I: Hyperbolicity, stability and geometry

Robin Pemantle

Current Developments in Mathematics, 18 November 2011

Two related concepts

These lectures are about two related properties of polynomials, **hyperbolicity** and **stability**.

Two related concepts

These lectures are about two related properties of polynomials, **hyperbolicity** and **stability**.

These notions have been at the heart of some fundamental results in very different fields, dating from 60 years ago to the present.

Two related concepts

These lectures are about two related properties of polynomials, **hyperbolicity** and **stability**.

These notions have been at the heart of some fundamental results in very different fields, dating from 60 years ago to the present.

I will survey the development and key uses of these concepts, devoting the first lecture largely to hyperbolicity and the second to stability.

Definitions

A homogeneous polynomial p of degree m is said to be **hyperbolic** in direction $\mathbf{x} \in \mathbb{R}^{\mathbf{d}}$ if $\mathbf{p}(\mathbf{y} + \mathbf{i}\mathbf{x}) \neq 0$ for all $\mathbf{y} \in \mathbb{R}^{\mathbf{d}}$.

A polynomial $\mathbf{q}$ is said to be **stable** if $\mathbf{q}(\mathbf{z}) \neq 0$ whenever each coordinate $\mathbf{z_j}$ is in the strict upper half plane.

Proposition 1 (hyperbolicity vs. stability)

A real homogeneous polynomial is stable if and only if it is hyperbolic in every direction in the positive orthant. □

Definitions

A homogeneous polynomial $\mathbf{p}$ of degree $\mathbf{m}$ is said to be **hyperbolic** in direction $\mathbf{x} \in \mathbb{R}^{\mathbf{d}}$ if $\mathbf{p}(\mathbf{y} + \mathbf{i}\mathbf{x}) \neq 0$ for all $\mathbf{y} \in \mathbb{R}^{\mathbf{d}}$.

A polynomial $\mathbf{q}$ is said to be **stable** if $\mathbf{q}(\mathbf{z}) \neq 0$ whenever each coordinate $\mathbf{z_j}$ is in the strict upper half plane.

Proposition 1 (hyperbolicity vs. stability)

A real homogeneous polynomial is stable if and only if it is hyperbolic in every direction in the positive orthant. □

Notation: $\mathbf{d}$ will denote the number of variables and $\mathbf{m}$ the degree.

What you are missing, part I:

Definition of hyperbolicity in the inhomogeneous case

What you are missing, part I:

Definition of hyperbolicity in the inhomogeneous case

Note: once we have restricted to the homogenous case, we may also assume without loss of generality that $\mathbf{p}$ is real: for homogeneous polynomials, hyperbolicity implies that some multiple is real.

Real roots

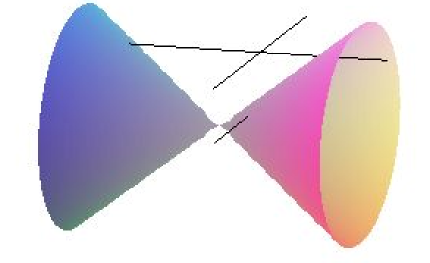
Proposition 2 (location of zeros)

$\mathbf{q}(\mathbf{y} + \mathbf{ix}) \neq 0$ for all $\mathbf{y}$ if and only if $\mathbf{q}(\mathbf{x}) \neq 0$ and $\mathbf{z} \mapsto \mathbf{q}(\mathbf{y} + \mathbf{zx})$ has all real zeros when $\mathbf{y}$ is real.

PROOF: Absorb the real part of $\mathbf{zx}$ into $\mathbf{y}$. □

Example: Lorentzian quadratic

Let $\mathbf{p}$ be the Lorentzian quadratic $\mathbf{t}^2 - \mathbf{x}_2^2 - \cdots - \mathbf{x}_\mathbf{d}^2$, where we have renamed $\mathbf{x}_1$ as "$\mathbf{t}$" because of its interpretation as the time axis in spacetime; then $\mathbf{p}$ is hyperbolic in every timelike direction, that is, for each direction $\mathbf{x}$ with $\mathbf{p}(\mathbf{x}) > 0$.



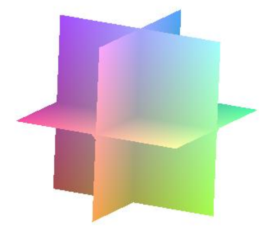
The time axis is left-right

Example: coordinate planes

The coordinate function $\mathbf{x_j}$ is hyperolic in direction $\mathbf{y}$ if and only if $\mathbf{y_j} \neq 0$ (this is true for any linear polynomial).

It is obvious from the definition that the product of polynomials hyperbolic in direction $\mathbf{y}$ is again hyperbolic in direction $\mathbf{y}$.

It follows that $\prod_{\mathbf{j}=1}^{\mathbf{d}} \mathbf{x_j}$ is hyperbolic in every direction not contained in a coordinate plane, that is, in every open orthant.



More examples

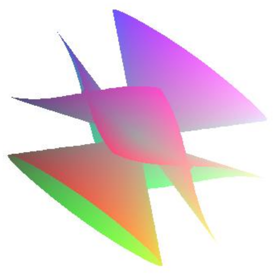
Early works developing the theory of hyperbolic functions seem to treat the Lorentzian quadratic as the only motivating example, though they discuss a few others to show that the theory is more general. The generality turned out to be useful in contexts that were only dreamed of much later. Along with these contexts came new examples.

We won't have time here to discuss two sources of examples, namely lacunas and self-concordant barrier functions. We will, however, discuss the example of rational Taylor series. It turns out that any polynomial, when localized at a point on the boundary of its amoeba, is hyperbolic. More on these notions later, but here is a picture.

Example: Fortress polynomial

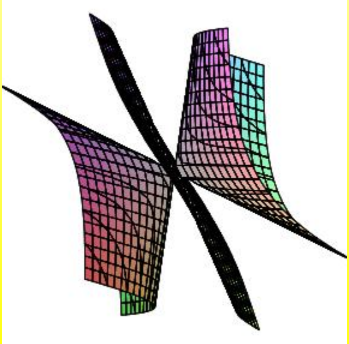
$$\mathbf{w}^4 - \mathbf{u}^2\mathbf{w}^2 - \mathbf{v}^2\mathbf{w}^2 + \frac{9}{25}\mathbf{u}^2\mathbf{v}^2$$

is the projective localization of the denominator (cleaned up a bit) of the so-called Fortress generating polynomial. It follows from [BP11, Proposition 2.12] that this polynomial is hyperbolic.



Another example from combinatorics

This is from a 1-parameter family of hyperbolic polynomials. It is irreducible (the collar in the middle is not a flat plane) except for one parameter value.



Ubiquity of zeros of polynomials

"The one contribution of mine that I hope will be remembered has consisted in pointing out that all sorts of problems of combinatorics can be viewed as problems of the location of the zeros of certain polynomials..."

– Gian-Carlo Rota (1985)

cited by Borcea, Brändén and Liggett (2009)

Ubiquity of zeros of polynomials

"The one contribution of mine that I hope will be remembered has consisted in pointing out that all sorts of problems of combinatorics can be viewed as problems of the location of the zeros of certain polynomials..."

– Gian-Carlo Rota (1985)

cited by Borcea, Brändén and Liggett (2009)

Also: PDE's/harmonic analysis, probability, number theory, ...

Organization

The first lecture will be more geometric. The important properties of hyperbolic functions are related to convexity and to cones of hyperbolicity. Applications are to propagation of wave-like PDE's, inverse Fourier transforms, and their application to analytic combinatorics.

The second lecture has a more algebraic flavor. Closure properties of the class of stable polynomials play a large role. Many of the applications concern determinants.

Hyperbolicity

Major uses of hyperbolicity

Highlights:

Major uses of hyperbolicity

Highlights:

* Gårding's necessary and sufficient condition for stability of the evolution of a wave-like PDE (1951).

Major uses of hyperbolicity

Highlights:

* Gårding's necessary and sufficient condition for stability of the evolution of a wave-like PDE (1951). See also Petrovsky (1947).

Major uses of hyperbolicity

Highlights:

* Gårding's necessary and sufficient condition for stability of the evolution of a wave-like PDE (1951). See also Petrovsky (1947).
* Atiyah-Bott-Gårding's computation of inverse Fourier transforms and lacunas for homogeneous rational functions (1970).

Major uses of hyperbolicity

Highlights:

* Gårding's necessary and sufficient condition for stability of the evolution of a wave-like PDE (1951). See also Petrovsky (1947).
* Atiyah-Bott-Gårding's computation of inverse Fourier transforms and lacunas for homogeneous rational functions (1970).
* Interior methods for convex programming (1997).

Major uses of hyperbolicity

Highlights:

* Gårding's necessary and sufficient condition for stability of the evolution of a wave-like PDE (1951). See also Petrovsky (1947).
* Atiyah-Bott-Gårding's computation of inverse Fourier transforms and lacunas for homogeneous rational functions (1970).
* Interior methods for convex programming (1997).
* Asymptotics of Taylor coefficients for rational functions (2011).

Properties

Properties of hyperbolic polynomials:

Properties

Properties of hyperbolic polynomials:

* Directions of hyperbolicity are partitioned into convex cones (Proposition 4);

Properties

Properties of hyperbolic polynomials:

* Directions of hyperbolicity are partitioned into convex cones (Proposition 4);
* if $\mathbf{x}$ and $\mathbf{y}$ are in the same cone then the roots of $\mathbf{q}(\mathbf{y} + \mathbf{tx})$ are nonpositive.

Properties

Properties of hyperbolic polynomials:

* Directions of hyperbolicity are partitioned into convex cones (Proposition 4);
* if $\mathbf{x}$ and $\mathbf{y}$ are in the same cone then the roots of $\mathbf{q}(\mathbf{y} + \mathbf{tx})$ are nonpositive.
* Cones of hyperbolicity for localizations of $\mathbf{q}$ can be arranged into a semi-continuously varying family (Theorem 11).

Properties

Properties of hyperbolic polynomials:

* Directions of hyperbolicity are partitioned into convex cones (Proposition 4);
* if $\mathbf{x}$ and $\mathbf{y}$ are in the same cone then the roots of $\mathbf{q}(\mathbf{y} + \mathbf{tx})$ are nonpositive.
* Cones of hyperbolicity for localizations of $\mathbf{q}$ can be arranged into a semi-continuously varying family (Theorem 11).
* The space plane can be deformed into the forward cone (Theorem 12 and the construction immediately following).

What you are missing, part II:

Properties of hyperbolicity in the inhomogeneous case

Stable evolution of PDE's

Let $\mathbf{q}$ be a polynomial in $\mathbf{d}$ variables and denote by $\mathbf{D_q}$ the operator $\mathbf{q}(\partial/\partial\mathbf{x})$ obtained by replacing each $\mathbf{x_i}$ by $\partial/\partial\mathbf{x_i}$.

Stable evolution of PDE's

Let $\mathbf{q}$ be a polynomial in $\mathbf{d}$ variables and denote by $\mathbf{D_q}$ the operator $\mathbf{q}(\partial/\partial \mathbf{x})$ obtained by replacing each $\mathbf{x_i}$ by $\partial/\partial \mathbf{x_i}$.

Let $\mathbf{r}$ be a vector in $\mathbb{R}^{\mathbf{d}}$, let $\mathbf{H_r}$ be the hyperplane orthogonal to $\mathbf{r}$, and consider the equation

$$\mathbf{D_q}(\mathbf{f}) = 0 \qquad (1)$$

in the halfspace $\{\mathbf{r} \cdot \mathbf{x} \geq 0\}$ with boundary conditions specified on $\mathbf{H_r}$ (typically, $\mathbf{f}$ and its first $\mathbf{d} - 1$ normal derivatives).

Stable evolution of PDE's

Let $\mathbf{q}$ be a polynomial in $\mathbf{d}$ variables and denote by $\mathbf{D_q}$ the operator $\mathbf{q}(\partial/\partial \mathbf{x})$ obtained by replacing each $\mathbf{x_i}$ by $\partial/\partial \mathbf{x_i}$.

Let $\mathbf{r}$ be a vector in $\mathbb{R}^{\mathbf{d}}$, let $\mathbf{H_r}$ be the hyperplane orthogonal to $\mathbf{r}$, and consider the equation

$$\mathbf{D_q}(\mathbf{f}) = 0 \qquad (1)$$

in the halfspace $\{\mathbf{r} \cdot \mathbf{x} \geq 0\}$ with boundary conditions specified on $\mathbf{H_r}$ (typically, $\mathbf{f}$ and its first $\mathbf{d} - 1$ normal derivatives).

We say that (1) evolves stably in direction $\mathbf{r}$ if convergence of the boundary conditions to 0 implies convergence of the solution to 0.

Stable evolution of PDE's

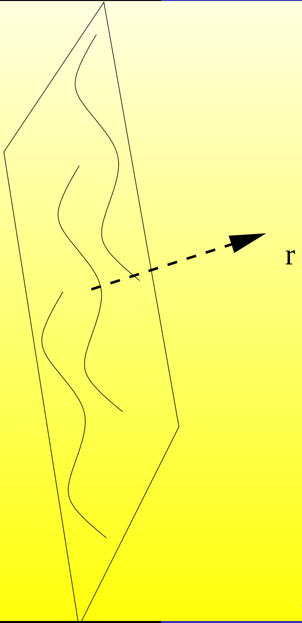
Let $\mathbf{q}$ be a polynomial in $\mathbf{d}$ variables and denote by $\mathbf{D_q}$ the operator $\mathbf{q}(\partial/\partial\mathbf{x})$ obtained by replacing each $\mathbf{x_i}$ by $\partial/\partial\mathbf{x_i}$.

Let $\mathbf{r}$ be a vector in $\mathbb{R}^{\mathbf{d}}$, let $\mathbf{H_r}$ be the hyperplane orthogonal to $\mathbf{r}$, and consider the equation

$$\mathbf{D_q}(\mathbf{f}) = 0 \tag{1}$$

in the halfspace $\{\mathbf{r}\cdot\mathbf{x} \geq 0\}$ with boundary conditions specified on $\mathbf{H_r}$ (typically, $\mathbf{f}$ and its first $\mathbf{d}-1$ normal derivatives).

We say that (1) evolves stably in direction $\mathbf{r}$ if convergence of the boundary conditions to 0 implies convergence of the solution to 0. Convergence here means uniform convergence of the function and its derivatives on compact sets.


r

Gårding's Theorem

Theorem 3 ([Går51, Theorem III])

The equation $\mathbf{D_q f} = 0$ *evolves stably in direction* $\mathbf{r}$ *if and only if* $\mathbf{q}$ *is hyperbolic in direction* $\mathbf{r}$.

Gårding's Theorem

Theorem 3 ([Går51, Theorem III])

The equation $\mathbf{D_q f} = 0$ *evolves stably in direction* $\mathbf{r}$ *if and only if* $\mathbf{q}$ *is hyperbolic in direction* $\mathbf{r}$.

Let us see why this should be true.

Gårding's Theorem

Theorem 3 ([Går51, Theorem III])

The equation $\mathbf{D_q f} = 0$ *evolves stably in direction* $\mathbf{r}$ *if and only if* $\mathbf{q}$ *is hyperbolic in direction* $\mathbf{r}$.

Let us see why this should be true.

We begin with the observation that if $\xi \in \mathbf{C^d}$ is any vector with $\mathbf{q}(\xi) = 0$ then $\mathbf{f}_\xi(\mathbf{x}) := \exp(\mathbf{i}\xi \cdot \mathbf{x})$ is a solution to $\mathbf{D_q f} = 0$. (Our solutions are allowed to be complex but live on $\mathbb{R}^\mathbf{d}$.)

Forward direction

Stability implies hyperbolicity:

Forward direction

STABILITY IMPLIES HYPERBOLICITY:

Assume WLOG that $\mathbf{r} = (0, \ldots, 0, 1)$.

Forward direction

STABILITY IMPLIES HYPERBOLICITY:

Assume WLOG that $\mathbf{r} = (0, \ldots, 0, 1)$.

Suppose $\mathbf{q}$ is not hyperbolic. Then at least one line parallel to $\mathbf{r}$ has a pair of complex roots, meaning that there is a ξ with $\mathbf{q}(\xi) = 0$ and $\xi = (\mathbf{a}_1, \ldots, \mathbf{a}_{\mathbf{d}-1}, \mathbf{a}_\mathbf{d} \pm \mathbf{bi})$, with $\{\mathbf{a}_\mathbf{i}\}$ and $\mathbf{b}$ real and $\mathbf{b} \neq 0$.

Forward direction

STABILITY IMPLIES HYPERBOLICITY:

Assume WLOG that $\mathbf{r} = (0, \ldots, 0, 1)$.

Suppose $\mathbf{q}$ is not hyperbolic. Then at least one line parallel to $\mathbf{r}$ has a pair of complex roots, meaning that there is a ξ with $\mathbf{q}(\xi) = 0$ and $\xi = (\mathbf{a}_1, \ldots, \mathbf{a}_{\mathbf{d}-1}, \mathbf{a}_\mathbf{d} \pm \mathbf{bi})$, with $\{\mathbf{a}_\mathbf{i}\}$ and $\mathbf{b}$ real and $\mathbf{b} \neq 0$.

Picking $\mathbf{b} < 0$, the function $\mathbf{f}_\xi$ grows exponentially in direction $\mathbf{r}$. For large λ, the function $\mathbf{f}_{\lambda\xi}$ grows even faster.

Forward direction

STABILITY IMPLIES HYPERBOLICITY:

Assume WLOG that $\mathbf{r} = (0, \ldots, 0, 1)$.

Suppose $\mathbf{q}$ is not hyperbolic. Then at least one line parallel to $\mathbf{r}$ has a pair of complex roots, meaning that there is a ξ with $\mathbf{q}(\xi) = 0$ and $\xi = (\mathbf{a_1}, \ldots, \mathbf{a_{d-1}}, \mathbf{a_d} \pm \mathbf{bi})$, with $\{\mathbf{a_i}\}$ and $\mathbf{b}$ real and $\mathbf{b} \neq 0$.

Picking $\mathbf{b} < 0$, the function $\mathbf{f}_\xi$ grows exponentially in direction $\mathbf{r}$. For large λ, the function $\mathbf{f}_{\lambda\xi}$ grows even faster.

Sending $\lambda \to \infty$, we may take initial conditions going to zero such that $\mathbf{f}_{\lambda\xi}(\mathbf{r}) = 1$ for all λ.

Argument for the reverse direction

The other direction is harder, but here is a hand-waving argument.

HYPERBOLICITY IMPLES STABILITY:

Argument for the reverse direction

The other direction is harder, but here is a hand-waving argument.

HYPERBOLICITY IMPLES STABILITY:

Suppose $\mathbf{q}$ is hyperbolic. For every real $\mathbf{r}' = (\mathbf{r}_1, \ldots, \mathbf{r}_{\mathbf{d}-1})$ in frequency space there are $\mathbf{d}$ real values of $\mathbf{r_d}$ such that $\mathbf{q}(\mathbf{r}) := \mathbf{q}(\mathbf{r}', \mathbf{r_d}) = 0$. For each such $\mathbf{r}$, the function $\mathbf{f_r}$ is a solution to $\mathbf{D_q}(\mathbf{f}) = 0$ traveling unitarily.

Argument for the reverse direction

The other direction is harder, but here is a hand-waving argument.

HYPERBOLICITY IMPLES STABILITY:

Suppose $\mathbf{q}$ is hyperbolic. For every real $\mathbf{r}' = (\mathbf{r}_1, \ldots, \mathbf{r}_{\mathbf{d}-1})$ in frequency space there are $\mathbf{d}$ real values of $\mathbf{r}_\mathbf{d}$ such that $\mathbf{q}(\mathbf{r}) := \mathbf{q}(\mathbf{r}', \mathbf{r}_\mathbf{d}) = 0$. For each such $\mathbf{r}$, the function $\mathbf{f}_\mathbf{r}$ is a solution to $\mathbf{D}_\mathbf{q}(\mathbf{f}) = 0$ traveling unitarily.

These $\mathbf{d}$ solutions are a unitary basis for the space $\mathbf{V}_{\mathbf{r}'}$ that they span, of solutions to (1) that restrict on $\mathbf{H}_\xi$ to $\mathbf{e}^{\mathbf{i}\mathbf{r}'\cdot\mathbf{x}}$, and the same is true at any later time.

Heuristic

HANDWAVE: Because the vector space has dimension $\mathbf{d}$ over any spatial frequency, we can believe we have all the solutions. They all evolve unitarily. Thus, writing any boundary conditions $\mathbf{f} = \mathbf{g}_0, \mathbf{f}' = \mathbf{g}_1, \ldots, \mathbf{f}^{(\mathbf{d}-1)} = \mathbf{g}_{\mathbf{d}-1}$ as an integral of functions $\mathbf{f}_\mathbf{r}$, unitary evolution implies a Parseval-type relation, meaning that small boundary conditions will lead to small values at any positive time.

Discussion

- That hyperbolicity is the "right" condition for stability of PDE's is not in question: Gårding's criterion is necessary and sufficient.

Discussion

- That hyperbolicity is the "right" condition for stability of PDE's is not in question: Gårding's criterion is necessary and sufficient.
- Hyperbolicity was used in a very direct way, implying $q(t) := q(r' + tr)$ has d real roots for any r'.

Discussion

- That hyperbolicity is the "right" condition for stability of PDE's is not in question: Gårding's criterion is necessary and sufficient.
- Hyperbolicity was used in a very direct way, implying $q(t) := q(r' + tr)$ has d real roots for any r'.
- The actual proof is dozens of pages and beyond our scope here, but at its heart is the construction of the **Riesz kernel**, to which we will return shortly.

Cones of hyperbolicity

Before going on, we need a little more of the theory of hyperbolic polynomials.

Cones of hyperbolicity

Before going on, we need a little more of the theory of hyperbolic polynomials.

Proposition 4 (cones of hyperbolicity)

Let p be real and homogeneous and denote the zero set of p by $\mathcal{V}$. If K is a connected component of $\mathbb{R}^d \setminus \mathcal{V}$ containing a direction of hyperbolicity for p, then every $x \in K$ is a direction of hyperbolocity for p. $\square$

Cones of hyperbolicity

Before going on, we need a little more of the theory of hyperbolic polynomials.

Proposition 4 (cones of hyperbolicity)

Let p be real and homogeneous and denote the zero set of p by $\mathcal{V}$. If K is a connected component of $\mathbb{R}^d \setminus \mathcal{V}$ containing a direction of hyperbolicity for p, then every $x \in K$ is a direction of hyperbolocity for p. □

The component of $\mathcal{V}^c$ containing x is called a **cone of hyperbolocity** of p and is denoted $K(p, x)$.

Example

Example 5 (Lorentzian quadratic)

Let $p = t^2 - x_2^2 - \cdots - x_d^2$ or any other nondegenerate quadratic with signature $(1, d-1)$. Then p is hyperbolic in direction x if and only if x is timelike, meaning that $p(x) > 0$. The two cones of hyperbolicity for p are the forward and backward cones (timelike vectors with x_1 respectively positive and negative).

Example

Example 5 (Lorentzian quadratic)

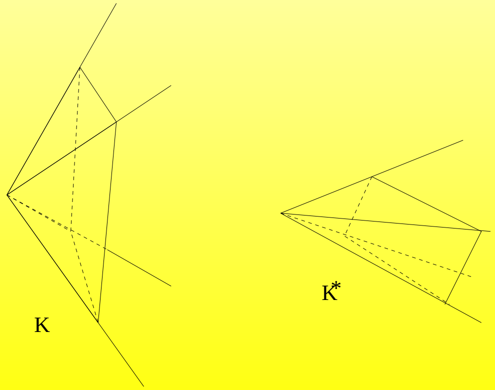
Let $p = t^2 - x_2^2 - \cdots - x_d^2$ or any other nondegenerate quadratic with signature $(1, d-1)$. Then p is hyperbolic in direction x if and only if x is timelike, meaning that $p(x) > 0$. The two cones of hyperbolicity for p are the forward and backward cones (timelike vectors with x_1 respectively positive and negative).

Example 6 (coordinate planes)

Let $p = \prod_{j=1}^{d} x_j$. Each monomial x_j is hyperbolic and its cones are two open halfspaces. The product of hyperbolic functions is hyperbolic and the cones are just the intersections. Consequently, p is hyperbolic with the 2^d orthants as cones of hyperbolicity.

Dual cones

Let $K \subseteq \mathbb{R}^d$ be a cone and let K^* denote the dual cone, that is the set of all y such that $x \cdot y \geq 0$ for all $x \in K$.



Riesz kernel

Let K be a cone of hyperbolicity for the homogeneous polyomial q of degree m and let K^* denote its dual cone.

Theorem 7 (Riesz kernel)

The function

$$Q(r,\alpha) := (2\pi)^{-d} \int_{\mathbb{R}^d} q(x+iy)^{-\alpha} \exp[r \cdot (x+iy)] \, dy$$

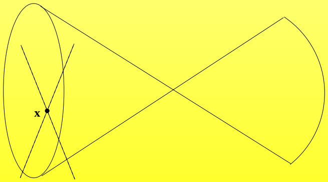
is well defined when $m \cdot \mathrm{Re}\{\alpha\} > d$ *and is independent of the choice of* $x \in K$*. For any* α*,* $Q(r,\alpha)$ *is defined in the sense of distributions and is always supported on the dual cone* K^**.*

Proof

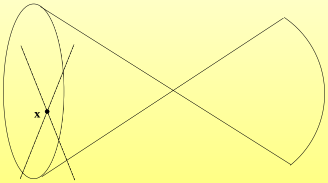
PROOF: Hyperbolicity in direction x implies that $q(x+iy)$ does not vanish for any y [use $q(-y+ix) \neq 0$ and homogeneity], from which an easy estimate is

$$|q(x+iy)^{-\alpha}| \leq C_x |y|^{-m \cdot \mathrm{Re}\,\{\alpha\}}$$

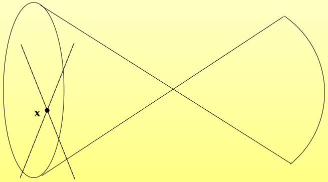
and convergence of the integral for $m \cdot \mathrm{Re}\,\{\alpha\} > d$ follows.



The space $x + i\mathbb{R}^d$ does not intersect the complex variety $\{q = 0\}$.



By holomorphicity we may deform x within the connected component K without changing the integral. If $y \cdot x < 0$ for some $x' \in K$ then deforming x to $\lambda x'$ and sending λ to infinity shows that the integral vanishes, proving that Q is supported on K^*.



By holomorphicity we may deform x within the connected component K without changing the integral. If $y \cdot x < 0$ for some $x' \in K$ then deforming x to $\lambda x'$ and sending λ to infinity shows that the integral vanishes, proving that Q is supported on K^*.

To extend to all α we use the first of the following facts.

Properties of Q

Recall:

$$Q(r,\alpha) := (2\pi)^{-d} \int_{\mathbb{R}^d} q(x+iy)^{-\alpha} \exp[r \cdot (x+iy)]\, dy$$

Let $E(r)$ denote $Q(r,1)$. Then

$$\begin{aligned} D_q Q(r,\alpha) &= Q(r,\alpha-1) \quad (\alpha \neq 1) \\ Q(\cdot,\alpha) * Q(\cdot,\beta) &= Q(\cdot,\alpha+\beta) \\ D_q E &= \delta \end{aligned}$$

where δ is the delta function at the origin.

Properties of Q

Recall:

$$Q(r, \alpha) := (2\pi)^{-d} \int_{\mathbb{R}^d} q(x + iy)^{-\alpha} \exp[r \cdot (x + iy)] \, dy$$

Let $E(r)$ denote $Q(r, 1)$. Then

$$\begin{aligned} D_q Q(r, \alpha) &= Q(r, \alpha - 1) \qquad (\alpha \neq 1) \\ Q(\cdot, \alpha) * Q(\cdot, \beta) &= Q(\cdot, \alpha + \beta) \\ D_q E &= \delta \end{aligned}$$

where δ is the delta function at the origin.

Note: the Riesz kernel is used to complete the (non-handwaving) proof of Gårding's theorem. Specifically, if $D_q f = 0$ then $f = (I - I_q D_q) P_\xi(f)$ where P_ξ is a continuous function of the boundary values and I_q is convolution with the Riesz kernel.

Consequences

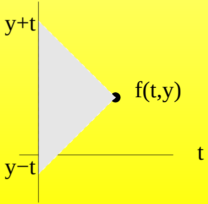
Consequences of these facts: f is constructed from the boundary conditions by convolution with E and E is supported on K^*, so $f(x)$ depends only on the boundary conditions on the intersection of H_ξ with $x - K^*$.

Propagation of two-dimensional wave equation

Example 8 (2-D wave equation)

Let f solve $f_{tt} - f_{yy} = 0$ with boundary conditions $f(0,y) = g(y)$ and $f_t(0,y) = h(y)$. Then an explicit formula for f in the right half plane is given by

$$f(t,y) = \frac{1}{2}\left[f(0,y+t) + f(0,y-t) + \int_{y-t}^{y+t} f'(0,u)\,du\right].$$



The Riesz kernel is so named because Riesz already computed it for Lorentzian quadratics (the only hyperbolic quadratics) in 1947.

The Riesz kernel is so named because Riesz already computed it for Lorentzian quadratics (the only hyperbolic quadratics) in 1947.

The fundamental solution E is computed in by $\int q^{-1} e^{r \cdot (x+iy)} \, dy$, however one would like a more explicit form.

The Riesz kernel is so named because Riesz already computed it for Lorentzian quadratics (the only hyperbolic quadratics) in 1947.

The fundamental solution E is computed in by $\int q^{-1} e^{r \cdot (x+iy)} \, dy$, however one would like a more explicit form.

In fact for some polynomials the Riesz kernel may be represented as an explicitly computable rational or algebraic integral.

The Riesz kernel is so named because Riesz already computed it for Lorentzian quadratics (the only hyperbolic quadratics) in 1947.

The fundamental solution E is computed in by $\int q^{-1} e^{r \cdot (x+iy)} \, dy$, however one would like a more explicit form.

In fact for some polynomials the Riesz kernel may be represented as an explicitly computable rational or algebraic integral.

To do so, [ABG70] turned the integral into a homogeneous integral. This required some more geometry of hyperbolic functions.

What you are missing, part III:

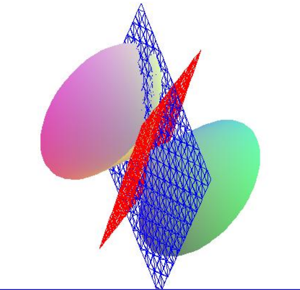
Application of the theory of hyperbolic functions to the construction of self-concordant barrier functions in convex programming.

Let p be a homogeneous polynomial and let $m = m_x(p)$ denote the degree of vanishing of p at x. Denote by $\text{loc}(p, x)$ the m-homogeneous part of $p(x + \cdot)$ which we call the **localization**. By definition if $p(x) \neq 0$ then $\text{loc}(p, x)$ is a nonzero constant.

Let p be a homogeneous polynomial and let $m = m_x(p)$ denote the degree of vanishing of p at x. Denote by $\text{loc}(p, x)$ the m-homogeneous part of $p(x + \cdot)$ which we call the **localization**. By definition if $p(x) \neq 0$ then $\text{loc}(p, x)$ is a nonzero constant.

Example 9

If $p = x_1^2 - \sum_{j=2}^d x_j^2$ is the Lorentzian quadratic, vanishing on the light cone $\{p = 0\}$, then the hyperplanes tangent to the cone are the vanishing sets of localizations of p.



Family of cones

Proposition 10 (family of local cones)

If p is hyperbolic then the functions $\text{loc}(p, x)$ are also. If C is a cone of hyperbolicity of p then each $\text{loc}(p, x)$ has a cone of hyperbolicity containing C. Denote this by $K^{p,C}(x)$.

Family of cones

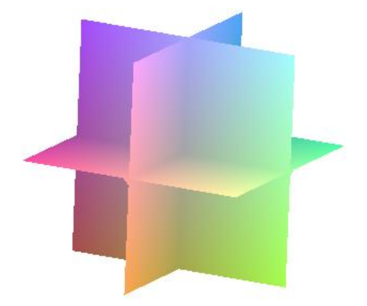
Proposition 10 (family of local cones)

If p is hyperbolic then the functions $\text{loc}(p, x)$ are also. If C is a cone of hyperbolicity of p then each $\text{loc}(p, x)$ has a cone of hyperbolicity containing C. Denote this by $K^{p,C}(x)$.

Recall that if y is a hyperbolic direction for $\text{loc}(p, x)$ then $q(x + ty) \neq 0$ for $t > 0$ (all roots are negative real). Thus y "points away from $\mathcal{V}$. By picking one of the cones of hyperbolicity at each x, we have in effect chosen a forward orientation from $\{p = 0\}$ into its complement.

Picture of orientation

For example, at each point in the intersection of j of the planes, choose the 2^{-j}-space containing the positive orthant.



Stratified behavior

In the previous examples, the cones $K^{p,C}(x)$ did not vary continuously with x but they did vary semi-continuously in the sense that they can only drop down at a limit point. This turns out to be true in general. The following result occupies Section 5 of [ABG70].

Stratified behavior

In the previous examples, the cones $K^{p,C}(x)$ did not vary continuously with x but they did vary semi-continuously in the sense that they can only drop down at a limit point. This turns out to be true in general. The following result occupies Section 5 of [ABG70].

Theorem 11 (semi-continuity)

Let C be any cone of hyperbolicity of p. Then the family of cones $K^{p,C}$ is semicontinuous in the sense that

$$\liminf_{x' \to x} K^{p,C}(x') \supseteq K^{p,C}(x) .$$

Stratified behavior

In the previous examples, the cones $K^{p,C}(x)$ did not vary continuously with x but they did vary semi-continuously in the sense that they can only drop down at a limit point. This turns out to be true in general. The following result occupies Section 5 of [ABG70].

Theorem 11 (semi-continuity)

Let C be any cone of hyperbolicity of p. Then the family of cones $K^{p,C}$ is semicontinuous in the sense that

$$\liminf_{x' \to x} K^{p,C}(x') \supseteq K^{p,C}(x) .$$

This is used to establish:

Vector field

Theorem 12 ([ABG70])

As x varies, suppose each $K^{p,C}(x)$ contains some vector v with $r \cdot v > 0$. Then there is a continuous, 1-homogeneous section $x \mapsto v(x)$ such that $v(x) \in K^{p,C}(x)$ and $r \cdot v(x) > 0$ for all x.

Vector field

Theorem 12 ([ABG70])

As x varies, suppose each $K^{p,C}(x)$ contains some vector v with $r \cdot v > 0$. Then there is a continuous, 1-homogeneous section $x \mapsto v(x)$ such that $v(x) \in K^{p,C}(x)$ and $r \cdot v(x) > 0$ for all x.

PROOF: By hypothesis, for each x there is a v.

Vector field

Theorem 12 ([ABG70])

As x varies, suppose each $K^{p,C}(x)$ contains some vector v with $r \cdot v > 0$. Then there is a continuous, 1-homogeneous section $x \mapsto v(x)$ such that $v(x) \in K^{p,C}(x)$ and $r \cdot v(x) > 0$ for all x.

PROOF: By hypothesis, for each x there is a v. By semi-continuity, this v works for all x' in some neighborhood of x.

Vector field

Theorem 12 ([ABG70])

As x varies, suppose each $K^{p,C}(x)$ contains some vector v with $r \cdot v > 0$. Then there is a continuous, 1-homogeneous section $x \mapsto v(x)$ such that $v(x) \in K^{p,C}(x)$ and $r \cdot v(x) > 0$ for all x.

PROOF: By hypothesis, for each x there is a v. By semi-continuity, this v works for all x' in some neighborhood of x. By compactness (we are working in projective space), finitely many of these neighborhoods cover.

Vector field

Theorem 12 ([ABG70])

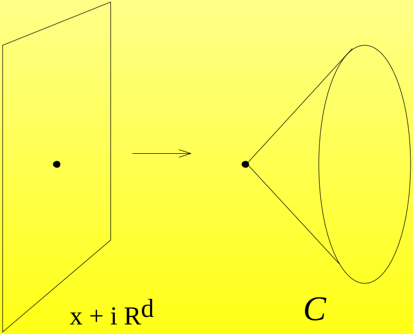
As x varies, suppose each $K^{p,C}(x)$ contains some vector v with $r \cdot v > 0$. Then there is a continuous, 1-homogeneous section $x \mapsto v(x)$ such that $v(x) \in K^{p,C}(x)$ and $r \cdot v(x) > 0$ for all x.

PROOF: By hypothesis, for each x there is a v. By semi-continuity, this v works for all x' in some neighborhood of x. By compactness (we are working in projective space), finitely many of these neighborhoods cover. By convexity, we may piece these together with a partition of unity while staying inside $K^{p,C}(x)$ at each point x. □

Consequences

I.

This is used by [ABG70] to deform the chain $x + i\mathbb{R}^d$ over which the inverse Fourier transform is integrated into a conical chain $\mathcal{C}$.



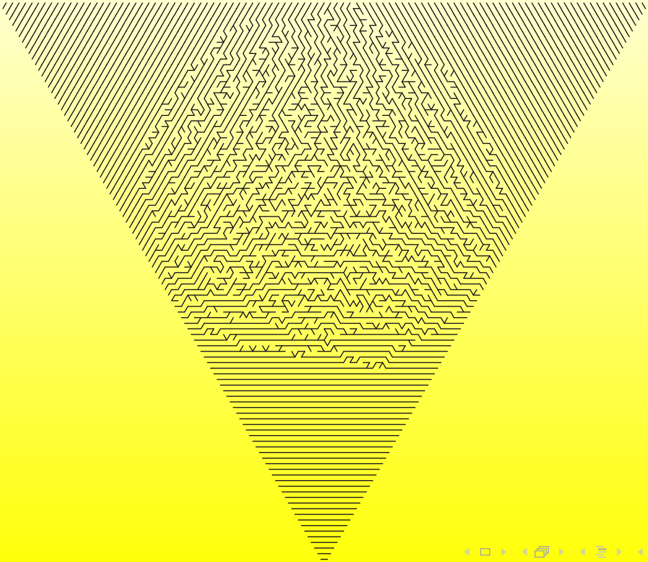
application to multivariate generating functions

II.

Let us see how this was used in [BP11] to compute asymptotics of the Taylor series for

$$\begin{aligned} F(x,y,z) &= \frac{1}{(1-Z)(3-X-Y-Z-XY-XZ-YZ+3XYZ)} \\ &= \sum_r a(r,s,t) X^r Y^s Z^t \qquad (2) \end{aligned}$$

where $a(r,s,t)$ is the probability of the **cube grove** of order $r+s+t$ having a horizontal edge at barycentric coordinate (r,s,t).

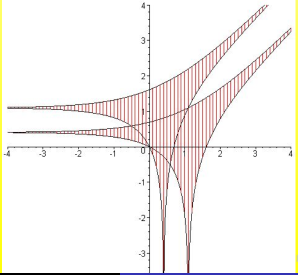


Amoebas

To see how one applies the [ABG70] theory to rational generating functions, we need one more definition. The **amoeba** of a polynomial q is the image of its zero set under the coordinatewise log-modulus map

$$(z_1, \ldots, z_d) \mapsto (\log |z_1|, \ldots, \log |z_d|) .$$

The connected components of the complement of any amoeba are open convex sets.

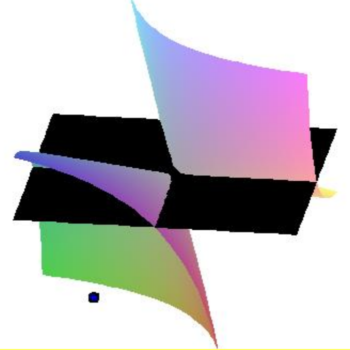


Cauchy's integral formula

The method works for any rational function P/Q. Any Laurent expansion of P/Q is convergent on some component B of the complement of amoeba(Q) and its coefficients are given there by Cauchy's formula, where x is any point in B:

$$a_{rst} = (2\pi i)^{-3} \int_{x+iT^d} e^{-r \cdot z} f(z)\, dz\,. \tag{3}$$

Here we have changed to logarithmic coordinates, so $f(z) := F(e^{z_1}, \ldots, e^{z_d})$ and $T^d := (\mathbb{R}/2\pi\mathbb{Z})^d$.



The imaginary fiber through any point in a cone of hyperbolicity, such as the point shown, does not intersect the zero set of q.

Hyperbolicity on the amoeba boundary

Let x be any point on the common boundary of amoeba(Q) and one of the components B of its complement. Let $q := Q \circ \exp$ denote Q in logarithmic coordinates.

Hyperbolicity on the amoeba boundary

Let x be any point on the common boundary of amoeba(Q) and one of the components B of its complement. Let $q := Q \circ \exp$ denote Q in logarithmic coordinates.

Theorem 13 (hyperbolicity on the amoeba boundary)

The polynomial $\operatorname{loc}(q, x)$ *is hyperbolic and has a cone of hyperbolicity,* C*, containing the geometric tangent cone* $\tan_x(B)$.

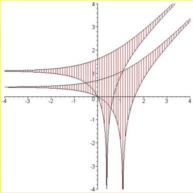
Hyperbolicity on the amoeba boundary

Let x be any point on the common boundary of amoeba(Q) and one of the components B of its complement. Let $q := Q \circ \exp$ denote Q in logarithmic coordinates.

Theorem 13 (hyperbolicity on the amoeba boundary)

The polynomial $\text{loc}(q, x)$ *is hyperbolic and has a cone of hyperbolicity,* C*, containing the geometric tangent cone* $\tan_x(B)$*.*

If x is the origin then C is the directions between 5:00 and 10:00



projective integral

It follows that the integral computing a_{rst} can be pushed onto a cone pointing outward from x.

projective integral

It follows that the integral computing a_{rst} can be pushed onto a cone pointing outward from x.

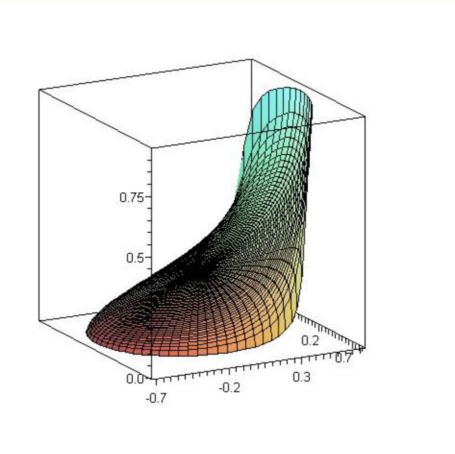
By the apparatus in [ABG70], one then has

$$a_r \sim E(r), \qquad E(r) = \widehat{\left(\frac{1}{Q}\right)},$$

where $\frac{1}{Q}$ is the inverse Fourier transform of $1/q$, otherwise known as the fundamental solution to the wave equation $D_q E = \delta$.

When Q is the denominator of (2), this gives

$$a_{rst} \sim \frac{1}{\pi} \arctan \left(\frac{\sqrt{2(rs + rt + st) - (r^2 + s^2 + t^2)}}{r + s - t} \right) .$$



$$a_{rst} \sim \frac{1}{\pi} \arctan \left(\frac{\sqrt{2(rs + rt + st) - (r^2 + s^2 + t^2)}}{r + s - t} \right) .$$

END PART I

II: Applications of stability in probability and combinatorics

Robin Pemantle

Current Developments in Mathematics, 18 November 2011

Outline

- Univariate stable functions

Outline

Outline

I Univariate stable functions

II Multivariate stable functions

III Negatively dependent random variables

Outline

I Univariate stable functions

II Multivariate stable functions

III Negatively dependent random variables

IV Multi-affine stability / strong Rayleigh property

Outline

I Univariate stable functions

II Multivariate stable functions

III Negatively dependent random variables

IV Multi-affine stability / strong Rayleigh property

Plan: go through I and II catalogue style: many statements, few proofs, and hitting highlights rather than results that build on each other. Then, for III and IV, try to give a coherent development.

Univariate stable functions

Real roots

A univariate stable polynomial f is by definition one with no roots in the open upper half plane.

Real roots

A univariate stable polynomial f is by definition one with no roots in the open upper half plane.

If f is real, then the set of roots is invariant under conjugation, so f has no roots in the lower half plane either, hence has all real roots.

Real roots

A univariate stable polynomial f is by definition one with no roots in the open upper half plane.

If f is real, then the set of roots is invariant under conjugation, so f has no roots in the lower half plane either, hence has all real roots.

If, additionally, the coefficients of f are nonnegative, then all roots of f are in $(-\infty, 0]$. Polynomials whose roots are all real and nonpositive have useful properties. Let us denote this class of univariate polynomials by RR.

Proposition 14

If $f \in \mathrm{RR}$ *then* $f/f(1)$ *is the probability generating function for a sum of independent Bernoulli random variables.* $\square$

Proposition 14

If $f \in \mathrm{RR}$ *then* $f/f(1)$ *is the probability generating function for a sum of independent Bernoulli random variables.* □

Proposition 15 (Pólya frequency criterion, Edrei 1953)

A polynomial with nonnegative real coefficients is in RR *if and only if its sequence of coefficients* $(a_0, \ldots, a_d)$ *is a* **Pólya frequency sequence**, *meaning that all the minors of the matrix* (a_{n-k}) *have nonnegative determinant.*

Proposition 14

If $f \in \mathrm{RR}$ then $f/f(1)$ is the probability generating function for a sum of independent Bernoulli random variables. □

Proposition 15 (Pólya frequency criterion, Edrei 1953)

A polynomial with nonnegative real coefficients is in RR *if and only if its sequence of coefficients $(a_0, \ldots, a_d)$ is a* **Pólya frequency sequence**, *meaning that all the minors of the matrix (a_{n-k}) have nonnegative determinant.*

Corollary 16 (log-concavity)

The coefficient sequence of any polynomial in RR *is log-concave.* □

Ultra-log concavity

In fact the coefficients of any $f \in$ RR are **ultra-logconcave**, meaning that $\{a_k / \binom{d}{k}\}$ is log-concave. These inequalities are due to Newton (1707).

Ultra-log concavity

In fact the coefficients of any $f \in \text{RR}$ are **ultra-logconcave**, meaning that $\{a_k / \binom{d}{k}\}$ is log-concave. These inequalities are due to Newton (1707).

Perhaps the single most useful theorem about the class of (complex) stable polynomials is that it is closed under coefficient-wise multiplication by **Pólya-Schur multiplier sequences**. Say that a sequence $\{\lambda(0), \lambda(1), \ldots\}$ is a **multiplier sequence** if

$$f = \sum_{n=0}^{\infty} a_n x^n \in \text{RR implies } T(f) := \sum_{n=0}^{\infty} \lambda(n) a_n x^n \in \text{RR} \, .$$

Pólya-Schur Theorem

Theorem 17 (Pólya-Schur, 1914)

Let $\phi(z) := \sum_n \lambda(n) z^n / n!$ *be the exponential generating function for the sequence* λ*. The following are equivalent.*

(i) λ *is a multiplier sequence.*

(ii) ϕ *is entire and either* $\phi(z)$ *or* $\phi(-z)$ *is the uniform limit on compact sets of polynomials in* RR.

(iii) *Either* $\phi(z)$ *or* $\phi(-z)$ *is entire and can be written as* $Cz^n e^{az} \prod_{k=1}^{\infty} (1 + \alpha_k z)$ *for a summable sequence of nonnegative numbers* $\{\alpha_k\}$.

(iv) *For all integers* $n > 0$*, the polynomial* $T[(1+z)^n]$ *is hyperbolic with zeros all of the same sign.*

Multivariate stability

Multivariate stability

Multivariate stability

Multivariate stability

The presentation owes a debt to David Wagner's recent survey article in the AMS Bulletin [Wag11].

Multivariate stability

Multivariate stability

The presentation owes a debt to David Wagner's recent survey article in the AMS Bulletin [Wag11].

Definition 18

Recall: a complex polynomial q in d variables is said to be **stable** *if $q(z_1, \ldots, z_d) = 0$ implies not all coordinates z_j are in the open upper half plane.*

Easy properties

Proposition 19 (easy closure properties)

The class of stable polynomials is closed under the following.

(a) *Products: f and g are stable implies fg is stable;*

Easy properties

Proposition 19 (easy closure properties)

The class of stable polynomials is closed under the following.

(a) *Products: f and g are stable implies fg is stable;*

(b) *Index permutations: f is stable implies $f(x_{\pi(1)}, \ldots, x_{\pi(d)})$ is stable where $\pi \in \mathcal{S}_d$;*

Easy properties

Proposition 19 (easy closure properties)

The class of stable polynomials is closed under the following.

(a) *Products: f and g are stable implies fg is stable;*

(b) *Index permutations: f is stable implies $f(x_{\pi(1)}, ..., x_{\pi(d)})$ is stable where $\pi \in \mathcal{S}_d$;*

(c) *Diagonalization: f is stable implies $f(x_1, x_1, x_3, \ldots, x_d)$ is stable;*

Easy properties

Proposition 19 (easy closure properties)

The class of stable polynomials is closed under the following.

(a) *Products: f and g are stable implies fg is stable;*

(b) *Index permutations: f is stable implies $f(x_{\pi(1)}, \ldots, x_{\pi(d)})$ is stable where $\pi \in \mathcal{S}_d$;*

(c) *Diagonalization: f is stable implies $f(x_1, x_1, x_3, \ldots, x_d)$ is stable;*

(d) *Specialization: if f is stable and $\mathrm{Im}\,(a) \geq 0$ then $f(a, x_2, \ldots, x_d)$ is stable;*

Easy properties

Proposition 19 (easy closure properties)

The class of stable polynomials is closed under the following.

(a) *Products: f and g are stable implies fg is stable;*

(b) *Index permutations: f is stable implies $f(x_{\pi(1)}, \ldots, x_{\pi(d)})$ is stable where $\pi \in \mathcal{S}_d$;*

(c) *Diagonalization: f is stable implies $f(x_1, x_1, x_3, \ldots, x_d)$ is stable;*

(d) *Specialization: if f is stable and $\mathrm{Im}(a) \geq 0$ then $f(a, x_2, \ldots, x_d)$ is stable;*

(e) *Inversion: if the degree of x_1 in f is m and f is stable then $x_1^m f(-1/x_1, x_2, \ldots, x_d)$ is stable;*

Differentiation

Lemma 20 (differentiation)

If f is stable then $\partial f / \partial x_j$ is either stable or identically zero.

Differentiation

Lemma 20 (differentiation)

If f is stable then $\partial f / \partial x_j$ is either stable or identically zero.

PROOF: Fix any values of $\{x_i : i \neq j\}$ in the upper half plane. As a function of x_j, f has no zeros in the upper half plane. By the Gauss-Lucas theorem, the zeros of f' are in the convex hull, therefore not in the upper half plane. □

Differentiation

Lemma 20 (differentiation)

If f is stable then $\partial f / \partial x_j$ is either stable or identically zero.

PROOF: Fix any values of $\{x_i : i \neq j\}$ in the upper half plane. As a function of x_j, f has no zeros in the upper half plane. By the Gauss-Lucas theorem, the zeros of f' are in the convex hull, therefore not in the upper half plane. □

The next property, Wagner calls an "astounding" recent generalization of the Pólya-Schur theorem.

Multivariate Pólya-Schur theorem

This characterizes not just multiplier sequence but all $\mathbb{C}$-linear maps preserving stability. To restrict to multiplier sequences, take $T(x^\alpha) = \lambda(\alpha) x^\alpha$.

Multivariate Pólya-Schur theorem

This characterizes not just multiplier sequence but all $\mathbb{C}$-linear maps preserving stability. To restrict to multiplier sequences, take $T(x^\alpha) = \lambda(\alpha) x^\alpha$.

Theorem 21 ([BB09, Theorem 1.3])

The $\mathbb{C}$-linear map $T : \mathbb{C}[x] \to \mathbb{C}[x]$ preserves stable polynomials if and only if either its range is scalar multiples of a single stable polynomial or the series

$$\sum_{\alpha \in (\mathbb{Z}^+)^d} (-1)^{|\alpha|} T(x^\alpha) \frac{y^\alpha}{\alpha!}$$

is a uniform limit on compact sets of stable polynomials in $\mathbb{C}[x, y]$.

What you are missing, part IV: any hint of the proof

The crucial step is to establish a criterion reminiscent to criterion (iv) in the univariate case (that $T[(1+z)^n]$ always has real roots of the same sign):

A power series $\sum_\alpha P_\alpha(x) y^\alpha$ whose coefficients are polynomials in x is in the closure of stable polynomials in $\mathbb{C}[x,y]$ if and only if for all $\beta \in (\mathbb{Z}^+)^d$,

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

is stable in $\mathbb{C}[x,y]$.

Determinants

A number of theorems and conjectures about stable polynomials have to do with determinants. I will describe one classical result, one recent result and one conjecture that is still open.

Determinants

A number of theorems and conjectures about stable polynomials have to do with determinants. I will describe one classical result, one recent result and one conjecture that is still open.

1 Stability of $\det(A + x_1 B_1 + \cdots + x_d B_d)$.

Determinants

A number of theorems and conjectures about stable polynomials have to do with determinants. I will describe one classical result, one recent result and one conjecture that is still open.

1 Stability of $\det(A + x_1 B_1 + \cdots + x_d B_d)$.
2 Real roots of the mixed determinant $\det(xA, -B)$.

Determinants

A number of theorems and conjectures about stable polynomials have to do with determinants. I will describe one classical result, one recent result and one conjecture that is still open.

1. Stability of $\det(A + x_1 B_1 + \cdots + x_d B_d)$.
2. Real roots of the mixed determinant $\det(xA, -B)$.
3. Nonnegative coefficients of the polynomial $\lambda \mapsto \mathrm{Tr}(A + \lambda B)^n$.

Positive definite matrices

1\.

A classical example of hyperbolicity already cited in [Går51] that if A is Hermitian and B is nonnegative definite then $t \mapsto \det(A + tB)$ has only real zeros.

Positive definite matrices

1.

A classical example of hyperbolicity already cited in [Går51] that if A is Hermitian and B is nonnegative definite then $t \mapsto \det(A + tB)$ has only real zeros.

The multivariate generalization of this is that if A is Hermitian and $B_1, \ldots, B_d$ are positive definite, then

$$f := \det(A + x_1 B_1 + \cdots + x_d B_d) \qquad (4)$$

is a real stable polynomial in $x_1, \ldots, x_d$. Note: if A is also positive definite, then f has positive coefficients.

Easy proof

The proof is more or less the same as the proof of hyperbolicity of $\det(A + tB)$.

Fix the real part of $x_1, \ldots, x_d$, all positive, and remove a factor of the positive definite square root of $Q := \sum \mathrm{Re}\,\{x_j\} B_j$ on both the right and the left to obtain $\det Q \det(iI + H)$ where H is Hermitian (subsuming $Q^{-1/2} A Q^{-1/2}$ as well as the similar term with A replaced by the sum of $\mathrm{Im}\,\{x_i\} B_i$. The eigenvalues of H are real, hence cannot equal $-i$. $\square$

Mixed determinants

2.

If A and B are $n \times n$ matrices, define the **mixed determinant**

$$\det(A, B) := \sum_{S \subseteq [n]} \det(A|_S) \det(B|_{S^c}) .$$

The definition for k matrices instead of two is analogous, substituting a partition into k parts for $\{S, S^c\}$.

Mixed determinants

2.

If A and B are $n \times n$ matrices, define the **mixed determinant**

$$\det(A, B) := \sum_{S \subseteq [n]} \det(A|_S) \det(B|_{S^c}) .$$

The definition for k matrices instead of two is analogous, substituting a partition into k parts for $\{S, S^c\}$.

Conjecture (Johnson's conjecture)

If A is Hermitian and B is positive definite then $\det(xB, -A)$ *has only real roots.*

This (and much more) was recently proved by Borcea and Brändén [BB08].

BMV conjecture

3.

Conjecture (Bessis-Moussa-Villani 1975)

If A is Hermitian and B is nonnegative definite then $\lambda \to \mathrm{Tr}(\exp(A - \lambda B))$ is the Laplace transform of a positive measure on $[0, \infty)$.

BMV conjecture

3.

Conjecture (Bessis-Moussa-Villani 1975)

If A is Hermitian and B is nonnegative definite then $\lambda \to \mathrm{Tr}(\exp(A - \lambda B))$ is the Laplace transform of a positive measure on $[0, \infty)$.

This has been proved for 2×2 matrices but is still open for all sizes greater than 2. This was shown in 2004 to be equivalent to the following.

> *For all nonnegative definite matrices A and B and all integers $n > 0$, the polynomial $\lambda \mapsto \mathrm{Tr}(A + \lambda B)^n$ has nonnegative coefficients.*

Negative dependence

Negatively dependent random variables

Binary-valued random variables

Let $\mathcal{B}_n := \{0, 1\}^n$ denote the Boolean lattice of rank n. The joint law of n binary random variables is a measure μ on $\mathcal{B}_n$. The probability generating function $f = f_\mu$ is given by

$$f_\mu(x_1, \ldots, x_n) := \sum_{\omega \in \mathcal{B}_n} \mu(\omega) \prod_{j=1}^{n} x_j^{\omega_j} = \mathbb{E} x^\omega .$$

Binary-valued random variables

Let $\mathcal{B}_n := \{0,1\}^n$ denote the Boolean lattice of rank n. The joint law of n binary random variables is a measure μ on $\mathcal{B}_n$. The probability generating function $f = f_\mu$ is given by

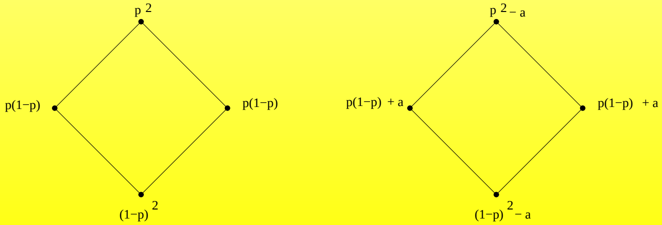
$$f_\mu(x_1, \ldots, x_n) := \sum_{\omega \in \mathcal{B}_n} \mu(\omega) \prod_{j=1}^{n} x_j^{\omega_j} = \mathbb{E} x^\omega \,.$$

Probability generating functions for measures on $\mathcal{B}_n$ all share two properties: they are **multi-affine**, meaning that no variable appears with a power greater than one, and their coefficients are real and nonnegative.

Negative correlation

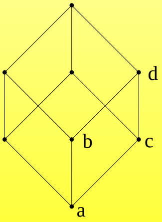
Example 22 (n=2)

The function $f(x,y) := p^2xy + p(1-p)x + p(1-p)y + (1-p)^2$ *generates two IID coin flips with success probability* p*. The function* $f(x,y) + a(xy - x - y + 1)$ *generates two exchangeable* p*-coins that are positively correlated if* $a < 0$ *and negatively correlated if* $a > 0$*.*



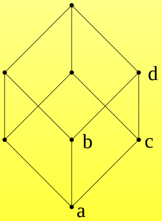
Lattice conditions

A 4-tuple (a, b, c, d) of the Boolean lattice $\mathcal{B}_n$ is a **diamond** if b and c cover a and if d covers b and c, where x covers y if $x \geq y$ and $x \geq u \geq y$ implies $u = x$ or $u = y$.



Lattice conditions

A 4-tuple (a, b, c, d) of the Boolean lattice $\mathcal{B}_n$ is a **diamond** if b and c cover a and if d covers b and c, where x covers y if $x \geq y$ and $x \geq u \geq y$ implies $u = x$ or $u = y$.



Say that μ satisfies the **positive lattice condition** if $\mu(b)\mu(c) \leq \mu(a)\mu(d)$ for every diamond (a, b, c, d). The reverse inequality is called the **negative lattice condition**.

FKG

The positive lattice condition is very useful, due to the following result.

Theorem 23 (FKG)

If μ satisfies the positive lattice condition then μ is positively associated and the projection of μ to any smaller set of variables satisfies both these conditions as well.

Here, **positively associated** means that

$$\mathbb{E}_\mu fg \geq (\mathbb{E}_\mu f) \ (\mathbb{E}_\mu g)$$

whenever f and g are both monotone increasing on μ_n.

Negative association

Every function is positively correlated with itself. Thus, to define negative association, we need to do more than reverse the inequality.

Negative association

Every function is positively correlated with itself. Thus, to define negative association, we need to do more than reverse the inequality.

We say that μ is **negatively associated** if $\mathbb{E}_\mu fg \leq (\mathbb{E}_\mu f)(\mathbb{E}_\mu g)$ whenever f and g are monotone increasing on $\mathcal{B}_n$ and in addition there is a set $S \subseteq \{1, \ldots, n\}$ such that f depends only on $\{\omega_j : j \in S\}$ and g depends only on $\{\omega_j : j \notin S\}$.

Negative association

Every function is positively correlated with itself. Thus, to define negative association, we need to do more than reverse the inequality.

We say that μ is **negatively associated** if $\mathbb{E}_\mu fg \leq (\mathbb{E}_\mu f)(\mathbb{E}_\mu g)$ whenever f and g are monotone increasing on $\mathcal{B}_n$ and in addition there is a set $S \subseteq \{1, \ldots, n\}$ such that f depends only on $\{\omega_j : j \in S\}$ and g depends only on $\{\omega_j : j \notin S\}$.

Historically, a number of notions of negative dependence have been defined. The weakest is pairwise negative correlation. Negative association is the strongest one that was suspected to hold in many examples.

In search of a theory

Unfortunately, the negative lattice condition does not imply negative association. In fact the NLC is not closed under passing to subsets.

In search of a theory

Unfortunately, the negative lattice condition does not imply negative association. In fact the NLC is not closed under passing to subsets.

Up until [BBL09] there was no satisfactory theory of negative dependence. The paper [Pem00] attempted, but failed, to find one. Say that μ has property h-NLC$^+$ if every measure obtained from μ by ignoring a subset of the variables or applying an external field has the NLC. Here an **external field** means multiplying each $\mu(\omega)$ by $\prod_j \lambda_j^{\omega_j}$ and renormalizing.

In search of a theory

Unfortunately, the negative lattice condition does not imply negative association. In fact the NLC is not closed under passing to subsets.

Up until [BBL09] there was no satisfactory theory of negative dependence. The paper [Pem00] attempted, but failed, to find one. Say that μ has property h-NLC$^+$ if every measure obtained from μ by ignoring a subset of the variables or applying an external field has the NLC. Here an **external field** means multiplying each $\mu(\omega)$ by $\prod_j \lambda_j^{\omega_j}$ and renormalizing.

Conjecture ([Pem00])

h-NLC$^+$ implies negative association.

Theory found!

It turns out that the "right" condition is not h-NLC$^+$. Say that the measure μ on $\mathcal{B}_n$ is **strong Rayleigh** if its generating function $f = f_\mu$ is stable.

Theorem 24 ([BBL09])

If μ is strong Rayleigh then μ is negatively associated.

Theory found!

It turns out that the "right" condition is not h-NLC$^+$. Say that the measure μ on $\mathcal{B}_n$ is **strong Rayleigh** if its generating function $f = f_\mu$ is stable.

Theorem 24 ([BBL09])

If μ is strong Rayleigh then μ is negatively associated.

In a short while I will give some indication of how this is proved.

More about strong Rayleigh measures

Strong Rayleigh measures have more nice properties. For example, they are closed under partial symmetrization.

More about strong Rayleigh measures

Strong Rayleigh measures have more nice properties. For example, they are closed under partial symmetrization.

Let μ_{ij} denote μ with the indices i and j transposed. If μ is SR then for any i, j and any $\theta \in [0, 1]$ the measure $\theta\mu + (1 - \theta)\mu_{ij}$ is SR. This is [BBL09, Theorem 4.20] and it leads to the very nice result:

More about strong Rayleigh measures

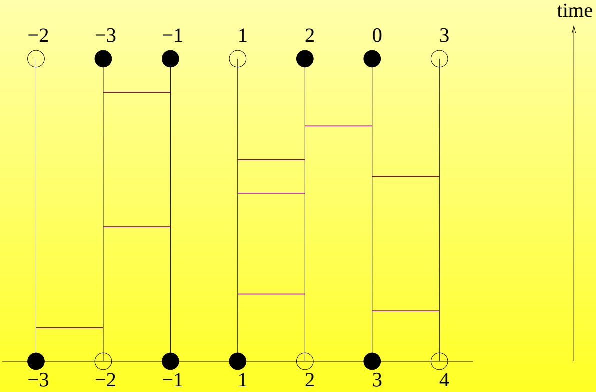
Strong Rayleigh measures have more nice properties. For example, they are closed under partial symmetrization.

Let μ_{ij} denote μ with the indices i and j transposed. If μ is SR then for any i, j and any $\theta \in [0, 1]$ the measure $\theta\mu + (1 - \theta)\mu_{ij}$ is SR. This is [BBL09, Theorem 4.20] and it leads to the very nice result:

Theorem 25 (exclusion dynamics)

Begin with a configuration in $\mathcal{B}_n$. For each i, j, swap the values ω_i and ω_j at some rate β_{ij}. Then the law of the configuration at any time t is SR.

Nearest neighbor exclusion process on $\mathbb{Z}$



Complex geometry

I will not be proving these two results, but I will quote the authors as to one aspect the methodology.

> *Clearly [this result] is real stable in nature. However, in order to establish it, we have to consider the (wider) complex stable context.*

Complex geometry

I will not be proving these two results, but I will quote the authors as to one aspect the methodology.

> *Clearly [this result] is real stable in nature. However, in order to establish it, we have to consider the (wider) complex stable context.*

In fact, they prove that stability of f_μ implies stability of $f_{\theta\mu+(1-\theta)\mu_{ij}}$ when μ is any *complex-valued* measure. It is important that f be multi-affine, but evidently not that f be real.

Another case of extending beyond $\mathbb{R}^+$

The property h-NLC$^+$ turns out to be equivalent to the (ordinary) Rayleigh property defined as the following inequality for all positive vectors x:

$$\forall i, j \qquad \frac{\partial f}{\partial x_i}(x) \, \frac{\partial f}{\partial x_j}(x) \geq f(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \, . \tag{5}$$

For multi-affine real polynomials, stability is equivalent to (5) for all $x \in \mathbb{R}^n$. Thus SR differs from h$=$NLC$^+$ in that the inequality is required for all x rather than for positive x. This should perhaps be called h-NLC$^\pm$ because it is h-NLC with closure under external fields both positive and negative!

Multi-affine stable functions

The theory of multi-affine stable functions

Polarization

Let f be a polynomial of degree m in one variable and define the polarization of f to be the result of replacing x^j be the normalized elementary symmetric function $\binom{m}{j}^{-1} e_j(x_1, \ldots, x_m)$. If f is a probability generating function then the event $\{X = j\}$ has been replaced by the event $\{\sum_{1=1}^m X_1 = j\}$.

Lemma 26 (polarization)

1. If f is univariate stable then its polarization is stable.

2. If f is multivariate stable then the analogous polarization, replacing x_i^j in each monomial by $e_j(x_{i1}, \ldots, x_{i\ell})$, is stable.

Polarization

Let f be a polynomial of degree m in one variable and define the polarization of f to be the result of replacing x^j be the normalized elementary symmetric function $\binom{m}{j}^{-1} e_j(x_1, \ldots, x_m)$. If f is a probability generating function then the event $\{X = j\}$ has been replaced by the event $\{\sum_{1=1}^m X_1 = j\}$.

Lemma 26 (polarization)

1. If f is univariate stable then its polarization is stable.

2. If f is multivariate stable then the analogous polarization, replacing x_i^j in each monomial by $e_j(x_{i1}, \ldots, x_{i\ell})$, is stable.

(In fact part 1 is if and only if)

Symmetric homogenization

The polarization lemma is proved via the classical Grace-Walsh-SzegHo coincidence theorem.

Symmetric homogenization

The polarization lemma is proved via the classical Grace-Walsh-SzegHo coincidence theorem.

It is then used by [BBL09] to prove the following result.

Symmetric homogenization

The polarization lemma is proved via the classical Grace-Walsh-SzegHo coincidence theorem.

It is then used by [BBL09] to prove the following result.

Suppose μ is a measure on $\mathcal{B}_n$ and define the **symmetric homogenization** μ_{sh} of μ to be the measure on μ_{2n} that is symmetric on $x_{d+1}, \ldots, x_{2n}$, restricts to μ when $x_1, \ldots, x_n$ are set equal to 1, and is n-homogeneous. In other words, to pick from μ_{sh}, first sample $X_1, \ldots, X_n$ from μ, then if these sum to k, choose $n - k$ indices uniformly from $n + 1$ to $2n$, set those variables equal to 1 and the rest of $X_{n+1}, \ldots, X_{2n}$ to zero.

sh preserves SR

Lemma 27

If μ is strongly Rayleigh then so is μ_{sh}.

sh preserves SR

Lemma 27

If μ is strongly Rayleigh then so is μ_{sh}.

PROOF:

sh preserves SR

Lemma 27

If μ is strongly Rayleigh then so is μ_{sh}.

PROOF: First one shows that the usual homogenization of a polynomial, multiplying each monomial by an appropriate power of x_{d+1}, preserves stability.

sh preserves SR

Lemma 27

If μ is strongly Rayleigh then so is μ_{sh}.

PROOF: First one shows that the usual homogenization of a polynomial, multiplying each monomial by an appropriate power of x_{d+1}, preserves stability. This follows from facts about hyperbolicity found in Gårding's original paper.

sh preserves SR

Lemma 27

If μ is strongly Rayleigh then so is μ_{sh}.

PROOF: First one shows that the usual homogenization of a polynomial, multiplying each monomial by an appropriate power of x_{d+1}, preserves stability. This follows from facts about hyperbolicity found in Gårding's original paper. Denoting this homogenization by μ_*, we see that μ_{sh} is the polarization of μ_*, so the result follows from the Polarization lemma. □

SR implies NA

Proof that strong Rayleigh measures are negatively associated:

SR implies NA

PROOF THAT STRONG RAYLEIGH MEASURES ARE NEGATIVELY ASSOCIATED:

1. Pass to μ_{sh}.

SR implies NA

PROOF THAT STRONG RAYLEIGH MEASURES ARE NEGATIVELY ASSOCIATED:

1. Pass to μ_{sh}.
2. Observe that SR implies Rayleigh which implies pairwise negative correlation.

SR implies NA

PROOF THAT STRONG RAYLEIGH MEASURES ARE NEGATIVELY ASSOCIATED:

1. Pass to μ_{sh}.
2. Observe that SR implies Rayleigh which implies pairwise negative correlation.
3. The original proof by Feder and Mihail [FM92] of negative association for **balanced matroids** now goes through, with homogeneity of μ_{sh} providing the "balance" property.

□

M. Atiyah, R. Bott, and L. Gårding.
Lacunas for hyperbolic differential operators with constant coefficients, I.
Acta Mathematica, 124:109–189, 1970.

J. Borcea and P. Brändén.
Applications of stable polynomials to mixed determinants.
Duke J. Math., 143:205–223, 2008.

J. Borcea and P. Brändén.
The Lee-Yang and Pólya-Schur programs, I: linear operators preserving stability.
Invent. Math., 177:541–569, 2009.

J. Borcea, P. Brändén, and T. Liggett.
Negative dependence and the geometry of polynomials.
J. AMS, 22:521–567, 2009.

Y. Baryshnikov and R. Pemantle.
Asymptotics of multivariate sequences, part iii: quadratic points.
Adv. Math., 228:3127–3206, 2011.

T. Feder and M. Mihail.
Balanced matroids.
Annual ACM Symposium on Theory of Computing, pages 26–38, 1992.

L. Gårding.
Linear hyperbolic partial differential equations with constant coefficients.
Acta Math., 85:1–62, 1951.

R. Pemantle.
Toward a theory of negative dependence.
J. Math. Phys., 41:1371–1390, 2000.

David G. Wagner.
Multivariate stable polynomials: theory and application.
Bull. AMS, 48:53–84, 2011.