Journal of Advances in Science and Technology
Vol. 12, Issue No. 24, November-2016, ISSN 2230-9659

An Analysis upon Various Applications of Zeros of Complex Univariate and Multivariate Polynomials: A Review

Mithilesh Kumari Jain*

Research Scholar, OPJS University, Rajasthan

Abstract – Problems in many different areas of mathematics reduce to questions about the zeros of complex univariate and multivariate polynomials. Recently, several significant and seemingly unrelated results relevant to theoretical computer science have benefited from taking this route: they rely on showing, at some level, that a certain univariate or multivariate polynomial has no zeros in a region. This is achieved by inductively constructing the relevant polynomial via a sequence of operations which preserve the property of not having roots in the required region. The goal of this article is to present this viewpoint and to convey why the study of zeros is a natural, powerful, and versatile tool. It is meant to be a gentle introduction for the essential techniques underlying these results, is largely self-contained and aimed at a broad theory audience.

---------------------------♦-----------------------------

INTRODUCTION

Consider the following results relevant in theoretical computer science:

1. The permanent of an *n* x *n* stochastic matrix is at least $n!/n^n$. (This has been used to show that every d-regular graph on *n* vertices has a traveling salesman tour of length at most.)

2. The polynomial time approximation algorithm for the Traveling Salesman Problem on undi-rected. unweighted graphs with approximation ratio $3/2-\varepsilon$, for some constant $\varepsilon > 0$, USD.

3. The seminal result by Lee and Yang in statistical physics that shows the lack of phase transition in the Ising model, and the mean magnetization of the Ising model and the average size of a matching in the monomer-dimer model are both #P-hard to compute.

4. For every *d,* there is an infinite sequence of d-regular bipartite Ramanujan graphs, whose adjacency matrices have all nontrivial eigenvalues bounded by $2\sqrt{d-1}$.

Every transitive graph with *m* edges and *n* vertices can be partitioned into $O(m/n)$ edge- disjoint subgraphs of size $O(n)$, each of which approximates the cuts of the original graph up to a constant factor. This is a special case of a spectral discrepancy theorem about par-titioning sets of vectors in $\mathbb{C}^n$, which also resolves the Kadison-Singer problem in operator theory.

While the above problems and results seem unrelated, their solutions share a common thread: they all rely on showing, at some level, that a certain univariate or multivariate polynomial has no zeros in a region of $\mathbb{C}^n$ (e.g., the upper complex half-plane, or the unit disk). This is achieved by inductively constructing the relevant polynomial via a sequence of operations which *preserve* the property of not having roots in the required region.

For instance, when the coefficients of the polynomial are real and the region of no zeros is the upper complex-half plane, the polynomial is called real stable and this property is preserved under operations such as multiplication, taking derivatives and specialization to real values. While there are extensive and difficult characterizations of real stable polynomials. the above properties of real stable polynomials are rather simple to prove and, surprisingly, are sufficient for the applications listed above. Moreover, when the polynomials are of combinatorial origin, these operations have clear algebraic and combinatorial interpretations. Thus, there is a robust way to encode many kinds of combinatorial objects as polynomials, and to draw useful conclusions from their analytic properties. More generally, this serves as evidence against the stereotype that the roots of polynomials are brittle and ill-behaved (which is the case under unnatural operations such as perturbing the coefficients), and therefore difficult to exploit.

Roughly, these principles are evident in the applications listed above as follows: In this paper, closure of real stability under taking derivatives allows one to lower bound the value of the permanent of a doubly stochastic matrix. In this study, using the fact that the polynomials corresponding to the max-entropy probability distributions on spanning trees are real-stable, robust and novel negative correlation and anti-concentration properties of them are established. The result in EU relies on

a stability result for derivatives of the partition function of the Ising model and extends the famous Lee-Yang theorem, real stability allows one to relate the behavior of one polynomial to the behavior of a sum of polynomials leading to a new existence argument. Lastly, this theory allows the authors to control the evolution of roots of a polynomial under the application of differential operators.

One may argue that some of the applications above have alternative proofs that do not require this machinery. However, the fact remains that understanding the zeros of the relevant polynomials is important, and. in certain cases, has led to major progress in problems of interest. Moreover, with dramatic progress in the mathematics of this area, such techniques have recently reached a certain maturity which makes them ripe for applications. Thus, we feel that there is need to communicate the essential techniques underlying these results, in a largely self-contained manner, to a broad theory audience, and that is the goal of this article. For more in depth exposition of techniques.

MULTIVARIATE POLYNOMIALS

Recall that $f(z_1,\dots,z_n)$ is said to be real stable if $f\in\mathbb{R}[z_1,\dots,z_n]$ and no root of it lies in $\mathscr{H}^n$

It seems harder to show that a multivariate polynomial is real stable. The first lemma reduces the problem of checking real stability of a multivariate polynomial to checking real-rootedness of a set of univariate polynomials, and turns out to be quite effective.

Lemma. A multivariate polynomial $f(z_1,\dots,z_n)\in\mathbb{R}[z_1,\dots,z_n]$ is stable if and only if for all $v\in\mathbb{R}^n$ and all $u\in\mathbb{R}^n_{>0}$, the univariate polynomial $f(v+tu)$ is real-rooted.

Proof. Suppose that $f(v+tu)$ is real-rooted for all $v\in\mathbb{R}^n$ and all $u\in\mathbb{R}^n_{>0}$, but f is not real stable.

The latter implies that there is an $a=(a_1,\dots,a_n)\in\mathscr{H}^n$ such that $f(a)=0$. Let $v\stackrel{\text{def}}{=}\Re(a)$ and $u\stackrel{\text{def}}{=}\Im(a)$. Since $a\in\mathscr{H}^n, u_i>0$ for all i. But then $f(a)=f(v+\iota u)=0$ and, hence, ι is a root of $f(v+tu)$ which contradicts the real-rootedness of $f(v+tu)$

For the other direction, suppose that there are $v\in\mathbb{R}^n$ and $u\in\mathbb{R}^n_{>0}$ and a $t=t_1+\iota t_2$ suchthat $f(v+tu)=0$. Since complex roots of a univariate polynomial appear in conjugates, we may assume that $t_2>0$. Thus, the imaginary part of each component of $v+tu$ is strictly positive contradicting the fact that f is real stable.

Using the lemma above, several multivariate polynomials can be shown to be real stable. Perhaps the simplest such polynomial (which can be seen to be real stable without appealing to the lemma above) is $\sum_i a_i z_i$ when $a_i\geq 0$ for all i. Since a root of a polynomial that is a product of two polynomials is a root of at least one of those two polynomials, the polynomial $\prod_j\sum_i a_{ij}z_i$ is also real stable. A bit more non-trivially, the following important class of polynomials arising from determinants can be shown to be real stable.

Lemma. Let $A_1,\dots,A_n\in\mathbb{R}^{m\times m}$ be positive definite matrices0 and B be a symmetric m x m real matrix. Then the polynomial $f(z_1,\dots,z_n)\stackrel{\text{def}}{=}\det(z_1A_1+\dots+z_nA_n+B)$ is real stable.

Proof. By Lemma it is sufficient to prove that for all $v\in\mathbb{R}^n$ and $u\in\mathbb{R}^n_{>0}$, $f(v+tu)$ is real-rooted. This is the same as showing that

$$\det\left(B+\sum_{i=1}^n v_iA_i+t\sum_{i=1}^n u_iA_i\right)$$

is real-rooted. Since $A_i\succ 0$ and $u_i>0$ for all i, $\sum_{i=1}^n u_iA_i\succ 0$. Thus, letting $M\stackrel{\text{def}}{=}\sum_{i=1}^n u_iA_i$, we need to show that

$$\det\left(M^{-1/2}\left(B+\sum_{i=1}^n v_iA_i\right)M^{-1/2}+tI\right)$$

This latter is true because $M^{-1/2}(B+\sum_{i=1}^n v_iA_i)M^{-1/2}$ is symmetric and every real-symmetric has all real eigenvalues. To see this, if A is a real symmetric matrix and λ isan eigenvalue with an eigenvector v, then $Av=\lambda v$. 'Conjugating both sides we obtain that $v^\star A^\top=\bar{\lambda}v^\star$, where $v^\star$ is the conjugate transpose of v. Hence, $v^\star Av=\bar{\lambda}v^\star v$, since A is symmetric. Thus, $\lambda|v|^2=\bar{\lambda}|v|^2$ which implies that $\lambda=\bar{\lambda}$. Thus, $\lambda\in\mathbb{R}$.

The above lemma can be proved in the setting when A_is are positive semi-definite (PSD) as opposed to being positive-definite. This is quite useful for applications. However, extending Lemma requires the following theorem from complex analysis whose proof is beyond the scope of the current article.

Theorem (Hurwitz). Let $\{f_k\}_{k\geq 0}$ be a sequence of Ω -stable polynomials over $z_1,\ldots,z_n$ for a connected and open set Ω that uniformly converge to a polynomial f over compact subsets of Ω. Then f is Ω -stable.

To use this theorem for a matrix A_i which is just guaranteed to be PSD one approximates each A_i by a sequence of matrices $A_i+\frac{1}{2^k}I$ which are positive definite and converge to A_i ask goes to infinity.

One can ask if all real stable polynomials arise from such determinants. This is the content of the Lax Conjecture and the interested reader is referred to.

LOWER BOUNDING THE PERMANENT

As a simple but powerful application of the closure properties we show how, starting with simple polynomials, we can argue about non-trivial (and computationally intractable) objects such as the permanent of a matrix. For a matrix $A=(a_{ij})_{i\in[n],j\in[n]}$ with real entries, its permanent is defined to be $\mathrm{per}(A)\stackrel{\text{def}}{=}\sum_{\sigma\in S_n}\prod_{i=1}^{n}a_{i\sigma(i)}$.

Consider the polynomial $f_A(z_1,\ldots,z_n)\stackrel{\text{def}}{=}\prod_{i=1}^{n}\sum_{j=1}^{n}a_{ij}z_j$, and note that f_A is clearly real stable.

Moreover, it follows from a repeated application of Lemma that, for any $1\leq i<n$, the polynomial

$$g_i(z_1,\ldots,z_i)\stackrel{\text{def}}{=}\partial^{(i+1,\ldots,n)}f_A(z_1,\ldots,z_i,0,\ldots,0)$$

is real stable. Note that $g_0=\partial^{(1,\ldots,n)}f_A(0,\ldots,0)=\mathrm{per}(A)$. If all entries of A are nonnegative, then it follows from Lemma that, for any fixed positive $b_1,\ldots,b_{i-1}$,

$$g_{i-1}(b_1,\ldots,b_{i-1})=\partial_i g_i(b_1,\ldots,b_{i-1},0)\geq\left(\frac{d_i-1}{d_i}\right)^{d_i-1}\frac{g_i(b_1,\ldots,b_i)}{b_i},$$

Where d_i is the degree of the polynomial $g_i(b_1,\ldots,b_{i-1},z_i)$. Fixing $s_1,s_2,\ldots,s_{i-1}$, let s_i be defined to be $\arg\inf_{t>0}\frac{g_i(s_1,\ldots,s_{i-1},t)}{t}$.

Thus, applying the above inequality for $i=0$ to $n-1$ and letting $d\stackrel{\text{def}}{=}\max_{i=1}^{n}d_i$, we obtain that $\mathrm{per}(A)=g_0$, which is at least

$$\left(\frac{d-1}{d}\right)^{d-1}\frac{g_1(s_1)}{s_1}\geq\cdots\geq\left(\frac{d-1}{d}\right)^{(d-1)n}\frac{g_n(s_1,\ldots,s_n)}{\prod_{i=1}^{n}s_i}=\left(\frac{d-1}{d}\right)^{(d-1)n}\frac{f_A(s_1,\ldots,s_n)}{\prod_{i=1}^{n}s_i}.$$

Since $\frac{f_A(s_1,\ldots,s_n)}{\prod_{i=1}^{n}s_i}\geq\inf_{b_1>0,\ldots,b_n>0}\frac{f_A(b_1,\ldots,b_n)}{\prod_{i=1}^{n}b_i}$, we need to calculate $\inf_{b_1>0,\ldots,b_n>0}\frac{f_A(b_1,\ldots,b_n)}{\prod_{i=1}^{n}b_i}$. It turns out that when A is a doubly stochastic matrix, then this quantity can be lower bounded by 1. Recall that a matrix A is said to be doubly stochastic matrix, i.e., $a_{ij}\geq 0$ and, $\sum_{i=1}^{n}a_{ij}=1$ for all j and $\sum_{j=1}^{n}a_{ij}=1$ for all f. To see the claim, observe that for any positive $b_1,\ldots,b_n$,

$$f_A(b_1,\ldots,b_n)=\prod_{i=1}^{n}\sum_{j=1}^{n}a_{ij}b_j\overset{\text{AM-GM};\sum_j a_{ij}=1}{\geq}\prod_{i=1}^{n}\prod_{j=1}^{n}b_j^{a_{ij}}=\prod_{j=1}^{n}\prod_{i=1}^{n}b_j^{a_{ij}}\overset{\sum_i a_{ij}=1}{=}\prod_{j=1}^{n}b_j.$$

Thus, when A is doubly stochastic,. Noting that we have proved the van der Waerden conjecture. $\inf_{b_1>0,\ldots,b_n>0}\frac{f_A(b_1,\ldots,b_n)}{\prod_{i=1}^{n}b_i}\geq 1\left(\frac{d-1}{d}\right)^{d-1}\geq\frac{1}{e}$,

Theorem. For a $n\times n$ doubly stochastic matrix A, $\mathrm{per}(A)\geq\left(\frac{1}{e}\right)^n$

As a corollary, let $G=(V,W,E)$ be a k-regular bipartite graph with $|V|=|W|=n$. Let A be the matrix with rows indexed by V, columns by W, and $a_{ij}=1$ if $(i,j)\in E$. Then, $\frac{1}{k}A$ is doubly stochastic and, hence, $\mathrm{per}(A)\geq\left(\frac{k}{e}\right)^n$. Note that per(A) counts exactly the number of perfect matchings in G.

PROBABILITY MEASURES AND REAL STABILITY

In this section we study probability distributions over $\{0,1\}^n$ by looking at their generating function. For a distribution μ, the generating function is the multivariate affine polynomial

$$g_\mu\stackrel{\text{def}}{=}\sum_{S\subseteq[n]}\mu(S)\prod_{i\in S}z_i=\sum_{S\subseteq[n]}\mu(S)z^S.$$

If g_μ is real stable, then one can derive a host of properties of μ by appealing to the closure properties enjoyed by real stable polynomials. In this case, μ is said to be strongly Rayleigh.

Definition . A measure μ over $\{0,1\}^n$ is said to be strongly Rayleigh if its generating function $\sum_{S\subseteq[n]}\mu(S)z^S$ is real stable.

Strongly Rayleigh measures satisfy the strongest forms of negative dependence, a consequence of which is the concentration of measure for a sum of random variables drawn from a such a measure. As a starting point, we prove the pairwise negative correlation property of strongly Rayleigh measures.

Definition. A measure μ is said to be pairwise negatively correlated if $\sum_{S:S\supseteq\{i\}}\mu(S)\sum_{T:T\supseteq\{j\}}\mu(T)\geq\sum_{S:S\supseteq\{i,j\}}\mu(S)$ for all $i\neq j.$ In terms of polynomials, this condition is equivalent to

$$\partial_i g_\mu(1,1,\ldots,1)\partial_j g_\mu(1,1,\ldots,1)\geq g_\mu(1,1,\ldots,1)\partial^{(i,j)}g_\mu(1,1,\ldots,1).$$

In fact, for strongly Rayleigh measures, one can show something stronger: the condition holds for all $(a_1,\ldots,a_n)\in\mathbb{R}^n$ rather than just the vector $(1,\ldots,1)$. This property, in fact, implies the strong Rayleigh measures but we just prove the forward direction.

Lemma. If $f\in\mathbb{R}[z_1,\ldots,z_n]$ is affine and stable, then $\partial_i f(a_1,\ldots,a_n)\partial_j f(a_1,\ldots,a_n)\geq f(a_1,\ldots,a_n)\partial^{(i,j)}f(a_1,\ldots,a_n)$ for all $(a_1,\ldots,a_n)\in\mathbb{R}^n$ Thus, the lemma implies that strongly Rayleigh measures are pair wise negatively correlated.

Proof Fix $(a_1,\ldots,a_n)\in\mathbb{R}^n$ and let $g(s,t)\stackrel{\text{def}}{=}f(a_1,\ldots,a_i+s,\ldots,a_j+t,\ldots,a_n)$. It follows from

Lemma along with Theorem that $g(s,t)$ is stable. On the other hand, since f is affine,

$$g(s,t)=f(a_1,\ldots,a_n)+s\partial_i f(a_1,\ldots,a_n)+t\partial_j f(a_1,\ldots,a_n)+st\partial^{(i,j)}f(a_1,\ldots,a_n).$$

Since $g(s,t)$ is stable, for any $s=s_1+is_2$ such that $g(s,t)=0, s_2\leq 0.$ If $g(s,t)=0,$ then (dropping $(a_1,\ldots,a_n)$ for the easy of reading).

$$f+s_1\partial_i f-s_2\partial^{(i,j)}f=0$$

$$s_2\partial_i f+\partial_j f+s_1\partial^{(i,j)}f=0.$$

Multiplying the first equation by $-\partial^{(i,j)}f$ and the second by $\partial_i f$ and adding them, we obtain

$$-f\partial^{(i,j)}f+s_2(\partial^{(i,j)}f)^2+s_2(\partial_i f)^2+\partial_i f\partial_j f=0.$$

Since, $s_2\leq 0,$ this implies that $f\partial^{(i,j)}f\leq\partial_i f\partial_j f$ completing the proof.

REAL STABILITY AND INTERLACING

In this section we show how real stability can be used to show that a bound on the largest root of a sum of a certain family of polynomials implies the same bound on the largest root of one polynomial in the family. This novel technique is central to the proof of the existence of Ramanujan graphs of all degrees [OS], and the resolution of the Kadison-Singer problem.

More formally, the family of polynomials we consider is $\{f_\sigma(z)\}_{\sigma\in\{-1,1\}^n}$ where each polyno mial in the family is of the same degree and has a positive leading coefficient. For $p=(p_1,\ldots,p_n)\in[0,1]^n$, we define a random polynomial $f_{(x_1,\ldots,x_n)}$ where x_i is an independent Bernoulli random variable which is 1 with probability p_i and -1 with probability $1-p_i.$ Assume, the seemingly strong hypothesis, that this family polynomials satisfies, for every $p\in[0,1]^n$ the polynomial $\mathbb{E}_{(p_1,\ldots,p_n)}[f_{(x_1,\ldots,x_n)}]$ is real-rooted. Such a family is shown to have the property that if the largest root of $\Sigma_\sigma f_\sigma$ is bounded by ρ, then there is a σ such that the largest root of f_σ is also bounded by ρ. This is captured in the following theorem.

Theorem .Suppose $\{f_\sigma(z)\}_{\sigma\in\{-1,1\}^n}$ is a family of real-rooted polynomials with positive leading coefficients where all have the same degree. Then, there is a σ such that the largest root of $f_\sigma(z)$ is at most the largest root of $\Sigma_{\sigma\in\{-1,1\}^n}f_\sigma(z).$ While we do not go into the the proof of the hypothesis for any specific family in this article, we mention that the real-rootedness of $\mathbb{E}_{(p_1,\ldots,p_n)}[f_{(x_1,\ldots,x_n)}]$ is shown by constructing a suitable starting multivariate polynomial that is real stable and then applying a carefully chosen sequence of closure properties such as the ones presented in paper We start the proof of Theorem for a family which satisfies the above hypothesis by observing the following implication of the hypothesis.

Lemma . Under the hypothesis, for any fixing $\sigma_1,\ldots,\sigma_k$ any convex combination of

$\sum_{\sigma_{k+1},\ldots,\sigma_n}f_{\sigma_1,\ldots,\sigma_k,1,\sigma_{k+1},\ldots\sigma_n}$ and $\sum_{\sigma_{k+1},\ldots,\sigma_n}f_{\sigma_1,\ldots,\sigma_k,-1,\sigma_{k+1},\ldots\sigma_n}$ are real-rooted.

Proof. For a parameter $\lambda\in[0,1],$ set $p_{k+1}\stackrel{\text{def}}{=}\lambda$ and $p_{k+2}=\cdots=p_n=1/2$ and $p_i\stackrel{\text{def}}{=}\frac{1+\sigma_i}{2}$ for $1\leq i\leq$ k. It follows that
$\mathbb{E}_{(\sigma_1,\ldots,\sigma_n)\leftarrow\mu_p}[f_\sigma(z)]=\lambda\sum_{\sigma_{k+1},\ldots,\sigma_n}f_{\sigma_1,\sigma_2,\ldots,\sigma_k,1}+(1-\lambda)\sum_{\sigma_{k+1},\ldots,\sigma_n}f_{\sigma_1,\sigma_2,\ldots,\sigma_k,-1}$
which is real-rooted by the hypothesis.

CONCLUSION

The conclusion of the above lemma is interesting because if any convex combination of two univariate polynomials with leading positive coefficients is real-rooted, then they have a common interlacing. Two real-rooted polynomials $f(z)$ and $g(z)$ of the same degree (d) are said to interlace if their roots alternate. More formally, if $a_1\leq\cdots\leq a_d$ are roots of f and $b_1\leq\cdots\leq b_d$ are the roots of g, then

$$a_1\leq b_1\leq a_2\leq b_2\leq\cdots\leq a_d\leq b_d$$ or
$$b_1\leq a_1\leq b_2\leq a_2\leq\cdots\leq b_d\leq a_d.$$

Further, if there is a polynomial which interlaces with both $f(z)$ and $g(z)$, we say that they have a common interlacing. The following lemma can be proved by showing that, if one looks at the intervals corresponding to the successive roots of each polynomial and order them from left to right, the corresponding intervals have non-empty intersection. This is a consequence of the fact that two interlacing polynomials with positive leading coefficients cannot differ in the number of roots they have in any interval of the form $[a,\infty)$ by more than 1. We omit the elementary but somewhat tedious proof.

REFERENCES

Adam Marcus, Daniel A. Spielman, and Nikhil Srivastava (2013). Interlacing families I: Bipartite Ramanujan graphs of all degrees. arXiv preprint arXiv:1304.4132.

Adam Marcus, Daniel A. Spielman and Nikhil Srivastava (2013). Interlacing families II: Mixed characteristic polynomials and the Kadison-Singer problem. arXiv preprint arXiv:1306.3969.

Adrian Lewis, Pablo Parrilo, and Motakuri Ramana (2005). The lax conjecture is true. Proceedings of the American Mathematical Society, 133(9): pp. 2495–2499.

C. N. Yang and T. D. Lee (1952). Statistical theory of equations of state and phase transitions. I. Theory of condensation. Physical Review, 87(3): pp. 404–409.

Julius Borcea and Petter Br¨and´en (2009). The Lee-Yang and P´olya-Schur programs. I. linear operators preserving stability. Inventiones mathematicae, 177(3): pp. 541–569.

Julius Borcea and Petter Br¨and´en (2009). The Lee-Yang and P´olya-Schur programs. II. theory of stable polynomials and applications. Communications on Pure and Applied Mathematics, 62(12): pp. 1595–1631.

Julius Borcea, Petter Br¨and´en, and Thomas Liggett (2009). Negative dependence and the geometry of polynomials. Journal of the American Mathematical Society, 22(2): pp. 521 567.

Leonid Gurvits (2006). Hyperbolic polynomials approach to van der Waerden/Schrijver-valiant like conjectures: sharper bounds, simpler proofs and algorithmic applications. In Proceedings of the thirty-eighth annual ACM symposium on Theory of computing, pp. 417–426. ACM.

Maria Chudnovsky and Paul Seymour (2007). The roots of the independence polynomial of a clawfree graph. Journal of Combinatorial Theory, Series B, 97(3): pp. 350–357.

Robin Pemantle (2012). Hyperbolicity and stable polynomials in combinatorics and probability. arXiv preprint arXiv:1210.3231.

Shayan Oveis Gharan, Amin Saberi, and Mohit Singh (2011). A randomized rounding approach to the traveling salesman problem. In FOCS, pp. 550–559.

Steve Fisk (2006). Polynomials, roots, and interlacing. arXiv preprint math/0612833.

Tobias M¨omke and Ola Svensson (2011). Approximating graphic TSP by matchings. In FOCS, pages 560–569.

Tom´as Feder and Milena Mihail (1992). Balanced matroids. In Proceedings of the twenty-fourth annual ACM symposium on Theory of computing, pages 26–38. ACM.

Corresponding Author

Mithilesh Kumari Jain*

Research Scholar, OPJS University, Rajasthan

E-Mail – rohitkumarjangra1@gmail.com