CrossMark

On the Structure of Positive Semi-Definite Finite Rank General Domain Hankel and Toeplitz Operators in Several Variables

Fredrik Andersson[1] · **Marcus Carlsson[1]**

Abstract Multivariate versions of the Kronecker theorem in the continuous multivariate setting has recently been published, that characterize the generating functions that give rise to finite rank multidimensional Hankel and Toeplitz type operators defined on general domains. In this paper we study how the additional assumption of positive semi-definite affects the characterization of the corresponding generating functions. We show that these theorems become particularly transparent in the continuous setting, by providing elegant if-and-only-if statements connecting the rank with sums of exponential functions. We also discuss how these operators can be discretized, giving rise to an interesting class of structured matrices that inherit desirable properties from their continuous analogs. In particular we describe how the continuous Kronecker theorem also applies to these structured matrices, given sufficient sampling. We also provide a new proof for the Carathéodory-Fejér theorem for block Toeplitz matrices, based on tools from tensor algebra.

Keywords Hankel · Toeplitz · Finite rank · Kronecker theorem · Carathéodory-Fejér theorem · Sums of exponentials

Mathematics Subject Classification Primary 47B35 · 15A03; Secondary 15A69 · 15B05 · 15B48 · 33B10 · 47A13

✉ Marcus Carlsson
marcus.carlsson@math.lu.se

Fredrik Andersson
fa@maths.lth.se

[1] Centre for Mathematical Sciences, Lund University, Box 118, 22100 Lund, Sweden

1 Introduction

The connection between low-rank Hankel and Toeplitz operators and matrices, and properties of the functions that generate them play a crucial role for instance in frequency estimation [7,32,46–48], system identification [14,16,31,33] and approximation theory [4–6,8–10,42]. The reason for this is that there is a connection between the rank of such an operator and its generating function being a sum of exponential functions, where the number of terms is connected to the rank of the operator (Kronecker's theorem). Moreover, adding the condition of positive semi-definite imposes further restrictions on the sum of exponentials (Caratheódory-Fejér's and Fischer's theorem), a result which underlies e.g. Pisarenko's famous method for frequency estimation [43].

We provide corresponding theorems in the multidimensional setting. In contrast to the one dimensional situation, the multidimensional framework provides substantial flexibility in how to define these operators. Whereas most previous research on multidimensional Hankel and Toeplitz type operators considers "generating functions/sequences" f that are defined on product domains, we here consider a framework where f is defined on an open connected and bounded domain Ω in $\mathbb{R}^d$ (or discretizations thereof). Besides providing beautiful new theorems, it is our hope that the new results in this paper will pave the way for applications in multidimensional frequency estimation/approximation/compression, in analogy with the use of Toeplitz and Hankel matrices in the one dimensional setting. For this reason, we present results both in the continuous and discretized setting, and discuss how they influence each other.

To present the key ideas, we here focus mainly on the continuous theory since it is more transparent. "General domain Hankel (Toeplitz) operators" is a class of integral operators whose kernel $K(x, y)$ is of the form $K(x, y) = f(x + y)$ or $K(x, y) = f(x - y)$[1], and f is the so called "generating function". Their precise definition also depends on an auxiliary domain Ω on which f is defined, we postpone detailed definitions to Sect. 2.2. We denote by Γ_f a generic general domain Hankel operator and by Θ_f their Toeplitz counterparts (see Fig. 1 for an example of a discretized Γ_f). These operators were introduced in [3] where it is shown that if Γ_f or Θ_f has rank equal to $K < \infty$, then f is necessarily an exponential polynomial;

$$f(x) = \sum_{j=1}^{J} p_j(x) e^{\zeta_j \cdot x} \tag{1.1}$$

where $J \leq K$ (assuming no cancellation), p_j are polynomials in $x = (x_1, \ldots, x_d)$, $\zeta_j \in \mathbb{C}^d$ and $\zeta_j \cdot x$ denotes the standard scalar product

$$\zeta_j \cdot x = \sum_{m=1}^{d} \zeta_{j,m} x_m.$$

---

[1] These have previously been referred to as truncated correlation/convolution operators, but we feel that the new name is more suggestive. In one dimension, these objects include Wiener–Hopf operators and Toeplitz/Hankel operators on the Paley–Wiener space.
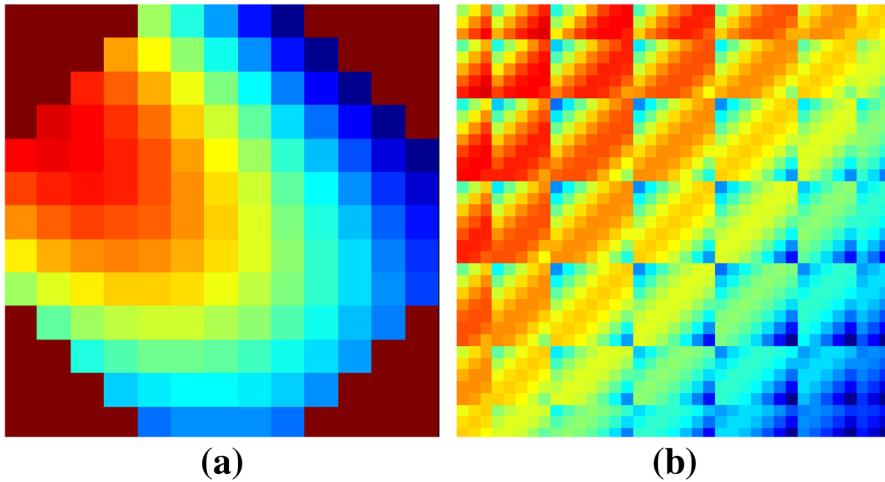
(a)   **(b)**

Fig. 1   a "Generating sequence" defined on a disc Ω; b The matrix realization of the corresponding "general domain Hankel operator" (see Sect. 4.1 for further details)

Conversely, any such exponential polynomial gives rise to finite rank Γ_f and Θ_f respectively, and there is a method to determine the rank given the generating function (1.1). Most notably, the rank equals K if f is of the form

$$f(x) = \sum_{k=1}^{K} c_k e^{\zeta_k \cdot x}, \tag{1.2}$$

where $c_k \in \mathbb{C}$ (assuming that there is no cancelation in (1.2)).

The main topic of this paper is the study of how the additional condition that Γ_f or Θ_f be positive semi-definite (PSD) affects the generating function f. We prove that Θ_f then has rank K if and only if f is of the form

$$f(x) = \sum_{k=1}^{K} c_k e^{i\xi_k \cdot x} \tag{1.3}$$

where $c_k > 0$ and $\xi_k \in \mathbb{R}^d$ (Theorem 7.1), which in a certain sense is an extension of Carathéodory-Fejér's theorem on PSD Toeplitz matrices. Correspondingly, Γ_f is PSD and has rank K if and only if f is of the form

$$f(x) = \sum_{k=1}^{K} c_k e^{\xi_k \cdot x} \tag{1.4}$$

where again $c_k > 0$ and $\xi_k \in \mathbb{R}^d$ (Theorem 8.1). Similar results for Hankel matrices date back to work of Fischer [22].

The only of the above results that has a simple counterpart in the finite dimensional discretized multivariable setting is the Carathéodory-Fejér's theorem, which has been observed previously in [53] (concerning block Toeplitz matrices). In this paper we provide a general result on tensor products, which can be used to "lift" structure results in one-dimension to the multi-dimensional setting. We use this to give an alternative proof of the discrete Carathéodory-Fejér theorem, which subsequently is used to prove the continuous counterpart.

Fischer's theorem on the other hand has no neat version in the multivariable finite dimensional setting, but has been generalized to so called small Hankel operators on $\ell^2(\mathbb{N}^d)$ in [44], a paper which also contains a result analog to (1.4).

However, the product domain setting is rather restrictive and not always a natural choice. Whereas one-dimensional generating functions necessarily are defined on an interval, there is an abundance of possible regions to define their corresponding multidimensional cousins. Despite this, the majority of multivariate treatments of these issues are set either directly in a block-Toeplitz/Hankel setting, or rely on tensor products. In both cases the corresponding domain of definition Ω of the generating function/sequence is a square (or multi-cube), but for concrete applications to multidimensional frequency estimation, the available data need not be naturally defined on such a domain. In radially symmetric problems, a circle may be more suitable or, for certain boundary problems, a triangle might be more appropriate.

Concerning analogs of the above results for the discretized counterparts of Θ_f and Γ_f, we show in this paper how to construct "structured matrices" that approximate their continuous counterparts, and hence can be expected to inherit these desirable properties, given sufficient sampling rate. We give simple conditions on the regularity of f and Ω needed for this to be successful. This gives rise to an interesting class of structured matrices, which we call "general domain Hankel/Toeplitz matrices". As an example, in Fig. 1 we have a "generating sequence" f on a discretized disc, together with a plot of its general domain Hankel matrix.

The paper is organized as follows. In the next section we review the theory and at the same time introduce the operators we will be working with in the continuous setting (Sect. 2.2). The short Sect. 3 provides a tool from tensor algebra, and also introduce useful notation for the discrete setting. Section 4 discuss how to discretize the Γ_f's and Θ_f's, and we discuss particular cases such as block Toeplitz and Hankel matrices. In Sect. 5 we prove the Caratheodory-Fejér theorem in the discrete (block) setting. Section 6 shows that the discrete operators approximate the continuous counterparts, given sufficient sampling rate, and we discuss Kronecker's theorem. Sections 7 and 8 consider structure results for f under the PSD condition, first for Θ_f's and then for Γ_f's. In the last section, we extend the above results to the corresponding operators on unbounded domains.

2 Review of the Field

A Toeplitz matrix is a matrix that is constant on the diagonals, i.e. the matrix elements satisfy $a_{k,j} = a_{k+1,j+1}$ for all indices k, j such that the above formula is well defined. A sequence f such that $a_{k,j} = f_{k-j}$ is called its generating sequence. Hankel matri-

ces on the other hand are constant on the anti-diagonals; $a_{k,j} = a_{k+1,j-1}$; and the sequence f such that $a_{k,j} = f_{k+j}$ is called its generating sequence. Naturally, the set of subindices for f depends on whether we are dealing with Hankel or Toeplitz matrices (and also if the upper left element is taken as $a_{1,1}$ or $a_{0,0}$), but this is not of importance here and hence we do not specify it.

Suppose that the generating sequence of either a Hankel matrix H or a Toeplitz matrix T (of size $N \times N$) is a "discretized exponential polynomial"

$$f_n = \sum_{j=1}^{J} p_j(n)\lambda_j^n \tag{2.1}$$

(where $\lambda_j \in \mathbb{C}$ are distinct) of "cardinality"

$$K = \sum_{j=1}^{J} (\deg p_j + 1) \tag{2.2}$$

strictly less than N. Based on the theory of Vandermonde-matrices, one can show that the rank of either H or T equals K, and that the polynomials p_j and the λ_j's are unique. The converse statement is not true; consider for example the Hankel matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.3}$$

Clearly, the rank is 2 but the generating sequence $(1, 0, 0, 0, 0, 0, 0, 0, 1)$ is not of the form (2.1) with $J = 1$ or 2. However, in terms of applications this doesn't matter because of the following stronger statement: If T or H has rank $K < N$ then its generating sequence is "generically" of the form

$$f_n = \sum_{k=1}^{K} c_k \lambda_k^n, \tag{2.4}$$

a fact which underlies the famous ESPRIT frequency estimation algorithm [46].

The above claims are certainly well known to specialists, but very hard to find in the literature. The book [28], which has two sections devoted entirely to the topic of the rank of finite Toeplitz and Hankel matrices, gives a number of exact theorems relating the rank with the "characteristic" of the corresponding matrix, which is a set of numbers related to when determinants of certain sub-matrices vanish. It is possible to deduce representations of the form (2.1) (under certain additional assumptions) from these results, but this is never stated explicitly. Another viewpoint has been taken by Mourrain et al. [11,17,36,37], in which, loosely speaking, these matrices are analyzed using projective algebraic geometry and the 1 to the bottom right corresponds to the

point ∞. The book [41] deals exclusively with the infinite dimensional case, and generalizations thereof. For completeness we provide outlines of proofs of the claims made earlier in the appendix, based on results in [18,28].

In either case, the complexity of the theory does not reflect the relatively simple interaction between rank and exponential sums, as indicated in the introduction. There are however a few exceptions in the discrete setting. Kronecker's theorem says that for a Hankel operator (i.e. an infinite Hankel matrix acting on $\ell^2(\mathbb{N})$), the rank is K if and only if the generating sequence is of the desired form (2.1) (0^0 defined as 1), with the restriction that $|\lambda_j| < 1$ if one is only interested in bounded operators, see e.g. [13,29,30,41]. Also, it is finite rank and PSD if and only if the generating sequence is of the form (2.4) with $c_k > 0$ and $\lambda_k \in (-1, 1)$, a result which also has been extended to the multivariable (tensor product) setting [44]. In contrast, there are no finite rank *bounded* Toeplitz operators (on $\ell^2(\mathbb{N})$). If boundedness is not an issue, then a version of Kronecker's theorem holds for Toeplitz operators as well [18].

Adding the PSD condition for a Toeplitz matrix yields a simple result which is valid (without exceptions) for finite matrices. This is the essence of what usually is called the Carathéodory-Fejér theorem. The result was used by Pisarenko [43] to construct an algorithm for frequency estimation. Since then, this approach has rendered a lot of related algorithms, for instance the MUSIC method [47]. We reproduce the statement here for the convenience of the reader. For a proof see e.g. Theorem 12 in [2] or Section 4 in [26]. Other relevant references include [1,15].

Theorem 2.1 *Let T be a finite $(N + 1) \times (N + 1)$ Toeplitz matrix with generating sequence $(f_n)_{n=-N}^{N}$. Then T is PSD and* $\mathsf{Rank}\, T = K \leq N$ *if and only if f is of the form*

$$f(n) = \sum_{k=1}^{K} c_k \lambda_k^n \tag{2.5}$$

where $c_k > 0$ and the λ_k's are distinct and satisfy $|\lambda_k| = 1$.

The corresponding situation for Hankel matrices H is not as clean, since (2.3) is PSD and has rank 2, but do not fit with the model (2.5) for $c_k > 0$ and real λ_k's. Results of this type seems to go back to Fischer [22], and we will henceforth refer to statements relating the rank of PSD Hankel-type operators to the structure of their generating sequence/function, as "Fischer-type theorems" (see e.g. Theorem 5 [2] or [22]). Corresponding results in the full rank case can be found e.g. in [50].

We end this subsection with a few remarks on the practical use of Theorem 2.1. For a finitely sampled signal, the autocorrelation matrix can be estimated by H^*H where H is a (not necessarily square) Hankel matrix generated by the signal. This matrix will obviously be PSD, but in general it will not be Toeplitz. However, under the assumption that the λ_k's in (2.5) are well separated, the contribution from the scalar products of the different terms will be small and might therefore be ignored. Under these assumptions on the data, the matrix H^*H is PSD and approximately Toeplitz, which motivates the use of the Carathéodory-Fejér theorem as a means to retrieve the λ_k's.

2.1 Toeplitz and Hankel Operators on the Paley–Wiener Space

The theory in the continuous case is much "cleaner" than in the discrete case. In this section we introduce the integral operator counterpart of Toeplitz and Hankel matrices, and discuss Kronecker's theorem in this setting.

Given a function on the interval $[-2, 2]$, we define the truncated convolution operator $\Theta_f : L^2([-1, 1]) \to L^2([-1, 1])$ by

$$\Theta_f(g)(x) = \int f(x - y)g(y)\,dy. \tag{2.6}$$

Replacing $x - y$ by $x + y$ we obtain the a truncated correlation operator which we denote by Γ_f. Following [45], we refer to these operators as Toeplitz and Hankel operators on the Paley–Wiener space (although in [3] they were called finite interval convolution/correlation operators). It is easy to see that if we discretize these operators, i.e. replace integrals by finite sums, then we get Toeplitz and Hankel matrices, respectively. More on this in Sect. 4.1.

Kronecker's theorem (as formulated by Rochberg in [45]) then states that $\mathsf{Rank}\,\Theta_f = K$ (and $\mathsf{Rank}\,\Gamma_f = K$) if and only if f is of the form

$$f(x) = \sum_{j=1}^{J} p_j(x)e^{\zeta_j x} \tag{2.7}$$

where p_j are polynomials and $\zeta_j \in \mathbb{C}$. Moreover, the rank of Θ_f (or Γ_f) equals the cardinality

$$K = \sum_{j=1}^{J} (\deg p_j + 1). \tag{2.8}$$

However, functions of the form

$$f(x) = \sum_{k=1}^{K} c_k e^{\zeta_k x}, \quad c_k, \zeta_k \in \mathbb{C} \tag{2.9}$$

are known to be dense in the set of all generating functions giving rise to rank K finite interval convolution operators. Hence, the general form (2.7) is hiding the following simpler statement, which often is of practical importance. Θ_f *generically has rank K if and only if f is a sum of K exponential functions* (see the Appendix for an outline of a proof of this claim). The corresponding statement is false in several variables, which is shown in [3]. The polynomial factors appear in the limit if two frequencies in (2.9) approach each other and interfere destructively, e.g.

$$x = \lim_{\epsilon \to 0^+} \frac{e^{\epsilon x} - 1}{\epsilon}. \tag{2.10}$$

This can heuristically explain why these factors do not appear when adding the PSD condition, since the functions on the right of (2.10) give rise to one large positive and one large negative eigenvalue.

2.2 General Domain Hankel and Toeplitz Integral Operators in Several Variables

Given any (square integrable) function f on an open connected and bounded set Ω in $\mathbb{R}^d$, $d \geq 1$, the natural counterpart to the operator (2.6) is the *general domain Toeplitz integral operator* $\Theta_f : L^2(\Upsilon) \to L^2(\Xi)$ defined by

$$\Theta_f(g)(x) = \int_\Upsilon f(x - y)g(y)\, dy, \quad x \in \Xi, \tag{2.11}$$

where Ξ and Υ are connected open bounded sets such that

$$\Omega = \Xi - \Upsilon = \{x - y : x \in \Xi, \ y \in \Upsilon\}. \tag{2.12}$$

In [3] such operators are studied, (albeit under the name general domain truncated convolution operator), and their finite rank structure is completely characterized. It is easy to see that Θ_f has rank K whenever f has the form

$$f(x) = \sum_{k=1}^{K} c_k e^{\zeta_k \cdot x}, \quad x \in \Omega, \tag{2.13}$$

where the $\zeta_1, \ldots, \zeta_K \in \mathbb{C}^d$ are assumed to be distinct and all c_k's are non-zero.

The reverse direction is however not as neat as in the one-dimensional case. It is true that the rank is finite only if f is an exponential polynomial (i.e. the multidimensional analogue of (2.7), see Theorem 4.4 in [3]), but there is no counterpart to the simple formula (2.8). However, Section 5 (in [3]) gives a complete description of how to determine the rank given the generating function f explicitly, Sect. 7 gives results on the generic rank based on the degree of the polynomials that appear in f, we also provide lower bounds on the rank, and Sect. 8 investigates the fact that polynomial coefficients seem to appear more frequently in the multidimensional setting. Section 9 contains an investigation related to boundedness of these operators for the case of unbounded domains, which we will treat briefly in Sect. 9 of the present paper.

If we instead set $\Omega = \Xi + \Upsilon$ then we may define the *general domain Hankel integral operator* (called truncated correlation operator in [3])

$$\Gamma_f(g)(x) = \int_\Upsilon f(x + y)g(y)\, dy, \quad x \in \Xi. \tag{2.14}$$

This is the continuous analogue of finite Hankel (block) matrices. As in the finite dimensional case, there is no real difference between Γ_f and Θ_f regarding the finite rank structure. In fact, one turns into the other under composition with the "trivial"

operator $\iota(f)(x) = f(-x)$, and thus all statements concerning the rank of one can easily be transferred to the other. We remark however that composition with ι does not preserve PSD, and hence separate proofs are needed in this situation. Finally, we remark that the choice $\Upsilon = \Xi = \mathbb{R}_+^d$ gives what is known as "small Hankel operators". The study of their boundedness and related topics have received a lot of attention, see e.g. [21,34,35].

2.3 Other Multidimensional Versions

The usual multidimensional framework is that of block-Hankel and block-Toeplitz matrices, tensor products, or so called "small Hankel operators" on $\ell^2(\mathbb{N}^d)$. In all cases, the generating sequence f is forced to live on a product domain. For example, in [52] they consider the generating sequences of the form (1.2) (where x is on a discrete grid) and give conditions on the size of the block Hankel matrices under which the rank is K, and in [53] it is observed that the natural counterpart of the Carathéodory-Fejér theorem can be lifted by induction to the block Toeplitz setting. For the full rank case, factorizations of these kinds of operators have been investigated in [20,49]. Extensions to multi-linear algebra are addressed for instance in [38–40]. Rank deficient block Toeplitz matrices also play an important role in [23].

Concerning "small Hankel operators", in addition to [44] we wish to mention [27] where a formula for actually determining the rank appears, although this is based on reduction over the dimension and hence not suitable for non-product domains.

There is some heuristic overlap between [3] and [24,25]. In [24] they consider block Hankel matrices with polynomial generating function, and obtain results concerning their rank (Theorem 4.6) that overlap with Propositions 5.3, Theorem 7.4 and Proposition 7.7 of [3] for the 2d case. Proposition 7 in [25] is an extension to 2d of Kronecker's theorem for infinite block Hankel matrices (not truncated), which can be compared with Theorem 4.4 in [3].

In the discrete setting, the work of Mourrain et al. considers a general domain context, and what they call "quasi Toeplitz/Hankel matrices" correspond to what here is called "general domain Toeplitz/Hankel matrices" (we stick to this term since we feel it is more informative). See e.g. Section 3.5 in [37], where such matrices are considered for solving systems of polynomial equations. In [11], discrete multidimensional Hankel operators (not truncated) are studied, and Theorem 5.7 is a description of the rank of such an operator in terms of decompositions of related ideals. Combined with Theorem 7.34 of [17], this result also implies that the generating sequence must be of the form (2.1). (See also Section 3.2 of [36], where similar results are presented.) These results can be thought of as a finite dimensional analogue (for product domains) of Theorem 1.2 and Proposition 1.4 in [3]. Theorem 5.9 gives another condition on certain ideals in order for the generating sequence to be of the simpler type, i.e. the counterpart of (1.2) instead of (1.1). In Sect. 6 of the same paper they give conditions for when these results apply also to the truncated setting, based on rank preserving extension theorems. Similar results in the one-variable setting is found in Section 3 of [18].

Finally, we remark that the results in this paper concerning finite rank PSD Hankel operators partially overlap heuristically with results of [44] and those found in Section

4 in [36], where the formula (2.4) is found in the (non-truncated) discrete environment. In the latter reference they subsequently provide conditions under which this applies to the truncated setting.

With these remarks we end the review and begin to present the new results of this paper. For the sake of introducing useful notation, it is convenient to start with the result on tensor products, which will be used to "lift" the one-dimensional Carathéodory-Fejér theorem to the multidimensional discrete setting.

3 A Property of Tensor Products

Let $U_1, \ldots, U_d$ be finite dimensional linear subspaces of $\mathbb{C}^n$. Then $\otimes_{j=1}^d U_j$ is a linear subspace of $\otimes_{j=1}^d \mathbb{C}^n$, and the latter can be identified with the set of $\mathbb{C}$-valued functions on $\{1, \ldots, n\}^d$. Given $f \in \otimes_{j=1}^d \mathbb{C}^n$ and $x \in \{1, \ldots, n\}^d$, we will write $f(x)$ for the corresponding value. For fixed $x = (x_1, \ldots, x_{d-1}) \in \{1, \ldots, n\}^{d-1}$ we define vectors

$$f_1(x) = \Big(f(j, x_1, \ldots, x_{d-1}) \Big)_{j=1}^n, \ f_2(x) = \Big(f(x_1, j, x_2, \ldots, x_{d-1}) \Big)_{j=1}^n, \ etc.,$$

i.e. the vectors obtained from f by fixing all but one variable (and collecting the $d-1$ fixed variables in x). We refer to these vectors as "probes" of f. If $f \in \otimes_{j=1}^d U_j$ then it is easy to see that all probes f_j of f will be elements of U_j, $j = 1, \ldots, d$. The following theorem states that the converse is also true.

Theorem 3.1 *If all possible probes $f_j(x)$ of a given $f \in \otimes_{j=1}^d \mathbb{C}^n$ lie in U_j, then $f \in \otimes_{j=1}^d U_j$.*

Proof First consider the case $d = 2$. Let $V \subset \otimes_{j=1}^2 \mathbb{C}^n$ consist of all f with the property stated in the theorem. This is obviously linear and $U_1 \otimes U_2 \subset V$. If we do not have equality, we can pick an f in V which is orthogonal to $U_1 \otimes U_2$. At least one probe $f_1(k)$ must be a non-zero element u_1 of U_1. Given any $u_2 \in U_2$ we have

$$\langle u_1 \otimes u_2, f \rangle = \sum_{j=1}^n u_{2,j} \langle u_1, f_1(j) \rangle = \Big\langle u_2, \sum_{i=1}^n \overline{u_{1,i}} f_2(i) \Big\rangle. \tag{3.1}$$

From the middle representation and the choice of u_1, we see that at least one value of the vector $\sum_{i=1}^n \overline{u_{1,i}} f_2(i)$ is non-zero. Moreover this is a linear combination of probes $f_2(i)$, and hence an element of U_2. But then we can pick $u_2 \in U_2$ such that the scalar product (3.1) is non-zero, which is a contradiction to the choice of f. The theorem is thus proved in the case $d = 2$.

The general case now easily follows by induction on the dimension, noting that $\otimes_{j=1}^d \mathbb{C}^n$ can be identified with $\mathbb{C}^n \otimes (\otimes_{j=2}^d \mathbb{C}^n)$ and that $\otimes_{j=1}^d U_j$ under this identification turns into $U_1 \otimes (\otimes_{j=2}^d U_j)$. □

4 General Domain Toeplitz and Hankel Operators and Matrices

The operators in the title arise as discretizations of general domain Toeplitz/Hankel integral operators. These become "summing operators", which can be represented as matrices in many ways, which we describe in the next section.

4.1 Discretization

For simplicity of notation, we here discretize using an integer grid, since grids with other sampling lengths (these are considered in Sect. 6.1) can be obtained by first dilating the respective domains. Let Ξ, Υ be any open connected and bounded domains in $\mathbb{R}^d$, and let f be a bounded function defined on $\Omega = \Xi - \Upsilon$. We will throughout the paper use bold symbols for discrete objects, and normal font for their continuous analogues. Set

$$\mathbf{\Upsilon} = \{\mathbf{x} \in \mathbb{Z}^d : \mathbf{x} \in \Upsilon\},$$

make analogous definition for $\Xi/\mathbf{\Xi}$ and define $\mathbf{\Omega} = \mathbf{\Upsilon} - \mathbf{\Xi}$. We let $\mathbf{\Theta}_f$ denote what we call a *general domain Toeplitz (summing) operator*

$$\mathbf{\Theta}_f(g)(\mathbf{x}) = \sum_{\mathbf{y} \in \mathbf{\Upsilon}} f(\mathbf{x} - \mathbf{y}) g(\mathbf{y}), \quad \mathbf{x} \in \mathbf{\Xi}, \tag{4.1}$$

where g is an arbitrary function on $\mathbf{\Upsilon}$. We will talk of $\mathbf{\Theta}_f$ as a discretization of the corresponding integral operator Θ_f, introduced in Sect. 2.2, more on this in Sect. 6.1.

We may of course represent g as a vector, by ordering the entries in some (non-unique) way. More precisely, by picking any bijection

$$o_y : \{1, \ldots, |\mathbf{\Upsilon}|\} \to \mathbf{\Upsilon} \tag{4.2}$$

we can identify g with the vector $\tilde{g}$ given by

$$(\tilde{g}_j)_{j=1}^{|\mathbf{\Upsilon}|} = g(o_y(j)).$$

Letting o_x be an analogous bijection for $\mathbf{\Xi}$, it is clear that $\mathbf{\Theta}_f$ can be represented as a matrix, where the (i, j)'th element is $f(o_x(i) - o_y(j))$. Such matrices will be called "general domain Toeplitz matrices", see Fig. 2 for a small scale example. We make analogous definitions for Γ_f and denote the corresponding discrete operator by $\mathbf{\Gamma}_f$. We refer to this as a "general domain Hankel (summing) operator" and its matrix realization as "general domain Hankel matrix". An example of such is shown in Fig. 1.
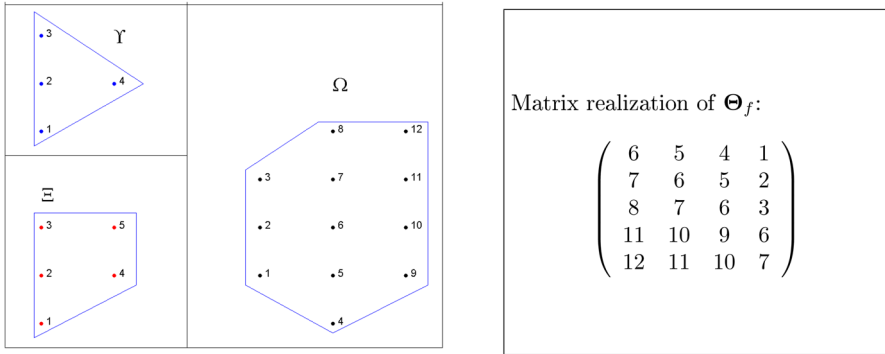
Fig. 2 *Left* Domains Ξ, Υ, and $\Omega = \Xi - \Upsilon$ with the points numbered lexicographically. *Right* Illustration of where the numbered points in Ω show up in the corresponding matrix realization of Θ_f

4.2 Block Toeplitz and Hankel Matrices

If we let Ξ and Υ be multi-cubes and the ordering bijections be the lexicographical order, then the matrix realization Θ_f of (4.1) becomes a block Toeplitz matrix. These are thus special cases of the more general operators considered above. Similarly, block Hankel matrices arise when representing Γ_f in the same way.

For demonstration we consider $\Xi = \Upsilon = \{-1, 0, 1\}^3$ so $\Omega = \{-2, \ldots, 2\}^3$. The lexicographical order then orders $\{-1, 0, 1\}^3$ as

$$(-1, -1, -1), \ (-1, -1, 0), \ (-1, -1, 1), \ (-1, 0, -1), \ (-1, 0, 0), \ldots, (1, 1, 1).$$

The matrix-realization T of a multidimensional Toeplitz operator Θ_f then becomes

$$T = \begin{pmatrix}
T_{f_3(0,0)} & T_{f_3(0,-1)} & T_{f_3(0,-2)} & T_{f_3(-1,0)} & T_{f_3(-1,-1)} & T_{f_3(-1,-2)} & T_{f_3(-2,0)} & T_{f_3(-2,-1)} & T_{f_3(-2,-2)} \\
T_{f_3(0,1)} & T_{f_3(0,0)} & T_{f_3(0,-1)} & T_{f_3(-1,1)} & T_{f_3(-1,0)} & T_{f_3(-1,-1)} & T_{f_3(-2,1)} & T_{f_3(-2,0)} & T_{f_3(-2,-1)} \\
T_{f_3(0,2)} & T_{f_3(0,1)} & T_{f_3(0,0)} & T_{f_3(-1,2)} & T_{f_3(-1,1)} & T_{f_3(-1,0)} & T_{f_3(-2,2)} & T_{f_3(-2,1)} & T_{f_3(-2,0)} \\
T_{f_3(1,0)} & T_{f_3(1,-1)} & T_{f_3(1,-2)} & T_{f_3(0,0)} & T_{f_3(0,-1)} & T_{f_3(0,-2)} & T_{f_3(-1,0)} & T_{f_3(-1,-1)} & T_{f_3(-1,-2)} \\
T_{f_3(1,1)} & T_{f_3(1,0)} & T_{f_3(1,-1)} & T_{f_3(0,1)} & T_{f_3(0,0)} & T_{f_3(0,-1)} & T_{f_3(-1,1)} & T_{f_3(-1,0)} & T_{f_3(-1,-1)} \\
T_{f_3(1,2)} & T_{f_3(1,1)} & T_{f_3(1,0)} & T_{f_3(0,2)} & T_{f_3(0,1)} & T_{f_3(0,0)} & T_{f_3(-1,2)} & T_{f_3(-1,1)} & T_{f_3(-1,0)} \\
T_{f_3(2,0)} & T_{f_3(2,-1)} & T_{f_3(2,-2)} & T_{f_3(1,0)} & T_{f_3(1,-1)} & T_{f_3(1,-2)} & T_{f_3(0,0)} & T_{f_3(0,-1)} & T_{f_3(0,-2)} \\
T_{f_3(2,1)} & T_{f_3(2,0)} & T_{f_3(2,-1)} & T_{f_3(1,1)} & T_{f_3(1,0)} & T_{f_3(1,-1)} & T_{f_3(0,1)} & T_{f_3(0,0)} & T_{f_3(0,-1)} \\
T_{f_3(2,2)} & T_{f_3(2,1)} & T_{f_3(2,0)} & T_{f_3(1,2)} & T_{f_3(1,1)} & T_{f_3(1,0)} & T_{f_3(0,2)} & T_{f_3(0,1)} & T_{f_3(0,0)}
\end{pmatrix}$$

where e.g.

$$T_{f_3(0,0)} = \begin{pmatrix}
f(0, 0, 0) & f(0, 0, -1) & f(0, 0, -2) \\
f(0, 0, 1) & f(0, 0, 0) & f(0, 0, -1) \\
f(0, 0, 2) & f(0, 0, 1) & f(0, 0, 0)
\end{pmatrix}$$

Note that this matrix has a Toeplitz structure on 3 levels, since each 3×3-block of the large matrix above is Toeplitz, and these blocks themselves form a 3×3 Toeplitz structure.

4.3 Exponential Sums

We pause the general development to note some standard facts that will be needed in what follows. Fix $N \in \mathbb{N}$, and for $j = 1, \ldots, d$ let Φ_j be a set of at most $2N$ numbers in $\mathbb{C}$. Set $\Phi = \Phi_1 \times \cdots \times \Phi_d$.

Proposition 4.1 *The set $\{e^{\zeta \cdot x} : \zeta \in \Phi\}$ is linearly independent as functions on $\{-N, \ldots, N\}^d$.*

Proof If $d = 1$ the result is standard, see e.g. Proposition 1.1 in [18] or [12, Sec. 3.3]. For $d > 1$, the function $e^{\zeta \cdot x} = e^{\zeta_1 x_1} \ldots e^{\zeta_d x_d}$ is a tensor. The desired conclusion now follows from the $d = 1$ case and standard tensor product theory.

We now set $\Upsilon = \Xi = \{-N, \ldots, N\}^d$, and let $\Omega = \{-2N, \ldots, 2N\}^d$ in accordance with Sect. 4.1. Consider functions f on Ω given by

$$f(x) = \sum_{k=1}^{K} c_k e^{\zeta_k \cdot x}. \tag{4.3}$$

We say that the representation (4.3) is reduced if all ζ_k's are distinct and the corresponding coefficients c_k are non-zero.

Proposition 4.2 *Let Φ be as before. Let the function f on $\{-2N, \ldots, 2N\}^d$ be of the reduced form (4.3) where each ζ_k is an element of Φ. Then*

$$\mathsf{Rank}\,\Theta_f = \mathsf{Rank}\,\Gamma_f = K.$$

Proof Pick a fixed $\zeta \in \mathbb{C}^d$ and consider $f(x) = e^{\zeta \cdot x}$ then

$$\Theta_f(g)(x) = \sum_{y \in \Upsilon} e^{\zeta \cdot x} e^{-\zeta \cdot y} g(y) = e^{\zeta \cdot x} \langle g, \overline{e^{-\zeta \cdot y}} \rangle, \tag{4.4}$$

which has rank 1. For a general f of the form (4.3) the rank will thus be less than or equal to K. But Proposition 4.1 implies that the set $\{e^{\zeta_k \cdot x}\}_{k=1}^{K}$ is linearly independent as functions on Ξ. Thus the rank will be precisely K, as desired. The argument for Γ_f is analogous. $\square$

We end this section with a technical observation concerning 1 variable.

Proposition 4.3 *Let f be a vector of length $m > n + 1$ and $K < n$. Let $\zeta_1, \ldots, \zeta_K$ be fixed and suppose that each sub-vector of f with length $n + 1$ can be written of the form (4.3), then f can be written in this form as well.*

Proof Consider two adjacent sub-vectors with overlap of length n. On this overlap the representation (4.3) is unique, due to Proposition 4.1. The result now easily follows.

5 The Multidimensional Discrete Carathéodory-Fejér Theorem

Throughout this section, let $\boldsymbol{\Upsilon}$, $\boldsymbol{\Xi}$ and $\boldsymbol{\Omega}$ be as in Sects. 4.2 and 4.3, i.e. multi-cubes centered at 0. The following theorem was first observed in [53], but using a completely different proof.

Theorem 5.1 *Set $\boldsymbol{\Xi} = \boldsymbol{\Upsilon} = \{-N, \dots, N\}^d$. Given f on $\{-2N, \dots, 2N\}^d$, suppose that $\boldsymbol{\Theta}_f$ is PSD and has rank K where $K \leq 2N$. Then f can be written as*

$$f(\boldsymbol{x}) = \sum_{k=1}^{K} c_k e^{i\xi_k \cdot \boldsymbol{x}} \tag{5.1}$$

where $c_k > 0$ and $\xi_k \in \mathbb{R}^d$ are distinct and unique. Conversely, if f has this form then $\boldsymbol{\Theta}_f$ is PSD with rank K.

The proof is based on the following simple observation about PSD matrices. Let $\mathsf{Ran}\ A$ denote the range of a matrix A, and $\mathsf{Ker}\ A$ its kernel.

Proposition 5.2 *Let*

$$\begin{pmatrix} A & B \\ B^* & C \end{pmatrix} \tag{5.2}$$

be a PSD matrix, where A, B, C are matrices (with dimensions compatible with (5.2)). Then

$$\mathsf{Ran}\ B \subset \mathsf{Ran}\ A.$$

Proof Note that the orthogonal complement of $\mathsf{Ran}\ B$ equals $\mathsf{Ker}\ B^*$. Since $A = A^*$ it suffices to show that $\mathsf{Ker}\ A \subset \mathsf{Ker}\ B^*$. Suppose that this is not the case and let $x \in \mathsf{Ker}\ A$ be such that $B^*x = y \neq 0$. For $t \in \mathbb{R}$ arbitrary we have

$$\left\langle \begin{pmatrix} A & B \\ B^* & C \end{pmatrix} \begin{pmatrix} x \\ ty \end{pmatrix}, \begin{pmatrix} x \\ ty \end{pmatrix} \right\rangle = 2t\mathsf{Re}\ \langle B^*x, y \rangle + t^2 \langle Cy, y \rangle = 2t\|y\|^2 + t^2 \langle Cy, y \rangle$$

Since $y \neq 0$ this expression takes negative values for some t, which is a contradiction. $\qquad\square$

Proof of Theorem 5.1 First assume that $\boldsymbol{\Theta}_f$ is PSD and has rank K. Let T be a block Toeplitz representation of $\boldsymbol{\Theta}_f$, as described in Sect. 4.2. Recall that the Toeplitz matrix $T_{f_d(\boldsymbol{0})}$ is the $(2N+1) \times (2N+1)$ sub-matrix on the diagonal of T, (and $\boldsymbol{0} \in \mathbb{Z}^{d-1}$). This is clearly PSD and of some rank $J_d \leq K$, so by the classical Carathéodory-Fejér theorem (Theorem 2.1), $f_d(\boldsymbol{0})$ can be represented by

$$f_d(\boldsymbol{0}) = \sum_{k=1}^{J_d} c_k e^{i\xi_k^d x}, \quad \boldsymbol{x} \in \{-2N, \dots, 2N\} \tag{5.3}$$

with $\xi_k^d \in \mathbb{R}$. We identify functions on $\{-N \ldots N\}$ with $\mathbb{C}^{2N+1}$ in the obvious way, and define $U_d \subset \mathbb{C}^{2N+1}$ by

$$U_d = \mathsf{Span} \{e^{i\xi_1^d \boldsymbol{x}}, \ldots, e^{i\xi_{J_d}^d \boldsymbol{x}}\}.$$

The analogous subspace of $\mathbb{C}^{4N+1}$ will be called U_d^{ext}. Note that $f_d(\mathbf{0}) \in U_d^{ext}$ by (5.3) and that

$$\mathsf{Ran}\, T_{f_d(\mathbf{0})} = U_d, \tag{5.4}$$

which follows easily from the proof of Proposition 4.2. Set $\Phi_d = \{\xi_1^d, \ldots, \xi_{J_d}^d\}$.

Fix $\boldsymbol{j} \in \{-2N, \ldots, 2N\}^{d-1}$ with $\boldsymbol{j} \leq \mathbf{0}$ in the lexicographical order. By restricting T to a suitable subspace, it is clear that the matrix

$$\begin{pmatrix} T_{f_d(\mathbf{0})} & T_{f_d(\boldsymbol{j})} \\ T_{f_d(-\boldsymbol{j})} & T_{f_d(\mathbf{0})} \end{pmatrix}$$

is PSD (see the example in Sect. 4.2). Hence $T^*_{f_d(\boldsymbol{j})} = T_{f_d(-\boldsymbol{j})}$ and Proposition 5.2 and (5.4) then imply that $\mathsf{Ran}\, T_{f_d(\boldsymbol{j})} \subset U_d$. It follows that all sub-vectors of $f_d(\boldsymbol{j})$ of length $2N + 1$ are in U_d, and thus

$$f_d(\boldsymbol{j}) \in U_d^{ext} \tag{5.5}$$

by Proposition 4.3. Moreover the relation $T^*_{f_d(\boldsymbol{j})} = T_{f_d(-\boldsymbol{j})}$ implies that $f_d(-\boldsymbol{j}) = \overline{\mathrm{flip}\, f_d(\boldsymbol{j})}$ where the operation flip v reverses the order of the vector v. Since flip $U_d^{ext} = U_d^{ext}$, it follows that (5.5) holds for all $\boldsymbol{j}$.

By choosing a different ordering and repeating the above argument, we conclude that for each l ($1 \leq l \leq d$), there is a corresponding subspace U_l^{ext} of dimension $J_l \leq K$ such that all possible probes $f_l(\cdot)$ are in U_l^{ext}. Let $\xi_k \in \mathbb{R}^d$ be an enumeration of all $J = J_1 J_2 \ldots J_d$ multi-frequencies $\Phi_1 \times \cdots \times \Phi_d$. The corresponding J exponential functions span $\otimes_{j=1}^d U_j$. By Theorem 3.1 we can thus write

$$f(x) = \sum_{k=1}^J c_k e^{i\xi_k \cdot \boldsymbol{x}}. \tag{5.6}$$

However, by Proposition 4.2, precisely K of the coefficients c_k are non-zero. This is (5.1). The uniqueness of the multi-frequencies is immediate by Proposition 4.1 (applied with $N := 2N$). The linear independence of these functions also gives that the coefficients are unique. To see that c_k is positive, ($1 \leq k \leq K$), just pick a function g on Ξ which is orthogonal to all other $e^{i\xi_j \cdot \boldsymbol{x}}$, $j \neq k$. Using the formula (4.4) it is easy to see that

$$0 \leq \langle \boldsymbol{\Theta}_f(g), g \rangle = c_k |\langle g, e^{i\xi_k \cdot \boldsymbol{x}} \rangle|^2, \tag{5.7}$$

and the first statement is proved.

For the converse, let f be of the form (5.1). Then $\boldsymbol{\Theta}_f$ has rank K by Proposition 4.2 and the PSD property follows by the fact that

$$0 \leq \langle \boldsymbol{\Theta}_f(g), g \rangle = \sum_{k=1}^{K} c_k |\langle g, e^{i\xi_k \cdot \boldsymbol{x}} \rangle|^2, \tag{5.8}$$

in analogy with (5.7).   □

It is possible to extend this result to more general domains as considered in Sect. 4.1. However, such extensions are connected with some technical conditions, which are not needed in the continuous case. Moreover, in the next section we will show that the discretizations of Sect. 4.1 capture the essence of their continuous counterparts, given sufficient sampling. For these reasons we satisfy with stating such extensions for the continuous case, see Sect. 7.

The above proof could also be modified to apply to block Hankel matrices, but since Fischer's theorem is connected with preconditions to rule out exceptional cases, the result is not so neat. (It does however provide alternative proofs to the results in [44] concerning small Hankel operators). Again, we here present only the cleaner continuous version, see Sect. 8.

6 The Multidimensional Discrete Kronecker Theorem

If we want to imitate the proof of Theorem 5.1 in Kronecker's setting, i.e. without the PSD assumption, then we have to replace (5.3) (a sum of exponentials) with (2.7) (a sum of exponentials with polynomial coefficients). With suitable modifications, the whole argument goes through up until (5.6), where now the ξ_k's can lie in $\mathbb{C}^d$ and c_k also can be polynomials. However, the key step of reducing the (J-term) representation (5.6) to the (K-term) representation (5.1), via Proposition 4.2, fails. Thus, the only conclusion we can draw is that f has a representation

$$f(\boldsymbol{x}) = \sum_{j=1}^{J} p_j(\boldsymbol{x}) e^{\zeta_j \cdot \boldsymbol{x}}, \quad \boldsymbol{x} \in \Omega, \tag{6.1}$$

where $J \leq K$, but we have very little information on the amount of terms in each p_j. This is a fundamental difference compared to before. In [3] examples are presented of general domain Hankel and Toeplitz integral operators, whose generating function is a single polynomial p, where Γ_p has rank K much lower than the amount of monomials needed to represent p. It is also not the case that these polynomials necessarily are the limit of functions of the form (2.13) (in a similar way as (2.10)), and hence we can not dismiss these polynomials as "exceptional". To obtain similar examples in the finite dimensional setting considered here, one can just discretize the corresponding Γ_p found in [3] (as described in Sect. 4.1).

Nevertheless, in the continuous setting (i.e. for operators of the form Θ_f and Γ_f, c.f. (2.11) and (2.14)) the correspondence between rank and the structure of f is

resolved in [3]. In particular it is shown that (either of) these operators have finite rank if and only if f is an exponential polynomial, and that the rank equals K if f is of the (reduced) form

$$f = \sum_{k=1}^{K} c_k e^{\zeta_k \cdot x}. \tag{6.2}$$

We now show that these results apply also in the discrete setting, given that the sampling is sufficiently dense. For simplicity of notation, we only consider the case Γ_f from now on, but include the corresponding results for Θ_f in the main theorems.

6.1 Discretization

Let bounded open domains Υ, Ξ be given, and let $l > 0$ be a sampling length parameter. Set

$$\Upsilon_l = \{ nl \in \Upsilon : n \in \mathbb{Z}^d \},$$

(c.f. (4.1)), make analogous definition for Ξ_l and define $\Omega_l = \Upsilon_l + \Xi_l$. We denote the cardinality of Υ_l by $|\Upsilon_l|$, and we define $\ell^2(\Upsilon_l)$ as the Hilbert space of all functions g on Υ_l and norm

$$\|g\|_{\ell^2} = \sqrt{\sum_{y \in \Upsilon_l} |g(y)|^2}.$$

We let $\Gamma_{f,l} : \ell^2(\Upsilon_l) \to \ell^2(\Xi_l)$ denote the summing operator

$$\Gamma_{f,l}(g)(x) = \sum_{y \in \Upsilon_l} f(x + y)g(y), \quad x \in \Xi_l.$$

When l is understood from the context, we will usually omit it from the notation to simplify the presentation. It clearly does not matter if f is defined on $\Xi + \Upsilon$ or $\Xi_l + \Upsilon_l$, and we use the same notation in both cases. We define $\Theta_{f,l}$ in the obvious analogous manner. Note that in Sects. 4 and 5 we worked with Θ_f, which with the new notation becomes the same as $\Theta_{f,1}$.

Proposition 6.1 *There exists a constant $C > 0$, depending only on Ξ, such that*

$$\|\Gamma_{f,l}\| \leq C l^{-d/2} \|f\|_{\ell^2(\Omega_l)}.$$

Proof By the Cauchy–Schwarz inequality we clearly have

$$|\Gamma_{f,l}(g)(x)| \leq \|f\|_{\ell^2(\Omega_l)} \|g\|_{\ell^2(\Upsilon_l)}$$

for each $x \in \Xi_l$. If we let $|\Xi_l|$ denote the amount of elements in this set, it follows that

$$\|\mathbf{\Gamma}_{f,l}\| \leq \|f_{\mathbf{\Omega}_l}\|_{\ell^2(\mathbf{\Omega}_l)} |\mathbf{\Xi}_l|^{1/2}.$$

Since Ξ is a bounded set, it is clear that $|\mathbf{\Xi}_l| l^d$ is bounded by some constant, and hence the result follows. □

Theorem 6.2 *Let Ξ and Υ be bounded open and connected and let f be a continuous function on $\Xi + \Upsilon$. Then*

$$\mathsf{Rank}\,\mathbf{\Gamma}_{f,l} \leq \mathsf{Rank}\,\Gamma_f.$$

Similarly, $\mathsf{Rank}\,\mathbf{\Theta}_{f,l} \leq \mathsf{Rank}\,\Theta_f$ *for any continuous f on $\Xi - \Upsilon$.*

Proof Given $\mathbf{y} \in \Upsilon_l$ and $t \leq l$ let $C_{\mathbf{y}}^{l,t}$ denote the multi-cube with center $\mathbf{y}$ and sidelength t, i.e. $C_{\mathbf{y}}^{l,t} = \{y \in \mathbb{R}^d : |y - \mathbf{y}|_\infty < t/2\}$, where $|\cdot|_\infty$ denotes the supremum norm in $\mathbb{R}^d$. Choose t_0 such that $\sqrt{d}t_0/2 < \mathsf{dist}(\Upsilon_l, \partial\Upsilon)$. For $t < t_0$ we then have that the set $\{e_{\mathbf{y}}^{l,t}\}_{\mathbf{y}\in\Upsilon_l}$ defined by $e_{\mathbf{y}}^{l,t} = t^{-d/2}\mathbf{1}_{C_{\mathbf{y}}^{l,t}}$ is orthonormal in $L^2(\Upsilon)$. We make analogous definitions for Ξ_l. Clearly $\ell^2(\Upsilon_l)$ is in bijective correspondence with $\mathsf{Span}\,\{e_{\mathbf{y}}^{l,t}\}_{\mathbf{y}\in\Upsilon_l}$ via the canonical map $P^{l,t}$, i.e. $P^{l,t}(\delta_{\mathbf{y}}) = e_{\mathbf{y}}^{l,t}$ where $\delta_{\mathbf{y}}$ is the "Kronecker δ-function". Let $Q^{l,t}$ denote the corresponding map $Q^{l,t} : \ell^2(\Xi_l) \to L^2(\Xi)$.

Now, clearly $\mathsf{Rank}\,Q^{l,t*}\Gamma_f P^{l,t} \leq \mathsf{Rank}\,\Gamma_f$ and

$$\frac{1}{t^d}\langle Q^{l,t*}\Gamma_f P^{l,t}\delta_{\mathbf{y}}, \delta_{\mathbf{x}}\rangle = \frac{1}{t^{2d}} \int_{|x-\mathbf{x}|_\infty < t/2} \int_{|y-\mathbf{y}|_\infty < t/2} f(x+y)\,dy\,dx.$$

If we denote this number by $\tilde{f}^t(\mathbf{x} + \mathbf{y})$, we see that $\frac{1}{t^d}Q^{l,t*}\Gamma_f P^{l,t} = \mathbf{\Gamma}_{\tilde{f}^t,l}$. It follows that $\mathsf{Rank}\,\mathbf{\Gamma}_{\tilde{f}^t,l} \leq \mathsf{Rank}\,\Gamma_f$. Since f is continuous, it is easy to see that $\lim_{t\to 0^+}\tilde{f}^t(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}+\mathbf{y})$, which implies that $\lim_{t\to 0^+}\mathbf{\Gamma}_{\tilde{f}^t,l} = \mathbf{\Gamma}_{f,l}$, and the proof is complete. □

6.2 From Discrete to Continuous

Our next result says that for sufficiently small l, the inequality in Theorem 6.2 is actually an equality. This needs some preparation. Given $\mathbf{y} \in \Upsilon_l$ we abbreviate $C_{\mathbf{y}}^{l,l}$ by $C_{\mathbf{y}}^l$, i.e. the multi-cube with center $\mathbf{y}$ and sidelength l. Set $\Upsilon_l^{int} = \{\mathbf{y} \in \Upsilon_l : C_{\mathbf{y}} \subset \Upsilon\}$, i.e. the set of those $\mathbf{y}$'s whose corresponding multicubes are not intersecting the boundary. Moreover, for each $\mathbf{y} \in \Upsilon_l$, set

$$e_{\mathbf{y}}^l = \begin{cases} l^{-d/2}\mathbf{1}_{C_{\mathbf{y}}^l}, & \text{if } \mathbf{y} \in \Upsilon_l^{int} \\ 0, & \text{else} \end{cases}$$

We now define $P^l : \ell^2(\Upsilon_l) \to L^2(\Upsilon)$ via $P^l(\delta_{\mathbf{y}}) = e_{\mathbf{y}}^l$. Note that this map is only a partial isometry, in fact, $P^{l*}P^l$ is the projection onto $\mathsf{Span}\,\{\delta_{\mathbf{y}} : \mathbf{y} \in \Upsilon_l^{int}\}$,

and $P^l P^{l*}$ is the projection in $L^2(\Upsilon)$ onto the corresponding subspace. We make analogous definitions for Ξ_l, denoting the corresponding partial isometry by Q^l. Set

$$N_l = N_l(\Upsilon) = |\Upsilon_l \setminus \Upsilon_l^{int}|,$$

i.e. N_l is the amount of multi-cubes $C_{\boldsymbol{y}}^l$ intersecting the boundary of Υ, and note that $N_l = \dim \mathsf{Ker}\, P^l$. Since Υ is bounded and open, it is easy to see that $|\Upsilon_l^{int}|$ is proportional to $1/l^d$. We will say that the boundary of a bounded domain Υ is *well-behaved* if

$$\lim_{l \to 0^+} l^d N_l = 0. \tag{6.3}$$

In other words, $\partial \Upsilon$ is well behaved if the amount of multi-cubes $C_{\boldsymbol{y}}^l$ properly contained in Υ asymptotically outnumbers the amount that are not. The next proposition implies that most decent domains have well-behaved boundaries.

Proposition 6.3 *Let Υ be a bounded domain with Lipschitz boundary. Then $\partial \Upsilon$ is well behaved.*

Proof By definition, for each point $x \in \partial \Upsilon$ one can find a local coordinate system such that $\partial \Upsilon$ locally is the graph of a Lipschitz function from some bounded domain in $\mathbb{R}^{d-1}$ to $\mathbb{R}$, see e.g. [51] or [19], Sec. 4.2. It is not hard to see that each such patch of the boundary can be covered by a collection of balls of radius l, where the amount of such balls is bounded by some constant times $1/l^{d-1}$. Since $\partial \Upsilon$ is compact, the same statement applies to the entire boundary. However, it is also easy to see that one ball of radius l can not intersect more than 3^d multi-cubes of the type $C_{\boldsymbol{y}}^l$, and henceforth N_l is bounded by some constant times $1/l^{d-1}$ as well. The desired statement follows immediately.   □

We remark that all bounded convex domains have well behaved boundaries, since such domains have Lipschitz boundaries, (see e.g. [19, Sec. 6.3]). Also, note that the above proof yielded a faster decay of $N_l l^d$ than necessary, so most "natural" domains will have well-behaved boundaries. We are now ready for the main theorem of this section:

Theorem 6.4 *Let the boundaries of the bounded open and connected domains Υ and Ξ be well behaved, and let f be a continuous function on $cl(\Xi + \Upsilon)$. Then*

$$\Gamma_f = \lim_{l \to 0^+} l^d Q^l \Gamma_{f,l} P^{l*}. \tag{6.4}$$

For f continuous and defined on $cl(\Xi - \Upsilon)$ we analogously have $\Theta_f = \lim_{l \to 0^+} l^d Q^l \Theta_{f,l} P^{l*})$.

Proof We first establish that $P^l P^{l*}$ converges to the identity operator I in the *SOT*-topology. Let $g \in L^2(\Upsilon)$ be arbitrary, pick any $\epsilon > 0$ and let $\tilde{g}$ be a continuous function on $cl(\Upsilon)$ with $\|g - \tilde{g}\| < \epsilon$. Then

$$\|g - P^l P^{l*} g\| \leq \|g - \tilde{g}\| + \|\tilde{g} - P^l P^{l*} \tilde{g}\| + \|P^l P^{l*}(\tilde{g} - g)\|.$$

Both the first and the last term are clearly $\leq \epsilon$, whereas it is easy to see that the limit of the middle term as $l \to 0^+$ equals 0, since $\tilde{g}$ is continuous on $cl(\Upsilon)$ and the boundary is well-behaved. Since ϵ was arbitrary we conclude that $\lim_{l \to 0^+} P^l P^{l*} g = g$, as desired. The corresponding fact for Q^l is of course then also true.

Now, since Γ_f is compact by Corollary 2.4 in [3], it follows by the above result and standard operator theory that

$$\Gamma_f = \lim_{l \to 0^+} Q^l Q^{l*} \Gamma_f P^l P^{l*},$$

and hence it suffices to show that

$$0 = \lim_{l \to 0^+} \|Q^l Q^{l*} \Gamma_f P^l P^{l*} - l^d Q^l \mathbf{\Gamma}_{f,l} P^{l*}\| = \lim_{l \to 0^+} \|Q^l (Q^{l*} \Gamma_f P^l - l^d \mathbf{\Gamma}_{f,l}) P^{l*}\|.$$

Since Q^l and P^{l*} are contractions, this follows if

$$\lim_{l \to 0^+} \|Q^{l*} \Gamma_f P^l - l^d \mathbf{\Gamma}_{f,l}\| = 0. \tag{6.5}$$

By the Tietze extension theorem, we may suppose that f is actually defined on $\mathbb{R}^n$ and has compact support there. In particular it will be equicontinuous. Now, to establish (6.5), let $g = g_1 + g_2 \in \ell^2(\mathbf{\Upsilon}_l)$ be arbitrary, where $\mathsf{supp}\, g_1 \subset \mathbf{\Upsilon}_l^{int}$ and $\mathsf{supp}\, g_2 \subset \mathbf{\Upsilon}_l \setminus \mathbf{\Upsilon}_l^{int}$. By definition, $P^l g_2 = 0$ so $Q^{l*} \Gamma_f P^l g_2 = 0$ whereas

$$|l^d \mathbf{\Gamma}_{f,l} g_2(\mathbf{x})| \leq l^d \|f\|_\infty N_l(\Upsilon)^{1/2} \|g_2\|,$$

by the Cauchy–Schwarz inequality. Thus

$$|(Q^{l*} \Gamma_f P^l - l^d \mathbf{\Gamma}_{f,l}) g_2(\mathbf{x})| \leq l^d \|f\|_\infty N_l(\Upsilon)^{1/2} \|g_2\|. \tag{6.6}$$

We now provide estimates for g_1. Given $\mathbf{x} \in \mathbf{\Xi}_l$ and $\mathbf{y} \in \mathbf{\Upsilon}_l$, set

$$\tilde{f}(\mathbf{x} + \mathbf{y}) = \frac{1}{l^{2d}} \int_{|x - \mathbf{x}|_\infty < l/2} \int_{|y - \mathbf{y}|_\infty < l/2} f(x + y) \, dy \, dx$$

and note that

$$\tilde{f}(\mathbf{x} + \mathbf{y}) = \frac{1}{l^d} \langle Q^{l*} \Gamma_f P^l \delta_{\mathbf{y}}, \delta_{\mathbf{x}} \rangle$$

whenever $\mathbf{x} \in \mathbf{\Xi}_l^{int}$ and $\mathbf{y} \in \mathbf{\Upsilon}_l^{int}$. As in the proof of Theorem 6.2 it follows that $Q^{l*} \Gamma_f P^l g_1(\mathbf{x}) = l^d \mathbf{\Gamma}_{\tilde{f},l} g_1(\mathbf{x})$ for $\mathbf{x} \in \mathbf{\Xi}_l^{int}$. For such $\mathbf{x}$ we thus have

$$|(Q^{l*} \Gamma_f P^l - l^d \mathbf{\Gamma}_{f,l}) g_1(\mathbf{x})| = |l^d \mathbf{\Gamma}_{\tilde{f}-f,l} g_1(\mathbf{x})| \leq l^d \|f - \tilde{f}\|_{\ell^2(\mathbf{\Omega}_l)} \|g_1\| \tag{6.7}$$

by Cauchy–Schwarz, and for $x \in \Xi \setminus \Xi_l^{int}$ we get

$$|(Q^{l*}\Gamma_f P^l - l^d \Gamma_{f,l}) g_1(x)| = |l^d \Gamma_{f,l} g_1(x)| \leq l^d \|f\|_\infty |\Upsilon_l|^{1/2} \|g_1\| \qquad (6.8)$$

due to the definition of Q^l. Combining (6.6)–(6.8) we see that

$$\|(Q^{l*}\Gamma_f P^l - l^d \Gamma_{f,l}) g\| \leq \|(Q^{l*}\Gamma_f P^l - l^d \Gamma_{f,l}) g_1\| + \|(Q^{l*}\Gamma_f P^l - l^d \Gamma_{f,l}) g_2\|$$
$$\leq |\Xi_l^{int}|^{1/2} l^d \|f - \tilde{f}\|_{\Omega_l} \|g_1\| + N_l(\Xi)^{1/2} l^d \|f\|_\infty |\Upsilon_l|^{1/2} \|g_1\|$$
$$+ |\Xi_l|^{1/2} l^d \|f\|_\infty N_l(\Upsilon)^{1/2} \|g_2\|.$$

Since Ξ and Υ are bounded sets, $|\Xi_l|$ and $|\Upsilon_l|$ are bounded by some constant C times $1/l^d$, and as $\|g_1\| \leq \|g\|$ and $\|g_2\| \leq \|g\|$, it follows that

$$\|(Q^{l*}\Gamma_f P^l - l^d \Gamma_{f,l})\| \leq C^{1/2} l^{d/2} \|f - \tilde{f}\|_{\ell^2(\Omega_l)} + C^{1/2} N_l(\Xi)^{1/2} l^{d/2} \|f\|_\infty$$
$$+ C^{1/2} l^{d/2} \|f\|_\infty N_l(\Upsilon)^{1/2}.$$

By Proposition 6.3 the last two terms go to 0 as l goes to 0. The same is true for the first term by noting that $l^{d/2} \|f - \tilde{f}\|_{\ell^2(\Omega_l)} \leq \|f - \tilde{f}\|_{\ell^\infty(\Omega_l)} l^{d/2} |\Omega_l|^{1/2}$ and

$$\lim_{l \to 0^+} \|f - \tilde{f}\|_{\ell^\infty(\Omega_l)} = 0,$$

which is an easy consequence of the equicontinuity of f. Thereby (6.5) follows and the proof is complete.   □

In particular, we have the following corollary. Note that the domains need not have well-behaved boundaries.

Corollary 6.5 *Let Υ and Ξ be open, bounded and connected domains, and let f be a continuous function on $cl(\Xi + \Upsilon)$. We then have*

$$\mathsf{Rank}\, \Gamma_f = \lim_{l \to 0^+} \mathsf{Rank}\, \Gamma_{f,l} \qquad (6.9)$$

Similarly, if f is continuous on $cl(\Xi - \Upsilon)$ we have $\mathsf{Rank}\, \Theta_f = \lim_{l \to 0^+} \mathsf{Rank}\, \Theta_{f,l}$.

Proof By Propositions 5.1 and 5.3 in [3], the rank of Γ_f is independent of Υ and Ξ. Combining this with Theorem 6.2, it is easy to see that it suffices to verify the corollary for any open connected subsets of Υ and Ξ. We can thus assume that their boundaries are well-behaved. By Theorem 6.4 and standard operator theory we have

$$\mathsf{Rank}\, \Gamma_f \leq \liminf_{l \to 0^+} \mathsf{Rank}\, l^d Q^l \Gamma_{f,l} P^{l*} = \liminf_{l \to 0^+} \mathsf{Rank}\, Q^l \Gamma_{f,l} P^{l*}$$
$$\leq \liminf_{l \to 0^+} \mathsf{Rank}\, \Gamma_{f,l}.$$

On the other hand, Theorem 6.2 gives

$$\limsup_{l \to 0^+} \mathsf{Rank} \, \mathbf{\Gamma}_{f,l} \leq \mathsf{Rank} \, \Gamma_f.$$

$\square$

7 The Multidimensional Continuous Carathéodory-Fejér Theorem

In the two final sections we investigate how the PSD-condition affects the structure of the generating functions. This condition only makes sense as long as

$$\Xi = \Upsilon,$$

which we assume from now on. In this section we show that the natural counterpart of Carathéodory-Fejér's theorem holds for general domain Toeplitz integral operators Θ_f, and in the next we consider Fischer's theorem for general domain Hankel integral operators.

Theorem 7.1 *Suppose that $\Xi = \Upsilon$ is open bounded and connected, $\Omega = \Xi - \Upsilon$, and $f \in L^2(\Omega)$. Then the operator Θ_f is PSD and has finite rank K if and only if there exist distinct $\xi_1, \ldots, \xi_K \in \mathbb{R}^d$ and $c_1, \ldots, c_K > 0$ such that*

$$f = \sum_{k=1}^{K} c_k e^{i\xi_k \cdot x}. \tag{7.1}$$

Proof Suppose first that Θ_f is PSD and has finite rank K. By Theorem 4.4 in [3], f is an exponential polynomial (i.e. can be written as (6.1)). By uniqueness of analytic continuation, it suffices to prove the result when $\Xi = \Upsilon$ are neighborhoods of some fixed point x_0. By a translation, it is easy to see that we may assume that $x_0 = 0$. We consider discretizations $\mathbf{\Theta}_{f,l}$ of Θ_f where l assume values 2^{-j}, $j \in \mathbb{N}$. For j large enough, (beyond J say), the operator $\mathbf{\Gamma}_{f,2^{-j}}$ has rank K (Corollary 6.5) and Theorem 5.1 applies (upon dilation of the grids). We conclude that for $j > J$ the representation (7.1) holds (on $\mathbf{\Omega}_{2^{-j}} = \mathbf{\Xi}_{2^{-j}} - \mathbf{\Upsilon}_{2^{-j}}$) but the ξ_k's may depend on j. However, since each grid $\mathbf{\Omega}_{2^{-j-1}}$ is a refinement of $\mathbf{\Omega}_{2^{-j}}$, Proposition 4.1 guarantees that this dependence on j may only affect the ordering, not the actual values of the set of ξ_k's used in (7.1). We can thus choose the order at each stage so that it does not depend on j. Since f is an exponential polynomial, it is continuous, so taking the limit $j \to \infty$ easily yields that (7.1) holds when x is a continuous variable as well.

Conversely, suppose that f is of the form (7.1). Then Θ_f has rank K by Proposition 4.1 in [3] (see also the remarks at the end of Sect. 2.2). The PSD condition follows by the continuous analogue of (5.8). $\square$

8 The Multidimensional Continuous Fischer Theorem

Theorem 8.1 *Suppose that $\Xi = \Upsilon$ is open bounded and connected, $\Omega = \Xi + \Upsilon$, and $f \in L^2(\Omega)$. The operator Γ_f is PSD and has finite rank K if and only if there exist distinct $\xi_1, \ldots, \xi_K \in \mathbb{R}^d$ and $c_1, \ldots, c_K > 0$ such that*

$$f = \sum_{k=1}^{K} c_k e^{\xi_k \cdot x}. \tag{8.1}$$

We remark that the continuous version above differs significantly from the discrete case, even in one dimension, since the sequence $(\lambda^n)_{n=0}^{2N}$ generates a PSD Hankel matrix for all $\lambda \in \mathbb{R}$ (even negative values), whereas the base e^{ξ_k} is positive in (8.1). Recall also the example (2.3), which does not fit in the discrete version of (8.1).

Proof Surprisingly, the proof is rather different than that of Theorem 7.1. First suppose that Γ_f is PSD and has finite rank K. Then f is an exponential polynomial, i.e. has a representation (6.1), by Theorem 4.4 in [3]. Suppose that there are non-constant polynomial factors in the representation (6.1), say $p_1(x)e^{\zeta_1 \cdot x}$. Let N be the maximum degree of all polynomials $\{p_j\}_{j=1}^{J}$. Pick a closed subset $\tilde{\Xi} \subset \Xi$ and $r > 0$ such that $\mathrm{dist}(\tilde{\Xi}, \mathbb{R}^d \backslash \Xi) > 2r$. Pick a continuous real valued function $g \in L^2(\mathbb{R}^d)$ with support in $\tilde{\Xi}$ that is orthogonal to the monomial exponentials

$$\{x^\alpha e^{\zeta_j \cdot x}\}_{|\alpha| \leq N, 1 \leq j \leq J} \backslash \{e^{\zeta_1 \cdot x}\}$$

(where $\alpha \in \mathbb{N}^d$ and we use standard multi-index notation), but satisfies $\langle g, e^{\zeta_1 \cdot x} \rangle = 1$, (that such a function exists is standard, see e.g. Proposition 3.1 in [3]). A short calculation shows that

$$\langle \Gamma_f g(\cdot - z), g(\cdot - w) \rangle = p_1(z + w)e^{\zeta_1 \cdot (z+w)} \tag{8.2}$$

whenever $|z|, |w| < r$. Since p_1 is non-constant, there exists a unit length $\nu \in \mathbb{R}^d$ such that $q(t) = p_1(r\nu t)$ is a non-constant polynomial in t. Set $\zeta = r\zeta_1 \cdot \nu$. Consider the operator $A : L^2([0, 1]) \rightarrow L^2(\Xi)$ defined via

$$A(\phi) = \int_0^1 \phi(t)g(x - r\nu t).$$

Clearly $A^*\Gamma_f A$ is PSD. It follows by (8.2) and Fubini's theorem that

$$\langle A^*\Gamma_f A(\phi), \psi \rangle = \int_0^1 \int_0^1 p_1(rvt + rvs)e^{\zeta_1 \cdot (rvt+rvs)} \phi(t)dt\overline{\psi(s)}ds$$
$$= \int_0^1 \int_0^1 q(t+s)e^{\zeta(t+s)}\phi(t)dt\overline{\psi(s)}ds.$$

With $h(t) = q(t)e^{\zeta t}$, it follows that the operator $\Gamma_h : L^2([0,1]) \to L^2([0,1])$ is PSD. Since Γ_h is self adjoint it is easy to see that $h(t+s) = \overline{h(s+t)}$, (either by repeating arguments from Sect. 6, or by standard results from integral operator theory). In particular h is real valued. This clearly implies that $\zeta \in \mathbb{R}$. Now consider the operator $B : L^2([0,1]) \to L^2([0,1])$ defined by $B(g)(t) = e^{-\zeta t}g(t)$. As before we see that $B^*\Gamma_h B = \Gamma_q$, and this operator is PSD. Given $0 < \epsilon < 1/2$, define $C_\epsilon : L^2([0,1/2]) \to L^2([0,1])$ by $C_\epsilon(g)(t) = \frac{g(t-\epsilon)-g(t)}{\epsilon}$, (where we identify functions on $[0,1/2]$ with functions on $\mathbb{R}$ that are identically zero outside the interval). It is easy to see that

$$C_\epsilon^*\Gamma_q \mathbb{C}_\epsilon = \Gamma_{\epsilon^{-2}(q(\cdot+2\epsilon)-2q(\cdot+\epsilon)+q(\cdot))},$$

in particular it is PSD. Since q is a polynomial, it is easy to see that $(q(\cdot+2\epsilon)-2q(\cdot+\epsilon)+q(\cdot))/\epsilon^2$ converges uniformly on compacts to q''. By simple estimates based on the Cauchy–Schwarz inequality (see e.g. Proposition 2.1 in [3]), it then follows that the corresponding sequence of operators converges to $\Gamma_{q''}$ (acting on $L^2([0,1/2])$), which therefore is PSD. Continuing in this way, we see that we can assume that q is of degree 1 or 2, where Γ_q acts on an interval $[0,3l]$ where $3l$ is a power of $1/2$. We first assume that the degree is 2, and parameterize $q(t) = a + b(t/l) + c(t/l)^2$. Performing the differentiation trick once more, we see that Γ_c is PSD on some smaller interval, which clearly means that $c > 0$. Now pick $g \in L^2([0,l])$ such that $\langle g, 1 \rangle = 1$, $\langle g, t \rangle = 0$, $\langle g, t^2 \rangle = 0$, and consider $D : \mathbb{C}^3 \to L^2([0,3l])$ defined by

$$D((c_0, c_1, c_2)) = c_0 g(\cdot) + c_1 g(\cdot - l) + c_2 g(\cdot - 2l).$$

By (8.2), the matrix representation of $D^*\Gamma_q D$ is

$$M = \begin{pmatrix} q(0) & q(l) & q(2l) \\ q(l) & q(2l) & q(3l) \\ q(2l) & q(3l) & q(4l) \end{pmatrix} = \begin{pmatrix} a & a+b+c & a+2b+4c \\ a+b+c & a+2b+4c & a+3b+9c \\ a+2b+4c & a+3b+9c & a+4b+16c \end{pmatrix},$$

which then is PSD. However, a (not so) short calculation shows that the determinant of M equals $-8c^3$ which is a contradiction, since it is less than 0 (recall that $c > 0$). We now consider the case of degree 1, i.e. $c = 0$ and $b \neq 0$. As above we deduce that the matrix

$$M = \begin{pmatrix} q(0) & q(l) \\ q(l) & q(2l) \end{pmatrix} = \begin{pmatrix} a & a+b+c \\ a+b+c & a+2b+4c \end{pmatrix},$$

has to be PSD, which contradicts the fact that its determinant is $-b^2$.

By this we finally conclude that there can be no polynomial factors in the representation (6.1). By the continuous version of Proposition 4.2 (see Proposition 4.1 in [3]), we conclude that f is of the form (6.2), i.e. $f = \sum_{k=1}^{K} c_k e^{\zeta_k \cdot x}$. From here the proof is easy. Repeating the first steps, we conclude that $\zeta_k \cdot v \in \mathbb{R}$ for all $v \in \mathbb{R}^d$, by which we conclude that ζ_k are real valued. We therefore call them ξ_k henceforth. With this at hand we obviously have

$$\langle \mathbf{\Gamma}_f(g), g \rangle = \sum_{k=1}^{K} c_k |\langle g, e^{\xi_k \cdot x} \rangle|^2 \qquad (8.3)$$

for all $g \in L^2(\Xi)$, whereby we conclude that $c_k > 0$.

For the converse part of the statement, let f be of the form (8.1). That Γ_f has rank K has already been argued (Proposition 4.1 in [3]) and that Γ_f is PSD follows by (8.3). The proof is complete. □

9 Unbounded Domains

For completeness, we formulate the results form the previous two sections for unbounded domains. Γ_f is defined precisely as before, i.e. via the formula (2.14), except that we now have to assume that $f(x + \cdot)$ is in $L^2(\Upsilon)$ for every $x \in \Xi$ and vice versa, $f(\cdot + y) \in L^2(\Xi)$ for every $y \in \Upsilon$ (see definition 1.1 in [3]). Obviously, analogous definitions/restrictions apply to Θ_f as well. The main difficulty with unbounded domains is that exponential polynomials then can give rise to unbounded operators. Following [3], we address this by assuming that Ω is convex and we let Δ_Ω[2] denote the set of directions $\vartheta \in \mathbb{R}^d$ such that the orthogonal projection of Ω on the half line $[0, \infty) \cdot \vartheta$ is a bounded set, and we let $int(\Delta_\Omega)$ denote its interior.

Theorem 9.1 *Let $\Xi = \Upsilon \subset \mathbb{R}^d$ be convex domains, set $\Omega = \Xi + \Upsilon$ and let f be a function on Ω such that $f(x + \cdot) \in L^2(\Upsilon) \,\forall x \in \Xi$ and $f(\cdot + y) \in L^2(\Xi) \,\forall y \in \Upsilon$. Then Γ_f is bounded, PSD and has finite rank if and only if f is of the form (8.1) and $\xi_k \in int(\Delta_\Omega)$ for all k.*

Proof This follows by straightforward modifications of the proofs in Section 9 of [3], so we satisfy with outlining the details. The "if" direction is easy so we focus on the "only if". We restrict the operator Γ_f to functions living on a subset (see Theorem 9.1 [3]) to obtain a new operator to which Theorem 8.1 above applies. From this we deduce that f locally has the form (8.1). That this formula then holds globally is an immediate consequence of uniqueness of real analytic continuation, combined with the observation that Ω is connected. Finally, the restriction on the ξ_k's is immediate by Theorem 9.3 in [3]. □

The corresponding situation for general domain Toeplitz integral operators is quite different. We first note that $\Theta_f : L^2(\Upsilon) \to L^2(\Xi)$ is bounded if and only if $\Gamma_f :$

---

[2] It was called Θ_Ω in [3], see Section 9.

$L^2(-\Upsilon) \to L^2(\Xi)$ is bounded, as mentioned in Sect. 2.2 and further elaborated on around formula (1.2) in [3]. With this, we immediately obtain the following theorem.

Theorem 9.2 *Let $\Xi, \Upsilon \subset \mathbb{R}^d$ be convex domains, set $\Omega = \Xi - \Upsilon$ and let f be a function on Ω such that $f(x - \cdot) \in L^2(\Upsilon) \, \forall x \in \Xi$ and $f(\cdot - y) \in L^2(\Xi) \, \forall y \in \Upsilon$. Then Θ_f is bounded and has finite rank if and only if f is an exponential polynomial (i.e. $f(x) = \sum_{j=1}^{J} p_j(x) e^{\zeta_j \cdot x}$) and $\mathsf{Re} \, \zeta_j \in int(\Delta_\Omega)$ for all j.*

However, if now again we let $\Xi = \Upsilon$ and we additionally impose PSD, the proof of Theorem 9.1 combined with Theorem 7.1 shows that $\zeta_j = i\xi_j$ for some $\xi \in \mathbb{R}^d$. However, Theorem 9.2 then forces $0 = \mathsf{Re} \, \zeta_j \in int(\Delta_\Omega)$, which can only happen if $\Delta_\Omega = \mathbb{R}^d$, since it is a cone. This in turn is equivalent to Ω being bounded, so we conclude that

Theorem 9.3 *Let $\Xi = \Upsilon \subset \mathbb{R}^d$ be convex unbounded domains, set $\Omega = \Xi - \Upsilon$ and let f be as in Theorem 9.2. Then Θ_f is bounded and PSD if and only if $f \equiv 0$.*

10 Conclusions

Multidimensional versions of the Kronecker, Carathéodory-Fejér and Fischer theorems are discussed and proven in discrete and continuous settings. The former relates the rank of general domain Hankel and Toeplitz type matrices and operators to the number of exponential polynomials needed for the corresponding generating functions/sequences. The latter two include the condition that the operators be positive semi-definite. The multi-dimensional versions of the Carathéodory-Fejér theorem behave as expected, while the multi-dimensional versions of the Kronecker theorem generically yield more complicated representations, which are clearer in the continuous setting. Fischer's theorem also exhibits a simpler structure in the continuous case than in the discrete. We also show that the discrete case approximates the continuous, given sufficient sampling.

Appendix

In this appendix we motivate some of the claims in Sect. 2. Given a sequence f (defined on $\mathbb{Z}$, say) we let H_f^N be the Hankel matrix $(f_{j+k})_{j,k=0}^{N}$ and we let T_f^N be the Toeplitz matrix $(f_{k-j})_{j,k=0}^{N}$.

Theorem 10.1 *Let f be of the form*

$$f_n = \sum_{j=1}^{J} p_j(n)\lambda_j^n, \quad \lambda_j \neq \lambda_{j'} \quad for\ all\ j \neq j' \tag{10.1}$$

with cardinality $K = \sum_{j=1}^{J}(\deg p_j + 1)$. For any $N \geq K$, we then have

$$\mathsf{Rank}\ H_f^N = \mathsf{Rank}\ T_f^N = K.$$

Proof Let R^T denote the transpose of a given matrix R. It follows from the calculations in the proof of Theorem 1.2 of [18] that H_f^N has a factorization RDR^T where R is a "generalized Vandermonde matrix" with K columns and D is a "block diagonal matrix" whose blocks have a special structure (see the bottom of page 23, [18]). For the present purposes we only need to know that D is invertible, which is immediate from its structure. Moreover, the columns of R are linearly independent (Proposition 1.1 of [18]). That $\mathsf{Rank}\ H_f^N = K$ now follows by basic linear algebra.

The corresponding statement for T_f^N follows by unitary equivalence similarly to the comments at the end of Sect. 2.2. Indeed, let ι^N be the Hankel matrix with ones on the main anti-diagonal and zeroes elsewhere. It is clear that ι is unitary and that $T_f^N = H_g^N \iota$ with $f_{n-N} = g_n$ for $n = 0, \ldots, 2N$. Thus $\mathsf{Rank}\ T_f^N = \mathsf{Rank}\ H_g^N$ and it is also easy to see that the cardinality of f and g coincide. $\qquad\square$

The article [18] is primarily concerned with when the converse of Theorem 10.1 holds. We note that this is the case whenever $\mathsf{Rank}\ H_f^N = \mathsf{Rank}\ H_f^{N-1} \leq N$, see Theorem 3.1 of [18].

Our next concern is to validate the statements concerning the generic form (2.4) of f, i.e. that

$$f_n = \sum_{k=1}^{K} c_k \lambda_k^n, \tag{10.2}$$

holds "generically" given that $\mathsf{Rank}\ H_f^N = K \leq N$ or $\mathsf{Rank}\ T_f^N = K \leq N$. We only outline the details in the Hankel case, the Toeplitz case being an easy consequence as in the previous proof. We adapt the concept of generic as introduced in Definition 7.2 of [3]. Briefly, this says that when dealing with a set $\mathcal{M}$ that is a union of manifolds of possibly different dimensions, a property holds generically if the set $\mathcal{M}_F$ where the property fails is a union of manifolds of lower dimension than the maximum dimension of the components of $\mathcal{M}$.

Now let $\mathcal{M}$ be the set of all sequences f such that $\mathsf{Rank}\ H_f^N \leq K$. First note that by considering c_k and λ_k as variables in $\mathbb{C}$, the expression (10.2) gives rise to a manifold of dimension $2K$ except at degenerate points (e.g. when two λ_k's coincide). In a similar way, we can consider generating functions of the more general form (10.1) as parts of manifolds where the coefficients of each p_j as well as the λ_j's are variables. By counting the number of free variables, it is clear that this gives rise to a manifold of lower dimension than $2K$. It remains to prove that the generating sequences $f \in \mathcal{M}$

that are not of the form (10.1), also are part of manifolds of dimension lower than $2K$. This follows from the below proposition, which characterizes all sequences giving rise to rank K Hankel matrices. Let $\{e_l\}_{l \in \mathbb{Z}}$ denote the canonical basis of $\ell^2(\mathbb{Z})$. We adopt the convention from [18] to define $n^j 0^n = e_j$.

Proposition 10.2 *If* $\mathsf{Rank}\ H_f^N = K \leq N$, *then there exists a number* $1 \leq k \leq K$ *such that*

$$f_n = \sum_{j=1}^{J} p_j(n)\lambda_j^n + \sum_{l=1}^{k} c_l e_{2N+1-l}, \quad n = 0, \ldots, 2N \tag{10.3}$$

where $\sum_{j=1}^{J}(\deg p_j + 1) + k = K$.

Remark Note that, arguing as before, the number of free variables in the representation (10.3) at most is $2(K - k) + k = 2K - k$ (with equality precisely when there are no non-constant polynomials present in the representation (10.3)). The amount of free variables is thus less than $2K - 1$ unless $k = 0$, which concludes the argument preceding the proposition.

Proof Let $\mathsf{Rank}\ H_f^N$ have (r, k)-characteristic, as defined in Section 10, [28]. By definition $\mathsf{Rank}\ H_f^r = \mathsf{Rank}\ H_f^{r-1} = r$, and hence $(f_n)_{n=0}^r$ is of the form (10.1) with cardinality r, by Theorem 3.1 in [18]. This expression can be used to define an alternative sequence $\tilde{f}$, such that $\mathsf{Rank}\ _{\tilde{f}}^M = r$ for all $M \geq r$. By the uniqueness statement in Theorem 9.2 of [28] and the definition of the $(r, k)-$characteristic, the representation (10.3) follows. Finally, $r + k = K$ by Theorem 11.1 of [28]. $\square$

References

1. Adamyan, V.M., Arov, D.Z., Krein, M.G.: Infinite Hankel matrices and generalized Carathéodory-Fejer and Riesz problems. Funct. Anal. Appl. **2**(1), 1–18 (1968)
2. Akhiezer, N.I., Krein, M.G.: Some questions in the theory of moments. AMS, Providence (1962)
3. Andersson, F., Carlsson, M.: On General Domain Truncated Correlation and Convolution Operators with Finite Rank. Integral Equ. Oper. Theory **82**(3), 339–370 (2015)
4. Andersson, F., Carlsson, M., de Hoop, M.: Nonlinear approximation of functions in two dimensions by sums of exponential functions. Appl. Comput. Harmon. Anal. **29**(2), 156–181 (2010)
5. Andersson, F., Carlsson, M., de Hoop, M.: Nonlinear approximation of functions in two dimensions by sums of wave packets. Appl. Comput. Harmon. Anal. **29**(2), 198–213 (2010)
6. Andersson, F., Carlsson, M., de Hoop, M.: Sparse approximation of functions using sums of exponentials and AAK theory. J. Approx. Theory **163**(2), 213–248 (2011)
7. Andersson, F., Carlsson, M., Tourneret, J.-Y., Wendt, H.: A new frequency estimation method for equally and unequally spaced data. Signal Process. IEEE Trans. **62**(21), 5761–5774 (2014)
8. Beylkin, G., Monzón, L.: On generalized gaussian quadratures for exponentials and their applications. Appl. Comput. Harmon. Anal. **12**(3), 332–373 (2002)
9. Beylkin, G., Monzón, L.: On approximation of functions by exponential sums. Appl. Comput. Harmon. Anal. **19**(1), 17–48 (2005)
10. Beylkin, G., Monzón, L.: Approximation by exponential sums revisited. Appl. Comput. Harmon. Anal. **28**(2), 131–149 (2010)
11. Brachat, J., Comon, P., Mourrain, B., Tsigaridas, E.: Symmetric tensor decomposition. Linear Algebra Appl. **433**(11–12), 1851–1872 (2010)
12. Brauer, F., Nohel, J.A.: Ordinary Differential Equations: a First Course. Wa Benjamin, California (1973)

13. Chui, C.K., Chen, G.: Discrete H^∞ optimization, with applications in signal processing and control systems, vol. 26. Springer-Verlag, Berlin (1997)
14. Chui, C.K., Li, X., Ward, J.D.: System reduction via truncated Hankel matrices. Math. Control, Signals Syst. **4**(2), 161–175 (1991)
15. Ciccariello, S., Cervellino, A.: Generalization of a theorem of Carathéodory. J. Phys. A Math. Gen. **39**(48), 14911 (2006)
16. Doyle, J.C., Glover, K., Khargonekar, P.P., Francis, B.A.: State-space solutions to standard H^2 and H^∞ control problems. Autom. Control IEEE Trans. **34**(8), 831–847 (1989)
17. Elkadi, M., Mourrain, B.: Introduction à la résolution des systèmes polynomiaux, vol. 59. Springer Science & Business Media, Berlin (2007)
18. Ellis, R.L., Lay, D.C.: Factorization of finite rank Hankel and Toeplitz matrices. Linear Algebra App. **173**, 19–38 (1992)
19. Evans, L.C., Gariepy, R.F.: Measure theory and fine properties of functions, vol. 5. CRC Press, USA (1991)
20. Feldmann, S., Heinig, G.: Vandermonde factorization and canonical representations of block Hankel matrices. Linear Algebra Appl. **241**, 247–278 (1996)
21. Ferguson, S.H., Lacey, M.T.: A characterization of product BMO by commutators. Acta Math. **189**(2), 143–160 (2002)
22. Fischer, E.: Über das Carathéodorysche problem, potenzreihen mit positivem reellen teil betreffend. Rendiconti del Circolo Matematico di Palermo **32**(1), 240–256 (1911)
23. Geronimo, J.S., Woerdeman, H.J.: Positive extensions, fejér-riesz factorization and autoregressive filters in two variables. Ann. Math. **160**(3), 839–906 (2004)
24. Golyandina, N., Usevich, K.: 2d-extension of singular spectrum analysis: algorithm and elements of theory. Matrix Methods: Theory, Algorithms, Applications, pp. 450–474. World Scientific Publishing, Singapore (2010)
25. Golyandina N., Usevich, K.: An algebraic view on finite rank in 2D-SSA. In Ermakov, S., Melas, V., Pepelyshev, A. (Eds.) Proceedings of the 6th St. Petersburg Workshop on Simulation, vol 1, pp. 308–313. VVM com. Ltd., July (2009)
26. Grenander, U., Szegö, G.: Toeplitz Forms and Their Applications, vol. 321. Univ of California Press, California (1958)
27. Gu, C.: Finite rank hankel operators on the polydisk. Linear Algebra Appl. **288**, 269–281 (1999)
28. Iokhvidov, I.S.: Hankel and Toeplitz matrices and forms: algebraic theory. Birkhauser, Cambridge (1982)
29. Katayama, T.: Subspace methods for system identification. Springer Science & Business Media, Berlin (2006)
30. Kronecker, L.: Leopold Kronecker's Werke. Bände I-V. Chelsea Publishing Co., New York (1968)
31. Kung, S., Lin, D.W.: Optimal Hankel-norm model reductions: multivariable systems. Autom. Control IEEE Trans. **26**(4), 832–852 (1981)
32. Kung, S.-Y., Arun, K.S., Rao, D.V.: State-space and singular-value decomposition-based approximation methods for the harmonic retrieval problem. JOSA **73**(12), 1799–1811 (1983)
33. Kung, S.-Y., Lin, D.W.: A state-space formulation for optimal Hankel-norm approximations. Autom. Control IEEE Trans. **26**(4), 942–946 (1981)
34. Lacey, M., Terwilleger, E.: Hankel operators in several complex variables and product BMO. Houston J. Math. **35**(1), 159–183 (2009)
35. Lacey, M., Terwilleger, E., Wick, B.: Remarks on product BMO. Proc. Am. Math. Soc. **134**(2), 465–474 (2006)
36. Lasserre, J.-B., Laurent, M., Mourrain, B., Rostalski, P., Trébuchet, P.: Moment matrices, border bases and real radical computation. J. Symb. Comput. **51** (2013)
37. Mourrain, B., Pan, V.Y.: Multivariate polynomials, duality, and structured matrices. J. Complex **16**(1), 110–180 (2000)
38. Papy, J.-M., De Lathauwer, L., Van Huffel, S.: Exponential data fitting using multilinear algebra: the single-channel and multi-channel case. Numer Linear Algebra Appl. **12**(8), 809–826 (2005)
39. Papy, J.-M., De Lathauwer, L., Van Huffel, S.: A shift invariance-based order-selection technique for exponential data modelling. IEEE Signal Process. Lett. **14**(7), 473 (2007)
40. Papy, J.-M., De Lathauwer, L., Van Huffel, S.: Exponential data fitting using multilinear algebra: the decimative case. J. Chemom. **23**(7–8), 341–351 (2009)
41. Peller, V.: Hankel operators and their applications. Springer-Verlag, New York (2003)

42. Peter, T., Plonka, G.: A generalized Prony method for reconstruction of sparse sums of eigenfunctions of linear operators. Inverse Probl. **29**(2), 025001 (2013)
43. Pisarenko, V.F.: The retrieval of harmonics from a covariance function. Geophys. J. Int. **33**(3), 347–366 (1973)
44. Power, S.C.: Finite rank multivariable hankel forms. Linear Algebra Appl. **48**, 237–244 (1982)
45. Rochberg, R.: Toeplitz and Hankel operators on the Paley-Wiener space. Integral Equ. Oper. Theory **10**(2), 187–235 (1987)
46. Roy, R., Kailath, T.: ESPRIT-estimation of signal parameters via rotational invariance techniques. IEEE Trans. Acoust. Speech Signal Process. **37**(7), 984–995 (1989)
47. Schmidt, R.O.: Multiple emitter location and signal parameter estimation. Antennas Propag. IEEE Trans. **34**(3), 276–280 (1986)
48. Stoica, P., Moses, R.: Spectral analysis of Signals. Prentice-Hall, USA (2005)
49. Tismenetsky, M.: Factorizations of Hermitian block Hankel matrices. Linear Algebra Appl. **166**, 45–63 (1992)
50. Tyrtyshnikov, E.E.: How bad are hankel matrices? NumerischeMathematik **67**(2), 261–269 (1994)
51. Verchota, G.: Layer potentials and regularity for the Dirichlet problem for Laplace's equation in Lipschitz domains. J. Funct. Anal. **59**(3), 572–611 (1984)
52. Yang, H.H., Hua, Y.: On rank of block hankel matrix for 2-d frequency detection and estimation. Signal Process. IEEE Trans. **44**(4), 1046–1048 (1996)
53. Yang, Z., Xie, L., Stoica P.: Generalized vandermondedecomposition and its use for multi-dimensional super-resolution. In: IEEE International Symposium on Information Theory (ISIT), HongKong, China (2015)