Math 270: Geometry of Polynomials **Fall 2015**

Lecture 12: Hyperbolic Polynomials, Interlacers, and Sums of Squares

Lecturer: Nima Anari *Scribe: Joe Kileel*

12.1 Hyperbolic Polynomials

We begin with one of the main definitions of this lecture.

Definition 12.1. Let $f \in \mathbb{R}[x_1, \ldots, x_n]_d$ be a homogeneous polynomial of degree d in n variables with real coefficients, and let $e \in \mathbb{R}^n$. We say f is **hyperbolic** with respect to e if $f(e) \neq 0$ and for all $a \in \mathbb{R}^n$ the univariate polynomial $f(te - a) \in \mathbb{R}[t]$ has only real roots. Its **hyperbolicity cone** is:

$$C_e(f) := \{a \in \mathbb{R}^n : f(te - a) \neq 0 \text{ when } t \leq 0\}.$$

It is clear $C_e(f)$ is invariant under scaling by $\mathbb{R}_{>0}$ and open. Further, this cone is convex. Its closure is basic semialgebraic with all faces exposed, but we shall not elaborate on these points. The first examples of hyperbolic polynomials are familiar from optimization.

Example 12.2. Let $f = \prod_{i=1}^n x_i \in \mathbb{R}[x_1, \ldots, x_n]_n$ and $e = (1, \ldots, 1)$. Then f is hyperbolic with respect to e and the hyperbolicity cone $C_e(f)$ is the positive orthant $\mathbb{R}^n_{>0}$.

Example 12.3. Let $X = (x_{ij})$ be an $m \times m$ symmetric matrix of $n = \binom{m+1}{2}$ indeterminantes. Let $f = \det(X) \in \mathbb{R}[x_{ij}]_m$ and $e = I$ be the identity matrix. Then f is hyperbolic with respect to e and the hyperbolicity cone $C_e(f)$ is the set of positive definite $m \times m$ matrices. To prove this, note that here $f(te - a) = \det(tI - A)$ is the characteristic polynomial of the real symmetric matrix A, so has all real roots, namely the eigenvalues of A.

A hyperbolic program is the task of minimizing a real linear functional over an affine slice of a hyperbolicity cone $C_e(f)$. Thus hyperbolic programs include linear programs and semidefinite programs. Hyperbolic programs are a useful unifying framework in optimization because they can be efficiently solved by interior point methods applied to the barrier function $-\log(f)$. A natural question is whether or not hyperbolic programs are strictly more general than semidefinite programs. This remains a fundamental open problem.

Conjecture 12.4 (Generalized Lax Conjecture)**.** Let $f \in \mathbb{R}[x_1, \ldots, x_n]_d$ be hyperbolic with respect to $e \in \mathbb{R}^n$. Do there exist real symmetric matrices $A_1, \ldots, A_n$ of the same size such that:

$$C_e(f) = \{a \in \mathbb{R}^n : \sum_{i=1}^n a_i A_i \succ 0\} \,?$$

We now discuss various algebraic versions of this conjecture, in decreasing order of naiveté. Throughout $f \in \mathbb{R}[x_1, \dots, x_n]_d$ is hyperbolic with respect to e.

Algebraic Lax version 1: Do there exist $d \times d$ symmetric matrices $A_1, \dots, A_n \in \mathrm{Sym}_2(\mathbb{R}^d)$ such that $\sum_{i=1}^n e_i A_i \succ 0$ and $f = \det(\sum_{i=1}^n x_i A_i)$?

Answer: Generically no, because of the following dimension count. Nuij [8] showed the set of hyperbolic polynomials in $\mathbb{R}[x_1, \dots, x_n]_d$ has nonempty interior, and the ambient vector space has dimension $\binom{n+d-1}{d}$. On the other hand, the space $(\mathrm{Sym}_2(\mathbb{R}^d))^{\times n}$ has dimension $n\binom{d+1}{2}$. For fixed d, the first count $O(n^d)$ dominates the second count $O(n)$. An explicit counterexample is $f = x_1^2 - x_2^2 - \dots - x_n^2$ and $e = (1, 0, \dots, 0)$ when $n \geq 4$.

Next lecture concerns the Helton-Vinnikov theorem [4], which is that algebraic Lax version 1 holds true in the special case of $n = 3$ variables.

Algebraic Lax version 2: Do there exist a integer $k > 0$ and $dk \times dk$ symmetric matrices $A_1, \dots, A_n \in \mathrm{Sym}_2(\mathbb{R}^{dk})$ such that $\sum_{i=1}^n e_i A_i \succ 0$ and $f^k = \det(\sum_{i=1}^n x_i A_i)$?

Answer: Not always. Brändén [3] showed the basis generating function of the Vámos matroid is a counterexample. The Vámos matroid is rank 4 on 8 elements with 65 out of its 70 four element subsets independent. It is noted for not being realizable over any field, because it violates the Ingleton rank inequalities. Brändén's proof uses this fact. See the end of this lecture for another possible proof. For contrast, dropping the hyperbolicity and positive-definiteness conditions, note that it is classical that for all $p \in \mathbb{R}[x_1, \dots, x_n]_d$ there do exist $k > 0$ and $A_1, \dots, A_n \in \mathrm{Sym}_2(\mathbb{R}^{dk})$ such that $p^k = \det(\sum_{i=1}^n x_i A_i)$.

Algebraic Lax version 3: Do there exist a polynomial $g \in \mathbb{R}[x_1, \dots, x_n]_{d'}$ hyperbolic with respect to e and $(d + d') \times (d + d')$ symmetric matrices $A_1, \dots, A_n \in \mathrm{Sym}_2(\mathbb{R}^{d+d'})$ such that $C_e(f) \subseteq C_e(g)$, $\sum_{i=1}^n e_i A_i \succ 0$ and $fg = \det(\sum_{i=1}^n x_i A_i)$?

Answer: Open. This is equivalent to the Generalized Lax Conjecture. In [5], Kummer verifies that (a specialization of) the basis generating function of the Vámos matroid does not give a counterexample to algebraic Lax version 3.

Remark 12.5. It is interesting to consider a quantitative version of the Generalized Lax Conjecture: how large must the size of A_i, equivalently the size of the semidefinite program, be taken? See [9].

12.2 Interlacers

Most of the rest of this lecture will be spent relating the hyperbolicity cone $C_e(f)$ to other cones. This is made possible by considering polynomials that interlace f. We introduce this concept next.

Definition 12.6. For univariate polynomials $f, g \in \mathbb{R}[t]$ with all real roots and $\deg(g) = \deg(f) - 1$, let $\alpha_1 \leq \dots \leq \alpha_d$ and $\beta_1 \leq \dots \leq \beta_{d-1}$ be the roots of f and g respectively. We say that g **interlaces** f if $\alpha_i \leq \beta_i \leq \alpha_{i+1}$ for all $i = 1, \dots, d-1$.

Example 12.7. If $f \in \mathbb{R}[t]$ is real-rooted, then the derivative f' interlaces f.

Definition 12.8. For univariate polynomials $f, g \in \mathbb{R}[t]$, define the **Wronskian** by the formula $W[f,g] := f'g - fg'$.

Lemma 12.9. *For univariate polynomials $f, g \in \mathbb{R}[t]$ with deg$(g) =$ deg$(f) - 1$, g interlaces f if and only if $W[f,g]$ is a nonnegative polynomial or a nonpositive polynomial.*

Proof. By a limit argument, we may reduce to the case that the product fg has only simple roots. Now let $\alpha_1 < \ldots < \alpha_d$ and $\beta_1 < \ldots < \beta_{d-1}$ be the roots of f and g respectively.

The d polynomials $\frac{f}{t-\alpha_i}$ form a basis for the space of inhomogeneous polynomials in $\mathbb{R}[t]$ of degree $\leq d-1$. So we may (uniquely) write $g = \sum_{i=1}^{d} b_i \big(\frac{f}{t-\alpha_i}\big)$ for $b_i \in \mathbb{R}$. It is an exercise that g interlaces f if and only if b_i all have the same sign and are nonzero. Now compute:

$$\frac{W[f,g]}{f^2} = -\frac{d}{dt}\Big(\frac{g}{f}\Big) = \sum_{i=1}^{d} \frac{b_i}{(t-\alpha_i)^2}.$$

It follows that $W[f,g]$ has constant sign if and only if b_i all have the same sign. □

Definition 12.10 (Multivariate interlacing)**.** For multivariate homogeneous polynomials $f, g \in \mathbb{R}[x_1, \ldots, x_n]$ with deg$(g) =$ deg$(f) - 1$ both hyperbolic with respect to $e \in \mathbb{R}^n$, we say that g **interlaces** f with respect to e if for all $a \in \mathbb{R}^n$ the univariate restriction $g(te-a)$ interlaces $f(te-a)$.

Example 12.11. Let $f \in \mathbb{R}[x_1, \ldots, x_n]$ be hyperbolic with respect to $e \in \mathbb{R}^n$. Then the directional derivative $D_e(f) = \sum_{i=1}^{n} e_i \frac{\partial}{\partial x_i} f$ is hyperbolic with respect to e and interlaces f.

For the remainder of this section, $f \in \mathbb{R}[x_1, \ldots, x_n]_d$ is hyperbolic with respect to $e \in \mathbb{R}^n$. Without loss of generality, we assume $f(e) > 0$. Our aim here is to indicate the logical relationship between results worked out in detail in Section 2 of [6].

Definition 12.12. The **set of interlacers** of f with respect to e equals:

$$\text{Int}_e(f) := \{g \in \mathbb{R}[x_1, \ldots, x_n]_{d-1} : \ g \text{ interlaces } f, \ g(e) > 0\}.$$

Here $\text{Int}_e(f)$ is a closed, convex cone. The following two lemmas, and partial converses of them, allow the links between $C_e(f)$, interlacers and nonnegative polynomials.

Lemma 12.13. *For any g and h in $\text{Int}_e(f)$, the product $g.h$ is nonnegative on the real variety $\mathcal{V}_{\mathbb{R}}(f)$.*

Proof. This can be shown by univariate restriction. □

Lemma 12.14. *For $a \in C_e(f)$, the product $D_e(f).D_a(f)$ is nonnegative on $\mathcal{V}_{\mathbb{R}}(f)$.*

Proof. This can be shown by a homotopy argument using that $C_e(f)$ is convex. □

Our first theorem today comes next. It relates $C_e(f)$ to interlacers.

Theorem 12.15. *The closure of the hyperbolicity cone $\overline{C_e(f)} = \{a \in \mathbb{R}^n : \ D_a(f) \in \text{Int}_e(f)\}$. This expresses $\overline{C_e(f)}$ as the preimage under the linear map $a \mapsto D_a(f)$ of the cone $\text{Int}_e(f)$.*

Sketch. ($\subseteq$) Let $a \in C_e(f)$. By Lemma 12.14, $D_e(f).D_a(f)$ is nonnegative on $\mathcal{V}_{\mathbb{R}}(f)$. By Example 12.11, a partial converse of Lemma 12.13, and a limit argument, [6] proves $D_a(f) \in \text{Int}_e(f)$. $\square$

Our second theorem comes next. It relates $\text{Int}_e(f)$ to nonnegative polynomials.

Theorem 12.16. *The cone of interlacers of f with respect to e equals:*

$$\text{Int}_e(f) = \{g \in \mathbb{R}[x_1, \dots, x_n]_{d-1} : \ D_e(f)g - fD_e(g) \geq 0\}.$$

This expresses $\text{Int}_e(f)$ as the preimage under the linear map $g \mapsto D_e(f)g - fD_e(g)$ of the closed, conex cone of nonnegative polynomials inside $\mathbb{R}[x_1, \dots, x_n]_{2d-2}$.

Sketch. ($\subseteq$) By Lemma 12.9 and univariate restriction, it follows that $D_e(f)g - fD_e(g)$ is a nonnegative polynomial or a nonpositive polynomial on $\mathbb{R}^n$. By Example 12.11 and Lemma 12.13, the polynomial $D_e(f)g$ is nonnegative at $\mathcal{V}_{\mathbb{R}}(f)$. So, evaluating at roots of f shows $D_e(f)g - fD_e(g)$ is a nonnegative polynomial on $\mathbb{R}^n$. $\square$

Combining Theorems 12.15 and 12.16, we arrive at the output of this section.

Corollary 12.17. *Let $f \in \mathbb{R}[x_1, \dots, x_n]_d$ be hyperbolic with respect to $e \in \mathbb{R}^n$, and assume $f(e) > 0$. Then the cone $\overline{C_e(f)}$ equals the preimage under the linear map $a \mapsto D_e(f)D_a(f) - fD_e(D_a(f))$ of the cone of nonnegative polynomials inside $\mathbb{R}[x_1, \dots, x_n]_{2d-2}$.*

12.3 Sums of Squares

Nonnegative polynomials are well-studied not least because of the following problem from 1900.

Hilbert's 17^{th} problem: Let $h \in \mathbb{R}[x_1, \dots, x_n]_{2d'}$ be a nonnegative polynomial. Do there exist polynomials $h_1, \dots, h_t \in \mathbb{R}[x_1, \dots, x_n]_{d'}$ such that h equals the sum of squares:

$$h = \sum_{i=1}^{t} h_i^2 \ ?$$

The answer is negative, and Motzkin found the explicit counterexample $h(x, y, z) = z^6 + x^4z^2 + x^2y^4 - 3x^2y^2z^2$ in 1966. However, we have the following positive result from 1927 due to Artin.

Theorem 12.18. *Let $h \in \mathbb{R}[x_1, \dots, x_n]_{2d'}$ be a nonnegative polynomial. There exist rational functions $h_1, \dots, h_t \in \mathbb{R}(x_1, \dots, x_n)$ such that h equals the sum of squares:*

$$h = \sum_{i=1}^{t} h_i^2.$$

In 1995, Reznick [10] showed that, in the case h is a strictly positive polynomial, the rational functions h_i in Artin's theorem may be taken to have common denominator $(x_1 + \ldots + x_n)^N$ for some positive integer N on which he gives upper bounds.

In view of Corollary 12.17, we can make the following inner approximations to $\overline{C_e(f)}$, where f is hyperbolic with respect to $e \in \mathbb{R}^n$.

Definition 12.19. For each even nonnegative integer N, define the following closed convex cone:

$$C_e^{\text{SOS}(N)}(f) := \Big\{a \in \mathbb{R}^n : \Big(D_e(f)D_a(f) - fD_e(D_a(f))\Big)(x_1 + \ldots + x_n)^N \text{ is a SOS of real polynomials}\Big\}.$$

By the preceding discussion, we have inclusions:

$$C_e^{\text{SOS}(0)}(f) \subseteq C_e^{\text{SOS}(2)}(f) \subseteq C_e^{\text{SOS}(4)}(f) \subseteq \ldots\ldots \subseteq \overline{C_e(f)}.$$

It is a 2012 result of Netzer and Sanyal [7] that, if $\mathcal{V}_{\mathbb{R}}(f)$ is smooth, we also have the equality $\overline{C_e(f)} = \bigcup_N C_e^{\text{SOS}(N)}(f)$. Theoretically, this gives an algorithm to decide membership in $\overline{C_e(f)}$.

Remark 12.20. Given $h \in \mathbb{R}[x_1, \ldots, x_n]_{2d'}$. Let $\mathbf{v}$ be a column vector consisting of all $\binom{n+d'-1}{d'}$ monomials in $\mathbb{R}[x_1, \ldots, x_n]_{d'}$ in some order. Then there exists a positive semidefinite matrix A such that $h = \mathbf{v}^t A\mathbf{v}$ if and only if h is a sum-of-squares of real polynomials. In particular, for a given N, deciding membership in $C_e^{\text{SOS}(N)}(f)$ can be cast as deciding feasibility of a semidefinite program.

Proposition 12.21. *Let $f \in \mathbb{R}[x_1, \ldots, x_n]_d$ be a hyperbolic polynomial with respect to $e \in \mathbb{R}^n$ such that algebraic Lax version 1 holds true. In other words, there exist symmetric matrices $A_1, \ldots, A_n \in$ $\text{Sym}_2(\mathbb{R}^n)$ such that $\sum_{i=1}^n e_i A_i \succ 0$ and $f = \det(\sum_{i=1}^n x_i A_i)$. Then $\overline{C_e(f)} = C_e^{\text{SOS}(0)}(f)$.*

Proof. Write $M(x) := \sum_{i=1}^n x_i A_i$. Let $a \in \overline{C_e(f)}$ and compute:

$$D_e \det(M) = \det(M)\,\text{tr}\big((D_e M)M^{-1}\big)$$

$$D_a\, D_e \det(M) = \det^2(M)\Big(\text{tr}(D_a M)\, M^{-1}\, (D_e M)\, M^{-1}\Big) + \det^2(M)\,\text{tr}\Big((D_a M)M^{-1}\Big)\,\text{tr}\big((D_e M)M^{-1}\big).$$

So:

$$\begin{aligned} D_e(f)D_a(f) - fD_e(D_a(f)) &= \det^2(M)\,\Big(\text{tr}(D_a M)\, M^{-1}\, (D_e M)\, M^{-1}\Big) \\ &= \text{tr}\Big((D_a M)\, M^{\text{adj}}\, (D_e M)\, M^{\text{adj}}\Big). \end{aligned}$$

Here M^{adj} denotes the adjugate matrix of M consisting of $(n-1) \times (n-1)$ minors of M up to sign. Since $f = \det(M(x))$, the hyperbolicity cone $\overline{C_e(f)}$ equals the spectrahedron $\{b \in \mathbb{R}^n : \sum_{i=1}^n b_i A_i \succeq 0\}$. In particular, $\sum_{i=1}^n a_i A_i \succeq 0$, and so we may write $D_a M = \sum_{i=1}^n a_i A_i = \sum_i \mu_i \mu_i^t$ for some $\mu_i \in \mathbb{R}^n$. Similarly write $D_e M = \sum_{i=1}^n e_i A_i = \sum_j \lambda_j \lambda_j^t$ for some $\lambda_j \in \mathbb{R}^n$. The above trace equals:

$$\sum_{i,j} \text{tr}\Big(\mu\,\mu_i^t\, M^{\text{adj}}\, \lambda_j\, \lambda_j^t\, M^{\text{adj}}\Big) = \sum_{i,j} \text{tr}\Big((\lambda_j^t\, M^{\text{adj}}\, \mu)^t\, (\lambda_j^t\, M^{\text{adj}}\, \mu)\Big),$$

which is a sum-of-squares. The result follows from Corollary 12.17. ☐

It follows that if $f \in \mathbb{R}[x_1, \ldots, x_n]_d$ is a hyperbolic polynomial with respect to e such that algebraic Lax version 2 holds true, in other words a power f^k admits a definite determinantal representation, then $\overline{C_e(f)} = C_e^{\mathrm{SOS}(0)}(f)$ for some N, since $\overline{C_e(f)} = \overline{C_e(f^k)}$.

Example 12.22. A conceptually easier proof that the basis generating function of the Vámos matroid $f_{\text{Vám}} \in \mathbb{R}[x_1, \ldots, x_8]_4$ is a counterexample to algebraic Lax version 2 than Brändén's orginal proof might go as follows. By matroid theory [2], $f_{\text{Vám}}$ is real stable, that is, hyperbolic with hyperbolicity cone containing the positive orthant $\mathbb{R}^8_{>0}$. By Corollary 12.17 and Proposition 12.21, it suffices to check $D_{e_7}(f)D_{e_8}(f) - fD_{e_7}(D_{e_8}(f))$ is not a sum-of-squares, where $e_i \in \mathbb{R}^8_{>0}$ is the i^{th} standard basis vector. By Remark 12.20, this is equivalent to checking infeasibility of a 120×120 SDP with 1716 linear equations. Would this computation terminate in software like [11]?

References

[1] J. BOCHNAK, M. COSTE, M.-F. ROY, *Real algebraic geometry*, Springer, 1998

[2] P. BRÄNDÉN, *Polynomials with the half-plane property and matroid theory.* Adv. Math., **216(1)**, (2007) 302–320.

[3] P. BRÄNDÉN, *Obstructions to determinantal representability.* Adv. Math., **226(2)**, (2011) 1202–1212.

[4] J.W. HELTON, V. VINNIKOV, *Linear matrix reprsentation of sets.* Comm. Pure Appl. Math., **60(5)**, (1997) 350–377.

[5] M. KUMMER, *A note on the hyperbolicity cone of the specialized Vámos polynomial.* `arXiv:1306.4483v2`.

[6] M. KUMMER, D. PLAUMANN, C. VINZANT, *Hyperbolicity, interlacers and sums of squares.* Special issue of "Lifts of Convex Sets" Math. Prog. B, **153(1)**, (2015) 223–245.

[7] T. NETZER, R. SANYAL, *Smooth hyperbolicity cones are spectrahedral shadows.* Special issue "Lifts of Convex Sets in Optimization" Math. Prog. B, **153(1)**, (2015) 213–221.

[8] W. NUIJ, *A note on hyperbolic polynomials.* Math. Scand., **23**, (1969) 69–72.

[9] P. PARRILO, J. SAUNDERSON, *Polynomial-sized semidefinite reprsentations of erivative relaxations of spectahedral cones.* Math. Prog. A, **153(2)**, (2015), 309–331.

[10] B. REZNICK, *Uniform denominators in Hilbert's seventeenth problem.* Math. Z., **220(1)**, (1995) 75–97.

[11] J.F. STURM, *SeDuMi 1.02: a MATLAB toolbox for optimization over symmetric cones*. Optim. Methods Softw. **11(12)**, (1999) 625–653.

[12] D.G. WAGNER, *Multivariate stable polynomials: theory and applications.* Bull. Amer. Math. Soc., **48** (2011), 53–84.