

Lecture 2: Closure Properties and Applications

Lecturer: Shayan Oveis Gharan

August 24th

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications.*

We can continue this list and try to discover more and more closure properties. Borcea and Brändén [BB09a; BB09b] proved a remarkable result characterizing all linear operators that are stability preserving. Here, we don't discuss their theorem in full generality but we discuss one of the main applications of their theorem which is being used in almost all applications.

Let $T : \mathbb{R}[z_1, \dots, z_n] \rightarrow \mathbb{R}[z_1, \dots, z_n]$ be an (differential) operator on polynomials with real coefficients defined as follows:

$$\sum_{\alpha, \beta \in \mathbb{N}_{\geq 0}} c_{\alpha, \beta} z^\alpha \partial^\beta$$

For example, $1 - z_1 \partial z_2$. In the above $z^\alpha = \prod_{i=1}^n z_i^{\alpha_i}$ and $\partial^\beta = \prod_{i=1}^n \partial_{z_i}^{\beta_i}$.

Define $F_T \in \mathbb{R}[z_1, \dots, z_n, w_1, \dots, w_n]$

$$F_T(z_1, \dots, z_n, w_1, \dots, w_n) = \sum_{\alpha, \beta \in \mathbb{N}_{\geq 0}} c_{\alpha, \beta} z^\alpha (-w)^\beta.$$

Note that F_T is a polynomial with $2n$ variables.

Theorem 2.1. *T is an stability preserver operator, i.e., it maps any real stable polynomial to another real stable polynomial, iff F_T is real stable.*

For example, the operator $1 - \partial_{z_1}$ is stability preserver, because $1 + w_1$ is real stable. Also, $1 + z_2 \partial_{z_1}$ and $1 - \partial_{z_1} \partial_{z_2}$ are stability preserver.

For a non-example, $1 + \partial_{z_1} \partial_{z_2}$ and $1 - \partial_{z_1} \partial_{z_2} \partial_{z_3}$ are not stability preserver. This is because even though $z_1 z_2$ is real stable, $(1 + \partial_{z_1} \partial_{z_2})(z_1 z_2) = z_1 z_2 + 1$ is not.

As a direct consequence, we show that the Multi-Affine-Part (MAP) operator is stability preserving. Given a polynomial p , $\text{MAP}(p)$ zeros out every monomial of p with a square and keeps all multilinear monomials. For example,

$$\text{MAP}(1 + 2x + xy + x^2y + z^3) = 1 + 2x + xy.$$

Lemma 2.2. *MAP is a stability preserving operator.*

Proof. Given a polynomial $p \in \mathbb{R}[z_1, \dots, z_n]$. We can write MAP as the following differential operator

$$\text{MAP}(p) = \prod_{i=1}^n (1 + z_i \partial_{z_i})|_{z_1 = \dots = z_n = 0}.$$

Now, observe that $(1 + z_i \partial_{z_i})$ is stability preserving. Furthermore, setting all $z_i = 0$ is also stability preserving. □

Application to Roots of Multivariate Matching Polynomial. Given a graph $G = (V, E)$, our first application is to show that the matching polynomial

$$\mu_G(z_1, \dots, z_n) = \sum_M (-1)^{|M|} \prod_{i \text{ sat in } M} z_i \quad (2.1)$$

is real stable. First, observe that

$$p(z_1, \dots, z_n) = \prod_{\{i,j\} \in E} (1 - z_i z_j)$$

is real stable. This is because $1 - z_i z_j$ is real stable and product of real stable polynomials is also real stable. Now, observe that $\mu_G = \text{MAP}(p)$. So, μ_G is real stable.

As a consequence the univariate matching polynomial

$$\mu_G(t) = \sum_M (-1)^{|M|} t^{2|M|}$$

is real rooted. This follows by symmetrizing the polynomial in (2.1), i.e., set $z_i = t$ for all i and noting that any univariate real stable polynomial is real rooted.

Note that in some notes the matching polynomial is defined as $\sum_M (-1)^{|M|} t^{n-2|M|}$. This polynomial is also real rooted as it is the same as $t^n \mu_G(1/t)$ and real rooted polynomials are closed under inversion.

2.1 Gurvits' Machinery

Given a matrix $M \in \mathbb{R}_{\geq 0}^{n \times n}$ the permanent of M is defined as follows:

$$\text{per}(M) = \sum_{\sigma \in \mathbb{S}_n} \prod_{i=1}^n M_{i, \sigma_i}.$$

This quantity is #P-hard to compute exactly so one can only appeal to approximation algorithms. The problem has been studied for years. After much effort, Jerrum, Sinclair and Vigoda [JSV04] managed to design a randomized $1 \pm \epsilon$ approximation algorithm for this problem that runs in time polynomial in $n, 1/\epsilon$.

Designing a deterministic algorithm with the same approximation factor remains a *fundamental open problem*. This question relates to several fundamental problem in TCS under the umbrella of polynomial identity testing (PIT). It is believed that one can simulate any polynomial time randomized algorithm with a polynomial time deterministic algorithm. Permanent is perhaps the most fundamental problem which we have no deterministic approach for.

After a long line of works—see Linial, Samarodnitsky, Widgerson [LSW00], Barvinok [Bar97], Gurvits [GS02; Gur06], Gurvits, Samarodnitsky [GS14] and Anari, Rezaei [AR19]—the best deterministic approximation factor known is $(\sqrt{2})^n$. Note that if we can design a 2^{n^ϵ} approximation algorithm for any $\epsilon > 0$ then it is possible to improve the approximation factor to $1 \pm \epsilon$.

In this section, we will prove the following result of Gurvits:

Theorem 2.3 (Gurvits [Gur06]). *Let $p \in \mathbb{R}[z_1, \dots, z_n]$ be a real stable polynomial with non-negative coefficients. Then,*

$$e^{-n} \inf_{z > 0} \frac{p(z_1, \dots, z_n)}{z_1 \dots z_n} \leq \partial_{z_1} \dots \partial_{z_n} p|_{z=0} \leq \inf_{z > 0} \frac{p(z_1, \dots, z_n)}{z_1 \dots z_n}.$$

In other words, the above theorem shows that we can approximate the coefficient of the monomial $z_1 \dots z_n$ in the real stable polynomial p (with non-negative coefficient) with the mathematical program $\inf_{z > 0} \frac{p(z_1, \dots, z_n)}{z_1 \dots z_n}$.

Convexity. First we show that $\inf_{z>0} \frac{p(z_1, \dots, z_n)}{z_1 \dots z_n}$ can be formulated as a convex programming problem. The idea is to do a change of variables $e^{y_i} \leftrightarrow z_i$. Note that since $z_i > 0$ this is a valid change of variables. Furthermore to minimize the ratio we can equivalently minimize the log of the ratio. So, we claim,

$$\log \frac{p(e_1^y, \dots, e_n^y)}{e^{y_1} \dots e^{y_n}} = \log p(e_1^y, \dots, e_n^y) - \sum_{i=1}^n y_i$$

is convex. To see this, it is enough to show $\log p(e^{y_1}, \dots, e^{y_n})$ is convex, as the second term is linear. To see this, recall that log-sum-exp is a convex function i.e.,

$$\log \sum_{i=1}^n a_i e^{(b_i, y)}$$

is convex as long as $a_1, \dots, a_n \geq 0$ and $b_i \in \mathbb{R}^n$ for all i .

2.1.1 Applications

Now, we show how to use this theorem to get an e^n approximation for permanent. Consider the following polynomial

$$p(z_1, \dots, z_n) = \prod_{i=1}^n \sum_{j=1}^n M_{i,j} z_j. \quad (2.2)$$

First, observe that p is real stable since all entries of M are nonnegative and the product of real stable polynomials is real stable. Second, observe the coefficient of the monomial $z_1 \dots z_n$ is exactly $\text{per}(M)$, i.e.,

$$\partial_{z_1} \dots \partial_{z_n} p |_{z=0} = \text{per}(M).$$

Therefore, [Theorem 2.3](#) gives an e^n approximation algorithm to the $\text{per}(M)$.

Next, we use this theorem to prove the van-der-Waerden conjecture:

Theorem 2.4. Let $M \in \mathbb{R}_{\geq 0}^{n \times n}$ be a doubly stochastic matrix, i.e., the sum of the entries in every row and column is exactly 1. Then, $\text{per}(M) \geq e^{-n}$.

Proof. By [Theorem 2.3](#) we can write

$$\text{per}(M) \geq e^{-n} \inf_{z>0} \frac{p(z_1, \dots, z_n)}{z_1 \dots z_n}$$

where p is the same polynomial as in (2.2). So it is enough to show $\inf_{z>0} \frac{p(z_1, \dots, z_n)}{z_1 \dots z_n} \geq 1$. To see this, we use the weighted AM-GM inequality.

Theorem 2.5 (Weighted AM-GM Inequality). Let $a_1, \dots, a_n \geq 0$ and $\lambda_1, \dots, \lambda_n \geq 0$ such that $\sum_i \lambda_i = 1$. Then,

$$\sum_i \lambda_i a_i \geq \prod_i a_i^{\lambda_i}.$$

By the above theorem we can write

$$p(z_1, \dots, z_n) = \prod_{i=1}^n \sum_{j=1}^n M_{i,j} z_j \geq \prod_{i=1}^n \prod_{j=1}^n z_j^{M_{i,j}} = \prod_{j=1}^n z_j^{\sum_{i=1}^n M_{i,j}} = \prod_{j=1}^n z_j,$$

as desired. The inequality follows by (weighted) AM-GM and the fact that every row of M adds up to 1 and the last equality follows by the fact that every column of M adds up to 1. \square

2.1.2 Proof of Theorem

Before proving [Theorem 2.3](#) let us first gain some intuition. The right side of the inequality is trivial and it holds for any polynomial with non-negative coefficients. This is because if we let $p = \sum_{\alpha \in \mathbb{N}_{\geq 0}} c_{\alpha} z^{\alpha}$, then

$$\inf_{z>0} \frac{p(z_1, \dots, z_n)}{z_1 \dots z_n} = c_{1,1,\dots,1} + \inf_{z>0} \sum_{\alpha \in \mathbb{N}_{\geq 0}} c_{\alpha} \frac{z^{\alpha}}{z_1 \dots z_n} \geq c_{1,1,\dots,1},$$

as desired.

Now, the left side of the inequality crucially uses the fact that p is real stable. For example, consider the non-stable polynomial $z_1^2 + z_2^2$. Then,

$$\inf_{z>0} \frac{z_1^2 + z_2^2}{z_1 z_2} = \inf_{z>0} \frac{z_1}{z_2} + \frac{z_2}{z_1} \geq 2,$$

where as the coefficient of $z_1 z_2$ is 0.

We start the proof of [Theorem 2.3](#) by proving a univariate version of the theorem. This not only will serve as the base of the induction, but will also give us some clues on how to approach the main theorem.

Lemma 2.6. *For any real rooted polynomial $f \in \mathbb{R}[t]$ with nonnegative coefficients,*

$$f'(0) \geq \frac{1}{e} \inf_{t>0} f(t)/t.$$

Proof. First, we show that f is a log-concave function.

Fact 2.7. *Let $f \in \mathbb{R}[t]$ be a real rooted polynomial with nonnegative coefficients then $\log f$ is a concave function of $\mathbb{R}_{\geq 0}$.*

This is because (dropping the constant term for convenience):

$$\log f = \log \prod_{i=1}^n (t + \alpha_i) = \sum_{i=1}^n \log(t + \alpha_i),$$

where $-\alpha_i$'s are roots of f . Since f has all positive coefficients, it is non-zero over the positive reals so $\alpha_i \geq 0$ for all i . The claim follows because \log is a concave function and the sum of concave functions is concave.

Next, we prove the lemma. Since $\log f$ is concave, for any $t \geq 0$ we can write,

$$\log f(t) \leq \log f(0) + t(\log f(0))'$$

Therefore, for any $t \geq 0$ we get

$$\log \frac{f(t)}{t} \leq \log f(0) + t \frac{f'(0)}{f(0)} - \log t$$

Setting $t = f(0)/f'(0)$ in the RHS we get

$$\inf_{t>0} \log \frac{f(t)}{t} \leq \log f(0) + 1 - \log \frac{f(0)}{f'(0)} = 1 + \log f'(0)$$

as desired. □

Having proved the lemma we are ready to prove [Theorem 2.3](#). We prove by induction on n . Let $q(z_1, \dots, z_{n-1}) = \partial_{z_n} p|_{z_n=0}$. Note that by closure properties of real stable polynomials q is a real stable polynomial. Furthermore, it has non-negative coefficients. Then,

$$\partial_{z_1} \dots \partial_{z_n} p|_{z=0} = \partial_{z_1} \dots \partial_{z_{n-1}} q|_{z=0} \geq e^{-(n-1)} \inf_{z_1, \dots, z_{n-1} > 0} \frac{q(z_1, \dots, z_{n-1})}{z_1 \dots z_{n-1}} \quad (2.3)$$

where the inequality follows by IH. Say the infimum in the RHS is attained at a point z_1^*, \dots, z_{n-1}^* ¹. Define $f(z_n) = p(z_1^*, \dots, z_{n-1}^*, z_n)$. Then, observe that

$$q(z_1^*, \dots, z_{n-1}^*) = f'(0) \geq \frac{1}{e} \inf_{z_n > 0} \frac{f(z_n)}{z_n} = \frac{1}{e} \inf_{z_n > 0} \frac{p(z_1^*, \dots, z_{n-1}^*, z_n)}{z_n}, \quad (2.4)$$

where the inequality follows by the fact that f is a real stable polynomial with non-negative coefficients. Therefore,

$$\partial_{z_1} \dots \partial_{z_n} p|_{z=0} \geq e^{-(n-1)} \frac{q(z_1^*, \dots, z_{n-1}^*)}{z_1^* \dots z_{n-1}^*} \geq e^{-n} \inf_{z_n > 0} \frac{p(z_1^*, \dots, z_{n-1}^*, z_n)}{z_1^* \dots z_{n-1}^* z_n} \geq e^{-n} \inf_{z > 0} \frac{p(z_1, \dots, z_n)}{z_1 \dots z_n}$$

where the first inequality follows by (2.3) and the second inequality follows by (2.4). This finishes the proof of [Theorem 2.3](#).

Remark 2.8. *The above proof is fairly general; it works for any polynomial whose univariate restrictions are log-concave. Gurvits has studied many generalizations and extensions of the above proof such as using it to estimate mixed volume of convex bodies [Gur06].*

2.2 Log Concavity of Real Stable Polynomials

An immediate generalization of the [Lemma 2.6](#) is the following theorem:

Theorem 2.9. *Let $p \in \mathbb{R}[z_1, \dots, z_n]$ be a homogeneous real stable polynomial. Then, p is log-concave over $\mathbb{R}_{>0}^n$.*

This theorem can be seen as a generalization of the well-known fact that $\det(P)$ is log-concave over the space of PSD matrices.

Proof. In the HW we will see that all coefficients of p must be non-negative so $\log p$ is well defined over $\mathbb{R}_{>0}^n$. It is enough to show that $\log p$ is concave along any interval in the positive orthant. Let $a, b \in \mathbb{R}_{>0}^n$, $b \in \mathbb{R}^n$, and consider the line $a + tb$ where for any $t \in [0, 1]$, $a + tb \in \mathbb{R}_{>0}^n$. We show that $\log p(a + tb)$ is concave. Say p is k -homogeneous, then

$$p(a + tb) = p(t(a/t + b)) = t^k p(a/t + b).$$

Since $a \in \mathbb{R}_{>0}^n$, and $p(\cdot)$ is stable, $p(a/t + b)$ is real rooted. Write $p(a/t + b) = p(a) \prod_{i=1}^k (t - \lambda_i)$ where $\lambda_1, \dots, \lambda_k$ are the roots.

Then, we have

$$p(a/t + b) = p(a) \prod_{i=1}^k (1/t - \lambda_i).$$

¹to be more precise we need to do a ϵ - δ argument but here we avoid that for the simplicity of the argument.

So,

$$p(a + tb) = t^k p(a/t + b) = p(a) \prod_{i=1}^k (1 - t\lambda_i).$$

We claim that for all $1 \leq i \leq k$, $\lambda_i < 1$. Otherwise, for some $t \in [0, 1]$, $p(a + tb) = 0$, but since $a + tb \in \mathbb{R}_{>0}^n$, $p(a + tb) > 0$ which is a contradiction. Therefore,

$$\log p(a + tb) = \log p(a) + \sum_{i=1}^k \log(1 - t\lambda_i).$$

The theorem follows by the fact that $\log(1 - t\lambda)$ is a concave function of t for $t \in [0, 1]$ when $\lambda < 1$. \square

Remark 2.10. *If p is real stable but not homogeneous but it has positive coefficients, it is still log-concave over $\mathbb{R}_{\geq 0}^n$. It turns out that in this case one can homogenize p and apply the above theorem.*

References

- [AR19] N. Anari and A. Rezaei. “A Tight Analysis of Bethe Approximation for Permanent”. In: *FOCS*. Ed. by D. Zuckerman. IEEE Computer Society, 2019, pp. 1434–1445 (cit. on p. 2-2).
- [Bar97] A. I. Barvinok. “Computing Mixed Discriminants, Mixed Volumes, and Permanents”. In: *Discret. Comput. Geom.* 18.2 (1997), pp. 205–237 (cit. on p. 2-2).
- [BB09a] J. Borcea and P. Brändén. “The Lee-Yang and Pólya-Schur programs. I. Linear operators preserving stability”. In: *Inventiones mathematicae* 177.541 (2009) (cit. on p. 2-1).
- [BB09b] J. Borcea and P. Brändén. “The Lee-Yang and Pólya-Schur Programs. II. Theory of Stable Polynomials and Applications”. In: *Communications on Pure and Applied Mathematics* 62.12 (2009). QC 20100525, pp. 1595–1631 (cit. on p. 2-1).
- [GS02] L. Gurvits and A. Samorodnitsky. “A Deterministic Algorithm for Approximating the Mixed Discriminant and Mixed Volume, and a Combinatorial Corollary”. In: *Discret. Comput. Geom.* 27.4 (2002), pp. 531–550 (cit. on p. 2-2).
- [GS14] L. Gurvits and A. Samorodnitsky. “Bounds on the Permanent and Some Applications”. In: *FOCS*. 2014, pp. 90–99 (cit. on p. 2-2).
- [Gur06] L. Gurvits. “Hyperbolic polynomials approach to Van der Waerden/Schrijver-Valiant like conjectures: sharper bounds, simpler proofs and algorithmic applications”. In: *STOC*. Ed. by J. M. Kleinberg. ACM, 2006, pp. 417–426 (cit. on pp. 2-2, 2-5).
- [JSV04] M. Jerrum, A. Sinclair, and E. Vigoda. “A Polynomial-Time Approximation Algorithm for the Permanent of a Matrix with Nonnegative Entries”. In: *J. ACM* 51.4 (July 2004), pp. 671–697 (cit. on p. 2-2).
- [LSW00] N. Linial, A. Samorodnitsky, and A. Wigderson. “A Deterministic Strongly Polynomial Algorithm for Matrix Scaling and Approximate Permanents”. In: *Combinatorica* 20 (2000), pp. 545–568 (cit. on p. 2-2).