x + sd)$. Then*

$$M(x + sd) = \#\{i : s_i \leq s\} \; .$$

Proof. Because $M(x + sd)$ is independent of $e \in \Lambda_{++}$, in proving the lemma we may assume $e = x$, in which case we need only consider the hyperbolic polynomial obtained by restricting p to the subspace spanned by x and d. However, every hyperbolic polynomial in two variables is of the form $x \mapsto \prod a_i^T x$ for some vectors $a_i \in \mathbb{R}^2$ (this easily follows from two facts: (i) every complex homogeneous polynomial in two variables is of the form $x \mapsto \prod a_i^T x$, where $a_i \in \mathbb{C}^2$; (ii) if a_i is not a (complex) multiple of a real vector then $\{x \in \mathbb{R}^2 : a_i^T x = 0\}$ contains only the origin.) With this, verification of the lemma becomes entirely straightforward. $\square$

Proof of Proposition 3.12.1. Assume $e \in \Lambda_{++}$. For arbitrary z and $0 \leq \ell \leq n-1$, let $\mathrm{M}_e^{(\ell)}(z)$ denote the number of non-negative roots (counting multiplicities) for $t \mapsto p_e^{(\ell)}(x + te)$. Additionally, for a vector d and value $s \geq 0$, let $\mathrm{N}_e^{(\ell)}(x, d, s)$ denote the number of roots (counting multiplicities) in the interval $[-s, s]$ for the univariate polynomial $\bar{s} \mapsto p_e^{(\ell)}(x + \bar{s}d)$; if the univariate polynomial is identically zero, let $\mathrm{N}_e^{(\ell)}(x, d, s) = \infty$.

If $x \in \Lambda_{++}$ (hence $x \in \Lambda_{++,e}^{(\ell)}$ for all $0 \leq \ell \leq n-1$), $d \neq 0$ and $s \geq 0$, Lemma 3.12.2 can be applied with $p_e^{(\ell)}$ in place of p, and applied with $-d$ as well as with d, yielding

$$\mathrm{N}_e^{(\ell)}(x, d, s) = \mathrm{M}_e^{(\ell)}(x + sd) + \mathrm{M}_e^{(\ell)}(x - sd) \; . \tag{3.12.7}$$

On the other hand, for any z, the interlacing of the roots of $t \mapsto p_e^{(\ell)}(z + te)$ and its derivative $t \mapsto p_e^{(\ell+1)}(z + te)$ gives $\mathrm{M}_e^{(\ell)}(z) \geq \mathrm{M}_e^{(\ell+1)}(z)$, and thus,

$$\ell \leq j \quad \Rightarrow \quad \mathrm{M}_e^{(\ell)}(z) \geq \mathrm{M}_e^{(j)}(z) \; . \tag{3.12.8}$$

From (3.12.7) and (3.12.8) follows

$$(x \in \Lambda_{++}) \wedge (\ell \leq j) \quad \Rightarrow \quad \mathrm{N}_e^{(\ell)}(x,d,s) \geq \mathrm{N}_e^{(j)}(x,d,s) \;. \tag{3.12.9}$$

Assume, now, that x, d, j and $k(j)$ satisfy the hypothesis of the proposition. Observe that definitions readily give $k(j) = \mathrm{N}_e^{(j)}(x,d,0)$. Moreover, proving the proposition amounts precisely to showing

$$\mathrm{N}_e^{(\ell)}(x,d,0) \geq \mathrm{N}_e^{(j)}(x,d,0) \quad \text{for } \ell = 0,\ldots,j \;. \tag{3.12.10}$$

For $\epsilon > 0$, let $x(\epsilon) := x + \epsilon$, a point in Λ_{++}. Each polynomial $\bar{s} \mapsto p_e^{(\ell)}(x(\epsilon) + \bar{s}d)$ $(0 \leq \ell \leq n-1)$ is not identically zero and has only real roots (indeed, by homogeneity, the roots are the reciprocals of the non-zero roots for $t \mapsto p_e^{(\ell)}(d + tx(\epsilon))$). Thus, since

$$\begin{array}{c}\text{bounded roots of a univariate polynomial}\\ \text{vary continuously in the coefficients}\\ \text{(so long as the polynomial does not become identically zero)}\end{array} \tag{3.12.11}$$

we have for $0 \leq \ell \leq j$ that either $\mathrm{N}_e^{(\ell)}(x,d,0) = \infty$ (i.e., $\bar{s} \mapsto p_e^{(\ell)}(x + \bar{s}d) \equiv 0$), or

$$\begin{aligned}
\mathrm{N}_e^{(\ell)}(x,d,0) &= \lim_{s\downarrow 0}\left(\lim_{\epsilon\downarrow 0}\left(\mathrm{N}_e^{(\ell)}(x(\epsilon),d,s)\right)\right)\\
&\geq \lim_{s\downarrow 0}\left(\lim_{\epsilon\downarrow 0}\left(\mathrm{N}_e^{(j)}(x(\epsilon),d,s)\right)\right) \quad \text{(by (3.12.9))}\\
&\geq \lim_{s\downarrow 0}\left(\mathrm{N}_e^{(j)}(x,d,s/2)\right) \quad \text{(again using (3.12.11))}\\
&= \mathrm{N}_e^{(j)}(x,d,0)\;,
\end{aligned}$$

thereby establishing (3.12.10) and hence completing the proof. □

References

[1] H. H. Bauschke, O. Güler, A. S. Lewis and H. S. Sendov, Hyperbolic polynomials and convex analysis, *Canadian Journal of Mathematics* **53** (2001) no. 3, 470–488.

[2] J. Borcea and P. Brändén, Multivariate Pólya-Schur classification problems in the Weyl algebra, *Proceedings of the London Mathematical Society* (2010) advance access online.

[3] P. Brändén, Obstructions to determinantal representability, preprint available at arXiv.org.

[4] C.B. Chua, The primal-dual second-order cone approximations algorithm for symmetric cone programming, *Foundations of Computational Mathematics* **7** (2007), no. 3, 271-302.

[5] A.V. Fiacco and G.P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, SIAM, Philadelphia, PA, 1990.

[6] L. Gårding, An inequality for hyperbolic polynomials, *Journal of Mathematics and Mechanics* **8** (1959) no. 6, 957–965.

[7] O. Güler, Hyperbolic polynomials and interior point methods for convex programming, *Mathematics of Operations Research* **22** (1997), no. 2, 350–377.

[8] L. Gurvits, Van der Waerden/Schrijver-Valiant like conjectures and stable (aka hyperbolic) homogeneous polynomials: one theorem for all, *Electronic Journal of Combinatorics* **15** (2008) #R66.

[9] F.R. Harvey and H.B. Lawson Jr., Hyperbolic polynomials and the Dirichlet problem, preprint available at arXiv.org.

[10] J. Helton and V. Vinnikov, Linear matrix inequality representations of sets, *Communications of Pure and Applied Mathematics* 60 (2007) 654-674.

[11] A.S. Lewis, P.A. Parrilo, M.V. Ramana, The Lax conjecture is true, *Proceedings of the American Mathematical Society* **133** (2005) no. 9, 2495–2499.

[12] Y. Nesterov and A. Nemirovskii, *Interior-point polynomial algorithms in convex programming*, SIAM, Philadelphia, PA, 1994.

[13] J. Renegar, Hyperbolic programs, and their derivative relaxations, *Foundations of Computational Mathematics* **6** (2006) no. 1, 59-79.

[14] Y. Zinchenko, *The local behavior of the Shrink-Wrapping algorithm for linear programming*, Ph.D. thesis, Cornell University, 2005.

[15] Y. Zinchenko, On hyperbolicity cones associated with elementary symmetric polynomials, *Optimization Letters* **2** (2008) no. 3, 389-402.

[16] Y. Zinchenko, Shrink-wrapping trajectories for linear programming, preprint available at optimization-online.org.

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