On the zeros of eigenpolynomials of Hermitian Toeplitz matrices

Florian Bünger

accepted for publication April 27th, 2014

Abstract. This article refines a result of Delsarte, Genin, Kamp [5], [6], regarding the number of zeros on the unit circle of eigenpolynomials of complex Hermitian Toeplitz matrices and generalized Caratheodory representations of such matrices. This is achieved by exploring a key observation of Schur [20] stated in his proof of a famous theorem of Carathéodory [2]. In short, Schur observed that companion matrices corresponding to eigenpolynomials of Hermitian Toeplitz matrices H define isometries with respect to (spectrum shifted) submatrices of H. Looking at possible normal forms of these isometries leads directly to the results. This geometric, conceptual approach can be generalized to Hermitian or symmetric Toeplitz matrices over arbitrary fields. Furthermore, as a byproduct, Iohvidov's law in the jumps of the ranks and the connection between the Iohvidov parameter and the Witt index are established for such Toeplitz matrices.

Mathematics Subject Classification (2010). Primary 15B05; Secondary 15B57.

Keywords. Hermitian Toeplitz matrices, zeros of eigenpolynomials, Iohvidov parameter.

1. Introduction

A famous theorem of Carathéodory [2] states that for arbitrary complex numbers $a_1, ..., a_n$, $n \in \mathbb{N}$ [1], not all zero, there exist uniquely determined data $m \in \mathbb{N}$ with $m \leq n$, pairwise distinct $\varepsilon_1, ..., \varepsilon_m \in \mathbb{C}$ of modulus one and positive real numbers $r_1, ..., r_m$ such that

$$a_i = \sum_{j=1}^{m} r_j \varepsilon_j^i \tag{1.1}$$

[1] $\mathbb{N} := \{1, 2, 3, ...\}$, $\mathbb{N}_0 := \{0, 1, 2, 3, ...\}$

for all $i \in \{1,...,n\}$. Carathéodory proved his theorem by means of geometric considerations on convex bodies. Soon after Carathéodory, Fischer [7], Schur [20] and also Frobenius [8] gave algebraic proofs of this theorem. From the very beginning the connection between Carathéodory's theorem and Hermitian Toeplitz matrices was clear: Define

$$\mu := \mu(a_1,...,a_n) := r_1 + ... + r_m > 0,$$

$R := \operatorname{diag}(r_1,...,r_m) \in \mathbb{C}^{m,m}$ and the Vandermonde matrix

$$V := \begin{bmatrix} 1 & 1 & \dots & 1 \\ \varepsilon_1 & \varepsilon_2 & \dots & \varepsilon_m \\ \varepsilon_1^2 & \varepsilon_2^2 & \dots & \varepsilon_m^2 \\ \vdots & \vdots & \vdots & \vdots \\ \varepsilon_1^n & \varepsilon_2^n & \dots & \varepsilon_m^n \end{bmatrix} \in \mathbb{C}^{n+1,m}. \tag{1.2}$$

Then, the Hermitian Toeplitz matrix

$$H := H(\mu, a_1,...,a_n) := \begin{bmatrix} \mu & a_1 & a_2 & ... & a_n \\ \overline{a_1} & \mu & a_1 & a_2 & ... \\ \overline{a_2} & \overline{a_1} & \mu & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \overline{a_n} & ... & \overline{a_2} & \overline{a_1} & \mu \end{bmatrix} \in \mathbb{C}^{n+1,n+1}.$$

is positive semidefinite and admits the representation $H = \overline{V} R V^T$. Therefore, the vector $p = (p_0,...,p_m,0,...,0)^T \in \mathbb{C}^{n+1}$ consisting of the coefficients of the polynomial

$$p(x) := \prod_{i=1}^{m}(x - \varepsilon_i) = \sum_{i=0}^{m} p_i x^i \tag{1.3}$$

is contained in the kernel of H since $V^T p = (p(\varepsilon_1),...,p(\varepsilon_m))^T = 0$, and $p(x)$ is the uniquely determined nonzero monic (eigen-)polynomial of smallest degree with this property. Conversely, given an arbitrary Hermitian Toeplitz matrix

$$H = H(a_0,...,a_n) \in \mathbb{C}^{n+1,n+1}, \quad a_0,...,a_n \in \mathbb{C},$$

which is not a diagonal matrix, i.e., $a_1,...,a_n$ not all zero, then, necessarily,

$$\lambda := a_0 - \mu(a_1,...,a_n)$$

is the smallest eigenvalue of H, and, if $m \leq n$, ε_i and r_i are chosen according to Carathéodory's theorem for $a_1,...,a_n$, then H admits the so-called Carathéodory representation

$$H = \overline{V} R V^T + \lambda I_{n+1} \ , \tag{1.4}$$

where R and V are defined as before. From the previous also follows that $p(x)$ as defined in (1.3) is the uniquely determined nonzero monic eigenpolynomial of smallest degree corresponding to the smallest eigenvalue λ of H, and is therefore proved to have simple roots on the unit circle. By replacing

H by $-H$, the same holds true for the uniquely determined eigenpolynomial of smallest degree corresponding to the largest eigenvalue of H. This special root distribution of eigenpolynomials corresponding to the extremal eigenvalues of Hermitian Toeplitz matrices is widely discussed and repeatedly reproved in the literature, surely also because of its direct applications in the area of signal processing, see [4], [11], [10], [15], [16], [18], [19]. A good survey article on this subject is from Genin [9]. There it is stated that the following result from Delsarte, Genin and Kamp [5] (see also [6]) is the most general known one regarding the number of roots on the unit circle of any eigenpolynomial corresponding to any eigenvalue of a given complex Hermitian Toeplitz matrix:

Theorem 1.1 (Delsarte, Genin, Kamp). *Let $H \in \mathbb{C}^{n+1,n+1}$, $n \in \mathbb{N}_0$, be a Hermitian Toeplitz matrix with eigenvalues $\lambda_0 \leq \lambda_1 \leq ... \leq \lambda_n$, and let λ_s, $s \in \{0, ..., n\}$, be one of them with multiplicity m and Iohvidov parameter k. The index s shall be chosen such that either $s = 0$ or $\lambda_s > \lambda_{s-1}$. Then, any eigenpolynomial $p_s(x)$ corresponding to λ_s has at least $|n - m - 2s + 1|$ and at most $n - 2k$ zeros on the complex unit circle.*

Recall that for complex Hermitian matrices geometric and algebraic multiplicities of an eigenvalue coincide, i.e.,

$$\dim\ker(H - \lambda_s I_{n+1}) = m = \max\{i \in \mathbb{N} \mid (x - \lambda_s)^i \text{ divides } \det(xI_{n+1} - H)\} .$$

Both bounds stated in Theorem 1.1 are sharp in the sense that examples exist that attain them. The upper bound $n - 2k$ involves the less commonly known Iohvidov parameter k of an eigenvalue λ of a Hermitian, non-diagonal Toeplitz matrix $H = H(a_0, ..., a_n) \in \mathbb{C}^{n+1,n+1}$. It is defined as follows: Let m be the multiplicity of λ and let $r \in \{0, ..., n\}$ be maximal subject to λ being not an eigenvalue of the principal submatrix $H_r := H(a_0, ..., a_r)$, i.e., $H_r - \lambda I_{r+1}$ is regular. If such an r does not exist, set $r := -1$. It follows from results of Iohvidov [14], that $n - m - r$ is an even, non-negative integer, wherefore the Iohvidov parameter

$$k := \frac{1}{2}(n - m - r) \tag{1.5}$$

is well-defined. It is known, (see, for example, [5]) that any eigenpolynomial $q(x)$ of H corresponding to λ has the form

$$q(x) = x^k p(x) s(x) \ , \tag{1.6}$$

where $p(x)$ is the uniquely determined monic eigenpolynomial of degree $r+1$ of H_{r+1} corresponding to the eigenvalue λ and $s(x)$ is an arbitrary polynomial of degree at most $m-1$. Since $\deg(p(x)s(x)) \leq r+m = n-2k$, this is clearly an upper bound for the number of zeros on the unit circle of $q(x)$. Thus, actually, only the lower bound $|n - m - 2s + 1|$ given in Theorem 1.1 is non-trivial.

Delsarte, Genin and Kamp obtain this bound by considering two-variable Levinson polynomials and exploring their properties. One purpose of this article is to give a different, in my opinion more conceptual proof, exploiting the observation of Schur [20] that companion matrices of eigenpolynomials of Hermitian Toeplitz matrices define isometries of Hermitian forms defined by certain Toeplitz submatrices. Looking at the signature of these Hermitian forms and the possible normal forms of those isometries immediately gives the result. Moreover, this approach additionally allows to deduce statements not only on the total number of roots on the unit circle but also on their multiplicities. Recall that also the original result of Carathéodory involves a statement on multiplicities, namely that all roots of the eigenpolynomial of smallest degree corresponding to the smallest/largest eigenvalue of a non-diagonal Hermitian Toeplitz matrix are simple. This fact is by no means trivial compared to just proving that all roots of that eigenpolynomial are located on the unit circle. Furthermore, the perspective of normal forms of Schur's isometries does not rely on the field of complex numbers and draws direction to describing the eigenpolynomial structure of Hermitian or symmetric Toeplitz matrices over arbitrary fields. This will be explained in the sequel.

The main result of this article for the classical complex Hermitian case refining Theorem 1.1 reads as follows:

Theorem 1.2 (Main Theorem for complex Hermitian Toeplitz matrices). *Let $H \in \mathbb{C}^{n+1,n+1}$, $n \in \mathbb{N}_0$, be a Hermitian Toeplitz matrix with eigenvalues $\lambda_0 \leq \lambda_1 \leq \ldots \leq \lambda_n$, and let $\lambda = \lambda_s$, $s \in \{0, \ldots, n\}$, be one of them with multiplicity m and Iohvidov parameter k. The index s shall be chosen such that either $s = 0$ or $\lambda_s > \lambda_{s-1}$. Furthermore, let $p(x)$ be the monic eigenpolynomial of smallest degree corresponding to λ. The distinct roots of $p(x)$ on the unit circle are denoted by $\alpha_1, \ldots, \alpha_a$, $a \in \mathbb{N}_0$, and their multiplicities by $m_1, \ldots, m_a$. The remaining non-zero roots of $p(x)$ occur in conjugate pairs $\{\beta_1, \overline{\beta_1}^{-1}\}, \ldots, \{\beta_b, \overline{\beta_b}^{-1}\}$, $b \in \mathbb{N}_0$, where β_i and $\overline{\beta_i}^{-1}$ have the same multiplicity n_i, $i = 1, \ldots, b$. Then,*

$$a \geq |\{i \in \{1, \ldots, a\} \mid m_i \text{ odd }\}| \geq |n - m - 2s + 1| \ , \quad (1.7)$$

$$\sum_{i=1}^{a} \left\lfloor \frac{m_i}{2} \right\rfloor + \sum_{i=1}^{b} n_i \leq \frac{n - m + 1 - |n - m - 2s + 1|}{2} - k \ . \quad (1.8)$$

Note that in the extremal cases $s = 0$ (smallest eigenvalue) and $s = n - m + 1$ (largest eigenvalue) holds

$$\frac{n - m + 1 - |n - m - 2s + 1|}{2} = 0$$

wherefore (1.8) implies $k = 0 = b$, and $m_i = 1$ for all $i = 1, \ldots, a$, i.e., all roots of $p(x)$ lie on the unit circle and are simple. This is the classical result following from Carathéodory's theorem. Note also, that by Equation (1.6) Theorem 1.2 gives (sharp) lower bounds for the multiplicities of the roots on the unit circle of any eigenpolynomial.

Next, we want to elucidate the geometrical background of (1.7),(1.8): The spectrum shifted Hermitian Toeplitz matrix $G := H - \lambda_s I_{n+1}$ has $n_- := s$

negative eigenvalues $\lambda_i - \lambda_s$, $i = 0, ..., s-1$, $n_+ := n+1-s-m$ positive eigenvalues, $\lambda_i - \lambda_s$, $i = s+m, ..., n$, and $n_0 := m$ is the dimension of the kernel, i.e., the multiplicity of the eigenvalue zero. The triple (n_-, n_0, n_+) is the index of inertia of G, where n^+ and n_- are the positive and negative index of inertia, respectively. The difference $\mathrm{sign(G) := n_+ - n_-}$ is the signature of G. By Witt's decomposition theorem applied to the Hermitian form defined by G there is an up to isometry uniquely determined G-orthogonal decomposition $\mathbb{C}^{n+1} = A \textcircled{\perp} W \textcircled{\perp} \ker(G)$, where A is G-anisotropic (meaning here positive or negative definite) with dimension

$$\dim A = |\mathrm{sign(G)}| = |\mathrm{n_+ - n_-}| = |\mathrm{n - m - 2s + 1}| \ , \tag{1.9}$$

and W is a so-called hyperbolic space of even dimension

$$\begin{aligned}\dim W &= \dim \mathbb{C}^{n+1} - \dim \ker(G) - \dim A \\ &= n+1-n_0 - |n_+ - n_-| = n+1-m-|n-m-2s+1| \ . \end{aligned}\tag{1.10}$$

The number

$$\mathrm{ind}(W) := \frac{\dim W}{2} = \frac{n+1-m-|n-m-2s+1|}{2} \tag{1.11}$$

is the so-called Witt index of W which is the dimension of a maximal totally G-isotropic subspace of W. Then, the maximal dimension of a totally G-isotropic subspace of the whole space $\mathbb{C}^{n+1}$ clearly is

$$\mathrm{ind}(G) := \mathrm{ind}(W) + n_0 \ .$$

Neglecting the addend n_0, which is the dimension of the G-radical $\ker(G)$, $\mathrm{ind}(W)$ is the essential part of the Witt index of G. Thus, in the light of (1.9) and (1.10), Inequality (1.7) means that the number of roots on the unit circle of the eigenpolynomial $p(x)$ of smallest degree corresponding to the eigenvalue λ_s is bounded from below by $|\mathrm{sign(H - \lambda_s I_{n+1})}|$ and Inequality (1.8) describes how their multiplicities m_i are related to $\mathrm{ind}(W) - k$, the difference of the essential part of the Witt index and the Iohvidov parameter of $H - \lambda_s I_{n+1}$. (Clearly, dropping the Iohvidov parameter $k \geq 0$, gives the coarser upper bound $\mathrm{ind}(W) = \frac{n+1-m-|n-m-2s+1|}{2}$.)

In the extremal cases $s = 0$ and $s = n$, $|\mathrm{sign(H - \lambda_s I_{n+1})}| = \mathrm{n + 1 - m}$ is maximal which clarifies the symmetry of both cases. In general, for the considered purpose, for any eigenvalue λ of H of multiplicity m_λ,

$$t_\lambda := |\mathrm{sign(H - \lambda I_{n+1})}| = \mathrm{n + 1 + m_\lambda - 2 \cdot ind(H - \lambda I_{n+1})} \ , \tag{1.12}$$

which is the dimension of the anisotropic subspace (the core form) in a Witt decomposition of $H - \lambda I_{n+1}$, can be seen as the more direct, appropriate label for λ than its position s in the ordered chain of all eigenvalues which relies on the ordering of the real numbers. In this regard, recall that the index of inertia as defined here for complex Hermitian matrices is not available for arbitrary fields. In contrast, Witt's decomposition theorem remains true for Hermitian and also for symmetric or symplectic [2] matrices over arbitrary fields $\mathbb{K}$ (with characteristic $\mathrm{char}(\mathbb{K}) \neq 2$ in the symmetric, non-symplectic case). Accordingly, the Witt index is always well-defined in these cases. Since, in such general cases, algebraic and geometric multiplicities of eigenvalues may differ, m_λ must be taken as the geometric multiplicity, i.e., $m_\lambda := \dim\ker(H - \lambda I_{n+1})$. See Section 2 for basic facts in geometric algebra like Witt decomposition, Witt index, etc. In this spirit, we obtain the following general main theorem from which Theorem 1.2 follows as a special case.

Theorem 1.3 (Main Theorem for Hermitian or symmetric Toeplitz matrices over arbitrary fields).
Let $\mathbb{K}$ be a field, let α be an involutory field automorphism of $\mathbb{K}$, i.e., $\alpha^2 = \mathrm{id}_\mathbb{K}$, and let $H \in \mathbb{K}^{n+1,n+1}$, $n \in \mathbb{N}_0$, be an α-Hermitian or symmetric Toeplitz matrix. We assume $\alpha \neq \mathrm{id}_\mathbb{K}$ for the Hermitian case and $\alpha = \mathrm{id}_\mathbb{K}$ and $\mathrm{char}(K) \neq 2$ for the symmetric case. Thus, $\alpha(H_{j,i}) = H_{i,j} = H_{i',j'}$ for all $i, j, i', j' \in \{0, ..., n\}$ with $|i - j| = |i' - j'|$. Now, let $\lambda \in \mathbb{K}$ be an eigenvalue of H with $\alpha(\lambda) = \lambda$ and geometric multiplicity m, and let $p(x) \in \mathbb{K}[x]$ be the uniquely determined monic eigenpolynomial of smallest degree corresponding to λ Then, $p(x)$ admits a prime factor decomposition

$$p(x) = x^k \prod_{i=1}^{a} q_i(x)^{m_i} \prod_{j=1}^{b} (r_j(x) r_j^*(x) \alpha(r_j(0))^{-1})^{n_j} \tag{1.13}$$

where $a, b, k \in \mathbb{N}_0$, $m_i, n_j \in \mathbb{N}$, $q_i(x), r_j(x), r_j^*(x)\alpha(r_j(0))^{-1} \in \mathbb{K}[x]\backslash\{x\}$ are pairwise distinct, monic prime polynomials, and the $q_i(x)$ are α-symmetric, $i = 1, ..., a$, $j = 1, ..., b$. Here, for a polynomial $s(x) = \sum_{i=0}^{d} s_i x^i \in \mathbb{K}[x]$ of degree $d \geq 0$, the α-reciprocal polynomial is defined by $s^*(x) := x^d \alpha(s)(x^{-1}) = \sum_{i=0}^{d} \alpha(s_i) x^{d-i}$, and $s(x)$ is α-symmetric if $s^*(x)$ is a scalar multiple of $s(x)$. Since $(H - \lambda I_{n+1})^T = \alpha(H - \lambda I_{n+1})$ is α-Hermitian respectively symmetric,

$$t := n + 1 + m - 2 \cdot \mathrm{ind}(H - \lambda I_{n+1}) \tag{1.14}$$

is well-defined. It satisfies

$$\sum_{\substack{i=1 \\ m_i \text{ odd}}}^{a} \deg(q_i(x)) \geq t \ , \tag{1.15}$$

$$\begin{aligned} \sum_{i=1}^{a} \left\lfloor \frac{m_i}{2} \right\rfloor \deg(q_i(x)) + \sum_{j=1}^{b} n_j \deg(r_j(x)) &\leq \frac{n - m + 1 - t}{2} - k \\ &= \mathrm{ind}(H - \lambda I_{n+1}) - m - k \ . \end{aligned} \tag{1.16}$$

The reason why in the Hermitian case $\alpha \neq \mathrm{id}_\mathbb{K}$ only eigenvalues λ with $\alpha(\lambda) = \lambda$ are considered is that $H - \lambda I_{n+1}$ shall again be a Hermitian Toeplitz

[2]A Matrix $S \in \mathbb{K}^{n,n}$ is symplectic, if $v^T S v = 0$ for all $v \in \mathbb{K}^n$. This is equivalent to $S^T = -S$ and $S_{i,i} = 0$ for $i = 1, ..., n$. If $\mathrm{char}(\mathbb{K}) \neq 2$, then the latter condition follows from the former and can therefore be dropped in that case.

matrix for which, in particular, $\mathrm{ind}(H-\lambda I_{n+1})$ is well defined. For $\mathbb{K}=\mathbb{C}$ and Hermitian H, all eigenvalues of H are real wherefore that condition is no restriction. In this case (1.7), (1.8) of Theorem 1.2 follow from (1.15), (1.16).

We remark that it seems not much promising to consider more general divisors $c(x)$ of the characteristic polynomial of H with $\alpha(c)(x)=c(x)$ instead of just linear ones like $x-\lambda$ because, even though $c(H)^T=\alpha(c(H))$ means that $c(H)$ is again Hermitian or symmetric, $c(H)$ will in general not be Toeplitz which is essential for the proof of Theorem 1.3. Also, it does not make much sense to include symplectic Toeplitz matrices, as for $\lambda\neq 0$, $H-\lambda I_{n+1}$ will not be symplectic. Indeed, for $\mathrm{char}(K)=2$, $H-\lambda I_{n+1}$ would be symmetric but Witt's decomposition theorem does not hold true for symmetric, non-symplectic forms in characteristic 2 so that $\mathrm{ind}(H-\lambda I_{n+1})$ would be not well-defined. Nevertheless, some aspects of symplectic Toeplitz matrices will be considered on the way to the proof of Theorem 1.3. Finally, we remark that it will be shown that the definition of the Iohvidov parameter k, see (1.5), appearing without further explanation in (1.13) as exponent of x, and the eigenspace representation (1.6) remain valid for the general situation of Theorem 1.3 so that (1.15), (1.16) can again be adapted to arbitrary eigenpolynomials corresponding to λ.

Continuing Theorem 1.3, the following Theorem 1.4 formulates a generalization of the classical Caratheodory representation (1.4) for arbitrary eigenvalues λ with $\alpha(\lambda)=\lambda$.

Theorem 1.4 (Generalized Caratheodory representation).
The notation of Theorem 1.3 *is kept and k is the Iohvidov-Parameter of $H-\lambda I_{n+1}$. Suppose that all roots of $p(x)$ are contained in $\mathbb{K}$, i.e. $q_i(x)=x-\varepsilon_i$ and $r_j(x)=x-\delta_j$ for suitable $\varepsilon_i,\delta_j\in\mathbb{K}\backslash\{0\}$ such that $\alpha(\varepsilon_i)\varepsilon_i=1\neq\alpha(\delta_j)\delta_j$ and such that all $\varepsilon_i,\delta_j,\alpha(\delta_j)^{-1}$, $i=1,...,a$, $j=1,...,b$, are pairwise distinct. Thus, $p(x)=x^k\prod_{i=1}^a(x-\varepsilon_i)^{m_i}\prod_{j=1}^b((x-\delta_j)(x-\alpha(\delta_j)^{-1}))^{n_j}$. Set $r:=\deg(p(x))-k-1=\sum_{i=1}^a m_i+2\sum_{j=1}^b n_j-1=n-m-2k$. If $r>-1$, let $V\in\mathbb{K}^{n+1-k,r+1}$ be the confluent Vandermonde matrix*

$$
\begin{aligned}
V &= [V_1,...,V_a,W_1,W_1^*,...,W_b,W_b^*]\ , && (1.17)\\
V_i &:= V(\varepsilon_i,n+1-k,m_i)\in\mathbb{K}^{n+1-k,m_i}\ ,\ i=1,...,a,\\
W_j &:= V(\delta_j,n+1-k,n_j)\in\mathbb{K}^{n+1-k,n_j}\ ,\ j=1,...,b,\\
W_j^* &:= V(\alpha(\delta_j)^{-1},n+1-k,n_j)\in\mathbb{K}^{n+1-k,n_j}\ ,\ j=1,...,b,
\end{aligned}
$$

where for arbitrary $\mu\in\mathbb{K}$ and $s,t\in\mathbb{N}$, $V(\mu,s,t)\in\mathbb{K}^{s,t}$ is defined by

$$V(\mu,s,t)_{i,j}=\begin{cases}\binom{i}{j}\mu^{i-j} & ,\ \text{if } i\geq j,\ ^3\\ 0 & ,\ \text{else},\end{cases} \tag{1.18}$$

for $i=0,...,s-1$, $j=0,...,t-1$. Let $H_{n-k}\in\mathbb{K}^{n+1-k,n+1-k}$ denote the upper left submatrix of H of order $n+1-k$. Then,

$$H_{n-k}=\alpha(V)RV^T+\lambda I_{n+1-k} \tag{1.19}$$

for some regular block diagonal matrix $R = \operatorname{diag}(Q_1, ..., Q_a, R_1, ..., R_b)$ *with* $Q_i \in \mathbb{K}^{m_i,m_i}$, $\alpha(Q_i)^T = Q_i$, $R_j = \begin{bmatrix} 0 & \alpha(S_j)^T \\ S_j & 0 \end{bmatrix} \in \mathbb{K}^{2n_j,2n_j}$, $S_j \in \mathbb{K}^{n_j,n_j}$. *If* $r = -1$, *then* $H_{n-k} = \lambda I_{n+1-k}$.

We recall the meaning of $V(\mu, s, t)$ defined in (1.18) under the simplifying assumption $\operatorname{char}(\mathbb{K}) = 0$: For an arbitrary polynomial

$$q(x) = \sum_{i=0}^{s-1} q_i x^i \in \mathbb{K}[x]$$

of degree at most $s-1$ with coefficient vector $q := (q_0, q_1, ..., q_{s-1})$ holds

$$qV(\mu, s, t) = (q(\mu), q'(\mu), \frac{1}{2}q''(\mu), ..., \frac{1}{j!}q^{(j)}(\mu), ..., \frac{1}{(t-1)!}q^{(t-1)}(\mu))$$

where

$$q^{(j)}(x) := \sum_{i=j}^{s-1} \frac{i!}{(i-j)!} q_i x^{i-j}$$

is the formal j-th derivative of $q(x)$. Thus, if $s > t$, then $qV(\mu, s, t) = 0$ if, and only if, μ is a t-fold root of $q(x)$.

Corollary 1.5. *If* $k = 0$ *and* $H \neq \lambda I_{n+1}$ *(i.e.* $r > -1$*) in Theorem* 1.4, *then*

$$H = \overline{V} R V^T + \lambda I_{n+1} \; . \tag{1.20}$$

If moreover all roots of $p(x)$ *are simple, i.e.* $m_i = 1 = n_j$ *for all* i, j, *then* V *is the usual Vandermonde matrix*

$$V = \begin{bmatrix} 1 & \dots & 1 & 1 & 1 & \dots & 1 & 1 \\ \varepsilon_1 & \dots & \varepsilon_a & \delta_1 & \alpha(\delta_1)^{-1} & \dots & \delta_b & \alpha(\delta_b)^{-1} \\ \varepsilon_1^2 & \dots & \varepsilon_a^2 & \delta_1^2 & \alpha(\delta_1)^{-2} & \dots & \delta_b^2 & \alpha(\delta_b)^{-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ \varepsilon_1^n & \dots & \varepsilon_a^n & \delta_1^n & \alpha(\delta_1)^{-n} & \dots & \delta_b^n & \alpha(\delta_b)^{-n} \end{bmatrix}$$

and R *becomes*

$$R = \operatorname{diag}\left(r_1, ..., r_a, \begin{bmatrix} 0 & \alpha(s_1) \\ s_1 & 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 & \alpha(s_b) \\ s_b & 0 \end{bmatrix}\right)$$

for nonzero $r_i, s_j \in \mathbb{K}$ *with* $\alpha(r_i) = r_i$. *Thus, the first matrix row of* (1.20) *reads componentwise*

$$h_0 := H_{0,0} = \lambda + \sum_{i=1}^{a} r_i + \sum_{j=1}^{b} s_j + \alpha(s_j) \tag{1.21}$$

$$h_l := H_{0,l} = \sum_{i=1}^{a} r_i \varepsilon_i^l + \sum_{j=1}^{b} s_j \delta_j^l + \alpha(s_j)\alpha(\delta_j)^{-l} \; , \; l = 1, ..., n. \tag{1.22}$$

---

[3]For $x, y \in \mathbb{N}_0$ with $x \geq y$, $z := \binom{x}{y} \in \mathbb{N}$ is interpreted as an element of $\mathbb{K}$ in the usual way as $\sum_{i=1}^{z} 1_{\mathbb{K}}$ which might be zero if $\operatorname{char}(\mathbb{K})$ divides z.

For $\mathbb{K} = \mathbb{C}$ and Hermitian H, (1.21) and (1.22) are equivalent to Theorem 5 of [6]. [4] In this case, and if λ is the smallest eigenvalue of H, then, by Theorem 1.2, as noted before, $k = 0 = b$ and $H - \lambda I_{n+1}$ is non-negative definite wherefore $r_i > 0$ for all i. Thus (1.22) becomes the classical Result (1.1) of Caratheodory. (If $H = \lambda I_{n+1}$, i.e. if $r = -1$, then $a = 0 = b$ and (1.22) still holds true.)

The article is organized as follows: In Section 2 normal forms of isometries of classical Hermitian, symmetric and symplectic forms over arbitrary fields are revisited. Section 3 fixes a notation and states some simple properties of Toeplitz matrixces which are needed in the subsequent sections. Section 4 establishes an extension theorem for Hermitian or symmetric or symplectic Toeplitz matrices over arbitrary fields. Originally, this was proved by Iohvidov [14] for complex Hermitian Toeplitz matrices. In section 5 Schur's isometry related to the monic eigenpolynomial of smallest degree in the kernel of a singular Hermitian or symmetric or symplectic Toeplitz matrix is introduced. In Section 6 the proof of Theorem 1.3 is given by simply looking at normal forms of Schur's isometries. As already mentioned, Theorem 1.2 follows as a special case. Finally, Theorem 1.4 is proved.

2. Normal forms of Isometries

In this section we shortly state the classification of all possible normal forms of isometries of finite-dimensional, regular, Hermitian, symmetric or symplectic spaces. The description follows Huppert's book [13].

The derivation of normal forms of finite-dimensional linear mappings is one of the main topics in undergraduate linear algebra lectures and is also taught to students of applied sciences, wherefore linear normal forms are broadly known also to engineers. Also metric structures like Hermitian or Euclidean inner products on finite-dimensional vector spaces and corresponding isometries, i.e., linear mappings that preserve the given metric structure, belong to the basic linear algebra education. In view of the importance of symmetry considerations in all areas of natural sciences, this is well justified. Even though near at hand, classifying the possible normal forms of isometries does not belong to the usual undergraduate linear algebra canon and is, for general fields $\mathbb{K}$ and regular unitary, symmetric or symplectic sesquilinear forms, even hardly known to mathematicians. In my opinion, this is quite a gap in basic education. Theorems 2.1 and 2.2 below state this general classification. The casual reader or those who are anyway only interested in the classical case of complex numbers, which is probably of most importance in

[4] Ciccariello and Cervellino [3] published a slightly different version of Theorem 5 of [6]. They also consider the case of Iohvidov-Parameter $k = 0$ which is Condition (i) in their "Generalized Carathéodory theorem" on page 14916. Their main improvement seems to be the replacement of the assumption that all roots of $p(x)$ are simple by a second Condition (ii) based on their Lemma A which assures this and also that all roots are unimodular, i.e. $b = 0$. Actually, this Lemma A is not new as they proposed but follows directly from a Theorem of Herglotz and Krein, see [14] Theorem A.I.2, p.194.

view of applications, may skip these theorems and just take note of Theorem 2.3.

For formulating the stated theorems, we need to introduce some mainly commonly known facts and standard notation: In the sequel $\mathbb{K}$ always denotes a field, for example, $\mathbb{K} = \mathbb{C}$, the field of complex numbers, or $\mathbb{K} = \mathbb{R}$, the field of real numbers, or $\mathbb{K} = \mathbb{F}_q$, the finite field with q elements, where q is some prime power. Furthermore, V denotes a finite-dimensional vector space over $\mathbb{K}$ of some dimension $n \in \mathbb{N}$, and α denotes a field automorphism of $\mathbb{K}$, i.e., a bijection of $\mathbb{K}$ which satisfies $\alpha(ab) = \alpha(a)\alpha(b)$ and $\alpha(a+b) = \alpha(a) + \alpha(b)$ for all $a, b \in \mathbb{K}$. An α-sesquilinear form on V is a function $f : V \times V \to \mathbb{K}$ that satisfies

a) $f(u+v, w) = f(u, w) + f(v, w)$
b) $f(u, v+w) = f(u, v) + f(u, w)$
c) $f(au, v) = \alpha(a)f(u, v)$
d) $f(u, av) = af(u, v)$

for all $u, v, w \in V$ and all $a \in \mathbb{K}$. It is called orthosymmetric if $f(u, v) = 0$ implies $f(v, u) = 0$ for all $u, v \in V$, i.e., if the orthogonality relation

$$u \perp v \quad :\Leftrightarrow \quad f(u, v) = 0$$

is symmetric, which will be assumed from know on.

Two nonempty subsets U, W of V are called orthogonal if $u \perp w$ for all $u \in U$, $w \in W$, and we write $U \perp W$ in this case. If two subspaces U, W of V satisfy $U \perp W$ and $U \cap W = \{0\}$, then their direct sum $U \oplus W$ is also denoted by $U \textcircled{\perp} W$ in order to stress the orthogonality relation. The orthogonal complement of a nonempty subset U of V is defined by $U^\perp := \{v \in V \mid v \perp U\}$. Recall that $U^\perp$ is always a subspace of V. If U is a subspace, then $\dim U^\perp = \dim V - \dim U$. The subspace U is called regular, if its radical $\mathrm{rad}(U) := U \cap U^\perp$ is the zero space. In this case holds $V = U \textcircled{\perp} U^\perp$. The form f is called regular, if the radical $\mathrm{rad}(f) := \mathrm{rad}(V) = V^\perp = \{0\}$. As it is generally known, f is orthosymmetric if, and only if, one of the following cases holds: .

a) $\alpha = \mathrm{id}_\mathbb{K}$ and f is symmetric, i.e $f(u, v) = f(v, u)$ for all $u, v \in V$.
b) $\alpha = \mathrm{id}_\mathbb{K}$ and f is symplectic, i.e $f(u, u) = 0$ for all $u \in V$.
c) $\alpha^2 = \mathrm{id}_\mathbb{K} \neq \alpha$ and there is a nonzero $a \in \mathbb{K}$ such that the α-sesquilinear form g defined by $g(u, v) := a \cdot f(u, v)$, $u, v \in V$, is α-Hermitian, i.e., $g(u, v) = \alpha(g(v, u))$ for all $u, v \in V$.
d) $\alpha \neq \mathrm{id}_\mathbb{K}$ and $\dim \mathrm{rad}(V) \geq \dim V - 1$. (negligible, trivial case)

Recall that for symplectic f always holds $f(u, v) = -f(v, u)$ for all $u, v \in V$, and, if the characteristic of $\mathbb{K}$, denoted by $\mathrm{char}(\mathbb{K})$, is distinct from 2, then this property is equivalent to f being symplectic. If $\mathrm{char}(\mathbb{K}) = 2$, this is not true, but note, since $-1 = 1$ in this case, a symplectic form is always a special kind of a symmetric form. For $\mathrm{char}(K) \neq 2$, the property $f(u, v) = -f(v, u)$ for all $u, v \in V$ is known as skew-symmetry. Thus, in this case, symplectic forms are also called skew-symmetric forms. Analogously, for $\alpha^2 = \mathrm{id}_\mathbb{K} \neq \alpha$

and $\mathrm{char}(\mathbb{K}) \neq 2$ forms satisfying $\alpha(f(u,v)) = -f(v,u)$ are called skew-Hermitian, but, since there exists always a nonzero $a \in \mathbb{K}$ with $\alpha(a) = -a$, the form $a \cdot f$ is Hermitian so that c) applies.

From now on we assume that either $\alpha = \mathrm{id}_{\mathbb{K}}$ and f is symmetric or symplectic or $\alpha^2 = \mathrm{id}_{\mathbb{K}} \neq \alpha$ and f is Hermitian (meaning α-Hermitian). In matrix notation this translates to

$$\alpha(G)^T = \varepsilon G$$

where

$$G := (f(v_i, v_j)_{0 \le i,j \le n-1}) \in \mathbb{K}^{n,n}$$

is the Gramian matrix of f corresponding to some basis $v_0, ..., v_{n-1}$ of V, and $\varepsilon := -1$ if $\mathrm{char}(\mathbb{K}) \neq 2$ and f is symplectic, and $\varepsilon := 1$ otherwise. The Gramian matrix G is symmetric or symplectic or α-Hermitian if, and only if, f has this property.

A vector $u \in V$ is called isotropic if $f(u,u) = 0$, i.e., $u \perp u$, and anisotropic if $f(u,u) \neq 0$. A subspace U of V is called totally isotropic if $f(u_1, u_2) = 0$ for all $u_1, u_2 \in U$, and it is called anisotropic if all its nonzero vectors are anisotropic. [5] A regular subspace H of V is a hyperbolic space if it admits a decomposition $H = H_1 \oplus H_2$ into two totally isotropic subspaces H_1 and H_2. Then, necessarily, $\dim H_1 = \dim H_2$, and therefore $\dim H = 2 \dim H_1$ is even. In case of $\dim H_i = 1$, $i = 1, 2$, H is called a hyperbolic plane. In general, a hyperbolic space H always admits an orthogonal decomposition into $m := \frac{1}{2} \dim H$ hyperbolic planes, and, conversely, every direct orthogonal sum of hyperbolic planes is a hyperbolic space.

An orthosymmetric α-sesquilinear form f is called classical if one of the following conditions holds:

a) f is symmetric and $\mathrm{char}(K) \neq 2$.
b) f is Hermitian.
c) f is symplectic.

A classical form always admits a Witt decomposition

$$V = H \textcircled{\perp} A \textcircled{\perp} \mathrm{rad}(f)$$

into a hyperbolic subspace H, an anisotropic subspace A and the radical $\mathrm{rad}(f)$. This decomposition is up to isometry unique. Therefore, the dimensions of all maximal totally isotropic subspaces of V are equal, and this common dimension is called the Witt index of f (or V), denoted by

$$\mathrm{ind}(f) := \mathrm{ind}(V) = \dim \mathrm{rad}(f) + \frac{1}{2} \dim H \ .$$

The restriction of f to A (or even A itself) is called the core form of f.

[5]The zero space $\{0\}$ is the only subspace of V which is both, totally isotropic and anisotropic.

If $\mathbb{K} = \mathbb{C}$ and if f is Hermitian with Gramian matrix G having index of inertia (n_-, n_0, n_+), then

$$\begin{aligned} \dim \operatorname{rad}(f) &= n_0 = \dim \ker(G) \ , \\ \operatorname{ind}(G) := \operatorname{ind}(f) &= \min\{n_+, n_-\} + n_0 \ , \\ \dim A &= |\mathrm{sign(G)}| = |\mathrm{n}_+ - \mathrm{n}_-| \ , \end{aligned}$$

and f restricted to A is positive/negative definite if sign(G) is positive/negative.

A linear isomorphism φ of V is an isometry of f, if

$$f(\varphi(u), \varphi(v)) = f(u, v)$$

for all $u, v \in V$. In matrix notation this reads

$$\alpha(M)^T G M = G$$

where M is the matrix of φ and G is the Gramian matrix of f with respect to some common basis of V. For regular G this identity implies that the characteristic polynomial char(φ) and the minimal polynomial mip(φ) of φ are α-symmetric. Recall from Theorem 1.3 that a polynomial

$$p(x) = \sum_{i=0}^{m} p_i x^i \in K[x] \ , \quad p_m \neq 0,$$

of degree $m \in \mathbb{N}_0$ is said to be α-symmetric if the α-reciprocal polynomial $p^*(x)$ of $p(x)$ defined by

$$p^*(x) := x^m \alpha(p)(x^{-1}) = \sum_{i=0}^{m} \alpha(p_i) x^{m-i} \tag{2.1}$$

is a scalar multiple of $p(x)$. A φ-invariant subspace of V will be called a φ-(sub)module, and the whole vector space V is φ-decomposable if there exist nontrivial φ-submodules U and W of V such that $V = U \perp W$. If this is not the case, V is called orthogonally φ-indecomposable. Clearly, V admits always a decomposition $V = V_1 \perp V_2 \perp \ldots \perp V_m$ into orthogonally φ-indecomposable φ-submodules $V_1, \ldots, V_m$, $m \in \mathbb{N}$. Therefore, having a classification of normal forms of isometries in mind, it is sufficient to restrict to orthogonally indecomposable modules.

Theorem 2.1 (Normal forms of isometries, part 1).
Let V be a finite-dimensional vector space over a field $\mathbb{K}$, let α be an involutory field automorphism of $\mathbb{K}$, i.e., $\alpha^2 = \mathrm{id}_{\mathbb{K}}$, and let V be endowed with a regular, orthosymmetric α-sesquilinear form $f : V \times V \to \mathbb{K}$. Furthermore, let φ be an isometry of V with respect to f, and suppose that V is an orthogonally indecomposable φ-module. Then, V and φ fit to exactly one of the following types:

T1. *$V = V_1 \oplus V_2$ for indecomposable φ-modules V_1, V_2, $\mathrm{mip}(\varphi|_{V_1}) = p(x)^t$, $\mathrm{mip}(\varphi|_{V_2}) = p^*(x)^t$, $p(x) \in \mathbb{K}[x]$ is irreducible, $\gcd(p(x), p^*(x)) = 1$, $t \in \mathbb{N}$. (Then, necessarily, V_1 and V_2 are totally isotropic subspaces which means that V is a hyperbolic space.)* [6]

T2. *$V = V_1 \oplus V_2$ for indecomposable φ-modules V_1, V_2, $\operatorname{mip}(\varphi|_{V_1}) = p(x)^t = \operatorname{mip}(\varphi|_{V_2})$, $p(x) \in \mathbb{K}[x]$ irreducible and α-symmetric, $t \in \mathbb{N}$.*

T3. *V is an indecomposable φ-module, $\operatorname{mip}(\varphi) = p(x)^t$, $p(x) \in \mathbb{K}[x]$ irreducible and α-symmetric, $t \in \mathbb{N}$.*
(Then, necessarily, $p(\varphi)^{\lceil \frac{t}{2} \rceil} V = \ker(p(\varphi)^{\lfloor \frac{t}{2} \rfloor})$ is a totally isotropic subspace of V.)

Now, if it is additionally assumed that the sesquilinear form f is either symmetric, symplectic or α-unitary, which are the most common cases, then Theorem 2.1 can be refined. (Recall that $\alpha = \mathrm{id}_{\mathbb{K}}$ in the symmetric or symplectic case.)

Theorem 2.2 (Normal forms of isometries, part 2).
Let $p(x) \in \mathbb{K}[x]$ be monic and irreducible.

1. *If $\gcd(p(x), p^*(x)) = 1$, then, for any $t \in \mathbb{N}$, indecomposable φ-modules V of type T1 with $\operatorname{mip}(\varphi) = p(x)p^*(x)^t$ can be constructed.*
2. *If $p(x)$ is α-symmetric and if (V, φ) is of type T2 with $\operatorname{mip}(\varphi) = p(x)^t$, $t \in \mathbb{N}$, then $\alpha = 1$ and $p(x) = x \pm 1$. If additionally $\operatorname{char}(\mathbb{K}) \neq 2$, then the following holds true:*
   i) *If f is symmetric, then t must be even, and for all even t modules of this type can be constructed.*
   ii) *If f is symplectic, then t must be odd, and for all odd t modules of this type can be constructed.*
   ii) *There exist totally isotropic φ-submodules V_1 and V_2 of V such that $V = V_1 \oplus V_2$. In particular, V is a hyperbolic space.*
3. *Let $p(x)$ be α-symmetric. If $\alpha \neq 1$ or $p(x) \neq x \pm 1$, then, for any $t \in \mathbb{N}$, indecomposable φ-modules V of type T3 with $\operatorname{mip}(\varphi) = p(x)^t$ can be constructed. If $\alpha = 1$ and $p(x) = x \pm 1$ and $\operatorname{char}(\mathbb{K}) \neq 2$, then the following holds true:*
   i) *If f is symmetric, then t must be odd, and for all odd t modules of this type can be constructed.*
   ii) *If f is symplectic, then t must be even, and for all even t modules of this type can be constructed.*

   If $\alpha = 1$ and $p(x) = x \pm 1$ and $\operatorname{char}(\mathbb{K}) = 2$, then t must be even or equal to 1. Moreover:
   i) *For all even t symplectic modules (and hence symmetric modules) of this type can be constructed.*
   ii) *For the trivial case $t = 1 = \dim V$ holds $\varphi = \pm\mathrm{id}_V$ and clearly a symmetric, non-symplectic module of this type exists.* [7]

Proofs of Theorems 2.1, 2.2 are given in [13], §8, Theorems 8.7 and 8.9. For the special case $\mathbb{K} = \mathbb{C}$ and $\alpha(a) := \overline{a}$, $a \in \mathbb{C}$, these theorems imply:

[6] Here, and in Theorem 2.2, minimal polynomials are understood to be unique up to a non-zero scalar multiple.

[7] This trivial case was neglected in [13], Theorem 8.9.

Theorem 2.3 (Normal forms of unitary isometries for $\mathbb{K}=\mathbb{C}$).
Let $H \in \mathbb{C}^{n,n}$, $n \in \mathbb{N}_0$, be a regular, Hermitian matrix with indices of positive and negative inertia n^+ and n^-, respectively, and Witt index $\operatorname{ind}(H) := \min\{n_+, n_-\}$*. Furthermore, let $M \in \mathbb{C}^{n,n}$ be a unitary isometry of V with respect to H, i.e., $\overline{M}^T HM = H$. Assume that $V := \mathbb{C}^n$ cannot be decomposed into proper H-orthogonal, M-invariant subspaces. Then, one of the following cases holds true:*

T1. *$V = V_1 \oplus V_2$ for M-indecomposable, M-invariant subspaces V_1, V_2 with* $\operatorname{mip}(M|_{V_1}) = (x-a)^t$, $\operatorname{mip}(M|_{V_2}) = (x-\overline{a}^{-1})^t$ *for some $a \in \mathbb{C}\backslash\{0\}$ with $a\overline{a} \neq 1$ and $t := \frac{n}{2}$. Necessarily V_1 and V_2 are totally isotropic subspaces which implies $t = n^+ = n^- = \operatorname{ind}(H)$.*

T3. *V is M-indecomposable with* $\operatorname{mip}(M) = (x-a)^n$ *for some $a \in \mathbb{C}\backslash\{0\}$ with $a\overline{a} = 1$, i.e., a lies on the unit circle. Moreover,*

$$\operatorname{ind}(H) = \min\{n^+, n^-\} = \left\lfloor \frac{n}{2} \right\rfloor \ .$$

In particular, H is positive or negative definite, if, and only if, $n = 1$ and $H = a$ for some a on the unit circle.

Proof. It only remains to prove $\min\{n^+, n^-\} = \lfloor \frac{n}{2} \rfloor =: t$ in case T3. Clearly, $\operatorname{ind}(H) \leq t$. Set $s := n - t \geq \frac{n}{2}$. The subspace $U := \ker((M - aI_n)^t) = (M - aI_n)^s V$ has dimension $\dim U = t$ and is totally isotropic because for $u, v \in V$ holds

$$\begin{aligned}&\overline{(M-aI_n)^s u}^T H (M-aI_n)^s v \\ &= \overline{u}^T H(-\overline{a}M^{-1})^s (M - \overline{a}^{-1} I_n)^s (M - aI_n)^s v \\ &= \overline{u}^T H(-\overline{a}M^{-1})^s (M - aI_n)^{2s} v = 0 \ .\end{aligned}$$

Thus $\operatorname{ind}(H) \geq \dim U = t$, and therefore $\operatorname{ind}(H) = t$. □

3. Classical Toeplitz forms: Notation and elementary properties

For the rest of the article the following notation is fixed:

$\mathbb{K}$ is a field, α is an involutory field automorphism of $\mathbb{K}$, i.e., $\alpha^2 = \mathrm{id}_{\mathbb{K}}$, and $H \in \mathbb{K}^{n+1,n+1}$, $n \in \mathbb{N}_0$, is an α-Hermitian or symmetric or symplectic Toeplitz matrix of order $n+1$. We assume $\alpha \neq \mathrm{id}_{\mathbb{K}}$ for the Hermitian case, $\alpha = \mathrm{id}_{\mathbb{K}}$ for the symmetric or symplectic case and $\operatorname{char}(K) \neq 2$ for the symmetric, non-symplectic case, unless otherwise stated. Furthermore, set $\varepsilon := -1$ if H is symplectic and $\operatorname{char}(K) \neq 2$ and $\varepsilon := 1$ otherwise so that $\alpha(H)^T = \varepsilon H$. Under these assumptions, the α-sesqulinear form

$$f = f_H : \mathbb{K}^{n+1} \times \mathbb{K}^{n+1} \to \mathbb{K}, \quad f(v,w) := \alpha(v)^T H w$$

with Gramian matrix H is a classical one and therefore admits an up to isometry unique Witt decomposition. In this context H will also be called a classical Toeplitz matrix. The property $\alpha(H)^T = \varepsilon H$ is equivalent to

$\alpha(f(v,w)) = \varepsilon f(w,v)$ for all $v,w \in \mathbb{K}^{n+1}$. The elements of the first row of H are denoted by $h_0, h_1, ..., h_n \in \mathbb{K}$ so that the (i,j)-entry of H is $H_{i,j} = h_{j-i}$ for $i,j \in \{0,...,n\}$ where $h_{-k} := \varepsilon\alpha(h_k)$ for $k \in \{1,...,n\}$. For $r \in \{0,...,n\}$,

$$H_r := (h_{i-j})_{0\le i,j\le r} \in \mathbb{K}^{r+1,r+1}$$

denotes the upper left submatrix of order $r+1$ of H. Next, set

$$J := (\delta_{i,n-i})_{0\le i,j\le n} = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix} \in \mathbb{K}^{n+1,n+1}.$$

For an arbitrary matrix $A = (a_{i,j})_{0\le i,j\le n} \in \mathbb{K}^{n+1,n+1}$ and $i,j \in \{0,...,n\}$ holds $(J^T AJ)_{i,j} = A_{n-i,n-j}$, i.e., $J^T AJ$ is the matrix obtained by reflecting all entries at the center of the matrix. An essential property of the Toeplitz matrix H is

$$J^T HJ = \varepsilon\alpha(H) = H^T. \tag{3.1}$$

In the symmetric case this means $J^T HJ = H$, i.e., H is centrosymmetric, wherefore J is an isometry with respect to the classical form f defined by H. In the Hermitian case holds $\alpha(J)^T HJ = J^T HJ = \alpha(H)$, i.e. H is α-centrosymmetric, wherefore, in general, J is not an isometry with respect to H. Finally, for symplectic, non-symmetric H holds $\mathrm{char}(K) \ne 2$ and $J^T HJ = -H$, i.e. H is skew centrosymmetric, and J is not and isometry of H.

As already indicated by the notation of eigenpolynomials, it will be very convenient to identify a vector $v = (v_0,...,v_n)^T \in \mathbb{K}^n$ with the corresponding polynomial $v(x) := (1,x,...,x^n)\cdot v = \sum_{i=0}^n v_i x^i$. For example, we may write $Hv(x)$ for the matrix vector product Hv. Now, define

$$v^{(*,n)} := \alpha(Jv) = \alpha(v_n, v_{n-1}, ..., v_0)^T. \tag{3.2}$$

Then

$$v^{(*,n)}(x) = \sum_{i=0}^{n} \alpha(v_{n-i})x^i = x^n\alpha(v)(x^{-1}) = x^{n-\deg(v(x))}v^*(x)\ , \tag{3.3}$$

where $v^*(x) := x^{\deg(v(x))}\alpha(v)(x^{-1})$ is the α-reciprocal polynomial of $v(x)$ like defined in (2.1). Now, (3.1) implies

$$f(v,w) = f(w^{(*,n)}, v^{(*,n)}) \tag{3.4}$$

for all $v,w \in \mathbb{K}^{n+1}$. A trivial but useful consequence of (3.4) is:

Remark 3.1. The radical of f, which is the kernel of H, is α-symmetric, meaning $v \in \mathrm{rad}(f)$ if, and only if, $v^{(*,n)} \in \mathrm{rad}(f)$.

Another simple property of Toeplitz forms is their shift invariance:

Remark 3.2. For $u,v \in \mathbb{K}^{n+1}$ with $\deg(u(x)), \deg(v(x)) < n$ holds

$$f(u(x), v(x)) = f(xu(x), xv(x))\ .$$

Proof. For $\tilde{u} := (u_0, ..., u_{n-1})^T$ and $\tilde{v} := (v_0, ..., v_{n-1})^T$ holds $u = (\tilde{u}^T, 0)^T$, $v = (\tilde{v}^T, 0)^T$ and

$$f(u(x), v(x)) = \alpha(\tilde{u}^T, 0) \begin{bmatrix} H_{n-1} & * \\ * & * \end{bmatrix} \begin{bmatrix} \tilde{v} \\ 0 \end{bmatrix} = \alpha(0, \tilde{u}^T) \begin{bmatrix} * & * \\ * & H_{n-1} \end{bmatrix} \begin{bmatrix} 0 \\ \tilde{v} \end{bmatrix}$$
$$= f(xu(x), xv(x)) \ .$$

□

In the subsequent section the ranks and the kernels of the upper left submatrices $H_r \in \mathbb{K}^{r+1,r+1}$ of H, $r \leq n$, will be investigated. For simplicity of notation, subspaces and vectors of $\mathbb{K}^{r+1}$ on which H_r acts will be identified with their natural embeddings into larger spaces $\mathbb{K}^{r+i}$, $i = 0, ..., n-r$, by simply adding trailing zeros to the components. Thus, for example, a vector $u = (u_0, u_1)^T \in \mathbb{K}^2$ is, depending on the context, identified with $u' := (u_0, u_1, 0)^T \in \mathbb{K}^3$ or $u'' := (u_0, u_1, 0, 0)^T \in \mathbb{K}^4$. In this manner subspaces of $\mathbb{K}^r$ and, in particular, $\mathbb{K}^r$ itself are considered as subspaces of $\mathbb{K}^{r+i}$. In the sequel we will tacitly work with this slight ambiguity. Note carefully, that intrinsic metric properties of subspaces U of $\mathbb{K}^r$ with respect to the form f_{H^r} defined by H_r like being totally isotropic or regular or hyperbolic remain valid under the stated natural embedding of U into a larger space $\mathbb{K}^{r+i}$ and simultaneous extension of the classical form to $f_{H^{r+i}}$ defined by H_{r+i}. For technical reasons we formally define H_{-1} to be an empty matrix with no (zero) rows and columns. Formally, we write $H_{-1} \in \mathbb{K}^{0,0}$ and define $\det H_{-1} := 1$, which is consistent with Leibniz' product formula for the determinant, and wherefore H_{-1} is considered as "regular". Moreover, we define $\mathrm{rank(H_{-1})} := 0$, $\dim\ker(H_{-1}) := 0$, $\mathrm{ind}(H_{-1}) := 0$.

4. Iohvidov's extension Theorems for classical Toeplitz forms over arbitrary fields

In this section results of Iohvidov [14], Chap. III, §15, on the jumps in the ranks of complex Hermitian Toeplitz matrices will be established for symmetric or Hermitian or symplectic Toeplitz matrices over arbitrary fields. In this course, the Iohvidov parameter (see (1.5)) will be defined in this general setting and it will also be shown that the eigenpolynomial representation (1.6) remains valid for $\lambda = 0$ with geometric multiplicity m, i.e., for m-dimensional kernel.

Even though Iohvidov considered only complex Hermitian Toeplitz matrices, possibly almost all of his arguments can be carried over to arbitrary fields and also to symmetric and symplectic Toeplitz matrices. But, of course, this cannot be assumed or stated without a detailed proof, especially because Iohvidov's exposition, although very well organized and furnished with accurate proofs, is quite lengthy since it is based on direct calculations with Toeplitz matrices, their specific entries and minors. Thus, repeating Iohvidov's arguments would be quite tiresome and inappropriate for a publication. Therefore, I will reprove the needed parts in a different, less computational, more structural way by means of geometric algebra. As a byproduct this will demonstrate the connection between the Iohvidov parameter and the Witt index which is new to my knowledge. In short, the Iohvidov parameter k of the classical Toeplitz matrix H (see the Notation fixed in Section 3) is nothing but the difference of the Witt index of H and that of its largest regular upper left submatrix H' and the dimension of the kernel of H which is the dimension of the radical of H:

$$k = \operatorname{ind}(H) - \operatorname{ind}(H') - \dim\ker(H)\ . \tag{4.1}$$

Thinking in terms of Witt decompositions this implies that Witt decompositions of H and H' simply differ by k additional hyperbolic planes and a radical of dimension $\dim\ker(H)$. [8]

Theorem 4.1 (Iohvidov's law on the jumps in the rank). *Let $r \in \{-1, 0, 1, .., n\}$ be maximal subject to $\det H_r \neq 0$, i.e., $H_r \in \mathbb{K}^{r+1,r+1}$ is the largest regular upper left submatrix of H. Furthermore, let $m := \dim\ker(H)$ so that $\rho := \mathrm{rank(H) = n + 1 - m}$.*

a) *$\rho - (r + 1) = n - m - r$ is even so that the Iohvidov parameter*

$$k := \frac{n - m - r}{2}$$

is a well-defined non-negative integer. If H is singular, then also H_{r+1} is singular with one-dimensional kernel. In this case, the uniquely determined monic non-zero eigenpolynomial $p(x)$ lying in the kernel of H_{r+1} is α-symmetric and has degree $r + 1$. Moreover,

$$\begin{aligned}\ker(H) &= \operatorname{span}\{x^{k+i}p(x) \mid 0 \le i \le m-1\}\\ &= \{x^k p(x)s(x) \mid s(x) \in \mathbb{K}[x], \deg(s(x)) \le m-1\}\end{aligned}$$

and

$$\begin{aligned}U_1 &:= \operatorname{span}\{x^i p(x) \mid 0 \le i < k\}\\ U_2 &:= \operatorname{span}\{x^{m+k+i}p(x) \mid 0 \le i < k\}\end{aligned}$$

are totally isotropic, k-dimensional subspaces such that

$$\mathbb{K}^{n+1} = ((\mathbb{K}^{r+1} \operatorname{\bigcirc\!\!\!\!\perp} U_1) \oplus U_2) \operatorname{\bigcirc\!\!\!\!\perp} \ker(H).$$

Furthermore, the two-dimensional subspaces

$$W_i := \operatorname{span}\{x^i p(x), x^{k+m+i}p(x)\}\ , \quad i = 0, ..., k-1\ ,$$

spanned by the isotropic vectors $x^i p(x), x^{k+m+i}p(x)$ are hyperbolic planes which are not necessarily pairwise orthogonal but satisfy $U_1 \oplus U_2 = W_0 \oplus ... \oplus W_{k-1}$.

[8] In lofty language, this means for the symmetric case that the quadratic forms corresponding to H and H' are equivalent in the Witt group of the coefficient field $\mathbb{K}$ modulo radical.

b) $\mathrm{rank}(\mathrm{H_{r+i}}) = \mathrm{r}+1$ *for* $i = 0, ..., k+m$ *and* $\mathrm{rank}(\mathrm{H_{r+k+m+i}}) = \mathrm{r}+1+2\mathrm{i}$ *for* $i = 1, ..., k$*. In words: During the first* $k+m$ *extension steps of the regular matrix* H_r*, the rank remains the same, namely* $r+1$*, while during the final* k *extension steps the rank increases by two in each step.*

c) *If* $\mathbb{K}^{r+1} = A \oplus W$ *is a Witt decomposition for* H_r *with anisotropic subspace* A *and hyperbolic subspace* W*, then there is a hyperbolic subspace* U *of* $\mathbb{K}^{n+1}$ *with* $\dim U = 2k$ *such that* $\mathbb{K}^{n+1} = A \oplus W \oplus U \oplus \ker(H)$ *is a Witt decomposition for* H*. In particular,*

$$\mathrm{ind}(H) = \frac{1}{2}(\dim W + \dim U) + \dim\ker(H) = \mathrm{ind}(H_r) + k + m.$$

This means that a Witt decomposition for H_r *differs from one for* H *only by a radical of dimension* m *and a hyperbolic subspace the Witt index of which is the Iohvidov parameter.* [9]

Before proving this Theorem we remark that the definition of the parameter r follows Genin [9]. It differs from Iohvidov's definition [14] by 1, i.e., $r_{\mathrm{Iohvidov}} = r+1$. For reasons of certain notational benefits we work (like Genin [9]) with a given Toeplitz matrix H of order $n+1$ while Iohvidov uses a Toeplitz matrix of order n. In order to keep the final formula $k = \frac{n-m-r}{2}$ for the Iohvidov parameter unchanged, the shift in the parameter r is necessary.

The following trivial linear algebra statements grouped in the remark below will be used for the proof of Theorem 4.1:

Remark 4.2. Let V be a finite-dimensional $\mathbb{K}$-vector space with $\dim V \geq 1$, let $\varphi : V \to V$ be a linear mapping, let W be a hyperplane of V, i.e., $\dim W = \dim V - 1$, and let $pr_W : V \to W$ be a linear projection from V to W, i.e., $pr_W(V) = W$ and $pr_W(w) = w$ for all $w \in W$. Then, for the linear mapping

$$\psi := pr_W \circ \varphi|_W : W \to W, \quad w \mapsto pr_W(\varphi(w))$$

the following holds:

a) $\ker(\psi) = (\ker(\varphi) \cap W) \oplus U$ for a subspace U with $\dim U \leq 1$
b) $\dim\ker(\varphi) + 1 \geq \dim\ker(\psi) \geq \dim(\ker(\psi) \cap \ker(\varphi)) \geq \dim\ker(\varphi) - 1$
c) $\mathrm{rank}(\varphi) \geq \mathrm{rank}(\psi) \geq \mathrm{rank}(\varphi) - 2$
d) $\dim\ker(\psi) < \dim\ker(\varphi) \quad \Leftrightarrow \quad \ker(\psi) \subsetneqq \ker(\varphi)$

Proof.
a) $\varphi(\ker(\psi)) \subseteq \ker(pr_W)$ and $\dim\ker(pr_W) = 1$ imply

$$\ker(\psi) = (\ker(\varphi) \cap \ker(\psi)) \oplus U = (\ker(\varphi) \cap W) \oplus U$$

for an at most one-dimensional subspace U of $\ker(\psi)$.

b) Part a) implies

$$\dim\ker(\psi) = \dim(\ker(\varphi) \cap W) + \dim(U) \leq \dim\ker(\varphi) + 1 \ .$$

[9] For $r = 0$, A and W may formally be considered as zero spaces.

Next, the dimension formula

$$\dim(A+B) = \dim A + \dim B - \dim(A \cap B)$$

for arbitrary subspaces A, B of V applied to $\ker(\varphi) \cap W = \ker(\varphi|_W) = \ker(\psi) \cap \ker(\varphi)$ yields

$$\begin{aligned}\dim(\ker(\varphi) \cap W) &= \dim\ker(\varphi) + \dim W - \dim(\ker(\varphi) + W) \\ &\geq \dim\ker(\varphi) + \dim W - \dim V = \dim\ker(\varphi) - 1\ .\end{aligned}$$

c) Clearly, $\mathrm{rank}(\varphi) \geq \mathrm{rank}(\psi)$ and from b) follows

$$\mathrm{rank}(\psi) = \dim \mathrm{W} - \dim\ker(\psi) \geq \dim \mathrm{V} - 1 - (\dim\ker(\varphi) + 1) = \mathrm{rank}(\varphi) - 2\ .$$

d) The implication "$\Leftarrow$" is trivial. For proving the other direction, suppose that $\dim\ker(\psi) < \dim\ker(\varphi)$. Then, by a) and b) $\dim U = 0$ wherefore $\ker(\psi) = \ker(\varphi) \cap W \subsetneqq \ker(\varphi)$. □

Remark 4.2 will be used to relate the kernel/rank of some submatrix $H_i \in \mathbb{K}^{i+1,i+1}$ of H, $0 \leq i < n$, to the kernel/rank of the "one-step-extension" $H_{i+1} \in \mathbb{K}^{i+2,i+2}$. Thus φ will be identified with H_{i+1}, ψ with H_i, $V := \mathbb{K}^{i+2}$, $W := \mathbb{K}^{i+1}$ considered as a hyperplane of V, and pr_W is the projection from V to W that sets the last component of a vector $v := (v_0, ..., v_i, v_{i+1})^T \in V$ to zero, i.e., $pr_W(v) = (v_0, ..., v_i, 0)^T$.

Proof of Theorem 4.1. a) First, the simple case $\rho = r + 1$ is considered. Then, $n - r = m$ and $k := \frac{n-r-m}{2} = 0$ is a non-negative integer. Suppose that H is singular. By Remark 4.2 b) (or for trivial reasons if $r = -1$), $\dim\ker(H_{i+1}) \leq \dim\ker(H_i) + 1$, for $i = r, .., n-1$. Since $\dim\ker(H_r) = 0$ and $n - r = m = \dim\ker(H_n)$, this implies $\dim\ker(H_{i+1}) = \dim\ker(H_i) + 1$ for $i = r, .., n-1$. By Remark 4.2 d), $p(x) \in \ker(H_{r+i}) \subsetneqq \ker(H_{r+i+1})$, $i = 1, ..., m-1$. By Remark 3.1, $\ker(H_{r+i})$ is α-symmetric, so that $x^{i-1}p(x) = \alpha(p(0))^{-1}p^{(*,r+i)}(x) \in \ker(H_{r+i})$, $i = 1, ..., m$. Thus

$$\ker(H) = \cup_{i=1}^{m} \ker(H_{r+i}) \supset \{x^{i-1}p(x) \mid i = 1, ..., m\}\ .$$

Clearly, $p(x), xp(x), x^2p(x), ..., x^{m-1}p(x)$ are m linearly independent polynomials and therefore span $\ker(H)$. (Recall that polynomials are identified with their coefficient vectors prolonged by trailing zeros up to the dimension of the considered vector space.) This proves a) for $\rho = r + 1$.

The general proof of a) proceeds by induction on n. First, consider the base case $n = 0$. If $\rho = 1 = n + 1$, then H is regular wherefore $r = n = 0$ and $\rho = r + 1$. If $\rho = 0$, then $H = 0$ wherefore $r = -1$ and again $\rho = r + 1$. Thus, in any case, $\rho = r + 1$ holds true, and this case was proved in the beginning.

Next, for the inductive step, let $n > 0$ and assume that the assertion is true for $\tilde{H} := H_{n-1}$. We may also assume that

$$\rho > r + 1 \tag{4.2}$$

because the case $\rho = r + 1$ is already solved. Set $\tilde{m} := \dim\ker\tilde{H}$, $\tilde{\rho} := \mathrm{rank}(\tilde{\mathrm{H}}) = \mathrm{n} - \tilde{\mathrm{m}}$. Since $r < \rho - 1 \leq n$, H_r is also the largest regular upper left submatrix of $\tilde{H}$ and H_{r+1} is also an upper left submatrix of $\tilde{H}$. Now, by the induction hypothesis

$$\tilde{k} := \frac{\tilde{\rho} - (r+1)}{2} = \frac{n - 1 - \tilde{m} - r}{2}$$

is a non-negative integer,

$$\ker(\tilde{H}) = \operatorname{span}\{x^{\tilde{k}+i} p(x) \mid 0 \leq i \leq \tilde{m} - 1\} ,$$

and

$$\tilde{U}_1 := \operatorname{span}\{x^i p(x) \mid 0 \leq i < \tilde{k}\} ,$$

$$\tilde{U}_2 := \operatorname{span}\{x^{\tilde{m}+\tilde{k}+i} p(x) \mid 0 \leq i < \tilde{k}\}$$

are totally isotropic, k-dimensional subspaces such that

$$\mathbb{K}^n = ((\mathbb{K}^{r+1} \operatorname{\textcircled{\perp}} \tilde{U}_1) \oplus \tilde{U}_2) \operatorname{\textcircled{\perp}} \ker(\tilde{H}). \tag{4.3}$$

Furthermore, the two-dimensional subspaces

$$W_i := \operatorname{span}\{x^i p(x), x^{\tilde{k}+\tilde{m}+i} p(x)\} , \quad i = 0, ..., \tilde{k} - 1 \tag{4.4}$$

spanned by the isotropic vectors $x^i p(x), x^{\tilde{k}+\tilde{m}+i} p(x)$ are hyperbolic planes satisfying $\tilde{U}_1 \oplus \tilde{U}_2 = W_0 \oplus ... \oplus W_{\tilde{k}-1}$.

The decomposition (4.3) implies, in particular, that $W := \mathbb{K}^{r+1} \oplus \tilde{U}_1 \oplus \tilde{U}_2$ is a regular subspace. Now, consider $W^{\perp_H}$, the orthogonal complement of W in the larger space $\mathbb{K}^{n+1}$. Clearly $\ker(\tilde{H}) \subseteq W^{\perp_H}$, and, since W is regular, we have

$$\mathbb{K}^{n+1} = W \operatorname{\textcircled{\perp}} W^{\perp_H} , \tag{4.5}$$

and $\dim W^{\perp_H} = n + 1 - \dim W = \dim \ker(\tilde{H}) + 1$. By Remark 4.2 b), $|m - \tilde{m}| \leq 1$.

Case 1: $m = \tilde{m}$**.** Assume that $\ker(H) \neq \ker(\tilde{H})$. Then, there must exist a $v \in \ker(H) \setminus \ker(\tilde{H})$. Therefore, $W^{\perp_H} = \ker(\tilde{H}) \operatorname{\textcircled{\perp}} \langle v \rangle$ [10] is totally isotropic and (4.5) implies

$$\ker(\tilde{H}) \subsetneq W^{\perp_H} = \operatorname{rad}(H) = \ker(H) ,$$

a contradiction, i.e., this case cannot occur.

Case 2: $m = \tilde{m} + 1$**.** Then, by (4.2), $\tilde{k} = \frac{n-m-r}{2} = \frac{\rho-(r+1)}{2} > 0$. Remark 4.2 d) yields $x^{\tilde{k}} p(x) \in \ker(\tilde{H}) \subsetneq \ker(H)$. Since $\ker(H)$ is α-symmetric (c.f. Remark 3.1), we have

$$q(x) := (x^{\tilde{k}} p(x))^{(*,n)} = x^{n-(\tilde{k}+r+1)} p^*(x) = \alpha(p(0)) x^{n-(\tilde{k}+r+1)} p(x) \in \ker(H).$$

But $n - (\tilde{k} + r + 1) = \tilde{k} + \tilde{m}$ and $\tilde{k} > 0$ imply

$$q(x) \in \tilde{U}_2 \cap \ker(H) \subseteq (\tilde{U}_1 \oplus \tilde{U}_2) \cap \ker(\tilde{H}) = \{0\},$$

a contradiction.

[10]For a vector v, $\langle v \rangle := \operatorname{span}\{v\} = \mathbb{K}v$ denotes the subspace spanned by v.

Case 3: $m = \tilde{m} - 1$**.** Then, $n - m - r = (n - 1 - \tilde{m} - r) + 2 = 2(\tilde{k} + 1)$ is even and $k := \frac{n-m-r}{2} = \tilde{k} + 1$ is a well-defined positive integer. Now, assume that there exists a $v \in \ker(H) \backslash \ker(\tilde{H})$. Then, like in case 1, it follows that $W^{\perp_H} = \ker(\tilde{H}) \textcircled{\perp} \langle v \rangle$ is totally isotropic and (4.5) implies

$$\ker(\tilde{H}) \subsetneqq W^{\perp_H} = \mathrm{rad}(H) = \ker(H) \ ,$$

wherefore $\tilde{m} < m$, a contradiction. Hence,

$$\ker(H) \subsetneqq \ker(\tilde{H}) = \mathrm{span}\{x^{\tilde{k}+i}p(x) \mid 0 \leq i \leq \tilde{m} - 1 = m\} \ .$$

Since $x^{m+2k-1}p(x) = x^{n-r-1}p(x)$ has degree n, clearly

$$\mathbb{K}^{n+1} = \mathbb{K}^n \oplus \langle x^{m+2k-1}p(x) \rangle \ . \tag{4.6}$$

For $i \in \{1, ..., m\}$ the shift invariance of Toeplitz forms (c.f. Remark 3.2) yields

$$f(x^{\tilde{k}+i}p(x), x^{m+2k-1}p(x)) = f(x^{\tilde{k}+i-1}p(x), x^{m+2k-2}p(x)) = 0$$

since $x^{m+2k-2}p(x) \in \mathbb{K}^n$ and $x^{\tilde{k}+i-1}p(x) \in \ker(\tilde{H})$. Therefore, $x^{\tilde{k}+i} \in \mathrm{rad}(f) = \ker(H)$ and, comparing dimensions, we conclude

$$\begin{aligned} \ker(H) &= \mathrm{span}\{x^{\tilde{k}+i}p(x) \mid 1 \leq i \leq m\} \\ &= \mathrm{span}\{x^{k+i}p(x) \mid 0 \leq i \leq m - 1\}. \end{aligned} \tag{4.7}$$

In particular, $x^{k-1}p(x) = x^{\tilde{k}}p(x) \in \ker(\tilde{H}) \backslash \ker H$ and (4.6) imply

$$f(x^{k-1}p(x), x^{m+2k-1}p(x)) \neq 0.$$

Hence, $W_{k-1} := \mathrm{span}\{x^{k-1}p(x), x^{m+2k-1}p(x)\}$ is a hyperbolic plane. Define

$$\begin{aligned} U_1 &:= \mathrm{span}\{x^i p(x) \mid 0 \leq i < k\} = \tilde{U}_1 \oplus \langle x^{k-1}p(x) \rangle \ , \\ U_2 &:= \mathrm{span}\{x^{m+k+i}p(x) \mid 0 \leq i < k\} = \tilde{U}_2 \oplus \langle x^{m+2k-1}p(x) \rangle \ . \end{aligned}$$

Clearly, since $x^{k-1}p(x) \in \ker(\tilde{H}) \subseteq \tilde{U}_1^{\perp} \backslash \tilde{U}_1$, U_1 is k-dimensional, totally isotropic subspace satisfying $\mathbb{K}^{r+1} \textcircled{\perp} U_1$. Also $\dim U_2 = k$, and using (4.3), (4.4) supplies the decompositions

$$\mathbb{K}^{n+1} = ((\mathbb{K}^{r+1} \textcircled{\perp} U_1) \oplus U_2) \textcircled{\perp} \ker(H) \quad , \quad U_1 \oplus U_2 = W_1 \oplus ... \oplus W_{k-1}.$$

It remains to show that U_2 is totally isotropic. For that, it suffices to show that the isotropic vector $x^{m+2k-1}p(x)$ is orthogonal to the totally isotropic subspace $\tilde{U}_2$ which follows immediately from the shift invariance of Toeplitz forms: Take a basis vector $x^{\tilde{k}+\tilde{m}+i}p(x) = x^{k+m+i}p(x) \in \tilde{U}_2$, $0 \leq i \leq \tilde{k} - 1$, Then,

$$f(x^{k+m+i}p(x), x^{m+2k-1}p(x)) = f(p(x), x^{k-1-i}p(x)) = \{0\}$$

because $p(x), x^{k-1-i}p(x) \in \tilde{U}_1$. This finishes the inductive step and the proof of a).

b) $\mathbb{K}^{r+1+i} \subseteq \mathbb{K}^{r+1} \textcircled{\perp} U_1 \textcircled{\perp} \ker(H)$ for $i = 0, ..., k + m$. Hence,

$$\dim\ker(H_{r+k+m}) = \dim U_1 + \dim\ker(H) = k + m$$

and therefore

$$r+1 = \mathrm{rank(H_r)} \le \mathrm{rank(H_{r+i})} \le \mathrm{rank(H_{r+k+m})}$$
$$= (r+1+k+m) - (k+m) = r+1\ ,$$

i.e., $\mathrm{rank(H_{r+i}) = r+1}$ for $i = 0, ..., k+m$. Now, using Remark 4.2 gives

$$\mathrm{rank(H_{r+k+m})} = r+1\ ,$$
$$\mathrm{rank(H_{r+k+m+i+1})} \le \mathrm{rank(H_{r+k+m+i})} + 2\ , \quad \mathrm{i = 0, ..., k-1},$$
$$\mathrm{rank(H)} = r+1+2k$$

which implies $\mathrm{rank(H_{r+k+m+i}) = r+1+2i}$ for $i = 0, ..., k$.

c) $A \perp W$ is regular so that $\mathbb{K}^{n+1} = (A \perp W) \perp (A \perp W)^\perp$. Since $(U_1 \perp \ker(H)) \subseteq (A \perp W)^\perp$ is a totally isotropic subspace of maximal dimension, Witt's decomposition theorem supplies a totally isotropic subspace U_2' with $\dim U_2' = \dim U_1 = k$ such that $U = U_1 \oplus U_2'$ is a hyperbolic space and $(A \perp W)^\perp = U \perp \ker(H)$ is a Witt decomposition of $(A \perp W)^\perp$. Thus

$$\mathbb{K}^{n+1} = A \perp W \perp U \perp \ker(H)$$

is a Witt decomposition of $\mathbb{K}^{n+1}$ with maximal hyperbolic space $W \perp U$ and anisotropic core form A. □

5. Schur's isometry

In the sequel we assume $n \ge 1$, that H is singular, and that the upper left submatrix

$$H_{n-1} = (h_{i-j})_{i,j=0,...,n-1} \in \mathbb{K}^{n,n}$$

of order n is non-singular. In this section we permit $\mathrm{char}(\mathbb{K}) = 2$ for symmetric, non-symplectic H.

Furthermore, let $v = (v_0, v_1, ..., v_n)^T \in \mathbb{K}^{n+1}$ denote a nonzero vector in the kernel of H, i.e., $Hv = 0$. Then $v_n \ne 0$, since otherwise $w := (v_0, ..., v_{n-1})^T$ would be a nonzero vector in the kernel of H_{n-1} contradicting its regularity. Thus, by dividing through $v_n \ne 0$, we can assume without loss of generality that $v_n = 1$. [11] Therefore, $Hv = 0$ reads row by row:

$$h_{n-i} = -\sum_{j=0}^{n-1} h_{j-i} v_j \quad , \quad i = 0, ..., n. \tag{5.1}$$

Setting $w := -(v_0, ..., v_{n-1})^T, g := (h_n, h_{n-1}, ..., h_1)^T \in \mathbb{K}^n$ this becomes in matrix vector notation

$$H_{n-1} w = g\ , \tag{5.2}$$
$$\varepsilon\alpha(g)^T w = h_0 = \varepsilon\alpha(h_0)\ . \tag{5.3}$$

[11] In the context of Theorem 4.1 a), $r = n-1$ and $p(x) = v(x)$ has degree $r+1 = n$.

Schur's [20] key observation is that the following companion matrix $C \in \mathbb{K}^{n,n}$ with characteristic polynomial $v(x) := \sum_{i=0}^{n} v_i x^i = x^n + \sum_{i=0}^{n-1} v_i x^i \in \mathbb{K}[x]$ is an isometry with respect to H_{n-1}, i.e., it satisfies $\alpha(C)^T H_{n-1} C = H_{n-1}$:

$$C := \left[\begin{array}{ccc|c} & & & -v_0 \\ \hline 1 & & & -v_1 \\ & \ddots & & \vdots \\ & & 1 & -v_{n-1} \end{array}\right] = \left[\begin{array}{c|c} 0 & \\ I_{n-1} & w \end{array}\right]. \tag{5.4}$$

This is easily verified by using (5.2), (5.3) and $\alpha(H_{n-1})^T = \varepsilon H_{n-1}$:

$$\begin{aligned} \left[\begin{array}{c|c} * & * \\ \hline * & \alpha(C)^T H_{n-1} C \end{array}\right] &= \left[\begin{array}{c} e_1^T \\ \hline \alpha(C)^T \end{array}\right] H_{n-1} \left[\begin{array}{c|c} e_1 & C \end{array}\right] \\ &= \left[\begin{array}{c} I_n \\ \hline \alpha(w)^T \end{array}\right] H_{n-1} \left[\begin{array}{c|c} I_n & w \end{array}\right] \\ &= \left[\begin{array}{c|c} H_{n-1} & g \\ \hline \alpha(w)^T H_{n-1} & \alpha(w)^T g \end{array}\right] \\ &= \left[\begin{array}{c|c} H_{n-1} & g \\ \hline \varepsilon\alpha(g)^T & h_0 \end{array}\right] = H \\ &= \left[\begin{array}{c|c} * & * \\ \hline * & H_{n-1} \end{array}\right]. \end{aligned} \tag{5.5}$$

Here $e_1 = (1, 0, ..., 0)^T \in \mathbb{K}^n$ denotes the first standard basis column vector of length n. Note that $\alpha(C)^T H_{n-1} C = H_{n-1}$ in particular means that C is regular which means that $v_0 = (-1)^n \det C \neq 0$. This can of course be seen directly for the same reason as $v_n \neq 0$ or by the fact that $v(x)$ is α-symmetric.

6. Proof of the main Theorems

Proof of Theorem 1.3. By replacing H by $H - \lambda I_{n+1}$, we may assume without loss of generality, that $\lambda = 0$. Define r and the Iohvidov parameter $k = \frac{n-m-r}{2}$ like in Theorem 4.1. From Theorem 4.1 it follows that $p(x) = x^k v(x)$, where $v(x) = x^{r+1} + \sum_{i=0}^{r} v_i x^i$ is the uniquely determined, α-symmetric, monic polynomial of degree $r+1$ that spans $\ker(H_{r+1})$. Thus, $v(x)$ admits a prime factor decomposition

$$v(x) = \prod_{i=1}^{a} q_i(x)^{m_i} \prod_{j=1}^{b} (r_j(x) r_j^*(x) \alpha(r_j(0))^{-1})^{n_j} \tag{6.1}$$

where $a, b \in \mathbb{N}_0$, $m_i, n_j \in \mathbb{N}$, $q_i(x), r_j(x), \alpha(r_j(0))^{-1} r_j^*(x) \in \mathbb{K}[x] \backslash \{x\}$ are pairwise distinct, monic prime polynomials, and $q_i(x)$ is α-symmetric, $i = 1, ..., a$, $j = 1, ..., b$. This supplies the prime factor decomposition of $p(x)$ given in (1.13). Let us first assume that $r > -1$. Consider Schur's isometry

C of H_r defined by

$$C := \left[\begin{array}{ccc|c} & & & -v_0 \\ \hline 1 & & & -v_1 \\ & \ddots & & \vdots \\ & & 1 & -v_r \end{array}\right] \in \mathbb{K}^{r+1,r+1}$$

which is a companion matrix with characteristic and minimal polynomial

$$\mathrm{char}(C) = \mathrm{mip}(C) = v(x) \ . \tag{6.2}$$

Considering the normal form of C according to Theorems 2.1 and 2.2 and using (6.1) and (6.2) yields the following decomposition of $\mathbb{K}^{r+1}$ with respect to H_r:

$$\mathbb{K}^{r+1} = [\ \textcircled{\perp}\,_{i=1}^{a} \ker(q_i(C)^{m_i})]\ \textcircled{\perp}\ [\ \textcircled{\perp}\,_{j=1}^{b} \ker(r_j(C)^{n_j}) \oplus \ker(r_j^*(C)^{n_j})]\ . \tag{6.3}$$

where $U_i := \ker(q_i(C)^{m_i})$, $W_j := \ker(r_j(C)^{n_j})$ and $W_j^* := \ker(r_j^*(C)^{n_j})$ are C-invariant, C-indecomposable subspaces with $\mathrm{mip}(C|_{U_i}) = q_i(x)^{m_i}$, $\mathrm{mip}(C|_{W_j}) = r_j(x)^{n_j}$, $\mathrm{mip}(C|_{W_j}) = r_j^*(x)^{n_j}$, respectively. Note, that (6.2) in particular means that modules of type T2 defined in Theorems 2.1 cannot occur. The regular spaces $W_j \oplus W_j^*$ are hyperbolic with totally isotropic subspaces W_j, W_j^* of dimension $\deg(r_j(x)^{n_j}) = n_j \deg(r_j(x))$, $j = 1, ..., b$. Also, each regular space U_i contains the totally isotropic subspace

$$\tilde{U}_i := \ker(q_i(C)^{\lfloor \frac{m_i}{2} \rfloor}) = q_i(C)^{\lceil \frac{m_i}{2} \rceil} U_i$$

with $\dim \tilde{U}_i = \lfloor \frac{m_i}{2} \rfloor \deg(q_i)$, $i = 1, ..., a$. Thus, the Witt index of H_r can be estimated from below by

$$\begin{aligned} \mathrm{ind}(H_r) &\geq \sum_{i=1}^{a} \mathrm{ind}(U_i) + \sum_{j=1}^{b} \mathrm{ind}(W_j \oplus W_j^*) \geq \sum_{i=1}^{a} \dim \tilde{U}_i + \sum_{j=1}^{b} \dim(W_j) \\ &= \sum_{i=1}^{a} \left\lfloor \frac{m_i}{2} \right\rfloor \deg(q_i) + \sum_{j=1}^{b} n_j \deg(r_j(x))\ . \end{aligned} \tag{6.4}$$

By Iohvidov's generalized Theorem 4.1, $\mathrm{ind}(H_r) = \mathrm{ind}(H) - k - m$ which, inserted for the left hand-side of (6.4), proves (1.16).

Recall, that t defined in (1.14) is the dimension of the core form of H denoted by $\mathrm{cdim}\, H$, i.e., $t = \dim A$ where A is the H-anisotropic subspace in a Witt decomposition $\mathbb{K}^{n+1} = A \textcircled{\perp} W \textcircled{\perp} \ker(H)$ with hyperbolic W. By Iohvidov's generalized Theorem 4.1, t is also the dimension of the core form of H_r. Denoting the dimension of the core form of a subspace U by

$$\mathrm{cdim}\, U := \dim U + \dim \mathrm{rad}(U) - 2\mathrm{ind}(U),$$

(6.3) implies

$$t = \operatorname{cdim} H = \operatorname{cdim} H_r \leq \sum_{i=1}^{a} \operatorname{cdim} U_i = \sum_{i=1}^{a} \dim U_i - 2\operatorname{ind}(U_i) \tag{6.5}$$

$$\leq \sum_{i=1}^{a} \dim U_i - 2\dim \tilde{U}_i = \sum_{\substack{i=1 \\ m_i \text{ odd}}}^{a} \dim U_i - 2\dim \tilde{U}_i = \sum_{\substack{i=1 \\ m_i \text{ odd}}}^{a} \deg(q_i(x)) \ .$$

This is (1.15) which finishes the the proof for $r > -1$.

In the degenerate case $r = -1$ both inequalities (6.4), (6.5) remain valid because then $a = 0 = b$ and $t = 0 = \operatorname{ind}(H_{-1})$ so that the left and right hand-sides of both inequalities are zero. □

As noted in the introduction, Theorem 1.2 follows as a special case.

In preparation for the proof of Theorem 1.4 we state the following well-known fact on confluent Vandermonde matrices. For the readers convenience, a proof is given in the appendix.

Remark 6.1. Let

$$p(x) = \prod_{i=1}^{m} (x - a_i)^{n_i} = x^n + \sum_{i=0}^{n-1} p_i x^i \in \mathbb{K}[x]$$

be a polynomial of degree $n \in \mathbb{N}$ which has m distinct roots $a_1, ..., a_n \in \mathbb{K}$ of multiplicities $n_1, ..., n_m \in \mathbb{N}$, respectively, with $n_1 + ... + n_m = n$. Then, the corresponding confluent Vandermonde matrix

$$V := [V(a_1, n, n_1), ..., V(a_m, n, n_m)] \in \mathbb{K}^{n,n}$$

is regular with determinant

$$\det V = \prod_{1 \leq i < j \leq m} (a_j - a_i)^{n_i n_j} \tag{6.6}$$

(see (1.18) for the definition of $V(a_i, n, n_i)$), and V transforms the companion matrix C with characteristic polynomial p to its Jordan normal form N:

$$C := \left[\begin{array}{cccc|c} & & & & -p_0 \\ \hline 1 & & & & -p_1 \\ & \ddots & & & \vdots \\ & & & 1 & -p_{n-1} \end{array}\right] \in \mathbb{K}^{n,n} \ ,$$

$$V^T C V^{-T} = N := \operatorname{diag}(J(a_1, n_1), ..., J(a_m, n_m)) \ ,$$

$$J(a_i, n_i) := \begin{bmatrix} a_i & & & \\ 1 & \ddots & & \\ & \ddots & a_i & \\ & & 1 & a_i \end{bmatrix} \in \mathbb{K}^{n_i, n_i} \ , \ i = 1, ..., m.$$

$[J(a_i, n_i)$ is the Jordan block of order n_i with eigenvalue a_i of algebraic multiplicity n_i and geometric multiplicity 1.]

Proof of Theorem 1.4. The notation used in the proof of Theorem 1.3 is kept. Recall that $\lambda = 0$ was assumed without loss of generality.

Let us again first consider the case $r > -1$ and let $V_r \in \mathbb{K}^{r+1,r+1}$ denote the upper square confluent Vandermonde submatrix of V. By the previous Remark 6.1, V_r is regular and transforms the companion matrix C (Schur's isometry of H_r) to its Jordan normal form N:

$$\begin{aligned} V_r^T C V_r^{-T} &= N \\ &:= \operatorname{diag}(J(\varepsilon_1, m_1), ..., J(\varepsilon_a, m_a), \\ &\qquad J(\delta_1, n_1), J(\alpha(\delta_1)^{-1}, n_1), ..., J(\delta_b, n_b), J(\alpha(\delta_b)^{-1}, n_b)) \ . \end{aligned}$$

Writing

$$V_r^{-T} = [A_1, ..., A_a, B_1, B_1^*, ..., B_b, B_b^*]$$

with $A_i \in \mathbb{K}^{r+1,m_i}$ and $B_j, B_j^* \in \mathbb{K}^{r+1,n_j}$, then the columns of A_i span $\ker(C - \varepsilon_i I_{r+1})^{m_i}$, the columns of B_j span $\ker(C - \delta_i I_{r+1})^{n_j}$, and the columns of B_j^* span $\ker(C - \alpha(\delta_j)^{-1} I_{r+1})^{n_j}$, $i = 1, ..., a$, $j = 1, ..., b$. Thus, by Theorem 2.1, the columns of A_i and $[B_j, B_j^*]$ span the H_r-orthogonally indecomposable H_r-regular C-modules which are H_r-perpendicular to each other. Moreover, the columns of $[B_j, B_j^*]$ span the hyperbolic space

$$\ker(C - \delta_j I_{r+1})^{n_j} \oplus \ker(C - \alpha(\delta_j)^{-1} I_{r+1})^{n_j}$$

of type T1 with H_r-totally isotropic, C-invariant subspaces $\ker(C - \delta_j I_{r+1})^{n_j}$ and $\ker(C - \alpha(\delta_j)^{-1} I_{r+1})^{n_j}$, respectively. This means

$$\alpha(V_r^{-1}) H_r V_r^{-T} = \operatorname{diag}(Q_1, ..., Q_a, R_1, ..., R_b) =: R \tag{6.7}$$

for regular blocks $Q_i \in \mathbb{K}^{m_i,m_i}$ with $\alpha(Q_i)^T = Q_i$ and

$$R_j = \begin{bmatrix} 0 & S_j \\ \alpha(S_j)^T & 0 \end{bmatrix} \in \mathbb{K}^{2n_j,2n_j} \ , \ S_j \in \mathbb{K}^{n_j,n_j} \ , \ i = 1, ..., a \ , \ j = 1, ..., b \ .$$

Bringing V_r to the right side of (6.7) gives

$$H_r = \alpha(V_r) R V_r^T \ . \tag{6.8}$$

First, the case $k = 0$ is considered and we have to prove

$$H = \alpha(V) R V^T \ . \tag{6.9}$$

Note that $n = r + m$ and $V \in \mathbb{K}^{n+1,r+1}$ in this case. By Theorem 4.1 the polynomial

$$v(x) = p(x) = \prod_{i=1}^{a} (x - \varepsilon_i)^{m_i} \prod_{j=1}^{b} ((x - \delta_j)(x - \alpha\delta_j{}^{-1}))^{n_j}$$

(cf. (6.1) with $k = 0$) of degree $r + 1$ fulfills

$$\ker(H) = \operatorname{span}\{x^i p(x) \mid 0 \le i \le m - 1\} \ .$$

But, for $0 \le i \le m-1$ also holds $V^T(x^i p(x)) = 0$ because by construction of the confluent Vandermonde matrix V all polynomials $q(x)$ of degree at most n that are divided by $p(x)$ fulfill by $V^T q(x) = 0$. Thus, using (6.8), for the ordered basis

$$(z_0, ..., z_n) := (1, x, ..., x^r, p(x), xp(x), ..., x^{m-1}p(x))$$

of $\mathbb{K}^{n+1}$ holds

$$\begin{aligned} \alpha(z_i)^T H z_j &= \begin{cases} \alpha(z_i)^T H_r z_j = \alpha(z_i)^T \alpha(V_r) R V_r^T z_j & \text{, if } i, j \le r, \\ 0 & \text{, else,} \end{cases} \\ &= \alpha(z_i)^T \alpha(V) R V^T z_j \end{aligned}$$

for all $i, j \in \{0, ..., n\}$. This implies (6.9).

Next, the case $k > 0$ is considered. By Theorem 4.1, $\tilde{H} := H_{n-k}$ has Iohvidov parameter $\tilde{k} = 0$ and kernel dimension $\tilde{m} = m + k$. Therefore,

$$\tilde{H} = H_{n-k} = \alpha(V) R V^T$$

follows from the previously considered case. This finishes the proof of Theorem 1.4 for $r > -1$. If $r = -1$, then Theorem 4.1 b) directly yields for $i := k + m$, $\mathrm{rank}(\mathrm{H_{n-k}} - \lambda \mathrm{I_{n-k+1}}) = \mathrm{r} + 1 = 0$, i.e. $H_{n-k} = \lambda I_{n-k+1}$. □

Appendix

Proof of Remark 6.1. Consider

$$\tilde{V} := [V(x_1, n, n_1), ..., V(x_m, n, n_m)] \in \mathbb{Z}[x_1, ..., x_m]^{n,n} \subseteq \mathbb{R}[x_1, ..., x_m]^{n,n}$$

with indeterminates $x_1, ..., x_m$. Then, by [12] or [1], for example,

$$\det \tilde{V} = \prod_{1 \le i < j \le m} (x_j - x_i)^{n_i n_j} \in \mathbb{Z}[x_1, ..., x_m] \ .$$

Since the mapping

$$\mathbb{Z}[x_1, ..., x_m] \to \mathbb{K}, \quad q(x_1, ..., x_m) \to q(a_1, ..., a_m)$$

is a ring homomorphism, $\det V = \det \tilde{V}(a_1, ..., a_m)$ holds true which proves (6.6).

The remaining part of the proof adapts [17]. Set

$$\tilde{p}(x_1, ..., x_m, x) := \prod_{i=1}^{m} (x - x_i)^{n_i} = \sum_{i=0}^{n} \tilde{p}_i(x_1, ..., x_m) x^i \in \mathbb{Z}[x_1, ..., x_m][x] \ ,$$

with suitable $\tilde{p}_i(x_1, ..., x_m) \in \mathbb{Z}[x_1, ..., x_m]$, $\tilde{p}_n := 1$, where x is another indeterminate distinct from $x_1, ..., x_m$. Furthermore, define

$$\tilde{C} := \left[\begin{array}{ccc|c} & & & -\tilde{p}_0(x_1, ..., x_m) \\ \hline 1 & & & -\tilde{p}_1(x_1, ..., x_m) \\ & \ddots & & \vdots \\ & & 1 & -\tilde{p}_{n-1}(x_1, ..., x_m) \end{array}\right] \in \mathbb{Z}[x_1, ..., x_m]^{n,n} \ ,$$

$$v[x] := V(x, n, 1) = (1, x, x^2, ..., x^{n-1})^T =\in \mathbb{Z}[x]^n.$$

Formal differentiation with respect to x will be denoted by ∂_x. Then,

$$V(x_i,n,n_i) = [v[x_i], \partial_x v[x_i], \frac{1}{2}\partial_x^2 v[x_i], ..., \frac{1}{j!}\partial_x^j v[x_i], ..., \frac{1}{(n_i-1)!}\partial_x^{n_i-1} v[x_i]] \ , \tag{6.10}$$

i.e.,

$$V_{[:,j]}(x_i,n,n_i) := \frac{1}{j!}\partial_x^j v[x_i]$$

is the j-th column of $V(x_i,n,n_i)$ for $i \in \{1,...,m\}$, $j \in \{0,...,n_i-1\}$. Now, since $\tilde{C}$ does not depend on x, the vector matrix product of the transpose of such a column (where x_i is replaced by x) and $\tilde{C}$ computes as

$$\begin{aligned} V_{[:,j]}(x,n,n_i)^T\tilde{C} &= \frac{1}{j!}(\partial_x^j v[x])^T\tilde{C} = \frac{1}{j!}\partial_x^j(v[x]^T\tilde{C}) \qquad (6.11)\\ &= \frac{1}{j!}\partial_x^j(x, x^2, ..., x^{n-1}, x^n - \tilde{p}(x_1,...,x_m,x))\\ &= \frac{1}{j!}\left(\partial_x^j(xv[x]^T - \tilde{p}(x_1,...,x_m,x)e_n)\right) \ , \ [e_n := [0,...,0,1] \in \mathbb{Z}^{1,n}]\\ &= \frac{1}{j!}\left(\partial_x^j(xv[x]^T) - \partial_x^j\tilde{p}(x_1,...,x_m,x)e_n\right)\\ &= \frac{1}{j!}\left(\sum_{k=0}^{j}\binom{j}{k}\partial_x^k x \cdot \partial_x^{j-k}v[x]^T - \partial_x^j\tilde{p}(x_1,...,x_m,x)e_n\right)\\ &= \begin{cases} xv[x]^T - \tilde{p}(x_1,...,x_m,x)e_n & \text{, if } j=0,\\ x\frac{1}{j!}\partial_x^j v[x]^T + \frac{1}{(j-1)!}\partial_x^{j-1}v[x]^T & \\ \quad -\frac{1}{j!}\partial_x^j\tilde{p}(x_1,...,x_m,x)e_n & \text{, if } j>0. \end{cases} \end{aligned}$$

Since x_i is an n_i-fold root of $\tilde{p}(x_1,...,x_m,x)$ considered as a polynomial in x, $(x-x_i)^{n_i-j}$ divides $\partial_x^j\tilde{p}(x_1,...,x_m,x)$, i.e.

$$\partial_x^j\tilde{p}(x_1,...,x_m,x) = (x-x_i)^{n_i-j}\cdot s_{i,j}(x_1,...,x_m,x) \tag{6.12}$$

for suitable $s_{i,j}(x_1,...,x_m,x) \in \mathbb{Z}[x_1,...,x_m,x]$. On the other hand,

$$\begin{aligned} \frac{1}{j!}\partial_x^j\tilde{p}(x_1,...,x_m,x) &= \sum_{k=j}^{n}\binom{k}{j}\tilde{p}_k(x_1,...,x_m)x^{k-j}\\ &=: t_j(x_1,...,x_m,x) \in \mathbb{Z}[x_1,...,x_m,x] \ . \end{aligned} \tag{6.13}$$

From (6.12) and (6.13) follows

$$(x-x_i)^{n_i-j}\cdot s_{i,j}(x_1,...,x_m,x) = j!\cdot t_j(x_1,...,x_m,x) \in \mathbb{Z}[x_1,...,x_m,x] \ .$$

Since no $k \in \mathbb{N}\backslash\{1\}$ divides $(x-x_i)$ in $\mathbb{Z}[x_1,...,x_m,x]$, necessarily $j!$ divides $s_{i,j}(x_1,...,x_m,x)$, i.e., $r_{i,j}(x_1,...,x_m,x) := \frac{1}{j!}s_{i,j}(x_1,...,x_m,x) \in \mathbb{Z}[x_1,...,x_m,x]$ and $\frac{1}{j!}\partial_x^j\tilde{p}(x_1,...,x_m,x) = (x-x_i)^{n_i-j}\cdot r_{i,j}(x_1,...,x_m,x)$. Therefore, (6.11) becomes

$$\begin{aligned} V_{[:,j]}(x,n,n_i)^T\tilde{C} &= xV_{[;,j]}(x,n,n_i) + V_{[:,j-1]}(x,n,n_i) \\ &\quad - (x-x_i)^{n_i-j} \cdot r_{i,j}(x_1,...,x_m,x)e_n \\ &= J(x,n_i,n_i)V_{[:,j]}(x,n,n_i)^T \\ &\quad - (x-x_i)^{n_i-j} \cdot r_{i,j}(x_1,...,x_m,x)e_n \end{aligned} \tag{6.14}$$

with $V[:,-1](x,n,n_i) := 0$ and

$$r_{i,0}(x_1,...,x_m,x) := \prod_{\substack{k=0\\k\neq i}}^{m}(x-x_k)^{n_k} = \frac{\tilde{p}(x_1,...,x_m,x)}{(x-x_i)^{n_i}} \ .$$

Since the insertion mapping

$$\mathbb{Z}[x_1,...,x_m,x] \to \mathbb{K},\ q(x_1,...,x_m,x) \mapsto q(a_1,...,a_m,a_i)$$

is a ring homomorphism, (6.14) is transformed by this mapping to

$$V_{[:,j]}(a_i,n,n_i)^T C = J(a_i,n_i,n_i)V_{[:,j]}(a_i,n,n_i)^T. \tag{6.15}$$

Ranging over all $i = 1,...,m$ this yields $V^T C = NV^T$, and since V is regular, as shown in the beginning, we finally derive $V^T CV^{-T} = N$. □

References

[1] A.C. Aitken, *Determinants and Matrices.* §50, Interscience Publishers Inc., 1962.

[2] C. Carathéodory, *Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Funktionen.* Rendiconti del Circolo Matematico di Palermo **32** (1911), 193–217.

[3] S. Ciccariello, A. Cervellino, *Generalization of a theorem of Caratheodory.* Journal of Physics A. Math. Gen. **39** (2006), 14911–14928.

[4] L. Data, S.D. Morgera, *Comments and Corrections on "On the Eigenvectors of Symmetric Toeplitz Matrices".* IEEE Transactions on Acoustics, Speech, and Signal Processing ASSP-**32**, no.2 (1984), 440–441.

[5] P. Delsarte, Y. Genin, Y. Kamp, *Parametric Toeplitz Systems.* Circuits, Systems and Signal Processing **3**, Issue 2 (1984), 207–223.

[6] P. Delsarte, Y. Genin, *Spectral Properties of Finite Toeplitz Matrices.* Springer Lecture Notes in Control and Information Sciences **58** (1984), 194–213.

[7] E. Fischer, *Über das Carathéodory'sche Problem, Potenzreihen mit positivem reellen Teil betreffend.* Rendiconti del Circolo Matematico di Palermo **32** (1911), 240–256.

[8] G. Frobenius, *Ableitung eines Satzes von C. Carathéodory aus einer Formel von Kronecker.* Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (1912), 16–31.

[9] Y. Genin, *A survey of the eigenstructure properties of finite Hermitian Toeplitz matrices.* Integral Equations and Operator Theory **10** (1987), 621–639.

[10] U. Grenander, G. Szegö, *Toeplitz Forms and Their Applications.* University of California Press, 1958.

[11] C. Gueguen, *Linear prediction in the singular case and the stability of eigen models.* Acoustics, Speech and Signal Processing, IEEE International Conference on ICASSP **6** (1981), 881–885.

[12] T.T. Ha and J.A. Gibson, *A Note on the Determinant of a Functional Confluent Vandermonde Matrix and Controllability.* Linear Algebra and its Applications **30** (1980), 69–75.

[13] B. Huppert, *Angewandte Lineare Algebra.* de Gruyter, 1990.

[14] I.S. Iohvidov, *Hankel and Toeplitz Matrices and Forms: Algebraic Theory.* Birkhäuser, 1982.

[15] J. Makhul, *On the Eigenvectors of Symmetric Toeplitz Matrices.* IEEE Transactions on Acoustic, Speech and Signal Processing ASSP **29**, no.4 (1981), 868–872.

[16] V.F. Pisarenko, *The Retrieval of Harmonics from a Covariance Function.* Geophys. J.R. astr. Soc. **33** (1973), 347–366.

[17] J. Brendan Quigley, *Lecture Notes on Differential and Difference equations.* Ch.4, University College Dublin, 2005.

[18] S.S. Reddi, *Eigenvector properties of Toeplitz matrices and their application to spectral analysis of time.* Signal Processing **7** (1984), 45–56.

[19] E. Robinson, *Statistical Communication and Detection.* Griffin, London, 1967.

[20] I. Schur, *Über einen Satz von C. Carathéodory.* Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (1912), 4–15.

Florian Bünger
accepted for publication April 27th, 2014
Institute for Reliable Computing
Hamburg University of Technology
Schwarzenbergstr. 95
D-21073 Hamburg
Germany

e-mail: `florian.buenger@tuhh.de`