Generalization of a theorem of Carathéodory

2006 J. Phys. A: Math. Gen. 39 14911

(http://iopscience.iop.org/0305-4470/39/48/006)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **39** (2006) 14911–14928 doi:10.1088/0305-4470/39/48/006

Generalization of a theorem of Carathéodory

S Ciccariello[1] **and A Cervellino**[2,3]

[1] Dipartimento di Fisica 'G Galilei', Università di Padova & Unità INFM, Via Marzolo 8, I-35131 Padova, Italy
[2] Laboratory for Neutron Scattering, ETH Zurich and PSI Villigen, CH-5232 Villigen PSI, Switzerland

E-mail: salvino.ciccariello@pd.infn.it and antonio.cervellino@psi.ch

Received 25 April 2006, in final form 16 October 2006
Published 15 November 2006
Online at stacks.iop.org/JPhysA/39/14911

Abstract
Carathéodory stated the conditions to be obeyed by n complex numbers $c_1, \ldots, c_n$ in order that they can be written uniquely in the form $c_p = \sum_{j=1}^{m} \rho_j {\epsilon_j}^p$ with $p = 1, \ldots, n$, ϵ_js being different unimodular complex numbers and ρ_js strictly positive numbers. We give the conditions to be obeyed for the former property to hold true if ρ_js are simply required to be real and different from zero. The number of the possible choices of the signs of ρ_js are at most equal to the number of the distinct eigenvalues of the Hermitian Toeplitz matrix whose (i, j) th entry is c_{j-i}, where c_{-p} by definition is the complex conjugate of c_p and $c_0 = 0$. This generalization is relevant to neutron scattering. Its proof is made possible by a lemma stating the necessary and sufficient conditions to be obeyed by the coefficients of a polynomial equation for all the roots to lie on the unit circle. This lemma is an interesting side result of our analysis.

PACS numbers: 61.12.Bt, 02.30.Zz, 89.70.+c, 02.10.Yn, 02.50.Ga

1. Introduction

Carathéodory's theorem [1] states that

Carathéodory's theorem. *n ($\geqslant 1$) complex number $c_1, \ldots, c_n$ and their complex conjugates, respectively denoted by $c_{-1}, \ldots, c_{-n}$, can always and uniquely be written as*

$$c_p = \sum_{j=1}^{m} \rho_j {\epsilon_j}^p, \qquad p = 0, \pm 1, \ldots, \pm n, \tag{1}$$

[3] On leave from Istituto di Cristallografia (CNR-IC), Via Amendola 122/O, I-70126 Bari, Italy.

with $\rho_j \in \mathbb{R}$, $\rho_j > 0$, $\epsilon_j \in \mathbb{C}$, $|\epsilon_j| = 1$, $\epsilon_j \neq \epsilon_k (j \neq k = 1, \ldots, m)$*, and* m *and* c_0 *uniquely determined by* c_j*s with* $j \neq 0$.

This theorem is very useful for the inverse scattering problem [2] as well as for the information theory [3, 4]. In the first case, this appears evident from the following remark. Writing ϵ_js as $\mathrm{e}^{\mathrm{i}2\pi x_j}$ with $0 \leqslant x_j < 1$, c_ps take the form $\sum_{j=1}^{m} \rho_j \mathrm{e}^{\mathrm{i}2\pi x_j p}$ so as they can be interpreted as the scattering amplitudes associated with the 'scattering vector' values $p = 0, \pm 1, \ldots, \pm n$ and generated by m point scatterers with 'charges' $\rho_1, \ldots, \rho_m$ respectively located at $x_1, \ldots, x_m$. One concludes that the theorem of Carathéodory allows us to determine the number of the scattering centres as well as their positions and charges from the knowledge of scattering amplitudes $c_1, \ldots, c_n$. Further, it ensures that the solution of this inverse problem is unique. However, in the case of neutron scattering [5], the charges of the scattering centres no longer have the same sign. Nonetheless, we have recently shown [6] that the so-called algebraic approach to solving the structure of an ideal crystal from its x-ray diffraction pattern [2] can also be applied to the case of neutron scattering. This result suggests that Carathéodory's theorem can be generalized so as to avoid the requirement that the sign of all the scattering charges be positive. This paper shows how to obtain this generalization according to the following plan. In the first of the three subsections of section 2, we review the proof of Carathéodory's theorem given by Grenander and Szegö [7]. In the second, we argue how this proof can be generalized to the case where charges ρ_js are no longer required to be positive and we state the generalized version of Carathéodory's theorem. In the third subsection, we report some numerical illustrations that also consider the case of non-diagonal charges recently considered by Ellys and Lay [8]. The results are so far presented for a practical minded reader. They are rigorously proved in section 3, while section 4 draws our final conclusions. The proof given in section 3 uses a lemma that states the conditions for all the roots of a polynomial equation to have unit modulus. We emphasize that these conditions are necessary and sufficient and only involve the coefficients of the considered polynomial equation in contradiction to presently known theorems [9, 10]. The proof of the lemma is given in appendix A. Appendices B and C are devoted to the derivation of other ancillary results.

2. Carathéodory's theorem and its generalization

In this section, we first review the proof of Carathéodory's theorem given by Grenander and Szegö [7]. Then, we show how this proof can be adapted to the case where the charges are no longer required to be positive. By this adaptation, the given c_js must obey further requirements that once specified lead to the generalized version of Carathéodory's theorem. In the third subsection, we numerically illustrate the above results.

2.1. Grenander and Szegö's proof

This proof of Carathéodory's theorem is based on an enlargement of the set of the n given complex numbers $c_1, \ldots, c_n$ to a set containing $(2n+1)$ complex numbers, still denoted as c_p with $-n \leqslant p \leqslant n$. Here, each c_{-p} with negative index is defined as the complex conjugate of the corresponding c_p, i.e. $c_{-p} \equiv \overline{c_p}$ with $p = 1, \ldots, n$. (Hereafter an overbar will always denote the complex conjugate.). The remaining value c_0, real by assumption, is determined as follows. Consider the $(n+1) \times (n+1)$ matrix (C) defined as

$$(C)_{r,s} \equiv C_{r,s} = c_{s-r}, \qquad r, s = 1, \ldots, (n+1). \tag{2}$$

This matrix is a Hermitian Toeplitz matrix [11] and its diagonal elements are equal to c_0. This value is chosen in such a way to ensure that matrix (C) is *singular* (i.e. $\det(C) = 0$) and the associated bilinear Hermitian form

$$u^{\dagger}(C)u \equiv \sum_{r,s=1}^{n+1} \overline{u_r} C_{r,s} u_s \tag{3}$$

is non-negative definite. (In equation (3) u is an $(n+1)$-dimensional complex vector.) To show that c_0 is unique one proceeds as follows [3]. Let $(\hat{C})$ denote the $(n+1) \times (n+1)$ matrix that has its non-diagonal elements equal to the correspondent elements of (C) and its diagonal ones equal to zero, i.e., with $r, s = 1, \ldots, (n+1)$,

$$\hat{C}_{r,s} \equiv \begin{cases} c_{s-r}, & \text{if } r \neq s, \\ 0, & \text{if } r = s. \end{cases} \tag{4}$$

This matrix is Hermitian. Then its eigenvalues $(\chi_j,\ j = 1, \ldots, n+1)$ are real and can be labelled so as to have $\chi_1 \leqslant \cdots \leqslant \chi_{n+1}$. Further, they are such that $\sum_{j=1}^{n+1} \chi_j = 0$ because the trace of $(\hat{C})$ is zero. Hence, $\chi_1 < 0$ and $\chi_{n+1} > 0$. One immediately realizes that matrix (C) is obtained by setting $c_0 = -\chi_1 > 0$ so that $(C) = (\hat{C}) - \chi_1(I)$, (I) being the unit matrix. In fact, the matrix $(\hat{C} - \chi_1 I)$ is a Hermitian Toeplitz matrix with its diagonal elements equal to $(-\chi_1)$. The secular equation of this matrix is

$$\det((C) - z(I)) = \det((\hat{C}) - \chi_1(I) - z(I)) = \det((\hat{C}) - (\chi_1 + z)(I)) = 0.$$

This equation is the same equation that determines the eigenvalues of $(\hat{C})$ if, instead of variable z, we use the shifted variable $z + \chi_1$. Thus, the eigenvalues of (C) are $0 = (\chi_1 - \chi_1) \leqslant (\chi_2 - \chi_1) \leqslant \cdots \leqslant (\chi_{n+1} - \chi_1)$ and matrix (C) is non-negative definite. (C) is uniquely determined by $c_1, \ldots, c_n$ because the remaining quantity c_0 depends on $c_1, \ldots, c_n$. Further, let μ_1 $(\geqslant 1)$ denote the multiplicity of eigenvalue χ_1 of $(\hat{C})$, then the rank of (C) is $(n + 1 - \mu_1)$ and this value gives the number of the addends present in equation (1), i.e. $m = (n + 1 - \mu_1)$. Exploiting the non-negative definiteness and the singularity of (C), Grenander and Szegö showed that (I) the $(2n + 1)$ equalities stated in (1) are satisfied; (II) the m complex numbers ϵ_j present in (1) are the roots of the following polynomial equation of degree m:

$$P_m(z) = D_m{}^{-1} \det \begin{pmatrix} c_0 & c_1 & \cdots & c_{m-1} & c_m \\ c_{-1} & c_0 & \cdots & c_{m-2} & c_{m-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{-m+1} & c_{-m+2} & \cdots & c_0 & c_1 \\ 1 & z & \cdots & z^{m-1} & z^m \end{pmatrix} = 0, \tag{5}$$

where D_m denotes the determinant of the left principal minor contained in the first m rows of (C) (clearly $D_m > 0$ because (C) is non-negative definite); (III) the roots are distinct and unimodular and (IV) ρ_js are strictly positive and given by

$$\rho_j = \frac{1}{P_m{}'(\epsilon_j)} \sum_{p=0}^{m-1} \beta_{j,p} c_p, \tag{6}$$

where the prime denotes the derivative and $\beta_{j,p}$s are the coefficients of the following polynomial:

$$P_m(z)/(z - \epsilon_j) \equiv \sum_{p=0}^{m-1} \beta_{j,p} z^p. \tag{7}$$

The important role played by equation (5) appears evident from (II)–(IV). For this reason, equation (5) will be called *resolvent equation* in the following.

The above procedure shows that it is always possible to associate a singular, unique, non-negative definite, Hermitian, Toeplitz matrix (C) of order $(n+1)$ to n $(\geqslant 1)$ given complex numbers $c_1, \ldots, c_n$ because these uniquely determine c_0. Denoting the rank of (C) by m, according to Carathéodory's theorem, all the elements of (C) can be written as specified in (1). After introducing the $m \times (n+1)$ Vandermonde $(\mathcal{V})$ matrix

$$\mathcal{V}_{i,j} \equiv \epsilon_i^{j-1}, \qquad |\epsilon_j| = 1, \quad i = 1, \ldots, m, \quad j = 1, \ldots, (n+1), \tag{8}$$

and the $m \times m$ positive-definite diagonal matrix $(\mathcal{Q})$

$$\mathcal{Q}_{i,j} \equiv \rho_i \delta_{i,j}, \qquad i, j = 1, \ldots, m, \tag{9}$$

from equations (2) and (1) follows that

$$\begin{aligned} C_{r,s} = c_{s-r} &= \sum_{i=1}^{m} \rho_i {\epsilon_i}^{s-r} = \sum_{i,j=1}^{m} {\epsilon_i}^{-r+1} \rho_i \delta_{i,j} {\epsilon_j}^{s-1} \\ &= \sum_{i,j=1}^{m} \overline{{\epsilon_i}^{r-1}} \rho_i \delta_{i,j} {\epsilon_j}^{s-1} = \sum_{i,j=1}^{m} \overline{\mathcal{V}_{i,r}} \mathcal{Q}_{i,j} \mathcal{V}_{j,s}. \end{aligned} \tag{10}$$

Hence, (C) can be written as

$$(C) = (\mathcal{V})^{\dagger} (\mathcal{Q}) (\mathcal{V}), \tag{11}$$

where $(\mathcal{V})^{\dagger}$ denotes the matrix Hermitian conjugate of $(\mathcal{V})$. This remark leads to the following:

Corollary of Carathéodory's theorem. *A non-negative definite, Hermitian Toeplitz matrix* (C) *of order* $n+1$ *and rank* m $(\leqslant n)$ *uniquely factorizes as in (11) in terms of matrices* $(\mathcal{V})$ *and* $(\mathcal{Q})$*, defined by (8) and (9).*

2.2. Formulation of the generalized Carathéodory theorem

We already stressed that, starting from n given complex numbers $c_1, \ldots, c_n$, the above choice of c_0 is the only one that yields a singular, non-negative definite, Hermitian Toeplitz matrix (C) of order $n+1$. It is evident that, if we choose $c_0 = -\chi_{n+1}$, the resulting matrix defined by equations (2) (for the moment denoted as (C_{n+1})) is singular and non-positive definite, while $-(C_{n+1})$ is singular and non-negative definite. Then, to the latter we can apply Grenander and Szegö's analysis and conclude

$$c_0 = -\chi_{n+1} = \sum_{j=1}^{m} {\rho'}_j \qquad \text{and} \qquad c_p = \sum_{j=1}^{m} {\rho'}_j {{\epsilon'}_j}^p, \qquad p = \pm 1, \ldots, \pm n$$

where ρ'_js are now strictly negative numbers, $m = (n+1-\mu_{n+1})$ with μ_{n+1} equal to the multiplicity of χ_{n+1}, and ϵ'_js distinct unimodular complex numbers equal to the roots of the resolvent equation generated by $-(C_{n+1})$. Of course, this result represents only a trivial generalization of Carathéodory's theorem. However, further choices of c_0 are still possible and the following considerations will show that some of these lead to a non-trivial generalization of Carathéodory's theorem. First, it is convenient to determine all the *distinct* Toeplitz matrices that can be obtained from (2) by choosing c_0 equal to the opposite of each eigenvalue of matrix $(\hat{C})$. To this aim, we shall denote the *distinct* eigenvalues of $(\hat{C})$ by $\hat{\chi}_1 < \cdots < \hat{\chi}_\nu$, their multiplicities by $\mu_1, \mu_2, \ldots, \mu_\nu$, and the number of the positive (negative) $\hat{\chi}_l$s by $p_{\hat{\chi}}$ $(n_{\hat{\chi}})$. We clearly have $\sum_{j=1}^{\nu} \mu_j = (n+1)$, and $\nu = (n_{\hat{\chi}} + p_{\hat{\chi}} + 1)$ or $\nu = (n_{\hat{\chi}} + p_{\hat{\chi}})$ depending on whether one of $\hat{\chi}_l$s is equal to zero or not. Consider now the ν matrices (C_l) of order $(n+1)$ defined, for $l = 1, \ldots, \nu$, as

$$(C_l)_{r,s} \equiv C_{l;r,s} \equiv \begin{cases} c_{s-r}, & \text{if } r \neq s \text{ and } r, s = 1, \ldots, n+1, \\ c_{l,0} = -\hat{\chi}_l, & \text{if } r = s = 1, \ldots, n+1. \end{cases} \tag{12}$$

The eigenvalues of matrix (C_l) are $(\chi_1 - \hat{\chi}_l), \ldots, (\chi_{n+1} - \hat{\chi}_l)$. Thus, (C_l) respectively has

$$\mu_l\ (\geqslant 1), \qquad N_{l,-} \equiv \sum_{i=1}^{l-1} \mu_i \qquad \text{and} \qquad N_{l,+} \equiv \sum_{i=l+1}^{\nu} \mu_i \tag{13}$$

null, negative and positive eigenvalues. Its rank is

$$m_l = (n + 1 - \mu_l) = N_{l,+} + N_{l,-}. \tag{14}$$

All the matrices defined by (12) are distinct, because they differ for the diagonal elements. Since $c_{l,0}$s and $c_{-1}, \ldots, c_{-n}$ are uniquely determined by $c_1, \ldots, c_n$, we have the important

Property A. *The ν matrices defined by (12) are the only distinct, singular, Hermitian, Toeplitz matrices of order $(n+1)$ that are generated by the n complex numbers $c_1, \ldots, c_n$.*

We always have $\nu \geqslant 2$ because the trace of $(\hat{C})$ is zero. Since matrices (C_1) and (C_ν) respectively are non-positive and non-negative defined, the sought for generalization of Carathéodory's theorem only concerns (C_l)s defined by (12) with $1 < l < \nu$. These matrices exist only if $\nu > 2$. Assuming this condition fulfilled, in the following we shall mainly refer to the last matrices omitting, whenever possible, index l for notational simplicity.

We face now the question: what are the conditions that must be obeyed by each of these matrices for the relevant complex numbers $c_{l,0}, c_{\pm 1}, \ldots, c_{\pm n}$ to be written in the form (1) relaxing the requirement that all the involved ρ_js are strictly positive?

The situation looks similar to that found in the case of Carathéodory's theorem and one might conclude that for each of the ν matrices defined by (12) the resolvent equation can be obtained from equation (5) using there the considered (C_l) matrix. Unfortunately, this matrix is neither non-negative nor non-positive definite so that the condition $D_{m_l} \neq 0$ is not ensured by the property that (C_l) has rank m. For this to happen one must require that the determinant of the minor of (C_l), formed by the first m rows and m columns of (C_l), be different from zero. This is the first constraint that must be obeyed by the considered (C_l). It can be formalized as

Constraint (i). *Each matrix (C_l) defined by (12) must be such that its rank m_l coincides with its principal rank.*

We define *principal rank* of a general matrix (A) as the highest of the rank values of all the possible *strictly principal* (left) minors of (A). Moreover, a *strictly principal* (left) minor (of order m) of (A) is a minor formed by m subsequent rows and m subsequent columns of (A), the rows and the column having the same index values, e.g. $p+1, p+2, \ldots, p+m$ with $0 \leqslant p \leqslant (n-m)$ and n equal to the order of (A).

Once the considered (C_l) obeys constraint (i), the existence of the polynomial $P_{m_l}(z)$ is ensured. But the polynomial is generated now by a matrix that is indefinite so that Grenander and Szegö's procedure cannot be used to conclude that the roots of the corresponding equation $P_{m_l}(z) = 0$ are unimodular and distinct. Thus, we must require that the coefficients of this polynomial equation obey a set of constraints in order that these properties of the roots are satisfied.

The constraints are specified by lemma A that will be reported after some preliminary definitions. First, we observe that a polynomial equation of degree N can generally be written as

$$P_N(z) \equiv \prod_{j=1}^{N}(z-\epsilon_j) = \sum_{n=0}^{N} a_n z^n = 0, \tag{15}$$

its coefficients a_n being related to the roots ϵ_j as

$$a_n \equiv (-)^{N-n} \sum_{1\leqslant j_1<j_2<\cdots<j_{N-n}\leqslant N} \epsilon_{j_1}\epsilon_{j_2}\cdots\epsilon_{j_{N-n}}, \qquad n=0,1,\ldots,N-1 \tag{16}$$

and $a_N = 1$. The following symmetric functions of the roots [12, 13]

$$\sigma_p \equiv \sum_{j=1}^{N} \epsilon_j{}^p, \qquad p = 0, 1, \ldots \tag{17}$$

can be expressed in terms of a_ns by solving the following equations:

$$\begin{array}{lllllll}
 & & & & & \sigma_0 = & N \\
 & & & & & a_N\sigma_1 = & -a_{N-1} \\
 & & & a_N\sigma_2 & & +a_{N-1}\sigma_1 = & -2a_{N-2} \\
 & & a_N\sigma_3 & +a_{N-1}\sigma_2 & & +a_{N-2}\sigma_1 = & -3a_{N-3} \\
\ldots & \ldots & \ldots & \ldots & \ldots & = & \ldots \\
a_N\sigma_{N-1}+ & \cdots & +a_4\sigma_3 & +a_3\sigma_2 & & +a_2\sigma_1 = & -(N-1)a_1 \\
a_N\sigma_{p+N}+ & \cdots & +a_2\sigma_{p+2} & +a_1\sigma_{p+1} & & +a_0\sigma_p = & 0, \qquad p=0,1,\ldots.
\end{array} \tag{18}$$

For negative ps, we set

$$\sigma_{-p} \equiv \overline{\sigma_p}, \qquad p = 1, \ldots, N, \tag{19}$$

and we introduce the $(N+1)\times(N+1)$ Hermitian Toeplitz matrix $(\mathcal{S})$ defined as

$$\mathcal{S}_{i,j} \equiv \sigma_{j-i}, \qquad i, j = 1, \ldots, N+1. \tag{20}$$

The lemma that specifies the constraints on the coefficients of polynomial equation (15) for this to have distinct and unimodular roots is the following:

Lemma A. *A polynomial equation of degree N (see equation (15)) has distinct and unimodular roots if and only if (a) its coefficients obey the conditions*

$$a_0 \neq 0 \qquad and \qquad \overline{a_m} = a_{N-m}/a_0 \qquad for \quad m = 0, \ldots, N, \tag{21}$$

and (b) Toeplitz matrix $(\mathcal{S})$, defined by (20), is non-negative definite and has rank N.

The lemma will be proved in appendix A. We can now state the

Generalized Carathéodory theorem. *Let (C_l) be one of the matrices defined by (12). If (C_l) is such that (i) its rank m_l coincides with its principal rank and (ii) the coefficients of the associated resolvent equation $P_{m_l}(z) = 0$ obey conditions (a) and (b) reported in lemma A, then the complex numbers $c_{l,0}, c_{\pm 1}, \ldots, c_{\pm n}$ that define (C_l) can be written in the form*

$$C_{l;r,r} \equiv c_{l,0} = \sum_{j=1}^{m} \rho_{l,j}, \qquad C_{l;r,s} \equiv c_{s-r} = c_{\pm p} = \sum_{j=1}^{m} \rho_{l,j}\epsilon_{l,j}^{\pm p}, \qquad p = 1, \ldots, n \tag{22}$$

$$\begin{array}{l}
\rho_j \in \mathbb{R}, \quad \rho_j \neq 0, \quad \epsilon_j \in \mathbb{C}, \quad |\epsilon_j| = 1, \\
\epsilon_j \neq \epsilon_i \quad \text{if} \quad j \neq i, \quad j, i = 1, \ldots, m,
\end{array} \tag{23}$$

with $m = m_l$, $N_{l,+}$ positive $\rho_{l,j}$s and $N_{l,-}$ negative $\rho_{l,j}$s m_l being defined by (14) and $N_{l,+}$ and $N_{l,-}$ by (13).

This generalized theorem, similarly to Carathéodory's, leads to the following:

Corollary of the generalized Carathéodory theorem. *A Hermitian Toeplitz matrix* (C) *of order* $n+1$ *and rank* $m \leqslant n$ *uniquely factorizes as*

$$(C) = (\mathcal{V})^{\dagger}(\mathcal{Q})(\mathcal{V}), \tag{24}$$

where $(\mathcal{V})$ *is defined by (8) and* $(\mathcal{Q})$ *is an* $m \times m$ *diagonal matrix with* $Q_{j,j} = \rho_j$, *iff* (C) *obeys conditions (i) and (ii) reported in the above theorem.*

2.3. Some numerical illustrations

Before proving these theorems, we report three numerical illustrations of the reported results as follows:

(1) The first refers to a case where it is impossible to express a set of c_js in terms of positive and negative ρ_js. Let $c_1 = 0, c_2 = 0, c_3 = 1$. The distinct eigenvalues of the associated matrix $(\hat{C})$ are -1, 0 and 1, with respective multiplicities 1, 2 and 1. The matrix (C_1), obtained by setting $c_{1,0} = 1$, is non-negative definite, has rank 3 and eigenvalues equal to 0, 1, 1, 2. Equations (22) and(23) with $l = 1$ are fulfilled with $\epsilon_{1,1} = 1, \epsilon_{1,2} = -\mathrm{e}^{\mathrm{i}\pi/3}, \epsilon_{1,3} = \mathrm{e}^{\mathrm{i}2\pi/3}$ and $\rho_{1,1} = \rho_{1,2} = \rho_{1,3} = 1/3$. The matrix (C_3), obtained by setting $c_{3,0} = -1$, is non-positive definite with rank equal to 3 and eigenvalues equal to $-2, -1, -1, 0$. The solutions are $\epsilon_{3,1} = -1, \epsilon_{3,2} = \mathrm{e}^{\mathrm{i}\pi/3}, \epsilon_{3,3} = -\mathrm{e}^{\mathrm{i}2\pi/3}$ and $\rho_{3,1} = \rho_{3,2} = \rho_{3,3} = -1/3$. Finally, matrix (C_2) is obtained by setting $c_{2,0} = 0$ and coincides with $(\hat{C})$. It is non-definite, has rank equal to 2 and principal rank equal to 0. Thus, condition (i) of the generalized Carathéodory theorem is not satisfied and it is impossible to write the given c_ps, namely 0, 0, 0 and 1 in the form (22), (23) with $m = 2$ as one can easily check.

(2) The second example considers the case where $c_1 = 1, c_2 = 0, c_3 = 1$. The distinct eigenvalues of the associated matrix $(\hat{C})$ are $-2, 0, 2$ with respective multiplicities 1, 2, 1. Setting $c_{2,0} = 0$, the resulting (C_2) matrix coincides with $(\hat{C})$. It is non-definite, its rank is 2 and equals its principal rank value. The generalized Carathéodory's theorem applies because the resolvent equation is $P_2(z) = z^2 - 1 = 0$. The solution is $\epsilon_{2,1} = 1$, $\epsilon_{2,2} = -1, \rho_{2,1} = 1/2$ and $\rho_{2,2} = -1/2$. With these values one easily checks that $c_{2,0} = \rho_{2,1} + \rho_{2,2} = 0$, $c_1 = \rho_{2,1}\epsilon_{2,1} + \rho_{2,2}\epsilon_{2,2} = 1$, $c_2 = \rho_{2,1}{\epsilon_{2,1}}^2 + \rho_{2,2}{\epsilon_{2,2}}^2 = 0$ and $c_3 = \rho_{2,1}{\epsilon_{2,1}}^3 + \rho_{2,2}{\epsilon_{2,2}}^3 = 1$.

   Setting $c_{1,0} = 2$, the resulting (C_1) is non-negative definite with rank 3. The solution is $\epsilon_{1,1} = 1, \epsilon_{1,2} = \mathrm{i}, \epsilon_{1,3} = -\mathrm{i}, \rho_{1,1} = 1, \rho_{1,2} = 1/2$ and $\rho_{1,3} = 1/2$. The last choice $c_{3,0} = -2$ defines a non-positive definite (C_3) matrix with rank equal to 3 with solution $\epsilon_{3,1} = -1, \epsilon_{3,2} = i, \epsilon_{3,3} = -i, \rho_{3,1} = -1, \rho_{3,2} = -1/2$ and $\rho_{3,3} = -1/2$.

(3) The last example corresponds to having $c_0 = \delta_0, c_1 = \delta_0 + \mathrm{i}\delta_1\mathrm{e}^{\mathrm{i}\varphi}$ and $c_2 = (\delta_0 + 2\mathrm{i}\delta_1)\mathrm{e}^{2\mathrm{i}\varphi}$ with δ_0, δ_1 and φ real numbers. The eigenvalues are 0 and $(3\delta_0/2)\left(1 \pm \sqrt{1 + 8\delta_1^2/\left(3\delta_0^2\right)}\right)$. Moreover, we assume that the values of δ_0 and δ_1 are such that the three eigenvalues are distinct. Hermitian matrix (C) is indefinite. It has rank and principal rank equal to 2. The resolvent equation is $P_2(z) = z^2 - 2\mathrm{e}^{\mathrm{i}\varphi}z + \mathrm{e}^{2\mathrm{i}\varphi} = 0$. Its coefficients $a_0 = \mathrm{e}^{2\mathrm{i}\varphi}$, $a_1 = -2\mathrm{e}^{\mathrm{i}\varphi}$ and $a_2 = 1$ obey conditions (21). From equations (18) and (19) follows that $\sigma_0 = 2, \sigma_1 = 2\mathrm{e}^{\mathrm{i}\varphi}$ and $\sigma_2 = 2\mathrm{e}^{2\mathrm{i}\varphi}$ so that resulting matrix $(\mathcal{S})$ has rank 1 instead of 2. (In fact, the reported equation has a unimodular root with multiplicity 2 since it can be written as $P(z) = (z - \mathrm{e}^{\mathrm{i}\varphi})^2 = 0$.) Thus, condition (ii) of the generalized Carathéodory theorem is violated and (C) cannot be written as $(\mathcal{V})^{\dagger}(Q)(\mathcal{V})$.

In this example, however, all the roots are unimodular. Then, according to theorem 3.4 of Ellis and Lay [8], (C) factorizes as

$$(C) = (\mathcal{V}')^{\dagger}(\Delta)(\mathcal{V}') \tag{25}$$

where $(\mathcal{V}')$ is a generalized confluent Vandermonde matrix generated by the distinct roots of the resolvent equation and (Δ) is a block-diagonal matrix with reversed upper triangular blocks. In fact, in our case, one finds that

$$\begin{aligned}\begin{pmatrix} \delta_0, & (\delta_0+\mathrm{i}\delta_1)\,\mathrm{e}^{\mathrm{i}\varphi}, & (\delta_0+2\mathrm{i}\delta_1)\,\mathrm{e}^{2\mathrm{i}\varphi} \\ (\delta_0-\mathrm{i}\delta_1)\,\mathrm{e}^{-\mathrm{i}\varphi}, & \delta_0, & (\delta_0+\mathrm{i}\delta_1)\,\mathrm{e}^{\mathrm{i}\varphi} \\ (\delta_0-2\mathrm{i}\delta_1)\,\mathrm{e}^{-2\mathrm{i}\varphi}, & (\delta_0-\mathrm{i}\delta_1)\,\mathrm{e}^{-\mathrm{i}\varphi}, & \delta_0 \end{pmatrix} \\ = \begin{pmatrix} 1, & 0 \\ \mathrm{e}^{-\mathrm{i}\varphi}, & \mathrm{e}^{-\mathrm{i}\varphi} \\ \mathrm{e}^{-2\mathrm{i}\varphi}, & 2\mathrm{e}^{-2\mathrm{i}\varphi} \end{pmatrix} \begin{pmatrix} \delta_0, & \mathrm{i}\delta_1 \\ -\mathrm{i}\delta_1, & 0 \end{pmatrix} \begin{pmatrix} 1, & \mathrm{e}^{\mathrm{i}\varphi}, & \mathrm{e}^{2\mathrm{i}\varphi} \\ 0, & \mathrm{e}^{\mathrm{i}\varphi}, & 2\mathrm{e}^{2\mathrm{i}\varphi} \end{pmatrix}.\end{aligned} \tag{26}$$

In proving factorization (25), Ellis and Lay required neither the hermiticity of Toeplitz matrix (C) nor semidefiniteness of (C). This last property amounts to relaxing the condition that ρ_js present in (1) are strictly positive. However, in proving this more general factorization, Ellis and Lay explicitly assumed that all the roots of the resolvent equation are unimodular without stating the conditions for the unimodularity to occur. Hence, the more general factorization (25) becomes fully proved if one adds lemma A to the considerations reported in [8]. We note that generalization (25) does not appear to be physically relevant because it is not yet clear the physical meaning of a 'charge' matrix (Δ) strictly non-diagonal.

3. Proof of the generalized Carathéodory theorem

We recall that this theorem states that the complex numbers $c_{l,0}, c_{\pm 1}, \ldots, c_{\pm n}$ that define (C_l), one of the matrices defined by (12), can be written in the form (22) and (23) with $m = m_l$, $N_{l,+}$ positive $\rho_{l,j}$s and $N_{l,-}$ negative $\rho_{l,j}$s, if (C_l) is such that (i) its rank m_l coincides with its principal rank and (ii) the coefficients of the resolvent polynomial $P_{m_l}(z) = 0$ obey conditions (a) and (b) reported in lemma A.

As already emphasized, its proof must be achieved by a procedure different from Grenander and Szegö's because matrices (C_l), defined in (12), are indefinite if $1 < l < \nu$. Our attention will now focus on one of these matrices and we shall omit index l for notational simplicity.

3.1. Necessity of condition (i)

We first show the necessity of condition (i) assuming that the coefficients of (C) can be written as in (22) and (23). In such a case, the $m \times (n+1)$ rectangular Vandermonde matrix $(\mathcal{V})$ defined by (8) and the $m \times m$ diagonal matrix $(\mathcal{Q})$ defined by (9) exist, equation (10) applies and the considered (C) can be written as in (11). Consider now the $p \times p$ minor of (C) formed by the rows $r_1 < \cdots < r_p$ and the columns $s_1 < \cdots < s_p$. By Bezout's theorem [14, 15] one finds that

$$\det\left(C^{\,s_1,\ldots,s_p}_{\,r_1,\ldots,r_p}\right) = \sum_{1\leqslant i_1<i_2<\cdots<i_p\leqslant m} \left[\prod_{j=1}^{p} \rho_{i_j}\right] \det\left(V^{\,s_1,\ldots,s_p}_{\,i_1,\ldots,i_p}\right) \det\left(\overline{V}^{\,r_1,\ldots,r_p}_{\,i_1,\ldots,i_p}\right) \tag{27}$$

where, adopting Gantmacher's notation [14], the lower and upper indices inside each determinant symbol denote the rows and the columns of the considered minors of (C), $(\mathcal{V})$

and $(\overline{\mathcal{V}})$. This expression makes it clear that the determinant of any minor of (C) of order $p > m$ is equal to zero because the order of (Q) is m.[4] For this reason, the rank of (C) cannot exceed m, i.e. the number of the different ϵ_js. At the same time, if $p = m$, the determinant of each $(m \times m)$ strictly principal minor $\left(C\,{}^{q+1,\ldots,q+m}_{q+1,\ldots,q+m}\right)$ of (C), with $0 \leqslant q \leqslant n+1-p$, is

$$\left[\prod_{j=1}^{m} \rho_j\right] \left| \prod_{1\leqslant i\leqslant j\leqslant m} (\epsilon_j - \epsilon_i) \right|^2. \tag{28}$$

In fact, if $p = m$, the sum present in (27) involves a single term and, due to (8), in $\det\left(V\,{}^{q+1,\ldots,q+m}_{1,\ldots,m}\right)$ we can factorize ϵ_1^q in the first row, ϵ_2^q in the second row and so on. The remaining matrix is a Vandermonde matrix so that

$$\det\left(V\,{}^{q+1,\ldots,q+m}_{1,\ldots,m}\right) = \prod_{l=1}^{m} \epsilon_l{}^q \prod_{1\leqslant i\leqslant j\leqslant m} (\epsilon_j - \epsilon_i).$$

In the same way, one shows that

$$\det\left(\bar{V}\,{}^{q+1,\ldots,q+m}_{1,\ldots,m}\right) = \prod_{l=1}^{m} \overline{\epsilon_l}{}^q \prod_{1\leqslant i\leqslant j\leqslant m} (\overline{\epsilon_j} - \overline{\epsilon_i}).$$

Finally, the unimodularity of ϵ_js yields expression (28) that is different from zero whatever q. In this way, the rank of (C) is equal to its principal rank and the necessity of condition (i) is proved[5]. This result and property (c) of Toeplitz matrices (see appendix C) imply that the first m rows or columns of (C) are linearly independent.

3.2. Necessity of condition (ii)

To prove the necessity of (ii), we must prove that the mth degree polynomial equation $P_m(z) = 0$ with roots equal to $\epsilon_1, \ldots, \epsilon_m$, distinct and unimodular by assumption, coincides with the resolvent equation defined by (5) and generated by the considered matrix (C). We shall adopt the same notation of equations (15) and (16), with the obvious change $N \to m$. Then, $a_m = 1$ and, due to the unimodularity of ϵ_js, $|a_0| = 1$. In order to determine remaining coefficients $a_0, \ldots, a_{m-1}$ of $P_m(z)$, we observe that, from the condition $P_m(\epsilon_i) = 0$, follows that $\epsilon_i^m = -\sum_{l=0}^{m-1} a_l \epsilon_i^l$. After multiplying the latter by $\epsilon_i{}^j$, with $j \in \mathbb{Z}$, and setting $q = m + j$, one finds that

$$\epsilon_i^q = -\sum_{l=0}^{m-1} a_l \epsilon_i^{l+q-m}, \qquad q \in \mathbb{Z} \quad \text{and} \quad i = 1, \ldots, m. \tag{29}$$

The substitution of these relations in (22), where for simplicity we put $c_0 = c_{l,0}$, yields

$$c_{s-r} = -\sum_{i=1}^{m} \rho_i \sum_{l=0}^{m-1} a_l \epsilon_i^{l+s-r-m} = -\sum_{l=0}^{m-1} a_l c_{s-r-(m-l)}. \tag{30}$$

[4] This is also evident from (24): by construction the rows of $(\mathcal{V})$ span the range space of (C) and (C) has rank at most equal to m.

[5] For the case of Hermitian non-negative Toeplitz matrices considered by Grenander and Szegö, their proof shows that the principal rank of these matrices is equal to their rank. An alternative proof was given by Goedkoop [16] introducing a finite-dimensional Hilbert space (see also [2]). Thus, for this kind of matrices, one has the interesting property that the rank is obtained by considering the strictly principal minors of increasing order till finding a singular minor. If the latter's order is $m + 1$ the rank of the matrix is m.

Taking $(s-r)=1,2,\ldots,m$, one obtains the following system of linear equations:

$$\begin{array}{llllllll} c_0a_0 & +c_1a_1 & +c_2a_2+ & \cdots & +c_{m-1}a_{m-1} & = & -c_m \\ c_{-1}a_0 & +c_0a_1 & +c_1a_2+ & \cdots & +c_{m-2}a_{m-1} & = & -c_{m-1} \\ . & . & . & . & \cdots & . & . \\ c_{-m+1}a_0 & +c_{-m+2}a_1 & +c_{-m+3}a_2+ & \cdots & +c_0a_{m-1} & = & -c_1. \end{array} \tag{31}$$

This uniquely determines coefficients $a_0,\ldots,a_{m-1}$ of the polynomial because the determinant of the coefficients is different from zero (see the last remark in section 3.1). Actually, the determinant coincides with D_m defined below equation (5). In this way we have shown that $P_m(z)=0$, the equation determined by ϵ_js, coincides with Grenander and Szegö's equation (5), i.e. with the resolvent equation generated by matrix (C), no longer required to be non-negative or non-positive definite. Since the roots of $P_m(z)=0$ are distinct and unimodular, the coefficients of this equation obey conditions (a) and (b) of lemma A. In this way, the necessity of (ii) is proved.

To complete the discussion, we report the explicit expressions of coefficients a_l in terms of the appropriate minors of (C). In fact, the solution of (31) is

$$a_l=(-)^{m-l}\det\left(C^{1,\ldots,l,l+2,\ldots,m+1}_{1,\ldots,m}\right)\big/D_m,\qquad l=0,\ldots,m \tag{32}$$

$$D_m\equiv\det\left(C^{\,1,\ldots,m}_{1,\ldots,m}\right)\neq 0, \tag{33}$$

where in (32) we let l assume the value m since it yields the known relation $a_m=1$. Furthermore, using (C.4) (with $n=m+1$, $p=1$ and $q=2$), from equation (32) with $l=0$ one finds

$$a_0=(-)^m\det\left(C^{2,\ldots,m+1}_{1,\ldots,m}\right)\big/D_m=(-)^m\mathrm{e}^{\mathrm{i}\theta_1}. \tag{34}$$

Hence, a_0 is unimodular as already evident from (16).

3.3. Sufficiency of (i) and (ii)

If the rank m of the considered matrix (C) coincides with its principal rank value, then equation (5) yields the resolvent equation $P_m(z)=0$. The coefficients of this equation are given by (32)–(34). They obey conditions (21) as shown in appendix B. Hence, condition (a) of lemma A is already fulfilled. From the previous coefficients one evaluates matrix $(\mathcal{S})$ by (17)–(20) (with $N\to m$). If matrix $(\mathcal{S})$ turns out to be non-negative and has rank m, then $P_m(z)=0$ has m distinct unimodular roots. After solving this equation, ρ_js can be determined by equations (6) and (7) since these also apply in the case of non-definite (C). Alternatively, ρ_js can be determined solving the system of m linear equations

$$\sum_{j=1}^{m}{\epsilon_j}^p\rho_j=c_p,\qquad p=0,\ldots,(m-1) \tag{35}$$

that follow from (22). (For notational simplicity, we still omit index l present in the definition of c_0.) These equations can also be written as

$$\sum_{j=1}^{m}\mathcal{V}^T_{p+1,j}\rho_j=c_p,\qquad p=0,\ldots,(m-1) \tag{36}$$

where $(\mathcal{V}^T)$ is the transpose of the $m \times m$ upper left principal minor of matrix (8). The formal solution of (36) is

$$\rho_j = \sum_{p=0}^{m-1} (\mathcal{V}^T)^{-1}{}_{j,p+1} c_p, \qquad j = 1, \ldots, m, \tag{37}$$

since $(\mathcal{V}^T)$ is certainly non-singular.

Finally, it must be proved that the numbers of ρ_js that turn out to be positive or negative are, respectively, equal to N_+ and N_- (again we omit index l). To this aim, consider the Hermitian bilinear form

$$\mathbb{C}_2[u] \equiv \sum_{r,s=1}^{n+1} \overline{u_r} C_{r,s} u_s, \qquad u_s \in \mathbb{C}. \tag{38}$$

This is immediately expressed by equation (24) in terms of the diagonal form

$$\mathbb{C}_2[u] = \sum_{p=1}^{m} \overline{v_p[u]} \rho_p v_p[u] \tag{39}$$

where

$$v_p[u] \equiv \sum_{s=1}^{n+1} \mathcal{V}_{p,s} u_s, \qquad p = 1, \ldots, m. \tag{40}$$

At the same time, since (C) is Hermitian, it can be diagonalized by a unitary transformation (U) and written as

$$(C) = (U)^\dagger (\chi)(U) \tag{41}$$

where

$$(\chi)_{r,s} = (\chi_r - \hat{\chi}_l)\delta_{r,s}, \qquad r, s = 1, \ldots, (n+1) \tag{42}$$

with χ_rs equal to the eigenvalues of $(\hat{C})$. As discussed at the beginning of this section, μ_l of $(\chi_r - \hat{\chi}_l)$s are equal to zero, $N_{l,+}$ are positive and $N_{l,-}$ negative. Therefore, we can first compact $(U^\dagger)$ by eliminating μ_l columns whose index corresponds to the rows of (χ) containing the zero eigenvalues and then we compact (χ) eliminating the rows and the columns containing the zero eigenvalues. Hereinafter (U) shall be an $m \times (n+1)$ rectangular matrix with orthonormal rows and (χ) an $m \times m$ diagonal and non-singular matrix. We set

$$w_p \equiv \sum_{s=1}^{n+1} U_{p,s} u_s, \qquad p = 1, \ldots, m \tag{43}$$

and consider w_ps as the arbitrary independent variables. Using (24) we can write

$$(U)^\dagger (\chi)(U) = (C) = (\mathcal{V})^\dagger (Q)(\mathcal{V}).$$

The row spaces of (U) and $(\mathcal{V})$ necessarily coincide with the (U) and $(\mathcal{V})$ right image spaces that in turn coincide with the eigenspace of (C) associated with the eigenvalue zero. There exists then a non-singular $m \times m$ matrix (R) such that $(\mathcal{V}) = (R)(U)$. Now, for any complex m-tuple $\boldsymbol{w} = (w_1, \ldots, w_m)$, we have $\boldsymbol{w}^\dagger (\chi)\boldsymbol{w} = \boldsymbol{w}^\dagger (R)^\dagger (Q)(R)\boldsymbol{w}$ which leads to $(\chi) = (R)^\dagger (Q)(R)$. Thus, (χ) and (Q) are related by a congruence, Sylvester's inertia law [17] applies and the number of positive (negative) ρ_ps coincides with the number of positive (negative) $(\chi_p - \hat{\chi}_l)$s. In this way, the sufficiency of (i) and (ii) is proved and the proof of the generalized Carathéodory theorem is complete.

The corollary of this theorem is easily proved by the previous considerations.

4. Conclusion

Summarizing, given n complex numbers $c_1, \ldots, c_n$ one considers the Hermitian Toeplitz matrix $(\hat{C})$ defined by (4). One evaluates its distinct eigenvalues, denoted by $\hat{\chi}_1 < \cdots < \hat{\chi}_\nu$ with mutiplicities $\mu_1, \ldots, \mu_\nu$. Setting $(C_l) \equiv (\hat{C}) - \hat{\chi}_l(I)$ with $l = 1, \ldots, \nu$, the resulting matrices with $l \neq 1$ and $l \neq \nu$ are indefinite, while (C_1) and (C_ν), respectively, are non-negative and non-positive definite. For each $l \in \{1, \ldots, \nu\}$, the complex numbers $-\hat{\chi}_l, c_1, \ldots, c_n$ can uniquely be written in the form (23) with m_l given by (14) iff (i) the rank and the principal rank of (C_l) are equal to m_l and (ii) the $(m_l + 1) \times (m_l + 1)$ matrix $(\mathcal{S}_l)$, defined by (17)–(20), is non-negative definite and has rank m_l.

One concludes that we have μ different ways for writing the n complex numbers $c_1, \ldots, c_n$ in the form (23), where μ denotes the number of (C_l) that obeys conditions (i) and (ii).

As a final remark we observe that, in theorems 2.2 and 3.4 of Ellis and Lay [8], the assumption that the resolvent has unimodular roots can be removed, thanks to lemma A. It can be substituted with the constructive requirements: (a) the coefficients of the resolvent equation obey equation (21) if the given Toeplitz matrix $(\mathcal{T})$ is not Hermitian (oppositely, the condition is already fulfilled as proved in appendix B), (b) if the discriminant of the resolvent equation is equal to zero, one algebraically eliminates [12, 13] all the multiple roots from the resolvent obtaining the lowest degree resolvent equation (i.e. the polynomial equation defined by the distinct roots of the outset resolvent), (c) by the coefficients of the (new) resolvent equation one constructs the relevant matrix $(\mathcal{S})$ and one checks its positive definiteness. In the only affirmative case generalized factorization (25) of $(\mathcal{T})$ is possible. This reduces to Carathéodory's generalized one if $(\mathcal{T})$ is Hermitian and the outset resolvent has no multiple roots.

Appendix A. Proof of the lemma on the unimodular roots

We shall now prove lemma A reported in section 2.2. It states the necessary and sufficient condition for all the zeros of a polynomial equation with complex coefficients to lie on the unit circle. It is stressed that the known theorems ensuring such property lean upon the existence of other polynomial with unimodular roots [9, 10], while lemma A only involves the coefficients of (15), the given polynomial equation.

We first observe that condition (a) of lemma A amounts to state that condition (21) are necessary and sufficient for the roots of (15) to obey $\overline{\epsilon_j} = 1/\epsilon_{i_j}$ for $j = 1, \ldots, N$ and $i_1, \ldots, i_N$ equal to a permutation of $\{1, \ldots, N\}$. This is proved in appendix B in the form of lemma B. Since the reported conditions on ϵ_js do neither ensure that the roots are distinct nor that they are unimodular, we must show that remaining condition (b) of lemma A is necessary and sufficient for the validity of the last two properties. To this aim, we start by observing that the noticed properties of ϵ_js allow us to extend definitions (17) to negative ps, so as to write

$$\sigma_p \equiv \sum_{j=1}^{N} {\epsilon_j}^p \qquad p = 0, \pm 1, \pm 2, \ldots, \tag{A.1}$$

and to convert definitions (19) into the equalities

$$\sigma_{-p} = \overline{\sigma_p}, \qquad p = 0, \pm 1, \pm 2, \ldots. \tag{A.2}$$

Furthermore, the last of relations (18) also holds true for negative p integers. In fact, the complex conjugate of this relation by (21) becomes

$$\begin{aligned} &a_0(\overline{a_N}\,\overline{\sigma_{p+N}} + \cdots + \overline{a_2}\,\overline{\sigma_{p+2}} + \overline{a_1}\,\overline{\sigma_{p+1}} + \overline{a_0}\,\overline{\sigma_p}) \\ &\qquad = a_0\sigma_{-p-N} + \cdots + a_{N-2}\sigma_{-p-2} + a_{N-1}\sigma_{-p-1} + a_N\sigma_{-p} = 0, \end{aligned} \tag{A.3}$$

and the statement is proved. The previous considerations show that all the matrix elements σ_ps of matrix $(\mathcal{S})$, defined by (20), are known in terms of $a_0, \ldots, a_N$.

We now prove the necessity of condition (b) of lemma A.

If $N\epsilon_j$s are unimodular and distinct, the $N \times (N+1)$ matrix $(\mathcal{V})$ defined by (8) exists. The assumed properties of ϵ_js ensure that $(\mathcal{S}) = (\mathcal{V}^\dagger)(\mathcal{V})$, that $\det(\mathcal{S}) = 0$ and that the rank of $(\mathcal{V})$ is N. These three properties imply that $(\mathcal{S})$ is a non-negative definite matrix of rank N and the necessity of condition (b) of lemma A is proved. We can apply Carathéodory's theorem to $(\mathcal{S})$ and conclude that σ_ps can uniquely be written as

$$\sigma_p = \sum_{j=1}^{N} \tau_j {\omega_j}^p, \qquad p = 0, \pm 1, \ldots, \pm N \tag{A.4}$$

with ω_js unimodular, distinct and roots of the resolvent equation generated by matrix $(\mathcal{S})$, i.e.,

$$Q_N(z) = {\Delta_N}^{-1} \det \begin{pmatrix} \sigma_0 & \sigma_1 & \cdots & \sigma_{N-1} & \sigma_N \\ \sigma_{-1} & \sigma_0 & \cdots & \sigma_{N-2} & \sigma_{N-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & z & \cdots & z^{N-1} & z^N \end{pmatrix} = 0, \tag{A.5}$$

where Δ_N denotes the determinant of the $N \times N$ upper left principal minor of $(\mathcal{S})$. The comparison of (A.4) with (A.1) and the uniqueness of the Carathéodory decomposition imply $\tau_1 = \cdots = \tau_N = 1$ and $(\omega_1, \ldots, \omega_n) = (\epsilon_1, \ldots, \epsilon_N)$. The last equality implies that $Q_N(z) = P_N(z)$.

To prove the sufficiency of condition (b) we must show that the roots of $P_N(z) = 0$ are distinct and unimodular if $(\mathcal{S})$ is a non-negative matrix and has rank N. In fact, from the last property follows that $\Delta_N \neq 0$, and from definition (A.1) and property (A.2) that

$$\begin{aligned} \Delta_N &= \det \begin{pmatrix} N & \sum_{j=1}^N \epsilon_j & \sum_{j=1}^N \epsilon_j^2 & \cdots & \sum_{j=1}^N \epsilon_j^{N-1} \\ \sigma_{-1} & \sigma_0 & \sigma_1 & \cdots & \sigma_{N-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sigma_{-N+1} & \sigma_{-N+2} & \sigma_{-N+3} & \cdots & \sigma_0 \end{pmatrix} \\ &= \sum_{j=1}^{N} \det \begin{pmatrix} 1 & \epsilon_j & {\epsilon_j}^2 & \cdots & {\epsilon_j}^{N-1} \\ \sigma_{-1} & \sigma_0 & \sigma_1 & \cdots & \sigma_{N-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sigma_{-N+1} & \sigma_{-N+2} & \sigma_{-N+3} & \cdots & \sigma_0 \end{pmatrix} \\ &= \sum_{1 \leqslant j_1, \ldots, j_N \leqslant N} \det \begin{pmatrix} 1 & \epsilon_{j_1} & {\epsilon_{j_1}}^2 & \cdots & {\epsilon_{j_1}}^{N-1} \\ {\epsilon_{j_2}}^{-1} & 1 & \epsilon_{j_2} & \cdots & {\epsilon_{j_2}}^{N-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ {\epsilon_{j_N}}^{-N+1} & {\epsilon_{j_N}}^{-N+2} & {\epsilon_{j_N}}^{-N+3} & \cdots & 1 \end{pmatrix}. \end{aligned} \tag{A.6}$$

The last expression can also be written as

$$\sum_{1 \leqslant j_1, \ldots, j_N \leqslant N} \frac{1}{\epsilon_{j_1}^0 \epsilon_{j_2}^1 \cdots {\epsilon_{j_N}}^{N-1}} \det \begin{pmatrix} 1 & \epsilon_{j_1} & \epsilon_{j_1}^2 & \cdots & {\epsilon_{j_1}}^{N-1} \\ 1 & \epsilon_{j_2} & \epsilon_{j_2}^2 & \cdots & {\epsilon_{j_2}}^{N-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & \epsilon_{j_N} & \epsilon_{j_N}^2 & \cdots & {\epsilon_{j_N}}^{N-1} \end{pmatrix}. \tag{A.7}$$

Within the sum the only terms with $j_1 \neq j_2 \neq \cdots \neq j_N$ differ from zero. In other words, the possible values of $\{j_1, \ldots, j_N\}$ correspond to the possible permutations of $\{1, \ldots, N\}$. The values of the corresponding determinants are $(-)^P \prod_{1 \leqslant i < j \leqslant N} (\epsilon_j - \epsilon_i)$ where P is the number of the transpositions required for passing from $\{j_1, \ldots, j_N\}$ to $\{1, \ldots, N\}$. One concludes that

$$\Delta_N = \prod_{1\leqslant i<j\leqslant N} (\epsilon_j - \epsilon_i)(1/\epsilon_j - 1/\epsilon_i). \tag{A.8}$$

Thus, $\Delta_N \neq 0$ ensures that the roots of $P_N(z) = 0$ are distinct. We show now that the resolvent of $(\mathcal{S})$, i.e. equation (A.5), coincides with $P_N(z)$. In fact, $Q_N(z)$ can be written as $Q_N(z) \equiv \sum_{p=0}^{N} q_p z^p = 0$ with

$$q_p \equiv \frac{(-1)^{N+p}}{\Delta_N}\det\begin{pmatrix} \sigma_0 & \cdots & \sigma_{p-1} & \sigma_{p+1} & \cdots & \sigma_{N-1} & \sigma_N \\ \sigma_{-1} & \cdots & \sigma_{p-2} & \sigma_p & \cdots & \sigma_{N-2} & \sigma_{N-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \sigma_{-N+1} & \cdots & \sigma_{-N+p} & \sigma_{-N+p+2} & \cdots & \sigma_0 & \sigma_1 \end{pmatrix}. \tag{A.9}$$

Manipulations similar to those performed in equations (A.6), (A.7) convert the determinant present in (A.9) into

$$\sum_{1\leqslant j_1,\ldots,j_N\leqslant N} \frac{1}{\epsilon_{j_1}^0\epsilon_{j_2}^1\cdots\epsilon_{j_N}{}^{N-1}}\det\begin{pmatrix} 1 & \cdots & \epsilon_{j_1}{}^{p-1} & \epsilon_{j_1}{}^{p+1} & \cdots & \epsilon_{j_1}{}^{N-1} & \epsilon_{j_1}{}^{N} \\ 1 & \cdots & \epsilon_{j_2}{}^{p-1} & \epsilon_{j_2}{}^{p+1} & \cdots & \epsilon_{j_2}{}^{N-1} & \epsilon_{j_2}{}^{N} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & \cdots & \epsilon_{j_N}{}^{p-1} & \epsilon_{j_N}{}^{p+1} & \cdots & \epsilon_{j_N}{}^{N-1} & \epsilon_{j_N}{}^{N} \end{pmatrix}.$$

Using the property that $\epsilon_j{}^N = -\sum_{p=0}^{N-1} a_p\epsilon_j{}^p$, the above expression becomes

$$\sum_{1\leqslant j_1,\ldots,j_N\leqslant N} \frac{-a_p}{\epsilon_{j_1}^0\epsilon_{j_2}^1\cdots\epsilon_{j_N}{}^{N-1}}\det\begin{pmatrix} 1 & \cdots & \epsilon_{j_1}{}^{p-1} & \epsilon_{j_1}{}^{p+1} & \cdots & \epsilon_{j_1}{}^{N-1} & \epsilon_{j_1}{}^{p} \\ 1 & \cdots & \epsilon_{j_2}{}^{p-1} & \epsilon_{j_2}{}^{p+1} & \cdots & \epsilon_{j_2}{}^{N-1} & \epsilon_{j_2}{}^{p} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & \cdots & \epsilon_{j_N}{}^{p-1} & \epsilon_{j_N}{}^{p+1} & \cdots & \epsilon_{j_N}{}^{N-1} & \epsilon_{j_N}{}^{p} \end{pmatrix},$$

and from (A.7) and (A.9) one concludes that $q_p = a_p$, $p = 0, \ldots, N$. In this way, the resolvent of $(\mathcal{S})$ coincides with $P_N(z)$. Consequently, ϵ_js are also unimodular because the assumed non-negativeness of $(\mathcal{S})$ and Carathéodory's theorem ensure that the unimodularity is true for the roots of resolvent $Q_N(z)$.

In this way lemma A is fully proved.

From the lemma follows, for instance, that the quadratic and cubic equations have distinct unimodular roots iff their coefficients are as follows:

$$\begin{aligned} &N = 2: \quad a_1 = \pm\rho\mathrm{e}^{\mathrm{i}\phi/2}, \quad a_0 = \mathrm{e}^{\mathrm{i}\phi} \quad \text{with} \quad 0 \leqslant \rho < 2 \quad \text{and} \quad \phi \in [0, 2\pi) \\ &N = 3: \quad a_2 = \rho\mathrm{e}^{\mathrm{i}(\phi-\psi)}, a_1 = \rho\mathrm{e}^{\mathrm{i}\psi}, a_0 = \mathrm{e}^{\mathrm{i}\phi} \quad \text{with} \\ &\quad \text{either } 0 \leqslant \rho \leqslant 1, \phi, \psi \in [0, 2\pi) \\ &\quad \text{or } 1 \leqslant \rho < 3, \phi \in [0, 2\pi), \quad (2\phi - \Phi(\rho)) < 3\psi < (2\phi + \Phi(\rho)) \end{aligned}$$

where $\Phi(\rho) \equiv \arccos[(\rho^4 + 18\rho^2 - 27)/8\rho^3]$.

An example of Hermitian Toeplitz matrix whose resolvent does not obey the conditions required by the lemma because the rank of (S) is smaller than N was given at the end of section 2.3.

We remark that lemma A could more generally be formulated as follows[6]:

Lemma A'. *A polynomial equation of degree N has unimodular roots iff (a) its coefficients obey conditions (21) and (b) Toeplitz matrix* $(\mathcal{S})$*, defined by (20), is non-negative definite. The number of the distinct roots is equal to M, the rank of* $(\mathcal{S})$.

This lemma is essentially proved as lemma A.

[6] We thank Dr Alessandro De Paris for having brought this point to our attention.

Appendix B. Properties of the resolvent coefficients

For completeness, we first report the proof of a well-known [9] lemma stating the conditions for a polynomial to be *self-reciprocal*. The lemma says

Lemma B. *The coefficients a_l of a polynomial equation of degree N obey*

$$\overline{a_n} = a_{N-n}/a_0, \qquad n = 0, \ldots, N, \quad with \quad a_0 \neq 0, \tag{B.1}$$

iff the roots of the equation are such that $\overline{\epsilon_j} = 1/\epsilon_{i_j}$ for $j = 1, \ldots, N$ and $i_1, \ldots, i_N$ equal to a permutation of $\{1, \ldots, N\}$.

To prove the necessity one starts from expression (16) of a_ns. The assumed property of the roots implies that $a_0 \neq 0$. Moreover, taking the complex conjugate of a_n one finds

$$\begin{aligned}\overline{a_n} &= (-)^{N-n} \sum_{1\leqslant j_1<\cdots<j_{N-n}\leqslant N} \overline{\epsilon_{j_1}} \cdots \overline{\epsilon_{j_{N-n}}} \\ &= (-)^{N-n} \sum_{1\leqslant j_1<\cdots<j_{N-n}\leqslant N} \frac{1}{\epsilon_{i_{j_1}} \cdots \epsilon_{i_{j_{N-n}}}} \\ &= \frac{(-)^{N-n}}{\epsilon_1 \cdots \epsilon_N} \sum_{1\leqslant i_1<\cdots<i_{N-n}\leqslant N} \frac{\epsilon_1 \cdots \epsilon_N}{\epsilon_{i_1} \cdots \epsilon_{i_{N-n}}} \\ &= \frac{(-)^{N-n}}{\prod_{j=1}^N \epsilon_j} \sum_{1\leqslant i_1<\cdots<i_n\leqslant N} \epsilon_{i_1} \cdots \epsilon_{i_n} = \frac{a_{N-n}}{a_0}.\end{aligned}$$

To prove the sufficiency one observes that

$$\begin{aligned}\overline{P_N(\overline{z})} &= \prod_{j=1}^{N}(z - \overline{\epsilon_j}) = \sum_{j=1}^{N} \overline{a_j} z^j = \sum_{j=1}^{N} \frac{a_{m-j}}{a_0} z^j \\ &= \frac{z^N}{a_0} \sum_{t=0}^{N} \frac{a_t}{z^t} = \frac{z^N}{a_0} \prod_{j=1}^{N} \left(\frac{1}{z} - \epsilon_j\right).\end{aligned} \tag{B.2}$$

The previous manipulations require that no root is equal to zero and this is ensured by the condition $a_0 \neq 0$. With $z = \overline{\epsilon_j}$, whatever j in $\{1, \ldots, m\}$, the first product in (B.2) vanishes. For the last to vanish it must result $1/\overline{\epsilon_j} = \epsilon_{i_j}$ and the property of the roots is recovered.

In passing it is noted that lemma B is true also when some roots have multiplicity greater than one.

We prove now the property mentioned at the beginning of section 3.3, namely:

Property B. *The coefficients of the resolvent equation, defined by equations (32)–(34), obey conditions (21).*

In fact, the unimodularity of a_0 is evident from (34). Since $a_0 \neq 0$, after substituting (32) into (21) one finds

$$\overline{\det\left(C^{1,\ldots,l,l+2,\ldots,m+1}_{1,\ldots,m}\right)} \det\left(C^{2,\ldots,m+1}_{1,\ldots,m}\right) = \det\left(C^{1,\ldots,m-l,m-l+2,\ldots,m+1}_{1,\ldots,m}\right) \det\left(C^{1,\ldots,m}_{1,\ldots,m}\right), \qquad l = 0, \ldots, m. \tag{B.3}$$

Taking $n = (m+1)$, $(j_1, \ldots, j_m) = (1, \ldots, m-l, m-l+2, \ldots, m+1)$ and $(i_1, \ldots, i_m) = (1, \ldots, m)$ in (C.3), one finds that

$$\det\left(C^{1,\ldots,m-l,m-l+2,\ldots,m+1}_{1,\ldots,m}\right) = \overline{\det\left(C^{1,\ldots,l,l+2,\ldots,m+1}_{2,\ldots,m+1}\right)}. \tag{B.4}$$

The rhs of (B.3) can be written as

$$\overline{\det\left(C^{1,\ldots,l,l+2,\ldots,m+1}_{2,\ldots,m+1}\right)}\det\left(C^{\ 2,\ldots,m+1}_{2,\ldots,m+1}\right) \tag{B.5}$$

by (B.4) and the property that all the $m \times m$ strictly principal minors of (C) coincide. From (C.7) follows that

$$\det\left(C^{\ 2,\ldots,m+1}_{2,\ldots,m+1}\right) = \det\left(\lambda^{\ 1,\ldots,m}_{2,\ldots,m+1}\right)\det\left(C^{2,\ldots,m+1}_{1,\ldots,m}\right),$$

and

$$\det\left(C^{1,\ldots,l,l+2,\ldots,m+1}_{2,\ldots,m+1}\right) = \det\left(\lambda^{\ 1,\ldots,m}_{2,\ldots,m+1}\right)\det\left(C^{1,\ldots,l,l+2,\ldots,m+1}_{1,\ldots,m}\right).$$

The substitution of the above two relations into (B.5) yields

$$\left|\det\left(\lambda^{\ 1,\ldots,m}_{2,\ldots,m+1}\right)\right|^2\overline{\det\left(C^{1,\ldots,l,l+2,\ldots,m+1}_{1,\ldots,m}\right)}\det\left(C^{2,\ldots,m+1}_{1,\ldots,m}\right),$$

that identically coincides with the left-hand side of (B.3) due to (C.8).

Appendix C. Some properties of Hermitian Toeplitz matrices

We list here a series of properties obeyed by a square Hermitian Toeplitz matrix (C) of order n and partly reported in [9].

(1) Its elements obey to

$$\overline{C}_{r,s} = C_{s,r} = c_{r-s} = \overline{c}_{s-r}, \qquad r, s = 1, \ldots, n, \tag{C.1}$$

so that all the elements of (C) contained in a line parallel to the main diagonal are equal.

(2) One has the reflection symmetry with respect to the second diagonal formalized by the condition

$$C_{r,s} = C_{n+1-s,n+1-r}. \tag{C.2}$$

(3) All the $(m \times m)$ strictly principal minors of (C), whatever the considered rows (and columns), are identical.
'The property is a consequence of (1).'

(4) For any choice of m rows $(1 \leqslant i_1 < \cdots < i_m \leqslant n)$ and m columns $(1 \leqslant j_1 < \cdots < j_m \leqslant n)$ it results

$$\det\left(C^{\ j_1,\ldots,j_m}_{i_1,\ldots,i_m}\right) = \det\left(C^{T\ \ (n+1-j_m),\ldots,(n+1-j_1)}_{(n+1-i_m),\ldots,(n+1-i_1)}\right) = \overline{\det\left(C^{\ (n+1-j_m),\ldots,(n+1-j_1)}_{(n+1-i_m),\ldots,(n+1-i_1)}\right)}. \tag{C.3}$$

'The first equality, where (C^T) denotes the transposed of (C), follows from property (2) and the second from the Hermiticity of (C).'

The following properties, that we think to be original, hold only true for Hermitian Toeplitz matrices having their rank equal to the principal one.

(5) If the principal rank of a Hermitian Toeplitz matrix (C) is equal to the rank m $(\leqslant n)$ of (C), any $(m \times m)$ strictly principal minor of (C) is non-singular.

'The property immediately follows from the definition of 'principal rank' of a matrix, reported in section 2.2 below statement (i), and (3)'.

(6) For any $(n \times n)$ Hermitian Toeplitz matrix of rank equal to its principal rank m ($\leqslant n$), the determinant of any of its minors formed by m *subsequent* rows and m *subsequent* columns is simply related by a phase factor to the determinant of the (strictly) principal minor contained in the considered rows or columns, i.e.,

$$\begin{aligned}
&\det\left(C\,{}^{q+1,\ldots,q+m}_{p+1,\ldots,p+m}\right) = \mathrm{e}^{\mathrm{i}\theta_{p-q}} \det\left(C\,{}^{p+1,\ldots,p+m}_{p+1,\ldots,p+m}\right),\\
&\det\left(C\,{}^{q+1,\ldots,q+m}_{p+1,\ldots,p+m}\right) = \mathrm{e}^{\mathrm{i}\theta_{p-q}} \det\left(C\,{}^{q+1,\ldots,q+m}_{q+1,\ldots,q+m}\right),\\
&\theta_{p-q} \in \mathbb{R}, \qquad p, q = 0, 1, \ldots, n-m.
\end{aligned} \tag{C.4}$$

'Clearly, if the first of the above two equalities is true the second is also true because of (3). To prove the first of equalities (C.4) one observes that (4) implies that any m distinct rows of (C) can be written as linear combinations of m other distinct rows (see, e.g., [12], chapter III). Hence, rows $(p+1), \ldots, (p+m)$ can be expressed in terms of rows $(q+1), \ldots, (q+m)$ as

$$C_{r,s} = \sum_{t=q+1}^{q+m} \lambda_{r,t} C_{t,s}, \qquad r = (p+1), \ldots, (p+m), \qquad s = 1, \ldots, n, \tag{C.5}$$

where $\lambda_{r,t}$s are suitable numerical coefficients. From these relations follows that

$$\det\left(C\,{}^{p+1,\ldots,p+m}_{p+1,\ldots,p+m}\right) = \det\left(\lambda\,{}^{q+1,\ldots,q+m}_{p+1,\ldots,p+m}\right) \det\left(C\,{}^{p+1,\ldots,p+m}_{q+1,\ldots,q+m}\right). \tag{C.6}$$

Due to (e) the left-hand side of (C.6) is different from zero so that both factors on the rhs are different from zero. The complex conjugation of (C.6), by the Hermiticity of (C), yields

$$\det\left(C\,{}^{p+1,\ldots,p+m}_{p+1,\ldots,p+m}\right) = \overline{\det\left(\lambda\,{}^{q+1,\ldots,q+m}_{p+1,\ldots,p+m}\right)} \det\left(C\,{}^{q+1,\ldots,q+m}_{p+1,\ldots,p+m}\right). \tag{C.7}$$

From equation (C.5) also follows that

$$\begin{aligned}
\det\left(C\,{}^{q+1,\ldots,q+m}_{p+1,\ldots,p+m}\right) &= \det\left(\lambda\,{}^{q+1,\ldots,q+m}_{p+1,\ldots,p+m}\right) \det\left(C\,{}^{q+1,\ldots,q+m}_{q+1,\ldots,q+m}\right)\\
&= \det\left(\lambda\,{}^{q+1,\ldots,q+m}_{p+1,\ldots,p+m}\right) \det\left(C\,{}^{p+1,\ldots,p+m}_{p+1,\ldots,p+m}\right),
\end{aligned}$$

where the last equality follows from (3). The substitution of the last equality into equation (C.6) and the fact that, by assumption, $\det\left(C\,{}^{p+1,\ldots,p+m}_{p+1,\ldots,p+m}\right) \neq 0$ imply that

$$\left|\det\left(\lambda\,{}^{q+1,\ldots,q+m}_{p+1,\ldots,p+m}\right)\right|^2 = 1 \qquad p, q = 0, \ldots, n-m, \tag{C.8}$$

and equation (C.4) is proved. That the phase factor depends on $p-q$ instead of (p, q) follows from the fact that the two determinants present in (C.4) do not change with the two substitutions $p \to p+1$ and $q \to q+1$ owing to (C.1).' An immediate consequence of (6) is the property that

(7) Any $(m \times m)$ minor formed by m subsequent rows and m subsequent columns of a Hermitian Toeplitz matrix with rank equal to its principal rank m is non-singular.

References

[1] Carathéodory C 1911 *Rend. Circ. Matem. Palermo* **32** 193
[2] Cervellino A and Ciccariello S 1996 *Riv. Nuovo Cimento* **19/8** 1
[3] Pisarenko V F 1973 *Geophys. J. R. Astron. Soc.* **33** 346
[4] Beylkin G and Monzon L 2002 *Appl. Comput. Harmon. Anal.* **12** 332

[5] Feigin L A and Svergun D I 1987 *Structure Analysis by Small-Angle X-ray and Neutron Scattering* (New York: Plenum)
[6] Cervellino A and Ciccariello S 2005 *Acta Crystallogr.* A **61** 494
[7] Grenander U and Szegö G 1984 *Toeplitz Forms and Their Applications* (New York: Chelsea) p 56
[8] Ellis R L and Lay D C 1992 *Linear Algebra Appl.* **173** 19
[9] Chen W 1995 *J. Math. Anal. Appl.* **190** 714
[10] McNamee J M 1999 *J. Comput. Appl. Math.* **110** 305
[11] Iohvidov I S 1982 *Hankel and Toeplitz Matrices and Forms: Algebraic Theory* (Boston, MA: Birkhauser)
[12] Caccioppoli R 1956 *Lezioni di Analisi Matematica* vol I (Napoli: Liguori)
[13] Cox D, Little J and O'Shea D 1998 *Using Algebraic Geometry* (Berlin: Springer)
[14] Gantmacher F R 1966 *Théorie des Matrices* vol I (Paris: Dunod)
[15] Mehta M L 1989 *Matrix Theory: Selected Topics and Useful Results* (Les Ulis: Les Editions de Physique)
[16] Goedkoop J A 1950 *Acta Crystallogr.* **3** 374
[17] Weisstein E W *Sylvester's Inertia Law* MathWorld—A Wolfram Web: http://mathworld.wolfram.com/CongruenceTransformation.html