



On the Roots of a Hyperbolic Polynomial Pencil

Victor Katsnelson¹

Received: 3 May 2016 / Accepted: 20 July 2016 / Published online: 2 August 2016 © Institute for Mathematical Sciences (IMS), Stony Brook University, NY 2016

Abstract Let $v_0(t), v_1(t), \dots, v_n(t)$ be the roots of the equation R(z) = t, where R(z) is a rational function of the form

$$R(z) = z - \sum_{k=1}^{n} \frac{\alpha_k}{z - \mu_k},$$

 μ_k are pairwise distinct real numbers, $\alpha_k > 0$, $1 \le k \le n$. Then for each real ξ , the function $e^{\xi \nu_0(t)} + e^{\xi \nu_1(t)} + \cdots + e^{\xi \nu_n(t)}$ is exponentially convex on the interval $-\infty < t < \infty$.

Keywords Hyperbolic polynomial pencil · Determinant representation · Exponentially convex functions

Mathematics Subject Classification 11C99 · 26C10 · 26C15 · 15A22 · 42A82

1 Roots of the Equation R(z) = t as Functions of t

In the present paper we discuss questions related to properties of roots of the equation

$$R(z) = t \tag{1.1}$$

☑ Victor Katsnelson victor.katsnelson@weizmann.ac.il; victorkatsnelson@gmail.com

¹ Department of Mathematics, Weizmann Institute, 7610001 Rehovot, Israel

as functions of the parameter $t \in \mathbb{C}$, where R is a rational function of the form

$$R(z) = z - \sum_{1 \le k \le n} \frac{\alpha_k}{z - \mu_k},\tag{1.2}$$

 μ_k are pairwise distinct real numbers, $\alpha_k > 0$, $1 \le k \le n$. We adhere to the enumeration agreement¹

$$\mu_1 > \mu_2 > \dots > \mu_n. \tag{1.3}$$

The function R is representable in the form

daf

$$R(z) = \frac{P(z)}{Q(z)},\tag{1.4}$$

where

$$Q(z) = (z - \mu_1) \cdot (z - \mu_2) \cdot \dots \cdot (z - \mu_n),$$
 (1.5)

$$P(z) \stackrel{\text{def}}{=} R(z) \cdot Q(z) \tag{1.6}$$

are monic polynomials of degrees

1

$$\deg P = n + 1, \quad \deg Q = n.$$
 (1.7)

Since $P(\mu_k) = -\alpha_k Q'(\mu_k) \neq 0$, the polynomials *P* and *Q* have no common roots. Thus the ratio in the right hand side of (1.4) is irreducible. The Eq. (1.1) is equivalent to the equation

$$P(z) - tQ(z) = 0. (1.8)$$

Since the polynomial P(z) - tQ(z) is of degree n + 1, the latter equation has n + 1 roots for each $t \in \mathbb{C}$.

The function *R* possess the property

$$\operatorname{Im} R(z) / \operatorname{Im} z > 0 \quad \text{if} \quad \operatorname{Im} z \neq 0. \tag{1.9}$$

Therefore if Im t > 0, all roots of the equation (1.1), which is equivalent to the Eq. (1.8), are located in the half-plane Im z > 0. Some of these roots may be multiple.

However if *t* is real, all roots of the Eq. (1.1) are real and simple, i.e. of multiplicity one. Thus for real *t*, the Eq. (1.1) has n + 1 pairwise distinct real roots $v_k(t)$: $v_0(t) > v_1(t) > \cdots > v_{n-1}(t) > v_n(t)$. Moreover for each real *t*, the poles μ_k of the function *R* and the roots $v_k(t)$ of the Eq. (1.1) are interlacing:

$$\nu_0(t) > \mu_1 > \nu_1(t) > \mu_2 > \nu_2(t) > \dots > \nu_{n-1}(t) > \mu_n > \nu_n(t), \quad \forall t \in \mathbb{R}.$$
(1.10)

¹ We assume that $n \ge 1$.

In particular for t = 0, the roots $v_k(0) = \lambda_k$ of the Eq. (1.1) are the roots of the polynomial *P*:

$$P(z) = (z - \lambda_0) \cdot (z - \lambda_1) \cdot \cdots \cdot (z - \lambda_n), \qquad (1.11)$$

$$\lambda_0 > \mu_1 > \lambda_1 > \mu_2 > \lambda_2 > \dots > \lambda_{n-1} > \mu_n > \lambda_n.$$
(1.12)

Since R'(x) > 0 for $x \in \mathbb{R}$, $x \neq \mu_1, \ldots, \mu_n$, each of the functions $v_k(t), k = 0, 1, \ldots, n$, can be continued as a single valued holomorphic function to some neighborhood of \mathbb{R} . However the functions $v_k(t)$ can not be continued as single-valued analytic functions to the whole complex *t*-plane. According to (1.4),

$$R'(z) = \frac{P'(z)Q(z) - Q'(z)P(z)}{Q^2(z)}.$$
(1.13)

The polynomial P'Q - Q'P is of degree 2n and is strictly positive on the real axis. Therefore this polynomial has *n* roots ζ_1, \ldots, ζ_n in the upper half-plane Im(z) > 0and *n* roots $\overline{\zeta_1}, \ldots, \overline{\zeta_n}$ in the lower half-plane Im(z) < 0. (Not all roots ζ_1, \ldots, ζ_n must be distinct.) The points ζ_1, \ldots, ζ_n and $\overline{\zeta_1}, \ldots, \overline{\zeta_n}$ are the critical points of the function *R*: $R'(\zeta_k) = 0$, $R'(\overline{\zeta_k}) = 0$, $1 \le k \le n$. The critical values $t_k = R(\zeta_k)$, $\overline{t_k} =$ $R(\overline{\zeta_k})$, $1 \le k \le n$, of the function *R* are the ramification points of the function v(t):

$$R(v(t)) = t \tag{1.14}$$

(Even if the critical points ζ' and ζ'' of *R* are distinct, the critical values $R(\zeta')$ and $R(\zeta'')$ may coincide.) We denote the set of critical values of the function *R* by \mathcal{V} :

$$\mathcal{V} = \mathcal{V}^+ \cup \mathcal{V}^-, \quad \mathcal{V}^+ = \{t_1, \dots, t_n\}, \quad \mathcal{V}^- = \{\overline{t_1}, \dots, \overline{t_n}\}.$$
(1.15)

Not all values t_1, \ldots, t_n must be distinct. However $\mathcal{V} \neq \emptyset$. In view of (1.9), Im $t_k > 0, 1 \le k \le n$. So

$$\mathcal{V}^+ \subset \{t \in \mathbb{C} : \operatorname{Im} t > 0\}, \quad \mathcal{V}^- \subset \{t \in \mathbb{C} : \operatorname{Im} t < 0\}.$$
(1.16)

Let *G* be an arbitrary simply connected domain in the *t*-plane which does not intersect the set \mathcal{V} . Then the roots of Eq. (1.1) are pairwise distinct for each $t \in G$. We can enumerate these roots, say $v_0(t)$, $v_1(t)$, ... $v_n(t)$, such that all functions $v_k(t)$ are holomorphic in *G*.

The strip S_h ,

$$S_h = \{t \in \mathbb{C} : |\operatorname{Im} t| < h\}, \text{ where } h = \min_{1 \le k \le n} \operatorname{Im} t_k,$$
(1.17)

does not intersect the set \mathcal{V} . So n + 1 single valued holomorphic branches of the function v(t), (1.14), are defined in the strip S_h . We choose such enumeration of these branches which agrees with the enumeration (1.10) on \mathbb{R} .

From (1.6) and (1.2) it follows that the polynomial P is representable in the form

$$P(z) = z Q(z) - \sum_{k=1}^{n} \alpha_k Q_k(z), \qquad (1.18a)$$

where

$$Q_k(z) = Q(z)/(z - \mu_k), \quad k = 1, 2, \dots, n.$$
 (1.18b)

2 Determinant Representation of the Polynomial Pencil P(z) - t Q(z)

The polynomial pencil P(z) - tQ(z) is *hyperbolic*: for each real *t*, all roots of the Eq. (1.8) are real.

Using (1.18), we represent the polynomial P(z) - tQ(z) as the characteristic polynomial det(zI - (A + tB)) of some matrix pencil, where A and B are self-adjoint $(n + 1) \times (n + 1)$ matrices, rank B = 1. We present these matrices explicitly.

Lemma 2.1 Let $A = ||a_{p,q}||$ and $B = ||b_{p,q}||$, $0 \le p, q \le n$, be $(n + 1) \times (n + 1)$ matrices with the entries

$$a_{0,0} = 0, \ a_{p,p} = \mu_p \quad for \ p = 1, 2, \dots, n,$$

$$a_{p,q} = 0 \quad for \ p = 1, 2, \dots, n, \ q = 1, 2, \dots, n, \ p \neq q,$$

$$a_{0,p} = \overline{a_{p,0}} \quad for \ p = 1, 2, \dots, n,$$
(2.1)

and

$$b_{0,0} = 1$$
, all other $b_{p,q}$ vanish. (2.2)

Then the equality

$$\det(zI - A - tB) = (z - t) \cdot Q(z) - \sum_{k=1}^{n} |a_{0,k}|^2 Q_k(z).$$
(2.3)

holds.

Proof The matrix zI - (A + tB) is of the form

$$zI - (A + tB) = \begin{bmatrix} z - t & -a_{0,1} & -a_{0,2} & \cdots & -a_{0,n-1} & -a_{0,n} \\ -\overline{a_{0,1}} & z - \mu_1 & 0 & \cdots & 0 & 0 \\ -\overline{a_{0,2}} & 0 & z - \mu_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -\overline{a_{0,n-1}} & 0 & 0 & \cdots & z - \mu_{n-1} & 0 \\ -\overline{a_{0,n}} & 0 & 0 & \cdots & 0 & z - \mu_n \end{bmatrix}$$

We compute the determinant of this matrix using the cofactor formula.

Comparing (1.18) and (2.3), we see that if the conditions

$$|a_{0,p}|^2 = \alpha_p, \quad p = 1, 2, \dots, n$$
 (2.4)

are satisfied, then the equality

$$P(z) - tQ(z) = \det(zI - A - tB)$$
(2.5)

holds for every $z \in \mathbb{C}$, $t \in \mathbb{C}$.

The following result is an immediate consequence of Lemma 2.1.

Theorem 2.2 Let *R* be a function of the form (1.2), where $\mu_1, \mu_2, \ldots, \mu_n$ are pairwise distinct real numbers and $\alpha_1, \alpha_2, \ldots, \alpha_n$ are positive numbers. Let *Q* and *P* be the polynomials related to the the function *R* by the equalities (1.5) and (1.18).

Then the pencil of polynomials P(z) - tQ(z) is representable as the characteristic polynomial of the matrix pencil A + tB, i.e. the equality (2.5) holds for every $z \in \mathbb{C}$, $t \in \mathbb{C}$, where B is the matrix with the entries (2.2), and the entries of the matrix A are defined by by (2.1) with

$$a_{0,p} = \sqrt{\alpha_p} \,\omega_p, \quad p = 1, 2, \dots, n,$$
 (2.6)

 ω_p are arbitrary² complex numbers of absolute value one:

$$|\omega_p| = 1, \quad p = 1, 2, \dots, n.$$
 (2.7)

Corollary 2.3 Let R, A, B be the same as in Theorem 2.2. For each $t \in \mathbb{C}$, the roots $v_0(t), v_0(t), \ldots, v_n(t)$ of the Eq. (1.2) are the eigenvalues of the matrix A + tB.

Lemma 2.4 Let R, A, B be the same as in Theorem 2.2, $v_0(t), v_0(t), \ldots, v_n(t)$ be the roots of the Eq. (1.2) and h(z) be an entire function. Then the equality

$$\sum_{k=0}^{n} h(\nu_k(t)) = \text{trace} \{h(A+tB)\}$$
(2.8)

holds for every $t \in \mathbb{C}$.

Proof We refer to Corollary 2.3. If v is an eigenvalue of some square matrix M, then h(v) is an eigenvalue of the matrix h(M). In (2.8), we interpret the trace of the matrix h(A + tB) as its *spectral trace*, that is as the sum of all its eigenvalues.

² We will use the freedom in choosing ω_p to prescribe signs \pm to the entries $a_{0,p}$.

3 Exponentially Convex Functions

Definition 3.1 A function f(t) on the interval a < t < b is said to belong to the class $W_{a,b}$ if f is continuous on (a, b) and if all forms

$$\sum_{r,s=1}^{N} f(t_r + t_s) \zeta_r \overline{\zeta_s} \quad (N = 1, 2, 3, \dots)$$
(3.1)

are non-negative for every choice of complex numbers $\zeta_1, \zeta_2, \ldots, \zeta_N$ and for every choice of real numbers t_1, t_2, \ldots, t_N assuming that all sums $t_r + t_s$ are within the interval (a, b).

The class $W_{a,b}$ was introduced by Bernstein (1928), see Sect. 15 there. Somewhat later, Widder also introduced the class $W_{a,b}$ and studied it. Bernstein called functions $f(x) \in W_{a,b}$ exponentially convex.

Properties of the class of exponentially convex functions

P1. If $f(t) \in W_{a,b}$ and $c \ge 0$ is a nonnegative constant, then $cf(t) \in W_{a,b}$. P2. If $f_1(t) \in W_{a,b}$ and $f_2(t) \in W_{a,b}$, then $f_1(t) + f_2(t) \in W_{a,b}$. P3. If $f_1(t) \in W_{a,b}$ and $f_2(t) \in W_{a,b}$, then $f_1(t) \cdot f_2(t) \in W_{a,b}$. P4. Let $\{f_n(t)\}_{1 \le n < \infty}$ be a sequence of functions from the class $W_{a,b}$. We assume that for each $t \in (a, b)$ there exists a limit $f(t) = \lim_{n \to \infty} f_n(t)$, and that $f(t) < \infty \forall t \in (a, b)$. Then $f(t) \in W_{a,b}$.

From the functional equation for the exponential function it follows that for each real number u, for every choice of real numbers $t_1, t_2, ..., t_N$ and complex numbers $\zeta_1, \zeta_2, ..., \zeta_N$, the equality holds

$$\sum_{r,s=1}^{N} e^{\xi(t_r+t_s)} \zeta_r \overline{\zeta_s} = \left| \sum_{p=1}^{N} e^{\xi t_p} \zeta_p \right|^2 \ge 0.$$
(3.2)

The inequality (3.2) can be stated as

Lemma 3.2 For each real number ξ , the function $e^{\xi t}$ of the variable t belongs to the class $W_{-\infty,\infty}$.

The term *exponentially convex function* is justified by the following integral representation for any function $f(t) \in W_{a,b}$.

Theorem 3.3 (The representation theorem) For the representation

$$f(x) = \int_{\xi \in (-\infty,\infty)} e^{\xi x} \sigma(d\xi) \quad (a < x < b)$$
(3.3)

to be valid, where $\sigma(d\xi)$ is a non-negative measure, it is necessary and sufficient that $f(x) \in W_{a,b}$.

The proof of the Representation Theorem can be found in Akhiezer (1965) (Theorem 5.5.4), and in Widder (1946) (Chapter 6, Theorem 21).

Corollary 3.4 *The representation* (3.3) *shows that* f(x) *is the value of a function* f(z) *holomorphic in the strip* a < Re z < b.

4 Herbert Stahl's Theorem

In the paper Bessis et al. (1975) a conjecture was formulated which is now commonly known as the BMV conjecture:

The BMV conjecture Let U and V be Hermitian matrices. Then the function

$$\varphi(t) = \operatorname{trace} \left\{ e^{U+tV} \right\} \tag{4.1}$$

of the variable t belongs to the class $W_{-\infty,\infty}$.

If the matrices U and V commute, the exponential convexity of the function φ , (4.1), is evident. In this case, the sum

$$\sum_{r,s=1}^{N} \varphi(t_r + t_s) \zeta_r \overline{\zeta_s} = \text{trace} \left\{ e^{U/2} \left(\sum_{r=1}^{N} e^{t_r V} \zeta_r \right) \left(\sum_{s=1}^{N} e^{t_s V} \zeta_s \right)^* (e^{U/2})^* \right\}$$

is non-negative because this sum is the trace of a non-negative matrix. The measure σ in the integral representation (3.3) of the function φ , (4.1), is an atomic measure supported on the spectrum of the matrix V.

In the general case, if the matrices U and V do not commute, the BMV conjecture remained an open question for longer than 40 years. In 2011, Herbert Stahl proved the BMV conjecture.

Theorem 4.1 (H. Stahl) Let U and V be Hermitian matrices.

Then the function $\varphi(t)$ defined by (4.1) belongs to the class $W_{-\infty,\infty}$ of functions exponentially convex on $(-\infty,\infty)$.

The first arXiv version of Stahl's Theorem appeared in Stahl (2011), the latest arXiv version—in Stahl (2012), the journal publication—in Stahl (2013).

The proof of Herbert Stahl is based on ingenious considerations related to Riemann surfaces of algebraic functions. In Eremenko (2015), a simplified version of the Herbert Stahl proof is presented.

We present a toy version of Theorem 4.1 which is enough for our goal.

Theorem 4.2 Let U and V be Hermitian matrices. We assume moreover that

- 1. All off-diagonal entries of the matrix U are non-negative.
- 2. The matrix V is diagonal.

Then the function $\varphi(t)$ defined by (4.1) belongs to the class $W_{-\infty,\infty}$.

Proof For $\rho \ge 0$, let $U_{\rho} = U + \rho I$, where I is the identity matrix. If ρ is large enough, then all entries of the matrix U_{ρ} are non-negative. Let us choose and fix such ρ . It is clear that

$$e^{U+tV} = e^{-\rho} e^{U_{\rho}+tV}.$$
(4.2)

We use the Lie product formula

$$e^{U_{\rho}+tV} = \lim_{m \to \infty} (e^{U_{\rho}/m} e^{tV/m})^m.$$
(4.3)

All entries of the matrix $e^{U_{\rho}/m}$ are non-negative numbers. Since matrix V is Hermitian, its diagonal entries are real numbers. Thus

$$e^{tV/m} = \operatorname{diag}(e^{tv_1/m}, e^{tv_2/m}, \dots, e^{tv_m/m}),$$

where v_1, v_2, \ldots, v_m are real numbers. The exponentials $e^{tv_j/m}$ are functions of t from the class $W_{-\infty,\infty}$. Each entry of the matrix $e^{U_\rho/m} e^{tV/m}$ is a linear combination of these exponentials with non-negative coefficients. According to the properties P1 and P2 of the class $W_{-\infty,\infty}$, the entries of the matrix $e^{U_\rho/m} e^{tV/m}$ are functions of the class $W_{-\infty,\infty}$. Each entry of the matrix $(e^{U_\rho/m} e^{tV/m})^m$ is a sum of products of some entries of the matrix $e^{U_\rho/m} e^{tV/m}$. According to the properties P2 and P3 of the class $W_{-\infty,\infty}$, the entries of the matrix $(e^{U_\rho/m} e^{tV/m})^m$ are functions of t belonging to the class $W_{-\infty,\infty}$. From the limiting relation (4.3) and from the property P4 of the class $W_{-\infty,\infty}$. From (4.2) it follows that all entries of the matrix e^{U_+tV} belong to the class $W_{-\infty,\infty}$. All the more, the function $\varphi(t) = \text{trace} \{e^{U+tV}\}$, which is the sum of diagonal entries of the matrix e^{U+tV} , belongs to the class $W_{-\infty,\infty}$.

5 Exponential Convexity of the Sum $e^{\xi v_0(t)} + \cdots + e^{\xi v_n(t)}$

Let ξ be a real number. Taking $h(z) = e^{\xi z}$ in Lemma 2.4, we obtain

Lemma 5.1 Let R be the rational function of the form (1.2), $v_0(t)$, $v_1(t)$, ..., $v_n(t)$ be the roots of the Eq. (1.1). Let A and B be the matrices (2.1), (2.6), (2.2) which appear in the determinant representation (2.5) of the matrix pencil P(z) - tQ(z).

Then the equality

$$\sum_{k=0}^{n} e^{\xi \, \nu_k(t)} = \operatorname{trace}\{e^{\xi A + t(\xi B)}\}$$
(5.1)

holds.

Now we choose ω_p in (2.6) so that all off-diagonal entries of the matrix $U = \xi A$ are non-negative: if $\xi > 0$, then $\omega_p = +1$, if $\xi < 0$, then $\omega_p = -1$, $1 \le p \le n$.

Applying Theorem 4.2 to the matrices $U = \xi A$, $V = \xi B$, we obtain the following result

Theorem 5.2 Let *R* be the rational function of the form (1.2), $v_0(t)$, $v_1(t)$, ..., $v_n(t)$ be the roots of the Eq. (1.1). Then for each $\xi \in \mathbb{R}$, the function

$$g(t,\xi) \stackrel{def}{=} \sum_{k=0}^{n} e^{\xi \, \nu_k(t)}$$
(5.2)

of the variable t belongs to the class $W_{-\infty,\infty}$.

Theorem 5.3 Let $f \in W_{u,v}$, where $-\infty \le u < v \le +\infty$. Let *R* be the rational function of the form (1.2), $v_0(t)$, $v_1(t)$, ..., $v_n(t)$ be the roots of the Eq. (1.1). Assume that for some $a, b, -\infty \le a < b \le +\infty$, the inequalities

$$u < v_k(t) < v, \quad a < t < b, \quad k = 0, 1, \dots, n$$
 (5.3)

hold.

Then the function

$$F(t) \stackrel{def}{=} \sum_{k=0}^{n} f(v_k(t))$$
(5.4)

belongs to the class $W_{a,b}$.

Proof According to Theorem 3.3, the representation

$$f(x) = \int_{\xi \in (-\infty,\infty)} e^{\xi x} \sigma(d\xi), \quad \forall x \in (u,v)$$

holds, where σ is a non-negative measure. Substituting $x = v_k(t)$ to the above formula, we obtain the equality

$$f(\nu_k(t)) = \int_{\xi \in (-\infty,\infty)} e^{\xi \nu_k(t)} \sigma(d\xi), \quad \forall t \in (a,b), \quad k = 0, 1, \dots, n.$$

Hence

$$F(t) = \int_{\xi \in (-\infty,\infty)} g(t,\xi) \,\sigma(d\xi), \quad \forall t \in (a,b).$$
(5.5)

Theorem 5.4 is a consequence of Theorem 5.2 and of the properties P1, P2, P4 of the class of exponentially convex functions. \Box

Example For $\gamma > 0$, the function $f(x) = e^{\gamma x^2}$ is exponentially convex on $(-\infty, \infty)$: $e^{\gamma x^2} = \int_{\xi \in (-\infty,\infty)} e^{\xi x} \sigma(d\xi)$, where $\sigma(d\xi) = \frac{1}{2\sqrt{\pi\gamma}} e^{-\xi^2/4\gamma} d\xi$.

Thus the function $F(t) = \sum_{k=0}^{n} e^{\gamma (v_k(t))^2}$ is exponentially convex on $(-\infty, \infty)$.

Remark 5.4 Familiarizing himself with our proof of Theorem 5.2, Alexey Kuznetsov (http://www.math.yorku.ca/~akuznets/) gave a new proof of a somewhat weakened version of this theorem. His proof is based on the theory of stochastic Lévy processes.

References

Ахиезер, Н.И. *Классическая: проблема моментов.* Физматгиз, Москва (1965) (in Russian). English Transl.: Akhiezer, N.I.: The Clasical Moment Problem. Oliver and Boyd, Edinburgh (1965)

Bernstein, S.N.: Sur les functions absolument monotones. Acta Math. 52, 1-66 (1928). (in French)

Bessis, D., Moussa, P., Villani, M.: Monotonic converging variational approximations to the functional integrals in quantum statistical mechanics. J. Math. Phys. **16**(11), 2318–2325 (1975)

Eremenko, A.: Herbert Stahl's proof of the BMV conjecture. Sb. Math. 206(1), 87-92 (2015)

Stahl, H.: Proof of the BMV conjecture. arXiv:1107.4875v1, pp. 1–56, 25 July 2011 (2011)

Stahl, H.: Proof of the BMV conjecture. arXiv:1107.4875v3, pp. 1-25, 17 August 2012

Stahl, H.: Proof of the BMV conjecture. Acta Math. 211, 255-290 (2013)

Widder, D.V.: Laplace Transform. Princeton University Press, Princeton (1946)