Around Descartes' rule of signs

joint with J. Forsgård, Vl. Kostov, and Dm. Novikov

Boris Shapiro

March 30, 2016

Topics to discuss

Main references

J. Forsgård, Vl. Kostov, and B. Shapiro, **Could René Descartes have known this?**, Exper. Math., 24:4, (2015) 438–448.

J. Forsgård, Dm. Novikov, and B. Shapiro, **A tropical analog of Descartes' rule of signs**, IMRN, to appear.

I think, therefore I am

Figure: Renatus Cartesius in person.

Some history

The father of modern philosophy and mathematician René Descartes (in Latin Renatus Cartesius) who spent most of his life in the Dutch Republic, died on 11 February 1650 in Stockholm, Sweden. He had been invited by Queen Christina of Sweden to tutor her. The cause of death was said to be pneumonia.

One theory claims that accustomed to working in bed until noon, he may have suffered damage to his health from Christina's study regime, which began early in the morning at 5 a.m.

Some history, cont.1

Another theory says that he might have been poisoned with arsenic for the following reason. At this time Queen Christina intended to convert to Catholicism and later she actually did that and abdicated her throne as Swedish law requires a Protestant ruler. The only Catholic with whom she had prolonged contact had been Descartes which might have caused the intense hatred by the Swedish Protestant clergy.

Some history, cont.2

On the other hand, another lead says that he might have been poisoned by a local Catholic priest who was afraid that Descartes' radical religious ideas might interfere with Christina's intention to convert. In any case already in 1663, the Pope placed his works on the Index Librorum Prohibitorum (Index of Prohibited Books).

As a Catholic in a Protestant nation, he was interred in a graveyard used mainly for unbaptized infants in Adolf Fredriks Kyrka in Stockholm. In 1667 his remains were taken to France and first buried in the Abbey of Sainte Geneviève.

Some history, cont.3

Although the National Convention in 1792 had planned to transfer his remains to the Panthéon, in 1819 they were moved to the Abbey of Saint-Germain-des-Prés in Paris. Two centuries later, they are still resting in a chapel of the latter abbey while his memorial, erected in the 18th century, remains in the Swedish church.

As if all that was not enough already, it seems that Descartes' body was shipped without his head due to the small size of the coffin sent for him from France, and his cranium was delivered to "Musée de l'Homme“ in Paris only in early 19-th century by the famous Swedish chemist J. Berzelius who bought it on an auction in Stockholm!

One wonders if there is any reasonable evidence behind any of these myths.

Intro

In 1637, Descartes published his ground-breaking philosophical and mathematical treatise "Discours de la méthode" where he explained how to solve all mathematical problems.

In particular, he described his famous rule of signs claiming that the number of positive roots of a real univariate polynomial does not exceed the number of sign changes in its sequence of coefficients.

In what follows we only consider real polynomials with all non-vanishing coefficients.

An arbitrary ordered sequence $\overline{\sigma} = (\sigma_0, \sigma_1, ..., \sigma_d)$ of $\pm$-signs is called a *sign pattern*.

Given a sign pattern $\overline{\sigma}$, we call its *Descartes' pair* $(p_{\overline{\sigma}}, n_{\overline{\sigma}})$ the pair of non-negative integers counting sign changes and sign preservations of $\overline{\sigma}$.

Definitions

The Descartes' pair of $\overline{\sigma}$ gives the upper bound on the number of positive and negative roots of any polynomial of degree d whose signs of coefficients are given by $\overline{\sigma}$. (Observe that, for any $\overline{\sigma}$, $p_{\overline{\sigma}} + n_{\overline{\sigma}} = d$.)

To any polynomial $q(x)$ with the sign pattern $\overline{\sigma}$, we associate the pair (pos_q, neg_q) giving the numbers of its positive and negative roots counted with multiplicities. Obviously the pair (pos_q, neg_q) satisfies the standard restrictions

$$pos_q \leq p_{\overline{\sigma}},\ pos_q \equiv p_{\overline{\sigma}} (\bmod 2),\ neg_q \leq n_{\overline{\sigma}},\ neg_q \equiv n_{\overline{\sigma}} (\bmod 2). \tag{1}$$

Main Problem

We call pairs (pos, neg) satisfying (1) *admissible* for $\overline{\sigma}$.

It turns out that not for every pattern $\overline{\sigma}$, all its admissible pairs (pos, neg) are realizable by polynomials with the sign pattern $\overline{\sigma}$.

Below we address this very basic question.

Problem

For a given sign pattern $\overline{\sigma}$, which admissible pairs (pos, neg) are realizable by polynomials whose signs of coefficients are given by $\overline{\sigma}$?

First results

To the best of our knowledge, this natural question was for the first time raised by B. Anderson, J. Jackson, M. Sitharam in 1998. Soon after that D. J. Grabiner found the first example of non-realizable combination for polynomials of degree 4.

Namely, he has shown that the sign pattern $(+,-,-,-,+)$ does not allow to realize the pair $(0,2)$ and the sign pattern $(+,+,-,+,+)$ does not allow to realize $(2,0)$. Observe that their Descartes' pairs are equal to $(2,2)$.

Proof

The argument is very simple. (Due to symmetry induced by $x \mapsto -x$ it suffices to consider only the first case.)

Observe that a fourth-degree polynomial with only two negative roots for which the sum of roots is positive could be factored as $a(x^2 + bx + c)(x^2 - sx + t)$ with $a, b, c, s, t > 0,\ s^2 < 4t$, and $b^2 \geq 4c$. The product of these factors equals $a(x^4 + (b - s)x^3 + (t + c - bs)x^2 + (bt - cs)x + ct)$.

To get the correct sign pattern, we need $b < s$ and $bt < cs$, which gives $b^2t < s^2c$ and thus $b^2/c < s^2/t$. But

$$b^2/c \geq 4 > s^2/t.$$

Notation

For any pair (d, k) of non-negative integers with $d - k \geq 0$; $d - k \equiv 0 \mod 2$, denote by $Pol_{d,k}$, the set of all monic real polynomials of degree d with k real simple roots.

Denote by $Pol_d(\widehat{\sigma}) \subset Pol_d$ the set (orthant) of all monic polynomials $p = x^d + a_1 x^{d-1} + ... + a_d$ whose coefficients $(a_1, ..., a_d)$ have the (shortened) sign pattern $\widehat{\sigma} = (\sigma_1, ..., \sigma_d)$ respectively. Finally, set $Pol_{d,k}(\widehat{\sigma}) = Pol_{d,k} \cap Pol_d(\widehat{\sigma})$.

To the best of our knowledge, for arbitrary (d, k) and $\widehat{\sigma}$, the topology of $Pol_{d,k}(\widehat{\sigma})$ has not been previously studied.

Notation, cont.

We have the natural $\mathbb{Z}_2 \times \mathbb{Z}_2$-action on the space of monic polynomials and on the set of all sign patterns respectively.

The first generator acts by reverting the signs of all monomials of odd degree (which for polynomials means $P(x) \to (-1)^d P(-x)$).

The second generator acts by reading the pattern backwards (which for polynomials means $P(x) \to x^d P(1/x)$). If one wants to preserve the set of monic polynomials one has to divide $x^d P(1/x)$ by its leading term.

(Up to some trivialities) the properties we will study below are invariant under this action.

Simple result

We start with the following simple result.

Theorem

(i) *If d is even, then $Pol_{d,0}(\widehat{\sigma})$ is nonempty if and only if $\sigma_d = +$ (i.e., the constant term is positive).*

(ii) *For any pair of positive integers (d, k) with $d - k \geq 0$ and $d - k \equiv 0 \mod 2$ and any sign pattern $\widehat{\sigma} = (\sigma_1, ..., \sigma_d)$, the set $Pol_{d,k}(\widehat{\sigma})$ is nonempty.*

Observation

Observe that, in general, the set $Pol_{d,k}(\widehat{\sigma})$ does not have to be connected.

The total number k of real zeros can be distributed between m positive and n negative in different ways satisfying the inequalities $m+n=k$, $m \leq p_{\widehat{\sigma}},\ n \leq n_{\widehat{\sigma}}$ and $m \equiv p_{\widehat{\sigma}} \mod 2$, $n \equiv n_{\widehat{\sigma}} \mod 2$.

On the other hand, some specific sets $Pol_{d,k}(\widehat{\sigma})$ must be connected. In particular, the following holds.

Proposition

Proposition

(i) *For any* d *and* $\widehat{\sigma}$, *the sets* $Pol_{d,d}(\widehat{\sigma})$ *and* $Pol_{d,0}(\widehat{\sigma})$ *are contractible. (The latter set is empty for odd* d*.)*
(ii) *For the (shortened) sign pattern* $\widehat{+} = (+, +, \dots, +)$ *consisting of all pluses, the set* $Pol_{d,k}(\widehat{+})$ *is contractible, for any* $k \le d, \ k \equiv d \mod 2$*. (The same holds for the shortened alternating sign pattern* $(-, +, -, \dots)$*.)*
(iii) *For any sign pattern* $\overline{\sigma} = (1, \widehat{\sigma})$ *with just one sign change, all sets* $Pol_{d,k}(\widehat{\sigma})$ *are non-empty. For* $k = d$ *(which is the case of real-rooted polynomials having one positive and* $d - 1$ *negative roots), this set is contractible.*

Non-realizability

Proposition

For d even, consider patterns satisfying the following three conditions:
(a) the sign of the constant term (i.e., the last entry) is $+$*;*
(b) the signs of all odd monomials are $+$*;*
(c) among the remaining signs of even monomials there are $\ell \geq 1$ *minuses (at arbitrary positions).*
Then, for any such sign pattern, the pairs $(2, 0), (4, 0), \dots,$ $(2\ell, 0)$, *and only they, are non-realizable.*

(Using the standard $\mathbb{Z}_2 \times \mathbb{Z}_2$-action one obtains more such examples.)

More Stuff

Proposition

Consider a sign pattern $\overline{\sigma}$ *with* 2 *sign changes, consisting of* m *consecutive pluses (including the leading* 1*) followed by* n *consecutive minuses and then by* p *consecutive pluses, where* $m+n+p=d+1$. *Then*
(i) *for the pair* $(0, d-2)$, *this sign pattern is not realizable if*

$$\kappa := \frac{d-m-1}{m} \cdot \frac{d-p-1}{p} \geq 4; \tag{2}$$

(ii) *the sign pattern* $\overline{\sigma}$ *is realizable with any pair of the form* $(2, v)$.

Up to degree 6

Theorem

(i) *Up to degree $d \leq 3$, for any sign pattern $\overline{\sigma}$, all admissible pairs (pos, neg) are realizable.*
(ii) *For $d = 4$, (up to the standard $\mathbb{Z}_2 \times \mathbb{Z}_2$-action) the only non-realizable combination is $(1, -, -, -, +)$ with the pair $(0, 2)$;*
(iii) *For $d = 5$, (up to the standard $\mathbb{Z}_2 \times \mathbb{Z}_2$-action) the only non-realizable combination is $(1, -, -, -, -, +)$ with the pair $(0, 3)$;*
(iv) *For $d = 6$, (up to the standard $\mathbb{Z}_2 \times \mathbb{Z}_2$-action) the only non-realizable combinations are $(1, -, -, -, -, -, +)$ with $(0, 2)$ and $(0, 4)$; $(1, +, +, +, -, +, +)$ with $(2, 0)$; $(1, +, -, -, -, -, +)$ with $(0, 4)$.*

degree 7

Trying to extend Theorem 2, we obtained a computer-aided classification of all non-realizable sign patterns and pairs for $d = 7$ and almost all for $d = 8$, see below.

Theorem

For $d = 7$, among the 1472 possible combinations of a sign pattern and a pair (up to the standard $\mathbb{Z}_2 \times \mathbb{Z}_2$-action), there exist exactly 6 which are non-realizable. They are:

$$(1, +, -, -, -, -, -, +) \quad (0, 5); \quad (1, +, -, -, -, -, +, +) \quad (0, 5);$$

$$(1, +, -, +, -, -, -, -) \quad (3, 0); \quad (1, +, +, -, -, -, -, +) \quad (0, 5);$$

$$\text{and,} \quad (1, -, -, -, -, -, -, +) \quad (0, 3), \ (0, 5).$$

degree 8

Theorem

For $d = 8$, among the 3648 possible combinations of a sign pattern and a pair (up to the standard $\mathbb{Z}_2 \times \mathbb{Z}_2$-action), there exist 13 which are known to be non-realizable. They are:

$(1, +, -, -, -, -, -, +, +) \quad (0, 6)$; $(1, -, -, -, -, -, -, +, +) \quad (0, 6)$;

$(1, +, +, +, -, -, -, -, +) \quad (0, 6)$; $(1, +, +, -, -, -, -, -, +) \quad (0, 6)$;

$(1, +, +, +, -, +, +, +, +) \quad (2, 0)$; $(1, +, +, +, +, +, -, +, +) \quad (2, 0)$;

$(1, +, +, +, -, +, -, +, +) \quad (2, 0)\ ,\ (4, 0)\ ;\ 1, -, -, -, +, -, -, -, +)$ $(0, 2)\ ,\ (0, 4)$; $(1, -, -, -, -, -, -, -, +) \quad (0, 2), (0, 4), (0, 6)$.

Unknown

Remark

For $d = 8$, there exist exactly 7 (up to the standard $\mathbb{Z}_2 \times \mathbb{Z}_2$-action) combinations of a sign pattern and a pair for which it is still unknown whether they are realizable or not. They are:

$$(1,+,-,+,-,-,-,+,+) \quad (4,0); \quad (1,+,-,+,-,+,-,-,+) \quad (4,0);$$

$$(1,+,+,-,-,-,-,+,+) \quad (0,6); \quad (1,+,+,-,-,+,-,+,+) \quad (4,0);$$

$$(1,+,+,+,-,+,-,-,+) \quad (4,0); \quad (1,+,-,+,-,-,-,-,+) \quad (4,0)$$

and $(0,4)$.

Main Conjecture

Conjecture

For an arbitrary sign pattern $\overline{\sigma}$, the only type of pairs (pos, neg) which can be non-realizable has either pos or neg vanishing. In other words, for any sign pattern $\overline{\sigma}$, each pair (pos, neg) satisfying (1) *with positive pos and neg is realizable.*

Rephrasing the above conjecture, we say that the only phenomenon implying non-realizability is that "real roots on one half-axis force real roots on the other half-axis". At the moment this conjecture is verified by computer-aided methods up to $d = 10$.

Comercial break

The famous Descartes' rule of signs claims that the number of positive roots of a real univariate polynomial does not exceed the number of sign changes in its sequence of coefficients.

In what follows, among other things, we suggest a conceptually new conjectural upper bound on the number of real roots of real univariate polynomial applicable in the situation when Descartes' rule of signs gives a trivial restriction.

Stay tuned!

basic facts

The following notion is borrowed from the classical Wiman-Valiron theory. A non-negative integer k is said to be a *central index* of a real polynomial

$$P(x) = \sum_{i=0}^{d} a_i x^i,$$

if there exists a number $x_k \geq 0$ such that

$$|a_k| x_k^k \geq \max_{i \neq k} |a_i| x_k^i. \tag{3}$$

An analog

The next notion is analogous to the central index. A non-negative integer k is called a *dominating index* of a real polynomial

$$P(x) = \sum_{i=0}^{d} a_i x^i,$$

if there exists a real number $x_k \geq 0$ such that

$$|a_k| x_k^k \geq \sum_{i \neq k} |a_i| x_k^i. \tag{4}$$

Comments

Notice that (3) is an analog of (4) if the right-hand side of (4) is interpreted as a tropical sum.

We will say that a polynomial P of degree d is *tropically real-rooted* if each integer $k = 0, \ldots, d$ is a central index of f.

Remarks

To relate property (4) to real-rootedness of univariate polynomials, we say that a real-rooted polynomial P is called *strongly real-rooted* if each polynomial obtained by an arbitrary sign change of the coefficients of $P(x)$ is real-rooted as well.

Proposition

A real polynomial P of degree d is strongly real-rooted if and only if every integer $k = 0, \dots, d$ is a dominating index of P.

By the (standard) *tropicalization* of a real polynomial $P(x) = \sum_{i=0}^{d} a_i x^i$ we mean the tropical polynomial given by:

$$tr_P(\xi) = \max_{0 \le i \le d} (i\xi + \ln|a_i|), \ \xi \in \mathbb{R}. \tag{5}$$

If $a_i = 0$, then the corresponding term in $tr_P(\xi)$ should be interpreted as $-\infty$, and thus it can be ignored when taking the maximum.

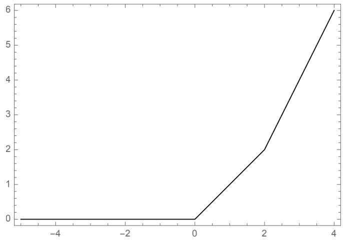

Figure: Standard tropicalization of $1 + x + e^{-2}x^2$.

Any corner of the graph of $tr_P(\xi)$, i.e., a value of ξ at which its slope changes, is called a *tropical root* of $tr_P(\xi)$.

We define *Descartes' multiplicity* of a tropical root ζ of tr_P to be one less than the number of terms of (5) for which the maximum in the right-hand side of (5) is attained at ζ.

With our definition of Descartes' multiplicity of a tropical root, the number of tropical roots of $tr_P(\xi)$ counted with multiplicities is one less than the number of central indices of P.

In particular, the number of tropical roots of $tr_P(\xi)$ is at most by one less than the number of monomials of P.

The latter circumstance is analogous to the fact that the number of real roots of P is at most one less than its number of monomials.

Let $k_0 \leq k_1 \leq \cdots \leq k_m$ be the central indices of P. Consider two sequences $\{\mathrm{sgn}(a_{k_i})\}_{0 \leq i \leq m}$ and $\{\mathrm{sgn}((-1)^{k_i} a_{k_i})\}_{0 \leq i \leq m}$. Take two consecutive central indices k_{i-1} and k_i of the polynomial P; to this pair we associate the tropical root $\xi_i = -\ln(a_{i-1}/a_i)/(k_{i-1} - k_i)$ of $tr_P(\xi)$.

If the difference $k_{i-1} - k_i$ is odd, then the pair (k_{i-1}, k_i) contributes a sign alternation in exactly one of the above sequences. In this case, we will say that ξ_i is a *positive* (respectively *negative*) *essential tropical root* of P.
If the difference $k_{i-1} - k_i$ is even, then either the pair (k_{i-1}, k_i) does not contribute a sign alternation in any of the above sequences, or it contributes a sign alternation in both.
In the former case we will say that ξ_i is a *non-essential tropical root* of P, and in the latter case we will say that ξ_i is a *positive-negative essential tropical root* of P.

By the number of *positive essential tropical roots of* P we mean the sum of the numbers of positive and positive-negative tropical roots of P.

Analogously, by the number of *negative essential tropical roots of* P we mean the sum of the numbers of negative and positive-negative tropical roots of P.

Finally by the total number of *essential tropical roots of* P we call the sum of the latter two numbers.

It is easy to see that the number of essential tropical roots of P is at most d.

Example

Take $P_1(x) = 1 + x^2$. The central indices of P_1 are $k_0 = 0$ and $k_1 = 2$. As $\ln|a_1| = \ln|0| = -\infty$, the polynomial P_1 has (with our definition of Descartes' multiplicity) exactly one simple tropical root. To count the number of positive and negative tropical roots of P_1 we need to count the number of sign alternations in the sequences $\{1, 1\}$ and $\{1, (-1)^2\} = \{1, 1\}$ respectively. That is, the number of essential tropical roots of P is equal to 0.

On the other hand, take $P_2(x) = 1 - x^2$. Similarly to P_1, the polynomial P_2 has one tropical root. However, to count the number of positive and negative tropical roots of P_2 we count the number of sign alternations in the sequences $\{1, -1\}$ and $\{1, -(-1)^2\} = \{1, -1\}$ respectively. That is, the number of essential tropical roots of P_2 is equal to 2.

In general, the number of real/positive/negative roots of a real polynomial $P(x)$ is not smaller or equal than the number of essential tropical/positive/negative roots of $tr_P(\xi)$.

Our goal is to multiply the coefficients of polynomials by some fixed positive factors so that the number of real/positive/negative roots will be smaller than that of the tropicalization after such multiplication.

Given an arbitrary triangular sequence

$$\lambda = \{\lambda_{k,d}\},\ 0 \le k \le d,\ d \in \mathbb{N}$$

of positive numbers, and a real univariate polynomial $P(x) = \sum_{k=0}^{d} a_k x^k$ of any degree d, we define its λ-*tropicalization* as

$$tr_P^{\lambda}(\xi) = \max_{0 \le k \le d} (k\xi + \ln|a_k| + \ln \lambda_{k,d}),\ \xi \in \mathbb{R}. \quad (6)$$

This is the standard tropicalization of $Q(x) = \sum_{i=0}^{d} \lambda_{k,d} a_k x^k$.

Definition

A finite sequence $\{\lambda_{k,d}\}_{0\le k\le d}$, of positive numbers is called a *degree d real-to-tropical root preserver* if for any polynomial P of degree d, the number of essential tropical roots of (6) is greater than or equal to the number of non-zero real roots of P.

A triangular sequence $\lambda = \{\lambda_{k,j}\}_{0\le k\le j,\ j\in\mathbb{N}}$ is called a *real-to-tropical root preserver* if, for each d, its finite subsequence $\{\lambda_{k,d}\}_{0\le k\le d}$, is a degree d real-to-tropical root preserver.

We recall that the *recession cone* of a set $X \subset \mathbb{R}^{d+1}$ is the largest pointed (i.e. including the origin) cone $C \subseteq \mathbb{R}^{d+1}$ such that if $x \in X$, then $x + c \in X$ for all $c \in C$. Our main result is as follows.

Theorem

The set $\Lambda_d \subset \mathbb{R}_+^{d+1}$ of all degree d real-to-tropical root preservers $\{\lambda_{k,d}\}_{0 \le k \le d}$ is a nonempty closed full-dimensional subset of $\mathbb{R}_+^{d+1}$.

Moreover, the recession cone of its logarithmic image $\mathrm{Ln}(\Lambda_d)$ coincides with the cone of all concave sequences of length $d + 1$.

(*Here for any $\Omega \subset \mathbb{R}_+^k$, by $\mathrm{Ln}(\Omega)$ we mean the set in $\mathbb{R}^k$ obtained by taking natural logarithms of points from Ω coordinatewisely.*)

Theorem 5 shows that there exist large families of real-to-tropical root preservers in each degree, and therefore large families of real-to-tropical root preserving triangular sequences.

Theorem

Assume that a sequence $\lambda = \{\lambda_{k,d}\}_{0\leq k\leq d}$ *of positive numbers satisfies the condition:*

$$\log \frac{\lambda_{k,d}^2}{\lambda_{k-1,d}\lambda_{k+1,d}} > 2\Delta_d := \frac{d^2}{4}\log 36d + (d+1)\log d + \log 4,\ 1\leq k\leq d \quad (7)$$

Then, for any real polynomial P*, the number of positive (negative) tropical roots of* tr_P^λ *is greater than or equal to the number of positive (negative) roots of* P*. In particular,* λ *is a real-to-tropical root preserver.*

On the other hand,

Theorem

There exists $c > 0$ with the following property. Assume that for some $k < d - 100$

$$\log \frac{\lambda_{j,d}^2}{\lambda_{j-1,d}\lambda_{j+1,d}} < 2c, \quad j = k, ..., k + 100. \tag{8}$$

Then there exists a polynomial P of degree d with positive coefficients such that tr_P^λ has three tropical roots, and P has four negative roots. In particular, $\{\lambda_{k,d}\}_{0 \le k \le d}$ cannot be a degree d real-to-tropical root preserver.

Main Conjecture

We finish with the following tantalizing conjecture. Consider the sequence $\lambda^\dagger$ given by

$$\lambda_k^\dagger := e^{-k^2},\ k = 0, 1, \dots .$$

We will denote by $tr_P^\dagger(\xi)$ the corresponding tropical polynomial associated to any real polynomial P, i.e.,

$$tr_P^\dagger(\xi) = \max_{0 \le k \le d} (k\xi + \ln|a_k| - k^2),\ \xi \in \mathbb{R}. \tag{9}$$

Conjecture (Conjectural tropical analog of Descartes' rule of signs)

For any real univariate polynomial $P(x)$, the number of its positive (negative) roots does not exceed the number of positive (negative) essential tropical roots of $tr_P^\dagger(\xi)$.

Thanks for your patience!

Congratulations, Vitya!

You saved Ch. 1 of my PhD thesis by your explanations

...