

MULTIVARIATE STABLE POLYNOMIALS: THEORY AND APPLICATIONS

DAVID G. WAGNER

In memoriam Julius Borcea.

ABSTRACT. Univariate polynomials with only real roots – while special – do occur often enough that their properties can lead to interesting conclusions in diverse areas. Due mainly to the recent work of two young mathematicians, Julius Borcea and Petter Brändén, a very successful multivariate generalization of this method has been developed. The first part of this paper surveys some of the main results of this theory of “multivariate stable” polynomials – the most central of these results is the characterization of linear transformations preserving stability of polynomials. The second part presents various applications of this theory in complex analysis, matrix theory, probability and statistical mechanics, and combinatorics.

1. INTRODUCTION.

I have been asked by the AMS to survey the recent work of Julius Borcea and Petter Brändén on their multivariate generalization of the theory of univariate polynomials with only real roots, and its applications. It is exciting work – elementary but subtle, and with spectacular consequences. Borcea and Brändén take center stage but there are many other actors, many of whom I am unable to mention in this brief treatment. Notably, Leonid Gurvits provides a transparent proof of a vast generalization of the famous van der Waerden Conjecture.

Space is limited and I have been advised to use “Bourbaki style”, and so this is an account of the essentials of the theory and a few of its applications, with complete proofs as far as possible. Some relatively straightforward arguments have been left as exercises to engage the reader, and some more specialized topics are merely sketched or even omitted. For the full story and the history and context of the subject one must go to the references cited, the references they cite, and so on. The introduction of [4], in particular, gives a good account of the genesis of the theory.

Here is a brief summary of the contents. Section 2 introduces stable polynomials, gives some examples, presents their elementary properties, and develops multivariate generalizations of two classical univariate results: the Hermite-Kekeya-Obreschkoff and Hermite-Biehler Theorems. We also state the Pólya-Schur Theorem characterizing “multiplier sequences”, as this provides an inspiration for much of the multivariate theory. Section 3 restricts attention to multiaffine stable polynomials: we present a characterization of multiaffine real stable polynomials by means

2000 *Mathematics Subject Classification*. Primary: 32A60; Secondary: 05A20, 05B35, 15A45, 15A48, 60G55, 60K35.

Research supported by NSERC Discovery Grant OGP0105392.

of parameterized quadratic inequalities, and characterize those linear transformations which take multiaffine stable polynomials to stable polynomials. In Section 4 we use parts of the forgoing for Borcea and Brändén's splendid proof of the Grace-Walsh-Szegő Coincidence Theorem. In Section 5, the Grace-Walsh-Szegő Theorem is used to extend the results of Section 3 from multiaffine to arbitrary stable polynomials. This culminates in an amazing multivariate generalization of the Pólya-Schur Theorem, the proof of which requires the development of a multivariate extension of the Szasz Principle (which is omitted, regretfully, for lack of space). Section 6 presents Borcea and Brändén's resolution of some matrix-theoretic conjectures of Johnson. Section 7 presents the derivation by Borcea, Brändén, and Liggett of negative association inequalities for the symmetric exclusion process, a fundamental model in probability and statistical mechanics. Section 8 presents Gurvits's sweeping generalization of the van der Waerden Conjecture. Finally, Section 9 briefly mentions a few further topics that could not be included fully for lack of space.

I thank Petter Brändén kindly for his helpful comments on preliminary drafts of this paper.

2. STABLE POLYNOMIALS.

We use the following shorthand notation for multivariate polynomials. Let $[m] = \{1, 2, \dots, m\}$, let $\mathbf{x} = (x_1, \dots, x_m)$ be a sequence of indeterminates, and let $\mathbb{C}[\mathbf{x}]$ be the ring of complex polynomials in the indeterminates \mathbf{x} . For a function $\alpha : [m] \rightarrow \mathbb{N}$, let $\mathbf{x}^\alpha = x_1^{\alpha(1)} \cdots x_m^{\alpha(m)}$ be the corresponding monomial. For $S \subseteq [m]$ we also let $\mathbf{x}^S = \prod_{i \in S} x_i$. Similarly, for $i \in [m]$ let $\partial_i = \partial/\partial x_i$, let $\boldsymbol{\partial} = (\partial_1, \dots, \partial_m)$, let $\boldsymbol{\partial}^\alpha = \partial_1^{\alpha(1)} \cdots \partial_m^{\alpha(m)}$ and let $\boldsymbol{\partial}^S = \prod_{i \in S} \partial_i$. The constant functions on $[m]$ with images 0 or 1 are denoted by $\mathbf{0}$ and $\mathbf{1}$, respectively. The \mathbf{x} indeterminates are always indexed by $[m]$.

Let $\mathcal{H} = \{z \in \mathbb{C} : \text{Im}(z) > 0\}$ denote the open upper half of the complex plane, and $\overline{\mathcal{H}}$ the closure of \mathcal{H} in \mathbb{C} . A polynomial $f \in \mathbb{C}[\mathbf{x}]$ is *stable* provided that either $f \equiv 0$ identically, or whenever $\mathbf{z} = (z_1, \dots, z_m) \in \mathcal{H}^m$ then $f(\mathbf{z}) \neq 0$. We use $\mathfrak{S}[\mathbf{x}]$ to denote the set of stable polynomials in $\mathbb{C}[\mathbf{x}]$, and $\mathfrak{S}_{\mathbb{R}}[\mathbf{x}] = \mathfrak{S}[\mathbf{x}] \cap \mathbb{R}[\mathbf{x}]$ for the set of *real stable* polynomials in $\mathbb{R}[\mathbf{x}]$. (Borcea and Brändén do not consider the zero polynomial to be stable, but I find the above convention more convenient.)

We rely on the following essential fact at several points.

Hurwitz's Theorem (Theorem 1.3.8 of [14]). *Let $\Omega \subseteq \mathbb{C}^m$ be a connected open set, and let $(f_n : n \in \mathbb{N})$ be a sequence of functions, each analytic and nonvanishing on Ω , which converges to a limit f uniformly on compact subsets of Ω . Then f is either nonvanishing on Ω or identically zero.*

Consequently, a polynomial obtained as the limit of a convergent sequence of stable polynomials is itself stable.

2.1. Examples.

Proposition 2.1 (Proposition 2.4 of [1]). *For $i \in [m]$, let A_i be an n -by- n matrix and let x_i be an indeterminate, and let B be an n -by- n matrix. If A_i is positive semidefinite for all $i \in [m]$ and B is Hermitian then*

$$f(\mathbf{x}) = \det(x_1 A_1 + x_2 A_2 + \cdots + x_m A_m + B)$$

is real stable.

Proof. Let \bar{f} denote the coefficientwise complex conjugate of f . Since $\overline{A_i} = A_i^\top$ for all $i \in [m]$, and $\overline{B} = B^\top$, it follows that $\bar{f} = f$, so that $f \in \mathbb{R}[\mathbf{x}]$. By Hurwitz's Theorem and a routine perturbation argument, it suffices to prove that f is stable when each A_i is positive definite. Consider any $\mathbf{z} = \mathbf{a} + i\mathbf{b} \in \mathcal{H}^m$, with $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$ and $b_i > 0$ for all $i \in [m]$ (abbreviated to $\mathbf{b} > \mathbf{0}$). Now $Q = \sum_{i=1}^m b_i A_i$ is positive definite, and hence has a positive definite square-root $Q^{1/2}$. Also note that $H = \sum_{i=1}^m a_i A_i + B$ is Hermitian, and that

$$f(\mathbf{z}) = \det(Q) \det(iI + Q^{-1/2} H Q^{-1/2}).$$

Since $\det(Q) \neq 0$, if $f(\mathbf{z}) = 0$ then $-i$ is an eigenvalue of $Q^{-1/2} H Q^{-1/2}$, contradicting the fact that this matrix is Hermitian. Thus, $f(\mathbf{z}) \neq 0$ for all $\mathbf{z} \in \mathcal{H}^m$. That is, f is stable. \square

Corollary 2.2. *Let Q be an n -by- m complex matrix, and let $X = \text{diag}(x_1, \dots, x_m)$ be a diagonal matrix of indeterminates. Then $f(\mathbf{x}) = \det(QXQ^\dagger)$ is real stable.*

Proof. Let $Q = (q_{ij})$, and for $j \in [m]$ let A_j denote the n -by- n matrix with hi -th entry $q_{hj}\bar{q}_{ij}$. That is, $A_j = Q_j Q_j^\dagger$ in which Q_j denotes the j -th column of Q . Since each A_j is positive semidefinite and $QXQ^\dagger = x_1 A_1 + \dots + x_m A_m$, the conclusion follows directly from Proposition 2.1. \square

2.2. Elementary properties. The following simple observation often allows multivariate problems to be reduced to univariate ones, as will be seen.

Lemma 2.3. *A polynomial $f \in \mathbb{C}[\mathbf{x}]$ is stable if and only if for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$ with $\mathbf{b} > \mathbf{0}$, $f(\mathbf{a} + \mathbf{b}t)$ is stable in $\mathfrak{S}[t]$.*

Proof. Since $\mathcal{H}^m = \{\mathbf{a} + \mathbf{b}t : \mathbf{a}, \mathbf{b} \in \mathbb{R}^m, \mathbf{b} > \mathbf{0}, \text{ and } t \in \mathcal{H}\}$, the result follows. \square

For $f \in \mathbb{C}[\mathbf{x}]$ and $i \in [m]$, let $\deg_i(f)$ denote the degree of x_i in f .

Lemma 2.4. *These operations preserve stability of polynomials in $\mathbb{C}[\mathbf{x}]$.*

- (a) **Permutation:** for any permutation $\sigma : [m] \rightarrow [m]$, $f \mapsto f(x_{\sigma(1)}, \dots, x_{\sigma(m)})$.
- (b) **Scaling:** for $c \in \mathbb{C}$ and $\mathbf{a} \in \mathbb{R}^m$ with $\mathbf{a} > \mathbf{0}$, $f \mapsto cf(a_1 x_1, \dots, a_m x_m)$.
- (c) **Diagonalization:** for $\{i, j\} \subseteq [m]$, $f \mapsto f(\mathbf{x})|_{x_i=x_j}$.
- (d) **Specialization:** for $a \in \mathcal{H}$, $f \mapsto f(a, x_2, \dots, x_m)$.
- (e) **Inversion:** if $\deg_1(f) = d$, $f \mapsto x_1^d f(-x_1^{-1}, x_2, \dots, x_m)$.
- (f) **Differentiation (or "Contraction"):** $f \mapsto \partial_1 f(\mathbf{x})$.

Proof. Parts (a,b,c) are clear. Part (d) is also clear in the case that $\text{Im}(a) > 0$. For $a \in \mathbb{R}$ apply part (d) with values in the sequence $(a + i2^{-n} : n \in \mathbb{N})$, and then apply Hurwitz's Theorem to the limit as $n \rightarrow \infty$. Part (e) follows from the fact that \mathcal{H} is invariant under the operation $z \mapsto -z^{-1}$. For part (f), let $d = \deg_1(f)$, and consider the sequence $f_n = n^{-d} f(nx_1, x_2, \dots, x_m)$ for all $n \geq 1$. Each f_n is stable and the sequence converges to a polynomial, so the limit is stable. Since $\deg_1(f) = d$, this limit is not identically zero. This implies that for all $z_2, \dots, z_m \in \mathcal{H}$, the polynomial $g(x) = f(x, z_2, \dots, z_m) \in \mathbb{C}[x]$ has degree d . Clearly $g'(x) = \partial_1 f(x, z_2, \dots, z_m)$. Let ξ_1, \dots, ξ_d be the roots of $g(x)$, so that $g(x) = c \prod_{h=1}^d (x - \xi_h)$ for some $c \in \mathbb{C}$. Since f is stable, $\text{Im}(\xi_h) \leq 0$ for all $h \in [d]$. Now

$$\frac{g'(x)}{g(x)} = \frac{d}{dx} \log g(x) = \sum_{h=1}^d \frac{1}{x - \xi_h}.$$

If $\text{Im}(z) > 0$ then $\text{Im}(1/(z - \xi_h)) < 0$ for all $h \in [d]$, so that $g'(z) \neq 0$. Thus, if $\mathbf{z} \in \mathcal{H}^m$ then $\partial_1 f(\mathbf{z}) \neq 0$. That is, $\partial_1 f$ is stable. \square

Of course, by permutation, parts (d,e,f) of Lemma 2.4 apply for any index $i \in [m]$ as well (not just $i = 1$). Part (f) is essentially the Gauss-Lucas Theorem: the roots of $g'(x)$ lie in the convex hull of the roots of $g(x)$.

2.3. Univariate stable polynomials. A nonzero univariate polynomial is real stable if and only if it has only real roots. Let f and g be two such polynomials, let $\xi_1 \leq \xi_2 \leq \dots \leq \xi_k$ be the roots of f , and let $\theta_1 \leq \theta_2 \leq \dots \leq \theta_\ell$ be the roots of g . These roots are *interlaced* if they are ordered so that $\xi_1 \leq \theta_1 \leq \xi_2 \leq \theta_2 \leq \dots$ or $\theta_1 \leq \xi_1 \leq \theta_2 \leq \xi_2 \leq \dots$. For each $i \in [\ell]$, let $\hat{g}_i = g/(x - \theta_i)$. If $\deg f \leq \deg g$ and the roots of g are simple, then there is a unique $(a, b_1, \dots, b_\ell) \in \mathbb{R}^{\ell+1}$ such that

$$f = ag + b_1 \hat{g}_1 + \dots + b_\ell \hat{g}_\ell.$$

Exercise 2.5. Let $f, g \in \mathfrak{S}_{\mathbb{R}}[x]$ be nonzero and such that fg has only simple roots, let $\deg f \leq \deg g$, and let $\theta_1 < \dots < \theta_\ell$ be the roots of g . The following are equivalent:

- (a) The roots of f and g are interlaced.
- (b) The sequence $f(\theta_1), f(\theta_2), \dots, f(\theta_\ell)$ alternates in sign (strictly).
- (c) In $f = ag + \sum_{i=1}^{\ell} b_i \hat{g}_i$, all of b_1, \dots, b_ℓ have the same sign (and are nonzero).

The *Wronskian* of $f, g \in \mathbb{C}[x]$ is $W[f, g] = f' \cdot g - f \cdot g'$. If $f = ag + \sum_{i=1}^{\ell} b_i \hat{g}_i$ as in Exercise 2.5 then

$$\frac{W[f, g]}{g^2} = \frac{d}{dx} \left(\frac{f}{g} \right) = \sum_{i=1}^{\ell} \frac{-b_i}{(x - \theta_i)^2}.$$

It follows that if f and g are as in Exercise 2.5(a) then $W[f, g]$ is either positive for all real x , or negative for all real x . Since $W[g, f] = -W[f, g]$ the condition that $\deg f \leq \deg g$ is immaterial. Any pair f, g with interlacing roots can be approximated arbitrarily closely by such a pair with all roots of fg simple. It follows that for any pair f, g with interlacing roots, the Wronskian $W[f, g]$ is either nonnegative on all of \mathbb{R} or nonpositive on all of \mathbb{R} .

Nonzero univariate polynomials $f, g \in \mathfrak{S}_{\mathbb{R}}[x]$ are in *proper position*, denoted by $f \ll g$, if $W[f, g] \leq 0$ on all of \mathbb{R} . For convenience we also let $0 \ll f$ and $f \ll 0$ for any $f \in \mathfrak{S}_{\mathbb{R}}[x]$; in particular $0 \ll 0$.

Exercise 2.6. Let $f, g \in \mathfrak{S}_{\mathbb{R}}[x]$ be real stable. Then $f \ll g$ and $g \ll f$ if and only if $cf = dg$ for some $c, d \in \mathbb{R}$ not both zero.

Hermite-Kakeya-Obreschkoff (HKO) Theorem (Theorem 6.3.8 of [14]). *Let $f, g \in \mathbb{R}[x]$. Then $af + bg \in \mathfrak{S}_{\mathbb{R}}[x]$ for all $a, b \in \mathbb{R}$ if and only if $f, g \in \mathfrak{S}_{\mathbb{R}}[x]$ and either $f \ll g$ or $g \ll f$.*

Hermite-Biehler (HB) Theorem (Theorem 6.3.4 of [14]). *Let $f, g \in \mathbb{R}[x]$. Then $g + if \in \mathfrak{S}[x]$ if and only if $f, g \in \mathfrak{S}_{\mathbb{R}}[x]$ and $f \ll g$.*

Proofs of HKO and HB. It suffices to prove these when fg has only simple roots.

For HKO we can assume that $\deg(f) \leq \deg(g)$. Exercise 2.5 shows that if the roots of f and g are interlaced then for all $a, b \in \mathbb{R}$, the roots of g and $af + bg$

are interlaced, so that $af + bg$ is real stable. The converse is trivial if $cf = dg$ for some $c, d \in \mathbb{R}$ not both zero, so assume otherwise. From the hypothesis, both f and g are real stable. If there are $z_0, z_1 \in \mathcal{H}$ for which $\operatorname{Im}(f(z_0)/g(z_0)) < 0$ and $\operatorname{Im}(f(z_1)/g(z_1)) > 0$, then for some $\lambda \in [0, 1]$ the number $z_\lambda = (1-\lambda)z_0 + \lambda z_1$ is such that $\operatorname{Im}(f(z_\lambda)/g(z_\lambda)) = 0$. Thus $f(z_\lambda) - ag(z_\lambda) = 0$ for some real number $a \in \mathbb{R}$. Since $f - ag$ is stable (by hypothesis) and $z_\lambda \in \mathcal{H}$, this implies that $f - ag \equiv 0$, a contradiction. Thus $\operatorname{Im}(f(z)/g(z))$ does not change sign for $z \in \mathcal{H}$. This implies Exercise 2.5(c): all the b_i have the same sign (consider f/g at the points $\theta_i + i\epsilon$ for $\epsilon > 0$ approaching 0). Thus, the roots of f and g are interlaced.

For HB, let $p = g + if$. Considering $ip = -f + ig$ if necessary, we can assume that $\deg f \leq \deg g$. If $f \ll g$ then Exercise 2.5(c) implies that $\operatorname{Im}(f(z)/g(z)) \leq 0$ for all $z \in \mathcal{H}$, so that $g + if$ is stable. For the converse, let $p(x) = c \prod_{i=1}^d (x - \xi_i)$, so that $\operatorname{Im}(\xi_i) \leq 0$ for all $i \in [d]$. Now $|z - \xi_i| \geq |\bar{z} - \xi_i|$ for all $z \in \mathcal{H}$ and $i \in [d]$, so that $|p(z)| \geq |p(\bar{z})|$ for all $z \in \mathcal{H}$. For any $z \in \mathcal{H}$ with $f(z) \neq 0$ we have

$$\left| \frac{g(z)}{f(z)} + i \right| \geq \left| \frac{g(\bar{z})}{f(\bar{z})} + i \right| = \left| \frac{g(z)}{f(z)} - i \right|,$$

and it follows that $\operatorname{Im}(g(z)/f(z)) \geq 0$ for all $z \in \mathcal{H}$ with $f(z) \neq 0$. Since fg has simple roots it follows that $g(x) + yf(x)$ is stable in $\mathfrak{S}[x, y]$. By contraction and specialization, both f and g are real stable. By scaling and specialization, $af + bg$ is stable for all $a, b \in \mathbb{R}$. By HKO, the roots of f and g are interlaced. Since $\operatorname{Im}(f(z)/g(z)) \leq 0$ for all $z \in \mathcal{H}$, all the b_i in Exercise 2.5(c) are positive, so that $W[f, g]$ is negative on all of \mathbb{R} : that is $f \ll g$. \square

For $\lambda : \mathbb{N} \rightarrow \mathbb{R}$, let $T_\lambda : \mathbb{R}[x] \rightarrow \mathbb{R}[x]$ be the linear transformation defined by $T_\lambda(x^n) = \lambda(n)x^n$ and linear extension. A *multiplier sequence (of the first kind)* is such a λ for which $T_\lambda(f)$ is real stable whenever f is real stable. Pólya and Schur characterized multiplier sequences as follows.

Pólya-Schur Theorem (Theorem 1.7 of [4]). *Let $\lambda : \mathbb{N} \rightarrow \mathbb{R}$. The following are equivalent:*

- (a) λ is a multiplier sequence.
- (b) $F_\lambda(x) = \sum_{n=0}^{\infty} \lambda(n)x^n/n!$ is an entire function which is the limit, uniformly on compact sets, of real stable polynomials with all roots of the same sign.
- (c) Either $F_\lambda(x)$ or $F_\lambda(-x)$ has the form

$$Cx^n e^{ax} \prod_{j=1}^{\infty} (1 + \alpha_j x),$$

in which $C \in \mathbb{R}$, $n \in \mathbb{N}$, $a \geq 0$, all $\alpha_j \geq 0$, and $\sum_{j=1}^{\infty} \alpha_j$ is finite.

- (d) For all $n \in \mathbb{N}$, the polynomial $T_\lambda((1+x)^n)$ is real stable with all roots of the same sign.

One of the main results of Borcea and Brändén's theory is a great generalization of the Pólya-Schur Theorem – a characterization of all stability preservers: linear transformations $T : \mathbb{C}[\mathbf{x}] \rightarrow \mathbb{C}[\mathbf{x}]$ such that $T(f)$ is stable whenever f is stable. (Also the analogous characterization of real stability preservers.) This is discussed in some detail in Section 5.3.

2.4. Multivariate analogues of the HKO and HB Theorems. By analogy with the univariate HB Theorem, polynomials $f, g \in \mathbb{R}[\mathbf{x}]$ are said to be in *proper position*, denoted by $f \ll g$, when $g + if \in \mathfrak{S}[\mathbf{x}]$. (As will be seen, this implies that $f, g \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}]$.) Thus, the multivariate analogue of Hermite-Biehler is a definition, not a theorem.

Proposition 2.7 (Lemma 1.8 and Remark 1.3 of [5]). *Let $f, g \in \mathbb{C}[\mathbf{x}]$.*

- (a) *If $f, g \in \mathbb{R}[\mathbf{x}]$ then $f \ll g$ if and only if $g + yf \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}, y]$.*
 (b) *If $0 \neq f \in \mathfrak{S}[\mathbf{x}]$ then $g + yf \in \mathfrak{S}[\mathbf{x}, y]$ if and only if for all $\mathbf{z} \in \mathcal{H}^m$,*

$$\operatorname{Im} \left(\frac{g(\mathbf{z})}{f(\mathbf{z})} \right) \geq 0.$$

Proof. If $g + yf \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}, y]$ then $g + if \in \mathfrak{S}[\mathbf{x}]$, by specialization. Conversely, assume that $h = g + if \in \mathfrak{S}[\mathbf{x}]$ with $f, g \in \mathbb{R}[\mathbf{x}]$, and let $z = a + ib$ with $a, b \in \mathbb{R}$ and $b > 0$. By Lemma 2.3, for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$ with $\mathbf{b} > \mathbf{0}$ we have $h(\mathbf{a} + \mathbf{b}t) \in \mathfrak{S}[t]$. By HB, $\widehat{f}(t) = f(\mathbf{a} + \mathbf{b}t)$ and $\widehat{g}(t) = g(\mathbf{a} + \mathbf{b}t)$ are such that $\widehat{f} \ll \widehat{g}$. By HKO, $c\widehat{f} + d\widehat{g} \in \mathfrak{S}_{\mathbb{R}}[t]$ for all $c, d \in \mathbb{R}$. By HKO again, the roots of $b\widehat{f}$ and of $\widehat{g} + a\widehat{f}$ are interlaced. Since $W[b\widehat{f}, \widehat{g} + a\widehat{f}] = bW[\widehat{f}, \widehat{g}] \leq 0$ on \mathbb{R} , it follows that $b\widehat{f} \ll \widehat{g} + a\widehat{f}$. Finally, by HB again, $\widehat{g} + (a + ib)\widehat{f} \in \mathfrak{S}[t]$. Since this holds for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$ with $\mathbf{b} > \mathbf{0}$, Lemma 2.3 implies that $g + (a + ib)f \in \mathfrak{S}[\mathbf{x}]$. Since this holds for all $a, b \in \mathbb{R}$ with $b > 0$, $g + yf \in \mathfrak{S}[\mathbf{x}, y]$. This proves part (a).

For part (b), first let $g + yf$ be stable. By specialization, g is also stable. If $g \equiv 0$ then there is nothing to prove. Otherwise, consider any $\mathbf{z} \in \mathcal{H}^m$, so that $f(\mathbf{z}) \neq 0$ and $g(\mathbf{z}) \neq 0$. There is a unique solution $z \in \mathbb{C}$ to $g(\mathbf{z}) + zf(\mathbf{z}) = 0$, and since $g + yf$ is stable, $\operatorname{Im}(z) \leq 0$. Hence, $\operatorname{Im}(g(\mathbf{z})/f(\mathbf{z})) = \operatorname{Im}(-z) \geq 0$. This argument can be reversed to prove the converse implication. \square

Exercise 2.8 (Corollary 2.4 of [4]). $\mathfrak{S}[\mathbf{x}] = \{g + if : f, g \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}] \text{ and } f \ll g\}$.

Here is the multivariate HKO Theorem of Borcea and Brändén.

Theorem 2.9 (Theorem 1.6 of [4]). *Let $f, g \in \mathbb{R}[\mathbf{x}]$. Then $af + bg \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}]$ for all $a, b \in \mathbb{R}$ if and only if $f, g \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}]$ and either $f \ll g$ or $g \ll f$.*

Proof. First assume that $f \ll g$, and let $a, b \in \mathbb{R}$ with $b > 0$. By Proposition 2.7(a), $g + yf \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}, y]$. By scaling and specialization, $bg + (a + i)f \in \mathfrak{S}[\mathbf{x}]$. By Proposition 2.7(a) again, $f \ll (af + bg)$. Thus $af + bg \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}]$ for all $a, b \in \mathbb{R}$. The case that $g \ll f$ is similar.

Conversely, assume that $af + bg \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}]$ for all $a, b \in \mathbb{R}$. Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$ with $\mathbf{b} > \mathbf{0}$, and let $\widehat{f}(t) = f(\mathbf{a} + \mathbf{b}t)$ and $\widehat{g}(t) = g(\mathbf{a} + \mathbf{b}t)$. By Lemma 2.3, $a\widehat{f} + b\widehat{g} \in \mathfrak{S}_{\mathbb{R}}[t]$ for all $a, b \in \mathbb{R}$. By HKO, for each $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$ with $\mathbf{b} > \mathbf{0}$, either $\widehat{f} \ll \widehat{g}$ or $\widehat{g} \ll \widehat{f}$.

If $\widehat{f} \ll \widehat{g}$ for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$ with $\mathbf{b} > \mathbf{0}$, then by HB, $\widehat{g} + i\widehat{f} \in \mathfrak{S}[t]$ for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$ with $\mathbf{b} > \mathbf{0}$. Thus $g + if \in \mathfrak{S}[\mathbf{x}]$ by Lemma 2.3, which is to say that $f \ll g$ (by definition). Similarly, if $\widehat{g} \ll \widehat{f}$ for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$ with $\mathbf{b} > \mathbf{0}$ then $g \ll f$.

It remains to consider the case that $f(\mathbf{a}_0 + \mathbf{b}_0t) \ll g(\mathbf{a}_0 + \mathbf{b}_0t)$ for some $\mathbf{a}_0, \mathbf{b}_0 \in \mathbb{R}^m$ with $\mathbf{b}_0 > \mathbf{0}$, and $g(\mathbf{a}_1 + \mathbf{b}_1t) \ll f(\mathbf{a}_1 + \mathbf{b}_1t)$ for another $\mathbf{a}_1, \mathbf{b}_1 \in \mathbb{R}^m$ with $\mathbf{b}_1 > \mathbf{0}$. For $0 \leq \lambda \leq 1$, let $\mathbf{a}_\lambda = (1 - \lambda)\mathbf{a}_0 + \lambda\mathbf{a}_1$ and $\mathbf{b}_\lambda = (1 - \lambda)\mathbf{b}_0 + \lambda\mathbf{b}_1$. Since roots of polynomials move continuously as the coefficients are varied continuously, there is a value $0 \leq \lambda \leq 1$ for which both $f(\mathbf{a}_\lambda + \mathbf{b}_\lambda t) \ll g(\mathbf{a}_\lambda + \mathbf{b}_\lambda t)$ and $g(\mathbf{a}_\lambda + \mathbf{b}_\lambda t) \ll f(\mathbf{a}_\lambda + \mathbf{b}_\lambda t)$. From Exercise 2.6, it follows that $cf(\mathbf{a}_\lambda + \mathbf{b}_\lambda t) = dg(\mathbf{a}_\lambda + \mathbf{b}_\lambda t)$

for some $c, d \in \mathbb{R}$ not both zero. Now $h = cf - dg \in \mathfrak{S}[\mathbf{x}]$ by hypothesis, and since $h(\mathbf{a}_\lambda + \mathbf{b}_\lambda t) \equiv 0$ identically, it follows that $h(\mathbf{a}_\lambda + i\mathbf{b}_\lambda) = 0$. Since $\mathbf{b}_\lambda > \mathbf{0}$ and h is stable, this implies that $h \equiv 0$, so that $cf = dg$ in $\mathfrak{S}[\mathbf{x}]$. In this case, both $f \ll g$ and $g \ll f$ hold. \square

For $f, g \in \mathbb{C}[\mathbf{x}]$ and $i \in [m]$, let $W_i[f, g] = \partial_i f \cdot g - f \cdot \partial_i g$ be the i -th Wronskian of the pair (f, g) .

Corollary 2.10 (Theorem 1.9 of [5]). *Let $f, g \in \mathbb{R}[\mathbf{x}]$. The following are equivalent:*

- (a) $g + if$ is stable in $\mathfrak{S}[\mathbf{x}]$, that is $f \ll g$;
- (b) $g + yf$ is real stable in $\mathfrak{S}_{\mathbb{R}}[\mathbf{x}, y]$;
- (c) $af + bg \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}]$ for all $a, b \in \mathbb{R}$, and $W_i[f, g](\mathbf{a}) \leq 0$ for all $i \in [m]$ and $\mathbf{a} \in \mathbb{R}^m$.

Proof. Proposition 2.7(a) shows that (a) and (b) are equivalent.

If (a) holds then Theorem 2.9 implies that $af + bg \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}]$ for all $a, b \in \mathbb{R}$. To prove the rest of (c), let $i \in [m]$ and $\mathbf{a} \in \mathbb{R}^m$, and let $\delta_i \in \mathbb{R}^m$ be the unit vector with a one in the i -th position. Since $f \ll g$, for any $\mathbf{b} \in \mathbb{R}^m$ with $\mathbf{b} > \mathbf{0}$ we have $f(\mathbf{a} + (\mathbf{b} + \delta_i)t) \ll g(\mathbf{a} + (\mathbf{b} + \delta_i)t)$ in $\mathfrak{S}_{\mathbb{R}}[t]$, from Proposition 2.7(a) and Lemma 2.3. By the Wronskian condition for univariate polynomials in proper position,

$$W[f(\mathbf{a} + (\mathbf{b} + \delta_i)t), g(\mathbf{a} + (\mathbf{b} + \delta_i)t)] \leq 0$$

for all $t \in \mathbb{R}$. Taking the limit as $\mathbf{b} \rightarrow \mathbf{0}$ and evaluating at $t = 0$ yields

$$W_i[f, g](\mathbf{a}) = W[f(\mathbf{a} + \delta_i t), g(\mathbf{a} + \delta_i t)]|_{t=0} \leq 0,$$

by continuity. Thus (a) implies (c).

To prove that (c) implies (b), let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$ with $\mathbf{b} = (b_1, \dots, b_m) > \mathbf{0}$, and let $a, b \in \mathbb{R}$ with $b > 0$. By Lemma 2.3, to show that $g + yf \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}, y]$ it suffices to show that $g(\mathbf{a} + \mathbf{b}t) + (a + ib)f(\mathbf{a} + \mathbf{b}t) \in \mathfrak{S}[t]$. From (c) it follows that $p = g + af$ and $q = bf$ are such that $\alpha p + \beta q \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}]$ for all $\alpha, \beta \in \mathbb{R}$. By Theorem 2.9, either $p \ll q$ or $q \ll p$. Now

$$\begin{aligned} W[q(\mathbf{a} + \mathbf{b}t), p(\mathbf{a} + \mathbf{b}t)] &= bW[f(\mathbf{a} + \mathbf{b}t), g(\mathbf{a} + \mathbf{b}t)] \\ &= b \sum_{i=1}^m b_i W_i[f, g](\mathbf{a} + \mathbf{b}t) \leq 0, \end{aligned}$$

by the Wronskian condition in part (c). Thus $q(\mathbf{a} + \mathbf{b}t) \ll p(\mathbf{a} + \mathbf{b}t)$, so that $p(\mathbf{a} + \mathbf{b}t) + iq(\mathbf{a} + \mathbf{b}t) \in \mathfrak{S}[t]$. Since $p + iq = g + (a + ib)f$, this shows that (c) implies (b). \square

Exercise 2.11 (Corollary 1.10 of [5]). Let $f, g \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}]$ be real stable. Then $f \ll g$ and $g \ll f$ if and only if $cf = dg$ for some $c, d \in \mathbb{R}$ not both zero.

Proposition 2.12 (Lemma 3.2 of [5]). *Let V be a \mathbb{K} -vector subspace of $\mathbb{K}[\mathbf{x}]$, with either $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.*

- (a) *If $\mathbb{K} = \mathbb{R}$ and $V \subseteq \mathfrak{S}_{\mathbb{R}}[\mathbf{x}]$ then $\dim_{\mathbb{R}} V \leq 2$.*
- (b) *If $\mathbb{K} = \mathbb{C}$ and $V \subseteq \mathfrak{S}[\mathbf{x}]$ then $\dim_{\mathbb{C}} V \leq 1$.*

Proof. For part (a), suppose to the contrary that $f, g, h \in V$ are linearly independent over \mathbb{R} (and hence not identically zero). By Theorem 2.9, either $f \ll g$ or $g \ll f$, and similarly for the other pairs $\{f, h\}$ and $\{g, h\}$. Renaming these polynomials as necessary, we may assume that $f \ll h$ and $h \ll g$. Now, for all $\lambda \in [0, 1]$ let $p_\lambda = (1 - \lambda)f + \lambda g$, and note that each $p_\lambda \neq 0$. By Theorem 2.9, for each $\lambda \in [0, 1]$ either $h \ll p_\lambda$ or $p_\lambda \ll h$. Since $p_0 = f \ll h$ and $h \ll g = p_1$, by continuity of

the roots of $\{p_\lambda : \lambda \in [0, 1]\}$ there is a $\lambda \in [0, 1]$ such that $h \ll p_\lambda$ and $p_\lambda \ll h$. But then, by Exercise 2.11, either $\{f, g\}$ is linearly dependent or h is in the span of $\{f, g\}$, contradicting the supposition.

For part (b), let $\text{Re}(V) = \{\text{Re}(h) : h \in V\}$. Then $\text{Re}(V)$ is a real subspace of $\mathfrak{S}_{\mathbb{R}}[\mathbf{x}]$, so that $\dim_{\mathbb{R}} \text{Re}(V) \leq 2$ by part (a). If $\dim_{\mathbb{R}} \text{Re}(V) \leq 1$ then $\dim_{\mathbb{C}} V \leq 1$. In the remaining case let $\{p, q\}$ be a basis of $\text{Re}(V)$ with $f = p + iq \in V$. By Corollary 2.10, $W_i[q, p](\mathbf{a}) \leq 0$ for all $i \in [m]$ and $\mathbf{a} \in \mathbb{R}^m$. Since p and q are not linearly dependent, there is an index $k \in [m]$ such that $W_k[q, p] \neq 0$.

Consider any $g \in V$. There are reals $a, b, c, d \in \mathbb{R}$ such that

$$g = (ap + bq) + i(cp + dq).$$

Since g is stable, $W_k[cp + dq, ap + bq](\mathbf{a}) = (ad - bc)W_k[q, p](\mathbf{a}) \leq 0$ for all $\mathbf{a} \in \mathbb{R}^m$. Since $W_k[q, p] \neq 0$, it follows that $ad - bc \geq 0$. Now, for any $v, w \in \mathbb{R}$, $g + (v + iw)f$ is in V . Since

$$g + (v + iw)f = (a + v)p + (b - w)q + i((c + w)p + (d + v)q),$$

this argument shows that $H = (a + v)(d + v) - (b - w)(c + w) \geq 0$ for all $v, w \in \mathbb{R}$. But

$$4H = (2v + a + d)^2 + (2w + c - b)^2 - (a - d)^2 - (b + c)^2,$$

so that $H \geq 0$ for all $v, w \in \mathbb{R}$ if and only if $a = d$ and $b = -c$. This implies that $g = (a + ic)f$, so that $\dim_{\mathbb{C}} V = 1$. \square

3. MULTIAFFINE STABLE POLYNOMIALS.

A polynomial f is *multiaffine* if each indeterminate occurs at most to the first power in f . For a set \mathcal{S} of polynomials, let \mathcal{S}^{MA} denote the set of multiaffine polynomials in \mathcal{S} . For multiaffine $f \in \mathbb{C}[\mathbf{x}]^{\text{MA}}$ and $i \in [m]$ we use the ‘‘ultrashorthand’’ notation $f = f^i + x_i f_i$ in which $f^i = f|_{x_i=0}$ and $f_i = \partial_i f$. This notation is extended to multiple distinct indices in the obvious way – in particular,

$$f = f^{ij} + x_i f_i^j + x_j f_j^i + x_i x_j f_{ij}.$$

3.1. A criterion for real stability. For $f \in \mathbb{C}[\mathbf{x}]$ and $\{i, j\} \subseteq [m]$, let

$$\Delta_{ij} f = \partial_i f \cdot \partial_j f - f \cdot \partial_i \partial_j f.$$

Notice that for $f \in \mathbb{C}[\mathbf{x}]^{\text{MA}}$,

$$\Delta_{ij} f = f_i^j f_j - f^j f_{ij} = W_i[f^j, f_j] = -W_i[f_j, f^j],$$

and

$$\Delta_{ij} f = f_i^j f_j^i - f^{ij} f_{ij}.$$

Theorem 3.1 (Theorem 5.6 of [8] and Theorem 3 of [16]). *Let $f \in \mathbb{R}[\mathbf{x}]^{\text{MA}}$ be multiaffine. The following are equivalent:*

- (a) f is real stable.
- (b) For all $\{i, j\} \subseteq [m]$ and all $\mathbf{a} \in \mathbb{R}^m$, $\Delta_{ij} f(\mathbf{a}) \geq 0$.
- (c) Either $m = 1$, or there exists $\{i, j\} \subseteq [m]$ such that f_i, f^i, f_j and f^j are real stable, and $\Delta_{ij} f(\mathbf{a}) \geq 0$ for all $\mathbf{a} \in \mathbb{R}^m$.

Proof. To see that (a) implies (b), fix $\{i, j\} \subseteq [m]$. Proposition 2.7(a) shows that $f_j \ll f^j$, and from the calculation above and Corollary 2.10, it follows that $\Delta_{ij}f(\mathbf{a}) = -W_i[f_j, f^j](\mathbf{a}) \geq 0$ for all $\mathbf{a} \in \mathbb{R}^m$.

We show that (b) implies (a) by induction on m , the base case $m = 1$ being trivial. For the induction step let f be as in part (b), let $a \in \mathbb{R}$, and let $g = f|_{x_m=a}$. For all $\{i, j\} \subseteq [m-1]$ and $\mathbf{a} \in \mathbb{R}^{m-1}$, $\Delta_{ij}g(\mathbf{a}) = \Delta_{ij}f(\mathbf{a}, a) \geq 0$. By induction, $g = f^m + af_m$ is real stable for all $a \in \mathbb{R}$; it follows that $af_m + bf^m \in \mathfrak{S}_{\mathbb{R}}[x_1, \dots, x_{m-1}]$ for all $a, b \in \mathbb{R}$. Furthermore, for all $j \in [m-1]$ and $\mathbf{a} \in \mathbb{R}^{m-1}$, $W_j[f_m, f^m](\mathbf{a}) = -\Delta_{jm}f(\mathbf{a}, 1) \leq 0$. This verifies condition (c) of Corollary 2.10 for the pair (f_m, f^m) , and it follows that $f = f^m + x_m f_m \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}]$, completing the induction.

It is clear that (a) and (b) imply (c) – we show that (c) implies (b) below. This is clear if $m \leq 2$, so assume that $m \geq 3$. To begin with, let $\{h, i, j\} \subseteq [m]$ be three distinct indices, and consider $\Delta_{ij}f$ as a polynomial in x_h . That is, $\Delta_{ij}f = A_{hij}x_h^2 + B_{hij}x_h + C_{hij}$ in which

$$\begin{aligned} A_{hij} &= f_{hi}^j f_{hj}^i - f_h^{ij} f_{hij} = \Delta_{ij}f_h, \\ B_{hij} &= f_{hi}^j f_j^{hi} - f_h^{ij} f_{ij}^h + f_i^{hj} f_{hj}^i - f^{hij} f_{hij}, \text{ and} \\ C_{hij} &= f_i^{hj} f_j^{hi} - f^{hij} f_{ij}^h = \Delta_{ij}f^h. \end{aligned}$$

If $\Delta_{ij}f(\mathbf{a}) \geq 0$ for all $\mathbf{a} \in \mathbb{R}^m$ then this quadratic polynomial in x_h :

$$\Delta_{ij}f(a_1, \dots, a_{h-1}, x_h, a_{h+1}, \dots, a_m)$$

has a nonpositive discriminant for all $\mathbf{a} \in \mathbb{R}^m$. That is, $D_{hij} = B_{hij}^2 - 4A_{hij}C_{hij}$ is such that $D_{hij}(\mathbf{a}) \leq 0$ for all $\mathbf{a} \in \mathbb{R}^m$.

It is a surprising fact that as a polynomial in $\{x_k : k \in [m] \setminus \{h, i, j\}\}$, D_{hij} is invariant under all six permutations of its indices, as is seen by direct calculation:

$$\begin{aligned} D_{hij} &= (f_h^{ij} f_{ij}^h)^2 + (f_i^{hj} f_{hj}^i)^2 + (f_j^{hi} f_{hi}^j)^2 + (f_{hij} f^{hij})^2 \\ &\quad - 2(f_h^{ij} f_{ij}^h f_i^{hj} f_{hj}^i + f_i^{hj} f_{hj}^i f_j^{hi} f_{hi}^j + f_j^{hi} f_{hi}^j f_h^{ij} f_{ij}^h) \\ &\quad - 2(f_h^{ij} f_{ij}^h + f_i^{hj} f_{hj}^i + f_j^{hi} f_{hi}^j) f^{hij} f_{hij} \\ &\quad + 4f_i^h f_j^i f_{hj}^i f_{hi}^j + 4f_h^{ij} f_i^{hj} f_j^{hi} f_{hij}. \end{aligned}$$

Now for the proof that (c) implies (b) when $m \geq 3$. Consider any $h \in [m] \setminus \{i, j\}$. Then

$$\Delta_{hi}f = A_{jhi}x_j^2 + B_{jhi}x_j + C_{jhi}$$

has discriminant $D_{jhi} = D_{hij}$. Since f_j and f^j are real stable, we have $A_{jhi}(\mathbf{a}) = \Delta_{hi}f_j(\mathbf{a}) \geq 0$ and $C_{jhi}(\mathbf{a}) = \Delta_{hi}f^j(\mathbf{a}) \geq 0$ for all $\mathbf{a} \in \mathbb{R}^m$. Since $\Delta_{ij}f(\mathbf{a}) \geq 0$ for all $\mathbf{a} \in \mathbb{R}^m$ it follows that $D_{jhi}(\mathbf{a}) = D_{hij}(\mathbf{a}) \leq 0$ for all $\mathbf{a} \in \mathbb{R}^m$. It follows that $\Delta_{hi}f(\mathbf{a}) \geq 0$ for all $\mathbf{a} \in \mathbb{R}^m$. (Note that if $B^2 - 4AC \leq 0$ and either $A = 0$ or $C = 0$, then $B = 0$.) A similar argument using the fact that f_i and f^i are real stable shows that $\Delta_{hj}f(\mathbf{a}) \geq 0$ for all $\mathbf{a} \in \mathbb{R}^m$.

It remains to show that $\Delta_{hk}f(\mathbf{a}) \geq 0$ for all $\mathbf{a} \in \mathbb{R}^m$ when $\{h, k\}$ is disjoint from $\{i, j\}$. We have seen that $\Delta_{hi}f(\mathbf{a}) \geq 0$ for all $\mathbf{a} \in \mathbb{R}^m$, and we know that both f_i and f^i are real stable. The argument above applies once more: $\Delta_{hi}f(\mathbf{a}) \geq 0$ for all $\mathbf{a} \in \mathbb{R}^m$, so that $D_{ihk}(\mathbf{a}) = D_{khi}(\mathbf{a}) \leq 0$ for all $\mathbf{a} \in \mathbb{R}^m$, and then since $A_{ihk}(\mathbf{a}) \geq 0$ and $C_{ihk}(\mathbf{a}) \geq 0$ for all $\mathbf{a} \in \mathbb{R}^m$ it follows that $\Delta_{hk}f(\mathbf{a}) \geq 0$ for all $\mathbf{a} \in \mathbb{R}^m$. Thus (c) implies (b). \square

3.2. Linear transformations preserving stability – multiaffine case.

Lemma 3.2 (Lieb-Sokal Lemma, Lemma 2.1 of [5]). *Let $g(\mathbf{x}) + yf(\mathbf{x}) \in \mathfrak{S}[\mathbf{x}, y]$ be stable and such that $\deg_i(f) \leq 1$. Then $g - \partial_i f \in \mathfrak{S}[\mathbf{x}]$ is stable.*

Proof. Since g is stable (by specialization to $y = 0$), there is nothing to prove if $\partial_i f \equiv 0$ identically, so assume otherwise (and hence that $f \not\equiv 0$). By permutation we can assume that $i = 1$. Since f is stable and $z_1, z \in \mathcal{H}$ imply that $z_1 - z^{-1} \in \mathcal{H}$, it follows that

$$yf(x_1 - y^{-1}, x_2, \dots, x_m) = -\partial_1 f(\mathbf{x}) + yf(\mathbf{x})$$

is stable. Proposition 2.7(b) implies that for all $\mathbf{z} \in \mathcal{H}^m$,

$$\operatorname{Im} \left(\frac{g(\mathbf{z}) - \partial_1 f(\mathbf{z})}{f(\mathbf{z})} \right) = \operatorname{Im} \left(\frac{g(\mathbf{z})}{f(\mathbf{z})} \right) + \operatorname{Im} \left(\frac{-\partial_1 f(\mathbf{z})}{f(\mathbf{z})} \right) \geq 0.$$

Thus, by Proposition 2.7(b) again, $g - \partial_1 f + yf$ is stable. Specializing to $y = 0$ shows that $g - \partial_1 f$ is stable. \square

Exercise 3.3 (Lemma 3.1 of [5]). Let $f \in \mathbb{C}[\mathbf{x}]^{\text{MA}}$ and $\mathbf{w} \in \mathcal{H}^m$. Then for all $\epsilon > 0$ sufficiently small, $(\mathbf{x} + \mathbf{w})^{[m]} + \epsilon f(\mathbf{x})$ is stable. (Here $(\mathbf{x} + \mathbf{w})^{[m]} = \prod_{i=1}^m (x_i + w_i)$.)

For a linear transformation $T : \mathbb{C}[\mathbf{x}]^{\text{MA}} \rightarrow \mathbb{C}[\mathbf{x}]$ of multiaffine polynomials, define the *algebraic symbol* of T to be the polynomial

$$T((\mathbf{x} + \mathbf{y})^{[m]}) = T \left(\prod_{i=1}^m (x_i + y_i) \right) = \sum_{S \subseteq [m]} T(\mathbf{x}^S) \mathbf{y}^{[m] \setminus S}$$

in $\mathbb{C}[x_1, \dots, x_m, y_1, \dots, y_m] = \mathbb{C}[\mathbf{x}, \mathbf{y}]$.

Theorem 3.4 (Theorem 1.1 of [5]). *Let $T : \mathbb{C}[\mathbf{x}]^{\text{MA}} \rightarrow \mathbb{C}[\mathbf{x}]$ be a linear transformation. Then T maps $\mathfrak{S}[\mathbf{x}]^{\text{MA}}$ into $\mathfrak{S}[\mathbf{x}]$ if and only if either*

- (a) $T(f) = \eta(f) \cdot p$ for some linear functional $\eta : \mathbb{C}[\mathbf{x}]^{\text{MA}} \rightarrow \mathbb{C}$ and $p \in \mathfrak{S}[\mathbf{x}]$, or
- (b) the polynomial $T((\mathbf{x} + \mathbf{y})^{[m]})$ is stable in $\mathfrak{S}[\mathbf{x}, \mathbf{y}]$.

Proof. First, assume (b) that $T((\mathbf{x} + \mathbf{y})^{[m]}) \in \mathfrak{S}[\mathbf{x}, \mathbf{y}]$ is stable. By inversion, it follows that $\mathbf{y}^{[m]} T((\mathbf{x} - \mathbf{y}^{-1})^{[m]})$ is also stable. Thus, if $f \in \mathfrak{S}[w_1, \dots, w_m]$ is stable then

$$\mathbf{y}^{[m]} T((\mathbf{x} - \mathbf{y}^{-1})^{[m]}) f(\mathbf{w}) = \sum_{S \subseteq [m]} T(\mathbf{x}^S) (-\mathbf{y})^S f(\mathbf{w})$$

is stable. If f is also multiaffine then repeated application of the Lieb-Sokal Lemma 3.2 (replacing y_i by $-\partial/\partial w_i$ for $i \in [m]$) shows that

$$\sum_{S \subseteq [m]} T(\mathbf{x}^S) \frac{\partial^S}{\partial \mathbf{w}^S} f(\mathbf{w})$$

is stable. Finally, specializing to $\mathbf{w} = \mathbf{0}$ shows that $T(f(\mathbf{x}))$ is stable. Thus, the linear transformation T maps $\mathfrak{S}[\mathbf{x}]^{\text{MA}}$ into $\mathfrak{S}[\mathbf{x}]$. This is clearly also the case if (a) holds.

Conversely, assume that T maps $\mathfrak{S}[\mathbf{x}]^{\text{MA}}$ into $\mathfrak{S}[\mathbf{x}]$. Then for any $\mathbf{w} \in \mathcal{H}^m$, $(\mathbf{x} + \mathbf{w})^{[m]} \in \mathfrak{S}[\mathbf{x}]^{\text{MA}}$, so that $T((\mathbf{x} + \mathbf{w})^{[m]}) \in \mathfrak{S}[\mathbf{x}]$.

First, assume that there is a $\mathbf{w} \in \mathcal{H}^m$ for which $T((\mathbf{x} + \mathbf{w})^{[m]}) \equiv 0$ identically. For any $f \in \mathbb{C}[\mathbf{x}]^{\text{MA}}$ let $\epsilon > 0$ be as in Exercise 3.3. Then $\epsilon T(f) = T((\mathbf{x} + \mathbf{w})^{[m]} + \epsilon f)$ is stable, so that $T(f)$ is stable. Thus, the image of $\mathbb{C}[\mathbf{x}]^{\text{MA}}$ under T is a \mathbb{C} -subspace of $\mathfrak{S}[\mathbf{x}]$. By Proposition 2.12(b), T has the form of case (a).

Secondly, if $T((\mathbf{x} + \mathbf{w})^{[m]}) \not\equiv 0$ for all $\mathbf{w} \in \mathcal{H}^m$ then, since each of these polynomials is in $\mathfrak{S}[\mathbf{x}]$, we have $T((\mathbf{x} + \mathbf{w})^{[m]})|_{\mathbf{x}=\mathbf{z}} \neq 0$ for all $\mathbf{z} \in \mathcal{H}^m$ and $\mathbf{w} \in \mathcal{H}^m$. This shows that $T((\mathbf{x} + \mathbf{y})^{[m]})$ is stable in $\mathfrak{S}[\mathbf{x}, \mathbf{y}]$, which is the form of case (b). \square

Theorem 3.4 has a corresponding real form – the proof is completely analogous.

Theorem 3.5 (Theorem 1.2 of [5]). *Let $T : \mathbb{R}[\mathbf{x}]^{\text{MA}} \rightarrow \mathbb{R}[\mathbf{x}]$ be a linear transformation. Then T maps $\mathfrak{S}_{\mathbb{R}}[\mathbf{x}]^{\text{MA}}$ into $\mathfrak{S}_{\mathbb{R}}[\mathbf{x}]$ if and only if either*

- (a) $T(f) = \eta(f) \cdot p + \xi(f) \cdot q$ for some linear functionals $\eta, \xi : \mathbb{R}[\mathbf{x}]^{\text{MA}} \rightarrow \mathbb{R}$ and $p, q \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}]$ such that $p \ll q$, or
- (b) the polynomial $T((\mathbf{x} + \mathbf{y})^{[m]})$ is real stable in $\mathfrak{S}_{\mathbb{R}}[\mathbf{x}, \mathbf{y}]$, or
- (c) the polynomial $T((\mathbf{x} - \mathbf{y})^{[m]})$ is real stable in $\mathfrak{S}_{\mathbb{R}}[\mathbf{x}, \mathbf{y}]$.

Proof. **Exercise 3.6.** \square

4. THE GRACE-WALSH-SZEGŐ COINCIDENCE THEOREM.

Let $f \in \mathbb{C}[x]$ be a univariate polynomial of degree at most m , and let $\mathbf{x} = (x_1, \dots, x_m)$ as usual. For $0 \leq j \leq m$, the j -th elementary symmetric function of \mathbf{x} is

$$e_j(\mathbf{x}) = \sum_{1 \leq i_1 < \dots < i_j \leq m} x_{i_1} \cdots x_{i_j} = \sum_{S \subseteq [m]: |S|=j} \mathbf{x}^S.$$

The m -th polarization of f is the polynomial obtained as the image of f under the linear transformation Pol_m defined by $x^j \mapsto \binom{m}{j}^{-1} e_j(\mathbf{x})$ for all $0 \leq j \leq m$, and linear extension. In other words, $\text{Pol}_m f$ is the unique multiaffine polynomial in $\mathbb{C}[\mathbf{x}]^{\text{MA}}$ that is invariant under all permutations of $[m]$ and such that $\text{Pol}_m f(x, \dots, x) = f(x)$. A circular region is a nonempty subset \mathcal{A} of \mathbb{C} that is either open or closed, and which is bounded by either a circle or a straight line.

Theorem 4.1 (Grace-Walsh-Szegő, Theorem 3.4.1b of [14]). *Let $f \in \mathbb{C}[x]$ have degree at most m and let \mathcal{A} be a circular region. If either $\deg(f) = m$ or \mathcal{A} is convex, then for every $\mathbf{z} \in \mathcal{A}^m$ there exists $z \in \mathcal{A}$ such that $\text{Pol}_m f(\mathbf{z}) = f(z)$.*

Figure 1 illustrates the Grace-Walsh-Szegő (GWS) Theorem for the polynomial $f(x) = x^5 + 10x^2 + 1$. The black dots mark the solutions to $f(x) = 0$. Any permutation of the red (grey) dots is a solution to $\text{Pol}_5 f(x_1, \dots, x_5) = 0$. By GWS, any circular region containing all the red dots must contain at least one of the black dots. The figure indicates the boundaries of several circular regions for which this condition is met.

The proof of GWS in this section is adapted from Borcea and Brändén [6].

4.1. Reduction to the case of stable polynomials. First of all, it suffices to prove GWS for open circular regions, since a closed circular region is the intersection of all the open circular regions which contain it. Second, it suffices to show that for any $g \in \mathbb{C}[x]$ of degree at most m , if $\deg(g) = m$ or \mathcal{A} is convex, and $\mathbf{z} \in \mathcal{A}^m$ is such that $\text{Pol}_m g(\mathbf{z}) = 0$, then there exists $z \in \mathcal{A}$ such that $g(z) = 0$. This implies the stated form of GWS by applying this special case to $g(x) = f(x) - c$, where $c = \text{Pol}_m f(\mathbf{z})$. Stated otherwise, it suffices to show that if $f(z) \neq 0$ for all $z \in \mathcal{A}$ then $\text{Pol}_m f(\mathbf{z}) \neq 0$ for all $\mathbf{z} \in \mathcal{A}^m$ (provided that either $\deg(f) = m$ or \mathcal{A} is convex).

Let \mathcal{M} be the set of Möbius transformations $z \mapsto \phi(z) = (az + b)/(cz + d)$ with $a, b, c, d \in \mathbb{C}$ and $ab - cd = \pm 1$. Then \mathcal{M} with the operation of functional composition is a group of conformal transformations of the Riemann sphere $\widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$,

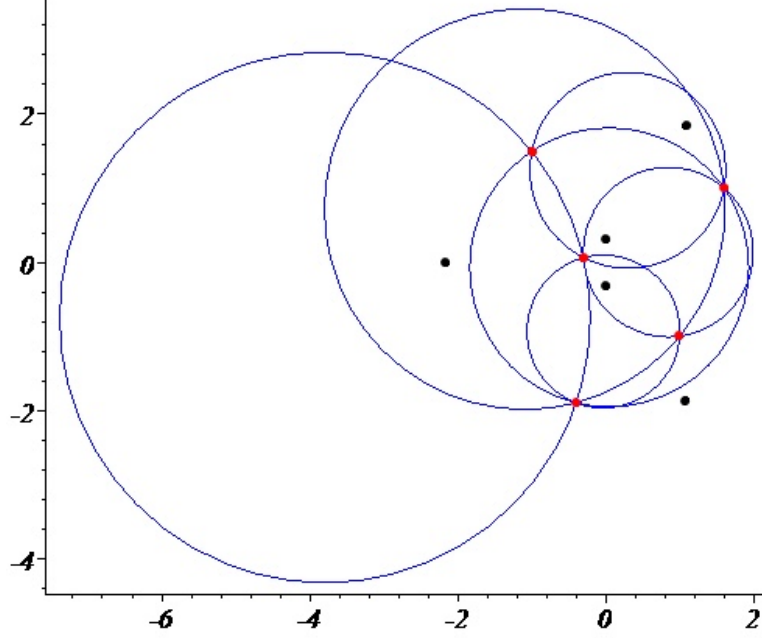


FIGURE 1. Illustration of the Grace-Walsh-Szegő Theorem.

and it acts simply transitively on the set of all ordered triples of distinct points of $\widehat{\mathbb{C}}$. Consequently, for any open circular region \mathcal{A} there is a $\phi \in \mathcal{M}$ such that $\phi(\mathcal{H}) = \{\phi(z) : z \in \mathcal{H}\} = \mathcal{A}$. We henceforth regard circular regions as subsets of $\widehat{\mathbb{C}}$. Note that an open circular region \mathcal{A} is convex if and only if it does not contain ∞ . (The point ∞ is on the boundary of any open half-plane.) In this case, if $\phi(z) = (az + b)/(cz + d)$ is such that $\phi(\mathcal{H}) = \mathcal{A}$ then $cz + d \neq 0$ for all $z \in \mathcal{H}$.

Given $0 \neq f \in \mathbb{C}[x]$ of degree at most m , consider the polynomial $\tilde{f}(x) = (cx + d)^m f((ax + b)/(cx + d))$. If either $\deg(f) = m$ or \mathcal{A} is convex, then f is nonvanishing on \mathcal{A} if and only if $\tilde{f}(z)$ is nonvanishing on \mathcal{H} . Also,

$$\text{Pol}_m \tilde{f}(\mathbf{x}) = \text{Pol}_m f(\phi(x_1), \dots, \phi(x_m)) \cdot \prod_{i=1}^m (cx_i + d).$$

Thus, to prove GWS it suffices to prove the following lemma.

Lemma 4.2. *Let $f \in \mathbb{C}[x]$ be a univariate polynomial of degree at most m . Then $\text{Pol}_m f$ is stable if and only if f is stable.*

Clearly, diagonalization implies that if $\text{Pol}_m f$ is stable then f is stable, so only the converse implication needs proof. This is accomplished in the following two easy steps.

4.2. Partial symmetrization. The group $\mathfrak{S}(m)$ of all permutations $\sigma : [m] \rightarrow [m]$ acts on $\mathbb{C}[\mathbf{x}]$ by the rule $\sigma(f)(x_1, \dots, x_m) = f(x_{\sigma(1)}, \dots, x_{\sigma(m)})$. Notice that

$$\sigma(\mathbf{x}^\alpha) = \prod_{i=1}^m x_{\sigma(i)}^{\alpha(i)} = \prod_{i=1}^m x_i^{\alpha \circ \sigma^{-1}(i)} = \mathbf{x}^{\alpha \circ \sigma^{-1}}.$$

For $\{i, j\} \subseteq [m]$, let τ_{ij} be the transposition that exchanges i and j and fixes all other elements of $[m]$.

Lemma 4.3. *Let $0 \leq \lambda \leq 1$ and $\{i, j\} \subseteq [m]$, and let $T_{ij}^{(\lambda)} = (1 - \lambda) + \lambda\tau_{ij}$. If $f \in \mathfrak{S}[\mathbf{x}]^{\text{MA}}$ is stable and multiaffine then $T_{ij}^{(\lambda)} f \in \mathfrak{S}[\mathbf{x}]^{\text{MA}}$ is stable and multiaffine.*

Proof. If f is multiaffine then $(1 - \lambda)f + \lambda\tau_{ij}(f)$ is also multiaffine. We apply Theorem 3.4 to show that $T = T_{ij}^{(\lambda)}$ preserves stability of multiaffine polynomials. By permutation we can assume that $\{i, j\} = \{1, 2\}$. The algebraic symbol of T is

$$T((\mathbf{x} + \mathbf{y})^{[m]}) = T((x_1 + y_1)(x_2 + y_2)) \cdot \prod_{i=3}^m (x_i + y_i).$$

Clearly, this is stable if and only if the same is true of $T((x_1 + y_1)(x_2 + y_2))$. Exercise 4.4 completes the proof. \square

Exercise 4.4. Use the results of Sections 2.4 or 3.1 to show that for $0 \leq \lambda \leq 1$, the polynomial

$$x_1 x_2 + ((1 - \lambda)x_1 + \lambda x_2)y_2 + (\lambda x_1 + (1 - \lambda)x_2)y_1 + y_1 y_2$$

is real stable.

4.3. Convergence to the polarization. Let $0 \neq f(x) \in \mathfrak{S}[x]$ be a univariate stable polynomial of degree at most m : say $f(x) = c(x - \xi_1) \cdots (x - \xi_n)$ in which $c \neq 0$, $n \leq m$, and $\xi_i \notin \mathcal{H}$ for all $i \in [n]$. Then the polynomial $F_0 \in \mathbb{C}[\mathbf{x}]$ defined by

$$F_0(x_1, \dots, x_m) = c(x_1 - \xi_1) \cdots (x_n - \xi_n)$$

is multiaffine and stable, and $F_0(x, \dots, x) = f(x)$. Let $\Sigma = (\{i_k, j_k\} : k \in \mathbb{N})$ be a sequence of two-element subsets of $[m]$, and for each $k \in \mathbb{N}$ let $T_k = T_{i_k j_k}^{(1/2)}$ and define $F_{k+1} = T_k(F_k)$. By induction using Lemma 4.3, each $F_k \in \mathfrak{S}[\mathbf{x}]^{\text{MA}}$ is multiaffine and stable, and $F_k(x, \dots, x) = f(x)$ for all $k \in \mathbb{N}$. We will construct such a sequence Σ for which $(F_k : k \in \mathbb{N})$ converges to $\text{Pol}_m f$.

Let $P \in \mathbb{C}[\mathbf{x}]^{\text{MA}}$ be multiaffine, say $P(\mathbf{x}) = \sum_{S \subseteq [m]} c(S)\mathbf{x}^S$. For $\{i, j\} \subseteq [m]$ let

$$\omega_{ij}(P) = \sum_{S \subseteq [m]} |c(S) - c(\tau_{ij}(S))|$$

be the ij -th imbalance of P , and let $\|P\| = \sum_{\{i, j\} \subseteq [m]} \omega_{ij}(P)$ be the total imbalance of P .

Exercise 4.5. (a) Let $(P_k : k \in \mathbb{N})$ be polynomials in $\mathbb{C}[\mathbf{x}]^{\text{MA}}$ for which there is a $p \in \mathbb{C}[x]$ such that $P_k(x, \dots, x) = p(x)$ for all $k \in \mathbb{N}$. If $\|P_k\| \rightarrow 0$ as $k \rightarrow \infty$, then $(P_k : k \in \mathbb{N})$ converges to a limit $P \in \mathbb{C}[\mathbf{x}]^{\text{MA}}$, and $\|P\| = 0$.

(b) For $P \in \mathbb{C}[\mathbf{x}]^{\text{MA}}$, $\|P\| = 0$ if and only if P is invariant under all permutations of $[m]$. Thus, in part (a) the limit is $P = \text{Pol}_m p$.

Exercise 4.6. Let $P \in \mathbb{C}[\mathbf{x}]^{\text{MA}}$, let $\{i, j\} \subseteq [m]$, and let $Q = T_{ij}^{(1/2)}P$.

- (a) Then $\omega_{ij}(Q) = 0$.
- (b) If $h \in [m] \setminus \{i, j\}$ then $\omega_{hi}(Q) \leq (\omega_{hi}(P) + \omega_{hj}(P))/2$, and similarly for $\omega_{hj}(Q)$.
- (c) If $\{h, k\} \subseteq [m] \setminus \{i, j\}$ then $\omega_{hk}(Q) = \omega_{hk}(P)$.
- (d) Consequently, $\|Q\| \leq \|P\| - \omega_{ij}(P)$.

Now we choose the sequence $\Sigma = (\{i_k, j_k\} : k \in \mathbb{N})$ as follows: for each $k \in \mathbb{N}$, $\{i_k, j_k\} \subseteq [m]$ is any pair of indices $\{i, j\}$ for which $\omega_{ij}(F_k)$ attains its maximum value. Then $\omega_{i_k j_k}(F_k) \geq \binom{m}{2}^{-1} \|F_k\|$, so that by Exercise 4.6(d) and induction on $k \in \mathbb{N}$,

$$\|F_{k+1}\| \leq \left(1 - \binom{m}{2}^{-1}\right) \|F_k\| \leq \left(1 - \binom{m}{2}^{-1}\right)^{k+1} \|F_0\|.$$

Thus, by Exercise 4.5, F_k converges to $\text{Pol}_m f$, the m -th polarization of f . Finally, since each F_k is stable (and the limit is a polynomial), Hurwitz's Theorem implies that $\text{Pol}_m f$ is stable. This completes the proof of Lemma 4.2, and hence of Theorem 4.1.

5. POLARIZATION ARGUMENTS AND STABILITY PRESERVERS.

For $\kappa \in \mathbb{N}^m$ and a set $\mathcal{S} \subseteq \mathbb{C}[\mathbf{x}]$ of polynomials, let $\mathcal{S}^{\leq \kappa}$ be the set of all $f \in \mathcal{S}$ such that $\deg_i(f) \leq \kappa(i)$ for all $i \in [m]$. Let

$$I(\kappa) = \{(i, j) : i \in [m] \text{ and } j \in [\kappa(i)]\}$$

and let $\mathbf{u} = \{u_{ij} : (i, j) \in I(\kappa)\}$ be indeterminates. For $f \in \mathbb{C}[\mathbf{x}]^{\leq \kappa}$, Let $\text{Pol}_{\kappa(i)}^{(i)} f$ denote the $\kappa(i)$ -th polarization of x_i in f : this is the image of f under the linear transformation $\text{Pol}_{\kappa(i)}^{(i)}$ defined by $x_i^j \mapsto \binom{\kappa(i)}{j}^{-1} e_j(u_{i1}, \dots, u_{i\kappa(i)})$ for each $0 \leq j \leq \kappa(i)$, and linear extension. Finally, the κ -th polarization of f is

$$\text{Pol}_{\kappa} f = \text{Pol}_{\kappa(m)}^{(m)} \circ \dots \circ \text{Pol}_{\kappa(1)}^{(1)} f.$$

This defines a linear transformation $\text{Pol}_{\kappa} : \mathbb{C}[\mathbf{x}]^{\leq \kappa} \rightarrow \mathbb{C}[\mathbf{u}]^{\text{MA}}$.

5.1. The real stability criterion revisited.

Proposition 5.1. *Let $\kappa \in \mathbb{N}^m$ and $f \in \mathbb{C}[\mathbf{x}]^{\leq \kappa}$. Then $\text{Pol}_{\kappa} f$ is stable if and only if f is stable.*

Proof. Diagonalization implies that if $\text{Pol}_{\kappa} f$ is stable then f is stable, so only the converse implication needs proof. Assume that f is stable, and let $z_{ij} \in \mathcal{H}$ for $(i, j) \in I(\kappa)$. By induction on m , repeated application of GWS shows that there are $\mathbf{z} = (z_1, \dots, z_m) \in \mathcal{H}^m$ such that

$$\text{Pol}_{\kappa} f(z_{ij} : (i, j) \in I(\kappa)) = f(\mathbf{z}).$$

Since f is stable it follows that $\text{Pol}_{\kappa} f$ is stable. \square

If $f \in \mathbb{R}[\mathbf{x}]^{\leq \kappa}$ then Theorem 3.1 applies to $\text{Pol}_{\kappa} f$. Thus, Proposition 5.1 bootstraps the real stability criterion from multiaffine to arbitrary polynomials. This is a typical application of the GWS Theorem.

5.2. Linear transformations preserving stability – polynomial case.

Theorem 5.2 (Theorem 1.1 of [5]). *Let $\kappa \in \mathbb{N}^m$, and let $T : \mathbb{C}[\mathbf{x}]^{\leq \kappa} \rightarrow \mathbb{C}[\mathbf{x}]$ be a linear transformation. Then T maps $\mathfrak{S}[\mathbf{x}]^{\leq \kappa}$ into $\mathfrak{S}[\mathbf{x}]$ if and only if either*

- (a) $T(f) = \eta(f) \cdot p$ for some linear functional $\eta : \mathbb{C}[\mathbf{x}]^{\leq \kappa} \rightarrow \mathbb{C}$ and $p \in \mathfrak{S}[\mathbf{x}]$, or
- (b) the polynomial $T((\mathbf{x} + \mathbf{y})^\kappa)$ is stable in $\mathfrak{S}[\mathbf{x}, \mathbf{y}]$.

Proof. Let $\mathbf{u} = \{u_{ij} : (i, j) \in I(\kappa)\}$, and define a linear transformation $\tilde{T} : \mathbb{C}[\mathbf{u}]^{\text{MA}} \rightarrow \mathbb{C}[\mathbf{x}]$ as follows. For every $A \subseteq I(\kappa)$, define $\alpha(A) : [m] \rightarrow \mathbb{N}$ by putting $\alpha(A, i) = |\{j \in [\kappa(i)] : (i, j) \in A\}|$ for each $i \in [m]$. Then for each $A \subseteq I(\kappa)$ define $\tilde{T}(\mathbf{u}^A) = T(\mathbf{x}^{\alpha(A)})$, and extend this linearly to all of $\mathbb{C}[\mathbf{u}]^{\text{MA}}$. Let $\Delta : \mathbb{C}[\mathbf{u}]^{\text{MA}} \rightarrow \mathbb{C}[\mathbf{x}]$ be the diagonalization operator defined by $\Delta(u_{ij}) = x_i$ for all $(i, j) \in I(\kappa)$, extended algebraically.

Notice that $T = \tilde{T} \circ \text{Pol}_\kappa$, and that $\tilde{T} = T \circ \Delta$. By Proposition 5.1 (and Lemma 2.4), it follows that T preserves stability if and only if \tilde{T} preserves stability. This is equivalent to one of two cases in Theorem 3.4.

In case (a), if $\tilde{T} = p \cdot \tilde{\eta}$ for some $p \in \mathfrak{S}[\mathbf{x}]$ and linear functional $\tilde{\eta} : \mathbb{C}[\mathbf{y}]^{\text{MA}} \rightarrow \mathbb{C}$ then $T = p \cdot (\eta \circ \text{Pol}_\kappa)$ is also in case (a). Conversely, if T is in case (a) then the same is true of \tilde{T} , by construction.

In case (b), let $\text{Pol}_\kappa^{(\mathbf{y})} : \mathbb{C}[\mathbf{y}]^{\leq \kappa} \rightarrow \mathbb{C}[\mathbf{v}]^{\text{MA}}$ denote the κ -th polarization of the \mathbf{y} variables. The symbols of T and \tilde{T} are related by

$$\tilde{T}((\mathbf{u} + \mathbf{v})^{I(\kappa)}) = (T \circ \Delta)((\mathbf{u} + \mathbf{v})^{I(\kappa)}) = \text{Pol}_\kappa^{(\mathbf{y})} T((\mathbf{x} + \mathbf{y})^\kappa),$$

and Proposition 5.1 shows that T is in case (b) if and only if \tilde{T} is in case (b). \square

5.3. Linear transformations preserving stability – transcendental case.

Exercise 5.3. Let $T : \mathbb{C}[\mathbf{x}] \rightarrow \mathbb{C}[\mathbf{x}]$ be a linear transformation.

- (a) Then $T : \mathfrak{S}[\mathbf{x}] \rightarrow \mathfrak{S}[\mathbf{x}]$ if and only if $T : \mathfrak{S}[\mathbf{x}]^{\leq \kappa} \rightarrow \mathfrak{S}[\mathbf{x}]$ for all $\kappa \in \mathbb{N}^m$.
- (b) Define $S : \mathbb{C}[\mathbf{x}, \mathbf{y}] \rightarrow \mathbb{C}[\mathbf{x}, \mathbf{y}]$ by $S(\mathbf{x}^\alpha \mathbf{y}^\beta) = T(\mathbf{x}^\alpha) \mathbf{y}^\beta$ and linear extension. If $T((\mathbf{x} + \mathbf{u})^\kappa)$ is stable for all $\kappa \in \mathbb{N}^m$ then $S((\mathbf{x} + \mathbf{u})^\kappa (\mathbf{y} + \mathbf{v})^\beta)$ is stable for all $\kappa, \beta \in \mathbb{N}^m$.

Let $\overline{\mathfrak{S}[\mathbf{x}]}$ denote the set of all power series in $\mathbb{C}[[\mathbf{x}]]$ that are obtained as the limit of a sequence of stable polynomials in $\mathfrak{S}[\mathbf{x}]$ which converges uniformly on compact sets. Theorem 5.4 is an astounding generalization of the Pólya-Schur Theorem. For $\alpha \in \mathbb{N}^m$, let $\alpha! = \prod_{i=1}^m \alpha(i)!$.

Theorem 5.4 (Theorem 1.3 of [5]). *Let $T : \mathbb{C}[\mathbf{x}] \rightarrow \mathbb{C}[\mathbf{x}]$ be a linear transformation. Then T maps $\mathfrak{S}[\mathbf{x}]$ into $\overline{\mathfrak{S}[\mathbf{x}]}$ if and only if either*

- (a) $T(f) = \eta(f) \cdot p$ for some linear functional $\eta : \mathbb{C}[\mathbf{x}] \rightarrow \mathbb{C}$ and $p \in \mathfrak{S}[\mathbf{x}]$, or
- (b) the power series

$$T(e^{-\mathbf{x}\mathbf{y}}) = \sum_{\alpha : [m] \rightarrow \mathbb{N}} (-1)^\alpha T(\mathbf{x}^\alpha) \frac{\mathbf{y}^\alpha}{\alpha!}$$

is in $\overline{\mathfrak{S}[\mathbf{x}, \mathbf{y}]}$

(Theorem 3.5 has a similar extension – see Theorems 1.2 and 1.4 of [5].)
For $\alpha \leq \beta$ in \mathbb{N}^m , let $(\beta)_\alpha = \beta! / (\beta - \alpha)!$, and for $\alpha \not\leq \beta$ let $(\beta)_\alpha = 0$.

Theorem 5.5 (Theorem 5.1 of [5]). *Let $F(\mathbf{x}, \mathbf{y}) = \sum_{\alpha \in \mathbb{N}^m} P_\alpha(\mathbf{x}) \mathbf{y}^\alpha$ be a power series in $\mathbb{C}[\mathbf{x}][[\mathbf{y}]]$ (so that each $P_\alpha \in \mathbb{C}[\mathbf{x}]$). Then $F(\mathbf{x}, \mathbf{y})$ is in $\overline{\mathfrak{S}[\mathbf{x}, \mathbf{y}]}$ if and only if for all $\beta \in \mathbb{N}^m$,*

$$\sum_{\alpha \leq \beta} (\beta)_\alpha P_\alpha(\mathbf{x}) \mathbf{y}^\alpha$$

is stable in $\mathfrak{S}[\mathbf{x}, \mathbf{y}]$.

(This implies the analogous result for real stability, since $\mathfrak{S}_{\mathbb{R}}[\mathbf{x}] = \mathfrak{S}[\mathbf{x}] \cap \mathbb{R}[\mathbf{x}]$.)

Exercise 5.6. Derive Theorem 5.4 from Theorems 5.2 and 5.5. (Hint: $T((\mathbf{x} + \mathbf{y})^\kappa)$ is stable if and only if $T((\mathbf{1} - \mathbf{xy})^\kappa)$ is stable.)

One direction of Theorem 5.5 is relatively straightforward.

Lemma 5.7 (Lemma 5.2 of [5]). *Fix $\beta \in \mathbb{N}^m$. The linear transformation $T : \mathbf{y}^\alpha \mapsto (\beta)_\alpha \mathbf{y}^\alpha$ on $\mathbb{C}[\mathbf{y}]$ preserves stability.*

Proof. By Theorem 5.2 and Exercise 5.3(a), it suffices to show that for all $\kappa \in \mathbb{N}^m$, the polynomial $T((\mathbf{y} + \mathbf{u})^\kappa)$ is stable. Now

$$T((\mathbf{y} + \mathbf{u})^\kappa) = \prod_{i=1}^m \left[\sum_{j=0}^{\kappa(i)} j! \binom{\kappa(i)}{j} \binom{\beta(i)}{j} y_i^j u_i^{\kappa(i)-j} \right],$$

so it suffices to show that for all $k, b \in \mathbb{N}$, the polynomial $f(t) = \sum_{j=0}^k j! \binom{k}{j} \binom{b}{j} t^j$ is real stable. Let $g(t) = (1 + d/dt)^k t^b$. One can check that $f(t) = t^b g(1/t)$. It thus suffices to show that $1 + d/dt$ preserves stability. For any $a \in \mathbb{N}$, $(1 + d/dt)(t + u)^a = (t + u + a)(t + u)^{a-1}$ is stable, and so Theorem 5.2 implies the result. \square

Now, let $F = F(\mathbf{x}, \mathbf{y})$ be as in the statement of Theorem 5.5, and let $(F_n : n \in \mathbb{N})$ be a sequence of stable polynomials $F_n(\mathbf{x}, \mathbf{y}) = \sum_{\alpha \in \mathbb{N}^m} P_{n,\alpha}(\mathbf{x}) \mathbf{y}^\alpha$ in $\mathfrak{S}[\mathbf{x}, \mathbf{y}]$ converging to F uniformly on compact sets. Fix $\beta \in \mathbb{N}^m$ and define a linear transformation $T : \mathbb{C}[\mathbf{x}, \mathbf{y}] \rightarrow \mathbb{C}[\mathbf{x}, \mathbf{y}]$ by $T(\mathbf{x}^\gamma \mathbf{y}^\alpha) = (\beta)_\alpha \mathbf{x}^\gamma \mathbf{y}^\alpha$ and linear extension. By Lemma 5.7 and Exercise 5.3, T preserves stability in $\mathfrak{S}[\mathbf{x}, \mathbf{y}]$. Thus, $(T(F_n) : n \in \mathbb{N})$ is a sequence of stable polynomials converging to $T(F)$. Since $T(F)$ is a polynomial the convergence is uniform on compact sets, and so Hurwitz's Theorem implies that $T(F)$ is stable.

The converse direction of Theorem 5.5 is considerably more technical, although the idea is simple. With F as in the theorem, for each $n \geq 1$ let

$$F_n(\mathbf{x}, \mathbf{y}) = \sum_{\alpha \leq n\mathbf{1}} (n\mathbf{1})_\alpha P_\alpha(\mathbf{x}) \frac{\mathbf{y}^\alpha}{n^\alpha}.$$

The sequence $(F_n : n \geq 1)$ converges to F , since for each $\alpha \in \mathbb{N}^m$, $n^{-\alpha} (n\mathbf{1})_\alpha \rightarrow 1$ as $n \rightarrow \infty$. Each F_n is stable, by hypothesis (and scaling). The hard work is involved with showing that the convergence is uniform on compact sets. To do this, Borcea and Brändén develop a very flexible multivariate generalization of the Szasz Principle [5, Theorem 5.6] – in itself an impressive accomplishment. Unfortunately, we have no space here to develop this result – see Section 5.2 of [5].

6. JOHNSON'S CONJECTURES.

Let $\mathcal{A} = (A_1, \dots, A_k)$ be a k -tuple of n -by- n matrices. Define the *mixed determinant* of \mathcal{A} to be

$$\text{Det}(\mathcal{A}) = \text{Det}(A_1, \dots, A_k) = \sum_{(S_1, \dots, S_k)} \prod_{i=1}^k \det A_i[S_i],$$

in which the sum is over all ordered sequences of k pairwise disjoint subsets of $[n]$ such that $[n] = S_1 \cup \dots \cup S_k$, and $A_i[S_i]$ is the principal submatrix of A_i supported on rows and columns in S_i . Let $A_i(S_i)$ be the complementary principal submatrix supported on rows and columns not in S_i , and for $j \in [n]$ let $A_i(j) = A_i(\{j\})$.

For example, when $k = 2$ and $A_1 = xI$ and $A_2 = -B$, this specializes to $\text{Det}(xA_1, -B) = \det(xI - B)$, the characteristic polynomial of B . In the late 1980s, Johnson made three conjectures about the $k = 2$ case more generally.

Johnson's Conjectures. *Let A and B be n -by- n matrices, with A positive definite and B Hermitian.*

- (a) *Then $\text{Det}(xA, -B)$ has only real roots.*
- (b) *For $j \in [n]$, the roots of $\text{Det}(xA(j), -B(j))$ interlace those of $\text{Det}(xA, -B)$.*
- (c) *The inertia of $\text{Det}(xA, -B)$ is the same as that of $\det(xI - B)$.*

In part (c), the *inertia* of a univariate real stable polynomial p is the triple $\iota(p) = (\iota_-(p), \iota_0(p), \iota_+(p))$ with entries the number of negative, zero, or positive roots of p , respectively.

In 2008, Borcea and Brändén [1] proved all three of these statements in much greater generality.

Theorem 6.1 (Theorem 2.6 of [1]). *Fix integers $\ell, m, n \geq 1$. For $h \in [\ell]$ and $i \in [m]$ let B_h and A_{hi} be n -by- n matrices, and let*

$$L_h = \sum_{i=1}^m x_i A_{hi} + B_h.$$

- (a) *If all the A_{hi} are positive semidefinite and all the B_h are Hermitian, then $\text{Det}(\mathcal{L}) = \text{Det}(L_1, \dots, L_\ell) \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}]$ is real stable.*
- (b) *For each $j \in [n]$, let $\mathcal{L}(j) = (L_1(j), \dots, L_\ell(j))$. With the hypotheses of part (a), the polynomial $\text{Det}(\mathcal{L}) + y \text{Det}(\mathcal{L}(j)) \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}, y]$ is real stable.*

Proof. Let $Y = \text{diag}(y_1, \dots, y_n)$ be a diagonal matrix of indeterminates. By Proposition 2.1, for each $h \in [\ell]$ the polynomial

$$\det(Y + L_h) = \sum_{S \subseteq [n]} \mathbf{y}^S \det L_h(S)$$

is real stable in $\mathfrak{S}_{\mathbb{R}}[\mathbf{x}, \mathbf{y}]$. By inversion of all the \mathbf{y} indeterminates, each

$$\det(I - YL_h) = \sum_{S \subseteq [n]} (-1)^{|S|} \mathbf{y}^S \det L_h[S]$$

is real stable. Since $\prod_{h=1}^{\ell} \det(I - YL_h)$ is real stable, contraction and specialization imply that

$$\text{Det}(\mathcal{L}) = (-1)^n \frac{\partial^n}{\partial y_1 \cdots \partial y_n} \prod_{h=1}^{\ell} \det(I - YL_h) \Big|_{\mathbf{y}=\mathbf{0}}$$

is real stable, proving part (a).

For part (b), let V be the n -by- n matrix with all entries zero except for $V_{jj} = y$. By part (a),

$$\text{Det}(V, L_1, \dots, L_h) = \text{Det}(\mathcal{L}) + y \text{Det}(\mathcal{L}(j))$$

is real stable. \square

Theorem 6.1 (with Corollary 2.10) clearly settles Conjectures (a) and (b).

Proof of Conjecture (c). Let A and B be n -by- n matrices with A positive definite and B Hermitian. Let $(\iota_-, \iota_0, \iota_+)$ be the inertia of $\det(xI - B)$. Let $f(x) = \text{Det}(xA, -B)$, and let (ν_-, ν_0, ν_+) be the inertia of f .

We begin by showing that $\nu_0 = \iota_0$. Since $\iota_0 = \min\{|S| : S \subseteq [n] \text{ and } \det(B(S)) \neq 0\}$, it follows that $\nu_0 \geq \iota_0$. The constant term of $f(x)$ is $(-1)^n \det(B)$, so that if $\iota_0 = 0$ then $\nu_0 = 0$. If $\iota_0 = k > 0$ then let $S = \{s_1, \dots, s_k\} \subseteq [n]$ be such that $\det(B(S)) \neq 0$. For $0 \leq i \leq k$ let $f_i(x) = \text{Det}(A(\{s_1, \dots, s_i\}), -B(\{s_1, \dots, s_i\}))$, so that $f_0(x) = f(x)$. By Theorem 6.1, the roots of f_{i-1} and of f_i are interlaced, for each $i \in [k]$. Thus,

$$\nu_0 = \iota_0(f_0) \leq \iota_0(f_1) + 1 \leq \iota_0(f_2) + 2 \leq \cdots \leq \iota_0(f_k) + k = k = \iota_0,$$

since $\iota_0(f_k) = 0$ because $\det(B(S)) \neq 0$. Therefore $\nu_0 = \iota_0$.

For any positive definite matrix A , $\text{Det}(xA, -B)$ is a polynomial of degree n . Suppose that A is such a matrix for which $\nu_+ \neq \iota_+$. Consider the matrices $A_\lambda = (1 - \lambda)I + \lambda A$ for $\lambda \in [0, 1]$. Each of these matrices is positive definite. From the paragraph above, each of the polynomials $g_\lambda(x) = \text{Det}(xA_\lambda, -B)$ has $\iota_0(g_\lambda) = \iota_0$. Since $\iota_+(g_0) = \iota_+ \neq \nu_+ = \iota_+(g_1)$ and the roots of g_λ vary continuously with λ , there is some value $\mu \in (0, 1)$ for which $\iota_0(g_\mu) > \iota_0$. This contradiction shows that $\nu_+ = \iota_+$, and hence $\nu_- = \iota_-$ as well. \square

Borcea and Brändén [1] proceed to derive many inequalities for the principal minors of positive semidefinite matrices, and some for merely Hermitian matrices. These are applications of inequalities valid more generally for real stable polynomials. The simplest of these inequalities are as follows.

For an n -by- n matrix A , the j -th *symmetrized Fisher product* is

$$\sigma_j(A) = \sum_{S \subseteq [n]: |S|=j} \det(A[S]) \det(A(S)).$$

and the j -th *averaged Fisher product* is $\hat{\sigma}_j(A) = \binom{n}{j}^{-1} \sigma_j(A)$. Notice that $\sigma_j(A) = \sigma_{n-j}(A)$ for all $0 \leq j \leq n$.

Corollary 6.2. *Let A be an n -by- n positive semidefinite matrix.*

(a) *Then $\hat{\sigma}_j(A)^2 \geq \hat{\sigma}_{j-1}(A) \hat{\sigma}_{j+1}(A)$ for all $1 \leq j \leq n - 1$.*

(b) *Also, $\hat{\sigma}_0(A) \leq \hat{\sigma}_1(A) \leq \cdots \leq \hat{\sigma}_{\lfloor n/2 \rfloor}$.*

(c) *If A is positive definite and $\det(A) = d$ then*

$$\frac{\hat{\sigma}_1(A)}{d} \geq \left(\frac{\hat{\sigma}_2(A)}{d} \right)^{1/2} \geq \left(\frac{\hat{\sigma}_3(A)}{d} \right)^{1/3} \geq \cdots \geq \left(\frac{\hat{\sigma}_n(A)}{d} \right)^{1/n} = 1.$$

Proof. It suffices to consider positive definite A . By Theorem 6.1, the polynomial $\text{Det}(xA, -A) = \sum_{j=0}^n (-1)^j \sigma_j(A) x^j$ has only real roots, and these roots are all positive. Part (a) follows from Newton's Inequalities [12, Theorem 51]. Part (a) and the symmetry $\sigma_j(A) = \sigma_{n-j}(A)$ for all $0 \leq j \leq n$ imply part (b). Part (c) follows from Maclaurin's Inequalities [12, Theorem 52]. \square

7. THE SYMMETRIC EXCLUSION PROCESS.

This section summarizes an application of stable polynomials to probability and statistical mechanics from a 2009 paper of Borcea, Brändén and Liggett [7].

Let Λ be a set of *sites*. A symmetric exclusion process (SEP) is a type of Markov chain with state space a subset of $\{0, 1\}^\Lambda$. In a state $S : \Lambda \rightarrow \{0, 1\}$, the sites in $S^{-1}(1)$ are *occupied* and the sites in $S^{-1}(0)$ are *vacant*. This is meant to model a physical system of particles interacting by means of hard-core exclusions. Such models come in many varieties – to avoid technicalities we discuss only the case of a finite system Λ and continuous time t . (The results of this section extend to countable Λ under a reasonable finiteness condition on the interaction rates.) Symmetry of the interactions turns out to be crucial, but particle number conservation is unimportant.

Let E be a set of two-element subsets of Λ . For each $\{i, j\} \in E$, let $\lambda_{ij} > 0$ be a positive real, and let $\tau_{ij} : \Lambda \rightarrow \Lambda$ be the permutation that exchanges i and j and fixes all other sites. Our SEP Markov chain \mathcal{M} proceeds as follows. Each $\{i, j\} \in E$ has a Poisson process “clock” of rate λ_{ij} , and these are independent of one another. With probability one, no two clocks ever ring at the same time. When the clock of $\{i, j\}$ rings, the current state S is updated to the new state $S \circ \tau_{ij}$. In other words, when the $\{i, j\}$ clock rings, if exactly one of the sites $\{i, j\}$ is occupied then a particle hops from the occupied to the vacant of these two sites.

Let $\Lambda = [m]$ and $\Omega = \{0, 1\}^\Lambda$, let φ_0 be an initial probability distribution on Ω , and let φ_t be the distribution of the state of \mathcal{M} , starting at φ_0 , after evolving for time $t \geq 0$. We are concerned with properties of the distribution φ_t that hold for all $t \geq 0$.

7.1. Negative correlation and negative association. Consider a probability distribution φ on Ω . An *event* \mathcal{E} is any subset of Ω . The probability of the event \mathcal{E} is $\Pr[\mathcal{E}] = \sum_{S \in \mathcal{E}} \varphi(S)$. An event \mathcal{E} is *increasing* if whenever $S \leq S'$ in Ω and $S \in \mathcal{E}$, then $S' \in \mathcal{E}$. For example, if K is any subset of Λ and \mathcal{E}_K is the event that all sites in K are occupied, then \mathcal{E}_K is an increasing event. Notice that this event has the form $\mathcal{E}_K = \mathcal{E}' \times \{0, 1\}^{\Lambda \setminus K}$ for some event $\mathcal{E}' \subseteq \{0, 1\}^K$. Two events \mathcal{E} and \mathcal{F} are *disjointly supported* when one can partition $\Lambda = A \cup B$ with $A \cap B = \emptyset$ and $\mathcal{E} = \mathcal{E}' \times \{0, 1\}^B$ and $\mathcal{F} = \{0, 1\}^A \times \mathcal{F}'$ for some events $\mathcal{E}' \subseteq \{0, 1\}^A$ and $\mathcal{F}' \subseteq \{0, 1\}^B$.

A probability distribution on Ω is *negatively associated* (NA) when $\Pr[\mathcal{E} \cap \mathcal{F}] \leq \Pr[\mathcal{E}] \cdot \Pr[\mathcal{F}]$ for any two increasing events that are disjointly supported. It is *negatively correlated* (NC) when $\Pr[\mathcal{E}_{\{i,j\}}] \leq \Pr[\mathcal{E}_{\{i\}}] \cdot \Pr[\mathcal{E}_{\{j\}}]$ for any two distinct sites $\{i, j\} \subseteq \Lambda$. Clearly NA implies NC.

It is useful to find conditions under which NC implies NA, since NC is so much easier to check. The following originates with Feder and Mihail, but many others have contributed their insights – see Section 4.2 of [7]. The *partition function* of

any $\varphi : \Omega \rightarrow \mathbb{R}$ is the real multiaffine polynomial

$$Z(\varphi) = Z(\varphi; \mathbf{x}) = \sum_{S \in \Omega} \varphi(S) \mathbf{x}^S$$

in $\mathbb{R}[\mathbf{x}]^{\text{MA}}$. If φ is nonzero and nonnegative, then for any $\mathbf{a} \in \mathbb{R}^\Lambda$ with $\mathbf{a} > \mathbf{0}$, this defines a probability distribution $\varphi^{\mathbf{a}} : \Omega \rightarrow [0, 1]$ by setting $\varphi^{\mathbf{a}}(S) = \varphi(S) \mathbf{a}^S / Z(\varphi; \mathbf{a})$ for all $S \in \Omega$.

Feder-Mihail Theorem (Theorem 4.8 of [7]). *Let \mathcal{S} be a class of nonzero non-negative functions satisfying the following conditions.*

- (i) *Each $\varphi \in \mathcal{S}$ has domain $\{0, 1\}^\Lambda$ for some finite set $\Lambda = \Lambda(\varphi)$.*
 - (ii) *For each $\varphi \in \mathcal{S}$, $Z(\varphi)$ is a homogeneous polynomial.*
 - (iii) *For each $\varphi \in \mathcal{S}$ and $i \in \Lambda(\varphi)$, $Z(\varphi)|_{x_i=0}$ and $\partial_i Z(\varphi)$ are partition functions of members of \mathcal{S} .*
 - (iv) *For each $\varphi \in \mathcal{S}$ and $\mathbf{a} \in \mathbb{R}^{\Lambda(\varphi)}$ with $\mathbf{a} > \mathbf{0}$, $\varphi^{\mathbf{a}}$ is NC.*
- Then for every $\varphi \in \mathcal{S}$ and $\mathbf{a} \in \mathbb{R}^{\Lambda(\varphi)}$ with $\mathbf{a} > \mathbf{0}$, $\varphi^{\mathbf{a}}$ is NA.*

7.2. A conjecture of Liggett and Pemantle. In the early 2000s, Liggett and Pemantle arrived independently at the following conjecture, now a theorem.

Theorem 7.1 (Theorem 5.2 of [7]). *If the initial distribution φ_0 of a SEP is deterministic (i.e. concentrated on a single state) then φ_t is NA for all $t \geq 0$.*

Proof. This amounts to finding a class \mathfrak{S} of probability distributions such that:

- (1) deterministic distributions are in \mathfrak{S} ,
- (2) being in \mathfrak{S} implies NA, and
- (3) time evolution of the SEP preserves membership in \mathfrak{S} .

Borcea, Brändén, and Liggett [7] identified such a class: φ is in \mathfrak{S} if and only if the partition function $Z(\varphi)$ is homogeneous, multiaffine, and real stable. (Notice that if φ is in \mathfrak{S} then $\varphi^{\mathbf{a}}$ is in \mathfrak{S} for all $\mathbf{a} \in \mathbb{R}^\Lambda$ with $\mathbf{a} > \mathbf{0}$, by scaling.) We proceed to check the three claims above.

Claim (1) is trivial, since if $\varphi(S) = 1$ then $Z(\varphi) = \mathbf{x}^S$, which is clearly homogeneous, multiaffine, and real stable.

To check claim (2) we verify the hypotheses of the Feder-Mihail Theorem. Hypotheses (i) and (ii) hold since $Z(\varphi)$ is multiaffine and homogeneous. By specialization and contraction, (iii) holds. To check (iv), let $\mathbf{a} \in \mathbb{R}^\Lambda$ with $\mathbf{a} > \mathbf{0}$, let $\{i, j\} \subseteq \Lambda$, and consider the probability distribution $\varphi^{\mathbf{a}}$ on Ω . The occupation probability for site i is

$$\Pr[\mathcal{E}_{\{i\}}] = \sum_{S \in \{0,1\}^\Lambda: S(i)=1} \frac{\varphi(S) \mathbf{a}^S}{Z(\varphi; \mathbf{a})} = a_i \frac{\partial_i Z(\varphi; \mathbf{a})}{Z(\varphi; \mathbf{a})},$$

and similarly for $\Pr[\mathcal{E}_{\{j\}}]$. Likewise, $\Pr[\mathcal{E}_{\{i,j\}}] = a_i a_j Z(\varphi; \mathbf{a})^{-1} \cdot \partial_i \partial_j Z(\varphi; \mathbf{a})$. Now

$$\Pr[\mathcal{E}_{\{i,j\}}] - \Pr[\mathcal{E}_{\{i\}}] \cdot \Pr[\mathcal{E}_{\{j\}}] = -\frac{a_i a_j}{Z(\varphi; \mathbf{a})^2} \cdot \Delta_{ij} Z(\varphi; \mathbf{a}) \leq 0,$$

by Theorem 3.1. Thus $\varphi^{\mathbf{a}}$ is NC. By the Feder-Mihail Theorem, every φ in \mathfrak{S} is NA.

To check claim (3) we need some of the theory of continuous time Markov chains. The time evolution of a Markov chain \mathcal{M} with finite state space Ω is governed by a one-parameter semigroup $T(t)$ of transformations of \mathbb{R}^Ω . For a function $F \in \mathbb{R}^\Omega$

and time $t \geq 0$ and state $S \in \Omega$, $(T(t)F)(S)$ is the expected value of F at time t , given that the initial distribution of \mathcal{M} is concentrated at S with probability one at time 0. In particular, $\varphi_t = T(t)\varphi_0$ for all $t \geq 0$, and all initial distributions φ_0 . In the case of the SEP we are considering, the infinitesimal generator \mathcal{L} of the semigroup $T(t)$ is given by

$$\mathcal{L} = \sum_{\{i,j\} \in E} \lambda_{ij} (\tau_{ij} - 1).$$

For each $\{i, j\} \in E$, this replaces each $S \in \Omega$ by $S \circ \tau_{ij}$ at the rate λ_{ij} .

In preparation for Section 7.3, it is useful to regard \mathcal{L} as an element of the real semigroup algebra $\mathfrak{A} = \mathbb{R}[\mathfrak{E}]$ of the semigroup \mathfrak{E} of all endofunctions $f : \Omega \rightarrow \Omega$ (with the operation of functional composition). The left action of \mathfrak{E} on Ω is extended to a left action of \mathfrak{A} on $\mathbb{C}[\mathbf{x}]$ as usual: for $f \in \mathfrak{E}$ and $S \in \Omega$, $f(\mathbf{x}^S) = \mathbf{x}^{f(S)}$, extended bilinearly to all of \mathfrak{A} and $\mathbb{C}[\mathbf{x}]$. A permutation $\sigma \in \mathfrak{S}(\Omega)$ is identified with the endofunction $f_\sigma : S \mapsto S \circ \sigma^{-1}$, so this action of \mathfrak{A} agrees with the action of $\mathfrak{S}(m)$ in Section 4.2. A left action of \mathfrak{A} on \mathbb{R}^Ω is defined by $Z(f(F)) = f(Z(F))$ for all $f \in \mathfrak{E}$ and $F \in \mathbb{R}^\Omega$, and linear extension. More explicitly, for $f \in \mathfrak{E}$, $F \in \mathbb{R}^\Omega$, and $S \in \Omega$,

$$(f(F))(S) = F(f^{-1}(S)) = \sum \{F(S') : S' \in \Omega \text{ and } f(S') = S\}.$$

Consider an element of \mathfrak{A} of the form $\mathcal{L} = \sum_{i=1}^N \lambda_i (f_i - 1)$ with all $\lambda_i > 0$. Let $\lambda_i \leq L$ for all $i \in [N]$, and let $K = \sum_{i=1}^N \lambda_i$. The power series

$$\exp(t\mathcal{L}) = e^{-Kt} \sum_{n=0}^{\infty} \frac{t^n}{n!} \left[\sum_{i=1}^N \lambda_i f_i \right]^n = \sum_{f \in \mathfrak{E}} P_f(t) \cdot f$$

in $\mathfrak{A}[[t]]$ is such that for each $f \in \mathfrak{E}$, $P_f(t) \in \mathbb{R}[[t]]$ is dominated coefficientwise by $\exp((LN - K)t)$. Thus $\exp(t\mathcal{L}) \in \mathfrak{A}[[t]]$ converges for all $t \geq 0$. The semigroup of transformations generated by \mathcal{L} is $\exp(t\mathcal{L})$.

To check claim (3) we will show that the semigroup $T(t)$ of the SEP preserves stability for all $t \geq 0$: that is, if $Z(\varphi_0)$ is stable then $Z(\varphi_t) = T(t)Z(\varphi_0)$ is stable for all $t \geq 0$. This reduces to the case of a single pair $\{i, j\} \in E$, as follows. If \mathcal{M}_1 and \mathcal{M}_2 are Markov chains on the same finite state space, with semigroups $T_1(t)$ and $T_2(t)$ generated by \mathcal{L}_1 and \mathcal{L}_2 , then the semigroup generated by $\mathcal{L}_1 + \mathcal{L}_2$ is

$$T(t) = \lim_{n \rightarrow \infty} [T_1(t/n)T_2(t/n)]^n,$$

by the Trotter product formula. By Hurwitz's Theorem, It follows that if $T_i(t)$ preserves stability for all $t \geq 0$ and $i \in \{1, 2\}$, then $T(t)$ preserves stability for all $t \geq 0$. By repeated application of this argument, in order to show that the SEP semigroup $T(t) = \exp(t\mathcal{L})$ preserves stability for all $t \geq 0$ it is enough to show that for each $\{i, j\} \in E$, $T_{ij}(t) = \exp(t\lambda_{ij}(\tau_{ij} - 1))$ preserves stability for all $t \geq 0$. Now, since $\tau_{ij}^2 = 1$,

$$T_{ij}(t) = \left(\frac{1 + e^{-2\lambda_{ij}t}}{2} \right) + \left(\frac{1 - e^{-2\lambda_{ij}t}}{2} \right) \cdot \tau_{ij}.$$

By Lemma 4.3, this preserves stability for all $t \geq 0$. This proves Theorem 7.1. \square

7.3. Further observations. In verifying the hypotheses of the Feder-Mihail Theorem we used the fact that if $f \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}]^{\text{MA}}$ is multiaffine and real stable, then $\Delta_{ij}f(\mathbf{a}) \geq 0$ for all $\{i, j\} \subseteq E$ and $\mathbf{a} \in \mathbb{R}^m$, by Theorem 3.1. In fact, we only needed the weaker hypothesis that $\Delta_{ij}f(\mathbf{a}) \geq 0$ for all $\{i, j\} \subseteq E$ and $\mathbf{a} \in \mathbb{R}^m$ with $\mathbf{a} > \mathbf{0}$. A multiaffine real polynomial satisfying this weaker condition is a *Rayleigh* polynomial. (This terminology is by analogy with the Rayleigh monotonicity property of electrical networks – see Definition 2.5 of [7] and the references cited there. Multiaffine real stable polynomials are also called *strongly Rayleigh*.) The class of probability distributions φ such that $Z(\varphi)$ is homogeneous, multiaffine, and Rayleigh meets all the conditions of the Feder-Mihail Theorem. It follows that all such distributions are NA.

Claim (2) above can be generalized in another way – the hypothesis of homogeneity can be removed, as follows. Let $\mathbf{y} = (y_1, \dots, y_m)$ and let $e_j(\mathbf{y})$ be the j -th elementary symmetric function of the \mathbf{y} . Given a multiaffine polynomial $f = \sum_{S \subseteq [m]} c(S)\mathbf{x}^S$, the *symmetric homogenization* of f is the polynomial $f_{\text{sh}}(\mathbf{x}, \mathbf{y}) \in \mathbb{C}[\mathbf{x}, \mathbf{y}]^{\text{MA}}$ defined by

$$f_{\text{sh}}(\mathbf{x}, \mathbf{y}) = \sum_{S \subseteq [m]} c(S)\mathbf{x}^S \binom{m}{|S|}^{-1} e_{m-|S|}(\mathbf{y}).$$

Note that f_{sh} is homogeneous of degree m , and $f_{\text{sh}}(\mathbf{x}, \mathbf{1}) = f(\mathbf{x})$.

Proposition 7.2 (Theorem 4.2 of [7]). *If $f \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}]^{\text{MA}}$ is multiaffine and real stable then $f_{\text{sh}} \in \mathfrak{S}_{\mathbb{R}}[\mathbf{x}, \mathbf{y}]^{\text{MA}}$ is homogeneous, multiaffine and real stable.*

(We omit the proof.)

Corollary 7.3 (Theorem 4.9 of [7]). *Let $\varphi : \Omega \rightarrow [0, \infty)$ be such that $Z(\varphi)$ is nonzero, multiaffine, and real stable. Then for all $\mathbf{a} \in \mathbb{R}^m$ with $\mathbf{a} > \mathbf{0}$, $\varphi^{\mathbf{a}}$ is NA.*

Proof. By Proposition 7.2, $Z_{\text{sh}}(\varphi; \mathbf{x}, \mathbf{y})$ is nonzero, homogeneous, multiaffine, and real stable. This is the partition function for $\psi : \{0, 1\}^{[2m]} \rightarrow [0, \infty)$ given by $\psi(S) = \binom{m}{|S \cap [m]|}^{-1} \varphi(S \cap [m])$. By claim (2) above, $\psi^{\mathbf{a}}$ is NA for all $\mathbf{a} \in \mathbb{R}^{2m}$ with $\mathbf{a} > \mathbf{0}$. By considering those $\mathbf{a} \in \mathbb{R}^{2m}$ for which $a_i = 1$ for all $m+1 \leq i \leq 2m$, it follows that $\varphi^{\mathbf{a}}$ is NA for all $\mathbf{a} \in \mathbb{R}^m$ with $\mathbf{a} > \mathbf{0}$. \square

Corollary 7.4 (Theorem 5.2 of [7]). *If the initial distribution φ_0 of a SEP is such that $Z(\varphi)$ is stable (but not necessarily homogeneous), then $Z(\varphi_t)$ is stable, and hence φ_t is NA, for all $t \geq 0$.*

It is natural to try extending these results to asymmetric exclusion processes. For $(i, j) \in \Lambda^2$ define $\mathbf{t}_{ij} \in \mathfrak{E}$ by $\mathbf{t}_{ij}(S) = S \circ \tau_{ij}$ if $S(i) = 1$ and $S(j) = 0$, and $\mathbf{t}_{ij}(S) = S$ otherwise, for all $S \in \Omega$. That is, \mathbf{t}_{ij} makes a particle hop from site i to site j , if possible. Let E be a set of ordered pairs in Λ^2 , and for $(i, j) \in E$ let $\lambda_{ij} > 0$. An asymmetric exclusion process is a Markov chain on Ω with semigroup $T(t) = \exp(t\mathcal{L})$ generated by something of the form

$$\mathcal{L} = \sum_{(i,j) \in E} \lambda_{ij} (\mathbf{t}_{ij} - 1).$$

By the argument for claim (3) above, in order to show that $T(t)$ preserves stability for all $t \geq 0$, it suffices to do so for the two-site semigroup $T_{\{1,2\}}(t) =$

$\exp(t\mathcal{L}_{\{1,2\}})$ generated by

$$\mathcal{L}_{\{1,2\}} = \lambda_{12}(\mathbf{t}_{12} - 1) + \lambda_{21}(\mathbf{t}_{21} - 1).$$

Exercise 7.5 (Strengthening Remark 5.3 of [7]). With the notation above, let $\lambda = \lambda_{12} + \lambda_{21}$, $\beta_{12} = \lambda_{12}/\lambda$, and $\beta_{21} = \lambda_{21}/\lambda$.

- (a) In \mathfrak{A} , $\mathbf{t}_{12} + \mathbf{t}_{21} = 1 + \tau_{12}$.
- (b) If ω is any word in $\{\mathbf{t}_{12}, \mathbf{t}_{21}\}^n$, then $\mathbf{t}_{12}\omega = \mathbf{t}_{12}$ and $\mathbf{t}_{21}\omega = \mathbf{t}_{21}$.
- (c) The semigroup generated by $\mathcal{L}_{\{1,2\}}$ is

$$T_{\{1,2\}}(t) = e^{-\lambda t} + (1 - e^{-\lambda t})(\beta_{12}\mathbf{t}_{12} + \beta_{21}\mathbf{t}_{21})$$

- (d) The semigroup $T_{\{1,2\}}(t)$ preserves stability for all $t \geq 0$ if and only if $\beta_{12} = \beta_{21} = 1/2$, in which case it reduces to the SEP (of rate $\lambda/2$).

Thus, even the slightest asymmetry ruins preservation of stability by the SEP!

Finally, we consider a SEP in which particle number is not conserved. For $i \in \Lambda$ define $\mathbf{a}_i, \mathbf{a}_i^* \in \mathfrak{E}$ as follows: for $S \in \Omega$ and $j \in \Lambda$, let $(\mathbf{a}_i(S))(j) = (\mathbf{a}_i^*(S))(j) = S(j)$ if $j \neq i$, and $(\mathbf{a}_i(S))(i) = 0$ and $(\mathbf{a}_i^*(S))(i) = 1$. That is, \mathbf{a}_i annihilates a particle at site i , and \mathbf{a}_i^* creates a particle at site i , if possible.

A SEP with particle creation and annihilation is a Markov chain on Ω with semigroup $T(t) = \exp(t\mathcal{L})$ generated by something of the form

$$\mathcal{L} = \sum_{\{i,j\} \in E} \lambda_{ij}(\tau_{ij} - 1) + \sum_{i \in \Lambda} [\theta_i(\mathbf{a}_i - 1) + \theta_i^*(\mathbf{a}_i^* - 1)],$$

in which the first sum is the generator of the SEP in Theorem 7.1 and $\theta_i, \theta_i^* \geq 0$ for each $i \in \Lambda$.

By the argument for claim (3) above, to show that this $T(t)$ preserves stability for all $t \geq 0$, it suffices to do so for the one-site semigroups generated by $\mathcal{L}_1 = \theta(\mathbf{a}_1 - 1)$ and $\mathcal{L}_1^* = \theta(\mathbf{a}_1^* - 1)$, respectively.

Exercise 7.6. The semigroups generated by \mathcal{L}_1 and \mathcal{L}_1^* are

$$T_1(t) = e^{-\theta t} + (1 - e^{-\theta t})\mathbf{a}_1 \quad \text{and} \quad T_1^*(t) = e^{-\theta t} + (1 - e^{-\theta t})\mathbf{a}_1^*,$$

respectively. Both $T_1(t)$ and $T_1^*(t)$ preserve stability.

Corollary 7.7. *If the initial distribution φ_0 of a SEP with particle creation and annihilation is such that $Z(\varphi)$ is stable, then $Z(\varphi_t)$ is stable, and hence φ_t is NA, for all $t \geq 0$.*

8. INEQUALITIES FOR MIXED DISCRIMINANTS.

This section summarizes a powerful application of stable polynomials from a 2008 paper of Gurvits [11].

We will use without mention the facts that log and exp are strictly increasing functions on $(0, \infty)$. A function $\rho : I \rightarrow \mathbb{R}$ defined on an interval $I \subseteq \mathbb{R}$ is *convex* provided that for all $a_1, a_2 \in I$, $\rho((a_1 + a_2)/2) \leq (\rho(a_1) + \rho(a_2))/2$. It is *strictly convex* if it is convex and equality holds here only when $a_1 = a_2$. A function $\rho : I \rightarrow \mathbb{R}$ is *(strictly) concave* if $-\rho$ is (strictly) convex. For example, for positive reals $a_1, a_2 > 0$ one has $(\sqrt{a_1} - \sqrt{a_2})^2 \geq 0$, with equality only if $a_1 = a_2$. It follows that $\log((a_1 + a_2)/2) \geq (\log(a_1) + \log(a_2))/2$, with equality only if $a_1 = a_2$. That is, log is strictly concave.

Jensen's Inequality (Theorem 90 of [12]). *Let $\rho : I \rightarrow \mathbb{R}$ be defined on an interval $I \subseteq \mathbb{R}$, let $a_i \in I$ for $i \in [n]$, and let $b_i > 0$ for $i \in [n]$ be such that $\sum_{i=1}^n b_i = 1$. If ρ is convex then*

$$\rho\left(\sum_{i=1}^n b_i a_i\right) \leq \sum_{i=1}^n b_i \rho(a_i).$$

If ρ is strictly convex and equality holds, then $a_1 = a_2 = \dots = a_n$.

For integer $d \geq 1$, let $G(d) = (1 - 1/d)^{d-1}$, and let $G(0) = 1$. Note that $G(1) = 0^0 = 1$, and that $G(d)$ is a strictly decreasing function for $d \geq 1$. For homogeneous $f \in \mathbb{R}[\mathbf{x}]$ with nonnegative coefficients, define the *capacity* of f to be

$$\text{cap}(f) = \inf_{\mathbf{c} > \mathbf{0}} \frac{f(\mathbf{c})}{c_1 \cdots c_m},$$

with the infimum over the set of all $\mathbf{c} \in \mathbb{R}^m$ with $c_i > 0$ for all $i \in [m]$.

Lemma 8.1 (Lemma 3.2 of [11]). *Let $f = \sum_{i=0}^d b_i x^i \in \mathbb{R}[x]$ be a nonzero univariate polynomial of degree d with nonnegative coefficients. If f is real stable then $b_1 = f'(0) \geq G(d)\text{cap}(f)$, and if $\text{cap}(f) > 0$ then equality holds if and only if $d \leq 1$ or $f(x) = b_d(x + \xi)^d$ for some $\xi > 0$.*

Proof. If $\text{cap}(f) = 0$ then there is nothing to prove, so assume that $\text{cap}(f) > 0$. If $d = 0$ then $f'(0) = b_1 = 0 = G(0)\text{cap}(f)$, and if $d = 1$ then $f'(0) = b_1 = G(1)\text{cap}(f)$, so assume that $d \geq 2$. If $f(0) = 0$ then $f'(0) = \lim_{c \rightarrow 0} f(c)/c \geq \text{cap}(f) > G(d)\text{cap}(f)$. Thus, assume that $d \geq 2$ and $f(0) = b_0 > 0$. We may rescale the polynomial so that $b_0 = 1$. Now there are $a_i > 0$ for $i \in [d]$ such that

$$f(x) = \prod_{i=1}^d (1 + a_i x),$$

and $b_1 = a_1 + \dots + a_d$. For any $c > 0$ we have

$$\frac{\log(\text{cap}(f)c)}{d} \leq \frac{\log(f(c))}{d} = \frac{1}{d} \sum_{i=1}^d \log(1 + a_i c) \leq \log\left(1 + \frac{b_1 c}{d}\right),$$

by Jensen's Inequality. It follows that $\text{cap}(f)c \leq (1 + b_1 c/d)^d$ for all $c > 0$. Let $g(x) = (1 + b_1 x/d)^d$. Elementary calculus shows that

$$\text{cap}(g) = \inf_{c > 0} \frac{g(c)}{c} = \frac{g(c_*)}{c_*} = \frac{b_1}{G(d)}, \quad \text{in which } c_* = \frac{d}{b_1(d-1)}.$$

Since $\text{cap}(f) \leq \text{cap}(g)$, this yields the stated inequality. If equality holds, then equality holds in the application of Jensen's Inequality, and so f has the stated form. \square

Lemma 8.2 (Theorem 4.10 of [11]). *Let $f \in \mathfrak{S}_{\mathbb{R}}[x_1, \dots, x_m]$ be real stable, with nonnegative coefficients, and homogeneous of degree m . Let $g = \partial_m f|_{x_m=0}$. Then*

$$\text{cap}(g) \geq G(\deg_m(f))\text{cap}(f).$$

Proof. We may assume that $d = \deg_m(f) \geq 1$. Let $c_i > 0$ for $i \in [m-1]$, and let $p_{\mathbf{c}}(x) = f(c_1, \dots, c_{m-1}, x)$. Since f has nonnegative coefficients, $p_{\mathbf{c}} \neq 0$. As in the proof of Lemma 2.4(f), $p_{\mathbf{c}}$ has degree d . By specialization, $p_{\mathbf{c}}$ is real stable. Lemma 8.1 implies that

$$g(\mathbf{c}) = p'_{\mathbf{c}}(0) \geq G(d)\text{cap}(p_{\mathbf{c}}) \geq G(d)\text{cap}(f)$$

for all $\mathbf{c} \in \mathbb{R}^{m-1}$ with $\mathbf{c} > \mathbf{0}$. If $m = 1$ then $g = \text{cap}(g)$ is a constant. If $m \geq 2$ then for any such \mathbf{c} let $b = (c_1 \cdots c_{m-1})^{-1/(m-1)}$. Since g is homogeneous of degree $m - 1$,

$$\frac{g(\mathbf{c})}{c_1 \cdots c_{m-1}} = g(bc_1, \dots, bc_{m-1}) \geq G(d)\text{cap}(f).$$

It follows that $\text{cap}(g) \geq G(d)\text{cap}(f)$. \square

Theorem 8.3 (Theorem 2.4 of [11]). *Let $f \in \mathfrak{S}_{\mathbb{R}}[x_1, \dots, x_m]$ be real stable, with nonnegative coefficients, and homogeneous of degree m . Let $\deg_i(f) = d_i$ and $e_i = \min\{i, d_i\}$ for each $i \in [m]$. Then*

$$\partial^1 f(\mathbf{0}) \geq \text{cap}(f) \prod_{i=2}^m G(e_i).$$

Proof. Let $g_m = f$ and let $g_{i-1} = \partial_i g_i|_{x_i=0}$ for all $i \in [m]$. By contraction and specialization, g_i is real stable for each $i \in [m]$. Notice that $g_0 = \partial^1 f(\mathbf{0}) = \text{cap}(g_0)$. By Lemma 8.2. $\text{cap}(g_{i-1}) \geq \text{cap}(g_i) \cdot G(\deg_i g_i)$ for each $i \in [m]$. But $\deg_i g_i \leq \deg_i f = d_i$, and $\deg_i g_i$ is at most the total degree of g_i , which is i . Hence $\deg_i g_i \leq e_i$, and thus $G(\deg_i g_i) \geq G(e_i)$. Thus $\text{cap}(g_{i-1}) \geq \text{cap}(g_i) \cdot G(e_i)$ for each $i \in [m]$. Combining these inequalities (and $G(e_1) = 1$) gives the result. \square

With the notation of Theorem 8.3, since $e_i \leq i$ for all $i \in [m]$ and $G(d)$ is a decreasing function of d , one has the inequality

$$\prod_{i=2}^m G(e_i) \geq \prod_{i=2}^m G(i) = \prod_{i=2}^m \left(\frac{i-1}{i} \right)^{i-1} = \frac{m!}{m^m}.$$

Thus, the following corollary is immediate.

Corollary 8.4. *Let $f \in \mathfrak{S}_{\mathbb{R}}[x_1, \dots, x_m]$ be real stable, with nonnegative coefficients, and homogeneous of degree m . Then*

$$\partial^1 f(\mathbf{0}) \geq \frac{m!}{m^m} \cdot \text{cap}(f).$$

Theorem 8.5 (Theorem 5.7 of [11]). *Let $f \in \mathfrak{S}_{\mathbb{R}}[x_1, \dots, x_m]$ be real stable, with nonnegative coefficients, and homogeneous of degree m . Equality holds in the bound of Corollary 8.4 if and only if there are nonnegative reals $a_i \geq 0$ for $i \in [m]$ such that*

$$f(\mathbf{x}) = (a_1 x_1 + \cdots + a_m x_m)^m.$$

(We omit the proof.)

Lemma 8.6 (Fact 2.2 of [11]). *Let $f \in \mathbb{R}[x_1, \dots, x_m]$ be homogeneous of degree m , with nonnegative coefficients. Assume that $\partial_i f(\mathbf{1}) = 1$ for all $i \in [m]$. Then $\text{cap}(f) = 1$.*

Proof. Let $f = \sum_{\alpha} b(\alpha) \mathbf{x}^{\alpha}$, so that if $b(\alpha) \neq 0$ then $|\alpha| = \sum_{i=1}^m \alpha(i) = m$. By hypothesis, for all $i \in [m]$, $\sum_{\alpha} b(\alpha) \alpha(i) = 1$. Averaging these over all $i \in [m]$ yields $f(\mathbf{1}) = \sum_{\alpha} b(\alpha) = 1$, so that $\text{cap}(f) \leq 1$. Conversely, let $\mathbf{c} \in \mathbb{R}^m$ with $\mathbf{c} > \mathbf{0}$.

Jensen's Inequality implies that

$$\begin{aligned} \log(f(\mathbf{c})) &= \log\left(\sum_{\alpha} b(\alpha)\mathbf{c}^{\alpha}\right) \\ &\geq \sum_{\alpha} b(\alpha)\log(\mathbf{c}^{\alpha}) = \sum_{i=1}^m \log(c_i) \sum_{\alpha} b(\alpha)\alpha(i) = \log(c_1 \cdots c_m). \end{aligned}$$

It follows that $\text{cap}(f) \geq 1$. \square

Example 8.7 (van der Waerden Conjecture). An m -by- m matrix $A = (a_{ij})$ is *doubly stochastic* if all entries are nonnegative reals and every row and column sums to one. In 1926, van der Waerden conjectured that if A is an m -by- m doubly stochastic matrix then $\text{per}(A) \geq m!/m^m$, with equality if and only if $A = (1/m)J$, the m -by- m matrix in which every entry is $1/m$. In 1981 this lower bound was proved by Falikman, and the characterization of equality was proved by Egorychev. These results follow immediately from Corollary 8.4 and Theorem 8.5, as follows. It suffices to prove the result for an m -by- m doubly stochastic matrix $A = (a_{ij})$ with no zero entries, by a routine limit argument. The polynomial

$$f_A(\mathbf{x}) = \prod_{j=1}^m (a_{1j}x_1 + \cdots + a_{mj}x_m)$$

is clearly homogeneous and real stable, with nonnegative coefficients and of degree m , and such that $\deg_i(f_A) = m$ for all $i \in [m]$. Since A is doubly stochastic, Lemma 8.6 implies that $\text{cap}(f_A) = 1$. Since

$$\text{per}(A) = \partial^{\mathbf{1}} f_A(\mathbf{0}),$$

Corollary 8.4 and Theorem 8.5 imply the results of Falikman and Egorychev, respectively. Gurvits [11] also uses a similar argument to prove a refinement of the van der Waerden conjecture due to Schrijver and Valiant – see also [13].

Given n -by- n matrices A_1, \dots, A_m , the *mixed discriminant* of $\mathbf{A} = (A_1, \dots, A_m)$ is

$$\text{Disc}(\mathbf{A}) = \partial^{\mathbf{1}} \det(x_1 A_1 + \cdots + x_m A_m) \Big|_{\mathbf{x}=\mathbf{0}}.$$

This generalizes the permanent of an m -by- m matrix $B = (b_{ij})$ by considering the collection of matrices $\mathbf{A}(B) = (A_1, \dots, A_m)$ defined by $A_h = \text{diag}(a_{h1}, \dots, a_{hm})$ for each $h \in [m]$. In this case one sees that

$$\det(x_1 A_1 + \cdots + x_m A_m) = f_B(\mathbf{x})$$

with the notation of Example 8.7, and it follows that $\text{Disc}(\mathbf{A}(B)) = \text{per}(B)$.

Example 8.8 (Bapat's Conjecture). Generalizing the van der Waerden conjecture, in 1989 Bapat considered the set $\Omega(m)$ of m -tuples of m -by- m matrices $\mathbf{A} = (A_1, \dots, A_m)$ such that each A_i is positive semidefinite with trace $\text{tr}(A_i) = 1$, and $\sum_{i=1}^m A_i = I$. For any doubly stochastic matrix B , $\mathbf{A}(B)$ is in this set. The natural conjecture is that for all $\mathbf{A} \in \Omega(m)$, $\text{Disc}(\mathbf{A}) \geq m!/m^m$, and equality is attained if and only if $\mathbf{A} = \mathbf{A}((1/m)J)$. This was proved by Gurvits in 2006 – again, it follows directly from Corollary 8.4 and Theorem 8.5. It suffices to prove the result for $\mathbf{A} \in \Omega(m)$ such that each A_i is positive definite, by a routine limit argument. By Proposition 2.1, for $\mathbf{A} \in \Omega(m)$, the polynomial

$$f_{\mathbf{A}}(\mathbf{x}) = \det(x_1 A_1 + \cdots + x_m A_m)$$

is real stable. Since each A_i is positive definite, all coefficients of $f_{\mathbf{A}}$ are nonnegative, $f_{\mathbf{A}}$ is homogeneous of degree m , and $\deg_i(f_{\mathbf{A}}) = m$ for all $i \in [m]$. Since $\mathbf{A} \in \Omega(m)$, Lemma 8.6 implies that $\text{cap}(f_{\mathbf{A}}) = 1$. Thus, $f_{\mathbf{A}}$ satisfies the hypothesis of Theorems 8.3 and 8.5, and since $\text{Disc}(\mathbf{A}) = \partial^1 f_{\mathbf{A}}(\mathbf{0})$, the result follows.

9. FURTHER DIRECTIONS.

9.1. Other circular regions. Let $\Omega \subseteq \mathbb{C}^m$. A polynomial $f \in \mathbb{C}[\mathbf{x}]$ is Ω -stable if either $f \equiv 0$ identically, or $f(\mathbf{z}) \neq 0$ for all $\mathbf{z} \in \Omega$. At this level of generality little can be said. If $\Omega = \mathcal{A}_1 \times \cdots \times \mathcal{A}_m$ is a product of open circular regions then there are Möbius transformations $z \mapsto \phi_i(z) = (a_i z + b_i)/(c_i z + d_i)$ such that $\phi_i(\mathcal{H}) = \mathcal{A}_i$ for all $i \in [m]$. The argument in Section 4.1 shows that $f \in \mathbb{C}[\mathbf{x}]$ is Ω -stable if and only if

$$\tilde{f} = (\mathbf{c}\mathbf{z} + \mathbf{d})^{\deg f} \cdot f(\phi_1(z_1), \dots, \phi_m(z_m))$$

is stable. In this way results about stable polynomials can be translated into results about Ω -stable polynomials for any Ω that is a product of open circular regions.

Theorem 6.3 of [5] is the Ω -stability analogue of Theorem 5.2. We mention only two consequences of this. Let $\mathcal{D} = \{z \in \mathbb{C} : |z| < 1\}$ be the open unit disc, and for $\theta \in \mathbb{R}$ let $\mathcal{H}_\theta = \{e^{-i\theta}z : z \in \mathcal{H}\}$. Thus $\mathcal{H}_0 = \mathcal{H}$, and $\mathcal{H}_{\pi/2}$ is the open right half-plane. A \mathcal{D}^m -stable polynomial is called *Schur stable*, and a $\mathcal{H}_{\pi/2}^m$ -stable polynomial is called *Hurwitz stable*.

Proposition 9.1 (Remark 6.1 of [5]). *Fix $\kappa \in \mathbb{N}^m$, and let $T : \mathbb{C}[\mathbf{x}]^{\leq \kappa} \rightarrow \mathbb{C}[\mathbf{x}]$ be a linear transformation. The following are equivalent:*

- (a) *T preserves Schur stability.*
- (b) *$T((\mathbf{1} + \mathbf{xy})^\kappa)$ is Schur stable in $\mathbb{C}[\mathbf{x}, \mathbf{y}]$.*

Proposition 9.2 (Remark 6.1 of [5]). *Fix $\kappa \in \mathbb{N}^m$, and let $T : \mathbb{C}[\mathbf{x}]^{\leq \kappa} \rightarrow \mathbb{C}[\mathbf{x}]$ be a linear transformation. The following are equivalent:*

- (a) *T preserves Hurwitz stability.*
- (b) *$T((\mathbf{1} + \mathbf{xy})^\kappa)$ is Hurwitz stable in $\mathbb{C}[\mathbf{x}, \mathbf{y}]$.*

9.2. Applications of Theorem 5.4. It is natural to consider a multivariate analogue of the multiplier sequences studied by Pólya and Schur. Let $\lambda : \mathbb{N}^m \rightarrow \mathbb{R}$, and define a linear transformation $T_\lambda : \mathbb{C}[\mathbf{x}] \rightarrow \mathbb{C}[\mathbf{x}]$ by $T_\lambda(\mathbf{x}^\alpha) = \lambda(\alpha)\mathbf{x}^\alpha$ for all $\alpha \in \mathbb{N}^m$, and linear extension. For which λ does T_λ preserve real stability? The answer: just the ones you get from the Pólya-Schur Theorem, and no more.

Theorem 9.3 (Theorem 1.8 of [4]). *Let $\lambda : \mathbb{N}^m \rightarrow \mathbb{R}$. Then T_λ preserves real stability if and only if there are univariate multiplier sequences $\lambda_i : \mathbb{N} \rightarrow \mathbb{R}$ for $i \in [m]$ and $\epsilon \in \{-1, +1\}$ such that*

$$\lambda(\alpha) = \lambda_1(\alpha(1)) \cdots \lambda_m(\alpha(m))$$

for all $\alpha \in \mathbb{N}^m$, and either $\epsilon^{|\alpha|}\lambda(\alpha) \geq 0$ for all $\alpha \in \mathbb{N}^m$, or $\epsilon^{|\alpha|}\lambda(\alpha) \leq 0$ for all $\alpha \in \mathbb{N}^m$.

Theorem 5.4 (and similarly Propositions 9.1 and 9.2) can be used to derive a wide variety of results of the form: such-and-such an operation preserves stability (or Schur or Hurwitz stability). Here is a short account of Hinkkanen's proof of the Lee-Yang Circle Theorem, taken from Section 8 of [6].

For $f, g \in \mathbb{C}[\mathbf{x}]^{\text{MA}}$, say $f = \sum_{S \subseteq [m]} a(S) \mathbf{x}^S$ and $g = \sum_{S \subseteq [m]} b(S) \mathbf{x}^S$, let

$$f \bullet g = \sum_{S \subseteq [m]} a(S) b(S) \mathbf{x}^S$$

be the *Schur-Hadamard product* of f and g .

Theorem 9.4 (Hinkkanen, Theorem 8.5 of [6]). *If $f, g \in \mathbb{C}[\mathbf{x}]^{\text{MA}}$ are Schur stable then $f \bullet g$ is Schur stable.*

Proof. Let $T_g : \mathbb{C}[\mathbf{x}]^{\text{MA}} \rightarrow \mathbb{C}[\mathbf{x}]^{\text{MA}}$ be defined by $f \mapsto f \bullet g$. By Proposition 9.1, to show that T_g preserves Schur stability it suffices to show that $T_g((1 + \mathbf{xy})^{[m]})$ is Schur stable. Clearly $T_g((1 + \mathbf{xy})^{[m]}) = g(x_1 y_1, \dots, x_m y_m)$ is Schur stable since $g(\mathbf{x})$ is. Hence T_g preserves Schur stability, and so $f \bullet g$ is Schur stable. \square

Theorem 9.5 (Lee-Yang Circle Theorem, Theorem 8.4 of [6]). *Let $A = (a_{ij})$ be a Hermitian m -by- m matrix with $|a_{ij}| \leq 1$ for all $i, j \in [m]$. Then the polynomial*

$$f(\mathbf{x}) = \sum_{S \subseteq [m]} \mathbf{x}^S \prod_{i \in S} \prod_{j \notin S} a_{ij}$$

is Schur stable. The diagonalization $g(x) = f(x, \dots, x)$ is such that $x^m g(1/x) = g(x)$, and it follows that all roots of $g(x)$ are on the unit circle.

Proof. For $i < j$ in $[m]$ let

$$f_{ij} = (1 + a_{ij} x_i + \overline{a_{ij}} x_j + x_i x_j) \prod_{h \in [m] \setminus \{i, j\}} (1 + x_h).$$

One can check that each f_{ij} is Schur stable. The polynomial $f(\mathbf{x})$ is the Schur-Hadamard product of all the f_{ij} for $\{i, j\} \subseteq [m]$. By Theorem 9.4, $f(\mathbf{x})$ is Schur stable. \square

Section 8 of [6] contains many many more results of this nature.

9.3. A converse to the Grace-Walsh-Szegő Theorem. The argument of Sections 4.2 and 4.3 can be used to prove the following.

Exercise 9.6. If $f \in \mathfrak{S}[\mathbf{x}]^{\text{MA}}$ is multiaffine and stable then

$$T_{\mathfrak{S}(m)}(f) = \frac{1}{m!} \sum_{\sigma \in \mathfrak{S}(m)} \sigma(f)$$

is multiaffine and stable.

This is in fact equivalent to the GWS Theorem, since for all $f \in \mathbb{C}[\mathbf{x}]^{\text{MA}}$, $T_{\mathfrak{S}(m)} f(\mathbf{x}) = \text{Pol}_m f(x, \dots, x)$. For which transitive permutation groups $G \leq \mathfrak{S}(m)$ does the linear transformation $T_G = |G|^{-1} \sum_{\sigma \in G} \sigma$ preserve stability? The answer: not many, and they give nothing new.

Theorem 9.7 (Theorem 6 of [9]). *Let $G \leq \mathfrak{S}(m)$ be a transitive permutation group such that T_G preserves stability. Then $T_G = T_{\mathfrak{S}(m)}$.*

9.4. Phase and support theorems. A polynomial $f \in \mathbb{C}[\mathbf{x}]$ has *definite parity* if every monomial \mathbf{x}^α occurring in f has total degree of the same parity: all are even, or all are odd.

Theorem 9.8 (Theorem 6.2 of [10]). *Let $f \in \mathbb{C}[\mathbf{x}]$ be Hurwitz stable and with definite parity. Then there is a phase $0 \leq \theta < 2\pi$ such that $e^{-i\theta} f(\mathbf{x})$ has only real nonnegative coefficients.*

The *support* of $f = \sum_{\alpha} c(\alpha) \mathbf{x}^\alpha$ is $\text{supp}(f) = \{\alpha \in \mathbb{N}^m : c(\alpha) \neq 0\}$. Let δ_i denote the unit vector with a one in the i -th coordinate, and for $\alpha \in \mathbb{Z}^m$ let $|\alpha| = \sum_{i=1}^m |\alpha(i)|$. A *jump system* is a subset $\mathcal{J} \subseteq \mathbb{Z}^m$ satisfying the following *two-step axiom*:

(J) If $\alpha, \beta \in \mathcal{J}$ and $i \in [m]$ and $\epsilon \in \{-1, +1\}$ are such that $\alpha' = \alpha + \epsilon \delta_i$ satisfies $|\alpha' - \beta| < |\alpha - \beta|$, then either $\alpha' \in \mathcal{J}$ or there exists $j \in [m]$ and $\varepsilon \in \{-1, +1\}$ such that $\alpha'' = \alpha' + \varepsilon \delta_j \in \mathcal{J}$ and $|\alpha'' - \beta| < |\alpha' - \beta|$.

Jump systems generalize some more familiar combinatorial objects. A jump system contained in $\{0, 1\}^m$ is a *delta-matroid*. A delta-matroid \mathcal{J} for which $|\alpha|$ is constant for all $\alpha \in \mathcal{J}$ is the set of bases of a *matroid*. For bases of matroids, the two-step axiom (J) reduces to the basis exchange axiom familiar from linear algebra: if $A, B \in \mathcal{J}$ and $a \in A \setminus B$, then there exists $b \in B \setminus A$ such that $(A \setminus \{a\}) \cup \{b\}$ is in \mathcal{J} .

Theorem 9.9 (Theorem 3.2 of [8]). *If $f \in \mathfrak{S}[\mathbf{x}]$ is stable then the support $\text{supp}(f)$ is a jump system.*

Recall from Section 7 that for multiaffine polynomials with nonnegative coefficients, real stability implies the Rayleigh property. A set system \mathcal{J} is *convex* when $A, B \in \mathcal{J}$ and $A \subseteq B$ imply that $C \in \mathcal{J}$ for all $A \subseteq C \subseteq B$.

Theorem 9.10 (Section 4 of [15]). *Let $f = \sum_{S \subseteq [m]} c(S) \mathbf{x}^S$ be multiaffine with real nonnegative coefficients, and assume that f is Rayleigh.*

- (a) *The support $\text{supp}(f)$ is a convex delta-matroid.*
- (b) *The coefficients are log-submodular: for all $A, B \subseteq [m]$,*

$$c(A \cap B)c(A \cup B) \leq c(A)c(B).$$

REFERENCES

1. J. Borcea and P. Brändén, *Applications of stable polynomials to mixed determinants: Johnson's conjectures, unimodality, and symmetrized Fischer products*, Duke Math. J. **143** (2008), 205–223.
2. J. Borcea and P. Brändén, *Lee–Yang problems and the geometry of multivariate polynomials*, Lett. Math. Phys. **86** (2008), 53–61.
3. J. Borcea and P. Brändén, *Pólya–Schur master theorems for circular domains and their boundaries*, Ann. of Math. **170** (2009), 465–492.
4. J. Borcea and P. Brändén, *Multivariate Pólya–Schur classification problems in the Weyl algebra*, to appear in Proc. London Math. Soc.
5. J. Borcea and P. Brändén, *The Lee–Yang and Pólya–Schur programs I: linear operators preserving stability*, Invent. Math. **177** (2009), 541–569.
6. J. Borcea and P. Brändén, *The Lee–Yang and Pólya–Schur programs II: theory of stable polynomials and applications* Comm. Pure Appl. Math. **62** (2009), 1595–1631.
7. J. Borcea, P. Brändén, and T.M. Liggett, *Negative dependence and the geometry of polynomials*, J. Amer. Math. Soc. **22** (2009), 521–567.
8. P. Brändén, *Polynomials with the half-plane property and matroid theory*, Adv. Math. **216** (2007), 302–320.

9. P. Brändén and D.G. Wagner, *A converse to the Grace–Walsh–Szegő theorem*, Math. Proc. Camb. Phil. Soc. **147** (2009), 447–453.
10. Y.-B. Choe, J.G. Oxley, A.D. Sokal, and D.G. Wagner, *Homogeneous polynomials with the half-plane property*, Adv. in Appl. Math. **32** (2004), 88–187.
11. L. Gurvits, *Van der Waerden/Schrijver–Valiant like conjectures and stable (aka hyperbolic) homogeneous polynomials: one theorem for all. With a corrigendum*, Electron. J. Combin. **15** (2008), R66 (26 pp).
12. G. Hardy, J.E. Littlewood, and G. Pólya, “Inequalities (Second Edition),” Cambridge U.P., Cambridge UK, 1952.
13. M. Laurent and A. Schrijver, *On Leonid Gurvits’ proof for permanents*, <http://homepages.cwi.nl/~lex/files/perma5.pdf>
14. Q.I. Rahman and G. Schmeisser, “Analytic Theory of Polynomials,” London Math. Soc. Monographs (N.S.) **26**, Oxford U.P., New York NY, 2002.
15. D.G. Wagner, *Negatively correlated random variables and Mason’s conjecture for independent sets in matroids*, Ann. of Combin. **12** (2008), 211–239.
16. D.G. Wagner and Y. Wei, *A criterion for the half-plane property*, Discrete Math. **309** (2009), 1385–1390.

DEPARTMENT OF COMBINATORICS AND OPTIMIZATION, UNIVERSITY OF WATERLOO, WATERLOO,
ONTARIO, CANADA N2L 3G1

E-mail address: dgwagner@math.uwaterloo.ca